



دانشگاه صنعتی شاهرود



قطب علمی
سامانه های پردازش دیجیتال

پنجمین کنفرانس پردازش سیگنال و

سیستم های هوشمند ایران

۲۷ و ۲۸ آذرماه ۱۳۹۸ دانشکده مهندسی کامپیوتر و دانشکده مهندسی برق
دانشگاه صنعتی شاهرود



تنظیم و صفحه آرایی: محمدمهدی بخشی و حامد غضنفری

دبیرخانه کنفرانس

شاهرود، دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر، دانشکده مهندسی برق

کد پستی: ۳۶۱۹۹۵۱۶۱ تلفن: ۰۲۳۳۲۳۰۰۳۳۸

تارنما: www.icspis.shahroodut.ac.ir و www.icspis.ir

پست الکترونیک: icspis@shahroodut.ac.ir



مجموعه مقالات فارسی

پنجمین کنفرانس

پردازش سیگنال و سیستم‌های هوشمند ایران (ICSPIS)

دانشگاه صنعتی شاهرود، شاهرود

۲۷ و ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود



قطب علمی
سامانه‌های پردازش دیجیتال



عنوان

صفحه

شناسایی هویت نویسنده بر مبنای ترکیب ویژگی‌های ساختاری و بافتی دست خط

علیرضا نقش، سید حسام الدین هاشمی، فاطمه سادات بهشتی نژاد، قاسم گنجی زاده..... ۸

شناسائی محدوده جاده مبتنی بر بینائی ماشین با استفاده از دوربین **Monocular**

راضیه عدالت حقی، مریم مست ظهوری..... ۱۴

کنترل بهینه چندهدفه ربات اسکارا با استفاده از الگوریتم رقابت استعماری

پدرام سرمدی، علیرضا صفا، سیاوش جمشیدی، فهیمه کم، رویا آتابایی..... ۲۰

طراحی کنترل کننده تناسبی-انتگرالی-مشتقی بهینه برای سیستم گوی معلق به کمک الگوریتم ژنتیک

بهنام شجاعی، علیرضا صفا، احمد بادلی، احمد میرتقی، پدرام سرمدی..... ۲۷

تلفیق داده مبتنی بر یادگیری عمیق برای سنجش طیف مشارکتی چندباند

سید سبحان عطاری، روزبه رجبی، یاسر عطار ایزی..... ۳۴

توسعه الگوریتم آلفا ماینر برای پشتیبانی از فرایندهایی با رخدادهای موازی

آزاده سلطانی، حمید فدیشهای..... ۴۰

توسعه الگوریتم **EFIM** برای کشف مجموعه اقلام با متوسط ارزش بالا

آزاده سلطانی، محمود سلطانی..... ۴۶

تحلیل همارجاعی در متن با استفاده از یادگیری عمیق و تکنیک تصمیم‌گیری چندمعیاره

سمیرا حورعلی، مرتضی زاهدی، منصور فاتح..... ۵۲

طراحی سیستم تشخیص نفوذ شبکه بر پایه ترکیب روش های انتخاب ویژگی

آرش ادیبی، پیمان ستوده..... ۵۹

مدل سازی عبارات ظاهر نشده با استفاده از یادگیری عمیق

امین قاضی زاهدی، مرتضی زاهدی، منصور فاتح..... ۶۶

موقعیت یابی داخلی با استفاده از بلوتوث کم انرژی و رویکرد مبتنی بر شبکه عصبی

نقیسه دهقان سلماسی، مرضیه قاضی اسداللهی، رضا عزمی، سید محمدحسین هاشمی نژاد..... ۷۳

بهبود آشکارسازی ناهنجاری در تصاویر ابرطیفی با بیان تنک و فیلتر گوسین

مرضیه بابائی، مریم ایمانی..... ۷۹

پیاده سازی الگوریتم **PCA** برای بهبود اندازه گیری عمق عیوب زیرسطحی در آلیاژ مس به روش جریان گردابی پالسی

علی جهانی سه قلعه..... ۸۶

- آشکارسازی خودکار خونریزیهای بسیار کوچک مغزی در تصاویر SWI مبتنی بر شبکه عصبی کانولوشن
 ۹۳..... زهره نادری اسرمی، محمدرضا حسن زاده، مهدی ازوجی
- جداسازی طیفی تصاویر ابرطیفی بر اساس ساختار تنک بلوکی
 ۹۹..... سید حسین موسوی آذرنگ، روزبه رجبی، هادی زبانی
- کمی سازی استرس با بهره گیری از ویژگی های ارتباطات سیگنال های حجم خون مویرگی و میزان تعریق پوست
 ۱۰۷..... عبدالعزیز عزیزی لاری، علی مالکی
- پیش بینی مدت زمان باقی مانده چرخه زمانی فرآیندهای سازمانی براساس جنبه های کنترلی و جریان داده
 ۱۱۴..... ایمان فیروزیان، مرتضی زاهدی، حمید حسن پور
- تحلیل مرجع مشترک مبتنی بر منابع دانش و شبکه عصبی
 ۱۲۱..... سمیرا حورعلی، مرتضی زاهدی، منصور فاتح
- ارائه یک معماری یادگیری عمیق برای شناسایی اعمال انسان در ویدئو
 ۱۲۷..... مهدی سوادى حسینی، فواد قادری
- تحلیل ارزش و بخش بندی مشتریان با استفاده از تکنیک های داده کاوی و RFM (مطالعه موردی شرکت پتروشیمی
 بجنورد)
 ۱۳۳..... نفیسه روحانی، احسان جعفری، شادی لنگری
- تحلیل رفتار گرافهای تصادفی قدرت در مسأله جانمایی بهینه واحدهای اندازه گیری فاز در شبکه های هوشمند الکتریکی
 ۱۴۲..... سید محمد شهرآئینی، احد الوندی، و شهلا خرمالی
- بهبود امنیت لایه فیزیکی شبکه مشارکتی با استفاده از انتخاب رلهی تقویت -و- ارسال در حضور چندین شنودگر
 ۱۵۰..... سحر پرکوک، محمد ترابی، سعید شکرالهی
- پیشبینی دمای اتاق با استفاده از یادگیری تقویتی
 ۱۵۷..... زهرا مرتضایی، مهسا قاسمی، منصور فاتح
- طبقه بندی تصاویر ابرطیفی با استفاده از پروفایل های توصیفی و گشتاورهای تغییر ناپذیر
 ۱۶۲..... زهرا رستمی، حسن قاسمیان، مریم ایمانی
- افزایش دایره واژگان و بهبود عملکرد واژه نامه احساسی Senti-N-Gram با استفاده از پسوندها و پیشوندهای منفی ساز
 ۱۶۹..... حامد زرگری، دانشجوی دکتری، مرتضی زاهدی، استادیار
- تشخیص نوع نویز در تصویر با استفاده از هیستوگرام و آنتروپی
 ۱۷۵..... زینب خدابخشی، حمید حسن پور، سکینه اسدی امیری

ارزیابی دو روش استخراج ویژگی مبتنی بر پردازش زمان-فرکانس سیگنال رادار با هدف تشخیص زمین خوردن در سالمندان

فاطمه تاجیک، ندا فرجی..... ۱۸۰

افزایش کیفیت تصاویر درماتوسکوپی ملانوما با استفاده مناسب از ضرایب موجک در ساختار یادگیری عمیق

هاله فاتح، محمد تشنه لب، منصور فاتح..... ۱۸۶

بهبود الگوریتم ترمیم تصویر بر اساس اطلاعات ساختار و بافت با استفاده از مفهوم Quad-Tree

کیمیا پیوندی، فرزین یغمایی..... ۱۹۳

یک روش سریع برای انتخاب بهترین گرافها جهت ترکیب به منظور افزایش دقت انتشار برچسب

امیرنامجوی، علیرضا بساقزاده..... ۱۹۹

بازشناسی دیداری اشیاء از سیگنال تک ثبت EEG

مجتبی یاوندحسینی، فواد قادری..... ۲۰۶

ادغام هرمی تصاویر با نورگیری چندگانه با ویژگیهای مبتنی بر رنگ و توزیع شدت روشنایی در دنباله تصاویر ورودی

عیلرضا اسدی، مهدی ازوجی..... ۲۱۳

قطعه‌بندی تومورهای مغزی گلیوما مبتنی بر تصاویر MRI با استفاده از یادگیری عمیق

مطهره آقالری، علی آقاگل زاده، مهدی ازوجی..... ۲۱۹



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه تبریز

۲۷ الی ۲۸ آذرماه ۱۳۹۸

شناسایی هویت نویسنده بر مبنای ترکیب ویژگی‌های ساختاری و بافتی دست خط

علیرضا نقش^۱، سید حسام الدین هاشمی^۲، فاطمه سادات بهشتی نژاد^۳، قاسم گنجی زاده^۴

^۱ استادیار، دانشکده مهندسی برق، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران

Naghsh.a@Pel.iaun.ac.ir

^۲ دانشجوی کارشناسی ارشد هوش مصنوعی و رباتیک، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد نجف آباد، نجف آباد، ایران

Hessam_2013@yahoo.com

^۳ دانشجوی کارشناسی ارشد هوش مصنوعی و رباتیک، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد نجف آباد، نجف آباد، ایران

Fa_beheshti@yahoo.com

^۴ دانشجوی کارشناسی ارشد، دانشکده علوم پایه، دانشگاه یزد، یزد، ایران

Gasemgangizadeh20@gmail.com

چکیده

در این پژوهش رویکردی برون خط و مستقل از متن مبتنی بر ترکیب ویژگی‌های بافتی و ساختاری به منظور شناسایی و تایید هویت نویسنده یک دست‌خط ارائه شده است. ویژگی‌های بافتی توانایی استخراج ویژگی‌های ساختاری از جمله فاصله بین کلمه‌ای یا فاصله بین خطی را ندارند، به همین منظور در این تحقیق سعی شده است علاوه بر معرفی یک ویژگی بافتی و ساختاری جدید، با ترکیب این دو دسته ویژگی، دقت سیستم پیشنهادی افزایش یابد. روش پیشنهادی ابتدا سه ویژگی ساختاری شامل متوسط فاصله بین کلمه‌ای، متوسط فاصله بین خطوط و توزیع توأم جهت و مساحت حلقه‌های بسته را استخراج می‌کند و سپس سه ویژگی توسعه یافته بر مبنای الگوی دودویی محلی را به منظور استخراج ویژگی بافتی از متن استخراج می‌کند. هر سند دست‌نویس به صورت ترکیب سه هیستوگرام از الگوهای دودویی موجود در تصویر بازنمایی می‌شود و در نهایت ترکیب ویژگی‌های ساختاری و بافتی به عنوان ویژگی نهایی به طبقه‌بند نزدیک‌ترین همسایه ارائه می‌شود تا شناسایی نویسنده انجام گیرد. روش پیشنهادی بر روی دو مجموعه پایگاه‌داده یکی به زبان فارسی و دیگری به زبان عربی مورد ارزیابی قرار گرفت که بالاترین دقت با ترکیب ویژگی‌های بافتی به صورت یکنواخت و ویژگی‌های ساختاری به دست آمد.

کلمات کلیدی

شناسایی هویت نویسنده، الگوی دودویی توسعه یافته، ویژگی ساختاری

سبک‌های نوشتاری و استخراج ویژگی‌های کارآمد از کاراکترها و کلمات در یک سند دست‌نویس می‌توان به شناسایی و یا تایید هویت نویسنده دست یافت.

با این حال مانند بسیاری از سیستم‌های هوشمند دیگر چالش‌هایی در شناسایی دست خط وجود دارند که دست‌یابی به یک دقت تشخیص بالا را با مشکل مواجه می‌کنند. به عنوان مثال عوامل بسیاری هستند که دست خط یک فرد را تحت تاثیر قرار می‌دهند، که از آن جمله می‌توان به محدودیت بیومکانیکی فردی، آموزش‌های مرتبط به دست خط، موقعیت روحی فرد، ابزار نوشتاری و حتی هدف از دست خط (تفاوت بین متن مربوط به یک لیست خرید و نامه رسمی) اشاره کرد. علاوه بر این دست‌نویس یک فرد بر پایه یک الگوی خاص قرار دارد که

۱- مقدمه

با رشد و گسترش توان کامپیوترها، استفاده از این ابزار قدرتمند در تحلیل اسناد متنی به وضوح دیده می‌شود. با اینکه که این سیستم‌های کامپیوتری، نمی‌توانند به طور کامل جایگزین تحلیل‌های انسانی شوند، اما نقش اساسی در کمک به انسان برای تحلیل اسناد متنی از جمله بخش بندی، کاهش فضای جستجو، مصورسازی و پردازش سریعتر ایفا می‌کنند. تشخیص و یا تایید هویت نویسنده از روی یک سند متنی دست‌نویس یکی از موضوعاتی است که همواره در اموری مانند جرم شناسی مورد توجه بوده است. شناسایی نویسنده دست خط بستگی به سبک دست خط بین افراد مختلف دارد و لذا با شناسایی تفاوت بین

۲- مرور پیشینه

با توجه به اهمیت موضوع شناسایی نویسنده از روی دست خط در دهه گذشته روش‌های بسیاری ارائه شده است. که تفاوت آن‌ها را می‌توان در دو بخش استخراج ویژگی و رویکرد دسته‌بندی عنوان کرد. به طور کلی ویژگی‌های استخراج شده از دست خط را می‌توان به دو دسته ویژگی‌های ساختاری^۴ و بافتی^۵ طبقه‌بندی کرد [۵]. ویژگی‌های ساختاری در سطح محلی و سراسری قابل استخراج هستند و هدف آن‌ها استخراج خصوصیات ساختاری متن از قبیل ارتفاع متوسط خطوط، فاصله خالی درون کلمه‌ای و برون کلمه‌ای، انحراف زاویه کلمه‌ای با سطح افق و مواردی از این قبیل است. با وجود اینکه ویژگی‌های ساختاری دارای تفسیر هندسی هستند با این حال استخراج این ویژگی‌ها زمان بر بوده و موجب افزایش بار محاسباتی سیستم می‌شود چرا که به طور معمول برای استخراج ویژگی‌های ساختاری نیاز به بخش‌بندی جمله، کلمه و یا حتی حروف می‌باشد. ویژگی‌های مبتنی بر بافت، متن را به عنوان یک سطح شامل بافت در نظر می‌گیرند و ویژگی‌های بافتی را در سطح بلوک یا کل متن استخراج می‌کنند. مزیت این روش‌ها نسبت به دسته اول سرعت بالای اجرا در استخراج ویژگی است.

در [۶] روشی برای شناسایی نویسنده مبتنی بر کاراکترهای دست‌نویس ارائه شده است که از ویژگی‌های ساختاری به منظور توصیف دست نوشته استفاده می‌کند. به منظور استخراج ویژگی، از ویژگی‌های نظیر نسبت طول به عرض کاراکتر، نقاط نهایی^۶ (تعداد، جهات و موقعیت نسبی) و گشتاورهای مرتبه دوم و سوم کاراکتر استفاده کرده‌اند. همچنین در گام طبقه‌بندی از یک شبکه عصبی شعاعی پایه استفاده شده است. نویسندگان کارایی روش پیشنهادی را بر روی یک پایگاه داده خود ساخته شامل ۲۰ نویسنده ارزیابی کردند و به دقت ۹۵/۴۵ درصد دست یافتند.

در [۷] نویسندگان از ۷۰ نویسنده با تحصیلات، سن و جنسیت مختلف برای جمع آوری یک پایگاه داده به زبان فارسی استفاده کردند و از هر یک از نویسندگان خواسته شد تا دو متن از پیش مشخص شده حاوی یک پاراگراف را بنویسند. به منظور استخراج ویژگی، عملیات مورفولوژی بستن، با استفاده از المان‌های ساختاری متفاوت به تصویر اعمال شده است. المان‌های ساختاری همگی خطوط با زوایای شیب متفاوت هستند که اعمال آن به تصویر نشان دهنده موقعیت‌هایی از تصویر دست نوشته است که آن المان ساختاری حضور دارد. پس از این مرحله تعداد اشیاء سیاه باقی مانده در تصویر حاصل از عملیات استخراج زاویه، طول و مساحت کل اشیاء سیاه باقی مانده از عملیات استخراج ویژگی، به عنوان ویژگی نهایی استفاده شده است. در نهایت از یک طبقه‌بند مدل مخفی مارکوف برای طبقه‌بندی استفاده شده است. روش پیشنهادی با ۱۲ زاویه برای استخراج ویژگی بیش‌ترین نرخ تشخیص برابر ۸۱ درصد را روی پایگاه داده جمع آوری شده ارائه کرده است.

شخص طی سالیان متوالی با تمرین به دست می‌آورد و با اینکه دارای شکل و منحنی مشابه هستند اما هیچ دو نمونه از دست نوشته یک شخص دقیقاً همانند نیستند. نتایج اخیر از شناسایی نویسنده در مباحث دادگاهی نشان می‌دهد که دست خط هر فرد صفات کافی از فرد را برای اهداف شناسایی هویت به همراه دارد [۱].

علاوه بر چالش‌های ذاتی دست خط که در بالا اشاره شد مواردی مانند بخش‌بندی پیش زمینه و پس زمینه در اسناد متنی با کیفیت پایین و یا محدود بودن متن موجود از نویسنده ناشناخته می‌تواند مشکلات کار را دو چندان کند [۱].

شناسایی و تایید هویت نویسنده یکی از زیر شاخه‌های علم بیومتریک است که به معنای شناسایی هویت یک فرد بر اساس ویژگی‌های فیزیکی و یا رفتاری وی است [۲]. منظور از ویژگی‌های فیزیکی، صفات کمی بیولوژیکی قابل اندازه‌گیری مانند اثر انگشت، عنبیه و یا چهره است و ویژگی‌های رفتاری به معنای صفات متمایزی است که فرد هنگام انجام برخی از رفتارها مانند راه رفتن، گفتار و یا نوشتن یک دست خط از خود نشان می‌دهد.

روش‌های تشخیص نویسنده از روی سند دست نوشته بر حسب رویه اکتساب داده ورودی به دو دسته عمده ایستا و پویا تقسیم‌بندی می‌شوند [۳]. در روش‌های ایستا کاربران دست نوشته‌های خود را بر روی یک کاغذ می‌نویسند و سپس با استفاده از یک اسکنر نوری یا یک دوربین، نسخه دیجیتالی از دست نوشته تهیه می‌شود و سیستم بیومتریک، هویت نویسنده را با تحلیل تصویر دیجیتالی دست نوشته مشخص می‌کند. این روش به نام برون خط^۱ نیز شناخته می‌شود. در روش‌های پویا، کاربر دست نوشته خود را از طریق دستگاه‌هایی همچون تابلت‌های حساس به فشار به صورت بلادرنگ ضبط می‌کند. در این حالت جهت حرکت دست، میزان فشار قلم و سرعت قلم نیز قابل دریافت است. وجود این اطلاعات موجب شده است روش‌های پویا نسبت به دست اول نتایج بهتری را از نظر دقت تشخیص ارائه دهند.

روش‌های تشخیص دست خط از جنبه دیگری نیز قابل بررسی و دسته‌بندی هستند. وابستگی و استقلال از متن^۲ یکی از مواردی است که در روش‌های تشخیص نویسنده از روی سند دست نوشته مورد توجه می‌باشد. در روش‌های وابسته به متن^۳ به نویسندگان یک متن (ثابت) داده شده که این متن می‌تواند شامل یک صفحه کامل یا کلمات منفرد و حتی کاراکترهای متفاوت باشد. در این حالت، تشخیص تنها بر مبنای یک متن شاخص انجام می‌گیرد. در مقابل روش‌های مستقل از متن نیاز به کاراکترها و متن ثابتی ندارند و به همین دلیل از نظر کاربرد گسترده هستند [۴].

در این پژوهش رویکردی ترکیبی برای تعیین و تایید هویت نویسنده یک سند دست‌نویسه به صورت برون خط و وابسته به متن ارائه شده است.

⁴ Structural

⁵ Textural

⁶ End Point

¹ Offline

² Text Independent Writer Identification

³ Text Dependent Writer Identification

نزدیک‌ترین همسایه ارائه می‌شود. در ادامه هر یک از زیر بخش‌های سیستم پیشنهادی به طور کامل تشریح خواهد شد.

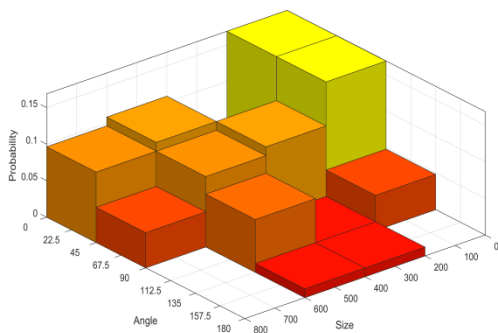
۳-۱- استخراج ویژگی‌های ساختاری

ویژگی‌های ساختاری اغلب به طور سراسری و از کل متن استخراج می‌شوند و اطلاعاتی را در رابطه با شکل کاراکترها ارائه می‌دهند. در این پژوهش از نحوه توزیع جهت و اندازه حلقه‌های درون متنی و همچنین فاصله بین کلمه‌ای و بین خطی به عنوان ویژگی‌های ساختاری استفاده شده است.

۳-۱-۱- توزیع جهت و اندازه حلقه‌های درون متن

منظور از حلقه در اینجا، کاراکترهایی با فرم بسته هستند، مانند کاراکترهای ف و م در فارسی می باشد. به منظور تشخیص حلقه‌ها سند دست نوشته از عملیات ریخت شناسی ریاضیاتی استفاده شده است. ابتدا تمامی حفره‌ها با استفاده از الگوریتم پر کردن^۸ پر شده و سپس تفاضل بین تصویر حاصل و تصویر باینری اولیه محاسبه می‌شود.

پس از استخراج حلقه‌ها، زاویه حلقه با محور افقی در چهار بخش 0 تا $\pi/4$ ، $\pi/4$ تا $\pi/2$ ، $\pi/2$ تا $3\pi/2$ و $3\pi/2$ تا π چندی سازی^۹ می‌شود همچنین مساحت هر حلقه از نظر تعداد پیکسل‌ها محاسبه و سپس در یکی از چهار محدوده 0 تا 200 ، 200 تا 400 ، 400 تا 600 و 600 تا 800 کوانتیزه می‌شود. در نهایت یک هیستوگرام دو بعدی که نحوه توزیع حلقه‌ها با توجه به جهت و اندازه را توصیف می‌کند، ساخته می‌شود. سپس مقادیر هیستوگرام بر مجموع کل مقادیر آن تقسیم می‌شود تا یک هیستوگرام نرمال به دست آید. در شکل ۱ یک نمونه هیستوگرام دو بعدی برای دست نوشته دیده می‌شود. به عنوان مثال این هیستوگرام نشان می‌دهد که نویسنده تمایل بیشتری به ترسیم حلقه‌های کوچک با سایز 0 تا 200 و جهت 0 تا 90 درجه داشته است.



شکل ۱: هیستوگرام دو بعدی توزیع حلقه‌ها با توجه به جهت و اندازه

در [۸] روشی برای شناسایی نویسنده به روش مستقل از متن از روی دست نوشته به زبان کانادایی (یکی از زبان‌های مرسوم در هند) ارائه شده که از یک پایگاه داده شامل ۲۵ نویسنده برای آموزش و تست استفاده شده است. بردار ویژگی شامل ویژگی‌های مکانی چند رزولوشنی بر پایه تبدیل رادون و تبدیل کسینوسی گسسته و همچنین ویژگی‌های ساختاری شامل نسبت طول به عرض و نرخ پیکسلی از هر کلمه است. منظور از نرخ پیکسلی تعداد پیکسل‌های روشن در تصویر مربوط به یک کلمه به کل پیکسل‌های آن است. در واقع روشی جدید برای ترکیب ویژگی‌های استخراج شده از دو یا چند کلمه ارائه شده است. همچنین برای طبقه‌بندی از طبقه‌بند نزدیک‌ترین همسایه با تکنیک ارزیابی متقابل^۷ پنج بخشی استفاده شده است. نرخ متوسط تشخیص $93/2$ درصد با استفاده از تنها یک کلمه و نرخ 100 درصد با استفاده از ترکیب ویژگی‌های سه یا بیشتر از سه کلمه به دست آمده است.

بررسی پژوهش‌های انجام شده نشان می‌دهد ویژگی‌های بافتی نتایج امیدوارکننده‌ای را ارائه کرده‌اند با این حال واضح است که ویژگی‌های بافتی قابلیت استخراج اطلاعات ساختاری را ندارند. هدف این پژوهش ارائه یک ویژگی ساختاری و یک ویژگی بافتی جدید از تصاویر دست نوشته و ترکیب آنها به منظور افزایش دقت نهایی است.

این مقاله به صورت زیر سازماندهی شده است: در بخش سوم جزئیات مربوط به روش پیشنهادی از جمله نحوه استخراج ویژگی‌های ساختاری، ویژگی‌های بافتی و طبقه‌بندی به طور کامل تشریح شده است. بخش چهارم به ارائه نتایج و آزمایشات انجام شده اختصاص داده شده است و در نهایت در بخش پنجم نتیجه‌گیری مطرح خواهد شد.

۳- روش پیشنهادی

همان طور که در بخش مروری بر ادبیات بیان شد الگوریتم‌های بسیاری مبتنی بر اطلاعات بافتی برای توصیف یک سند دسته نوشته ارائه شده است و دلیل این امر را می‌توان سهولت استخراج ویژگی‌های بافتی کارآمد بیان کرد. با این حال توسعه توصیف‌گرهای بافتی قوی‌تر و متمایزکننده‌تر همواره مورد توجه بوده است. از طرف عدم استخراج ویژگی‌های ساختاری از سند دست نوشته را می‌توان به عنوان یکی از معایب الگوریتم‌های استخراج بافت در نظر گرفت. چراکه این ویژگی‌ها قادر به استخراج ویژگی‌های سراسری از جمله فواصل بین خطوط و یا فواصل بین کلمات که می‌توانند اهمیت بالایی در تشخیص نویسنده داشته باشند نیستند. به منظور حل هر دو مساله، در این پژوهش از ترکیب یک ویژگی بافتی جدید با نام الگوی محلی دودویی توسعه یافته با ویژگی‌های ساختاری استفاده شده است.

سیستم پیشنهادی شامل ۳ گام کلی است. ابتدا مجموعه‌ای از پیش پردازش‌ها روی تصویر ورودی انجام می‌شود. سپس ویژگی‌های ساختاری و ویژگی‌های بافت از تصاویر پیش پردازش شده استخراج می‌شوند و در نهایت در گام سوم ویژگی‌های استخراج شده از هر تصویر جهت شناسایی نویسنده یا تایید هویت وی به طبقه‌بند

^۸ Fill Holes

^۹ Quantize

^۷ Cross Validation

اتسو^{۱۱} برای تبدیل سند دست نوشته به یک تصویر دودویی استفاده می‌شود. در گام بعد با استفاده از الگوریتم تشخیص مولفه‌های همبندی، هر یک از مولفه‌های همبند موجود در متن استخراج می‌شوند. فرآیند بعدی که به عنوان پیش پردازش جهت استخراج ویژگی‌های بافتی در نظر گرفته می‌شود حذف نویزهای ناشی از فرآیند آستانه‌گذاری و یا مربوط به نقطه، کاما و سایر مولفه‌هایی هستند که اطلاعات مفیدی در رابطه با دست نوشته فراهم نمی‌کنند و تأثیری در تعیین نویسنده ندارند. بدین منظور مولفه‌های همبندی با تعداد پیکسل‌های کم‌تر از ۲۰ عدد حذف خواهند شد.

بعد از استخراج مولفه‌های همبندی ابتدا مستطیل محدود کننده هر یک از مولفه‌های همبندی مشخص شده، سپس مرکز ثقل هر مولفه از طریق روابط (۳) و (۴) محاسبه می‌شود.

$$COG_x = \frac{1}{m \times n} \sum_{x=1}^N \sum_{y=1}^M I(x, y) \times x \quad (3)$$

$$COG_y = \frac{1}{m \times n} \sum_{x=1}^N \sum_{y=1}^M I(x, y) \times y \quad (4)$$

که در این رابطه $I(x, y)$ تصویر مولفه ی هم بند می باشد. پس از محاسبه مرکز ثقل هر مولفه هم‌بند، به گونه‌ای در کنار یکدیگر قرار می‌گیرند که مرکز ثقل آنها بر روی خط پایه^{۱۱} قرار گیرد. در این بخش متن تک خطی به طول ۲۵۶ پیکسل جدا می‌شود و خط اول را تشکیل می‌دهد. سپس میانگین ارتفاع خط اول (h) محاسبه شده و خط دوم در موقعیت $y_{newLine}$ قرار می‌گیرد.

۳-۲-۲- الگوی دودویی محلی توسعه یافته

عملگر LBP (local binary pattern) برای اولین بار توسط Ojala و همکارانش [۹] ارائه شد. طی آزمایش‌های انجام شده مشخص شد که عملگر LBP توانایی بالایی در بازنمایی بافت دارد.

همان‌طور که در رابطه (۷) مشخص شده است، ویژگی‌های محلی LBP معمولاً در یک همسایگی دایره‌ای متقارن، با مقایسه‌ی پیکسل‌های تصویر با همسایگی‌های آن استخراج می‌شود.

$$LBP(P, R) = \sum_{i=0}^{P-1} u(g_i - g_c) 2^i \quad (7)$$

در رابطه (۷)، p تعداد همسایگی‌ها، R شعاع همسایگی، g_i شدت روشنایی پیکسل همسایگی i ام ($i = 0, \dots, P - 1$)، g_c شدت روشنایی پیکسل مرکزی و $U(x)$ تابع پله‌ای است که به صورت رابطه (۸) تعریف می‌شود.

$$G(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

اگر از ۸ همسایگی برای استخراج الگوهای LBP استفاده شود در این صورت ۵۶ الگوی معنادار بکخواخت و دو الگو برای حالتی که شدت روشنایی همسایگی‌ها همگی بیش‌تر یا کم‌تر از پیکسل مرکزی است و یک الگوی برای بقیه حالت‌ها در نظر گرفته می‌شود. در مجموع پیکسل‌های تصویر با ۵۹ الگو برچسب گذاری می‌شوند. که هیستوگرام

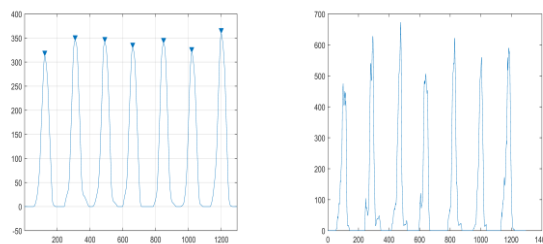
۳-۱-۲- فاصله بین خطوط و بین کلمات

برای استخراج فاصله بین خطوط و فاصله بین کلمات به ترتیب از دو مفهوم پروفایل افقی و پروفایل عمودی تصویر استفاده شده است. که توسط روابط (۱) و (۲) بیان می‌شود.

$$Profile_H(x) = \sum_{y=1}^M I(x, y) \quad (1)$$

$$Profile_v(y) = \sum_{x=1}^N I(x, y) \quad (2)$$

قله‌های ظاهر شده در پروفایل افقی شکل (۲)، متناظر با خطوط موجود در متن هستند. قله‌های برجسته ظاهر شده در پروفایل افقی متناظر با خطوط موجود در متن هستند با این حال همان‌طور که دیده می‌شود قله‌های کوچک بسیاری در پروفایل افقی ظاهر شده‌اند که یافتن قله‌های برجسته را با مشکل مواجه می‌کند به همین منظور در گام اول پروفایل افقی با استفاده از فیلتر میانگین هموار می‌شود و سپس قله‌های اصلی به دست می‌آیند. در نهایت میانگین فاصله بین قله‌های به دست آمده به عنوان ویژگی در نظر گرفته می‌شود.



پروفایل هموار شده

پروفایل اصلی

شکل ۲: پروفایل افقی به دست آمده از متن

در مرحله بعد هر یک از خطوط سند دست نوشته با استفاده از پروفایل افقی به صورت جداگانه، بخش بندی می‌شوند و سپس پروفایل عمودی برای هر خط به طور جداگانه محاسبه می‌شود. مکان‌هایی از پروفایل عمودی که دارای مقدار صفر است متناظر با فضای خالی بین کلمه‌ای می‌باشند. در نهایت میانگین بین فواصل خطوط و فضای خالی بین کلمات به عنوان ویژگی در نظر گرفته می‌شود.

۳-۲-۳- ویژگی بافت

از آنجا که هدف این پژوهش بر روی ویژگی‌ها بافتی متمرکز است، به منظور استخراج ویژگی‌های بافتی، به جای استفاده از تصویر اصلی از تصویر بافتی^{۱۰} استفاده می‌شود. ساخت تصویر بافتی باعث می‌شود نیازی به تصحیح زاویه خطوط برای نرمال‌سازی متن نباشد.

۳-۲-۱- ساخت تصویر بافتی

در ابتدا برای ساخت تصویر بافتی، به دلیل اینکه تصویر ورودی در قالب یک تصویر سطح خاکستری دریافت می‌شود، از آستانه گذاری

¹¹ Otsu

¹² Base Line

¹⁰ Texture Image

همسایه نزدیک خود دارد. طبقه‌بند k نزدیک‌ترین همسایه دارای فاز آموزش نیست و این از مزایای این طبقه‌بند محسوب می‌شود.

۴- نتایج و پیاده‌سازی

میزان کارآمدی روش پیشنهادی با توجه به ویژگی‌های بافتی مختلف و همچنین ترکیب ویژگی ساختاری با ویژگی بافتی پیشنهادی بر روی پایگاه داده IFN/ENIT با ۴۱۱ نویسنده مورد ارزیابی قرار گرفت. همچنین برای ارزیابی روش پیشنهادی بر روی متون به زبان فارسی با توجه به عدم وجود یک پایگاه داده قابل دسترس از ۲۰ نویسنده خواسته شد تا هر یک سه پاراگراف به زبان فارسی را بنویسند که در شکل (۳) نشان داده شده است.

ابوالقاسم نوری (راه ۲۱ خورشیدی - بندر شهید رجایی - ۲۹۷ خورشیدی دروش خراسان)
سخن سرای ایران و رسانه ساهانه جامع ملی ایران است.
نوری را بهترین رسانه برای نوشتن دروس دانشگاهی در زبان فارسی
پیشنهاد می‌دهم و استادان و دانشمندان نام برده است. سوره‌های نوری در این زمینه
بسیار درسی را آموختند و سوره نوری و در دانشگاه‌های مختلف همه دنیا مانند
دانشگاه کلمبیا، سوره صلوات قرار گرفت است.

شکل ۳: نمونه‌ای از سند متنی در پایگاه داده فارسی

نرخ تشخیص نویسنده برای هر یک از پایگاه داده‌ها به طور جداگانه در جدول (۱) و جدول (۲) نشان داده شده است. علاوه بر نرخ تشخیص $top1$ از نرخ تشخیص $top5$ و $top10$ نیز استفاده شده است. منظور از نرخ تشخیص top_k آن است که برچسب واقعی یک نمونه تست در یکی از k همسایگی نزدیک دیده شود. به عبارت دیگر برای مثال با فرض این که نمونه تست مربوط به کلاس $c1$ باشد، در این صورت اگر نزدیک‌ترین همسایه به نمونه تست، از کلاس $c2$ باشد از دید معیار $top1$ خطای طبقه‌بندی رخ داده است، با این حال اگر دومین نزدیک‌ترین همسایه مربوط به کلاس $c1$ باشد در این صورت از دید معیار $top2$ خطای طبقه‌بندی رخ نداده است. بدیهی است با افزایش مقدار k خطا کاهش پیدا می‌کند. نتایج نشان می‌دهد بر روی هر دو پایگاه داده روش پیشنهادی یعنی ترکیب ویژگی ساختاری با الگوی دودویی محلی توسعه یافته بالاترین نرخ تشخیص را ارائه می‌دهد.

جدول ۱: نرخ تشخیص برای پایگاه داده IFN/ENIT

ویژگی	Top 1	Top 5	Top 10
LBP	۷۶/۲۳	۷۹/۴۹	۸۳/۲۸
ELBP	۹۱/۵۶	۹۳/۳۶	۹۵/۳۶
ELBP+ Structural	۹۴/۲۳	۹۷/۵۶	۱۰۰

جدول ۲: نرخ تشخیص برای پایگاه داده فارسی

ویژگی	Top 1	Top 5	Top 10
LBP	۷۰/۰۰	۷۷/۲۳	۷۹/۵۸

نرمال شده تکرار این الگوها ویژگی‌های LBP را تشکیل می‌دهد. همچنین برای مقاوم‌سازی الگوها نسبت به دوران از کمینه مقدار الگوی دودویی تولید شده به صورت رابطه (۹) استفاده می‌شود:

$$LBP^{riu2}(P, R) = \begin{cases} \sum_{i=0}^{P-1} u(g_i - g_c) & \text{if } U(LBP(P, R)) \leq 2 \\ P + 1 & \text{otherwise} \end{cases} \quad (9)$$

برای همسایگی ۸ تایی تعداد الگوهای معنادار مقاوم به دوران ۱۰ عدد خواهد شد.

از آنجایی که LBP تنها رابطه بین نقطه مرکزی و همسایه‌های آن را کد می‌کند، در [۱۰] روشی تحت عنوان ELBP ارائه شده است که روابط مکانی در یک منطقه محلی را نیز کد می‌کند و در نتیجه حاوی اطلاعات مکانی می‌باشد. رویکرد ELBP شامل سه توصیف‌گر شبه LBP است: ELBP_CI، ELBP_NI و ELBP_RD که به ترتیب اطلاعات شدت روشنایی پیکسل مرکزی، پیکسل‌های مجاور آن و تفاوت‌های شعاعی را استخراج می‌کنند. در توصیف‌گر ELBP_CI شدت روشنایی پیکسل مرکزی از طریق آستانه‌گذاری به صورت رابطه (۱۰) محاسبه می‌شود و یک هیستوگرام دو ستونی برای کل تصویر به دست می‌آید.

$$ELBP_CI(X_c) = s(x_c - \beta) \quad (10)$$

که در این رابطه β میانگین کل تصویر است.

در توصیف‌گر ELBP_NI به جای استفاده از مقدار شدت روشنایی پیکسل مرکزی به عنوان مقدار آستانه، از متوسط شدت روشنایی پیکسل‌های همسایه برای آستانه‌گذاری و تولید الگوی دودویی استفاده می‌شود که در رابطه (۱۱) نشان داده شده است.

$$ELBP_NI_{r,p}(x_c) = \sum_{n=0}^{P-1} s(x_{n,r} - \beta_{r,p}) \times 2^n \quad (11)$$

که در این رابطه $\beta_{r,p} = \frac{1}{p} \sum_{n=0}^{P-1} x_{r,p,n}$ می‌باشد. در نهایت توصیف‌گر ELBP_RD از طریق آستانه‌گذاری براساس تفاضل شعاعی پیکسل‌ها به دست می‌آید که در رابطه (۱۲) آورده شده است.

$$ELBP_RD_{r,r-1,p}(x_c) = \sum_{n=0}^{P-1} s(x_{r,p,n} - x_{r-1,p,n}) \times 2^n \quad (12)$$

همانند الگوریتم LBP توصیف‌گرهای E LBP نیز می‌توانند دارای نسخه مقاوم به دوران و الگوهای یکنواخت باشند.

۳-۳- دسته بند نزدیک‌ترین همسایه

در مبحث بازشناسی الگو، الگوریتم k نزدیک‌ترین همسایه، یکی از الگوریتم‌های غیر پارامتری برای طبقه‌بندی و رگرسیون به شمار می‌آید. در هر دو حالت، ورودی شامل k نزدیک‌ترین همسایه از میان نمونه‌های آموزشی به نمونه مورد نظر برای طبقه‌بندی است. در این حالت یک نمونه جدید در کلاسی طبقه‌بندی می‌شود که بیش‌ترین رای را در میان k

صورت یک کد دودویی بیان کند. با این حال این ویژگی نقاط ضعفی دارد که منجر شده نسخه‌های توسعه‌یافته آن ارائه شود. در این تحقیق از الگوی دودویی محلی توسعه‌یافته برای توصیف یک سند دست‌نویس استفاده شد. علاوه بر این از با توجه به از دست رفتن ویژگی ساختاری در تصویر بافتی، به طور موازی وجداگانه سه ویژگی ساختاری نیز از تصویر دست‌نویس استخراج شد. در نهایت روش پیشنهادی بر روی دو پایگاه داده IFN/ENIT به زبان عربی و یک پایگاه داده جمع‌آوری شده به زبان فارسی ارزیابی شد. نتایج به دست آمده نشان می‌دهد روش پیشنهادی نسبت به تعداد نویسندگان در پایگاه داده تا حدی مقاوم است و بسیار کارآمدتر نسبت به الگوریتم استاندارد الگوی دودویی محلی عمل می‌کند.

مراجع

- [1] M. Yadav, A. Kumar, T. Patnaik, and B. Kumar, "A survey on offline signature verification," *International Journal of Engineering and Innovative Technology (IJEIT) Volume*, vol. ۲, pp. ۲۴۰-۲۳۷, ۲۰۱۳.
- [2] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*: Springer Science & Business Media, 2007.
- [3] S. M. Awaida and S. A. Mahmoud, "State of the art in off-line writer identification of handwritten text and survey of writer identification of Arabic text," *Educational Research and Reviews*, vol. ۷, p.445, ۲۰۱۲.
- [4] G. Ghiasi and R. Safabakhsh, "Offline text-independent writer identification using codebook and efficient code extraction methods," *Image and Vision Computing*, vol. ۳۱, pp. ۲۶۹-۲۷۹, 2013.
- [5] S. M. Idicula, "A survey on writer identification schemes," *Writer*, vol. ۲۵, p. ۱۵, 2011.
- [6] J. Ashok, "Writer Identification and Recognition Using Radial Basis Function ۱," 2010.
- [7] H. Khotanlou, "Writer Identity Recognition and Confirmation Using Persian Handwritten Texts," *Advances in Computer Science: An International Journal*, vol. ۴, pp. ۲۰-۲۴, 2015.
- [8] B. Dhandra and M. Vijayalaxmi, "A Novel Approach to Text Dependent Writer Identification of Kannada Handwriting," *Procedia Computer Science*, vol. ۴۹, pp. ۴۱-۴۳, 2015.
- [9] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. ۲۴, pp. ۹۸۷-۹۷۱, 2002.
- [10] L. Liu, L. Zhao, Y. Long, G. Kuang, and P. Fieguth, "Extended local binary patterns for texture classification," *Image and Vision Computing*, vol. ۳۰, pp. ۹۹-۸۶, 2012.
- [11] M. Bulacu, L. Schomaker, and A. Brink, "Text-independent writer identification and verification on offline arabic handwriting," *Ninth International Conference on Document Analysis and Recognition*, pp. ۷۷۳-۷۶۹, 2007.
- [12] M. N. Abdi and M. Khemakhem, "A model-based approach to offline text-independent Arabic writer identification and verification," *Pattern Recognition*, vol. ۴۸, pp. ۱۹۰۳-۱۸۹۰, 2015.

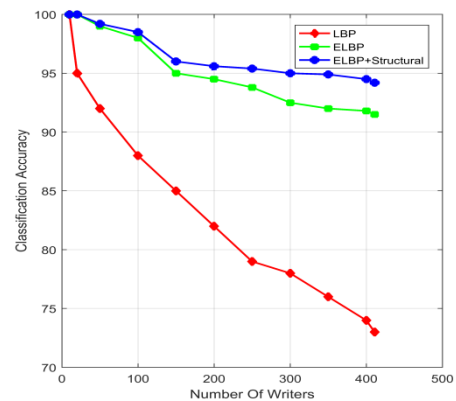
ELBP	۸۸/۲۵	۹۲/۶۳	۹۵/۰۱
ELBP+	۹۶/۲۳	۹۸/۵۹	۱۰۰
Structural			

همچنین نرخ تشخیص روش پیشنهادی بر روی پایگاه داده IFN/ENIT با دو رویکرد دیگر که اخیراً ارائه شده‌اند، مقایسه شده است که نتایج در جدول (۳) ارائه شده است.

جدول 3: مقایسه روش‌های مختلف از نظر نرخ طبقه‌بندی

نرخ شناسایی	تعداد نویسنده‌ها	زبان	پایگاه داده	روش
۸۸	۳۵۰	Arabic	IFN/E NIT	[۱۱]
۹۰	۴۱۱	Arabic	IFN/E NIT	[۱۲]
۹۴/۲۳	۴۱۱	Arabic	IFN/E NIT	روش پیشنهادی

شکل (۴) نشان می‌دهد که ELBP نسبت به تعداد نویسندگان پایداری بیشتری از خود نشان می‌دهد.



شکل ۴: نرخ تشخیص برای پایگاه داده IFN/ENIT بر حسب تعداد نویسندگان

۵- نتیجه گیری

شناسایی هویت یک سند دست‌نویس یکی از موضوعات مورد توجه در حوزه پزشکی قانونی و موارد دادگاهی است که توجه بسیاری از محققان در حوزه پردازش تصویر را در دهه گذشته به خود جلب کرده است. تحقیقات نشان داده است که ویژگی‌های بافتی با وجود فرایند استخراج ساده‌تر نسبت به ویژگی‌های ساختاری، قابلیت تفکیک بالاتری را برای شناسایی نویسنده یک سند دست‌نویس ارائه می‌دهند. یکی از ویژگی‌های بافتی مطرح در زمینه پردازش تصویر الگوی دودویی محلی است که می‌تواند روابط یک پیکسل با پیکسل‌های اطراف آن را به



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۸۹



دانشگاه تبریز

شناسائی محدوده جاده مبتنی بر بینائی ماشین با استفاده از دوربین Monocular

راضیه عدالت حقی^۱، مریم مست ظهوری^۲

^۱ دانشجوی کارشناسی ارشد، دانشکده برق، دانشگاه علم و صنعت، تهران
Saraedalat94@icloud.com

^۲ دکتری، گروه ریاضی، دانشگاه گیلان، رشت
Maryam_zohouri@yahoo.com
zohouri@phd.pnu.ac.ir

چکیده

تشخیص دقیق خطوط جاده یک مسئله مهم در سیستم‌های هشدار خروج خط و سیستم‌های ADAS¹ است که باعث افزایش ایمنی جهت کاهش تصادفات جاده‌ای می‌شود. هدف اصلی این سیستم‌ها تشخیص محدوده جاده و هشدار دادن به راننده زمانیکه خودرو از مسیر اصلی خود در حال انحراف یا انجام اقدامات اصلاحی برای جلوگیری از تصادف در شرایط خطرناک است، می‌باشد. سیستم تشخیص خط در وسایل نقلیه هوشمند از مباحث چالش برانگیز در بینائی ماشین می‌باشد. سنسور دوربین قابلیت بیشتری در تشخیص و تفکیک تصاویر دریافتی داشته و همچنین هزینه کمتری دارد. مبنای عملکرد خوب آشکار ساز، دارا بودن شناسائی دقیق و سریع در شرایط مختلف جاده و آب و هوا می‌باشد که تا حد امکان بتواند مطابق با شرایط حقیقی موجود باشد. در این مطالعه به شناسائی خطوط و محدوده جاده پرداخته‌ایم. آشکار ساز، از شبکه‌های عصبی، ACF و یا سایر روش‌ها استفاده کرده و با سرعت عمل بالا، شیء مورد نظر را شناسائی می‌کند. آشکار سازهای ACF نسبت به آشکار سازهایی که بر اساس شبکه‌های عصبی می‌باشند، عملکرد بهتری دارند. با اعمال دید BEW³ بر روی تصاویر دو بعدی دریافتی از دوربین و با بهره‌گیری از مدل‌های پارابولیک و آشکار ساز ACF، خطوط و خودرو شناسائی می‌شوند. از جمله مزایای روش ACF، سرعت و دقت بالای آن در مقایسه با شبکه‌های عصبی است.

کلمات کلیدی

بینائی ماشین، AFC، تشخیص محدوده جاده، دوربین Monocular.

۱- مقدمه

اندازی کرده که باعث افزایش ایمنی رانندگی می‌شود. در واقع این سیستم ایمنی جهت کاهش تصادفات جاده‌ای ایجاد شده است. این سیستم یک گام فعال مانند هشدار دادن به راننده یا انجام اقدامات اصلاحی برای جلوگیری از تصادف در شرایط خطرناک را فراهم می‌کند. تشخیص دقیق خطوط جاده یک مسئله مهم در سیستم‌های هشدار خروج خط و سیستم‌های کمک به راننده است. با این حال، خطوط مرزی همیشه به وضوح قابل مشاهده نیستند. به عنوان مثال، شرایط بد جاده، کم رنگ بودن خطوط علامت گذاری شده، اثرات محیطی مثل سایه‌های ناشی از اشیاء، درختان، وسایل نقلیه دیگر یا شرایط روشنایی مثل چراغ‌های

بسیاری از مردم هر ساله در تصادفات جاده‌ای به علت عدم توجه راننده جان خود را از دست می‌دهند. در حال حاضر حادثه‌های جاده‌ای به میزان زیادی افزایش یافته است. اغلب حوادث ناشی از عدم رعایت قوانین و رعب و وحشت راننده در حین رانندگی رخ می‌دهد. سیستم کمک راننده در ارائه ایمنی به رانندگان نقش مهمی دارد. در این سیستم‌ها بینائی ماشین به کمک راننده آمده و سیستمی خودکار را راه

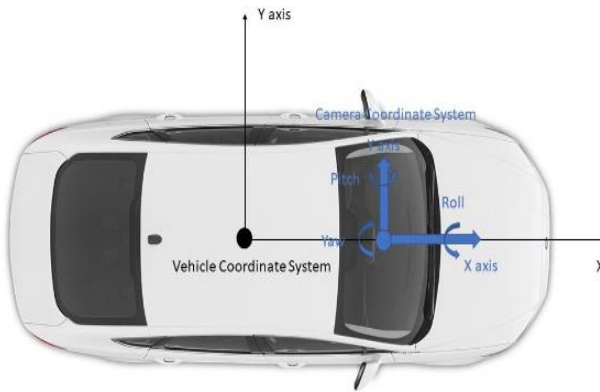
در این روش، شناسایی خودکار شیء و ردیابی آن توسط جعبه ابزار بینایی کامپیوتر^{xxxx} و جعبه ابزار خودکار رانندگی^o ایجاد می‌شود. در اینجا، از مجموعه داده‌های Caltech که شامل تصاویری از محیط خیابان‌های شهر می‌باشد استفاده شده است. نتایج حاصل از این مجموعه داده‌ها عملکرد بهتری در استخراج خطوط مرزی در شرایط مختلف محیطی مثل زمانی که سایه‌ها منجر به تغییرات روشنایی در جاده‌ها می‌شوند، نسبت به روش‌های معمول نشان می‌دهد. وسایل نقلیه خودران می‌توانند محیط را با استفاده از رادار (Radar)، لیدار (Lidar)، جی پی اس (GPS) و دوربین می‌توانند محیط را شناسایی کنند. هر نوع سنسور اطلاعات مختلفی را براساس ردیابی مسیرها ارائه می‌دهد. همچنین، میزان دقت در شرایط آب و هوایی مختلف، متفاوت می‌باشد. به عنوان مثال، یک دوربین دارای خاصیت تشخیص چهره و تفکیک پذیری خوب در لبه و ارتفاع است، اما عملکرد بدی در شب و در شرایط نامناسب آب و هوایی دارد. در مقایسه با آن، رادار و لیدار دارای عملکرد بهینه‌ای در شب می‌باشند و می‌توانند دقت بالایی را در جابه‌جایی سریع خودرو ارائه دهند. با این حال، آنها رزولوشن نامناسب در لبه و ارتفاع دارند. دوربین مونوکیولار (monocular) به همراه آشکار ساز ACF عملکرد بهتری در شناسایی خطوط و خودرو به صورت همزمان عرضه می‌کند [1].

۲- شناسایی محدوده جاده مبتنی بر بینایی ماشین

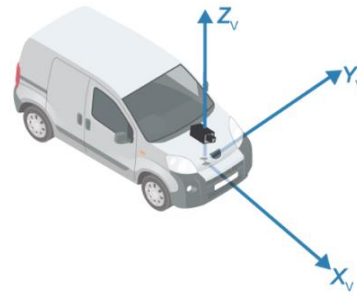
۲-۱- درک بینایی با استفاده از دوربین Monocular

دوربین مونوکیولار قابلیت بیشتری در تشخیص خطوط ممتد و غیرممتد و خطوط عابر پیاده داشته و همچنین هزینه محاسباتی کمتر و سائز دوربین کوچکتر می‌باشد بنابراین محل نصب دوربین آزادتر می‌باشد. در این مطالعه دوربین مونوکیولار جهت تشخیص و ردیابی خطوط به‌کار گرفته شده است و باید کالیبراسیون شود به همین دلیل نیاز به کالیبراسیون داخلی و بیرونی دوربین می‌باشد. دوربین مونوکیولار دارای کالیبراسیون راحتی می‌باشد و در وب سایت Caltech کالیبراسیون دوربین مورد استفاده در متلب ارائه شده است. علاوه بر آن با استفاده از دید BEW که دید دو بعدی می‌باشد و اعمال تابعی که مبتنی بر کنتراست پیکسل‌ها می‌باشد خطوط را شناسایی می‌کند و از آشکار ساز ACF جهت شناسایی خودرو استفاده شده که باعث بهبود عملکرد در این روش گردیده است. ادغام این دو باهم منجر به شناسایی سریع و دقیق شده بنابراین خودرو می‌تواند درست و به موقع تصمیم‌گیری کند. آشکار سازهای بررسی شده در این مطالعه به سه روش تقسیم شده‌اند. که در فصل دوم درک بینایی با استفاده از دوربین مونوکیولار، فصل سوم برجسب زدن اتومات شیء و فصل چهارم تقسیم بندی معنایی با استفاده از یادگیری عمیق بررسی شده‌اند [2].

۲-۲- روش‌های مبتنی بر بینایی

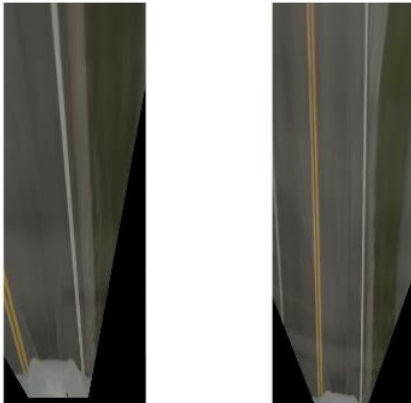


شکل ۳-۲(a): سیستم مختصات دوربین و خودرو



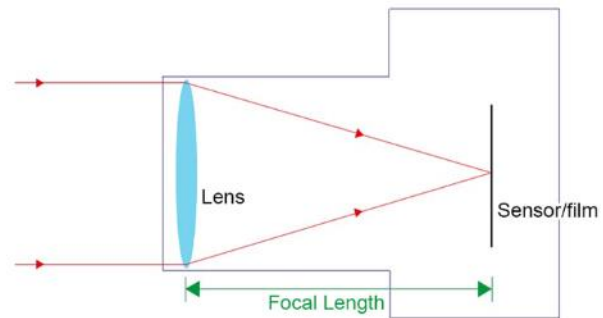
شکل ۱-۲(a): مکان قرارگیری دوربین در خودرو

ارتفاع نصب دوربین مستقیماً نسبت به زمین سنجیده می‌شود، اما زوایای دیگر نیاز به چرخش دستی براساس تصویر BEW دارند. شکل ۳-۲(b) پارامترهای بیرونی مورد نظر BEW را نشان می‌دهد. جاده و خطوط باید در دید صحیح BEW موازی با چرخش صحیح و زاویه نصب درست باشد.



شکل ۳-۲(b): دید BEW قبل (چپ) و بعد (راست) از نصب دوربین با زاویه درست

برنامه camera calibrator پارامترهای بیرونی دوربین را استخراج می‌کند و همان‌طور که شکل ۳-۲(c) نشان می‌دهد، پارامترهای بیرونی برنامه ماتریس نگاشت شده بین مختصات تصویر و مختصات دوربین می‌باشند. پارامترهای بیرونی مورد نیاز ماتریس انتقال بین مختصات خودرو و دوربین در شکل ۳-۲(a) نشان داده شده است.



شکل ۱-۲(b): توصیف طول کانونی

۳-۲- پیکر بندی دوربین

کالیبراسیون دوربین شامل تخمین پارامترهای ذاتی و بیرونی دوربین می‌باشد. پارامترهای ذاتی شامل طول کانونی، مرکز نوری دوربین و اندازه تصویر است. پارامترهای بیرونی شامل Yaw، Pitch، Roll است که می‌تواند یک ماتریس چرخش R و بردار انتقال را نیز ایجاد کند. پارامترهای داخلی برای تعیین رابطه بین مختصات دوربین و مختصات پیکسل هستند. پارامترهای بیرونی برای تعیین رابطه بین مختصات تصویر دوربین و مختصات دنیای واقعی است.

برنامه کالیبراسیون دوربین توسط پردازش تصویر و جعبه ابزار بینایی کامپیوتر در متلب جهت تعیین پارامترهای ذاتی استفاده می‌شود. در برنامه Camera Calibrator از دوربین با یک میدان دید (FOV) 95 درجه استفاده شده است یعنی زاویه‌ها کمتر از 95 درجه می‌باشند. سائز مربع در الگوی شطرنجی ۲۹ میلی متر به عنوان هدف کالیبراسیون ارائه شده است. برای نتیجه مطلوب، نیاز به جمع آوری تعدادی عکس می‌باشد که در اینجا از ده عکس استفاده شده است.

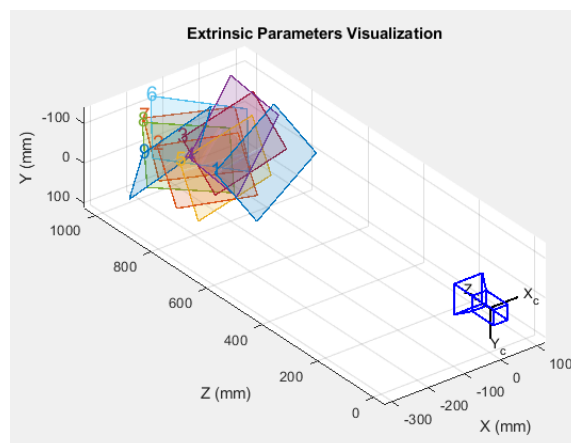
پارامترهای بیرونی دوربین در طی تمام آزمایشات متناسب با مکان نصب سنسور در وسیله نقلیه شامل ارتفاع نصب، Roll، Pitch و Yaw نصب می‌باشند که در شکل ۳-۲(a) نمایش داده شده است.

SegmentLaneMarkerRidge را فراهم می‌کند که از یک آشکارساز لبه، جهت استخراج ویژگی خطوط استفاده می‌کند. این تابع تصویر با دید پرنده را در حالت خاکستری به عنوان ورودی دریافت می‌کند، شیء BirdsEyeView در فاز نقشه برداری با چشم انداز معکوس و یک مقیاس که نشان دهنده عرض تقریبی ویژگی‌های خط برای ردیابی می‌باشد، ایجاد می‌شود و آخرین مقدار به تابع اجازه می‌دهد تا فیلتر مورد استفاده برای آستانه شدت کنتراست را تعیین کند. SegmentLaneMarkerRidge قابلیت دریافت یک آرگومان ورودی اضافی، حساسیت خط، یک فاکتور مقیاس غیر منفی را دارد که اجازه می‌دهد تعریف شود آیا یک مقدار نیاز به حفظ دارد یا خیر. این مقدار باعث بهبود در شناسایی و استخراج ویژگی‌ها می‌شود. در خروجی تابع یک تصویر باینری را با پیکسل‌های درست که اطلاعات مربوط به ویژگی‌های خط را نشان می‌دهد باز می‌گرداند که در شکل ۲(a)-۴ نشان داده شده است. پس از استخراج ویژگی، اتصالات مدل خط توسعه داده شده است.



شکل ۲(a)-۴: استخراج ویژگی خطوط

در این مرحله امکان ایجاد یک مدل پارامتری از خط شناسایی شده جهت تجسم ویژگی‌های استخراج شده در تصویر فراهم می‌شود. هدف اصلی این مرحله بدست آوردن نمایش سطح بالا و فشرده از مسیر است که می‌تواند برای تصمیم‌گیری مورد استفاده قرار گیرد. در این روش از تابع FindParabolicLaneBoundaries برای یافتن مدل خط مناسب استفاده می‌شود. این تابع از الگوریتم RANSAC برای یافتن خطوط مرزی استفاده می‌کند. همانطور که از نام تابع پیداست، مدل ایجاد شده مبتنی بر مدل پارابولیک است که متناسب با مجموعه‌ای از نقاط مرزی و یک عرض تقریبی است. نقاط مرزی منتخب فقط در صورتی که در مرز قرار گیرند، مربوط به فضای داخلی دید می‌باشند. مدل پارابولیک نهایی با استفاده از حداقل مربع‌های متناسب با نقاط فضای داخلی بدست آمده است. تابع در ورودی، نقاط منتخب را در مختصات وسیله نقلیه از فاز استخراج ویژگی‌ها دریافت می‌کند و آرایه‌ای از اشیاء ArabolicLaneBoundary را برای هر مدل فراهم می‌کند. آرایه برگشتی شامل سه ضریب $[a \ b \ c]$ از پارابولا مانند معادله چند جمله‌ای درجه دوم $ax^2 + bx + c$ است. علاوه بر این، دوام خط، نوع خط، حداقل و حداکثر موقعیت نقطه x مربوط به مرز محاسبه شده را نشان می‌دهد. سه پارامتر آخر برای رد کردن برخی از منحنی‌ها استفاده شده که می‌توانند از اکتشافات نامعتبر استفاده کنند. با محاسبه اختلاف بین حداقل و حداکثر x ، طول خط محاسبه می‌شود و با مقدار آستانه طول خط تعریف شده، مقایسه می‌شود. در صورت عدم رسیدن به حداقل آستانه مد نظر، خطوط یافت شده رد می‌شوند. همچنین برای



شکل ۲(c)-۳: برنامه‌های بیرونی توسط برنامه‌ها بر اساس مختصات تصویر و مختصات جهانی

۲-۴- شناسایی و ردیابی خطوط مرزی

در مرحله قبل، محدوده بینایی مورد نظر تعریف شد. این منطقه برای تبدیل تصاویر به تصاویر با دید BEW معرفی می‌شود، طوری که یک نمای مناسب و کافی از ناحیه جلوی خودرو و یک نمای جانبی برای دیدن خطوط جهت پیش بینی جاده میسر باشد. پس از استخراج منطقه دید مورد نظر، شیء Eye View Birds ساخته می‌شود تا بتواند تصویر اصلی را به تصویر برداری با دید پرنده با استفاده از تصویر برداری از چشم انداز معکوس انجام دهد. نتیجه این انتقال در شکل ۲(a)-۴ نمایش داده شده است. تصویر با دید پرنده به توابع اجازه می‌دهد که ویژگی و مدل خطوط استخراج شوند. استخراج ویژگی خطوط شامل شناسایی پیکسل‌هایی است که به خط سفید جاده تعلق دارند و به همراه حذف پیکسل‌های مربوط به غیر خطوط، در تصاویر با دید پرنده که از مرحله قبل استخراج شده است. شناسایی پیکسل‌های مربوطه با استفاده از رویکردی که مبتنی بر مشاهدات کنتراست پیکسل‌ها در مقایسه بین خطوط مارک شده و کف جاده می‌باشد، انجام می‌گیرد و شناخت خطوط با جستجوی پیکسل‌هایی که "مانند خط" هستند پیاده سازی می‌شود. این نوع پیکسل‌ها گروه‌هایی از نقاط با کنتراست رنگی بسیار متفاوت نسبت به نقاط مجاور در هر دو طرف خط می‌باشند. هدف اصلی این روش تشخیص لبه در یک تصویر می‌باشد و از این نظر بسیار حائز اهمیت است. این تکنیک به این جهت تشخیص لبه انتخاب شده است که ساده و اثر بخش می‌باشد و مبتنی بر میدان تنش $\nabla^2 I$ از مشتقات مرتبه اول ساخته شده است و قادر است پاسخی از جهت‌های گرادینان را دریافت کند که باعث می‌شود مقادیر غیر عادی مربوط به جهت‌های منحرف شده از جهت خط مورد انتظار را به آسانی حذف کند. از طرفی جهت بهبود تقسیم بندی ویژگی خطوط باید تصاویر با دید پرنده از RGB به خاکستری تبدیل شوند (شکل ۲(a)-۴). جعبه ابزار سیستم رانندگی خودکار تابعی با عنوان

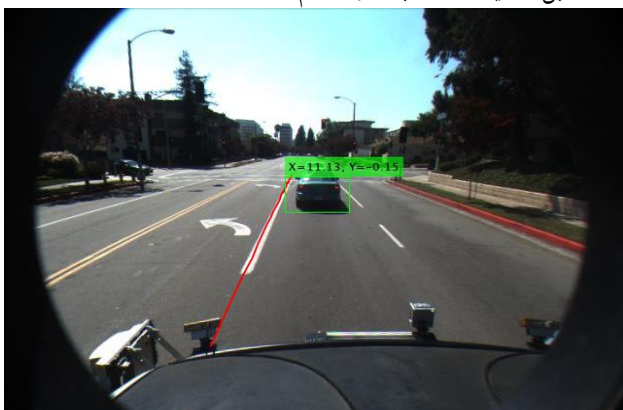
6 Pretained

7 Gaussian Mixture Model

8 Region-Based Convolutional Neural Networks

شکل ۲(a)-۵: شناسایی وسیله نقلیه

هر چند ردیاب ACF در مقایسه با سایر آشکارسازهای از پیش آموزش دیده کار بهتری انجام می‌دهد، اما هنوز مشکلاتی دارد. عملکرد آن به صورت پیوسته برای تمامی شرایط و تمامی فریم‌ها نمی‌باشد و برخی از فریم‌ها ممکن است برچسبی نداشته باشند. مشکل اصلی دیگر این است که برچسب مستطیل با اندازه خودرو سازگار نیست و ممکن است اندازه آن در فریم‌های مختلف متفاوت باشد. در نهایت با اعمال این دو آشکارساز به صورت همزمان به دو ویدئو متفاوت نتایج نشان داده شده در دو شکل ۲(b)-۵ و ۲(c)-۵ مشاهده می‌شود که با سرعت بالا و دقت قابل اطمینان محاسبات را انجام داده است.



شکل ۲(b)-۵: نتیجه آشکار ساز به ویدئو Caltech-cordova

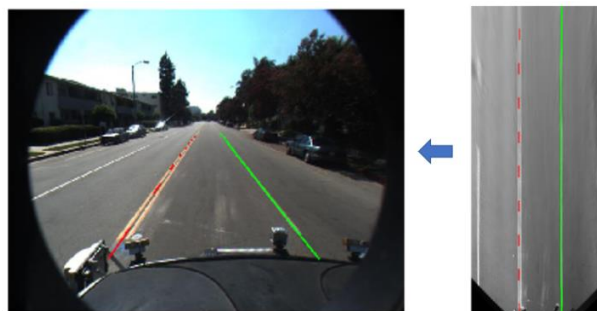


شکل ۲(c)-۵: نتیجه آشکار ساز به ویدئو Caltech-washington

سپاسگزاری

نویسندگان از آقای دکتر علی اصغر بهشتی شیرازی عضو محترم هیات علمی دانشکده برق دانشگاه علم و صنعت کمال تشکر را برای راهنمایی‌های ارزنده ایشان درباره این مقاله دارند.

رد کردن خطوط ضعیف، اندازه طول خط باید بیشتر از آستانه تعیین شده دیگر باشد. مدل‌های خط موجود در مختصات وسایل نقلیه و تصویر با دید پرنده و تصویر اصلی که از دوربین گرفته شده است در شکل ۲(b)-۴ نشان داده شده است. بنابراین نتایج بهتری از تشخیص خطوط در جاده‌هایی با خطوط مستقیم و نور مناسب یافت شده است در حالی که در جاده‌های متقاطع و جاده‌هایی با انحناء و دور زیاد محدودیت‌های چشمگیری مشاهده شده است.



شکل ۲(b)-۴: مدل خطوط

۲-۵- شناسایی وسایل نقلیه

جعبه ابزار رانندگی خودکار و جعبه ابزار بینایی کامپیوتر^{۴۴۴}، انواع آشکارسازهای از پیش آموزش دیده شامل ACF، GMM^۷، R-CNN^۸، Fast R-CNN و Faster R-CNN را ارائه می‌دهد. سه آشکار ساز آخر مبتنی بر روش‌های یادگیری عمیق هستند که در مقایسه با سایر روش‌ها نیاز به زمان پردازش بیشتری در مرحله آزمون دارند. نتایج حاصل از پیاده سازی آشکارساز GMM در مقایسه با آشکار سازهای ACF در نوزده آزمایش فیلم ترافیک کوتاه توسط آزمایشگاه رانندگی خودکار نشان داده است که دقیق نمی‌باشند. بنابراین ACF آشکارساز اصلی در این روش است. در شکل ۲(a)-۵ نتایج حاصل از آشکارساز ACF نمایش داده شده است.



مراجع

- [1] Hoang TM, Hong HG, Vokhidov H, Park KR., Antonietti, M. and Bergström, L. Road Lane Detection by Discriminating Dashed and Solid Road Lanes Using a Visible Light Camera Sensor. *Sensors(Basel)*, 16(8), E1313 (2016).
- [2] Amati N, Tonoli A, Bonfitto A. Combined lateral and longitudinal control for autonomous driving based on Model Predictive Control. *International Journal of Vehicle Mechanics and Mobility*, pp 261-279 (2014).
- [3] Niu M. Object Detection and Tracking for Autonomous Driving by MATLAB toolbox. The Ohio State University (2018).
- [4] Chen, Liang-Chieh et al. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation." *ECCV* (2018).
- [5] Brostow, G. J., J. Fauqueur, and R. Cipolla. "Semantic object classes in video: A high-definition ground truth database." *Pattern Recognition Letters*. Vol. 30, Issue 2, pp 88-97 (2009)
- [6] Brostow, Gabriel J., Jamie Shotton, Julien Fauqueur, and Roberto Cipolla. "Segmentation and Recognition Using Structure from Motion Point Clouds." *ECCV*. (2008)
- [7] Brostow, Gabriel J., Julien Fauqueur, and Roberto Cipolla. "Semantic Object Classes in Video: A High-Definition Ground Truth Database." *Pattern Recognition Letters*. (2008)
- [8] Torr, P. H. S., and A. Zisserman. "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry." *Computer Vision and Image Understanding*. Vol. 18, Issue 1, April, pp. 138-156 (2000).
- [9] M.A. Fischler and R.C. Bolles. Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Commun. ACM*, vol. 24, pg. 381-395 (1981).



کنترل بهینه چندهدفه ربات اسکارا با استفاده از الگوریتم رقابت استعماری

پدرام سرمدی، علیرضا صفا^۱، سیاوش جمشیدی، فهیمه کم، رویا آتابایی

^۱ استادیار گروه مهندسی برق، دانشکده مهندسی، دانشگاه گلستان، گرگان، ایران
آدرس پست الکترونیکی: a.safa@gu.ac.ir

چکیده

الگوریتم رقابت استعماری روش محاسبات تکاملی الهام گرفته از رفتارهای اجتماعی و سیاسی است که با تقلید از روند تکامل اجتماعی، اقتصادی و سیاسی کشورها و با مدل‌سازی ریاضی آن‌ها، ابزارهایی را در قالب مدونی ارائه می‌دهد که می‌توانند به حل مسائل پیچیده بهینه‌سازی کمک کنند. این الگوریتم جواب‌های مسئله بهینه‌سازی را در قالب کشورها نگریسته و سعی می‌کند در طی فرایندی تکرارشونده جواب‌ها را رفته‌رفته بهبود داده و در نهایت به جواب بهینه مسئله برسد. در این مقاله از این الگوریتم برای طراحی کنترل‌کننده بهینه برای یک بازوی ماهر صنعتی بهره گرفته شده است. رویکرد کنترلی اتخاذ شده مبتنی بر کنترل‌کننده تناسبی-انترگالی-مشتقی است. با توجه به ماهیت غیرخطی و توزیع بالای بازوی ماهر صنعتی، تنظیم پارامترهای کنترل‌کننده امری آسانی نیست چراکه اغلب روش‌های موجود در ادبیات کنترل برای سیستم‌های خطی نامتغیر با زمان تک‌ورودی-تک‌خروجی توسعه یافته‌اند. در این مقاله، با تبدیل مسئله طراحی کنترل‌کننده به یک مسئله بهینه‌سازی و با انتخاب یک تابع هزینه مناسب، فرآیند طراحی کنترل‌کننده به قسمی صورت می‌پذیرد که نه تنها سیستم حلقه-بسته پایداری را نتیجه دهد، بلکه خروجی‌های این سیستم‌ها از معیارهای کارایی مطلوبی برخوردار باشد. در این راستا ابتدا معادلات سینماتیکی و دینامیکی یک بازوی ماهر صنعتی با استفاده از روابط اولر-لاگرانژ استخراج می‌شود. سپس از الگوریتم رقابت استعماری برای طراحی کنترل‌کننده‌های بهینه برای این سیستم غیرخطی دو ورودی-دو خروجی کمک گرفته می‌شود. نتایج شبیه‌سازی حاکی از عملکرد مطلوب این کنترل‌کننده است.

کلمات کلیدی

الگوریتم رقابت استعماری چند هدفه، کنترل‌کننده تناسبی-مشتقی-انترگالی، بازوی ماهر صنعتی، سیستم غیرخطی چند متغیره.

۱- مقدمه

وافری برخوردار است و لزوم بهره‌گیری از یک رویکرد کنترلی مناسب احساس می‌گردد.

علی‌رغم این که در طی چند دهه گذشته راهکارهای کنترلی زیادی بسط و توسعه یافته‌اند ولی هنوز اغلب سیستم‌های صنعتی به دلیل سادگی ساختاری و کم‌هزینه بودن از نوع کنترل‌کننده‌های تناسبی-انترگالی-مشتقی (PID) هستند که تنها با سه پارامتر، کنترل انواع فرآیندها را ممکن می‌سازد [۲-۴].

راهکارهای متعددی برای تنظیم ضرایب کنترل‌کننده PID پیشنهاد شده است که از معروف‌ترین آن‌ها می‌توان به روش زیگلر-نیکولز^۲، کهن-کون^۳، مکان هندسی ریشه‌ها و روش‌های مبتنی بر منحنی بودی اشاره نمود [۵]. متأسفانه اغلب این روش‌ها برای سیستم‌های خطی نامتغیر با زمان تک‌ورودی-تک‌خروجی ارائه شده‌اند و توسعه آن برای سیستم‌های غیرخطی چند متغیره مقدور نیست. به‌ویژه، در ربات اسکارا

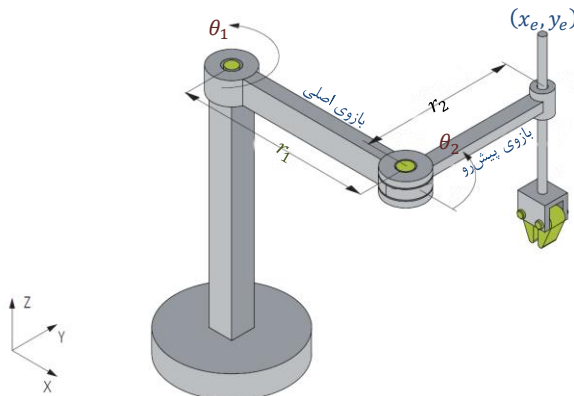
با پیشرفت فناوری و رفع نیازهای اولیه بشر، انسان خود را درگیر دغدغه‌های متعالی‌تری یافت. امروزه ساخت ربات‌ها برای آسان‌تر شدن زندگی و کمک به بالاتر رفتن سطح زندگی انسان‌ها و کمک به بالا بردن توانایی انسان از دغدغه‌های اصلی متخصصین در این حوزه است. یکی از این ربات‌ها صنعتی که توجه ویژه بدان شده است بازوی ربات انعطاف‌پذیر انتخابی^۱ یا به اختصار ربات اسکارا است [۱]. ساختار سری این بازو، موقعیت‌یابی دقیق در فرآیندهای صنعتی را مقدور می‌سازد. وظیفه اصلی اسکارا گرفتن اشیاء، تغییر موقعیت آن‌ها در فضای سه‌بعدی و سپس قرار دادن آن‌ها در مکانی دیگر است؛ بنابراین در این گونه بازوهای صنعتی دستیابی به دقت و سرعت بالا از اهمیت

² Ziegler-Nichols

³ Cohen-Coon

¹ Selective Compliance Assembly Robot Arm

با استفاده از روند محاسبات تکاملی توصیف‌شده، طراحی کنترل‌کننده به یک مسئله بهینه‌سازی ترجمه شده و از الگوریتم رقابت استعماری برای طراحی کنترل‌کننده‌های PID در راستای پایدارسازی سیستم حلقه-بسته و کاهش خطای حالت ماندگار، فراجش^{۱۰}، زمان خیز^{۱۱} و زمان نشست^{۱۲} بهره خواهیم برد. در این راستا، در بخش ۲ معادلات سینماتیکی و دینامیکی ربات استخراج شده و در بخش ۳، نحوه پیاده‌سازی الگوریتم رقابت استعماری برای حصول کنترل‌کننده



شکل ۱: شمایی از ربات اسکارا

بهینه تشریح خواهد شد. نتایج شبیه‌سازی در بخش ۴ بازتاب داده شده است. در بخش ۵ در برگزیده جمع‌بندی از کار حاضر ارائه شده است.

۲-۲- معادلات ربات اسکارا

در شکل ۱ شمایی از ربات مورد مطالعه آورده شده است که بازوهای اصلی و پیش‌رو در صفحه $x - y$ واقع شده است و بازوی کشویی^{۱۳} انتهایی ثابت شده است. در نتیجه حرکتی در راستای محور z خواهیم داشت و توصیف حرکت ربات تنها با استفاد از دو زاویه θ_1 و θ_2 ممکن خواهد بود. پارامترهای توصیف‌کننده این ربات در پیوست الف آورده شده است.

با توجه به هندسه ربات و شکل ۱ موقعیت مجری نهایی^{۱۴} (سینماتیک مستقیم) به صورت زیر قابل بیان خواهد بود:

$$x_e = r_1 \cos \theta_1 + r_2 \cos(\theta_1 + \theta_2) \quad (1)$$

$$y_e = r_1 \sin \theta_1 + r_2 \sin(\theta_1 + \theta_2) \quad (2)$$

همچنین زوایای θ_1 و θ_2 از روی موقعیت مجری نهایی، سینماتیک وارون، به صورت

$$\theta_2 = \arccos \left(\frac{(x_e^2 + y_e^2) - (r_1^2 + r_2^2)}{2r_1 r_2} \right) \quad (3)$$

$$\theta_1 = \operatorname{atan} \left(\frac{-r_2 \sin(\theta_2) x_e + (r_1 + r_2 \cos(\theta_2)) y_e}{r_2 \sin(\theta_2) + (r_1 + r_2 \cos(\theta_2)) x_e} \right) \quad (4)$$

محاسبه می‌شوند.

زمانی که یک محور حرکت می‌کند، گشتاور آن نیز بر روی محور دیگر تأثیر می‌گذارد. این پدیده که تحت عنوان تزویج مکانیکی^۴ شناخته می‌شود. همچنین خطی‌سازی سیستم و استفاده از راهکارهای کنترل خطی نتایج رضایت‌بخشی را در پی نخواهد داشت چراکه با حرکت سریع بازوی تأثیر پدیده تزویج مکانیکی بیشتر می‌شود و از اعتبار مدل خطی می‌کاهد. تنظیم دستی پارامترهای کنترل‌کننده نیز امری بسیار وقت‌گیر خواهد بود. راهکار دیگر پیشنهادی برای تنظیم پارامترهای کنترل PID استفاده از الگوریتم‌های تکاملی و فرا ابتکاری نظیر ژنتیک^۵ [۹-۶]، ازدحام ذرات^۶، [۱۰، ۱۱]، گرگ خاکستری^۷ [۱۲]، وال^۸ [۱۳] و ... است؛ اما در اغلب این پژوهش تمرکز بر روی پایدارسازی سیستم با کمینه کردن معیارهای عملکردی مانند انتگرال مربعات خطا (ISE) و یا انتگرال قدرمطلق خطا (IAE) است که منجر به یک سیستم حلقه-بسته پایدار با پاسخ کند شده‌اند. در این مقاله از الگوریتم رقابت استعماری^۹ (ICA) [۱۴] با پیشنهاد یک تابع هزینه جدید برای حصول پارامترهای کنترل‌کننده‌های PID استفاده شده است.

الگوریتم رقابت استعماری با مجموعه‌ای از راه‌حل‌های اولیه شروع می‌شود که جمعیت نامیده می‌شود. هر فرد در جمعیت به عنوان کشور شناخته می‌شود که توسط یک آرایه واقعی یا بردار نشان داده شده است. تمام کشورهای به دو دسته تقسیم می‌شوند: امپراتور و مستعمره. برخی از بهترین (قدرتمندترین) کشورها به عنوان امپراتورها انتخاب می‌شوند و بقیه نیز مستعمرات خواهند بود. مستعمره‌ها به امپراتورها با توجه به قدرتشان اختصاص می‌یابد. امپراتور و مستعمرات آن یک امپراتوری را تشکیل می‌دهند. قدرت هر امپراتوری بر هر دو قدرت امپراتور و قدرت مستعمرات آن متکی است. قدرت کل امپراتوری برابر است با قدرت امپراتور به علاوه درصد قدرت متوسط مستعمرات آن است. پس از تقسیم تمام مستعمرات در میان استعمارگرها، این مستعمرات شروع به حرکت به طرف کشور استعمارگر مرتبط می‌کنند. بعد از این، رقابت استعماری در میان تمام امپراتوری‌ها به منظور در اختیار گرفتن مستعمرات بیشتر آغاز می‌شود. هر امپراتوری که قادر به موفقیت در این مسابقه نباشد و نتواند قدرت خود را افزایش دهد از رقابت حذف خواهد شد. در رقابت، قدرت امپراتوری‌های قدرتمند به تدریج افزایش می‌یابد، برعکس قدرت ضعیف‌ترها کاهش خواهد یافت. در نتیجه، امپراتوری‌های ضعیف قدرت خود را از دست می‌دهند و در نهایت سقوط خواهند کرد. جنبش مستعمرات به سمت امپراتورهای مربوط به آن‌ها همراه با رقابت میان امپراتوری‌ها است. مکانیسم فروپاشی باعث همگرایی همه کشورهای به حالتی می‌شود که در آن فقط یک امپراتوری در جهان باقی می‌ماند و همه‌ی دیگر کشورها مستعمره‌ی آن امپراتوری هستند. در این جهان ایدئال، هر مستعمره دارای همان موقعیت و قدرت امپراتور است.

⁴ Mechanical Coupling

⁵ Genetic

⁶ Particle Swarm

⁷ Grey Wolf

⁸ Wall

⁹ Imperialist Competitive Algorithm

¹⁰ Overshoot

¹¹ Rise Time

¹² Settling Time

¹³ Prismatic

¹⁴ End Effector

بنابراین با استفاده از قضیه لیاپانوف ثابت می‌شود که $\lim_{t \rightarrow \infty} e_i(t) = 0$ و حالت‌های سیستم θ_i به مقدار مطلوب θ_{d_i} میل خواهند کرد.

۳- طراحی کنترل‌کننده مبتنی بر الگوریتم رقابت استعماری

در الگوریتم رقابت استعماری، هر کشوری یک راه‌حل خاص مسئله بهینه‌سازی است. هر کشور توسط یک آرایه واقعی یا بردار ارائه می‌شود.

برای یک مسئله N بعدی، این آرایه به این صورت تعریف می‌شود

$$\text{country} = [p_1, p_2, \dots, p_{N_{var}}] \quad (۸)$$

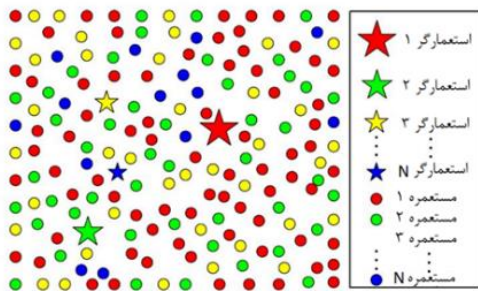
هزینه یک کشور با ارزیابی تابع هزینه متغیرهای $(p_1, p_2, \dots, p_{N_{var}})$ قابل مشاهده است.

$$\text{Cost} = f(\text{country}) = f(p_1, p_2, \dots, p_{N_{var}}) \quad (۹)$$

۳-۱- تولید امپراتوری‌های اولیه

ابتدا، جمعیت N_{pop} در فضای جستجو به طور تصادفی تولید شده.

سپس، N_{imp} از قدرتمندترین کشورها به عنوان امپراتورها و N_{col} از



شکل ۲: شکل‌گیری امپراتوری‌های اولیه

کشورهای باقی‌مانده انتخاب شده که همان مستعمرات خواهند بود و هرکدام متعلق به یک امپراتوری است. برای تشکیل امپراتوری‌های اولیه، به یک امپراتور خاص بر اساس قدرت خود مستعمرات اختصاص داده می‌شود. ابتدا تعداد مستعمره‌های یک امپراتوری متناسب با قدرت آن است. برای تقسیم مستعمرات در میان امپریالیست‌ها، هزینه نرمالیزه شده یک امپراتوری به این صورت تعریف می‌شود:

$$C_n = c_n - \max\{c_i\} \quad i = 1, 2, \dots, N_{imp} \quad (۱۰)$$

جایی که c_n هزینه امپراتوری n -ام و C_n هزینه نرمالیزه شده آن است. سپس قدرت نرمالیزه شده هر امپراتوری به صورت زیر قابل تعریف است:

$$P_n = \left| \frac{C_n}{\sum_{i=1}^{N_{imp}} C_i} \right| \quad (۱۱)$$

قدرت نرمالیزه شده امپراطوری نشان می‌دهد چه تعداد از مستعمرات باید توسط آن امپراتوری تسخیر شوند؛ بنابراین، تعداد اولیه مستعمرات یک امپراتوری برابر خواهد بود با

$$Nc_n = \text{round}\{P_n \cdot N_{col}\}, \quad (۱۲)$$

با فرض آن که از چرخ‌دنده‌ها برای افزایش گشتاور موتور متصل به بازوها به نسبت N_i و تعریف $J_m = J_m + J_{g_i}$ و استفاده از رابطه اویلر-لاگرانژ^{۱۵}

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} = Q_i, \quad i = 1, 2 \quad (۵)$$

که در آن \mathcal{L} متغیر لاگرانژ که تفاضل انرژی جنبشی و انرژی پتانسیل و Q_i نیروهای تعمیم‌یافته هستند معادله دینامیکی ربات اسکارا به صورت زیر خلاصه می‌گردد:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} C_{11} \\ C_{12} \end{bmatrix} = \begin{bmatrix} N_1 T_{m1} \\ N_2 T_{m2} \end{bmatrix} + \begin{bmatrix} T_{d1} \\ T_{d2} \end{bmatrix} \quad (۶)$$

جایی که

$$M_{11} = N_1^2 J_{m1} + J_1 + J_2 + \frac{m_1 r_1^2 + m_2 r_2^2 + 4m_2 r_1^2}{4} + m_2 r_1 r_2 (\cos(\theta_1 + \theta_2) \cos(\theta_1) + \sin(\theta_1 + \theta_2) \sin(\theta_1))$$

$$M_{12} = M_{21} = J_2 + \frac{m_2 r_2^2}{4} + \frac{m_2 r_1 r_2}{2} (\cos(\theta_1 + \theta_2) \cos(\theta_1) + \sin(\theta_1 + \theta_2) \sin(\theta_1))$$

$$M_{22} = N_2^2 J_{m2} + J_2 + \frac{m_2 r_2^2}{4} + \frac{m_2 r_1 r_2}{2} \left(\dot{\theta}_1 \dot{\theta}_2 + \frac{\dot{\theta}_2^2}{2} \right) \times [\cos(\theta_1 + \theta_2) \sin(\theta_1) - \sin(\theta_1 + \theta_2) \cos(\theta_1)]$$

$$C_{22} = \frac{m_2 r_1 r_2}{2} \dot{\theta}_1^2 [\sin(\theta_1 + \theta_2) \cos(\theta_1) - \cos(\theta_1 + \theta_2) \sin(\theta_1)]$$

و T_{d_i} بیانگر دینامیک‌های مدل‌نشده و اغتشاش‌ها وارده به رابط i ام و T_{m_i} گشتاور موتور متصل به رابط i ام است. دقت شود که از آنجایی که مرکز چارچوب جهانی بر روی عملگر بازوی اصلی قرار داده شده است، در معادله (۶) عبارت‌های شامل نیروهای گرانشی ظاهر نشده‌اند.

۲-۶- بیان مسئله کنترلی

هدف مسئله کنترلی طراحی کنترلی PID با ساختار

$$T_{mi} = \frac{1}{N_i} (K_{P_i} e_i + K_{I_i} \int e_i dt + K_{D_i} \dot{e}_i), \quad i = 1, 2 \quad (۷)$$

است به قسمی که برای سیستم مدل‌شده با معادله (۶) تضمین نمایند خطای تعقیب $e_i = \theta_i - \theta_{d_i}$ به ازای مسیر مطلوب θ_{d_i} به سمت صفر میل کند. همچنین سیستم حلقه-بسته از شاخص‌های کارایی مطلوبی نظیر حداقل فراجهد و پاسخ سریع برخوردار باشد.

نکته

با استفاده از تابع مثبت معین شعاعی بی‌کران $v = e_1^2/2 + e_2^2/2$ و مشتق‌گیری از v در امتداد مسیرهای سیستم (۶) می‌توان نشان داد که \dot{v} به ازای $K_{P_i} > 0$ ، $K_{I_i} > 0$ و $K_{D_i} > 0$ یک تابع منفی معین است؛

¹⁵ Euler-Lagrange Equation

هر دهه عددی بین صفر و یک تولید می‌شود، اگر عدد تولید شده کمتر از احتمال وقوع انقلاب باشد، موقعیت آن کشور به صورت تصادفی تغییر خواهد نمود.

۳-۴- تبادل موقعیت‌های استعمارگر و مستعمره

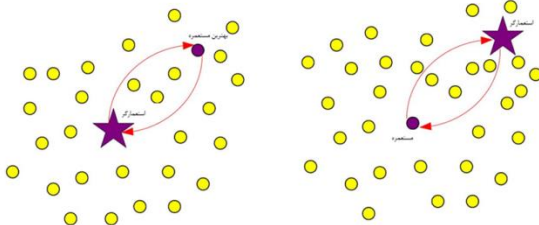
هنگامی که یک مستعمره به یک موقعیت جدید حرکت می‌کند، هزینه آن شاید بهتر از استعمارگر مربوطه گردد. در این حالت، استعمارگر و مستعمره موقعیت خود را معاوضه می‌کنند. این فرایند در شکل ۴ نشان داده شده است. سپس الگوریتم توسط استعمارگر جدید ادامه خواهد یافت و مستعمرات شروع به حرکت به سمت این موقعیت می‌کنند.

۳-۵- محاسبه قدرت کل امپراتوری

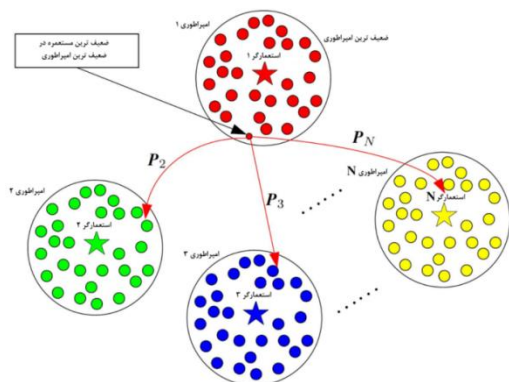
قدرت کل امپراتوری شامل دو بخش است. یکی قدرت استعمارگر و دیگری قدرت مستعمرات آن. در این دو بخش، قدرت استعمارگر تأثیر بیشتری بر قدرت کل دارد. قدرت کل یک امپراتوری به این صورت تعریف می‌شود

$$TC_n = C_n + \xi \cdot \frac{\sum_{i=1}^{NC_n} w_i}{NC_n}, \quad i = 1, 2, \dots, NC_n \quad (13)$$

جایی که TC_n کل هزینه امپراتوری n -ام است و w_i هزینه مستعمرات امپراتوری است. ξ یک عدد مثبت است که کمتر از 1 در نظر گرفته می‌شود. مقادیری بین 0.1 و 0.5 به ازای ξ می‌تواند نتایج رضایت بخشی را در بسیاری از مسائل به ارمغان آورد.



شکل 4: مبادله موقعیت بین استعمارگر و مستعمره



شکل 5: رقابت بین امپراتوری‌ها برای بدست آوردن ضعیف‌ترین مستعمره از ضعیف‌ترین امپراتوری

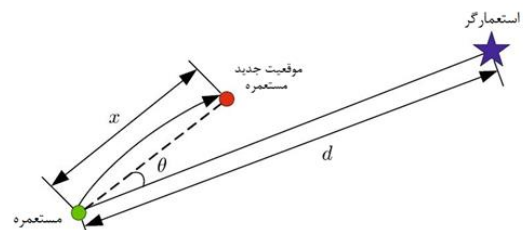
جایی که NC_n تعداد مستعمره اولیه امپراتوری n -ام است، N_{col} تعداد کل مستعمرات است. NC_n مستعمره به طور تصادفی انتخاب شده و به امپراطوری n -ام اختصاص داده می‌شود. این مستعمره‌ها با استعمارگر امپراتوری n -ام را تشکیل می‌دهند. از مکانیسم تقسیم مستعمره، امپراتوری‌های قوی‌تر تعداد بیشتری از مستعمره‌ها را دارا می‌گردند. شکل ۲ جمعیت اولیه هر امپراتوری را نشان می‌دهد. در این شکل، استعمارگر ۱ دارای بیشترین تعداد مستعمرات و قدرتمندترین است.

۳-۲- سیاست جذب

بعد از این که امپراتوری شکل گرفت، استعمارگر در هر امپراتوری تلاش می‌کند مستعمرات خود را بهبود بخشد. در الگوریتم ICA، این امر با حرکت همه مستعمره‌ها به سمت استعمارگر به دست می‌آید. جهت حرکت یک بردار از مستعمره به استعمارگر است. این حرکت در شکل ۳ نشان داده شده است که در آن مستعمره با واحدهای x به سمت استعمارگر حرکت می‌کند. x عدد تصادفی با توزیع یکنواخت به صورت زیر است:

$$x \sim U(0, \beta \times d), \quad (11)$$

که در آن d فاصله بین استعمارگر و مستعمره است، β یک عدد بزرگ‌تر از یک است. به منظور جستجوی نقطه‌هایی در اطراف



شکل 3: حرکت مستعمره به سمت استعمارگر

استعمارگر یک انحراف به جهت حرکت اضافه شده است. در شکل ۳، θ نشان‌دهنده این انحراف است. همچنین θ یک عدد تصادفی است با توزیع یکنواخت به صورت

$$\theta \sim U(-\gamma, \gamma), \quad (12)$$

که γ یک پارامتر تنظیم انحراف از جهت اصلی است. مقدار β و γ اختیاری است. مقداری در حدود 2 برای β و $\pi/4$ برای γ نتیجه مناسب در همگرایی پاسخ‌ها به کمینه سرتاسری ارائه می‌دهد.

۳-۳- انقلاب

بروز انقلاب تغییرات ناگهانی را در ویژگی‌های یک کشور ایجاد می‌کند. در الگوریتم رقابت استعماری، انقلاب با جابجایی تصادفی یک کشور مستعمره به یک موقعیت تصادفی جدید مدل‌سازی می‌شود. انقلاب از دیدگاه الگوریتمی باعث می‌شود کلیت حرکت تکاملی از گیر کردن در ده‌های محلی بهینگی نجات یابد که در بعضی موارد باعث بهبود موقعیت یک کشور شده و آن را به یک محدوده بهتری می‌برد. برای مدل‌سازی این پدیده سیاسی، به ازای هر کشور به صورت تصادفی در

۳-۶- رقابت بین استعمارگرها

رقابت بین استعمارگرها بین امپراتوری‌ها اتفاق می‌افتد چراکه هر امپراتوری خواهان داشتن مستعمره‌های دیگر امپراتوری و کنترل آن‌ها است. رقابت منجر به افزایش قدرت امپراتوری‌های قوی‌تر و کاهش قدرت ضعیف‌ترها می‌شود. در ICA، این رقابت با انتخاب یکی از ضعیف‌ترین مستعمرات از ضعیف‌ترین امپراتوری‌ها و رقابت بین همه امپراتوری‌ها برای داشتن این مستعمرات اجرا می‌شود. این روش در شکل ۵ نشان داده شده است.

در این رقابت، هر امپراتوری شانس داشتن مستعمرات را با توجه به قدرت کل آن دارد. احتمال تسخیر مستعمرات توسط امپراتوری به صورت زیر تعریف می‌شود

$$P_{P_n} = \left| \frac{NTC_n}{\sum_{i=1}^{N_{imp}} NTC_i} \right|^2 \quad (14)$$

جایی که N_{imp} تعداد کشورهای استعمارگر است، NTC_n هزینه کلی نرمالایز شده امپراتوری n -ام است که برابر با

$$NTC_n = TC_n - \max\{TC_i\}, i = 1, 2, \dots, N_{imp} \quad (15)$$

محاسبه می‌شود. به منظور تقسیم مستعمره‌های ذکر شده در امپراتوری‌ها مبتنی بر احتمال مالکیت آنها، ما بردار P را به شکل زیر تشکیل می‌دهیم

$$P = [P_{P1}, P_{P2}, \dots, P_{P_{imp}}] \quad (16)$$

سپس یک بردار R با همان اندازه بردار P به صورت

$$R = [r1, r2, \dots, rN_{imp}], r_i \sim U(0,1) \quad (17)$$

تعریف و بردار D به صورت زیر بدست می‌آید

$$D = P - R \quad (18)$$

امپراتوری که شاخص مربوطه آن در D حداکثر باشد ضعیف‌ترین مستعمره ضعیف‌ترین امپراتوری را اشغال خواهد کرد.

۳-۷- از بین بردن امپراتوری‌های بی‌قدرت

امپراتوری ضعیف در رقابت‌های استعمارگری سقوط خواهد کرد و مستعمرات آن‌ها در میان دیگر امپراتوری‌ها تقسیم خواهد شد. در مدل‌سازی مکانیسم سقوط، معیارهای مختلف می‌تواند برای در نظر گرفتن امپراتوری بی‌قدرت تعریف شود. به عنوان مثال می‌توان فرض نمود که یک امپراتوری زمانی که تمام مستعمراتش را از دست دهد فرو می‌ریزد و از بین می‌رود.

۳-۸- همگرایی

بعد از مدتی تمام امپراتوری‌ها به جز قدرتمندترین سقوط خواهند کرد و تمام مستعمرات تحت کنترل این امپراتوری منحصربه‌فرد خواهد بود. در این جهان ایدئال، تمام مستعمرات همراه با استعمارگر منحصربه‌فرد موقعیت و هزینه‌های مشابه خواهند داشت. در چنین شرایطی، رقابت متوقف شده و الگوریتم خاتمه یافته است.

۳-۹- نحوه استفاده از الگوریتم رقابت استعماری

برای تعیین ضرایب کنترل‌کننده‌های PID

در کنترل‌کننده‌های PID مورد استفاده برای کنترل ربات، ضریب تناسبی K_{P_i} ، ضریب انتگرالی K_{I_i} و ضریب مشتقی K_{D_i} به نحو مناسبی اختیار گردد که تضمین عملکرد مطلوبی را ارائه دهند، بنابراین، هر کشور در ICA به عنوان یک بردار شش بعدی مانند زیر نمایان می‌شود

$$Country = [K_{P1}, K_{I1}, K_{D1}, K_{P2}, K_{I2}, K_{D2}] \quad (19)$$

که دامنه کاوش برای همه این پارامتر بازوی $[0, 100]$ در نظر گرفته شده است.

۳-۱۰- تابع هزینه

به طور کلی، چند معیار عملکرد مبتنی بر خطای تعقیب مانند ISE و IAE می‌تواند به عنوان تابع هزینه برای تنظیم پارامترهای کنترل‌کننده PID استفاده می‌شود. با این حال، عملکرد معیارهای فوق دارای معایبی هستند. به عنوان مثال، به حداقل رساندن IAE و ISE ممکن است زمان نشست را طولانی کند.

در این مقاله، تابع هزینه به صورت

$$J = (1 - e^{-\alpha}) \sum_{i=1}^2 (M_{pi} + E_{ssi}) + e^{-\alpha} \sum_{i=1}^2 (t_{si} + t_{ri}) \quad (20)$$

در نظر گرفته شده است که در آن α یک عامل وزن، M_{pi} ، t_{ri} ، t_{si} و E_{ssi} به ترتیب فراجش، زمان خیز، زمان نشست و خطای حالت ماندگار است. با انتخاب مقداری کوچک α پاسخی سریع‌تر و مقداری بزرگ α پاسخ با فراجش کوچکتر و دقت بالا را نتیجه می‌دهد.

۴- نتایج شبیه‌سازی

در این مقاله تعداد کشورها برابر 60، تعداد امپراتوری‌های اولیه 20 و شرط خاتمه گذشت 100 دهه اختیار گردید. مدل ربات در محیط سیمولینک نرم‌افزار متلب مدل‌سازی گردید. در این مرحله فرض گردید که هدف رسیدن هر دو زاویه θ_1 و θ_2 به مقدار 30° است. از روش ODE3 با گام ثابت 0.01 برای شبیه‌سازی استفاده شد و در هر بار اجرای اسکریپت فایل مربوط به کدهای ICA و حصول ضرایب PID جدید، شبیه‌سازی به ازای این پارامترها انجام، تابع هزینه (۲۰) ارزیابی و مجدد به اسکریپت فایل ارسال می‌گشت.

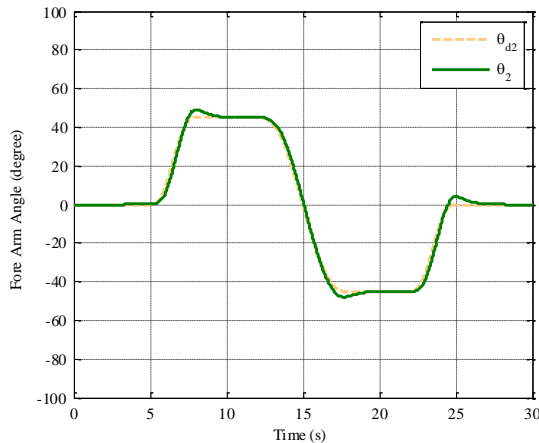
تابع هزینه مربوط به این الگوریتم در شکل ۶ و بهترین پاسخ‌ها بدست آمده در شکل ۷ ترسیم شده است. ضرایب کنترل‌کننده بهینه برابر با

$$K_{P1} = 100, K_{I1} = 0, K_{D1} = 24.0777 \quad (21)$$

$$K_{P2} = 100, K_{I2} = 0, K_{D2} = 88.3543$$

محاسبه می‌گردند. همچنین خلاصه‌ای از معیارهای عملکردی در جدول ۲ آورده شده است.

نتایج بدست آمده حاکی از عملکرد مطلوب ICA در تعیین پارامترهای کنترل‌کننده می‌باشد. پاسخ سیستم‌ها با فراجش ناچیز و سرعت بالا فرمان ورودی را بدون خطا دنبال می‌کند. برای ارزیابی بهتر

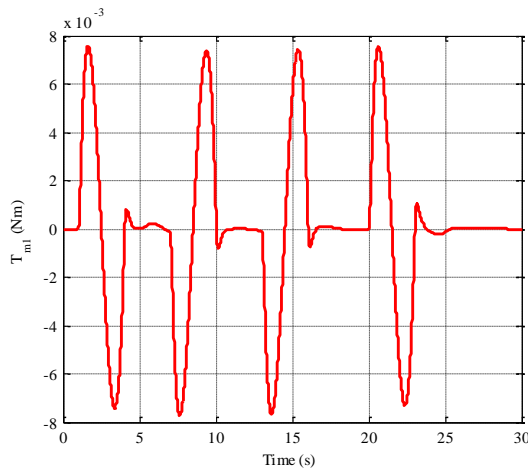


(ب)

شکل 8: خروجی ربات اسکارا با کنترل‌کننده بهینه مبتنی بر ICA به همراه مسیر مطلوب

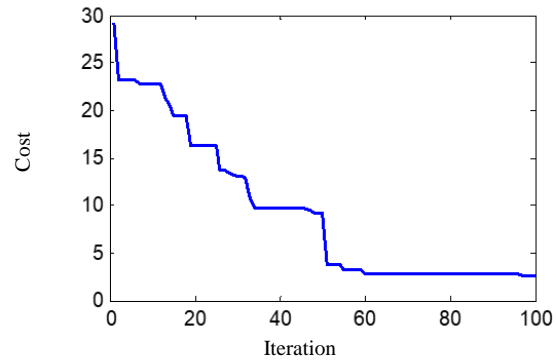
۵- نتیجه‌گیری

در این مقاله پایدارسازی یک بازوی ماهر سری صنعتی تحت عنوان ربات اسکارا مورد مطالعه قرار گرفت. برای پایدارسازی از راهکار کنترل‌کننده PID بهره گرفته شد. با عنایت به دینامیک غیرخطی ربات و چند ورودی-چند خروجی بودن آن حصول پارامترهای مناسب با استفاده از قواعد تجربی امری آسان نبود؛ لذا خواسته‌های مسئله، تعقیب فرمان ورودی، افزایش سرعت و کاهش فرجهش و خطای حالت ماندگار، به شکل یک تابع هزینه ترجمه گردید و طراحی کنترل‌کننده به یک مسئله بهینه‌سازی مبدل گشت. از الگوریتم رقابت استعماری برای حل این مسئله چند هدفه بهره گرفته شد. نتایج بدست آمده ضمن برخورداری از معیارهای کارایی مطلوب، تضمین می‌نمودند که دامنه گشتاور عملگرها از دامنه خود تجاوز نمایند. همچنین این کنترل‌کننده به خوبی توانستند بر اثر تزویج مکانیکی غلبه نمایند و کارایی الگوریتم ICA در تنظیم پارامترهای PID با شبیه‌سازی‌های که انجام گرفت به اثبات رسید.

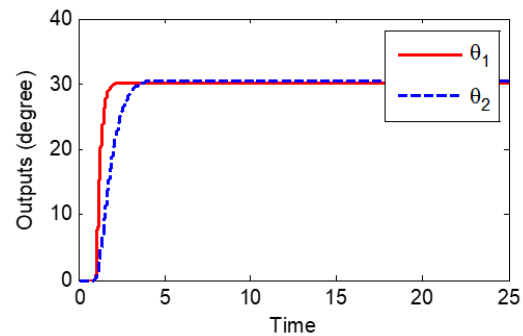


(الف)

کارایی کنترل‌کننده‌های پیشنهادی عملکرد آن‌ها در تعقیب متغیر با زمان در دستور کار قرار گرفت. زوایای ربات به همراه مسیر مطلوب در شکل ۸، گشتاورهای تولیدشده توسط موتورها در شکل ۹ و موقعیت مجری نهایی در شکل ۱۰ بازتاب داده شده‌اند که نتایج حاکی از عملکرد مطلوب کنترل‌کننده‌ها دارد.



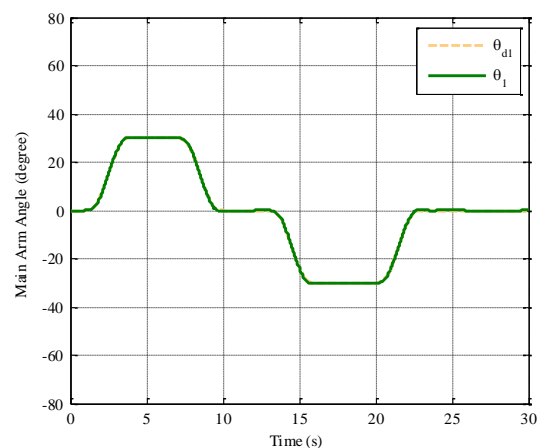
شکل ۶: تابع هزینه بعد گذشت ۱۰۰ دهه



شکل 7: خروجی ربات اسکارا به ازای ورودی پله با دامنه 30° با استفاده از کنترل‌کننده PID مبتنی بر ICA

جدول 4: شاخص‌های کارایی به ازای ورودی پله با دامنه 30°

خروجی	T_r (s)	T_s (s)	M_p	E_{ss}
θ_1	0.3852	1.6488	0.0342	0
θ_2	1.4086	3.28	0.0955	0

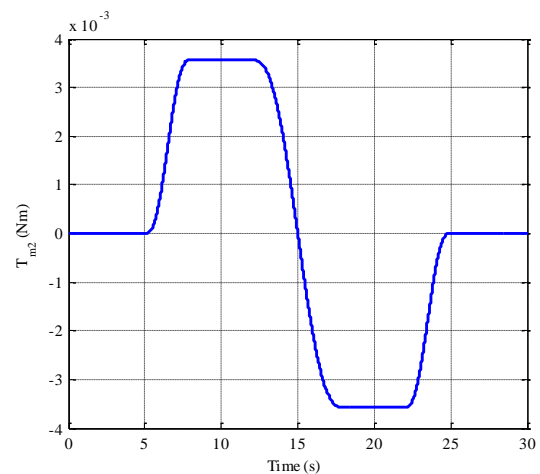


(الف)

200 mm	طول رابط پیش‌رو	r_2
$T_{mi} \leq 1.0$	گشتاور اعمالی از سمت موتور	T_{mi}

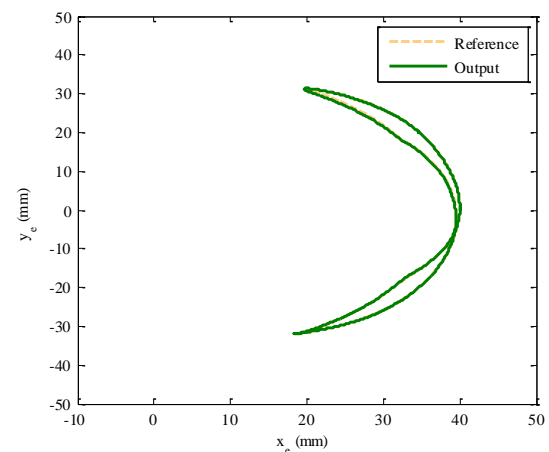
مراجع

- [1] Chi, J., H. Yu, and J. Yu, Hybrid tracking control of 2-DOF SCARA robot via port-controlled hamiltonian and backstepping. *IEEE Access*, 2018, **6**: p. 17354-17360.
- [2] Bauer, M., et al., The current state of control loop performance monitoring—A survey of application in industry. *Journal of Process Control*, 2016, **38**: p. 1-10.
- [3] Samad, T., A survey on industry impact and challenges thereof [technical activities]. *IEEE Control Systems Magazine*, 2017, **37**(1): p. 17-18.
- [4] Dotoli, M., et al., Advanced control in factory automation: a survey. *International Journal of Production Research*, 2017, **55**(5): p. 1243-1259.
- [5] Ribeiro, J., et al. Comparison of PID controller tuning methods: analytical/classical techniques versus optimization algorithms. in 2017 18th international Carpathian control conference (ICCC). 2017. IEEE.
- [6] Jiménez, R.L., A.G. Aguilar, and V.L.G. De Velasco, Close loop step test used for tuning PID controller by genetic algorithms. *Pistas Educativas*, 2018, **36**(112).
- [7] Harrag, A. and S. Messalti, Variable step size modified P&O MPPT algorithm using GA-based hybrid offline/online PID controller. *Renewable and Sustainable Energy Reviews*, 2015, **49**: p. 1247-1260.
- [8] Rout, M., et al. PID controller design for cruise control system using genetic algorithm. in 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). 2016. IEEE.
- [9] Vaishnavi, P. and G. Balasubramanian, Tuning of PID Controller for Cascade Unstable systems Using Genetic Algorithm. *International Journal of Pure and Applied Mathematics*, 2018, **118**(18): p. 2011-2019.
- [10] Mousa, M., M.A. Ebrahim, and M.M. Hassan, Stabilizing and swinging-up the inverted pendulum using PI and PID controllers based on reduced linear quadratic regulator tuned by PSO. *International Journal of System Dynamics Applications (IJSDA)*, 2015, **4**(4): p. 52-69.
- [11] Jain, R.V., M. Aware, and A. Junghare, Tuning of fractional order PID controller using particle swarm optimization technique for DC motor speed control. in 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES). 2016. IEEE.
- [12] Moura Oliveira, P., H. Freire, and E.S. Pires, Grey wolf optimization for PID controller design with prescribed robustness margins. *Soft Computing*, 2016, **20**(11): p. 4243-4255.
- [13] Hasanien, H.M., Whale optimisation algorithm for automatic generation control of interconnected modern power systems including renewable energy sources. *IET Generation, Transmission & Distribution*, 2017, **12**(3): p. 607-614.
- [14] Atashpaz-Gargari, E. and C. Lucas, Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. in 2007 IEEE congress on evolutionary computation. 2007. IEEE.
- [15] Das, M.T. and L. Canan Dülger, Mathematical modelling, simulation and experimental verification of a scara robot. *Simulation Modelling Practice and Theory*, 2005, **13**(3): p. 257-271.



(ب)

شکل 9: سیگنال‌های کنترلی: (الف) گشتاور تولیدشده توسط موتور متصل به بازوی اصلی، (ب) گشتاور موتور متصل به بازوی پیش‌رو



شکل 10: موقعیت مجری نهایی به همراه موقعیت مطلوب

پیوست الف: پارامترها و مقادیر نامی ربات اسکارا [۱۵]

مقدار نامی	توصیف	پارامتر
0.098 Kgm^2	ممان اینرسی رابط اصلی	J_1
0.0115 Kgm^2	ممان اینرسی رابط پیش‌رو	J_2
3.3	ممان اینرسی موتور	J_m
$\times 10^{-6} Kgm^2$	ممان اینرسی چرخ‌دنده موتور رابط اصلی	J_{g1}
0.002 Kgm^2	ممان اینرسی چرخ‌دنده متصل به موتور رابط پیش‌رو	J_{g2}
0.0005 Kgm^2	جرم رابط اصلی	m_1
1.90 Kg	جرم رابط پیش‌رو	m_2
0.93 Kg	نسبت چرخ‌دنده موتور رابط اصلی	N_1
90	نسبت چرخ‌دنده موتور رابط پیش‌رو	N_2
220	طول رابط اصلی	r_1
150 mm		



طراحی کنترل‌کننده تناسبی-انتگرالی-مشتقی بهینه برای سیستم گوی معلق به کمک الگوریتم ژنتیک

بهنام شجاعی، علیرضا صفا^۱، احمد بادلی، احمد میرتقی، پدram سرمدی

^۱ استادیار گروه مهندسی برق، دانشکده مهندسی، دانشگاه گلستان، گرگان، ایران.

آدرس پست الکترونیکی: a.safa@gu.ac.ir

چکیده

این مقاله راهکاری برای بهینه‌سازی ضرایب یک کنترل‌کننده تناسبی-انتگرالی-مشتقی (PID) با استفاده از الگوریتم ژنتیک با تمرکز به بهبود معیارهای کارایی حوزه زمان ارائه می‌دهد. سیستم مورد مطالعه در این پژوهش معلق‌سازی گوی با جریان هوا است. مسئله کنترلی، ثابت نگهداشتن گوی معلق داخل نازل در ارتفاع مشخص و یا تعقیب مرجع از پیش‌تعریف‌شده است. برای نیل به این مقصود، ابتدا معادله‌های دینامیکی سیستم گوی معلق بر اساس قوانین فیزیکی حاکم بر آن استخراج شده و یک تقریب خطی از سیستم ارائه می‌گردد. پارامترهای نامعلوم شناسایی شده و از الگوریتم ژنتیک برای تنظیم پارامترهای بهینه کنترل‌کننده PID کمک گرفته می‌شود. با انتخاب تابع هزینه مناسب ضرایب کنترل‌کننده به ازای کمترین فراجش و کمترین زمان نشست حاصل می‌شود. برای نشان دادن کارایی راهکار پیشنهادی مقایسه‌ای بین پاسخ سیستم‌ها حلقه-بسته با کنترل‌کننده بهینه حاصل از الگوریتم ژنتیک و رویکرد زیگلر-نیکولز شبیه‌سازی شده و از پلنت گوی معلق شرکت مهندسی کیا الکترونیک برای صحت‌سنجی نتایج به طور عملی استفاده گردید.

کلمات کلیدی

بهینه‌سازی، الگوریتم ژنتیک، کنترل‌کننده PID، معیارهای کارایی حوزه زمان، زیگلر-نیکولز.

۱- مقدمه

کنترل‌کننده تناسبی-انتگرالی-مشتقی (PID) یک راهکار کنترلی فراگیر است که در صنایع متفاوت مانند نفت و گاز، شیمیایی، پتروشیمی، کاغذ، غذا و ... نمود پیدا می‌کند [۱-۳]. این کنترل‌کننده قابلیت اطمینان و پایداری خود را در کنترل متغیرهای فرآیندهای مختلف مانند کنترل درجه حرارت، سطح، فشار، جریان، PH و ... ثابت کرده است [۴-۶]. عوامل دیگری که صنایع را برای انتخاب کنترل‌کننده PID جذب می‌کند می‌توان به کم‌هزینه بودن، نگهداری آسان و همچنین سادگی در ساختار کنترلی و فهم آسان آن عنوان نمود. با این حال، تنظیم نامناسب پارامترهای PID می‌تواند منجر به نوسانی و آهسته شدن پاسخ سیستم، پایداری ضعیف و در بدترین حالت ممکن از بین رفتن پایداری سیستم گردد [۷]. این امر باعث شده تا پژوهشگران در جستجوی روشی برای یافتن مناسب‌ترین پارامترهای

PID باشند. زیگلر-نیکولز^۱ [۸] و کهن-کون^۲ [۹] از جمله پیشگامان روش‌های تنظیم PID هستند. آن‌ها روش‌های تنظیم تجربی PID را بر اساس روش آزمون و خطا و منحنی پاسخ فرآیند پیشنهاد کرده‌اند. با این وجود، دشواری تنظیمات PID زمانی که سیستم پیچیده می‌شود مثلاً در سیستم‌های با مرتبه بالا، سیستم‌های تأخیردار، سیستم‌های غیرکمیته فاز و فرآیندهای غیرخطی، ممکن است افزایش یابد. به عنوان مثال، روش زیگلر-نیکولز ممکن است فراجش بالا، نوسانات زیاد و زمان نشست بیشتری را برای یک سیستم مرتبه بالا نتیجه دهد [۱۰] و روش کهن-کون فقط برای سیستم‌های دارای پاسخ پله S شکل راهکار ارائه می‌دهد [۹]. همچنین در اغلب راهکارهای ارائه‌شده تمرکزی بر روی معیارهای کارایی حوزه زمان نظیر فراجش، زمان خیز و زمان نشست نشده است. برای غلبه بر این مشکل‌ها در این پژوهش عملکرد

¹ Ziegler-Nichols

² Cohen-Coon

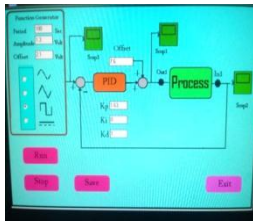
۳-۱- توصیف ریاضی سیستم

با استفاده از قانون دوم نیوتون معادله دینامیکی برای سیستم معلق‌سازی با جریان هوا قابل استخراج است. از آنجایی که تنها نیروهای وارد بر جسم معلق تأثیر جریان باد بر بالا بردن جسم و تأثیر نیروی جاذبه در جهت پایین هستند، معادله دینامیکی سیستم به صورت زیر بدست می‌آید:

$$m\ddot{z}(t) = \frac{1}{2} C_d \rho A (v_w - \dot{z}(t))^2 - mg \quad (2)$$



(ب)



(ج)



(الف)

شکل ۱۱: سیستم گوی معلق: (الف) گوی، دمنده و حسگر التراسونیک، (ب) مدار واسط بین کامپیوتر و سیستم گوی، (ج) واسط نرم‌افزاری برای تنظیم پارامترهای کنترل‌کننده و مرجع

که در آن m جرم جسم معلق، z ارتفاع جسم معلق درون لوله، ρ چگالی هوای داخل لوله، A مساحت جسم در معرض جریان هوای رو به بالا، v_w سرعت هوای داخل لوله، g شتاب گرانشی و C_d ضریب پسار^۴ است.

با تعریف $\alpha = 1/2 C_d \rho A$ معادله (۲) به صورت زیر بازنویسی می‌گردد:

$$\ddot{z}(t) = \frac{\alpha}{m} [(v_w - \dot{z}(t))^2 - g]. \quad (3)$$

نقطه تعادل این سیستم با در نظر گرفتن $\dot{z}(t) = \ddot{z}(t) = 0$ بدست می‌آید که برای این منظور سرعت هوا برای رسیدن به تعادل که با v_{eq} نمایش داده شده است، برابر خواهد بود با:

$$v_{eq}^2 = \frac{mg}{\alpha}. \quad (4)$$

معادله (۳) با استفاده از (۴) به صورت زیر بازنویسی می‌گردد:

$$\ddot{z} = g \left(\left(\frac{v_w - \dot{z}}{v_{eq}} \right)^2 - 1 \right). \quad (5)$$

با تعریف $x := \frac{v_w - \dot{z}}{v_{eq}}$ سمت راست تساوی (۵) به صورت

$$f(x) := g(x^2 - 1) \quad (6)$$

محاسبه می‌شود که با استفاده از تقریب تیلور حول نقطه $x = 1$ به صورت

الگوریتم ژنتیک (GA) [۱۱] در تنظیم پارامترهای PID بهینه برای بهبود معیارهای کارایی حوزه زمان مورد بررسی قرار گرفته است. عملکرد سیستم با استفاده از روش‌های تنظیم GA و روش زیگلر-نیکولز برای نشان دادن کارایی مورد مقایسه قرار گرفته است. در این راستا، این مقاله به صورت زیر سازمان یافته است.

در بخش ۲، مروری بر ساختار کنترل‌کننده خواهیم داشت. سیستم مورد مطالعه در بخش ۳ معرفی شده، معادلات ریاضی توصیف‌کننده آن استخراج و پارامترهای نامعلوم آن توسط آزمایش عملی حاصل می‌گردد. راهکار کنترلی پیشنهادی در بخش ۴ گنجانده شده است. نتایج شبیه‌سازی و آزمایش عملی در بخش ۵ بازتاب داده شده است و جمع‌بندی از کار حاضر در بخش ۶ ارائه گشته است.

۲- کنترل‌کننده PID

کنترل‌کننده PID از سه پارامتر مجزا تشکیل شده است: پارامتر تناسبی، انتگرالی و مشتقی که توسط سه متغیر K_p ، K_I و K_D مشخص می‌شوند. تنظیم این پارامترها پاسخ دینامیکی سیستم را بهبود داده و باعث کاهش فرآهش، از بین رفتن خطای حالت ماندگار و افزایش پایداری سیستم می‌گردد [۴]. تابع تبدیل یک کنترل‌کننده PID به صورت زیر در دسترس است:

$$C(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_I}{s} + K_D s. \quad (1)$$

زمانی که نقطه تنظیم تغییر کند، مقدار خطا بین نقطه تنظیم شده^۳ و مقدار خروجی واقعی محاسبه می‌شود. از سیگنال خطا $E(s)$ برای ایجاد اعمال تناسبی، انتگرالی و مشتقی استفاده می‌گردد. با وزن دهی به سیگنال‌های حاصل و جمع آن‌ها در کنترل‌کننده PID، سیگنال کنترلی $U(s)$ شکل می‌گیرد که به سیستم اعمال می‌شود. سپس سیگنال خروجی جدید بدست آمده و به کنترل‌کننده ارسال می‌شود و دوباره مقدار خطا محاسبه می‌شود و سیگنال کنترلی جدیدی به سیستم اعمال می‌شود و این چرخه تا صفر شدن خطای حالت ماندگار ادامه می‌یابد.

۳- سیستم گوی معلق

نمایی از سیستم مورد مطالعه در شکل ۱ بازتاب داده شده است. این سیستم از یک موتور DC به عنوان دمنده، یک گوی و یک لوله شفاف به ارتفاع ۵۰ سانتی‌متر تشکیل شده است. یک حسگر اولتراسونیک در انتهای لوله شفاف قرار گرفته که ارتفاع توپ را می‌سنجد. موقعیت اندازه‌گیری شده توسط مدار واسط به کامپیوتر ارسال شده و وارد نرم‌افزار مربوطه شده محاسبات مربوط به تولید سیگنال کنترلی در کامپیوتر پردازش و توسط مدار واسط به سیستم کنترلی ارسال می‌گردد. سرعت موتور DC توسط ولتاژ اعمالی به آن قابل کنترل است. در واقع ارتفاع توپ را سرعت چرخش موتور DC تعیین می‌کند. هدف این سیستم معلق نگه داشتن توپ در یک ارتفاع معین و یا دنبال کردن یک ورودی معین است.

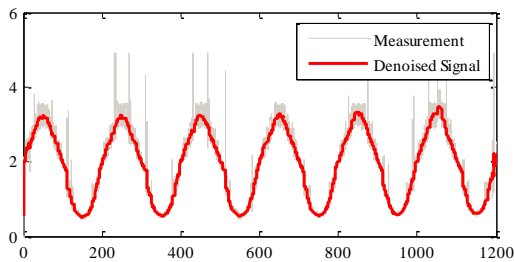
⁴ Drag Coefficient

³ Set point

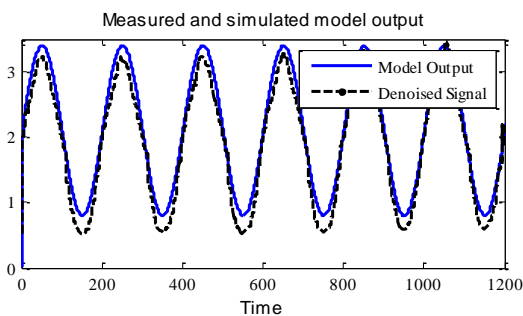
۴- طراحی کنترل‌کننده PID مبتنی بر الگوریتم ژنتیک

۴-۱- الگوریتم ژنتیک

در دهه شصت قرن بیستم، جان هالند الگوریتم ژنتیک را به عنوان یک روش بهینه‌سازی معرفی کرد [۶]. این الگوریتم یک روش بهینه‌سازی ابتکاری است که از مکانیسم انتخاب طبیعی که توسط چارلز داروین در کتاب منشأ انواع معرفی شد الهام گرفته شده است [۷]. این الگوریتم با تولید یک جمعیت تصادفی از کروموزوم‌ها که هرکدام نمایانگر یک جواب مسئله هستند، کار خود را آغاز می‌کند.



شکل 12: نمونه‌ای از خروجی مورد استفاده برای شناسایی با حذف نویز



شکل 13: سیگنال خروجی شبیه‌سازی شده و سیستم واقعی

پس این کروموزوم‌ها مورد ارزیابی قرار می‌گیرند و کارآمد بودن آن‌ها توسط تابع برازندگی ارزیابی می‌شود. بر اساس برازندگی هر کروموزوم و یک احتمال تعریف شده، یک گروه از جواب‌ها برای اعمال اپراتورهای ترکیب^۵ و جهش^۶ انتخاب می‌شوند. ترکیب اپراتوری است که از جواب‌های انتخاب شده به عنوان والد استفاده کرده و جواب‌های جدیدی را به عنوان فرزند به وجود می‌آورد. همچنین اپراتور جهش با ایجاد تغییرات تصادفی در تعدادی از جواب‌های والد، فرزندان جهش‌یافته‌ای را تولید می‌کند. این فرایندها به وجود آمدن جواب‌های کارآمدتر و یا جستجوی بیشتر فضای پاسخ را باعث می‌شوند. در نهایت پاسخ‌های ناکارآمد کنار گذاشته شده و این سه اپراتور، انتخاب، ترکیب و جهش، تا برقرار شدن شرط خاتمه تکرار می‌شوند.

$$f(x) = f(1) + \frac{df(1)}{dt}(x - 1) \quad (7)$$

خطی‌سازی می‌شود. پس داریم:

$$\ddot{z}(t) = 2g(x - 1) = \frac{2g}{v_{eq}}(v_2 - \dot{z} - v_{eq}) \quad (8)$$

بنابراین با تعریف $a = 2g/v_{eq}$ تابع تبدیل سیستم حول نقطه تعادل به صورت زیر محاسبه می‌گردد:

$$\frac{\Delta z(s)}{\Delta v(s)} = \frac{1}{s} \frac{a}{s + a} \quad (9)$$

که در آن Δz و Δv به ترتیب تغییرات ارتفاع توپ و تغییرات سرعت باد در نزدیکی نقطه تعادل است.

با مدل‌سازی دمنده توسط یک سیستم مرتبه اول، تابع تبدیل آن بین ولتاژ ورودی و سرعت باد به صورت زیر قابل توصیف خواهد بود:

$$\frac{\Delta v(s)}{u(s)} = \frac{k_v}{\tau s + 1} \quad (10)$$

K_v بهره ولتاژ ورودی u به سرعت باد و τ ثابت زمانی دمنده است.

در نهایت تابع تبدیل کل فرایند به صورت زیر محاسبه می‌شود:

$$G(s) = \frac{1}{s} \frac{ak_v}{(s + a)(\tau s + 1)} \quad (11)$$

۳-۲- شناسایی سیستم

با عنایت به مدل سیستم توصیف شده در (۱۱) شامل یک قطب در مبدأ است، سیستم حلقه-باز پایدار مرزی است؛ بنابراین پاسخ پله سیستم حلقه باز را به بالاترین ارتفاع لوله می‌رساند و شناسایی صحیح سیستم حلقه باز را ناممکن می‌سازد. برای فائق آمدن بر این چالش، شناسایی سیستم حلقه-بسته پایدار به ازای $k_p = 2$ به وسیله ابزار شناسایی سیستم در نرم‌افزار متلب در دستور کار قرار گرفت.

ولتاژ آستانه‌ای که سرعت باد نقطه تعادل را نتیجه می‌داد با آزمایش مقدار 7.6 ولت مشخص گردید. سیگنال‌های سینوسی با دامنه 1.3 و آفست 2.1 با دوره تناوب‌های مختلف از 10 تا 250 ثانیه به ورودی سیستم اعمال شد. نمونه‌گیری‌ها از داده‌های ورودی حداقل به ازای 5 برابر دوره تناوب صورت پذیرفت. دقت شود در سیستم مورد مطالعه هر ولتاژ ورودی مبین یک ارتفاع مشخصی از توپ است.

جهت شناسایی بهتر سیستم نویز از سیگنال خروجی آزمایش حذف و از سیگنال فیلتر شده جهت شناسایی سیستم استفاده شد که نمونه‌ای از سیگنال به همراه نمونه فیلتر شده آن در شکل ۳ نشان داده شده است.

در نهایت تابع تبدیل سیستم حلقه-بسته با استفاده از ورودی‌ها و خروجی‌های اندازه‌گیری شده به صورت زیر شناسایی شد:

$$G(s) = \frac{10.57}{s^3 + 13.24s^2 + 12.73s + 10.57} \quad (12)$$

برای صحت سنجی خروجی سیستم شناسایی در نرم‌افزار متلب و خروجی سیستم واقعی در شکل ۴ نمایش داده شده است که شباهت بالای 84% برای این دو سیگنال است. در نهایت سیستم حلقه باز برابر با $G(s) = \frac{0.4151}{s(0.82s+1)(0.958s+1)}$ بدست خواهد آمد.

⁵ Crossover

⁶ Mutation

۴-۲- استفاده از الگوریتم ژنتیک برای بدست آوردن

ضرایب PID

در این پژوهش تعداد جمعیت اولیه، کروموزوم‌ها 50 در نظر گرفته شد. همچنین نرخ ترکیب 0.7، نرخ جهش 0.2 و 100 بار تکرار الگوریتم که به عنوان شرط خاتمه انتخاب شده است. هر عضو جمعیت اولیه به صورت تصادفی در یک آرایه با سه عضو که توسط یک ماتریس 3×1 تعریف می‌شود، به وجود می‌آید. این اعداد مقداری بین (۰, ۱) دارند. هر عضو این ماتریس توسط تابع هزینه فراخوانی شده و عضو اول با ضریب 20 مقدار K_p ، عضو دوم با ضریب 10 مقدار K_i و عضو سوم با ضریب 1 مقدار K_d را تعیین می‌کند. این ضرایب بر اساس محدودیت‌های سیستم مورد مطالعه بدست آمده‌اند.

ورودی پله به سیستم داده شده و خروجی به ازای قانون PID با ضرایب بدست آمده از مرحله قبل تعیین می‌گردد. با تحلیل خروجی مقدار فراجاهش، زمان نشست و سایر پارامترهای حوزه زمان در دسترس خواهد بود. تابع هزینه به صورت زیر در نظر گرفته شده است:

$$J = \begin{cases} \omega_1 \times T_s + \omega_2 & \text{در صورت پایداری} \\ \times OV, & \\ 20000, & \text{در غیر این صورت} \end{cases} \quad (12)$$

یعنی در صورت که سیستم حلقه بسته ناپایدار باشد هزینه خیلی بزرگ به آن PID تعلق می‌گیرد و در صورت پایداری هزینه به صورت ضریب وزن داری از فراجاهش OV و زمان نشست T_s خواهد بود.

با مشخص شدن هزینه همه 50 عضو اولیه جمعیت، بر اساس هزینه بدست آمده به هریک از آن‌ها یک رتبه اختصاص داده و این پاسخ‌ها مرتب می‌شوند. سپس تعدادی به صورت تصادفی جهت اعمال ترکیب انتخاب می‌شوند.

اپراتور ترکیب انواع مختلفی دارد که با توجه به فضای پاسخ مسئله تعیین می‌شود. ترکیب مورد استفاده در این الگوریتم ترکیب یکنواخت است که به صورت زیر عمل می‌کند:

$$y1 = \alpha \times x1 + (1 - \alpha) \times x2;$$

$$y2 = \alpha \times x2 + (1 - \alpha) \times x1;$$

$x1$ و $x2$ والدین و $y1$ و $y2$ فرزند هستند. α یک متغیر تصادفی در بازه (۰, ۱) است که مقدار وابستگی هر فرزند را به والدین تعیین می‌کند. محدودیت این روش ترکیب در این است که فرزندان نهایتاً می‌توانند در همان بازه والدین باشند. بنابراین در صورتی که اعضای جمعیت همه در نزدیکی یک بهینه محلی باشند امکان خروج از آن بازه و رسیدن به بهینه کلی از بین می‌رود. برای حل این مشکل متغیری با نام δ را تعریف کرده و با انتخاب مقادیر مختلف برای این متغیر و اعمال آن به صورت زیر می‌توان این متغیر را به گونه‌ای تعیین کرد که فرزندان احتمال خروج از بازه والدین را به اندازه دلخواه داشته باشند. این متغیر معمولاً طول گام نامیده می‌شود. این متغیر می‌تواند هر مقداری داشته باشد اما برای حفظ ویژگی‌های مثبت والدین معمولاً مقدار کمی انتخاب می‌شود. طول گام در این پژوهش 0.1 انتخاب شد. نحوه اعمال آن به صورت زیر است:

$$\alpha = \text{unifrnd}(-\delta, 1 + \delta \text{size}(x1))$$

که منظور از unifrnd تابع با توزیع یکنواخت است.

پس از این مرحله نوبت به اپراتور جهش می‌رسد. این اپراتور یکی از اعضای هر کروموزوم به صورت تصادفی انتخاب کرده و یک متغیر تصادفی نرمال را به آن افزوده و آن را به صورت زیر تغییر می‌دهد:

$$j = \text{randi}([1 \text{ nVar}]);$$

$$y(j) = x(j) + \text{randn};$$

سپس همه فرزندان به وجود آمده توسط اپراتورهای ترکیب و جهش توسط تابع هزینه مورد بررسی قرار می‌گیرند و پس از مشخص شدن هزینه، به ترتیب از هزینه کم به زیاد مرتب می‌شوند. سپس مراحل انتخاب، ترکیب و جهش به ترتیب تکرار می‌شوند تا شرط خاتمه یا تکرار 100 نسل از فرزندان محقق شود. در شکل 4 فلوجارت این الگوریتم خلاصه شده است.

۵- نتایج آزمایش

در این بخش کارایی کنترل‌کننده‌های بهینه حاصل از الگوریتم ژنتیک مورد ارزیابی قرار می‌گیرند. برای این منظور 3 کنترل‌کننده با هدف (1) کاهش زمان نشست (T_s)، یعنی $\omega_1 = 1, \omega_2 = 0$ (2) کاهش فراجاهش (OV)، یعنی $\omega_1 = 0, \omega_2 = 1$ (3) کاهش هم‌زمان زمان نشست و فراجاهش، یعنی $\omega_1 = \omega_2 = 1$ طراحی شد. منحنی تابع هزینه برای این خواسته سوم در شکل 5 نشان داده شده است. همچنین در جدول 1 ضرایب بدست آمده برای این سه کنترل‌کننده آورده شده است.

برای ارزیابی کارایی کنترل‌کننده‌های پیشنهادی، از روش زیگر-نیکولز با رویکرد پاسخ فرکانسی استفاده شد. به عبارتی دقیق‌تر، از روش زیگر-نیکولز بر اساس پاسخ فرکانسی سیستم حلقه بسته با مشخص کردن محدوده پایداری تحت کنترل‌کننده صرفاً تناسبی استفاده شد. بهره تناسبی تا زمانی که سیستم به صورت حدی پایدار باشد، افزایش می‌یابد.

در سیستم مورد مطالعه بهره تناسبی به عنوان بهره نهایی K_u و متناظر با آن دوره فرکانسی نوسانات به عنوان دوره نهایی T_u شناخته می‌شود. این مقادیر برای سیستم شناسایی شده $K_u = 6.9516$ و $T_u = 3.5881$ محاسبه گردید. با در دسترس بودن پارامترهای K_u و T_u ، کنترل‌کننده مناسب با استفاده از روش زیگر-نیکولز ضرایب این کنترل‌کننده، PID_4 ، برابر خواهد بود با:

$$K_p = 3.2716, K_i = 1.1749, K_d = 2.2774.$$

۶- نتایج شبیه‌سازی

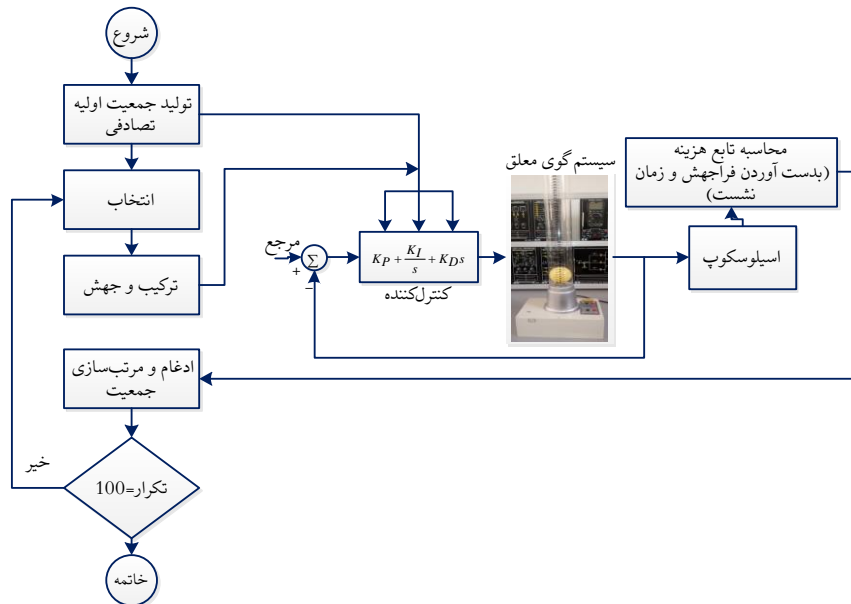
پاسخ پله سیستم حلقه-بسته با استفاده از چهار کنترل‌کننده طراحی شده در شکل 6 بازتاب داده شده است و در جدول 2 خلاصه‌ای از معیارهای کارایی این چهار کنترل‌کننده آورده شده است.

آن‌چنان که انتظار می‌رفت پاسخ پله گوی معلق با کنترل‌کننده زیگر-نیکولز نوسانی با فراجهشی بزرگ‌تر از بقیه نتیجه می‌دهد. کنترل‌کننده PID_1 کمترین زمان نشست و کنترل‌کننده PID_2 پاسخی بدون فراجاهش را ارائه می‌دهد. در کنترل‌کننده PID_3 در به ازای یک افزایش فراجاهش نسبت به PID_2 پاسخ سریع‌تری خواهیم داشت.

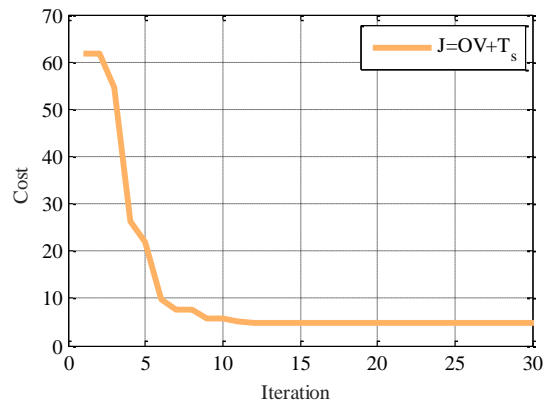
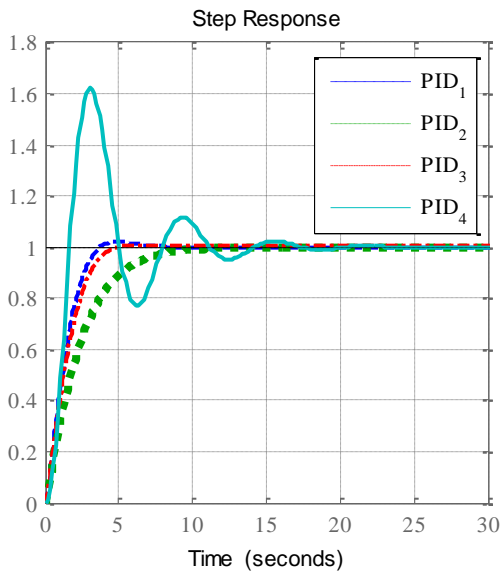
1.0193	0	0.8537	PID_2	$J = OV$
1.6306	0	1	PID_3	$J = T_s + OV$

جدول 5: پارامترهای کنترل‌کننده‌های بهینه بدست آمده از الگوریتم ژنتیک به ازای وزن‌های مختلف تابع هزینه

شاخص کارایی	نام کنترل‌کننده	K_D	K_I	K_P
$J = T_s$	PID_1	0.9997	0	1.8791



شکل 14: فلوچارت طراحی کنترل‌کننده PID مبتنی بر الگوریتم ژنتیک



شکل 15: تابع هزینه به ازای $J = OV + T_s$

شکل 16: نتایج شبیه‌سازی سیستم گوی معلق به ورودی پله

جدول 6: مقایسه معیارهای کارایی چهار کنترل‌کننده PID

کنترل‌کننده	PID_4	PID_3	PID_2	PID_1
زمان نشست	6.0	4.18	8.52	3.37
فراجهش	62.29	0.66	0	2.0

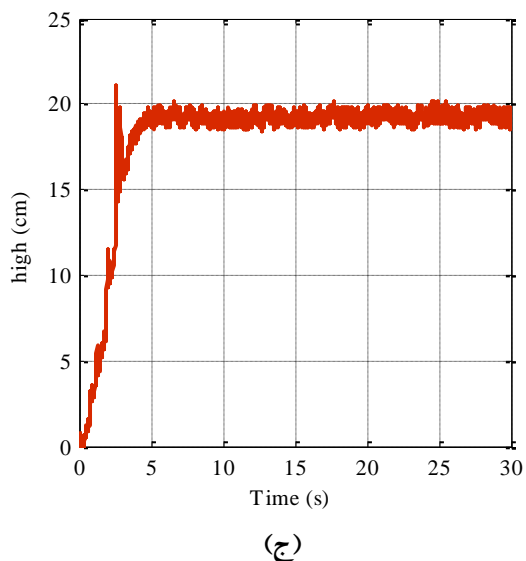
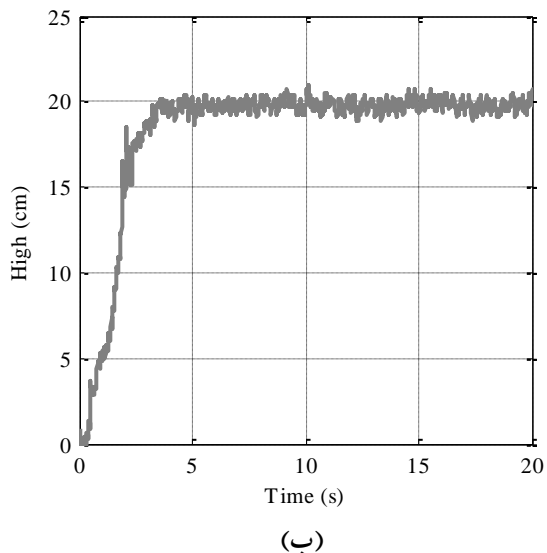
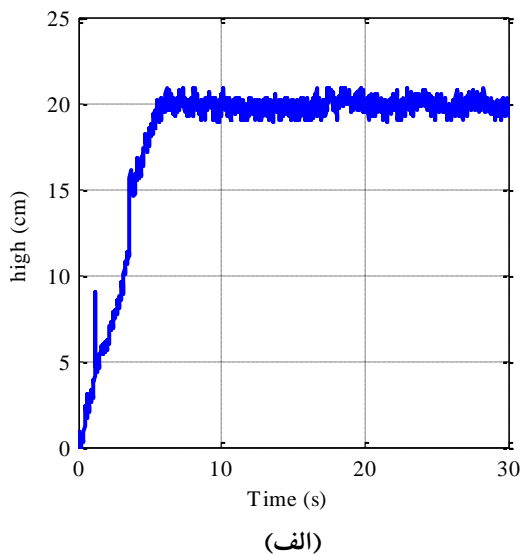
۶-۱- نتایج آزمون عملی

عملکرد چهار کنترل‌کننده به منظور معلق‌سازی گوی در ارتفاع 20 سانتی‌متری در آزمایشگاه مورد بررسی قرار گرفت. نتایج آزمون عملی

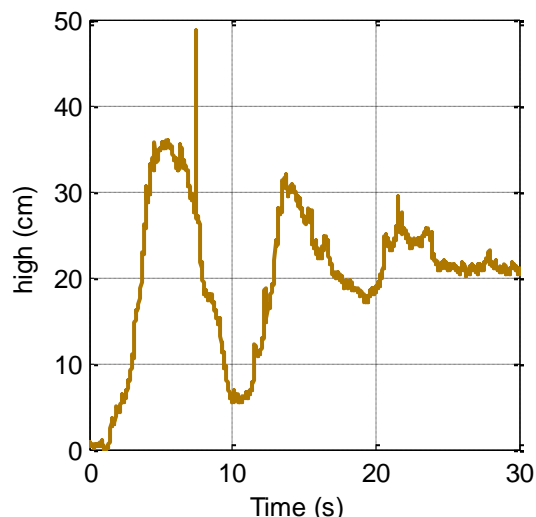
برای کنترل‌کننده‌های بهینه در شکل ۷ و کنترل‌کننده زیگلر-نیکولز در شکل ۸ ترسیم شده‌اند. نتایج این آزمون‌ها نیز مؤید کارایی کنترل‌کننده حاصل از الگوریتم ژنتیک هستند.

۷- جمع‌بندی

در این مقاله سیستم شناورسازی یک گوی توسط جریان هوا مورد مطالعه قرار گرفت. ابتدا معادلات غیرخطی توصیف‌کننده سیستم با استفاده از قوانین فیزیکی حاکم بر آن استخراج گردید. سپس مدل غیرخطی بدست آمده خطی‌سازی شد و با اعمال دسته‌ای از سیگنال‌های مشخص به ورودی سیستم و اندازه‌گیری خروجی‌های سیستم پارامترهای نامعلوم سیستم شناسایی گردید. سپس از الگوریتم ژنتیک برای حصول یک سیستم حلقه- بسته پایدار کمک گرفته شد. با انتخاب معیارهای کارایی متفاوت سه دسته کنترل‌کننده بهینه حاصل شد که مزیت هرکدام نسبت به روش زیگلر-نیکولز هم از طریق شبیه‌سازی و هم از طریق آزمون عملی مورد تأیید گرفت.



شکل 17: نتایج آزمون عملی کنترل‌کننده‌های بدست آمده از الگوریتم ژنتیک: (الف) هدف کاهش فرجهش، (ب) کاهش زمان نشست، (ج) کاهش فرجهش و زمان نشست



شکل 18: نتیجه آزمون عملی با پیاده‌سازی کنترل‌کننده زیگلر-نیکولز

مراجع

- [1] Bansal, H.O., R. Sharma, and P. Shreeraman, "PID controller tuning techniques: A review," *Journal of control engineering and technology*, 2012. 2(4): p. 168-176.
- [2] Panda, R.C., C.-C. Yu, and H.-P. Huang, "PID tuning rules for SOPDT systems: Review and some new results," *ISA transactions*, 2004. 43(2): p. 283-295.
- [3] De Moura Oliveira, "Modern heuristics review for PID control optimization," *A teaching experiment. in 2005 international conference on control and automation*. 2005. IEEE.
- [4] Koivo, H. and J. Tantt, "Tuning of PID Controllers: Survey of Siso and MIMO Techniques, in Intelligent tuning and adaptive control," 1991, Elsevier. p. 75-80.
- [5] Ang, K.H., G. Chong, and Y. Li, "PID control system analysis, design, and technology," *IEEE Transactions on Control Systems Technology*, 2005. 13(4): p. 559-576.
- [6] Åström, K.J. and T. Hägglund, "The future of PID control. *Control Engineering Practice*," 2001. 9(11): p. 1163-1175.
- [7] Åström, K.J. and T. Hägglund, *Automatic tuning of PID controllers*, 1988: Instrument Society of America (ISA).
- [8] Ziegler, J.G. and N.B. Nichols, "Optimum settings for automatic controllers," *Trans. ASME*, 1942. 64(11).
- [9] Cohen, G., "Theoretical consideration of retarded control," *Trans. ASME*, 1953. 75(1): p. 827-834.
- [10] Åström, K.J., et al., "Automatic tuning and adaptation for PID controllers-a survey," *Control Engineering Practice*, 1993. 1(4): p. 699-714.
- [11] Kramer, O., *Genetic algorithm essentials*. Vol. 679. 2017: Springer.
- [12] Kumar, M., et al., "Genetic algorithm: Review and application," *International Journal of Information Technology and Knowledge Management*, 2010. 2(2): p. 451-454.



تلفیق داده مبتنی بر یادگیری عمیق برای سنجش طیف مشارکتی چندباند

سید سبحان عطاری^۱، روزبه رجبی^۲، یاسر عطار ایزی^۳

^۱ دانشجوی کارشناسی ارشد، گروه مخابرات، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم،
attari.s@qut.ac.ir

^۲ استادیار، گروه مخابرات، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم،
rajabi@qut.ac.ir

^۳ استادیار، گروه مخابرات، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم،
attarizi@qut.ac.ir

چکیده

مساله سنجش طیف یکی از مسائل فراگیر در کاربرد های مختلف است. از جمله رادیوهای شناختی، شبکه های همزیست با اشتراک طیف و ... از طرفی مساله سنجش طیف با مشکلاتی نظیر محوشدگی، سایه افکنی و نقاط کور روبرو بوده که در طی زمان منجر به ظهور روش های سنجش طیف مشارکتی گردیده است. در سنجش طیف مشارکتی با استفاده از تلفیق اطلاعات چندین کاربر مشارکت کننده برای تعیین حضور کاربر اولیه در طیف مورد نظر تلاش می شود. تاکنون الگوریتم های تلفیق داده گوناگونی به منظور سنجش طیف مشارکتی ارائه شده اند که هر کدام مزایا و معایب خاص خود را دارند. در این مقاله، الگوریتمی مبتنی بر الگوریتم های یادگیری عمیق ارائه شده است که از لحاظ عملکرد بر روشهای پیشین برتری داشته و همچنین معایبی از روش های ارائه شده را برطرف نموده است. از جمله اینکه، در این روش تعداد کاربران مشارکت کننده در سنجش طیف، محدودیتی برای الگوریتم محسوب نمی شود و می تواند در پیاده سازی عملی شبکه های متحرک اشتراک طیف که در آنها تعداد کاربران مشارکت کننده کم و زیاد می شود، کمک کننده باشد.

کلمات کلیدی

رادیوی شناختی، سنجش طیف مشارکتی، شبکه های عصبی مصنوعی، یادگیری عمیق.

ضروری و دارای اولویت برای تحقیقات شبکه‌های مخابراتی نسل بعد است که بایستی مورد توجه قرار بگیرد.

یکی از اولین راه حل‌های امید بخش ارائه شده برای این چالش، استفاده از مفاهیم رادیوی شناختی می باشد که تقریباً دو دهه مورد بررسی گسترده جامعه پژوهشی قرار گرفته است. رادیوی شناختی این اجازه را می دهد تا دستگاه‌های بی سیم، طیف رادیویی را سنجش کنند، درباره وضعیت کانال‌های فرکانسی تصمیم بگیرند و پارامترهای ارتباطی خود را به گونه ای بازتنظیم کنند تا به کیفیت خدمات^۱ مورد نیاز دست یابند و به طور همزمان مصرف انرژی خویش را کمینه کنند. این دستگاه ها می‌توانند علاوه بر استفاده از باندهای بدون مجوز، از باندهای مجوزدار نیز، در زمانی که کاربران اولیه مجوزدار آنها فعال نیستند استفاده کنند و مانع از تداخل نامطلوب شوند [1].

۱- مقدمه

در پی افزایش چشمگیر تعداد دستگاه های بی سیم همراه و نیز مدیریت ناپویای طیف رادیویی، طیف رادیویی قابل دسترس دچار کمبودی روز افزون شده است. پیش بینی های رسمی حاکی از آن است که تا سال ۲۰۲۰ بیش از ۵۰ میلیارد دستگاه بی سیم به شبکه متصل می شوند، که احتمالاً همه آنها نیازمند دسترسی به اینترنت هستند. مدیریت ناپویای طیف رادیویی دیگر به اندازه کافی کارآمد نیست تا امکان دسترسی به تمام این دستگاه‌ها را فراهم آورد. در اثر تخصیص ناپویای طیف رادیویی، برخی از بخش‌های طیف رادیویی به شدت مورد استفاده قرار گرفته، در حالی که بخش هایی دیگر از آنها بدون استفاده یا کم استفاده باقی مانده اند. در نتیجه مساله کمبود طیف رادیویی یکی از موضوعات

به همین منوال به دلیل گسترش مفاهیم و کاربرد های الگوریتم های یادگیری عمیق، این الگوریتم ها نیز وارد عمل شدند. مقاله [8] به وسیله شبکه عمیق خود رمزنگار به صورت خودکار ویژگی های مناسب را برای مساله سنجش طیف مشارکتی استخراج کرده و سپس خروجی شبکه را به یک طبقه بند ماشین بردار پشتیبان وارد کرده است. مقاله [9] نیز با فرمول بندی مساله سنجش طیف مشارکتی به فرم ماتریسی از شبکه های عمیق کانولوشنی بدین منظور استفاده نموده است. اگرچه شبکه های عمیق به منظور سنجش طیف، فقط در دو مقاله فوق الذکر مورد استفاده واقع شده اند، اما در مسائل مرتبطی نظیر پیش بینی طیف و طبقه بندی سیگنال در رادیوهای شناختی به وفور مورد استفاده قرار گرفته اند.

در این مقاله ما به منظور حل مساله سنجش طیف مشارکتی چند باند از شبکه عصبی بازگشتی LSTM استفاده کرده و نتایج شبیه سازی را با روش های پیشین مقایسه خواهیم کرد. مزیت های روش پیشنهادی عبارتند از:

۱. انطباق پذیری شبکه پیشنهادی برای تعداد متفاوت کاربران اولیه که بر خلاف روش ارائه شده در [9] شبکه مورد استفاده برای ابعادی خاص از مساله آموزش داده شده بود.
۲. کارایی مستقل از نحوه شماره گذاری کاربران ثانویه، باز برخلاف مقاله [9] که نیاز به بازآرایی شماره گذاری کاربران ثانویه در صورت عدم اطلاع از موقعیت مکانی آنها بود.
۳. عدم نیاز به آموزش مجدد در هنگام تغییر محیط به دلیل نرمال سازی داده ها قبل از ارسال به گره مرکزی (برخلاف [5])
۴. قابلیت اجرای آشکارسازی سنجش طیف و تصمیم گیری علی-رغم عدم دریافت داده های یک یا چند کاربر در زمان مناسب. چارچوب بندی مقاله حاضر بدین شرح است: در فصل دوم به فرمول بندی مساله و مدل سیستم پرداخته می شود. فصل سوم جزئیات شبیه سازی صورت گرفته را بیان کرده و در فصل چهارم به بررسی نتایج شبیه سازی میپردازیم. در نهایت در فصل پنجم نتیجه گیری مقاله آورده می شود.

۲- صورت بندی مساله و مدل سیستم

در مساله سنجش طیف مشارکتی بحث شده، فرض شده است که یک کاربر اولیه با تعدادی کاربر ثانویه که تعدادشان در طول زمان متغیر است، در یک ناحیه مشخص به صورت متحرک و تصادفی پراکنده شده اند و N_B باند فرکانسی متفاوت هرکدام با پهنای باند W در اختیار شبکه اولیه قرار دارد و در هر لحظه از زمان از چندین باند فرکانسی به صورت تصادفی و همزمان استفاده می کند. توان ارسالی کاربر اولیه در هر یک از باندهای فرکانسی که آنها را اشغال کرده توان ثابت P می باشد و هر یک از باندهای اشغال شده به باند کناری خویش با ضریب η توان نشت می کند.

همچنین فرض می شود که در لحظه سنجش طیف، هیچ یک از کاربران ثانویه توانی ارسال نمی کنند (مشابه زمان سکوت)؛ در استاندارد IEEE 802.11 [9] تا مبدا تداخل ناشی از ارسال یکی از کاربران ثانویه بر روی نتیجه سنجش طیف تاثیر نامطلوب گذارد.

تکنولوژی رادیوی شناختی بستر بهبود کارایی طیفی را تحت سیاست فعلی تخصیص منابع طیفی تا حد ممکن فراهم می کند. کاربران ثانویه، نه تنها برای دسترسی فرصت طلبانه به حفره های فرکانسی بایستی به سنجش محیط طیفی پیرامون خویش بپردازند، بلکه نیاز دارند تا مرتباً حضور کاربر اولیه را در هنگام اشغال طیف های مجوزدار بسنجند تا هنگام حضور مجدد کاربر اولیه در باند اشغالی، فوراً مخابره را قطع نموده و کانال را خالی نمایند، زیرا کاربران اولیه از اولویت بالاتری برای استفاده از طیف های تخصیص یافته شان برخوردارند و کاربران ثانویه بایست تحت هیچ شرایطی در مخابره کاربران اولیه اختلال ایجاد نمایند. از این رو، سنجش طیف تکنولوژی کلیدی رادیوهای شناختی و شالوده اشتراک طیف است. و نیز کارایی هر دو شبکه اولیه و شبکه رادیوی شناختی، مرهون کارایی آشکارسازی در سنجش طیف می باشد. زیرا از طرفی نرخ بالای آشکارسازی غلط موجب جلوگیری از استفاده بهینه رادیوی شناختی شده و از طرف دیگر حساسیت پایین آشکارساز موجب اختلال در کارکرد شبکه اولیه می گردد [2].

در پی اینکه سایه افکنی و محو شدگی عمیق موجب می شوند تا نتایج سنجش توسط یک کاربر قابل اعتماد نباشد، در نتیجه، الگوریتم های سنجش طیف مشارکتی پیشنهاد شدند که به منظور بهبود قابلیت اطمینان آشکارساز، نتایج سنجش طیف چندین کاربر را با یکدیگر ترکیب می کنند. مقاله [3] الگوریتم های تلفیق داده OR ، AND و $K-RANK$ را پیشنهاد داده است، که در آن گره های مشارکت کننده تصمیمات خود را به مرکز تلفیق (که به عنوان گره مرکزی نیز شناخته می شود) ارسال می کنند و مرکز تلفیق، از الگوریتم های تلفیق برای تصمیم گیری نهایی خود استفاده می کند. الگوریتم هایی که در آنها، مرکز تلفیق تنها نتایج آشکارسازی را مدیریت می کند، الگوریتم تلفیق سخت نامیده می شود، که اجرای آن نسبتاً ساده و کم هزینه است. برعکس، در الگوریتم تلفیق نرم، گره های شرکت کننده، داده های تشخیص داده شده یا سایر اطلاعات مفید را به طور مستقیم به گره تلفیق انتقال می دهند. مرکز تلفیق مقدار زیادی اطلاعات غیر پردازش شده را مدیریت می کند که عملکرد و دقت الگوریتم تلفیق نرم را بهتر از روش تلفیق سخت می کند در حالی که سر بار تحمیلی بیشتری به شبکه دارد [2].

در این اثنا مساله سنجش طیف مشارکتی نیز همانند بسیاری مسائل دیگر که با داده سر و کار دارند مورد توجه علاقه مندان الگوریتم های یادگیری ماشین قرار گرفت. مقاله [4] از اولین مقالاتی بود که از الگوریتم های یادگیری ماشین برای سنجش طیف مشارکتی بهره جست و ضابطه ای خطی برای تلفیق داده ها مبتنی بر روش های بازشناسی الگو ارائه کرد. پس از آن مقاله [5] چند روش بدون نظارت و نظارت شده از الگوریتم های یادگیری ماشین را برای تمییز حالت خالی و اشغال بودن طیف در سنجش طیف مشارکتی به کار برد. مقاله [6] مساله سنجش طیف مشارکتی را با استفاده از الگوریتم ماشین بردار پشتیبان^۳ چند کلاسه صورت بندی کرد. مقاله [7] سنجش طیف مشارکتی فضایی-زمانی را به وسیله ماشین بردار پشتیبان با کمک پرتو دهی، صورت بندی و حل نمود.

$$y_i^j(m) = \begin{cases} \Gamma(q)x(m) + w_i^j(m) & \text{for } \mathcal{H}_1 \text{ \& } j \in B_P \\ \sqrt{\eta} \Gamma(q)x(m) + w_i^j(m) & \text{for } \mathcal{H}_1 \text{ \& } j \in B_A \\ w_i^j(m) & \text{otherwise} \end{cases} \quad (V)$$

که $x(\cdot)$ ، سیگنال ارسال شده توسط کاربر اولیه بوده، $w_i^j(\cdot)$ نویز دریافت شده توسط کاربر ثانویه i ام در باند z ام می باشد، B_P مجموعه تمامی باندهای اشغال شده توسط کاربر اولیه است و B_A مجموعه تمامی باندهای مجاور با اعضای مجموعه B_P می باشد. به علاوه مفروض است که انرژی ارسال سیگنال توسط کاربر اولیه در تمامی باندهای موجود مقدار ثابت P بوده و نویز $w_i^j(\cdot)$ نویز گوسی سفید جمع شونده با چگالی طیف توان N_0 باشد.

برای تحقق مقادیر آماره T_i^j در شبیه سازی از مدل آماری این آماره استفاده شده است. همانطور که در [11] نشان داده شده است این آماره دارای توزیع مربع کای می باشد اما با توجه به مقادیر متداول برای تعداد نمونه ها که عموماً بیش از ۲۰۰ هستند، عملاً طبق نظریه حد مرکزی میتوان توزیع آن را به صورت گوسی زیر در نظر گرفت:

$$T_i^j \sim \mathcal{N} \left(\sum_{m=1}^{N_S} \mathbb{E} [|y_i^j(m)|^2], \sum_{m=1}^{N_S} \text{Var} [|y_i^j(m)|^2] \right) \quad (A)$$

از طرفی در ادبیات مساله شبیه سازی آماری سیگنال ها، برای سیگنال های مدرن و OFDM، مدل گوسی یک مثل مناسب به شمار آمده است [12, 13]. و با در نظر گرفت مدل گوسی برای نمونه های سیگنال $y_i^j(m)$ میتوان نوشت:

$$\mathbb{E} [|y_i^j(m)|^2] = 2\sigma_w^2 + |h|^2(2\sigma_s^2) \quad (9)$$

$$\text{Var} [|y_i^j(m)|^2] = 4(\sigma_w^2 + |h|^2\sigma_s^2)^2 \quad (10)$$

در نتیجه برای آماره گوسی T_i^j خواهیم داشت:

$$T_i^j \sim \mathcal{N}(N_S(2\sigma_w^2 + 2\sigma_s^2|h|^2), 4N_S(\sigma_w^2 + \sigma_s^2|h|^2)^2) \quad (11)$$

نهایتاً به منظور تصمیم گیری تمامی کاربران ثانویه ای که موفق به سنجش طیف شده اند بایستی نتایج سنجش خود را به مرکز تلفیق داده (گره مرکزی) ارسال نموده تا گره مرکزی بتواند از تلفیق داده های کاربران ثانویه حضور کاربر ثانویه را در هر یک از باندهای طیفی آشکار کند. بدین منظور هر کاربر ثانویه داده های سنجش مربوط به باند های طیفی را به صورت یک بردار N_B بعدی درآورده و با توجه به مقادیر بیشینه و کمینه درایه های بردار مقادیر آن را بین ۱ و -۱ بهنجار نموده به مرکز تلفیق داده ارسال میکند. با این روش هم بار محاسبات مرکز تلفیق داده کاهش می یابد و هم میتوان با کدینگ مناسب سربار انتقال داده شبکه ثانویه کاهش داد. لذا کاربر ثانویه i ام بردار زیر را به گره مرکزی ارسال خواهد کرد:

مدل کانال بی سیم مخابراتی فرض شده، یک مدل احتمالاتی برای شبکه های رباتیک است که توسط [10] ارائه شده است. این مدل شامل شبیه سازی تلفات مسیر، محوشدگی چندمسیری^۵ و سایه افکنی^۶ به علاوه همبستگی مکانی دو بعدی محوشدگی چند مسیری و سایه افکنی می باشد. گرچه ما به دلیل شبیه سازی با رزولوشن پایین مکانی، محوشدگی چند مسیری را ناهمبسته شبیه سازی کرده ایم. در این مدل $\Gamma(q)$ توصیف کننده میزان تضعیف سیگنال دریافتی در گیرنده ای با موقعیت مکانی $q \in \mathbb{R}^2$ از فرستنده ای با موقعیت مکانی $q_b \in \mathbb{R}^2$ می باشد، و می توان آن را به صورت زیر نوشت:

$$\Gamma(q) = \Gamma_{PL}(q)\Gamma_{SH}(q)\Gamma_{MP}(q) \quad (1)$$

که در این عبارت، $\Gamma_{SH}(q)$ متغیر تصادفی مربوط به مولفه سایه افکنی و $\Gamma_{MP}(q)$ متغیر تصادفی بیانگر محوشدگی چند مسیری می باشند و $\Gamma_{PL}(q) = K_{PL}/\|q - q_b\|^{n_{PL}}$ تلفات وابسته به فاصله مسیر است و در آن K_{PL} ثابت تضعیف و n_{PL} نمای تضعیف تلفات مسیر می باشند. در این مدل متغیر تصادفی $\Gamma_{MP}(q)$ دارای توزیع رایس^۷ به شکل زیر است:

$$f_{\Gamma_{MP}}(x) = \frac{1}{(1 + K_{ric})e^{-K_{ric} - (1 + K_{ric})x}} I_0 \left(2\sqrt{xK_{ric}(K_{ric} + 1)} \right) \quad (2)$$

که در این عبارت $I_0(\cdot)$ تابع بسل اصلاح شده مرتبه صفر می باشد و K_{ric} عددی ثابت است. و $\Gamma_{SH}(q)$ نیز دارای توزیع لگ-نرمال می باشد. یعنی با فرض اینکه:

$$\xi(q) = 10 \log_{10}(\Gamma_{SH}(q)) \quad (3)$$

داریم:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\alpha}} e^{-x^2/2\alpha} \quad (4)$$

که در آن α واریانس سایه افکنی می باشد. همبستگی فضایی سایه افکنی نیز در مدل موجود، مشابه آنچه که در ادبیات مساله متداول است، به صورت نمایی مدل گردیده است و خود همبستگی $\xi(q)$ به شکل زیر بیان می شود:

$$\mathbb{E}\{\xi(q_1)\xi(q_2)\} = \rho(\|q_1 - q_2\|) = \alpha e^{-\|q_1 - q_2\|/\beta}, q_1, q_2 \in \mathbb{R}^2 \quad (5)$$

که در آن β عامل فاصله ناهمبستگی، بیانگر و کنترل کننده همبستگی فضایی سایه افکنی می باشد.

گرچه الگوریتم تلفیق داده ارائه شده مستقل از روش سنجش طیف کاربران ثانویه می باشد اما برای شبیه سازی مساله و ارزیابی کارکرد الگوریتم پیشنهادی، آشکارساز انرژی را برای تمامی کاربران ثانویه مد نظر قرار داده ایم. به طوریکه کاربر ثانویه i ام به طور مستقل در هر بازه زمانی Δt در باند z ام از N_B باند موجود به تعداد N_S نمونه برداری میکند و آماره زیر را تشکیل میدهد:

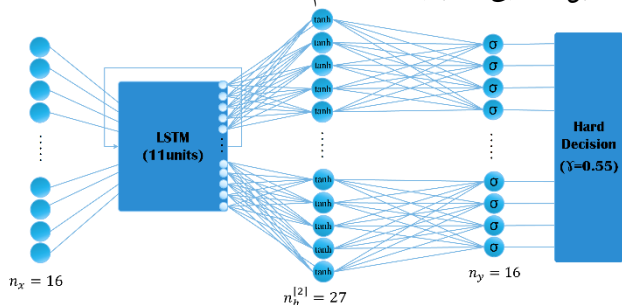
$$T_i^j = \sum_{m=1}^{N_S} |y_i^j(m)|^2 \quad (6)$$

که در آن $y_i^j(m)$ امین نمونه برداشته شده از سیگنال دریافتی کاربر ثانویه i ام در باند z ام می باشد. که میتوان آن را بر اساس فرض های صورت گرفته برای کانال بی سیم، تحت دو فرضیه حضور (\mathcal{H}_1) و عدم حضور (\mathcal{H}_0) به صورت زیر نوشت:

شبکه ای با پیچیدگی نسبتاً کمتری وجود داشته باشد و از لحاظ کارایی الگوریتم، مقاوم تر باشد. لذا این نوع شبکه برای مساله سنجش طیف از کارآمدی بهتری برخوردار است.

به طور ساده ساختار شبکه LSTM متشکل از یک نوار نقاله اطلاعات با سه دروازه است؛ دروازه فراموشی، دروازه ورودی و دروازه خروجی. دروازه فراموشی می آموزد تا اطلاعات ناکارآمد را از روی نوارنقاله حذف کند، دروازه ورودی می آموزد تا اطلاعات مورد نیاز را به نوار نقاله بیافزاید و در نهایت دروازه خروجی می آموزد تا اطلاعات مورد نیاز خروجی را از نوارنقاله بخواند.

ساختار شبکه پیشنهادی متشکل از سه لایه است، یک لایه LSTM به عنوان اولین لایه شبکه و دو لایه تمام متصل^۹ به منظور طبقه بند به دنبال آن. همانگونه که ذکر شد ابعاد بردارهای ورودی به لایه اول شبکه (LSTM) برابر ۱۶ می باشد. لایه LSTM، دارای ۱۱ واحد می باشد و لایه های دوم و سوم (تمام متصل) به ترتیب دارای ۲۷ و ۱۶ نورون می باشند. فعالساز لایه دوم تابع tanh و فعالساز لایه سوم تابع سیگموئید انتخاب گردیده است. به منظور یادگیری شبکه از تابع زیان آنتروپی متقابل دودویی^{۱۰} و بهینه سازی آدام^{۱۱} استفاده شده است.



شکل ۱۹ طرح شبکه عصبی پیشنهادی

۴- شبیه سازی و ارزیابی الگوریتم

برای شبیه سازی مساله به منظور تولید مصنوعی داده های سنجش طیف کاربران ثانویه، از نرم افزار متلب استفاده شده است. بدین ترتیب برای هر نمونه زمانی از داده ها، ابتدا تعداد کاربران ثانویه مشارکت کننده و مکانشان و باند های اشغال شده توسط کاربر اولیه به صورت تصادفی تعیین شدند؛ سپس تضعیف توان در کانال میان کاربر اولیه و کاربران ثانویه با توجه به مدل پیشنهادی [10] محاسبه شد و در نهایت بر اساس توزیع چگالی احتمال بیان شده در معادله (۱۱) برای آماره بردارهای T_i^j ، تولید و ذخیره سازی گردیدند. بدین منظور مقدار توان نویز برابر -154 dBm، توان ارسالی کاربر اولیه 23 dBm و محدوده حرکت کاربران اولیه و ثانویه یک مربع 300×300 متر لحاظ شده اند. تعداد باندهای اشتراکی و تعداد کاربران مشارکتی ۱۶ تا، پهنای باند هر کانال 10 MHz و ضریب نشست η برابر 20 dB - فرض شده اند. سپس شبکه عصبی سنجش طیف مشارکتی با کتابخانه کراس^{۱۲} و بستر نرم افزاری تنسورفلو^{۱۳} تحت زبان پایتون^{۱۴} پیاده سازی شده و داده های تولید شده به منظور یادگیری شبکه، به الگوریتم یادگیری داده شدند. نمودار یادگیری این الگوریتم که در شکل ۲۰ آورده شده است، نشان می دهند که شبکه عصبی به الگوریتم تلفیق داده مناسب همگرا شده است.

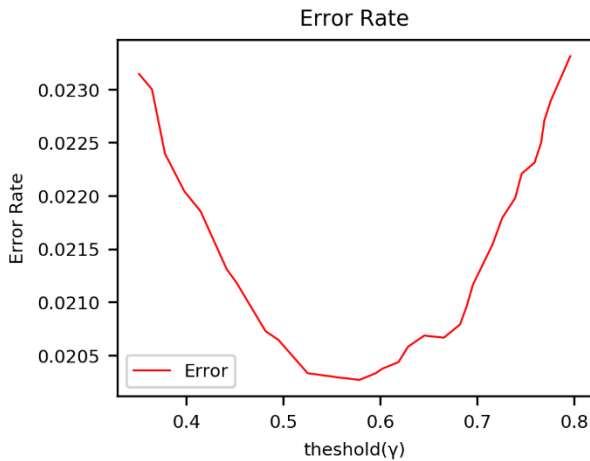
$$T_i \triangleq \frac{2}{\max_j\{T_i^j\} - \min_j\{T_i^j\}} \times \left(\begin{bmatrix} T_i^1 \\ T_i^2 \\ T_i^3 \\ \vdots \\ T_i^{N_B} \end{bmatrix} - \left(\frac{\max\{T_i^j\} + \min\{T_i^j\}}{2} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N_B \times 1} \right) \quad (12)$$

۳- الگوریتم تلفیق مبتنی بر LSTM

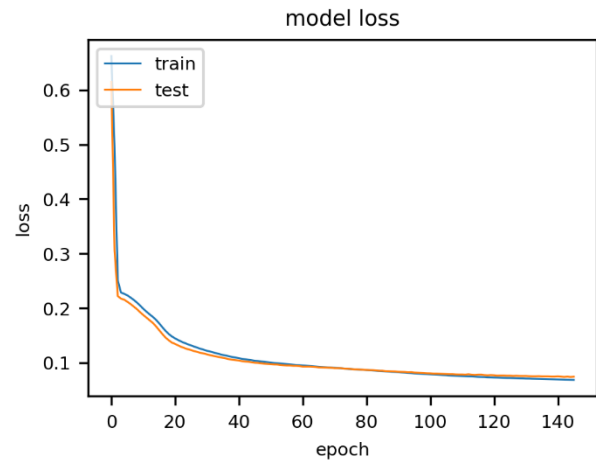
پس از دریافت داده های کاربران ثانویه توسط گره مرکزی، گره مرکزی بایستی نسبت به حضور یا عدم حضور کاربران اولیه در هریک از باندهای طیفی موجود تصمیم گیری نماید. بدین منظور در این مقاله یک الگوریتم مبتنی بر شبکه عصبی عمیق پیشنهاد شده است تا بتواند حضور کاربر اولیه را در تمامی باندهای موجود به طور همزمان آشکار کند. مزیت الگوریتم های مبتنی بر شبکه های عمیق نسبت به سایر الگوریتم های یادگیری ماشین این است که خود شبکه عمیق در حین عملیات یادگیری ویژگی های بهینه را جهت طبقه بندی می آموزد. لذا با این ویژگی از شبکه های عمیق می توان بر مساله بگرنج یافتن الگوریتم بهینه جهت سنجش طیف فائق آمد.

در سال های اخیر شبکه های یادگیری عمیق مورد توجه فراوان محققین قرار گرفته اند و این تحقیقات زمینه ساز پیشرفت های فراوان در زمینه های بینایی ماشین، تشخیص صدا، تشخیص اشیا، پردازش زبان های طبیعی و ... شده اند. دسته ای از این الگوریتم ها که برای تناسب بهتر با دنباله داده ها ظهور کرده اند الگوریتم های شبکه های عصبی بازگشتی^{۱۵} نامیده می شوند. این دسته از شبکه های عصبی قابلیت تطبیق با داده های با طول متغیر را دارا هستند. بدین معنی که نیازی نیست تا ابعاد داده های ورودی به این نوع از شبکه ها از پیش تعیین شده و ثابت باشد و می توانند با داده های ورودی با ابعاد متغیر تعامل کنند، لذا این دسته از شبکه های عصبی کاربرد فراوانی در مسائلی نظیر پردازش زبان های طبیعی، بازشناسی کلام و ... پیدا نموده اند. در عمل نیز در مساله سنجش طیف مشارکتی برای شبکه های بی سیم با یک چنین نوع از مساله ای مواجه هستیم. زیرا در طول زمان تعداد مشارکت کنندگان در سنجش طیف متغیر می باشد و این ایده اصلی ما برای استفاده از الگوریتم های شبکه عصبی بازگشتی برای مساله سنجش طیف است.

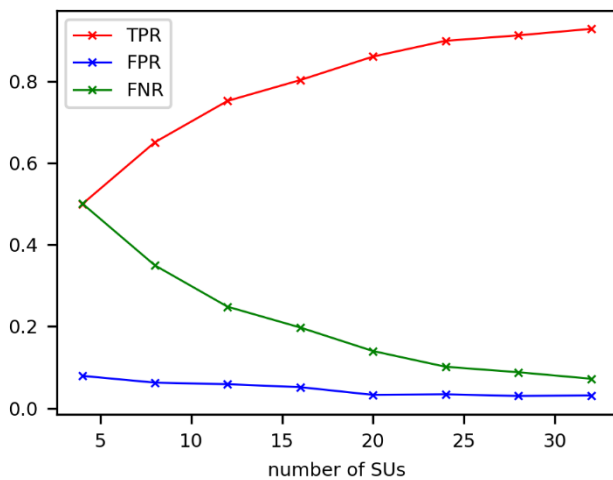
از جمله الگوریتم های موفق و پرکاربرد این دسته از شبکه های عصبی، شبکه های LSTM می باشند. شبکه LSTM، بر یکی از بزرگترین ضعف های شبکه های عصبی بازگشتی پیشین فائق آمده و بدین سبب محبوبیت دو چندانی نزد فعالین حوزه یادگیری ماشین به خصوص فعالین حوزه پردازش زبان های طبیعی کسب نمود. در شبکه های عصبی بازگشتی پیشین به دلیل پدیده ناپدید شدگی گرادیان، در حین یادگیری، وابستگی های طولانی مدت میان دنباله داده ها نادیده گرفته می شد. در واقع طبیعت LSTM این بود که اطلاعات را برای بازه های طولانی از زمان به یاد می آورد. و همین ویژگی LSTM موجب می شود تا برای داده هایی با دامنه تغییرات زیاد در طول دنباله ورودی (همچون تعداد مشارکت کنندگان در سنجش طیف)، نیاز به طراحی



شکل ۲۲ نمودار نرخ خطا به ازای حد آستانه

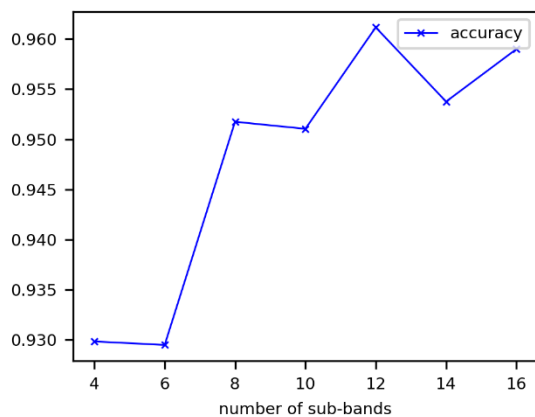


شکل ۲۰ نمودار یادگیری بر اساس معیار تلفات



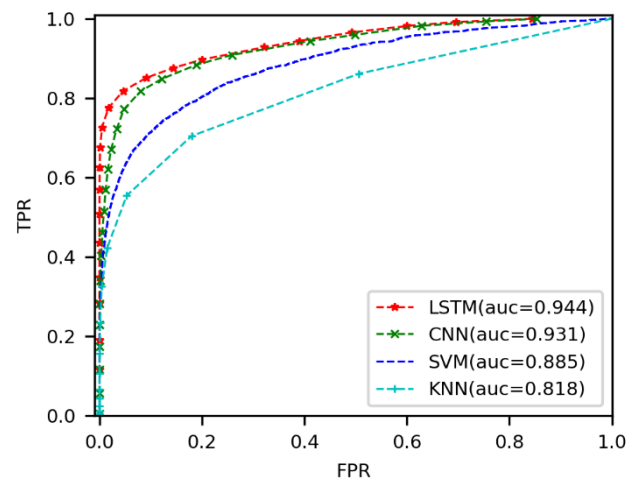
شکل ۲۳ نمودار پارامترهای آشکارسازی الگوریتم بر حسب تعداد کاربران ثانویه مشارکت کننده

همچنین با تغییر بعد لایه آخر شبکه عصبی، این الگوریتم را برای تعداد متفاوت کانال‌های اشتراکی ارزیابی کردیم. که نتیجه ارزیابی گویای تناسب پیچیدگی شبکه عصبی با ابعاد مساله سنسجش طیف تا ۱۶ باند همزمان می باشد. (شکل ۲۴ و شکل ۲۵)



شکل ۲۴ نمودار دقت الگوریتم بر حسب تعداد کانال‌های اشتراکی مورد سنسجش همزمان

به منظور ارزیابی کارایی شبکه عصبی طراحی شده، نمودار مشخصه عملکرد گیرنده^{۱۰} را برای الگوریتم پیشنهادی و الگوریتم‌های SVM، KNN و بدست آوردیم که در شکل ۲۱ قابل ملاحظه است. که بیانگر توانایی آشکارسازی صحیح با احتمال ۸۵ درصد به ازای هشدار اشتباه^{۱۶} ۱۰ درصد می باشد.



شکل ۲۱ نمودار مقایسه مشخصه عملکرد گیرنده (ROC)

این نمودار برتری الگوریتم پیشنهادی نسبت به الگوریتم‌های دیگر از نظر معیارهای احتمال هشدار غلط و احتمال خطای آشکارسازی را نشان می دهد.

همچنین با رسم نمودار نرخ خطا بر حسب حد آستانه آشکارسازی، اقدام به انتخاب حد آستانه بهینه نمودیم که نشان دهنده بهینگی حد آستانه ۰/۵۵ برای خروجی شبکه عصبی آشکارساز به ازای ۲ درصد خطای آشکارسازی می باشد. (شکل ۲۲)

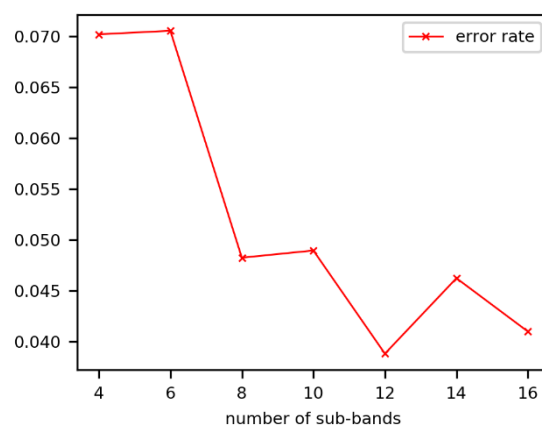
در مرحله بعد به منظور بررسی کارایی الگوریتم به ازای تعداد کاربران ثانویه مشارکت کننده مشخص، با تولید داده هایی به طول ثابت و مشخص (۴، ۸، ۱۲، ...، ۳۲)، به ارزیابی الگوریتم برای هر دسته از داده ها پرداخته و معیارهای آشکارسازی آنها را به ازای تعداد کاربران ثانویه رسم کردیم. (شکل ۲۳)

New Challenges, and Future Research Directions,” Sensors (Basel), vol. 19, no. 1, Jan. 2019.

- [2] F. Hu, B. Chen, and K. Zhu, “Full Spectrum Sharing in Cognitive Radio Networks Toward 5G: A Survey,” IEEE Access, vol. 6, pp. 15754–15776, 2018.
- [3] R. Tandra and A. Sahai, “Fundamental limits on detection in low SNR under noise uncertainty,” in 2005 International Conference on Wireless Networks, Communications and Mobile Computing, 2005, vol. 1, pp. 464–469 vol.1.
- [4] K. W. Choi, E. Hossain, and D. I. Kim, “Cooperative spectrum sensing under a random geometric primary user network model,” IEEE Transactions on Wireless Communications, vol. 10, no. 6, pp. 1932–1944, 2011.
- [5] K. M. Thilina, N. Saquib, and E. Hossain, “Machine Learning Techniques for Cooperative Spectrum Sensing in Cognitive Radio Networks,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 11, pp. 2209–2221, 2013.
- [6] O. P. Awe and S. Lambotharan, “Cooperative spectrum sensing in cognitive radio networks using multi-class support vector machine algorithms,” 2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS), pp. 1–7, 2015.
- [7] O. P. Awe and A. Deligiannis, “Spatio-Temporal Spectrum Sensing in Cognitive Radio Networks Using Beamformer-Aided SVM Algorithms,” pp. 25377–25388, 2018.
- [8] A. Subekti, H. F. Pardede, R. Sustika, and Suyoto, “Spectrum Sensing for Cognitive Radio using Deep Autoencoder Neural Network and SVM,” in 2018 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), 2018, pp. 81–85.
- [9] W. Lee, M. Kim, and D. Cho, “Deep Cooperative Sensing: Cooperative Spectrum Sensing Based on Convolutional Neural Networks,” IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 3005–3009, Mar. 2019.
- [10] A. Gonzalez-Ruiz, A. Ghaffarkhah, and Y. Mostofi, “A Comprehensive Overview and Characterization of Wireless Channels for Networked Robotic and Control Systems,” J. Robotics, vol. 2011, pp. 340372–340372, 2011.
- [11] S. Atapattu, C. Tellambura, and H. Jiang, *Energy detection for spectrum sensing in cognitive radio*. Springer, 2014.
- [12] A. Rudziński, “Normalized Gaussian approach to statistical modeling of OFDM signals,” Journal of Telecommunications and Information Technology, p. 54, Mar. 2014.
- [13] M. ElHassan, M. Crussière, J.-F. Hélar, Y. Nasser, and O. Bazzi, “A Novel Statistical Model of OFDM Signals with Optimal Tone Reservation PAPR Reduction for EVM Theoretical Analysis,” in 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2018, pp. 1–6.

زیر نویس‌ها

- ¹ Quality of Service
² Data Fusion
³ Support Vector Machine (SVM)
⁴ Quiet period
⁵ Multipath Fading
⁶ Shadowing
⁷ Rice Distribution
⁸ Recurrent neural network
⁹ Fully Connected (FC)
¹⁰ Binary cross entropy
¹¹ Adam optimizer
¹² Keras library
¹³ Tensorflow
¹⁴ Python
¹⁵ ROC curve
¹⁶ False alarm



شکل ۲۵ نمودار نرخ خطای الگوریتم برحسب تعداد کانال‌های اشتراکی مورد سنجش همزمان

به منظور مقایسه عملکرد این الگوریتم با سایر آشکارسازهای سنجش طیف مبتنی بر یادگیری ماشین، از معیار مساحت زیر منحنی مشخصه عملکرد گیرنده استفاده شده است که در جدول ۷ آمده است. همانطور که از نتایج جدول ۷ دیده می‌شود الگوریتم پیشنهادی از نظر عملکردی بر سایر الگوریتم‌های پیشین برتری دارد.

جدول ۷ مقادیر معیار مساحت زیر نمودار مشخصه عملکرد گیرنده برای هر یک از الگوریتم‌های نامبرده

Algorithms	AUC
KNN	0.818
SVM	0.885
CNN	0.931
LSTM	0.944

۵- نتیجه

در این مقاله به منظور تلفیق داده‌های سنجش طیف مشارکتی چند بانده از شبکه‌های عمیق LSTM استفاده شده است که اولاً موجب رفع معایبی از روش‌های ارائه شده پیشین شده و ثانیاً دقت آشکارسازی بیشتری نسبت به آنها ارائه می‌دهد. این الگوریتم از لحاظ معیار آشکارسازی مساحت زیر نمودار مشخصه عملکرد گیرنده به میزان ۰٫۹۴۴ دست یافته که بهبودی فاحش از نظر عملکرد آشکارسازی نسبت به الگوریتم‌های پیشین می‌باشد. همچنین دیگر محدودیت انعطاف ناپذیری نسبت به تغییر پیکربندی شبکه را ندارد. به علاوه کارایی این الگوریتم نسبت به تعداد کاربران مشارکت کننده و تعداد باندهای سنجش شونده نسبتاً مقاوم است.

این اولین کاری در حوزه سنجش طیف مشارکتی می‌باشد که از شبکه‌های عمیق بازگشتی بدین منظور استفاده کرده است. به منظور ادامه کار می‌توان الگوریتم مذکور را برای سنجش طیف زمانی-مکانی بکار برد و وابستگی‌های زمانی داده‌ها و مکان کاربران ثانویه را نیز در نظر گرفت.

مراجع

- [1] Y. Arjoun and N. Kaabouch, “A Comprehensive Survey on Spectrum Sensing in Cognitive Radio Networks: Recent Advances,



توسعه الگوریتم آلفا ماینر برای پشتیبانی از فرایندهایی با رخدادهای موازی

آزاده سلطانی^۱، حمید فدیشه‌ای^۲

^۱ گروه مهندسی کامپیوتر، دانشگاه بجنورد، بجنورد

a.soltani@ub.ac.ir

^۲ گروه مهندسی کامپیوتر، دانشگاه بجنورد، بجنورد

fadishei@ub.ac.ir

چکیده

فرایندکاوی عبارت است از تحلیل فرایندهای رخ داده شده در یک سازمان، که معمولاً در فایل گزارش رخ داده‌ها ثبت شده‌اند. فرایندکاوی شامل روشهایی برای استخراج مدل، بررسی سازگاری و انطباق فرایندهای واقعی با مدل پیش فرض و بهبود مدل فرایند است. در اکثر روشهای موجود، هر فرایند به صورت دنباله ای از رخ داده‌ها در نظر گرفته می‌شود که به صورت متوالی و پشت سر هم اجرا می‌شوند. این در حالیست که در بعضی از داده‌های واقعی، در مدل فرایند، رخ داده‌ها میتوانند موازی باشند. در این مقاله به طور خاص، بر الگوریتم آلفاماینر که یکی از روشهای استخراج مدل فرایند است، متمرکز شده‌ایم و آن را به گونه‌ای توسعه داده‌ایم که از رخدادهای موازی در یک فرایند پشتیبانی کند. برای ارزیابی روش جدید، از بسته نرم افزاری ProM استفاده کرده و الگوریتم آلفا ماینر پیاده سازی شده در آن را توسعه داده‌ایم. نتایج ارزیابی‌ها نشان می‌دهد روش پیشنهادی به خوبی توانسته است، فرایندهایی شامل رخدادهای موازی را مدل کند.

کلمات کلیدی

داده کاوی، فرایندکاوی، الگوریتم آلفاماینر، رخدادهای موازی.

۱- مقدمه

تجاری ترکیب کرده است. در سازمانها مدل پیش فرض، معمولاً مدل دلخواه سازمان است که توسط فرد خبره مشخص می‌شود، این در حالیست که روالی که به صورت واقعی اتفاق می‌افتد ممکن است متفاوت باشد و افراد از مدل پیش فرض تبعیت نکنند. یکی از وظایف فرایندکاوی استخراج مدل واقعی فرایندهاست. در این روشها، گزارش‌های موجود سیستم که در آنها، رخدادهای واقعی در سازمان همراه با زمان آنها ثبت شده است، تحلیل می‌گردد و یک مدل واقعی بر اساس فرایندهای واقعی استخراج می‌شود. با نمایش این مدل به مدیران و مشخص کردن گلوگاه‌ها^۳ و یا بررسی سازگاری این مدل با مدل پیش‌فرض، میتوان به بهبود مدل پیش‌فرض و همچنین بالابردن راندمان و کارایی سیستم کمک کرد. بنابراین هدف فرایندکاوی، تحلیلی مبتنی بر

امروزه با واژه‌های متعددی چون داده کاوی، داده‌های حجیم، هوش تجاری و فرایندکاوی سروکار داریم که تقریباً همه آنها به گونه‌ای بخش‌هایی از داده‌هایی که توسط سیستمهای اطلاعاتی جمع‌آوری شده‌اند را پردازش می‌کنند ولی تفاوت‌هایی هم با هم دارند. داده کاوی بر روی تحلیل مجموعه داده‌های بزرگ متمرکز می‌شود و مدیریت فرایند تجاری^۱ بر مدلسازی و بهبود فرایندهای تجاری متمرکز شده است. فرایندکاوی^۲ این دو مبحث را به هم متصل کرده است؛ به عبارت دیگر تحلیل داده را با مدلسازی، کنترل و بهبود فرایندهای

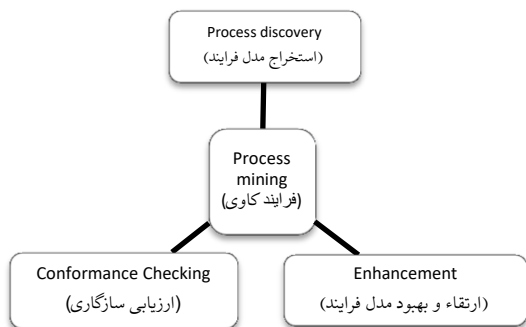
¹ Business Process Management

² Process Mining

³ Bottleneck

مورد چگونگی رخداد واقعی فرایند ها آغاز کنیم. سپس مدل واقعی را از روی این گزارشهای ثبت شده، استخراج نماییم [6]-[2].

- بررسی انطباق مدل^۸: به دلایل مختلفی ممکن است نیاز به ارزیابی سازگاری و انطباق دو مدل وجود داشته باشد. دلیل اول اینکه آیا مدل استخراج شده از گزارشهای ثبت شده با مدل پیش بینی شده و پیش فرض برابر هستند یا خیر. انحراف از مدل پیش فرض میتواند نمایانگر یک کلاهبرداری، ناکارآمدی در سیستم، طراحی بد روال انجام کار و یا روالهای منقضی شده باشد. دوم اینکه در بعضی روشهای استخراج مدل از انطباق مدل برای ارزیابی کارایی استفاده می شود. مثلا الگوریتمهای ژنتیک، از میزان انطباق برای انتخاب کاندیدها برای تولید نسل بعدی بهره می‌برند [12]-[8].
- بهبود^۹: با استخراج مدل واقعی، مقایسه آن با مدل پیش فرض، نمایش مدل برای افراد خبره و نمایش گلوگاه‌ها میتوان به بهبود و اصلاح جریان کار کمک کرد [1].



در این تحقیق به طور خاص بر روی مسئله استخراج فرایند متمرکز شده‌ایم. فایل گزارش رخدادها به عنوان ورودی به سیستم داده شده و مدل واقعی فرایندها استخراج میشود. شکل ۲ نمایی از چگونگی

شکل 26- شکل فرایند کاوی

استخراج مدل فرایند است. ورودی این متد، یک گزارش از رخدادها ثبت شده است، شبیه آنچه در شکل ۲ نشان داده شده است. هر نمونه (case) یک فرایند را نشان میدهد و هر رخداد (event یا task) یک فعالیت را نشان می‌دهند. فایل گزارش رخدادها، دنباله ای از این رخدادهاست که از نظر زمانی پشت سرهم اتفاق افتاده‌اند و در نهایت خروجی این متد یک مدل است که به شکلهای مختلف از جمله پتری نت یا گراف وابستگی ممکن است ایجاد شود. الگوریتمهای متعددی به منظور استخراج مدل فرایند تاکنون پیشنهاد شده‌اند، از جمله آلفا ماینر [2]، آلفا شارپ [13]، الگوریتم مبتنی بر منطقه [14]، روش مبتنی بر الگوریتم ژنتیک [4] و الگوریتم فازی [3]. این روشها برای

فرایند، بر روی داده‌های مربوط به وقایع و رخدادها است. منبع و ورودی داده، فایل گزارشی است که تمامی وقایع و رخدادها را سیستم در آن ثبت شده است. به طور معمول فایل گزارش وقایع^۴ به صورت متوالی رخدادها را ثبت می‌کنند که هر رخداد (event) یا واقعه، یک فعالیت (یک مرحله شناخته شده در فرایند) است که متعلق به یک نمونه (case) خاص است. برای هر رخداد اتفاق افتاده برچسب زمانی آن نیز ثبت خواهد شد. به طور دقیق‌تر می‌توان گفت که هر فایل گزارش ثبت شده، تشکیل شده از چندین نمونه که هر نمونه لیستی از وقایع اتفاق افتاده شده است.

به طور خلاصه می‌توان گفت وظایف اصلی فرایندکاوی عبارتند از: ۱- استخراج مدل فرایند واقعی ۲- بررسی سازگاری با مدل پیش فرض و ۳- بهبود مدل فرایند پیش فرض [1]. در این تحقیق بر مسئله استخراج مدل واقعی متمرکز شده‌ایم. روشهای متعددی برای استخراج مدل تا کنون ارائه شده است [6]-[2]. این مدلها تنها از فرایندهایی پشتیبانی میکنند که رخدادها مربوط به هر نمونه آن به صورت متوالی اتفاق افتاده باشند و به عبارت دیگر رخدادها موازی را پشتیبانی نمی‌کنند. این در حالیست که در دنیای واقعی ممکن است رخدادها موازی داشته باشیم. به طور مثال در فرایند انتخاب واحد دانشجویان، دروسی که در یک ترم انتخاب میشوند را میتوان به عنوان رخدادها موازی با هم در نظر گرفت. به عبارت دیگر در یک ترم دروس انتخاب شده نسبت به هم هیچ تاخر و تقدیمی ندارند؛ در صورتیکه ترمهای مختلف، زمان را در مدل فرایند مشخص می‌کنند. به طور مثال، انتخاب دروس در ترم اول، از نظر زمانی قبل از انتخاب دروس در ترم دوم رخ داده‌اند. در این مقاله الگوریتم پایه آلفا ماینر را به گونه ای توسعه داده‌ایم که بتواند رخدادها موازی را نیز پشتیبانی کند.

ساختار ادامه مقاله به این صورت است. در بخش ۲ به بیان مسئله و پیشینه تحقیق می‌پردازیم. بخش سوم، الگوریتم آلفا ماینر و روش پیشنهادی را توضیح می‌دهد، بخش چهارم شامل ارزیابی و نتایج پیاده سازی است و در نهایت بخش پنجم نتیجه تحقیق را نشان می‌دهد.

۲- بیان مسئله و کارهای انجام شده قبلی

میزان رشد داده‌های مربوط به رخدادها^۵ از قانون مور^۶ تبعیت میکند. به عبارت دیگر، مشابه تعداد ترانزیستورهای روی تراشه‌ها، ظرفیت حافظه‌ها و قدرت محاسباتی کامپیوترها، دنیای داده‌های دیجیتالی نیز به صورت نمایی رشد میکند و هر دو سال تقریباً دو برابر می‌شوند [7]. داده‌های مربوط به رخدادها، به کمک تکنیکهای فرایندکاوی قابل تحلیل هستند و می‌توان از آنها برای بهبود فرایندهای تجاری بهره برد. شکل ۱ وظایف مختلف فرایندکاوی را نشان می‌دهد. وظایف اصلی فرایندکاوی عبارتند از:

- استخراج مدل فرایند^۷: در این مسئله به جای شروع با طراحی جریان کاری^۸، میتوانیم با جمع‌آوری اطلاعاتی در

⁸ Workflow

⁹ Conformance Checking

¹⁰ Enhancement

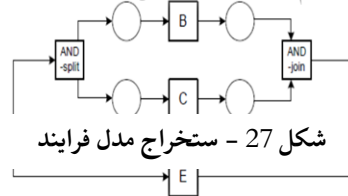
⁴ Event log

⁵ Event Data

⁶ Moore's Law

⁷ Process Discovery

case identi	task identi
case 1	task A
case 2	task A
case 3	task A
case 3	task B
case 1	task B
case 1	task C
case 2	task C
case 4	task A
case 2	task B
case 2	task D
case 5	task A
case 4	task C
case 1	task D
case 3	task C
case 3	task D
case 4	task B
case 5	task E
case 5	task D
case 4	task D



شکل ۲۷ - استخراج مدل فرایند

- ۳- $a \#_L b$ ، اگر و تنها اگر $a \not>_L b$ و $b \not>_L a$ به عبارت دیگر در هیچ فرایندی نه a بلافاصله قبل از b باشد و نه b بلافاصله قبل از a .
- ۴- $a \parallel_L b$ ، اگر و تنها اگر $a >_L b$ و $b >_L a$ مفهوم این ارتباط این است که در یک فرایند a رخداد بلافاصله قبل از b باشد و در فرایندی دیگر b بلافاصله قبل از a باشد. الگوریتم آلفا این حالت را به معنی موازی بودن دو رخداد در کل فایل گزارش در نظر می‌گیرد.
- به طور مثال، فرض کنید فایل گزارش L_1 به صورت $L_1 =$

جدول ۸- جدول footprint مربوط به فایل گزارش L_1 [۱]

$\{ \langle a, b, c, d \rangle, \langle a, c, b, d \rangle, \langle a, e, d \rangle \}$ تعریف شده باشد؛ به عبارت دیگر سه نمونه فرایند واقعی (trace) در فایل گزارش L_1 ثبت شده باشد. آنگاه $a >_{L_1} b$ ، $a >_{L_1} c$ ، $a >_{L_1} b$ ، $a \rightarrow_{L_1} b$ ، $e \#_{L_1} c$ ، $b \parallel_{L_1} c$ مثالهایی از انواع ارتباط بین رخدادها هستند. بین هر دو رخداد a و b یکی از انواع ارتباطهای ذکر شده باید وجود داشته باشد. بنابراین footprint هر فایل گزارش را میتوان به شکل یک ماتریس دو بعدی نشان داد. جدول ۱ نمونه ای از این ماتریس را برای مثال جاری نشان می‌دهد.

	a	b	c	d	e
a	$\#_{L_1}$	\rightarrow_{L_1}	\rightarrow_{L_1}	$\#_{L_1}$	\rightarrow_{L_1}
b	\leftarrow_{L_1}	$\#_{L_1}$	\parallel_{L_1}	\rightarrow_{L_1}	$\#_{L_1}$
c	\leftarrow_{L_1}	\parallel_{L_1}	$\#_{L_1}$	\rightarrow_{L_1}	$\#_{L_1}$
d	$\#_{L_1}$	\leftarrow_{L_1}	\leftarrow_{L_1}	$\#_{L_1}$	\leftarrow_{L_1}
e	\leftarrow_{L_1}	$\#_{L_1}$	$\#_{L_1}$	\rightarrow_{L_1}	$\#_{L_1}$

۳- روش پیشنهادی

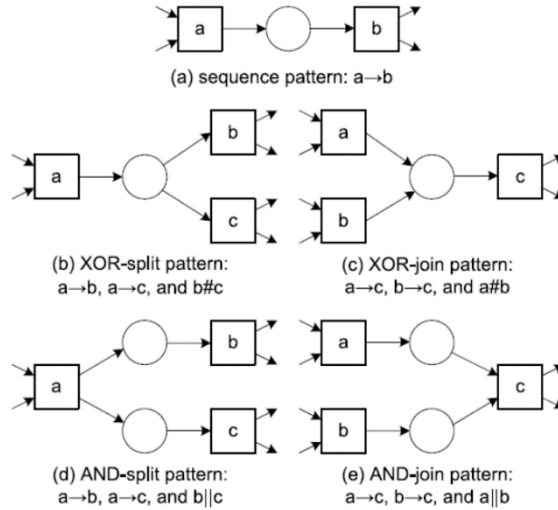
از آنجاییکه روش پیشنهادی بر پایه الگوریتم آلفاماینر است، ابتدا به معرفی مختصر این الگوریتم می‌پردازیم. الگوریتم آلفا ماینر به عنوان ورودی، گزارش رخدادها را در اختیار می‌گیرد و در فرایند واقعی را به صورت پتری نت استخراج می‌کند. در این الگوریتم چهار نوع ارتباط بین فعالیت‌ها تعریف شده است. اگر a و b دو فعالیت رخ داده شده در فایل گزارش L باشند، ارتباطهای مختلف بین a و b به این صورت است:

- ۱- $a >_L b$: اگر و تنها اگر یک نمونه فرایند (که در این مسئله به نام trace شناخته میشود) وجود داشته باشد به فرم $\sigma = t_1, t_2, \dots, t_n$ به گونه ای که $t_i = a$ و $t_{i+1} = b$.

مفهوم این نوع ارتباط این است که رخداد a حتما در یکی از فرایندهای نمونه، بلافاصله قبل از b آمده باشد.

- ۲- $a \rightarrow_L b$: اگر و تنها اگر $a >_L b$ و $a \not>_L a$ مفهوم این نوع ارتباط این است که در یک فرایند رخ داده شده، a رخداد بلافاصله قبل از b باشد و در هیچ فرایند دیگری b بلافاصله قبل از a نباشد.

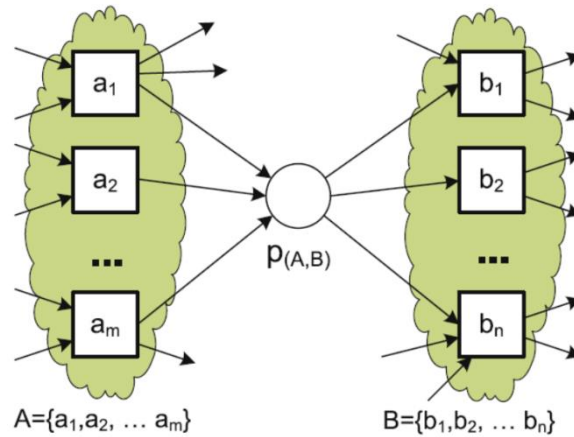
- ۴- مجموعه X_L را استخراج می‌کند. این مجموعه تمامی زوج مجموعه‌های A و B است که در هر زوج مجموعه، رخداد‌های موجود در A بلافاصله قبل از رخداد‌های B اتفاق افتاده‌اند و از طرفی عناصر A یا B بین خودشان ارتباط # وجود داشته باشد (قبل و بعد از هم اتفاق نیافتاده باشند).
- ۵- مجموعه Y_L را به دست می‌آورد که ماکزیمال مجموعه X_L است.
- ۶- مجموعه place ها P_L $(p(A, B))$ به دست می‌آید. به ازای هر زوج مجموعه‌ی (A, B) که عضو Y_L است، یک مکان $(p(A, B))$ تعریف می‌شود که رخداد‌های A را به B متصل می‌کند. رخداد‌های آغازین و پایانی هم به عنوان place به این مجموعه اضافه می‌شوند. شکل ۵ نمونه‌ای از یک place را نشان می‌دهد.
- ۷- در مرحله ۷ مجموعه پالها (F_L) ایجاد میشوند.
- ۸- خروجی این الگوریتم که یک مدل پتری نت به فرم $\alpha_L = (P_L, T_L, F_L)$ است استخراج می‌شود.



شکل 28- انواع الگوهایی که میتوان برای دنباله‌ای از فرایندها تعریف کرد [1].

- (1) $T_L = \{t \in T \mid \exists \sigma \in L \ t \in \sigma\}$
- (2) $T_I = \{t \in T \mid \exists \sigma \in L \ t = \text{first}(\sigma)\}$
- (3) $T_O = \{t \in T \mid \exists \sigma \in L \ t = \text{last}(\sigma)\}$
- (4) $X_L = \{(A, B) \mid A \subseteq T_L \wedge A \neq \emptyset \wedge B \subseteq T_L \wedge B \neq \emptyset \wedge \forall a \in A \forall b \in B \ a \rightarrow_L b \wedge \forall a_1, a_2 \in A \ a_1 \#_L a_2 \wedge \forall b_1, b_2 \in B \ b_1 \#_L b_2\}$
- (5) $Y_L = \{(A, B) \in X_L \mid \forall (A', B') \in X_L \ A \subseteq A' \wedge B \subseteq B' \implies (A, B) = (A', B')\}$
- (6) $P_L = \{p_{(A, B)} \mid (A, B) \in Y_L\} \cup \{i_L, o_L\}$
- (7) $F_L = \{(a, p_{(A, B)}) \mid (A, B) \in Y_L \wedge a \in A\} \cup \{(p_{(A, B)}, b) \mid (A, B) \in Y_L \wedge b \in B\} \cup \{(i_L, t) \mid t \in T_I\} \cup \{(t, o_L) \mid t \in T_O\}$
- (8) $\alpha(L) = (P_L, T_L, F_L)$

شکل 29- مراحل مختلف الگوریتم آلفا ماینر [1]

شکل 30- مکان $p(A,B)$ که دو مجموعه فعالیت A و B را به هم متصل می‌کند [1].

در این case به ترتیب اتفاق افتاده اند همراه با زمان رخداد آنها در این گزارش ثبت شده است. این مثال مربوط به دروسی است که توسط دانشجویان انتخاب می‌شوند. در این حالت هر دانشجو یک case و هر درس یک فعالیت و ترم مورد نظر هم، زمان رخداد فعالیت است. در جدول ۲ برای یک دانشجو اطلاعات چند ترم را نشان داده‌ایم. سه درس اول در ترمهای مختلف انتخاب شده اند و باید متوالی در نظر گرفته شوند و ۴ درس آخر، دو به دو همزمان انتخاب شده اند که انتظار داریم موازی در نظر گرفته شوند. شکل ۶ و ۷ به ترتیب خروجی آلفا ماینر پایه و آلفا ماینر اصلاح شده را نشان می‌دهند.

جدول 9- فایل گزارش ثبت شده که تنها دارای یک فرایند واقعی است.

Case#	نام رخداد	زمان وقوع
case1	ENGL 111	1-2005
case1	ENGL 112	2-2005
case1	ANTH 203	1-2006
case1	FMST 210	2-2006
case1	GEOG 102	2-2006
case1	PSYC 302	2-2007
case1	PSYC 361	2-2007

همانطور که در این دو شکل مشهود است روش آلفا ماینر پایه به زمان رخدادها اهمیتی نداده است و تنها بر اساس ترتیب وقوع فعالیت ها در گزارش ثبت شده، تصمیم گیری کرده است. در حالی که، روش پیشنهادی به خوبی توانسته است با در نظر گرفتن زمان وقوع، رخدادهای موازی را پشتیبانی کرده و این نقطه ضعف روش آلفا ماینر را برطرف نماید.

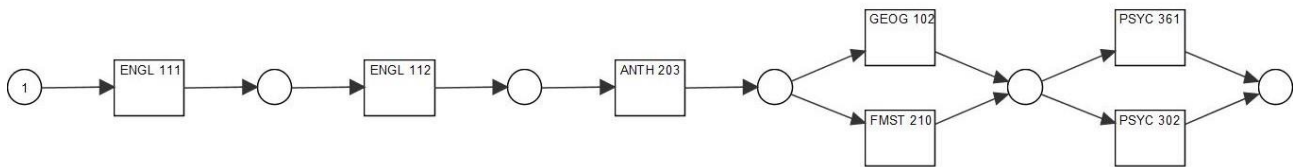
اگرچه گفته شده است که الگوریتم آلفا ماینر رخدادهای موازی را پشتیبانی می‌کند، اما منظور، موازی بودن رخدادها، بین case های مختلف است و نه برای یک case خاص. به عبارت دیگر در آلفا ماینر، رخداد الف می‌تواند در یک نمونه فرایند قبل از رخداد ب اتفاق بیفتد و در نمونه فرایند دیگر بعد از ب اتفاق بیفتد. اما آنچه ما انتظار داریم این است که در یک فرایند خاص هم بتوانیم، برای بعضی از رخدادها تقدم و تاخر قائل نشویم و آنها را موازی در نظر بگیریم. همانطور که توضیح داده شده، در تمام مراحل الگوریتم و پردازش فایل گزارش، تمام رخدادهای یک فرایند که در فایل ثبت شده است به صورت متوالی و پشت سر هم در نظر گرفته می‌شوند. به منظور توسعه الگوریتم آلفا، قبل از شروع الگوریتم، در مرحله ایجاد جدول footprint تغییری ایجاد کرده ایم. در این بخش اگر دو رخداد a و b در گزارش ثبت شده پشت سر هم باشند ولی زمان یکسان داشته باشند، $a >_{L_1} b$ در نظر گرفته نمیشود. در این حالت به دنبال اولین رخداد c ای هستیم که از نظر زمانی بلافاصله بعد از زمان a و b رخ داده باشد. سپس بین تمام رخدادهای x ای که در زمان یکسان و قبل از c رخ داده باشند رابطه $c >_{L_1} x$ را در جدول footprint قرار می‌دهیم. بقیه مراحل مانند قبل انجام خواهد شد.

۴- ارزیابی

برای ارزیابی عملکرد الگوریتم، از ابزار متن باز ProM استفاده کرده- ایم. در این ابزار، مجموعه‌ای کامل از الگوریتمهای فرایندکاوی به زبان جاوا پیاده سازی شده است. نسخه اصلی الگوریتم آلفا ماینر از این نرم افزار استخراج شده و تغییرات لازم، در راستای تحقق هدف پشتیبانی از رخدادهای موازی، به آن افزوده گردیده است. به منظور ارزیابی روش اصلاح شده، فایل گزارشی مشابه جدول ۲ به عنوان ورودی به هر دو الگوریتم داده شده است. برای نمایش بهتر در این فایل تنها اطلاعات مربوط به یک فرایند (case1) آمده است. فعالیت هایی که



شکل 31- خروجی الگوریتم آلفا ماینر پایه برای فایل نمونه جدول ۲



شکل ۳۲- خروجی الگوریتم آلفا ماینر اصلاح شده برای فایل نمونه جدول ۲

models,” *expert systems with applications*, vol. 65, pp. 194–211, 2016.

S. J. Leemans, D. Fahland, and W. M. Van der Aalst, “Scalable process discovery and conformance checking,” *Software & Systems Modeling*, vol. 17, no. 2, pp. 599–631, 2018. [12]

L. Wen, J. Wang, and J. Sun, “Mining invisible tasks from event logs,” in *Advances in Data and Web Management*, Springer, 2007, pp. 358–365. [13]

J. Carmona, J. Cortadella, and M. Kishinevsky, “A region-based algorithm for discovering Petri nets from event logs,” in *International Conference on Business Process Management*, 2008, pp. 358–373. [14]

۵- نتیجه

فرایندکاوی زمینه تحقیقاتی نسبتاً جدیدی است که دو فیلد داده کاوی و مدلسازی فرایندهای تجاری را با هم ترکیب کرده است و هدف آن بهبود فرایندهای واقعی است با تحلیل فرایندهای ثبت شده در فایل‌های گزارش. الگوریتم آلفا ماینر یکی از روشهای ارائه شده در فرایندکاوی است که هدفش کشف مدل فرایند به کمک فرایندهای ثبت شده است. در این الگوریتم فرض بر این است که تمام رخدادهای یک فرایند به صورت متوالی رخ میدهند که در دنیای واقعی اینگونه نیست. در این تحقیق با افزودن مفهوم رخدادهای موازی در یک فرایند این مشکل الگوریتم آلفا ماینر را برطرف کرده‌ایم. پیاده سازی و ارزیابی انجام شده به کمک ابزار متن باز ProM نشان میدهد که روش پیشنهادی توانسته است رخدادهای موازی را مدل کند.

مراجع

- [1] W. Van Der Aalst, *Process mining: discovery, conformance and enhancement of business processes*, vol. 2. Springer, 2011.
- [2] W. van der Aalst, T. Weijters, and L. Maruster, “Workflow mining: discovering process models from event logs,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 9, pp. 1128–1142, Sep. 2004.
- [3] C. W. Günther and W. M. Van Der Aalst, “Fuzzy mining—adaptive process simplification based on multi-perspective metrics,” in *International conference on business process management*, 2007, pp. 328–343.
- [4] A. K. A. de Medeiros, A. J. Weijters, and W. M. van der Aalst, “Genetic process mining: an experimental evaluation,” *Data Mining and Knowledge Discovery*, vol. 14, no. 2, pp. 245–304, 2007.
- [5] J. C. Buijs, B. F. Van Dongen, and W. M. van Der Aalst, “On the role of fitness, precision, generalization and simplicity in process discovery,” in *OTM Confederated International Conferences “On the Move to Meaningful Internet Systems”*, 2012, pp. 305–322.
- [6] A. J. Weijters and W. M. Van der Aalst, “Rediscovering workflow models from event-based data using little thumb,” *Integrated Computer-Aided Engineering*, vol. 10, no. 2, pp. 151–162, 2003.
- [7] W. M. Van der Aalst, “Process mining in the large: a tutorial,” in *European Business Intelligence Summer School*, 2013, pp. 33–76.
- [8] A. Rozinat and W. M. Van der Aalst, “Conformance checking of processes based on monitoring real behavior,” *Information Systems*, vol. 33, no. 1, pp. 64–95, 2008.
- [9] A. Adriansyah, B. F. van Dongen, and W. M. van der Aalst, “Towards robust conformance checking,” in *International Conference on Business Process Management*, 2010, pp. 122–133.
- [10] A. Adriansyah, B. F. van Dongen, and W. M. van der Aalst, “Conformance checking using cost-based fitness analysis,” in *2011 IEEE 15th International Enterprise Distributed Object Computing Conference*, 2011, pp. 55–64.
- [11] A. Burattin, F. M. Maggi, and A. Sperduti, “Conformance checking based on multi-perspective declarative process



۲۷ الی ۲۸ آذرماه ۱۳۹۸

توسعه الگوریتم EFIM برای کشف مجموعه اقلام با متوسط ارزش بالا

آزاده سلطانی^۱، محمود سلطانی^۲

^۱ گروه مهندسی کامپیوتر، دانشگاه بجنورد، بجنورد

a.soltani@ub.ac.ir

^۲ گروه مهندسی کامپیوتر، دانشگاه صنعتی قوچان، قوچان

m.soltani@qiet.ac.ir

چکیده

مسئله استخراج مجموعه اقلام با ارزش بالا توسعه‌ای از مسئله کشف قوانین انجمنی است که دو هدف را دنبال می‌کند؛ اول اینکه اقلام بتوانند بیش از یک بار در تراکنتها حضور داشته باشند و همچنین ارزش و اهمیت اقلام یکسان در نظر گرفته نشود. از نقاط ضعف روشهای ارائه شده برای این مسئله، این است که در آنها، مجموعه‌هایی با تعداد اقلام بیشتر، شانس بیشتری برای انتخاب شدن به عنوان مجموعه‌ای با ارزش دارند. در صورتیکه در دنیای واقعی اینگونه نیست و مجموعه‌هایی با تعداد اقلام بالا ارزش زیادی برای تصمیم‌گیری ندارند. به منظور حل این مشکل، الگوریتمهای متعددی معرفی شده‌اند که هدفشان یافتن مجموعه‌هایی با متوسط ارزش بالاست. اگرچه تلاشهای خوبی در این زمینه انجام پذیرفته است، هنوز نیاز به الگوریتمهایی با کارایی بالاتر وجود دارد. در این تحقیق الگوریتم EFIM که روشی کارا در استخراج مجموعه‌هایی با ارزش بالاست را به گونه‌ای توسعه داده ایم که بتواند مجموعه اقلام با متوسط ارزش بالا را بیابد. آزمایشهای انجام شده نشان میدهد روش پیشنهادی کارایی بهتری نسبت به روش MHAI دارد.

کلمات کلیدی

داده کاوی، الگوهای پرتکرار، الگوهای با متوسط ارزش بالا، الگوریتم EFIM.

۱- مقدمه

رشد روزافزون داده‌های دیجیتالی منجر به ظهور تکنیکهای متعدد داده‌کاوی گردیده است. یکی از این تکنیک‌ها کشف الگوهای پرتکرار است که اولین بار توسط آگراوال در سال ۱۹۹۳ معرفی شد [1] و از آن لحظه تا کنون تحقیقات گسترده‌ای در این زمینه انجام گرفته است [2]-[4]. در تعریف اولیه این مسئله، فرض شده است که هر قلم داده تنها یک بار می‌تواند در هر تراکنتش قرار بگیرد؛ همچنین ارزش و اهمیت اقلام باهم برابر است. این در حالیست که در دنیای واقعی یک قلم داده میتواند بیش از یک بار در یک تراکنتش اتفاق بیافتد و همچنین ارزش اقلام ممکن است متفاوت باشد به طور مثال در یک فروشگاه، یک مشتری میتواند از یک کالا چند عدد در سبد خرید خود قرار دهد و ارزش کالاها که مثلا میتواند قیمت آنها در نظر گرفت نیز با هم متفاوت است. برای حل این مشکل، زمینه تحقیقاتی جدیدی ارائه شد که هدفش استخراج مجموعه اقلام با ارزش است و نه صرفا پرتکرار. در مسئله کشف

الگوهای پرتکرار، بر اساس خاصیت بستار رو به پایین^۲ فضای جستجو به سادگی هرس می‌شد. به عبارت دیگر، بر اساس خاصیت apriori، اگر مجموعه‌ای کم تکرار باشد هیچ ابرمجموعه‌ای از آن نمیتواند پرتکرار باشد. بنابراین چنانچه در فضای جستجو به یک مجموعه با تعداد کمتر از حد آستانه برسیم میتوانیم زیر درخت آن مجموعه را در فضای جستجو هرس کنیم. این خاصیت برای ارزش مجموعه‌ها صادق نیست و ممکن است مجموعه‌ای کم ارزش باشد ولی با افزوده شدن یک قلم داده، با ارزش بالا، ارزش ابرمجموعه آن بیشتر از حداقل آستانه گردد. بنابراین برای حل مسئله کشف مجموعه اقلام با ارزش، مهمترین چالش هرس فضای جستجو است و محققان تلاش کرده‌اند با بالاتر برآورد کردن^۳ ارزش یک مجموعه به ویژگی‌ای دست یابند که خاصیت بستار رو به پایین را داشته باشد. اولین بار لین و همکارانش ویژگی TWU را معرفی کردند [5] که برابر است با مجموعه ارزش تراکنتهایی که شامل مجموعه x هستند و دارای خاصیت بستار رو به پایین است. در فاز اول تمام مجموعه‌هایی که دارای TWU بالاتر از حداقل

²Downward Closure Property

^۳ Overestimation

¹ item

مجموعه قلم تعریف می‌کنیم به عبارت دیگر: $X = \{i_1, i_2, \dots, i_k\}$ و آن را یک مجموعه‌ی k تایی می‌نامیم و اگر مجموعه قلم X ، زیرمجموعه‌ای از یک تراکنش دلخواه باشد ($X \subseteq T_q$) آنگاه می‌گوییم T_q شامل X است و یا X در T_q ظاهر شده است. برای هر قلم، دو مقدار ارزش بیرونی (مثال: سود واقعی یک قلم در فروشگاه) و ارزش درونی (مثال: تعداد قلم خریداری شده در یک سبد خرید) تعریف شده است. ارزش کل یک مجموعه قلم داده ($u(i_j, T_q)$) برابر است با حاصلضرب دو ارزش بیرونی و درونی آن. ارزش مجموعه قلم X در تراکنش T_q که با $u(X, T_q)$ نمایش داده می‌شود توسط فرمول زیر محاسبه می‌گردد.

$$u(X, T_q) = \sum_{i_j \in X \wedge X \subseteq T_q} q(i_j, T_q) \times pr(i_j)$$

ارزش مجموعه قلم X در پایگاه داده از جمع ارزش این مجموعه در تمام تراکنش‌ها حاصل می‌شود. این مقدار از فرمول زیر محاسبه می‌گردد.

$$u(X) = \sum_{X \subseteq T_q \wedge T_q \in D} u(X, T_q)$$

همچنین ارزش یک تراکنش که با $tu(T_q)$ نمایش داده می‌شود از مجموع ارزش تمام قلم داده‌های آن و توسط فرمول زیر بدست می‌آید.

متوسط ارزش مجموعه قلم k تایی در تراکنش T_q با $au(X, T_q)$ نمایش داده می‌شود و اینگونه محاسبه می‌گردد:

$$tu(T_q) = \sum_{i_j \in T_q} u(i_j, T_q)$$

$$au(X, T_q) = \frac{\sum_{i_j \in X \wedge X \subseteq T_q} q(i_j, T_q) \times pr(i_j)}{k}$$

در این فرمول k تعداد اقلام موجود در X است (به عبارت دیگر $|X| = k$).

در مسئله کشف اقلام با متوسط ارزش بالا، به دنبال یافتن تمام مجموعه اقلامی (HAUIs) هستیم که متوسط ارزش آنها در کل پایگاه داده D ، از مینیمم از پیش تعریف شده (minUtil) کمتر نباشد.

۲-۲- مرور کارهای گذشته

کشف مجموعه قلم‌های با ارزش بالا، توسعه‌ای از مسئله‌ی کشف الگوهای پرتکرار است. همانطور که در بخش قبل بیان شد، مجموعه‌های با ارزش، خاصیت بستار رو به پایین ندارند و نمیتوان فضای جستجو را در پروسه کشف آنها، به راحتی هرس کرد. زیرا اگر مجموعه قلمی دارای ارزش کمتر از مینیمم باشد ممکن است ابر مجموعه‌ای از آن، دارای ارزش بالاتر از مینیمم شود. به عبارت دیگر یک قلم داده‌ی با ارزش بالا به آن اضافه شود که ارزش کم اقلام قبلی را جبران کند. بنابراین، در الگوریتم‌های کشف مجموعه‌های با ارزش، یافتن معیاری که بتواند جایگزین ارزش مجموعه قلم شود و دارای خاصیت بستار رو به پایین باشد دارای اهمیت بالایی است.

آستانه هستند پیدا میشود سپس با محاسبه مقدار واقعی ارزش این مجموعه‌ها، در فاز دوم مجموعه‌هایی با ارزش بالا استخراج می‌شوند. الگوریتم‌های متعددی در این خصوص ارائه شده است [5]-[8].

از آنجاییکه ارزش یک مجموعه با مجموع ارزش اقلام آن برابر است، هر چه تعداد عناصر یک مجموعه بیشتر باشد ارزش کل آن بیشتر خواهد شد. بنابراین مجموعه‌هایی با طول بیشتر شانس بیشتری خواهند داشت که به عنوان الگوی با ارزش انتخاب شوند. برای برطرف کردن این مشکل، مسئله کشف الگوهای با متوسط ارزش بالا مطرح شد که با تقسیم کردن ارزش یک مجموعه بر تعداد اقلام آن، وابستگی به طول را از بین می‌بردند. الگوریتم‌های ارائه شده برای این مسئله جدید سه گروه تقسیم می‌شوند. دسته اول مبتنی بر الگوریتم apriori بوده و به صورت مرحله ای و با تولید مجموعه‌های کاندید الگوهای با متوسط ارزش بالا را تولید می‌کنند [9], [10]. این روش‌ها نیاز به پیمایش‌های متعدد پایگاه داده و تولید مجموعه‌های کاندید دارند. دسته دوم روش‌های مبتنی بر درخت هستند که با نگهداری مجموعه داده در درخت و پردازش درخت به صورت بازگشتی بدون نیاز به اسکن مجدد میتوانند الگوهای با ارزش را بیابند [11], [12]. گروه سوم، الگوریتم‌های مبتنی بر لیست هستند که با ذخیره اطلاعات لازم در لیست‌ها و ترکیب آنها به صورت اول عمق می‌توانند الگوهای با ارزش را بیابند [13]-[15]. با وجود ارائه شدن روش‌های متعدد، هنوز نیاز به الگوریتم‌هایی با کارایی بالاتر احساس میشود.

در این تحقیق الگوریتم [7] EFIM را گسترش داده‌ایم. این الگوریتم، روشی کارا در کشف مجموعه‌هایی با ارزش بالا است. هدف از روش پیشنهادی این است که با تغییر در مراحل مختلف الگوریتم EFIM و روش‌های هرس تعریف شده در آن، بتوان الگوهای با متوسط ارزش بالا را استخراج نمود. ساختار ادامه الگوریتم به شرح زیر است: بخش دوم به بیان مسئله و کارهای انجام پذیرفته قبلی می‌پردازد. روش پیشنهادی در فصل سوم توضیح داده خواهد شد. نتایج آزمایش‌ها و ارزیابی‌ها در بخش چهارم ارائه شده و در نهایت بخش پنجم شامل نتیجه‌گیری است.

۲- تعریف مسئله و کارهای انجام شده قبلی

در این فصل ابتدا مسئله کشف مجموعه اقلام با ارزش متوسط بالا تعریف شده و سپس به مرور کارهای انجام شده قبلی می‌پردازد.

۲-۱- بیان مسئله

فرض کنید $D = \{T_1, T_2, \dots, T_n\}$ مجموعه‌ی تراکنش‌ها و $I = \{i_1, i_2, \dots, i_m\}$ مجموعه‌ی کل اقلام مجزای موجود در مجموعه داده باشد. یک تراکنش دلخواه را با $T_q = \{i_1, i_2, \dots, i_l\}$ نشان می‌دهیم ($T_q \in D$) که دارای یک شناسه‌ی یکتا (q) می‌باشد که TID نامیده می‌شود. همچنین زیرمجموعه‌ای از اقلام را، یک

از ارزش X بیشتر می‌شود؛ در صورتیکه متوسط ارزش T_q از متوسط ارزش X کمتر خواهد بود. بنابراین در این دسته از الگوریتم‌ها از معیار ماکزیمم ارزش تراکنش^۵ (TMU) استفاده می‌شود. در هر مجموعه‌ای ماکزیمم ارزش از متوسط ارزش تمام زیرمجموعه‌هایش بیشتر است؛ بنابراین یک حد بالا برای متوسط ارزش محسوب می‌شود. بنابراین برای هرس کردن از حد بالای متوسط ارزش^۱ ($AUUB$) استفاده می‌شود که برابر است با مجموع ماکزیمم ارزش تراکنشهایی که مجموعه مورد نظر در آنها رخ داده است. برای این ویژگی خاصیت بستار رو به پایین برقرار است.

اولین الگوریتم برای حل مسئله‌ی کشف مجموعه‌های با متوسط ارزش بالا روش TPAU است [۱۰] که مبتنی بر Apriori است. این الگوریتم مشابه روش دو مرحله‌ای، ابتدا مجموعه‌های کاندید را، که حد بالای متوسط ارزش آنها از حداقل ارزش از پیش تعریف شده کمتر نیست را پیدا میکند و در مرحله بعد با اسکن مجدد متوسط ارزش واقعی را پیدا میکند. پیمایش‌های متعدد پایگاه داده و تولید حجم بالای مجموعه‌های کاندید از نقاط ضعف این روش است. روش HAU1-Growth یک روش درختی است [۱۱] که با نگهداری اطلاعات مربوط به اقلام و وزن آنها در تراکنشها نیاز به چندین پیمایش و همچنین تولید مجموعه‌های کاندید را از بین می‌برد. HAU1-Tree [۱۲] روش درختی دیگری است که با استفاده از شناسه‌ی تراکنشها به پایگاه داده‌ی تصویر شده در حافظه‌ی اصلی دسترسی پیدا می‌کند و در نتیجه دیگر نیازی به چندین پیمایش از پایگاه داده نخواهد داشت. روش‌های جدید دیگر همچون MHAI [۱۴] و EHAUPM [15] که همه آنها مبتنی بر لیست هستند نیز اخیراً معرفی شده‌اند.

۳- روش پیشنهادی

الگوریتم EFIM یک روش مبتنی بر لیست است که میتواند تمام مجموعه‌های با ارزش را استخراج کند. هدف از این تحقیق، توسعه‌ی الگوریتم EFIM است به گونه‌ای که بتواند مجموعه اقلام با متوسط ارزش بالا را بیابد. به این منظور در تعریف کران متوسط ارزش یک مجموعه و روش هرس این مجموعه تغییراتی اعمال کرده‌ایم. در این بخش به معرفی مراحل الگوریتم پیشنهادی می‌پردازیم که مبتنی بر EFIM است:

۳-۱- محاسبه AUUB

در ابتدای الگوریتم، با پیمایش پایگاه داده AUUB تمام اقلام، محاسبه شده و پایگاه داده بازبینی شده با حذف آیتم‌هایی که این مقدار در آنها از $minUtil$ کمتر باشد، ساخته می‌شود. در ادامه آیتم‌ها بر اساس مقدار AUUB مرتب می‌شوند (در مثالها برای

این معیار باید یک تخمین بیش از حد از ارزش مجموعه باشد که هرس کردن توسط آن خللی به درستی الگوریتم وارد نسازد. اولین بار لیو و همکارانش، مفهوم ارزش مبتنی بر تراکنش (TWU) را معرفی کردند [۵]. برای هر مجموعه قلم X ، TWU برابر است با مجموع ارزش تراکنشهایی که X عضو آنها است. این مقدار یک حد بالاست و همیشه بزرگتر یا مساوی ارزش واقعی مجموعه قلم X است؛ بنابراین اگر کمتر از مینیمم باشد میتوان گفت ارزش مجموعه قلم نیز کمتر از مینیمم است. همچنین این ویژگی دارای خاصی بستار رو به پایین است؛ هر توسعه Y از مجموعه قلم X ، فقط میتواند در همان تراکنشهایی رخ دهد که X در آنها ظاهر شده است؛ بنابراین مجموعه ارزش تراکنشهایی که Y در آنها رخ داده است، کمتر مساوی مجموع ارزش تراکنشهایی است که X در آنها رخ داده است. پس می‌توان گفت اگر مجموع ارزش تراکنشهایی که X در آنها رخ داده است کمتر از مینیمم باشد تمام ابرمجموعه‌های X نیز دارای مجموع ارزش تراکنش کمتر از مینیمم هستند و میتوانند هرس شوند. بنابراین لیو و همکارانش الگوریتمی دو مرحله‌ای معرفی کرده‌اند. در مرحله اول، تمام مجموعه‌های کاندید که TWU آنها کمتر از حداقل از پیش تعریف شده نباشد را بر اساس الگوریتم Apriori و با چندین گذر از پایگاه داده پیدا می‌کند. سپس در مرحله دوم، با پیمایشی دیگر ارزش واقعی این مجموعه‌ها را به دست آورده و مجموعه‌های با ارزش مشخص می‌شوند.

از آنجاییکه در تعریف مجموعه‌ی با ارزش بالا، توجهی به طول مجموعه نمی‌شود، مجموعه‌های طولانی‌تر شانس بیشتری پیدا می‌کنند که به عنوان مجموعه‌ی با ارزش بالا انتخاب شوند. در صورتیکه در دنیای واقعی مجموعه‌های کوتاه‌تر اهمیت بیشتری دارند. الگوریتم‌های متعددی معرفی شده که سعی کرده‌اند طول مجموعه را نیز در پروسه کشف الگوهای با ارزش در نظر بگیرند. به طور مثال در [۱۶] علاوه بر ارزش مجموعه، محدودیتی نیز روی طول مجموعه‌های خروجی اعمال می‌شود و تنها مجموعه‌هایی به عنوان الگوهای با ارزش بالا معرفی می‌شوند که طولشان از حداقل از پیش تعریف شده کمتر باشد. روش‌های متعدد دیگری نیز برای مقابله با این مشکل ارائه شده‌اند که اکثر آنها به جای ارزش یک مجموعه از متوسط ارزش آن مجموعه استفاده می‌کنند. به عبارت دیگر برای دخالت دادن طول مجموعه در پروسه کشف، ارزش کل مجموعه را بر طول آن نیز تقسیم می‌کنند. بر خلاف مسئله‌ی کشف الگوهای با ارزش که در آن، امکان استفاده از ارزش مبتنی بر تراکنش (TWU)، برای هرس کردن اولیه وجود داشت، در کشف مجموعه‌های با متوسط ارزش بالا، نمی‌توان از متوسط ارزش مبتنی بر تراکنش برای هرس استفاده نمود. زیرا متوسط ارزش یک تراکنش ممکن است از متوسط ارزش زیرمجموعه‌هایش کمتر باشد. به طور مثال فرض کنید، $T_q = \{a, b\}$ و $X = \{b\}$ ، و ارزش b از a بیشتر باشد، آنگاه ارزش T_q برابر است با مجموع ارزش a و b که

^۵ Transaction Maximum Utility

^۶ Average Utility Upper Bound

^۴ Transaction Weighted Utilization

در این فرمول $MRe(X, T_q)$ برابر است با مجموع ارزش اقلامی که پس از X در T_q قرار گرفته اند و افزودن ارزش آنها سبب افزایش متوسط ارزش $\frac{U(X, T_q) + MRe(X, T_q)}{|X| + MReN}$ خواهد شد و $MReN$ تعداد این اقلام است. به طور حتم ارزش این اقلام از متوسط ارزش کنونی باید بیشتر باشد. برای به دست آوردن $MRe(X, T_q)$ کافیست ارزش اقلامی که در T_q قرار دارند و قلم داده متناظرشان پس از X قرار دارد را به ترتیب نزولی مرتب کرده و تا زمانیکه افزودن این ارزش به ارزش کل سبب افزایش مقدار نهایی $\frac{U(X, T_q) + MRe(X, T_q)}{|X| + MReN}$ می‌شود، مقدار ارزش متناظر را به $MRe(X, T_q)$ اضافه نموده و به تعداد $MReN$ نیز یک واحد اضافه می‌کنیم.

اگر Z یک توسع دلخواه از X باشد که شامل قلم داده z هم باشد، بر اساس تعریف فوق می‌توان گفت $LAU(X, z) \geq AU(Z)$. بنابراین اگر متوسط ارزش محلی یک توسعه از X کمتر از \minUtil باشد میتوان تمام آن زیردرخت را از فضای جستجو هرس نمود.

۳-۴-۲- متوسط ارزش زیردرخت ^۹ (SAU)

متوسط ارزش زیر درخت برای مجموعه قلم داده X بر مبنای توسعه قلم $z \in E(X)$ اینگونه تعریف می‌شود:

$$SAU(X, z) = \sum_{T_q \in D, X \subseteq T_q, z \in T_q} \frac{U(\{X \cup z\}, T_q) + MRe(\{X \cup z\}, T_q)}{|X| + MReN}$$

$MRe(\{X \cup z\}, T_q)$ مشابه قبل به دست می‌آید با این تفاوت که ارزش اقلامی در محاسبه آن نقش دارند که پس از z قرار گرفته اند. بر اساس تعریف فوق می‌توان گفت $SAU(X, z) \geq AU(X \cup z)$. بنابراین اگر متوسط ارزش زیردرخت یک توسعه از X کمتر از \minUtil باشد میتوان زیردرخت $X \cup z$ را از فضای جستجو هرس نمود.

۳-۵-۵- مراحل الگوریتم پیشنهادی

الگوریتم اصلی به عنوان ورودی، حداقل ارزش تعریف شده توسط کاربر (\minUtil) و پایگاه داده تراکنشها (D) را دریافت می‌کند و به عنوان خروجی لیست تمام مجموعه اقلامی که شرط حداقل متوسط ارزش را دارا هستند را بر می‌گرداند.

در مرحله اول مجموعه قلم داده X را تهی در نظر می‌گیرد. به ازای هر قلم داده i عضو I ، متوسط ارزش محلی ($LAU(X, i)$) را محاسبه کرده و تمام اقلامی که متوسط ارزش محلی آنها بیشتر از حداقل \minUtil است را به مجموعه $Secondary(X)$ اضافه می‌کند.

$$Secondary(X) = \{i \mid i \in I, LAU(X, i) \geq \minUtil\}$$

نمایش بهتر از ترتیب حروف الفبا استفاده شده است). بر اساس این ترتیب می‌توان فضای جستجو را به صورت درختی در نظر گرفت. در این فضای جستجو به صورت اول عمق پیش رفته و در هر سطح، یک قلم داده به مجموعه X اضافه می‌شود. اقلامی که بتوان از آنها در توسعه مجموعه X استفاده کرد با $E(X)$ نشان داده می‌شوند. $E(X) = \{a \mid a \in I, a > b, \forall b \in X\}$

۳-۲- ایجاد تصویر و پرتو پایگاه داده^۷

به منظور کاهش حجم پایگاه داده و در نتیجه کاهش زمان محاسبه ارزش واقعی مجموعه ها، در EFIM، پرتو پایگاه داده ایجاد میشود و در هر سطح تنها تراکنشهایی نگهداشته می‌شوند که شامل اقلام $E(X)$ هستند. همچنین سایر اقلام نیز حذف می‌شوند. پرتو تراکنش T_q نسبت به مجموعه X اینگونه تعریف میشود: $X_{T_q} = \{i_x \in T_q \mid i_x \in E(X)\}$. بنابراین به مرور که در درخت جستجو پیش می‌رویم حجم پایگاه داده کاهش یافته و زمان جستجو کم می‌شود.

۳-۳- ادغام تراکنشها

با ایجاد پرتو از تراکنشها، طول آنها کاهش می‌یابد و احتمال ایجاد تراکنشهای که دارای مجموعه اقلام یکسان باشند، افزایش می‌یابد. در این مرحله، با ادغام تراکنشهایی با مجموعه اقلام یکسان، باز هم سعی می‌شود حجم پایگاه داده کاهش یابد. ادغام دو تراکنش، به معنی جایگزین کردن دو تراکنش یکسان با یک تراکنش است که دقیقاً دارای همان مجموعه قلم داده‌ها باشد و ارزش هر قلم داده نیز با جمع ارزشش در دو تراکنش قبلی برابر باشد. برای یافتن تراکنشهای یکسان، پایگاه داده بر اساس ترتیب تراکنشها مرتب شده، تراکنشهای پشت سرهم مقایسه می‌شوند و تراکنشهای برابر با هم ادغام می‌گردند.

۳-۴-۴- کرانهای جدید متوسط ارزش در الگوریتم

پیشنهادی

در الگوریتم پیشنهادی، کرانهای بالای ارزش، که در الگوریتم EFIM تعریف شده بود، اصلاح شده است که بتوانند کران بالای متوسط ارزش باشند. در این بخش به معرفی کرانهای جدید می‌پردازیم:

۳-۴-۱- متوسط ارزش محلی ^۸ (LAU)

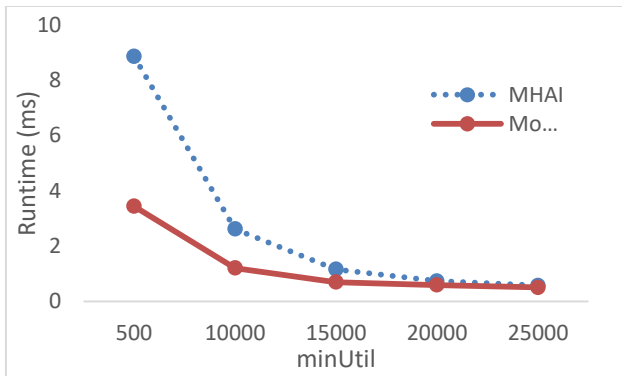
متوسط ارزش محلی برای مجموعه قلم داده X بر مبنای توسعه قلم $z \in E(X)$ اینگونه تعریف می‌شود:

$$LAU(X, z) = \sum_{T_q \in D, X \subseteq T_q, z \in T_q} \frac{U(X, T_q) + MRe(X, T_q)}{|X| + MReN}$$

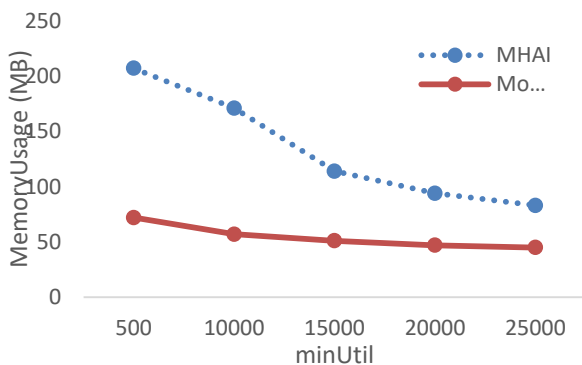
⁷ Database Projection

⁸ Local Average Utility

⁹ Sub-tree Average Utility



شکل 33- زمان اجرای دو الگوریتم MHAJ و Modified برای مجموعه داده Retail با minUtil های متفاوت



شکل 34- زمان اجرای دو الگوریتم MHAJ و Modified برای مجموعه داده Retail با minUtil های متفاوت

۵- نتیجه

کشف مجموعه اقلام با ارزش متوسط بالا، زیرمسئله کشف الگوهای با ارزش است که با تقسیم ارزش کل یک مجموعه بر طول آن توانسته است تاثیر طول مجموعه ها را بر روی ارزش نهایی آنها از بین ببرد. در این تحقیق، با گسترش الگوریتم EFIM که روشی برای کشف الگوهای با ارزش است توانسته ایم روشی برای کشف مجموعه اقلام با متوسط ارزش بالا ارائه دهیم که نسبت به روش MHAJ کارایی بهتری از نظر زمان اجرا و حافظه مصرفی دارد.

مراجع

- [1] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *Acm sigmod record*, 1993, vol. 22, pp. 207–216.
- [2] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th int. conf. very large data bases, VLDB*, 1994, vol. 1215, pp. 487–499.

مجموعه Secondary بر اساس AUUB مرتب میشود و تمام اقلامی که عضو این مجموعه نیستند از تراکنشها حذف می‌شوند. سپس به ازای تمام اقلام موجود در مجموعه Secondary، متوسط ارزش زیردرخت $(SAU(X, i))$ محاسبه می‌شود. تمام اقلامی که متوسط ارزش زیردرخت آنها بیشتر از حداقل minUtil است به مجموعه Primary(X) اضافه می‌شوند.

$Primary(X) = \{i | i \in Secondary, SAU(X, i) \geq minUtil\}$
در مرحله بعد به کمک این دو مجموعه یک تابع بازگشتی فراخوانی می‌شود که به کمک آن می‌توان مجموعه اقلام با متوسط ارزش بالا را به دست آورد.

در تابع بازگشتی ابتدا بر اساس هر عنصر i عضو Primary(X)، توسعه $Y = X \cup i$ را به دست آورده و متوسط ارزش آنرا محاسبه می‌کند. در صورتیکه کمتر از حداقل minUtil نباشد به مجموعه خروجی اضافه می‌گردد. سپس برای مجموعه Y دوباره مجموعه های Secondary و Primary تولید شده و تابع بازگشتی با این ورودیها دوباره فراخوانی می‌گردد.

۴- ارزیابی

در این بخش، روش پیشنهادی از نظر زمان اجرا و حافظه مصرفی با الگوریتم MHAJ مقایسه شده است. کد الگوریتم EFIM از پکیج [17] SPMF گرفته شده است. در این پکیج بیش از ۱۵۰ الگوریتم در زمینه داده کاوی و بخصوص کشف الگوهای پرتکرار و قوانین انجمنی به زبان جاوا پیاده سازی شده است. همچنین کد الگوریتم MHAJ بصورت کامل پیاده سازی شده است. جهت ارزیابی و مقایسه از مجموعه داده واقعی retail استفاده کرده ایم. این مجموعه داده شامل ۸۸۱۶۲ تراکنش و ۱۶۴۷۰ قلم داده است و متوسط طول هر تراکنش ۱۰٫۳ است. در آزمایش انجام شده، روش پیشنهادی و الگوریتم MHAJ را بر روی مجموعه داده Retail اعمال کرده ایم و حداقل آستانه برای ارزش (minUtil) را بین ۵۰۰۰ تا ۲۵۰۰۰ تغییر داده ایم. شکل‌های ۱ و ۲، زمان اجرا و حداکثر حافظه مصرفی الگوریتم پیشنهادی و روش MHAJ را در این آزمایش نشان می‌دهند. همانطور که مشهود است الگوریتم پیشنهادی هم از نظر زمان اجرا و هم از نظر حافظه مصرفی نسبت به MHAJ کارتر است.

Computer Information Systems and Industrial Management, 2014, pp. 33–42.

J. C.-W. Lin, T. Li, P. Fournier-Viger, T.-P. [13]

Hong, J. Zhan, and M. Voznak, “An efficient algorithm to mine high average-utility itemsets,” *Advanced Engineering Informatics*, vol. 30, no. 2, pp. 233–243, 2016.

U. Yun and D. Kim, “Mining of high average- [14]

utility itemsets using novel list structure and pruning strategy,” *Future Generation Computer Systems*, vol. 68, pp. 346–360, 2017.

J. C.-W. Lin, S. Ren, P. Fournier-Viger, and [15]

T.-P. Hong, “EHAUPM: Efficient high average-utility pattern mining with tighter upper bounds,” *IEEE Access*, vol. 5, pp. 12927–12940, 2017.

P. Fournier-Viger, J. C.-W. Lin, Q.-H. [16]

Duong, and T.-L. Dam, “FHM+: faster high-utility itemset mining using length upper-bound reduction,” in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, 2016, pp. 115–127.

P. Fournier-Viger, A. Gomariz, T. Gueniche, [17]

A. Soltani, C.-W. Wu, and V. S. Tseng, “SPMF: a java open-source pattern mining library,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3389–3393, 2014.

50] M. J. Zaki and K. Gouda, “Fast vertical [3] mining using diffsets,” in *Proc. 9th ACM SIGKDD*, 2003, pp. 326–335.

J. Han, J. Pei, Y. Yin, and R. Mao, “Mining [4] frequent patterns without candidate generation: A frequent-pattern tree approach,” *Data mining and knowledge discovery*, vol. 8, no. 1, pp. 53–87, 2004.

Y. Liu, W. Liao, and A. N. Choudhary, “A [5] Two-Phase Algorithm for Fast Discovery of High Utility Itemsets,” in *PAKDD*, 2005, vol. 3518, pp. 689–695.

P. Fournier-Viger, C.-W. Wu, S. Zida, and V. [6] S. Tseng, “FHM: faster high-utility itemset mining using estimated utility co-occurrence pruning,” in *International symposium on methodologies for intelligent systems*, 2014, pp. 83–92.

S. Zida, P. Fournier-Viger, J. C.-W. Lin, C.- [7] W. Wu, and V. S. Tseng, “EFIM: a fast and memory efficient algorithm for high-utility itemset mining,” *Knowledge and Information Systems*, vol. 51, no. 2, pp. 595–625, 2017.

C.-W. Lin, T.-P. Hong, and W.-H. Lu, “An [8] effective tree structure for mining high utility itemsets,” *Expert Systems with Applications*, vol. 38, no. 6, pp. 7419–7424, 2011.

T.-P. Hong, C.-H. Lee, and S.-L. Wang, [9] “Mining high average-utility itemsets,” in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, pp. 2526–2530.

T.-P. Hong, C.-H. Lee, and S.-L. Wang, [10] “Effective utility mining with the measure of average utility,” *Expert Systems with Applications*, vol. 38, no. 7, pp. 8259–8265, 2011.

C.-W. Lin, T.-P. Hong, and W.-H. Lu, [11] “Efficiently Mining High Average Utility Itemsets with a Tree Structure,” in *Intelligent Information and Database Systems*, 2010, pp. 131–139.

T. Lu, B. Vo, H. T. Nguyen, and T.-P. Hong, [12] “A new method for mining high average utility itemsets,” in *IFIP International Conference on*

۲۷ الی ۲۸ آذرماه ۱۳۹۸

تحلیل هم‌ارجاعی در متن با استفاده از یادگیری عمیق و تکنیک تصمیم‌گیری چندمعیاره

سمیرا حورعلی^۱، مرتضی زاهدی^۲، منصور فاتح^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
s.hourali@shahroodut.ac.ir

^۲ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
zahedi@shahroodut.ac.ir

^۳ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
mansour_fateh@shahroodut.ac.ir

چکیده

تحلیل هم‌ارجاعی یکی از گام‌های مهم در پردازش معنایی متن است که برای ارتقاء عملکرد سیستم‌هایی نظیر استخراج اطلاعات، خلاصه‌سازی متون، ترجمه ماشینی و پرسش و پاسخ کاربرد دارد. منظور از تحلیل هم‌ارجاعی شناسایی نامبری‌های موجود در متن اعم از ضمائر، عبارات اسمی و نامداری است که به یک موجودیت اشاره می‌کنند. هرچند طی چند دهه گذشته تحلیل هم‌ارجاعی همواره یک موضوع تحقیقاتی فعال بوده است؛ اما هنوز هم دقت آن در حد قابل قبول برای استفاده در کاربردهای درک متن نیست. انتخاب مرجع مناسب برای نامبری‌های موجود در متن تابع قواعد و منابع دانش مختلفی است که تطابق جنسیت، تعداد، ساختار رشته‌ای و سایر پارامترهای هم‌ارجاعی را تضمین کند. در این مقاله ابتدا توسط یادگیری عمیق نامبری‌های موجود در متن استخراج می‌شوند سپس توسط تکنیک تصمیم‌گیری چندمعیاره و روش وزن‌دهی بردار ویژه نامبری‌های هم‌ارجاعی استخراج می‌شوند. نتایج شبیه‌سازی نشان می‌دهد، روش پیشنهادی در مقایسه با سایر روش‌های مطرح و جدید در این زمینه عملکرد بهتری دارد. همچنین مقدار F1 روی پایگاه داده CONLL-2012 به میزان ۶/۳ درصد بهبود داده شده است.

کلمات کلیدی

تحلیل هم‌ارجاعی، نامبری، یادگیری عمیق، تصمیم‌گیری چندمعیاره، روش وزن‌دهی بردار ویژه.

۱- مقدمه

بحث فراهم می‌آورد. در ترجمه ماشینی، جایگزینی ضمائر با مراجع آن‌ها بسیار مهم است.

هرچند مسئله تحلیل مرجع یک موضوع تحقیقاتی فعال طی چهار دهه گذشته بوده است اما هنوز راه طولانی تا حل شدن کامل در پیش دارد. لبه دانش فعلی در زمینه این مسئله حداقل دارای سه نقطه ضعف است.

اول اینکه هیچ مدل محاسباتی دقیقی برای این مسئله وجود ندارد. رویکردهای جدید تحلیل هم‌ارجاعی [۲-۵] عموماً از روش‌های مبتنی بر یادگیری عمیق و یادگیری تقویتی استفاده می‌کنند اما نتوانسته‌اند مدل محاسباتی دقیقی برای حل این مسئله ارائه کنند. عدم استفاده از اطلاعات سطح خوشه و اطلاعات موجودیت‌ها

مسئله تحلیل هم‌ارجاعی یک موضوع جالب در پردازش زبان طبیعی است و در مسائلی مثل استخراج اطلاعات، خلاصه‌سازی متون، پاسخ به سؤالات و ترجمه ماشینی که در آن‌ها درک متن از اهمیت بالایی برخوردار است، کاربرد فراوان دارد. به‌عنوان مثال در کاربرد پاسخ به سؤالات، استفاده از تحلیل هم‌ارجاعی یکی از مراحل پیش‌پردازشی است که باعث بهبود بازخوانی^۱ می‌شود [۱]. در خلاصه‌سازی، تحلیل هم‌ارجاعی تأثیر زیادی در یافتن جملات پراهمیت در متن دارد و اطلاعاتی را در رابطه با موضوع اصلی

¹ Recall

است ضمیر، عبارات اسمی و موجودیت‌های نامدار توسط ساختار شبکه عصبی از پیش آموزش‌دیده^۶ استخراج شوند و در یافتن زنجیره‌های هم‌مرجع اطلاعات موجودیت‌ها نیز تأثیر داده شود. در ادامه این مقاله، در بخش دوم مدل در نظر گرفته شده برای حل مسئله را تشریح می‌کنیم. در بخش سوم نتایج شبیه‌سازی را ارائه می‌کنیم و در پایان، نتایج را شرح می‌دهیم.

۲- روش پیشنهادی

در این بخش ابتدا در فاز پیش‌پردازش چگونگی استخراج نامبری‌های موجود در متن شرح داده می‌شود؛ سپس طبق روش تصمیم‌گیری چندشاخصه ارائه شده و روش وزن‌دهی بردار ویژه نامبری‌های هم‌مرجع شناسایی و استخراج می‌شوند.

۲-۱- استخراج نامبری

روش BERT: در مدل پیشنهادی برای استخراج نامبری‌ها ابتدا توسط روش BERT [۱۶] تعبیه واژگان استخراج می‌شود. روش BERT در سایر زمینه‌های پردازش به نتایج ایده آلی دست یافته و طبق نتایج شبیه‌سازی نسبت به سایر روش‌های تعبیه واژگان نظیر ELMo و GloVe [۱۷، ۱۸] بهتر عمل می‌کند. BERT در واقع یک شبکه عصبی عمیق از پیش آموزش داده شده است که توسط مدل-های زبانی آموزش‌دیده و ویژگی‌های استخراج‌شده توسط این روش جهت بالا بردن کارایی روش پیشنهادی و شناسایی درست موجودیت‌های نامدار و نامبری‌ها مؤثر است. در تعبیه واژگان، هر واژه‌ای به بردار بازنمایشی^۷ نگاشت داده می‌شود. بنابراین هر واژه با بردار بازنمایشی که حاوی یکسری اطلاعات معنایی و ارتباطی از واژه است، نمایش داده می‌شود و بازنمایشی از واژگان با توجه به زمینه‌ی متنی که واژگان در آن قرار دارند، به دست می‌آید.

شبکه BiLSTM: پس از تعبیه واژگان باید دنباله کلمات از متن استخراج شود تا بررسی شود این دنباله‌ها نامبری هستند یا خیر. از آنجایی که شبکه‌های LSTM نسبت به RNN وابستگی‌های طولانی-تری را در یک دنباله از کلمات می‌توانند استخراج کنند، برای استخراج دنباله کلمات و ویژگی‌ها از شبکه‌های LSTM استفاده می‌شود. برای اینکه بتوان بازنمایشی برای هر واژه با توجه به کلمات قبل و بعد از آن بدست آورد، از LSTM دو جهته استفاده شده است. پس از BiLSTM از چندین لایه LSTM دیگر استفاده شده است، تا اطلاعات و ویژگی‌های بیشتر و مفیدتری از بردارهای بازنمایش استخراج شود، طبق نتایج شبیه‌سازی استفاده از این تکنیک در شناسایی نامبری‌ها مفید است. برای پیش‌بینی برچسب واژگان از یک لایه‌ی کاملاً متصل با تابع فعالیت Softmax استفاده شد، با این روش احتمال کلمه در هر یک از دسته‌ها (برچسب‌ها) بدست می‌آید.

باعث شده است این روش‌ها دقت پایین‌تری نسبت به روش پیشنهادی داشته باشند. دومین نقطه ضعف عدم دسترسی اکثریت سیستم‌های تحلیل مرجع کنونی به دانش پیچیده مورد نیاز برای تحلیل درست مرجع‌هاست، زیرا بسیاری از موارد هم‌ارجاعی تنها با کمک دانش جهان قابل حل هستند. سومین نقطه ضعف آن است که بیشتر کارهای یادگیری ماشین انجام شده تا بحال برای تحلیل مراجع در زبان انگلیسی بوده و اعمال این راه‌حل به دیگر زبان‌ها باعث بروز چالش در کار پوزحمت و زمان‌بر ایجاد اسناد حاشیه-نویسی می‌شود.

در حالت کلی روش‌های تحلیل هم‌ارجاعی به دو دسته مبتنی بر قاعده و مبتنی بر یادگیری ماشین تقسیم می‌شوند. در روش‌های مبتنی بر قاعده [۶-۹] موارد هم‌مرجع درون متن توسط مجموعه‌ای از قواعد دست‌نویس که توسط افراد خبره نوشته شده‌اند، استخراج می‌شوند. از مزایای این روش می‌توان به دقت بالا و سادگی طراحی اشاره کرد. اما قابلیت انعطاف این روش پایین است و باید برای هر زبان طبیعی، سیستم از ابتدا توسط افراد خبره طراحی شود.

روش‌های مبتنی بر یادگیری ماشین نیز به سه دسته روش‌های مبتنی بر یادگیری تقویتی [۱۰، ۴]، آماری [۱۱-۱۳] و مبتنی بر یادگیری عمیق [۲، ۳، ۱۴، ۱۵، ۱۶] تقسیم می‌شوند. در روش‌های آماری با ناظر به صورت دستی یا اتوماتیک داده‌های آموزشی برچسب‌گذاری شده‌اند. این داده‌ها برای یادگیری سیستم مورد استفاده قرار می‌گیرند. از طرف دیگر، در روش‌های بدون ناظر نیازی به داده‌های آموزشی نیست (یا داده‌ها آموزشی بسیار کمی مورد نیاز است) اما دقت این روش هنوز برای حل مسئله تحلیل مرجع پایین است. در روش‌های مبتنی بر یادگیری تقویتی عموماً، نامبری‌ها به عنوان حالت‌ها^۲ و اتصالات بین نامبری‌ها به عنوان عمل-ها^۳ در نظر گرفته شده‌اند؛ سپس طبق سیاست یادگیری تقویتی و تعریف توابع پاداش^۴ و خطا^۵ اتصالات هم‌ارجاعی تشخیص داده شده است. جدیدترین روش‌های ارائه‌شده در حیطه تحلیل هم-ارجاعی، روش‌های مبتنی بر یادگیری عمیق هستند. در این روش‌ها نامبری‌های موجود در متن توسط ساختار یادگیری عمیق استخراج شده‌اند سپس با تعریف توابع هزینه مختلف سیستم در جهت شناسایی درست زنجیره‌های هم‌مرجع، آموزش می‌بیند.

رویکرد این پژوهش برای حل مسئله تحلیل مرجع استفاده از دانش تصمیم‌گیری چندمعیاره است. معیارها یا ویژگی‌های استخراج شده برای نامبری‌ها توسط ساختار یادگیری عمیق در فاز پیش‌پردازش نقش مهمی در بیان معانی نامبری‌ها بر عهده دارند. در نظر گرفتن تمام ویژگی‌ها و وزن یا میزان اهمیت هر یک از آن‌ها برای یافتن نامبری‌های هم‌مرجع با نامبری موردنظر تأثیر زیادی در یافتن نامبری‌های هم‌مرجع دارد. همچنین در روش پیشنهادی سعی شده

² States

³ Actions

⁴ Reward

⁵ Punishment

⁶ Pre-trained neural network

⁷ Vector representation

- ۱- ماتریس M تشکیل داده می‌شود.
 - ۲- ماتریس $(M - \lambda I)$ تشکیل داده می‌شود.
 - ۳- دترمینان ماتریس $(M - \lambda I)$ را محاسبه کرده، مساوی صفر قرار داده و مقادیر λ محاسبه می‌شود.
- بزرگترین λ را λ_{max} می‌نامیم و آن را در رابطه (۷) قرار داده و مقادیر w_j ها محاسبه می‌شود.

$$(M - \lambda_{max} I) \cdot w = 0 \quad (7)$$

پس از محاسبه وزن ویژگی‌ها برای مراجع کاندید، طبق رابطه (۸) وزن مراجع جهت بررسی هم‌مرجع بودن یا نبودن با نامبری موردنظر به‌روز رسانی می‌شود.

$$w'_i = \frac{r_i w_i}{\sum_{i=1}^n r_i w_i} \quad (8)$$

رتبه‌بندی مراجع کاندید: پس از تشکیل ماتریس تصمیم و محاسبه وزن و مقدار ویژگی‌ها، مراجع کاندید باید برای نامبری موردنظر رتبه‌بندی شوند. جهت رتبه‌بندی مراجع از روش تصمیم‌گیری چندمعیاره [۱۹] ELECTRE استفاده می‌شود.

گام ۱- بی‌مقیاس کردن ماتریس تصمیم

در این مرحله سعی می‌شود معیارها با ابعاد مختلف به معیارهایی بدون بعد تبدیل شوند برای بی‌مقیاس کردن از رابطه (۹) استفاده می‌شود.

$$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^m r_{ij}^2}} \quad (9)$$

گام ۲- تعیین ماتریس تصمیم وزن‌دار نرمال شده

ماتریس تصمیم وزن‌دار از ضرب ماتریس تصمیم بی‌مقیاس شده در وزن معیارها که توسط روش بردار ویژه محاسبه شد، به دست می‌آید.

$$v_{ij} = w'_j * n_{ij} \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n \quad (10)$$

گام ۳- تشکیل مجموعه معیارها موافق و مخالف

برای هر زوج مرجع k و e مجموعه معیارها به دو زیر مجموعه موافق و مخالف تقسیم می‌شوند. مجموعه موافق (S_{ke}) مجموعه‌ای از معیارهایی است که در آن مرجع k به مرجع e ترجیح دارد و مجموعه مکمل آن مجموعه مخالف (D_{ke}) می‌باشد مجموعه معیارهای موافق برای معیارهای مثبت و منفی به ترتیب به‌صورت زیر تعریف می‌شوند.

$$S_{ke} = \{j | v_{kj} \geq v_{ej}\} \quad (11)$$

$$D_{ke} = \{j | v_{kj} \leq v_{ej}\}$$

مجموعه معیارهای مخالف برای مثبت و منفی به ترتیب به‌صورت زیر تعریف می‌شود.

$$D_{ke} = \{j | v_{kj} < v_{ej}\} = J - S_{ke} \quad (12)$$

$$D_{ke} = \{j | v_{kj} > v_{ej}\} = J - S_{ke}$$

گام ۴- تشکیل ماتریس توافق

ماتریس توافق یک ماتریس مربعی است که بعد آن تعداد مراجع می‌باشد. هر یک از درایه‌های این ماتریس، شاخص توافق بین دو مرجع نامیده می‌شود. مقدار این شاخص، از جمع وزن معیارهایی که در مجموعه موافق وجود دارند، به دست می‌آید. به عبارت دیگر

تابع خطا: تابع خطا برای محاسبه‌ی اختلاف مقدار پیش‌بینی شده از مقدار واقعی و استخراج نامبری‌ها از میان دنباله کلمات استخراج شده از شبکه BiLSTM تعریف می‌گردد. هدف نهایی شبکه بدست آوردن وزن‌هایی است که این تابع را به حداقل مقدار برساند. تابع خطا در این مقاله Cross Entropy در نظر گرفته شده است فرمول این تابع به‌صورت (۱) است.

$$L = -\text{LOG}(p(y)) \quad (1)$$

$p(y)$ در فرمول رابطه (۱)، خروجی شبکه است و احتمال دنباله‌ای از برجسب‌های y را نشان می‌دهد.

$$p(y) = \prod_{t=1}^m p_t[y_t] \quad (2)$$

m بیشترین تعداد واژه در یک جمله را نشان می‌دهد.

۲-۲- تشخیص نامبری‌های هم‌مرجع

پس از استخراج نامبری‌ها و ویژگی‌های آن‌ها، ماتریس تصمیم زیر برای مراجع (نامبری‌های) کاندید شده جهت بررسی هم‌مرجع بودن یا نبودن با نامبری مدنظر تشکیل می‌شود.

$$M = \begin{matrix} & \begin{matrix} f_1 & f_2 & \dots & f_n \end{matrix} \\ \begin{matrix} c_1 \\ c_2 \\ \dots \\ c_m \end{matrix} & \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} \end{matrix} \quad (3)$$

در ماتریس رابطه (۳) c_i مرجع کاندید i ام، f_j ویژگی j ام و r_{ij} مقدار استخراج شده گام قبل برای ویژگی j ام مرجع i ام است. وزن یا میزان اهمیت هر یک از ویژگی‌ها توسط روش بردار ویژه و به‌صورت زیر محاسبه می‌شود.

روش بردار ویژه جهت محاسبه وزن ویژگی‌ها: در این روش w_j ها به‌گونه‌ای تعیین می‌شوند که روابط زیر صادق باشند.

$$\begin{aligned} r_{11}w_1 + r_{12}w_2 + \dots + r_{1n}w_n &= \lambda \cdot w_1 \\ r_{21}w_1 + r_{22}w_2 + \dots + r_{2n}w_n &= \lambda \cdot w_2 \\ &\vdots \\ r_{n1}w_1 + r_{n2}w_2 + \dots + r_{nn}w_n &= \lambda \cdot w_n \end{aligned} \quad (4)$$

وزن عناصر:

$$w_j = \frac{1}{\lambda} \sum_{i=1}^n r_{ij} w_i \quad j = 1, 2, \dots, n \quad (5)$$

دستگاه معادلات فوق را به‌صورت رابطه (۶) می‌توان نشان داد:

$$M \cdot w = \lambda \cdot w \quad (6)$$

در رابطه (۶) M ماتریس تصمیم، w بردار ویژه و λ مقدار ویژه برای ماتریس A است. مراحل زیر جهت محاسبه w_j در روش بردار ویژه به‌صورت زیر است:

هر عضو ماتریس G نیز نشان‌گر رابطه تسلط مابین مراجع می‌باشد.

گام ۸- تشکیل ماتریس تسلط نهایی

ماتریس تسلط نهایی (H) از ضرب تک‌تک درایه‌های ماتریس تسلط موافق (F) در ماتریس تسلط مخالف (G) حاصل می‌شود.

$$h_{ke} = f_{ke} \cdot g_{ke} \quad (19)$$

گام ۹- انتخاب بهترین مرجع

ماتریس تسلط نهایی (H) ترجیحات جزئی مراجع را بیان می‌کند. به‌طور مثال، اگر مقدار h_{ke} برابر یک باشد بدین معناست که برتری مرجع k بر مرجع e در هر دو حالت موافق و مخالف قابل قبول است (یعنی برتری آن از حد آستانه موافقت بیشتر بوده و مخالفت و یا ضعف آن نیز از حد آستانه مخالفت کمتر است) ولیکن هنوز مرجع k شانس مسلط شدن توسط مراجع دیگر را دارد. مرجع ای باید انتخاب شود که بیشتر از آن مغلوب شود، تسلط داشته باشد و از این نظر می‌توان مراجع را رتبه‌بندی کرد.

۳- نتایج شبیه‌سازی

در این بخش روش پیشنهادی از جنبه‌های مختلف مورد بررسی قرار گرفته و با سایر روش‌ها مقایسه شده است. جهت بررسی کارایی روش پیشنهادی در این مقاله از دو پایگاه داده [۲۰] CONLL-2012 و [۲۱] English Gigaword استفاده شده است.

در جدول ۱ روش پیشنهادی طبق معیارهای MUC [۲۲]، B3 [۲۹]، CEAF_{φ4} [۲۴]، مقادیر دقت و بازخوانی و Avg.F1 (میانگین خروجی سه معیار MUC، B3 و CEAF_{φ4}) با روش‌های مطرح و جدید تحلیل هم‌ارجاعی مقایسه شده است. همان‌طور که جدول ۱ مشاهده می‌شود، روش پیشنهادی طبق همه معیارها بر روش‌های قبلی برتری دارد و مقدار F1 به میزان ۶/۳ درصد بهبود یافته است.

جدول ۱: نتایج ارزیابی روش پیشنهادی-پیکره CoNLL-2012

برای محاسبه شاخص توافق (C_{ke}) باید مرجع k و مرجع e مقایسه شده و مقدار آن از جمع وزن معیارهایی که مرجع k نسبت به مرجع e ترجیح دارد، به دست می‌آید. به زبان ریاضی، شاخص توافق از رابطه زیر محاسبه می‌شود.

$$C_{ke} = \sum_{j \in S_{ke}} w'_j \quad (13)$$

شاخص توافق، بیانگر میزان برتری مرجع k بر مرجع e بوده و مقدار آن از صفر تا یک تغییر می‌کند

گام ۵- تعیین ماتریس مخالف

ماتریس مخالف یک ماتریس مربعی می‌باشد که بعد آن تعداد مراجع می‌باشد. هر یک از درایه‌های این ماتریس، شاخص عدم توافق (مخالفت) بین دو مرجع نامیده می‌شود. مقدار این شاخص از رابطه زیر بدست می‌آید.

$$d_{ke} = \frac{\max_{j \in D_{ke}} |v_{kj} - v_{ej}|}{\max_{j \in J} |v_{kj} - v_{ej}|} \quad (14)$$

گام ۶- تشکیل ماتریس تسلط موافق

در مرحله چهارم نحوه محاسبه شاخص توافق (C_{ke}) بیان شد، هم‌اکنون در این مرحله یک مقدار معین برای شاخص توافق مشخص می‌شود که آن را آستانه موافقت می‌نامند و با C^- نشان داده می‌شود. آستانه موافقت از میانگین‌گیری شاخص‌های توافق (درایه‌های ماتریس توافق) به دست می‌آید. به زبان ریاضی مقدار آستانه موافقت از رابطه زیر محاسبه می‌شود.

$$\bar{C} = \sum_{k=1}^m \sum_{e=1}^m \frac{C_{ke}}{m(m-1)} \quad (15)$$

ماتریس تسلط موافق (F) با توجه به مقدار آستانه موافقت تشکیل می‌شود. اگر C_{ke} بزرگتر از C^- باشد، برتری مرجع k بر مرجع e قابل قبول است در غیر اینصورت مرجع k بر مرجع e برتری ندارد لذا درایه‌های ماتریس تسلط موافق از رابطه زیر تعیین می‌شود.

$$f_{ke} = \begin{cases} 1 & C_{ke} \geq \bar{C} \\ 0 & C_{ke} < \bar{C} \end{cases} \quad (16)$$

گام ۷- تشکیل ماتریس تسلط مخالف

ماتریس تسلط مخالف (G) مانند ماتریس تسلط موافق تشکیل می‌شود. بدین منظور ابتدا باید مقدار آستانه مخالفت (d^-) از میانگین‌گیری شاخص‌های مخالفت (درایه‌های ماتریس مخالف) محاسبه شود. به زبان ریاضی مقدار آستانه مخالفت از رابطه زیر محاسبه می‌شود.

$$\bar{d} = \sum_{k=1}^m \sum_{e=1}^m \frac{d_{ke}}{m(m-1)} \quad (17)$$

همان‌گونه که در مرحله پنجم بیان شد، مقدار شاخص مخالفت (d_{ke}) هر چه کمتر باشد بهتر است. زیرا میزان مخالفت (عدم توافق) برتری مرجع k بر مرجع e را بیان می‌کند. چنانچه d_{ke} از d^- بزرگتر باشد میزان مخالفت زیاد بوده و نمی‌توان از آن صرف‌نظر کرد بنابراین درایه‌های ماتریس تسلط مخالف (G) به‌صورت زیر محاسبه می‌شود.

$$g_{ke} = \begin{cases} 1 & d_{ke} > \bar{d} \\ 0 & d_{ke} \leq \bar{d} \end{cases} \quad (18)$$

داده و توانسته است ویژگی‌های بهتری نسبت به این روش ارائه دهد و در اکثر زمینه‌های پردازش متن نتایج قابل قبولی ارائه کند.

جدول ۲: نتایج ارزیابی تعبیه واژگان روی پیکره توسعه CoNLL-2012

	Avg. F1	Δ
مدل پیشنهادی (تعبیه واژه BERT)	۸۰/۲	
تعبیه واژه Word2Vec	۷۳/۴	-۶/۸
تعبیه واژه GloVe	۷۶/۲	-۴
تعبیه واژه ELMo	۷۸/۱	-۲/۱

در شکل ۱ افزایش میزان F1 تحت مقادیر دقت و بازخوانی جهت سنجش تشخیص موجودیت‌های نامدار روی پایگاه داده [۲۱] English Gigaword به تصویر کشیده شده است. مقادیر دقت و بازخوانی طبق روابط (۲۰) و (۲۱) محاسبه می‌شود. با توجه به شکل ۱ مشخص است، که با اضافه کردن بردار تعبیه‌ی واژگان، برچسب اجزای سخن و وزن‌دهی به ویژگی‌ها، دقت مقدار F1 افزایش یافته است. با بررسی مقادیر دقت و بازخوانی به این نتیجه رسیدیم که اگرچه با اضافه شدن یک لایه LSTM به مدل پیشین، دقت تغییری نکرد ولی فراخوانی افزایش چشمگیری داشت، بنابراین باز هم مقدار F1 افزایش یافته است.

$$(20) \quad \text{تعداد موجودیت‌های درست استخراج شده} \\ \text{تعداد کل موجودیت‌های استخراج شده} = \text{دقت}$$

بازخوانی

$$(21) \quad \frac{\text{تعداد موجودیت‌های درست استخراج شده}}{\text{تعداد کل موجودیت‌های درست}}$$

سایز حالت پنهان^۹ در شبکه‌های LSTM، ۳۰۰ در نظر گرفته شده است. مقادیر دیگری نیز مورد آزمایش قرار گرفت ولی این مقدار بهترین نتیجه را حاصل کرد. بنابراین، خروجی حاصل از شبکه‌ی BiLSTM برداری با ۶۰۰ بعد است که از اتصال خروجی شبکه‌ی روبه‌جلو و شبکه‌ی رو به عقب بدست می‌آید. یک لایه Dropout برای جلوگیری از بیش‌برازش^{۱۰} پس از لایه‌ی BiLSTM قرار دادیم. با این عمل برخی از ورودی‌ها به لایه‌ی بعد صفر می‌شوند. در اینجا Dropout، ۰٫۵، در نظر گرفته شده است، بدین معنا که در هر مرحله نیمی از ورودی‌ها در نظر گرفته نمی‌شوند. با یک بررسی جزئی‌تر می‌توان به این نتیجه رسید، که هرچه میزان اطلاعاتی که از ظاهر، معنا و ساختار واژگان در قالب بردارهای بازنمایش به شبکه وارد می‌شود، افزایش یابد، مقدار دقت نیز افزایش پیدا می‌کند و به تبع آن F1 نیز بیشتر می‌گردد. افزایش میزان اطلاعات از محتوای متنی که واژگان در آن قرار گرفته است با استفاده از BiLSTM و یک لایه LSTM بر روی خروجی آن، موجب افزایش فراخوانی می‌گردد و به تبع آن F1 نیز افزایش می‌یابد.

روش	MUC			B3			CEAF_φ4			Avg.F1
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	
کلارک و مینگ [۱۰]	۷۹/۲	۷۰/۴	۷۴/۶	۶۹/۹	۵۸/۰	۶۳/۴	۶۳/۵	۵۵/۵	۵۹/۲	۶۵/۷
سایرین [۲]	۸۱/۴	۷۹/۵	۸۰/۴	۷۲/۲	۶۹/۵	۷۰/۸	۶۸/۲	۶۷/۱	۶۷/۶	۷۳/۰
سایرین [۴]	۸۵/۴	۷۷/۹	۸۱/۴	۷۷/۹	۶۶/۴	۷۱/۷	۷۰/۶	۶۶/۳	۶۸/۴	۷۳/۸
پیشنهادی	۸۳/۸	۸۸/۵	۸۶/۰	۷۵/۴	۸۱/۴	۷۸/۳	۷۳/۴	۷۸/۸	۷۶/۰	۸۰/۱

دلیل این برتری در نظر گرفتن اطلاعات موجودیت‌ها در شبکه از پیش آموزش دیده BERT است. با در نظر گرفتن تعبیه واژگان توسط این روش اطلاعات معنایی بهتری در اختیار سیستم قرار گرفته و سیستم توانایی بالاتری در تشخیص درست نامبری‌ها پیدا می‌کند. همچنین توسط ساختار شبکه عصبی عمیق ارائه شده ویژگی‌های مفیدتری توسط شبکه برای نامبری‌ها استخراج شده و باعث می‌شود نامبری‌هایی که در روش‌های دیگر نادیده گرفته شده‌اند، در روش پیشنهادی در نظر گرفته شوند. همچنین توسط روند وزن‌دهی دقیق به ویژگی‌ها و در نظر گرفتن میزان اهمیت آن‌ها در مرحله تصمیم‌گیری و تولید نامبری‌های هم‌مرجع روش پیشنهادی توانسته است به دقت بالاتری نسبت به سایر روش‌ها دست یابد.

برای نمایش تأثیر نحوه استخراج تعبیه واژگان بر روی عملکرد روش پیشنهادی و مقدار Avg.F1، روش‌های مختلف تعبیه واژگان نظیر [۲۵] Word2Vec، [۱۷] GloVe، [۱۸] ELMo و BERT در جدول ۲ روی مجموعه توسعه پیکره CoNLL-2012 با یکدیگر مقایسه شده‌اند. همان‌طور که مشخص است کارکرد روش GloVe از روش Word2Vec اندکی بهتر است؛ زیرا در روش GloVe بردار اولیه بر اساس طول همجواری^۸ با شمارش تعداد ظاهر شدن دو کلمه کنار هم در این بازه، تشکیل می‌شود و در مرحله بعد، بر اساس نزدیکی یا دوری دو کلمه به یکدیگر، این اعداد، وزن دهی می‌شوند. روش ELMo از روش GloVe بهتر عمل می‌کند؛ زیرا در ساختار خود از مدل‌های زبانی دوجهته و شبکه BiLSTM استفاده کرده و ویژگی‌های نحوی و معنایی و زبان‌شناسی بهتری در اختیار سیستم قرار می‌دهد. در نهایت روش BERT بهتر از تمام روش‌های تعبیه واژگان عمل می‌کند زیرا این روش معماری روش ELMo را بهبود

⁹ Hidden State Size

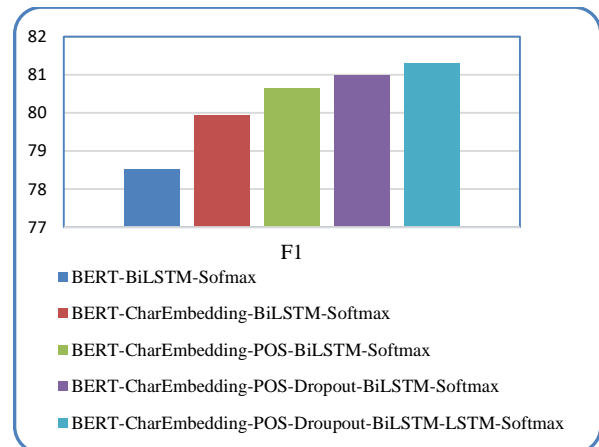
¹⁰ Over Fit

⁸ Window Size

روش‌های جدید تحلیل هم‌ارجاعی صورت گرفت. همچنین نشان داده شد که روش پیشنهادی با کمترین میزان خطا، مسئله تحلیل هم‌ارجاعی را به‌خوبی مدیریت می‌کند. سیستم‌های قبلی دقت و بازخوانی پایین‌تری نسبت به سیستم پیشنهادی داشتند و روی بعضی از مجموعه دادگان برای یک زبان خاص، عملکرد بهتری نشان می‌دادند. همچنین به دلیل استفاده از روش BERT در نظر گرفتن مدل‌های زبانی، نرخ تشخیص نامبری‌های درست در روش پیشنهادی نسبت به سایر روش‌ها بالاتر شده است. سیستم پیشنهادی عملکردی مناسب در انواع مختلف دادگان دارد و معیار Avg.F1 در مجموعه دادگان CoNLL-2012 ۶/۳ درصد افزایش یافته است. در آینده قصد داریم، مدل ارائه شده را با افزودن ویژگی و تکنیک‌های تصمیم‌گیری چندمعیاره فازی، جهت افزایش دقت تشخیص زنجیره‌های هم‌مرجع بهبود دهیم.

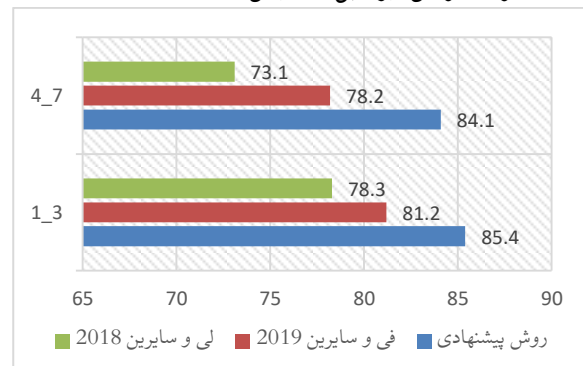
مراجع

- [1] T. S. Morton, "Using coreference for question answering," *Proceedings of the Workshop on Coreference and its Applications - CorefApp '99*, p. 85, 1999.
- [2] K. Lee, L. He, and L. Zettlemoyer, "Higher-order coreference resolution with coarse-to-fine inference," *arXiv preprint arXiv:1804.05392*, 2018.
- [3] B. Kantor and A. Globerson, "Coreference Resolution with Entity Equalization," in *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 2019, pp. 673–677.
- [4] H. Fei, X. Li, D. Li, and P. Li, "End-to-end Deep Reinforcement Learning Based Coreference Resolution," in *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 2019, pp. 660–665.
- [5] H. Peng, D. Khashabi, and D. Roth, "Solving hard coreference problems," *arXiv preprint arXiv:1907.05524*, 2019.
- [6] A. Haghighi and D. Klein, "Unsupervised coreference resolution in a nonparametric bayesian model," in *Proceedings of the 45th annual meeting of the association of computational linguistics*, 2007, pp. 848–855.
- [7] H. Lee, A. Chang, Y. Peirsman, N. Chambers, M. Surdeanu, and D. Jurafsky, "Deterministic coreference resolution based on entity-centric, precision-ranked rules," *Computational Linguistics*, vol. 39, no. 4, pp. 885–916, 2013.
- [8] A. Zeldes and S. Zhang, "When annotation schemes change rules help: A configurable approach to coreference resolution beyond ontonotes," in *Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2016)*, 2016, pp. 92–101.
- [9] K. Raghunathan et al., "A multi-pass sieve for coreference resolution," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 2010, pp. 492–501.
- [10] K. Clark and C. D. Manning, "Deep Reinforcement Learning for Mention-Ranking Coreference Models," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 2256–2262.
- [11] K. Clark and C. D. Manning, "Entity-centric coreference resolution with model stacking," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2015, vol. 1, pp. 1405–1415.
- [12] A. Rahman and V. Ng, "Supervised models for coreference resolution," in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2-Volume 2*, 2009, pp. 968–977.



شکل ۱: نمودار F1 در مدل‌های مختلف

در شکل ۲ دقت تشخیص نامبری روش پیشنهادی از میان دنباله کلمات و بر اساس تعداد کلمات موجود در دنباله با سایر روش مقایسه شده است. همان‌طور که مشاهده می‌شود، دقت روش پیشنهادی از سایر روش‌ها بالاتر است. در سایر روش‌ها دقت تشخیص نامبری با افزایش طول دنباله کلمات (۴ تا ۷ کلمه) کاهش چشمگیری دارد اما در روش پیشنهادی دقت تشخیص نامبری کاهش اندکی دارد. همچنین طبق ساختار ارائه شده نامبری‌های دیده نشده در مجموعه آموزش نیز قابل تشخیص هستند.



شکل ۲: نرخ تشخیص نامبری برحسب طول دنباله کلمات

۴- نتیجه‌گیری

در این مقاله، یک روش تحلیل هم‌ارجاعی بر پایه تصمیم‌گیری چندشاخصه و روش وزن‌دهی بردار ویژه روی متن ارائه شد. استخراج نامبری‌ها توسط ساختار شبکه عصبی عمیق باعث شد که روش پیشنهادی قدرت بالاتری نسبت به سایر روش‌ها در تشخیص نامبری‌ها داشته باشد. در این روش، انتخاب مرجع مناسب برای نامبری موردنظر بر اساس تمام معیارهای مؤثر در تحلیل هم‌ارجاعی انجام شد. با استفاده از ساختار تصمیم‌گیری چندشاخصه، دخیل کردن تمام معیارها در تصمیم‌گیری و در نظر گرفتن اطلاعات موجودیت‌ها، روش ارائه شده در مقایسه با روش‌های پیشین دقیق‌تر شده است. در این مقاله، مقایسه‌ی مناسبی بین روش پیشنهادی و

- [13] K. Clark and C. D. Manning, "Improving Coreference Resolution by Learning Entity-Level Distributed Representations," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2016, vol. 1, pp. 643–653.
- [14] K. Lee, L. He, M. Lewis, and L. Zettlemoyer, "End-to-end Neural Coreference Resolution," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 188–197.
- [15] S. Subramanian and D. Roth, "Improving Generalization in Coreference Resolution via Adversarial Training," in *Proceedings of the Eighth Joint Conference on Lexical and Computational Semantics (*SEM 2019)*, 2019, pp. 192–197.
- [16] M. Joshi, O. Levy, D. S. Weld, and L. Zettlemoyer, "BERT for Coreference Resolution: Baselines and Analysis," *arXiv preprint arXiv:1908.09091*, 2019.
- [17] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.
- [18] M. Peters *et al.*, "Deep Contextualized Word Representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 2018, vol. 1, pp. 2227–2237.
- [19] A. T. de Almeida, "Multicriteria decision model for outsourcing contracts selection based on utility function and ELECTRE method," *Computers & Operations Research*, vol. 34, no. 12, pp. 3569–3574, 2007.
- [20] S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, and Y. Zhang, "CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes," in *Joint Conference on EMNLP and CoNLL-Shared Task*, 2012, pp. 1–40.
- [21] "Annotated English Gigaword - Linguistic Data Consortium." [Online]. Available: <https://catalog.ldc.upenn.edu/LDC2012T21>. [Accessed: 12-Apr-2019].
- [22] M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman, "A model-theoretic coreference scoring scheme," in *Proceedings of the 6th conference on Message understanding*, 1995, pp. 45–52.
- [23] A. Bagga and B. Baldwin, "Algorithms for scoring coreference chains," in *The first international conference on language resources and evaluation workshop on linguistics coreference*, 1998, vol. 1, pp. 563–566.
- [24] X. Luo, "On coreference resolution performance metrics," in *Proceedings of the conference on human language technology and empirical methods in natural language processing*, 2005, pp. 25–32.
- [25] X. Rong, "word2vec parameter learning explained," *arXiv preprint arXiv:1411.2738*, 2014.



طراحی سیستم تشخیص نفوذ شبکه بر پایه ترکیب روش‌های انتخاب ویژگی

آرش ادیبی^۱، پیمان ستوده^۲

^۱ دانشجوی کارشناسی ارشد، مهندسی اتوماسیون و ابزار دقیق در صنعت نفت، دانشگاه شیراز
arash.adibi@ymail.com

^۲ استادیار، دانشکده مهندسی قدرت و کنترل، دانشگاه شیراز
psetoodeh@shirazu.ac.ir

چکیده

با توجه به رشد قابل ملاحظه استفاده از خدمات مبتنی بر شبکه در کاربردهای شخصی، سازمانی و صنعتی طی سال‌های اخیر و متناسب با آن، رشد دسترسی‌های غیرمجاز به این شبکه‌ها، لزوم حفظ امنیت اطلاعات و حضور سیستم‌های امنیت سایبری را نمایان نموده است. سیستم‌های تشخیص نفوذ به عنوان یکی از ابزارهای افزایش امنیت سایبری شبکه در دهه‌های اخیر مورد توجه قرار گرفته‌اند. در پژوهش حاضر با تمرکز بر روش‌های انتخاب ویژگی‌های مهم، به ارائه روشی جدید در طراحی این سیستم‌ها پرداخته ایم. در این شیوه، ویژگی‌های مهم مجموعه داده NSL-KDD، با ترکیب الگوریتم‌های Lasso، Gradient Boosting و Extra Tree انتخاب و سپس به کمک طبقه‌بند Random Forest، مدل ساخته می‌شود. برای پیاده‌سازی و ارزیابی میزان موفقیت روش پیشنهادی از زبان Python و در بستر Jupyter Notebook استفاده شده است. نتایج ارزیابی در مقایسه با سایر کارهای موفق نشان می‌دهد که دقت سیستم تشخیص نفوذ عملکرد بهتری داشته به طوری که دقت تشخیص ۹۹٫۷۹٪ برای حملات DoS، ۹۹٫۶۲٪ برای حملات Porbe، ۹۸٫۰۲٪ برای حملات R2L، و ۹۹٫۷۳٪ برای حملات U2R به دست می‌آید.

کلمات کلیدی

تشخیص نفوذ، انتخاب ویژگی، داده‌کاوی

۱- مقدمه

سال ۲۰۲۱ به میزان تقریبی ۵۱ درصد رشد خواهد داشت که در صورت عدم اختصاص راهکارهای نوین حفاظتی، امنیت اطلاعاتی کاربران و سیستم‌های تحت شبکه، به دلیل رشد دسترسی‌های غیرمجاز و حملات سایبری، به خطر خواهد افتاد [1]. علاوه بر تأمین امنیت کاربران در سطح شبکه‌هایی نظیر اینترنت که در دسترس عموم قرار دارد، مسائل رخ داده در منطقه خاورمیانه و حضور تأسیسات حساسی نظیر نیروگاه‌های اتمی و تلاش‌ها برای نفوذ به شبکه و آسیب‌رسانی به این تأسیسات در سال‌های اخیر، حضور سیستم‌های تشخیص نفوذ و رفتار غیرعادی را برای شبکه‌های دارای دسترسی خصوصی نیز حائز اهمیت می‌نماید [2]. در واقع، اجزای یک شبکه باید ارتباطی پایدار و ایمن داشته باشند و دسترسی به اطلاعات موجود در آن برای عوامل داخلی و خارجی غیرمجاز، میسر نباشد. در این راستا سیستم‌هایی با نام

با وجود پیشرفت‌های چشمگیر در به‌کارگیری سیستم‌های ایمن در بستر شبکه، سالانه گزارش‌هایی مبنی بر نفوذ غیرمجاز به شبکه‌ها منتشر می‌شود. بنابراین می‌توان نتیجه گرفت که سیستم‌ها همچنان به مکانیسم‌های بهبودیافته جهت مقابله با حملات پیشرفته نیاز خواهند داشت. به عنوان مثال، ساختارهای حفاظتی نوین در قالب سیستم‌های ناظر بر فعالیت‌های حاضر در شبکه و شناسایی رفتارهای ناهنجار در آن، می‌تواند به تکامل اقدامات حوزه امنیت اطلاعات در بستر شبکه کمک نماید. آمار رسمی بین‌المللی از رشد حضور کاربران سطح شبکه، میزان جرائم حوزه سایبری و افزایش بودجه کشورهای توسعه‌یافته جهت اقدامات حوزه امنیت سایبری، نشان از اهمیت بالایی این موضوع دارد. بر اساس گزارش شرکت آمریکایی سیسکو که در سال ۲۰۱۷ منتشر شده، ترافیک شبکه‌ها تا

متفاوتی از جمله تشخیص گفتار، پردازش زبان طبیعی، تحلیل بازار سهام و امنیت اطلاعات را دارا می‌باشد [13]. متناسب با حجم وسیع کاربردهای یادگیری ماشین، الگوریتم‌های متنوعی برای آن ارائه شده که با توجه به موضوع، شرایط و هدف مسئله به کار گرفته می‌شوند. الگوریتم‌های مختلف یادگیری ماشین، به‌طور کلی در چهار زیرگروه اصلی قرار می‌گیرند. یادگیری با ناظر^۶، نیمه نظارتی^۷، بدون ناظر^۸ و تقویتی^۹، چهار گروه مطرح شده می‌باشند و هر کدام روش‌های متنوعی جهت حل مسائل را شامل می‌شود [14].

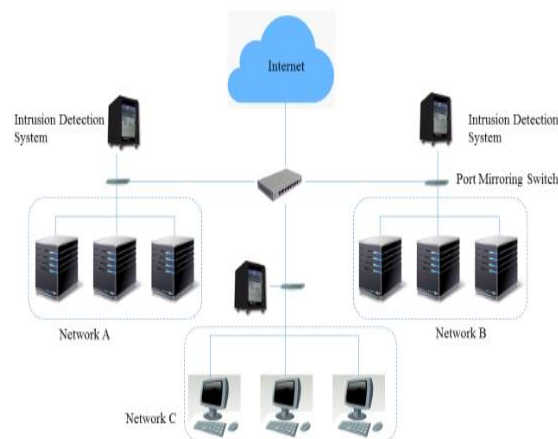
در این پژوهش قصد داریم تا با تمرکز بر مسئله امنیت سایبری شبکه‌ها و بررسی کاربرد روش‌های یادگیری ماشین در طراحی سیستم‌های تشخیص نفوذ مبتنی بر شناسایی رفتار غیرعادی و مقایسه آن با سیستم‌های فعلی، مسئله اعلام هشدارهای مبنی بر رفتار ناهنجار در شبکه را که می‌تواند به‌عنوان نفوذ شناخته شوند، تحت بررسی قرار دهیم.

در سال ۱۹۹۹، مجموعه داده KDD CUP به جهت استفاده در ایجاد سیستم‌های تشخیص نفوذ، طی نظارت بر شبکه نیروی هوایی آمریکا، تولید گردید. حجم مجموعه داده مذکور در حدود ۴ گیگابایت است و تقریباً پنج میلیون رکورد (۴,۸۹۸,۴۳۱ رکورد) را در بر می‌گیرد. هر رکورد شامل ۴۱ ویژگی و برجسب نرمال و حمله می‌باشد. در تحقیقاتی که طی سال ۲۰۰۵ صورت گرفت، تحلیلی بر این مجموعه داده انجام شد که سبب معرفی مجموعه داده بهینه‌شده NSL-KDD گردید [15]. مهم‌ترین خصوصیت مجموعه داده بهینه‌شده، حذف رکوردهای تکراری نسبت به مجموعه داده اصلی می‌باشد. تحقیق مذکور نشان داد که ۷۵ درصد داده‌های آموزش و ۷۸ درصد داده‌های تست، در مجموعه داده KDD تکراری می‌باشند. تکراری بودن داده‌ها، سبب می‌شود که در مرحله آموزش، الگوریتم یادگیری، به داده‌هایی که تکرار آن‌ها بیشتر است، بایاس گردد. اهمیت این موضوع در این است که ممکن است حملات خطرناک اما با تعداد کم وجود داشته باشد که در مرحله یادگیری نادیده گرفته شوند [15].

ما در این پژوهش از مجموعه داده بهینه‌شده NSL-KDD استفاده خواهیم نمود. به طور کلی مزایای مجموعه داده NSL-KDD نسبت به KDD-CUP به شرح زیر می‌باشد:

- رکوردهای تکراری حذف شده‌اند. لذا در فرایند آموزش، دسته‌بند، به سمت رکوردهای با تعداد بیشتر بایاس نخواهد شد.
- با توجه به حذف رکوردهای تکراری، نرخ تشخیص نهایی بهتر خواهد بود، زیرا حملات با تعداد رکورد کمتر نیز در فرایند آموزش لحاظ شده‌اند.
- با توجه به اینکه تعداد داده‌های رکورد به نرخ قابل قبولی رسیده‌اند، نیازی به انتخاب تصادفی از بین آن‌ها نیست و

تشخیص نفوذ (IDS^۱) مورد استفاده قرار می‌گیرند. IDS ها با پایش ترافیک شبکه، در تشخیص حملات و اعلام هشدار، نقش مؤثری ایفا کرده و به عنوان یک لایه حفاظتی شناخته می‌شوند [4]. طراحی سیستم‌های اندازه‌گیری و پایش ترافیک شبکه به منظور شناسایی و درک رفتار حملات سایبری و استفاده از دانش به‌دست‌آمده از این طریق، در تأمین امنیت شبکه‌ها، تبدیل به یک موضوع جذاب در جوامع علمی شده است [5] [6]. شکل ۱ معماری یک شبکه با حضور IDS و نحوه ارتباط آن با شبکه را به‌صورت نمادین نمایش داده است [7].



شکل ۳۵: معماری یک شبکه با حضور IDS [7]

IDS ها بر اساس دو اصل کلی طراحی شده‌اند. حالت اول، بر اساس اطلاعاتی که از نفوذهای تحلیل شده در گذشته به‌دست‌آمده و در طراحی IDS، استفاده شده است. در این حالت IDS ها به شناسایی نفوذهای غیرمجاز از طریق الگوهای شناخته‌شده می‌پردازند. دسته دوم، بر اساس پایشی که از شبکه موردنظر انجام می‌دهد، مدلی از وضعیت عادی آن استخراج نموده و بر اساس بروز اختلاف بین مدل به‌دست‌آمده و شرایط لحظه‌ای شبکه، می‌تواند هشدار وقوع رفتار غیرعادی را صادر کند [8]. از آنجاکه با توسعه علوم و فناوریهای مرتبط، روش‌های جدید نفوذ هم توسعه پیدا خواهند کرد، لذا دسته دوم IDS ها می‌توانند در مقابل روش‌های جدید نیز عملکرد مطلوبی داشته باشند [9] [10].

شرکت آمریکایی آی بی ام^۲، در جدیدترین گزارش خود، به‌کارگیری الگوریتم‌های هوش مصنوعی^۳ و یادگیری ماشین^۴ را به‌عنوان روش‌های نوین و کارآمد، جهت ارتقای امنیت سایبری سیستم‌ها معرفی کرده است [11]. به‌طور مشابه، شرکت پژوهشی آمریکایی گارتنر^۵، روش یادگیری ماشین را طی ۲ تا ۵ سال آینده، جزو پرکاربردترین روش‌های برتر حل مسئله در علوم مختلف عنوان کرده است [12]. یادگیری ماشین، قابلیت حل مسائل در حوزه‌های

¹ Intrusion Detection System

² IBM

³ Artificial Intelligence

⁴ Machine Learning

⁵ Gartner

⁶ Supervised Learning

⁷ Semi-supervised Learning

⁸ Unsupervised Learning

⁹ Reinforcement Learning

می‌توان بدون تولید مجموعه داده سبکتر فرایند طراحی سیستم تشخیص نفوذ را انجام داد [15].
مجموعه داده NSL-KDD از نظر ابعاد به صورت جدول ۱ می‌باشد.

جدول ۱- دسته‌بندی روش‌های انتخاب ویژگی [15]

داده‌ها	مجموعه داده آزمایش	مجموعه داده تست
حالت نرمال	۶۷۳۴۳	۹۷۱۱
حالت حمله	۵۸۶۳۰	۱۲۸۳۳
جمع	۱۲۵۹۷۳	۲۲۵۴۴

چالشی که با آن روبرو هستیم، عدم توازن یکسان حملات در مجموعه داده‌ها می‌باشد. به این معنی که تعداد حملات نوع DoS، بیشترین میزان را در مجموعه داده دارا می‌باشند، درحالی‌که سایر حملات نسبت به حملات این گروه، از پراکندگی کمتری برخوردار می‌باشند. این موضوع در فرایند یادگیری سیستم‌های تشخیص نفوذ، سبب بایاس الگوریتم یادگیری به سمت حملات با تعداد بیشتر خواهد شد. به جهت مقابله با این مشکل، کاربرد الگوریتم‌های انتخاب ویژگی اهمیت پیدا می‌کند. انتخاب ویژگی را می‌توان به عنوان فرآیند شناسایی ویژگی‌های مرتبط و حذف ویژگی‌های غیر مرتبط و تکراری، با هدف مشاهده زیرمجموعه‌ای از ویژگی‌ها که مسئله را به خوبی و با حداقل کاهش دقت در عملکرد کلی سیستم تشریح می‌کند، تعریف کرد.
مزایای انتخاب ویژگی در فرایند حل مسئله به کمک یادگیری ماشین عبارت‌اند از [17]:

- بهبود کارایی الگوریتم‌های یادگیری ماشین.
- کاهش داده کلی، محدود کردن نیازمندی‌های ذخیره‌سازی و احتمالاً کمک به کاهش هزینه‌ها.
- سادگی و قابلیت استفاده از مدل‌های ساده‌تر و کسب سرعت.

۲- پژوهش‌های مرتبط

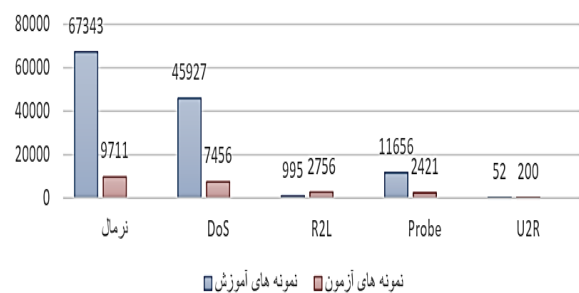
مقاله [18] با استفاده از روش انتخاب ویژگی بر اساس همبستگی^{۱۴}، اقدام به شناسایی ویژگی‌های مهم مجموعه داده NSL-KDD نموده است. سپس با استفاده الگوریتم درخت تصمیم J48، مرحله یادگیری مدل را انجام داده و سیستم نهایی تشخیص نفوذ را ارائه کرده است. نویسندگان به جهت بررسی اثر نوع طبقه بند، در دقت تشخیص، از دو طبقه بند دیگر با نام‌های ماشین بردار پشتیبان (SVM) و Naïve Bayes نیز استفاده نموده‌اند.

مقاله [19] نویسندگان با استفاده از روش بهینه‌شده الگوریتم ژنتیک^{۱۵}، به انتخاب ویژگی‌های مهم مجموعه داده NSL-KDD پرداخته‌اند. اهمیت استفاده از این روش جهت انتخاب ویژگی را در عملکرد بهتر طبقه بند در مرحله تشخیص حملات معرفی نموده‌اند. همچنین روش استفاده‌شده در مرحله یادگیری و ساخت مدل، SVM می‌باشد.

مقاله [20] روشی با نام جستجوی متوالی را به جهت انتخاب ویژگی‌های مهم مجموعه داده NSL-KDD معرفی نموده است. سیستم پیشنهادی تشخیص نفوذ، ابتدا با کلیه ویژگی‌ها طراحی شده و طبقه بند Naïve Bayes نیز مدل یادگیری ماشین را تولید می‌نماید. در مرحله بعد، الگوریتم انتخاب ویژگی جست‌وجوی متوالی، یک ویژگی را حذف کرده و سپس مجموعه جدید ویژگی‌ها

اکثر حملات تحت شبکه، به چهار گروه اصلی^{۱۰} DoS، R2L^{۱۱}، U2L^{۱۲} و PROBE^{۱۳} دسته‌بندی می‌شوند [16]. در R2L حملات سرویس‌هایی که کاربران قصد استفاده از آن‌ها را داشته باشند و یا در حال استفاده از آن باشند، توسط حملات DoS از دسترس خارج می‌شوند. در این نوع حملات، مهاجم با استفاده از نقاط آسیب‌پذیر سیستم، کنترل سیستم را به عنوان یک کاربر محلی به کمک شبکه، در اختیار می‌گیرد. در اغلب حملات R2L، مهاجم از طریق اینترنت به سیستم رایانه‌ای ورود پیدا می‌کند. به دست آوردن رمز عبور سیستم، مثالی از این نوع حمله می‌باشد. برخلاف حملات R2L که نفوذگ، از بیرون به سیستم نفوذ می‌کند، در حملات نوع U2R، نفوذگر یک حساب کاربری در سیستم هدف دارد. در این حالت، نفوذگر تلاش می‌کند تا امتیازات حساب کاربری خود را افزایش دهد تا بتواند کنترل سیستم را به دست بگیرد. حملات Probe را می‌توان با عنوان تجسس و پیمایش، جهت شناسایی نقاط آسیب‌پذیر معرفی نمود. در این حملات، نفوذگر تلاش می‌کند تا اطلاعاتی را در مورد سرویس‌ها و سیستم‌های شبکه استخراج نماید و به کمک اطلاعات به دست آمده، نقشه و طرحی از شبکه برای استفاده در حملات بعدی تولید نماید. در چهار گروه‌بندی مربوط به حملات، تعداد رکوردها در مجموعه داده NSL-KDD به صورت جدول ۲ می‌باشد.

جدول ۲- گروه‌بندی مجموعه داده NSL-KDD و نمایش ابعاد

¹⁰Denial-of-Service¹¹Root to Local attacks¹²User to Root attack¹³Probing attacks¹⁴ Correlation based Feature Selection method¹⁵ Optimal Genetic Algorithm

Author	Method used	Classifier used	Accuracy for Attack Classes			
			DOS	Prob e	R2L	U2R
(Dhanabal & Shantharajah 2015)	Correlation based Feature Selection method	J48	99.1 %	98.9 %	97.9 %	98.7 %
(Senthilnayaki et al. 2015)	Optimal Genetic Algorithm	SVM	99.15 %	99.08 %	96.50 %	97.03 %
(Zhang & Wang 2013)	Sequential search	Naïve Bayes	99.3 %	97.4 %	95.0 %	59.6 %
(Alazab et al. 2012)	Information gain	J48	99.7 %	97.8 %	91.3 %	97.2 %
(Mukherjee & Sharma 2012) [31]	Feature vitality based Method	Naïve Bayes	98.7%	98.8 %	96.1 %	64%
(Parsazad et al. 2012)	Correlation Coefficient	K-nearest neighbor	98.34 %	98.38 %	97.03 %	83.33 %
(Parsazad et al. 2012)	Fast feature Reduction	K-nearest neighbor	98.28 %	98.50 %	97.79 %	82.00 %
(Parsazad et al. 2012)	Least Square Regression Error	K-nearest neighbor	98.34 %	98.98 %	97.62 %	82.61 %
(Parsazad et al. 2012)	Maximal Information Compression Index	K-nearest neighbor	98.03 %	98.92 %	98.05 %	90.70 %

نتایج حاصل از مقایسه پژوهش‌های پیشین نشان می‌دهد که هر الگوریتم انتخاب ویژگی، در شناخت نوع خاصی از حملات موفق‌تر عمل می‌نماید. بر همین اساس در ادامه به معرفی روشی ترکیبی از روش‌های انتخاب ویژگی می‌پردازیم و سیستم تشخیص نفوذ بهینه‌تری را معرفی می‌کنیم.

۳- روش پیشنهادی

در سیستم پیشنهادی، از روش انتخاب ویژگی تعبیه شده^{۲۳} استفاده نموده ایم. این روش، انتخاب ویژگی را در فرآیند آموزش انجام می‌دهد. در این روش‌ها، جست‌وجو برای یک زیرمجموعه بهینه از ویژگی‌ها در مرحله ساخت دسته‌بند انجام می‌شود و می‌توان آن را به عنوان جست‌وجویی در فضای ترکیبی از زیرمجموعه‌ها و فرضیه‌ها دید. این روش‌ها قادر به ثبت وابستگی‌ها با هزینه‌های محاسباتی کمتری در مقایسه با بسته‌بندها می‌باشند. شکل ۲ نمایشی از عملکرد این روش را در یک نمای کلی از مسئله یادگیری ماشین نمایش می‌دهد.

را به طبقه بند داده و سیستم جدیدی تولید می‌شود. در هر مرحله، میزان دقت سیستم جدید با سیستم قبلی مقایسه می‌شود. اگر در اثر حذف هر ویژگی، دقت سیستم جدید، کاهش پیدا کند، بنابراین مشخص می‌شود که یک ویژگی مهم حذف گردیده است. با تکرار متوالی این روند، مجموعه‌ای از ویژگی‌های مهم شناسایی خواهند شد.

مقاله [21] با استفاده از روشی با نام بهره اطلاعاتی^{۱۶} و استفاده از طبقه بند درخت تصمیم J48، سیستم تشخیص نفوذی را معرفی کرده است. مجموعه داده مورداستفاده در این پژوهش -NSL-KDD بوده و نویسندگان استفاده از این مجموعه داده به جای استفاده از مجموعه داده KDD-CUP را با توجه به حذف رکوردهای تکراری، دارای اثر مثبت در افزایش نرخ دقت سیستم معرفی نموده‌اند. روش تشخیص ویژگی‌های مهم در روش معرفی شده در این پژوهش بر اساس دو پارامتر مدت زمان آموزش و نرخ تشخیص سیستم تشخیص نفوذ می‌باشد.

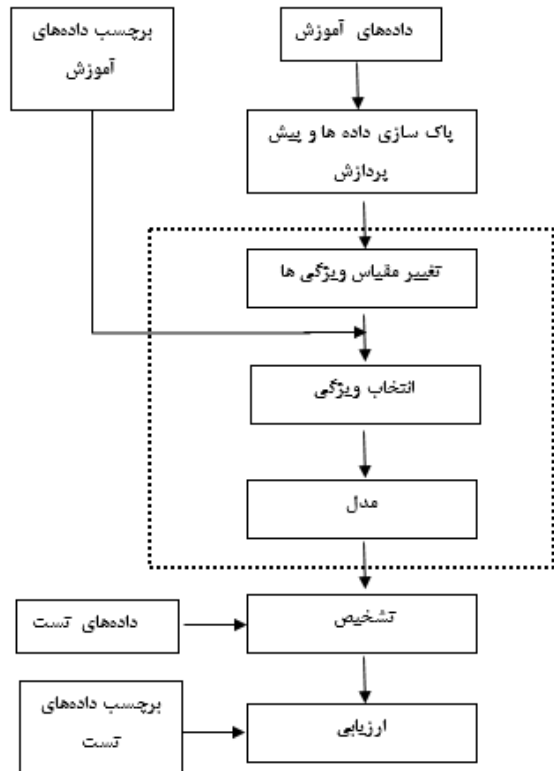
مقاله [22] با معرفی روشی تحت عنوان انتخاب ویژگی مبتنی بر ویژگی‌های حیاتی^{۱۷} و استفاده از طبقه بند Naïve Bayes سیستم تشخیص نفوذی را معرفی کرده‌اند. نویسندگان اهمیت استفاده از روش انتخاب ویژگی را کاهش ابعاد داده‌ها و همچنین استفاده از ویژگی‌های مهم داشته‌اند و بیان نموده‌اند که این عامل در افزایش سرعت عملکرد سیستم پیشنهادی، کاهش نیاز به فضای حافظه و همچنین افزایش دقت سیستم نهایی اثرگذار می‌باشد.

مقاله [23] با هدف معرفی روش سریع انتخاب ویژگی، چهار روش انتخاب ویژگی ضریب همبستگی^{۱۸}، کاهش سریع ویژگی‌ها^{۱۹}، حداقل خطای مربعات رگرسیون^{۲۰} و شاخص فشرده‌سازی حداکثر اطلاعات^{۲۱} را معرفی نموده است. در سیستم تشخیص نفوذ پیشنهادی، از روش نزدیک‌ترین همسایگی^{۲۲} در مرحله یادگیری و ساخت مدل استفاده شده است. انتخاب ویژگی با روش ضریب همبستگی، در واقع میزان همبستگی بین دوبه‌دوی ویژگی‌ها را بررسی کرده و ویژگی‌های دارای مشابهت بیشتر را در یک دسته قرار می‌دهد. روش حداقل مربعات رگرسیون خطا در واقع از مقایسه واریانس باقی‌مانده استفاده می‌کند. نویسندگان پس از بررسی سه روش فوق، روشی با نام کاهش سریع ویژگی‌ها را ارائه نموده‌اند. در این روش به جست‌جوی ویژگی‌های ضعیف‌تر پرداخته شده و با حذف آن‌ها، مجموعه‌ای شامل ویژگی برتر باقی می‌ماند. در جدول ۳، تعدادی از سیستم‌های معرفی شده در پژوهش‌های مرتبط را با توجه به میزان دقت سیستم پیشنهادی در شناسایی، معرفی نموده ایم.

جدول ۳- مقایسه پژوهش‌های پیشین

¹⁶ Information gain
¹⁷ Feature vitality based Method
¹⁸ Correlation Coefficient
¹⁹ Fast feature Reduction
²⁰ Least Square Regression Error
²¹ Maximal Information Compression Index
²² K-Nearest Neighbor

²³ Ensemble Method



شکل ۳- نمودار درختی فرایند طراحی سیستم تشخیص نفوذ پیشنهادی

پیش از ارزیابی عملکرد سیستم تشخیص نفوذ، نیاز داریم که ماتریس درهم ریختگی را تشکیل دهیم. به کمک این ماتریس می توان شاخص های ارزیابی سیستم یادگیری ماشین طراحی شده را محاسبه نمود. جدول ۵، محتوای ماتریس درهم ریختگی را به صورت نمادین نمایش می دهد.

		Predicted Class	
		Normal	Attack
Actual Class	Normal	True Negative (TN)	False Positive (FP)
	Attack	False Negative (FN)	True Positive (TP)

جدول ۵- ساختار ماتریس درهم ریختگی [25]

ماتریس درهم ریختگی به صورت شکل ۵ برای سیستم پیشنهادی بدست آمده است. هر ستون از این ماتریس، نمونه ای از مقدار پیش بینی شده را نمایش می دهد. در صورتی که هر سطر نمونه ای واقعی (درست) را در بر دارد.

الگوریتمی که روش های مبتنی بر مدل بر اساس آن عمل نموده و ویژگی های مهم را برای ما نمایان می کنند به صورت شکل ۲ می باشد.

General procedure for ensemble method.

Let D denote the original training data, k denote the number of base classifiers, and T be the test data.

```

for  $i = 1$  to  $k$  do
    Create training set  $D_i$  from  $D$ .
    Build a base classifier  $C_i$  from  $D$ .
end for
for each test record  $x \in T$  do
     $C^*(x) = \text{Vote}(C_1(x), C_2(x), \dots, C_k(x))$ 
end for
    
```

شکل ۲- استخراج ویژگی بر اساس روش های مبتنی بر مدل جایابی شده [24]

در مرحله انتخاب ویژگی برای هر گروه از حملات، از روش متفاوتی استفاده کرده ایم. در واقع، ابتدا با هر یک از چهار روش انتخاب ویژگی Lasso، Gradient Boosting، Extra tree، Random Forest، یک سیستم منفرد تشخیص نفوذ ایجاد کرده و سپس روشی که در تشخیص هر گروه از حملات نرخ دقت بیشتری کسب کرده را به عنوان روش برتر انتخاب کرده ایم. جدول ۴ نتیجه مقایسه را که شامل معرفی بهترین الگوریتم انتخاب ویژگی برای هر گروه از حملات می باشد، معرفی می کند. برای مرحله یادگیری نیز از روش Random Forest، استفاده نموده ایم.

جدول ۴- معرفی روش پیشنهادی جهت انتخاب ویژگی و دسته بند

	Feature selection method (select from model)	Classifier
DoS	Extra tree	Random forest
Probe	Lasso	Random forest
R2L	Gradient boosting	Random forest
U2R	lasso	Random forest

شمای کلی طراحی سیستم تشخیص نفوذ در شکل ۳ به صورت نمودار درختی ارائه شده است که شامل مرحله انتخاب ویژگی به عنوان پیش پردازش می باشد.

جدول ۶- ارزیابی عملکرد سیستم تشخیص نفوذ پیشنهادی نهایی

Feature selection method	Learning method: Random Forest Classifier			
	DoS	Probe	R2L	U2R
Extra tree				
Lasso				
Gradient boosting				
Lasso				
Accuracy	0.99796	0.99629	0.98023	0.99734
Precision	0.99893	0.99567	0.97479	0.96499
Recall	0.99638	0.99272	0.96903	0.84141
F-measure	0.99765	0.99418	0.97182	0.88730

۴- نتیجه گیری

سیستم پیشنهادی معرفی شده بر پایه ترکیب روش های انتخاب ویژگی Lasso، Gradient Boosting و Extra Tree از نظر دقت تشخیص، مناسب ارزیابی می شود، به طوریکه در شناخت حملات نوع DoS، Probe و U2R بیشترین دقت را نسبت به پژوهش های معرفی شده داشته و در مورد حملات R2L، با فاصله ۰،۳٪ دقت پایین تری را نسبت به بیشترین دقت سیستم های بررسی شده کسب نموده است. اما به عنوان یک سیستم منفرد، توانایی بالایی در شناخت مجموعه حملات را نسبت به سایر سیستم ها کسب کرده است. به عنوان پیشنهاد برای پژوهش های آتی می توان، عملکرد سیستم پیشنهادی حاضر را در حضور طبقه بند های مختلف بررسی نمود.

۵- مراجع

- [1] Cisco, "Cisco Visual Networking Index: Forecast and Methodology," 6 June 2017. [Online]. Available: https://www.cisco.com/c/dam/en/us/solutions/collateral/service_provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf.. [Accessed 26 4 2018].
- [2] S.M. Devine, "A process of defense – securing industrial control systems," Network Security, vol. 2017, no. 2, pp. 14-19, February 2017.
- [3] D. Alberto, "Peer-to-Peer Network Modeling for Adversarial Proactive Cyber Defenses," Master Thesis, Massachusetts Institute of Technology, Massachusetts, USA, 2017.
- [4] A.S. Shukla, R.Maurya, "Entropy-Based Anomaly Detection in a Network," Wireless Pers Commun, vol. 99, pp. 1487-1501, 2018.
- [5] W. Mazurczyk, M. Korczynski, K. Nakao, E. Kirda, C. Hesselman, K. Yoshioka, "Traffic Measurements for Cyber Security," IEEE Communications Magazine, vol. 55, no. 7, pp. 12-13, 2017.
- [6] P. Parameshwarappa, Z. Chen, A. Gangopadhyay, "Analyzing Attack Strategies Against Rule-Based Intrusion Detection Systems," in In Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking (Workshops ICDCN '18). ACM, New York, NY, USA, 2018.
- [7] M.H. Bhuyan, D.K. Bhattacharyya, J.K. Kalita, Network Traffic Anomaly Detection and Prevention : Concepts, Techniques, and Tools, Cham, Switzerland: Springer International Publishing

Confusion Matrices			
DoS		Probe	
Predicted attacks	0 1	Predicted attacks	0 2
Actual attacks		Actual attacks	
	0 9663 48		0 9476 235
	1 4382 3078		2 1490 931
R2L		U2R	
Predicted attacks	0 3	Predicted attacks	0 4
Actual attacks		Actual attacks	
	0 9711 0		0 9710 1
	3 2884 1		4 65 2

شکل ۵- ماتریس درهم ریختگی سیستم تشخیص نفوذ پیشنهادی

شاخص های دقت^{۲۴}، صحت^{۲۵}، Recall و F-measure را به جهت ارزیابی سیستم تشخیص نفوذ پیشنهادی به کمک ماتریس های درهم ریختگی ارائه شده در جدول ۵، طبق روابط ۱ الی ۴ محاسبه و نهایتاً در جدول ۶ نمایش داده شده است.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1_Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

با استفاده از شاخص دقت، می توان دریافت که آیا مدل درست آموزش دیده است یا خیر و کارایی آن به طور کلی چگونه بوده و تا چه اندازه خروجی را درست پیش بینی می کند. از معیار صحت به جهت تشخیص اینکه مدل نتیجه را مثبت (positive) تا چه اندازه درست پیش بینی نموده، استفاده می شود. زمانی که ارزش تشخیص های نادرست، بالا باشد، معیار Recall، معیار مناسبی خواهد بود. در صورتی که این معیار، مقدار کمی داشته باشد، در واقع نرخ تشخیص های اشتباه، زیاد است. معیار F1_Measure، یک معیار مناسب برای ارزیابی دقت یک آزمایش است [25]. این معیار Precision و Recall را با هم در نظر می گیرد. در جدول ۶، شاخص های ارزیابی سیستم تشخیص نفوذ پیشنهادی، برای هر گروه از حملات، نمایش داده شده است.

²⁴Accuracy

²⁵Precision

- AG, 2017.
- [8] S. Latha, Si. Prakash, "A survey on network attacks and Intrusion detection systems," in 4th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, India, 2017.
- [9] M.F. Umer, M. Sher, Y. Bi, "Flow-based intrusion detection: Techniques and challenges," *Computers & Security*, vol. 70, pp. 238-254, 2017.
- [10] I. Nevat, D. M. Divakaran, S. Member, S.G. Nagarajan, P. Zhang, "Anomaly Detection and Attribution in Networks With Temporally Correlated Traffic," *IEEE/ACM TRANSACTIONS ON NETWORKING (TON)*, vol. 26, no. 1, pp. 131-144, 2018.
- [11] "IBM Study: Responding to Cybersecurity Incidents Still a Major Challenge for Businesses," IBM, 14 3 2018. [Online]. Available: <http://www-03.ibm.com/press/us/en/pressrelease/53800.wss>. [Accessed 27 4 2018].
- [12] K. Panetta, "Top Trends in the Gartner Hype Cycle for Emerging Technologies, 2017," gartner, 2017. [Online]. Available: <https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/>. [Accessed 4 27 2018].
- [13] M. Stamp, *Introduction to Machine Learning with Applications in Information Security*, Florida: CRC Press, Taylor & Francis Group, 2018.
- [14] G. Hulten, *Building Intelligent Systems: A Guide to Machine Learning Engineering*, Washington, USA: Apress, 2018.
- [15] M. Tavallaei, E. Bagheri, W. Lu, A. Ghorbani, "A Detailed Analysis of the KDD Cup 99 Data Set," *Proceeding of the 3228 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA)*, pp. 53-58, 2009.
- [16] M. Onik, N.A. Zaben, H.P. Hoo, C. Kim, "A Novel Approach for Network Attack Classification Based on Sequential Questions," *Annals of Emerging Technologies in Computing (AETiC)*, vol. 2, no. 1, pp. 1-14, 2018.
- [17] N. AlNuaimi, M.M. Masud, M.A. Serhani, N. Zaki, "Streaming feature selection algorithms for big data: A survey," *Applied Computing and Informatics*, vol. 15, no. 1, pp. 13-29, 2019.
- [18] L. Dhanabal, S.P. Shantharajah, "A Study on NSL_KDD Dataset for Intrusion Detection System Based on Classification Algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 6, pp. 446-452, 2015.
- [19] B. Senthilnayagi, K. Venkatalakshmi, A. Kannan, "Intrusion Detection Using Optimal Genetic Feature Selection and SVM based Classifier," in 3rd International Conference on Signal Processing, Communication and Networking, Chennai, India, 2015.
- [20] F. Zhang, D. Wang, "An Effective Feature Selection Approach for Network Intrusion Detection," in IEEE Eighth International Conference on Networking, Architecture and Storage, Xi'an, China, 2013.
- [21] A. Alazab, M. Hobbs, J. Abawajy, M. Alazab, "Using Feature selection for intrusion detection system," in International Symposium on Communications and Information Technologies (ISCIT), Gold Coast, QLD, Australia, 2012.
- [22] S. Mukherjee, N. Sharma, "Intrusion Detection using Naive Bayes Classifier with Feature Reduction," *Procedia Technology*, vol. 4, pp. 119-128, 2012.
- [23] S. Parsazad, E. Saboori, A. Allahyar, "Fast Feature Reduction in intrusion detection datasets," in Proceedings of the 35th International Convention MIPRO, Opatija, Croatia, 2012.
- [24] H. Liu, H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Springer, 2013.
- [25] M. Navin, R. Pankaja, "Performance Analysis of Text Classification Algorithms using Confusion Matrix," *International Journal of Engineering and Technical Research (IJETR)*, vol. 6, no. 4, pp. 75-78, 2016.



مدل‌سازی عبارات ظاهر نشده با استفاده از یادگیری عمیق

امین قاضی زاهدی^۱، مرتضی زاهدی^۲، منصور فاتح^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

^۱ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

^۱ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

چکیده

عبارت کلیدی نشان‌دهنده و بیانگر مفاهیم اساسی برای یک متن است. در بسیاری از کاربردهای پردازش متن، نیاز به استخراج عبارات کلیدی با کیفیت مناسب است. در روش‌های قدیمی مدل‌سازی متن، معانی و مفاهیم مرتبط با متن اهمیت ویژه‌ای نداشته‌اند. در نتیجه ساختار اصلی مدل، بر اساس کلمات ظاهر شده در سند شکل گرفته است. طبق تحقیقات اخیر، اسناد هم‌خوشه اشتراک مناسبی در عبارات کلیدی مستقیماً ظاهر نشده در سند متنی دارند. در این تحقیق، مدلی جهت استخراج مفاهیم اساسی متن با استفاده از تخمین متون مشابه و اضافه کردن عبارات کلیدی به لایه‌های مخفی شبکه عمیق، ارائه شده است. این مدل‌سازی بر اساس تشابه خوشه‌ای با یک شبکه عمیق شکل گرفته است. این ساختار، سعی در کشف عبارات ظاهر نشده و سپس تخمین بهتر عبارات کلیدی با یک RNN را دارد. این روش به‌طور قابل توجهی مشکل عدم بازنمایی مفاهیم اساسی و عدم وجود عبارات کلیدی به‌طور مستقیم را حل می‌کند. روش پیشنهادی روی چهار دادگان متداول در این حوزه آزمایش شده است. نتایج این آزمایش‌ها، بهبود حدود ۱۴ درصدی نسبت به روش‌های معمول KEA و RNN را نشان می‌دهد.

کلمات کلیدی

استخراج عبارات کلیدی - شبکه عصبی عمیق - یادگیری دنباله به دنباله - خوشه‌بندی.

۱- مقدمه

داده نشده در متن را به‌عنوان عبارات کلیدی برچسب می‌زنند. در رویکرد دوم، غالباً برای استخراج موارد معنایی مشاهده نشده در متن از موارد هم‌معنی و مشابه کلمات متنی استفاده شده است. در حال حاضر شبکه‌های عصبی عمیق در حوزه وظایف مربوط به پردازش زبان طبیعی به کیفیت‌های مناسبی رسیده‌اند. این موارد مشتمل بر خلاصه‌سازی متون [1]، استخراج عبارات کلیدی [2]، سیستم پرسش و پاسخ [3] و تجزیه نحوی [4] است. این کاربردها اغلب به‌صورت دنباله‌ای از ورودی-خروجی در مدل‌سازی نمایش داده می‌شوند. از این رو یادگیری دنباله-به-دنباله (sequence-to-sequence) از رویکردهای جدی در آن‌ها است. یادگیری دنباله-به-دنباله رشته‌ای از ورودی را به رشته‌ای از خروجی نگاشت می‌کند. در این یادگیری، مدل احتمالاتی مبتنی بر حداکثر کردن احتمال بر روی دادگان ایجاد می‌شود [5]. از مهم‌ترین ویژگی‌های دنباله-به-دنباله عدم برابری طول دنباله ورودی با دنباله خروجی است. در سال‌های اخیر، روش‌های مبتنی بر یادگیری عمیق به علت موفقیت‌های کاربردی در یادگیری بدون نظارت، به‌طور گسترده‌ای

عبارات کلیدی^۱ جزئی از متن است که نمایشگر مفاهیم اصلی و موضوعات مهم متن است. این عبارات می‌توانند به‌وضوح در سند نمایش داده نشده باشد. طبق تحقیق این مطالعه، لغات و عبارات کلیدی نمایش داده نشده در متن، حدود ۵۰ درصد از کلمات کلیدی را شامل می‌شوند. این عبارات مستقیماً در متن وجود ندارد و فقط از لحاظ معنایی قابل استخراج است. این عبارات سهم مهمی در نمایش مفهوم یک سند دارد.

برای استخراج عبارات کلیدی دو رویکرد اصلی وجود دارد. در رویکرد اول، استخراج کلمات کلیدی مبتنی بر موارد قابل مشاهده در متن منبع و تخمین از روی مدل است. اما نویسندگان، علی‌الخصوص نویسندگان نشریات علمی از این رویکرد استفاده نمی‌کنند و غالباً موارد معنی‌دار و مفاهیم معنایی مستقیماً نمایش

^۱. Keyphrases

در ادامه این مقاله، ابتدا در بخش دوم پژوهش‌های پیشین بررسی شده‌اند. سپس، در بخش سوم چهارچوب مدل پیشنهادی ارائه شده است. در بخش چهارم، نتایج مدل پیشنهادی روی چهار دادگان مختلف ارائه شده‌اند. در بخش پنجم، نتیجه‌گیری نهایی و پیشنهاد برای کارهای آتی ارائه شده است.

۲- پیشینه و کارهای مرتبط

۲-۱- استخراج خودکار کلمات کلیدی

عبارت کلیدی، عبارتی مختصر و دقیق برای توصیف یک موضوع یا یک مبحث در سند است. الگوریتم‌های مختلفی برای استخراج عبارات کلیدی پیشنهاد شده‌اند. فرآیند استخراج عبارات کلیدی معمولاً به دو مرحله تقسیم می‌شود. اولین قدم ایجاد فهرستی از کاندیدهای عبارات با روش‌های اکتشافی است. روش‌های اصلی استخراج کاندیدها شامل حفظ توالی‌های کلمه است. این روش‌ها معمولاً با الگوهای Part-Of-Speech مطابقت دارند [12], [13]. برای حفظ توالی کلمات، استخراج N-GRAM های مهم و noun phrase ها ارائه شده‌اند [14], [15].

گام دوم برای استخراج کلمات کلیدی، رتبه‌بندی کاندیدها بر اساس اهمیت آن‌ها در سند است. پس از رتبه‌بندی، کاندیدهای برتر به‌عنوان عبارات کلیدی اعلام می‌شوند. در این گام، روش‌های یادگیری بانظارت و بدون نظارت، به‌طور گسترده استفاده می‌شوند. در روش‌های نظارت‌شده، این وظیفه به‌عنوان یک مسئله طبقه‌بندی باینری حل شده است و برای طبقه‌بندی مناسب، انواع مختلف روش‌های یادگیری و ویژگی‌های مختلف موردبررسی قرار گرفته‌اند [18]-[14].

در مورد رویکردهای بدون نظارت، ایده‌های اولیه شامل یافتن گره‌های مرکزی^۸ در گراف متن^۹ است [19], [20]. تشخیص عبارات کلیدی از خوشه‌های موضوعی^{۱۰} [21], [22] از کاربردهای رویکردهای بدون نظارت در استخراج کلمات کلیدی است.

به‌جز فرآیند دومرحله‌ای استخراج عبارات کلیدی، روش‌های دیگری برای این فرآیند وجود دارد. در برخی از این روش‌ها، دو مدل زبانی به‌عنوان معیار phraseness و عبارات بدون اطلاعات بیان‌شده است [24]. در مرجع [13] مدل الگوریتم عصبی مصنوعی مشترک، برای استخراج عبارات کلیدی از توییت‌ها پیشنهاد شده است که کاربرد دیگری از شبکه‌های عصبی عمیق در زمینه استخراج عبارات کلیدی است. این مدل، متمرکز بر چرچسب‌گذاری توالی‌ها است و قادر به پیش‌بینی عبارات کلیدی ظاهر نشده نیست.

مرجع [12] ایده بسیار نزدیکی به روش پیشنهادی در این مقاله، ارائه کرده است. در این تحقیق، از یک مدل ارتباط کلمه^{۱۱} استفاده

در حال استفاده هستند. این روش‌ها نقش مهمی در کاهش ابعاد جهت خوشه‌بندی دارند. برای نمونه اتوانکدر انباشته^۲ [6]، Deep CCA (DCCA) [7] و اتوانکدر پراکنده^۳ [8] نمونه‌های کاربردی از به‌کارگیری این نوع یادگیری هستند. به‌طور کلی رویکردهای استفاده از کاهش ابعاد با هدف پیش‌پردازش و فشردگی پیش از خوشه‌بندی، مفهومی از PCA و CCA را روی شبکه یادگیری عمیق ایجاد می‌کنند. این رویکردها موجب یادگیری نگاشتی غیرخطی از دامنه داده به فضاهای پنهان با ابعاد کم‌تر می‌شوند. این عملیات در شبکه‌های عصبی عمیق به‌عنوان مرحله پیش‌پردازش جداگانه تلقی می‌شوند و دقت مرحله خوشه‌بندی را افزایش می‌دهند. در واقع انتظار می‌رود که نمایه‌های پنهان داده‌های آموخته‌شده توسط این شبکه‌های عصبی عمیق به‌طور طبیعی برای خوشه‌بندی مناسب باشند. با این حال، صراحتاً هیچ اهداف ترویج^۴ خوشه‌ای در فرآیند یادگیری دخیل نیست. از این رو، شبکه‌های عصبی عمیق آموزش دیده، داده‌های کاهش‌یافته مناسب برای خوشه‌بندی را تضمین نمی‌کنند [9]. پس این نمایندگان پنهان^۵ دقت آموزش خوشه‌بندی را افزایش می‌دهند، اما اهداف ترویج خوشه‌ای^۶ به‌صراحت در فرآیند یادگیری دخیل نیستند. از این رو، همواره با کاهش ابعاد داده‌ها توسط شبکه‌های عمیق، نتایج خوشه‌بندی ارتقا نمی‌یابد [10], [11] و ارائه کاهش ابعاد هم‌زمان با خوشه‌بندی در یک ساختار، موجب بهبود خوشه‌بندی می‌شود. یک روش مؤثر مبتنی بر استفاده هم‌زمان از عملیات کاهش ابعاد و خوشه‌بندی مبتنی بر شبکه عصبی عمیق در [11] ارائه شده است.

هدف این مقاله، مدل‌سازی مفاهیم پنهان اصلی به‌طور مناسب و تخمین عبارات کلیدی با کیفیت مطلوب است. این مهم با خوشه‌بندی اولیه متون و نگاشت عبارت کلیدی به‌کل متون یک خوشه امکان‌پذیر شده است. به‌طور کلی این مقاله دارای دو بخش است. در ابتدا از یک شبکه عصبی عمیق خوشه‌بندی^۷ (DCN) جهت خوشه‌بندی اسناد مشابه استفاده شده است. این شبکه به کشف موارد ظاهر نشده کمک می‌کند. در ادامه این تحقیق، از یک تخمین‌گر عبارات کلیدی توسط آموزش یک مدل مبتنی بر RNN با آموزش دنباله-به-دنباله استفاده شده است. در این مدل، دنباله‌ای از عبارات متنی اسناد هم‌خوشه به‌عنوان ورودی آموزش داده شده است. در واقع، عبارات ظاهر نشده در یک متن، با مفهوم مشابه در اسناد هم‌خوشه دیگر ظاهر شده‌اند. با این فرآیند، این عبارات ظاهر نشده وارد چرخه آموزش مدل شده‌اند. نتایج آموزش‌ها توسط چهار دادگان مختلف بررسی شده است. این نتایج، نشان‌دهنده افزایش کیفیت شناسایی عبارات مستقیم ظاهر نشده در سند است. این عبارات بیشتر جنبه مفهومی دارند و مستقیماً در متن سند بکار نرفته‌اند.

². Stacked Auto Encoder (SAE)

³. Sparse Autoencoder

⁴. Clustering-Promoting

⁵. Latent Representations

⁶. Clustering-Promoting

⁷. Deep Clustering Network (DCN)

⁸. Central Nodes

⁹. Text Graph

¹⁰. Topical Clusters

¹¹. Word Alignment Model

داده‌ها، از ایده مشترک کاهش ابعاد و خوشه‌بندی استفاده شده است. در مراجع [26]، [11] تبدیلات غیرخطی قدرتمندی ارائه شده است. ایده اصلی این روش‌ها، اتصال یک ماژول خوشه‌بندی به لایه خروجی و یادگیری مشترک پارامترهای شبکه‌های عصبی عمیق و خوشه‌ها است. به‌طور مشخص، مسئله بهینه‌سازی به‌صورت رابطه (۳) مدل می‌شود.

$$\min_{W, \theta} \hat{L} = \sum_{i=1}^N q(\mathbf{f}(x_i; \mathbf{W}); \theta) \quad (3)$$

که $\mathbf{f}(x_i; \mathbf{W})$ خروجی شبکه و x_i داده نمونه است. \mathbf{W} پارامترهای شبکه را جمع می‌کند. θ پارامترهای برخی از مدل خوشه‌بندی را نشان می‌دهد. اگر رابطه (۱) برای K-means بکار رود، θ به‌عنوان مراکز M و انتساب به S_i است. در رابطه (۳)، $q(\cdot)$ نمایشگر برخی خوشه‌های از دست‌رفته است.

این ایده به نظر منطقی، اما مشکل‌ساز است. یک رویکرد بهینه محلی برای رابطه (۳)، $\mathbf{f}(x_i; \mathbf{W}) = 0$ و $\hat{L} = 0$ است. نوع دیگری از راهکارهای بدیهی جهت نگاشت مطلق نمونه داده‌ها به خوشه‌های محکم^{۱۹}، در نظر گرفتن مقداری ناچیز برای \hat{L} است. این مسئله می‌تواند بسیار ناخوشایند باشد زیرا هیچ تدارکی برای احترام به نمونه داده‌های x_i وجود ندارد [27].

این مدل‌سازی رابطه بین داده‌های مشاهده‌شده x_i و بازنمایی پنهان^{۲۰} از خوشه h_i را به‌صورت یک نگاشت غیرخطی بیان می‌کند. شبکه‌های عصبی عمیق توانایی خوبی برای تقریب تعداد متناسب پارامترها دارند. از این‌رو، روش پیشنهادی از این شبکه‌های عصبی عمیق به‌عنوان یک تابع نقشه‌برداری استفاده می‌شود.

۲-۳- مدل کدگذار-کدگشا در شبکه عصبی بازگشتی^{۲۱}

سیستم کدگذار-کدگشای مبتنی بر RNN به‌طور مناسبی برای انجام وظایف واقعی برای کاربردهای دنباله-به-دنباله عملی شده است [28]، [4]، [29]. در چارچوب کدگذار-کدگشا، توالی منبع $X = [x_1, \dots, x_{T_s}]$ به یک بردار طول ثابت c توسط رمزگذار RNN تبدیل می‌شود، یعنی:

$$\mathbf{h}_t = f(x_t, \mathbf{h}_{t-1}); \mathbf{c} = \phi(\{\mathbf{h}_1, \dots, \mathbf{h}_{T_s}\}) \quad (4)$$

که \mathbf{h}_t شامل وضعیت‌های RNN، \mathbf{c} بردار متن، \mathbf{F} تابع پویا و ϕ خلاصه‌ساز لایه‌های مخفی است. وضعیت نهایی نیز با \mathbf{h}_{T_s} نمایش داده می‌شود. در عمل مشخص شده است که RNN می‌تواند متشکل از LSTM [30] و یا GRU [28] باشد. رمزگشای RNN برای باز کردن یک بردار متن c به دنباله هدف از طریق مدل تخمین زیر استفاده می‌کند:

$$\mathbf{s}_t = f(y_{t-1}, \mathbf{s}_{t-1}, \mathbf{c}); p(y_t | y < t, X) = g(y_{t-1}, \mathbf{s}_{t-1}, \mathbf{c}); \quad (5)$$

شده است که سند را به عبارات کلیدی ترجمه می‌کند. این رویکرد مشکل شکاف واژگان^{۱۲} بین منبع و هدف را به درجه خاصی از بین می‌برد. با این حال، این مدل قادر به مقابله با مسئله شناخت معنایی نیست. علاوه بر این، این مدل با هدف پیدا کردن عنوان/خلاصه برای تعداد زیادی از نمونه‌های آموزشی ارائه شده است که ممکن است از هدف واقعی تولید عبارات کلیدی متفاوت باشد.

در مرجع [25] استخراج عبارات کلیدی ظاهر نشده با نام (عبارات غایب^{۱۳}) ارائه شده است. در این تحقیق، فرآیند کپی بر اساس مشابهت به‌دست‌آمده در ساختار مدل زبانی افزون شده است. این راهکار، کیفیت شناسایی عبارات ظاهر نشده را افزایش داده است.

۲-۲- خوشه‌بندی با DCN

مجموعه‌ای از نمونه‌های داده به‌صورت $\{x_i\}_{i=1, \dots, N}$ وجود دارد که $x_i \in \mathbb{R}^M$ است. وظیفه خوشه‌بندی، دسته‌بندی نمونه‌های داده N به K طبقه است. این رویکرد بسیار منطبق بر روش K-means است که در رابطه (۱) نمایش داده شده است:

$$\min_{M \in \mathbb{R}^{M \times K}, \{s_i \in \mathbb{R}^K\}} \sum_{i=1}^N \|x_i - M s_i\|_2^2 \quad (1)$$

که s_i یک بردار تخصیص^{۱۴} از نقطه داده i است که تنها یک علامت غیر صفر^{۱۵} دارد. $s_{i,j}$ ، نشان‌دهنده زامین عنصر s_i و ستون k ام از M است. خوشه‌بندی داده‌های با ابعاد بزرگ توسط K-means مناسب نیست. البته با استفاده از پیش‌پردازش و کاهش ابعاد داده‌ها، می‌توان خوشه‌بندی مناسب‌تری با K-means انجام داد. در این تحقیق با استفاده از PCA و NMF ابعاد داده‌ها کاهش یافته و سپس با K-means نتایج بهتری حاصل شده است.

اگرچه تصور می‌شود که خوشه‌های داده در دامنه‌های پنهان به‌خوبی جدا شده‌اند اما این کار توسط یک تابع بهینه تبدیل یافته امکان‌پذیر است. در مرجع [11] این مسئله بهینه‌سازی مشترک^{۱۶} به‌صورت رابطه (۲) نمایش داده شده است.

$$\min_{M, \{s_i\}, W, H} \|X - WH\|_F^2 + \lambda \sum_{i=1}^N \|h_i - M s_i\|_2^2 + r_1(H) + r_2(W) \quad (2)$$

که $X = [x_1, \dots, x_N]$ و $H = [h_1, \dots, h_N]$ و $\lambda \geq 0$ است که یک پارامتر برای توازن وفاداری داده‌ها^{۱۷} و ساختار خوشه پنهان است. رابطه (۲) شامل دو مرحله است. اولین مرحله کاهش ابعاد و دومین مرحله خوشه‌بندی پنهان است. ضرایب r_1 و r_2 جهت تنظیم^{۱۸} پارامترها هستند. جزئیات بیشتر در مرجع [11] وجود دارد. مدل داده‌ای $X \approx WH$ ممکن است بیش‌ازحد ساده باشد و روند تولید داده‌ها می‌تواند بسیار پیچیده‌تر از این تبدیل خطی باشد. از این‌رو، محققان تبدیلات غیرخطی قدرتمندی ارائه داده‌اند. به‌عنوان مثال در یک شبکه عصبی عمیق بازگشتی، برای مدل‌سازی فرآیند تولید

¹². Vocabulary Gaps

¹³. Absent Keyphrases

¹⁴. Assignment Vector

¹⁵. Non-Zero Element

¹⁶. Joint Optimization

¹⁷. Data Loyalty

¹⁸. Regularization

¹⁹. Tight Clusters

²⁰. Hidden Representation

²¹. RNN Encoder-Decoder Model

۱۹	۴۴,۳۱	۱۹,۲۷۵	Inspec
۲۱	۵۷,۹۹	۱۲,۲۹۶	SemEval
۱۴	۵۲,۲۶	۲,۴۶۱	Krapivin

۳-۱- مرحله اول: مدل‌سازی خوشه‌بندی

عمیق

علت استفاده از این مدل‌سازی، رابطه‌ای است که در داده‌های مشاهده شده x_i و بازنمایی پنهان دوستان خوشه‌ای h_i از طریق نگاشت غیرخطی صورت می‌گیرد.

$$h_i = f(x_i; W), f(\cdot; W): \mathbb{R}^M \rightarrow \mathbb{R}^R, \quad (۷)$$

که $f(x_i; W)$ تابع نگاشت و W مجموعه‌ای از پارامترها است. شبکه‌های عصبی عمیق توانایی تقریب نگاشت مداوم با استفاده از تعداد متناسبی از پارامترها را دارند. از این رو، در روش پیشنهادی، این شبکه عصبی عمیق به‌عنوان یک تابع نگاشت استفاده می‌شود. این روش تلاش می‌کند خوشه‌بندی و فرآیند یادگیری شبکه عصبی عمیق را هم‌زمان انجام می‌دهد. در حقیقت این مورد با در نظر گرفتن رابطه (۸) می‌تواند حل شود. راه‌حل کلیدی برای اجتناب از کاهش ابعاد خطی در بخش بازسازی قرار دارد. این مورد شامل یک دوره بازسازی^{۲۵} در کاهش ابعاد مشترک مبتنی بر یک شبکه عصبی عمیق و K-means است. رابطه (۸) تابع هزینه را نشان می‌دهد.

$$\min_{W, Z, M, \{s_i\}} \sum_{i=1}^N (l(g(f(x_i)), x_i) + \frac{\lambda}{2} \|f(x_i) - Ms_i\|_2^2) \quad (۸)$$

در رابطه (۸)، نماد $f(x_i; W)$ و $g(h_i; Z)$ به $g(h_i)$ و $f(x_i)$ ساده‌سازی شده است. تابع $l(\cdot): \mathbb{R}^M \rightarrow \mathbb{R}$ یک تابع اتلاف^{۲۶} خاص است که خطای بازسازی را اندازه‌گیری می‌کند. $\lambda \geq 0$ یک پارامتر تنظیم است که خطا در بازسازی را در مقایسه با بازنمایی‌های پنهان K-means متعادل می‌کند.

ساختار شبکه‌های رمزگشایی یک نسخه بازپرداخت^{۲۷} برای داده، از فضای مخفی است. برای هر دو شبکه کدگذاری و کدگشایی، از نورون‌های مبتنی بر فعال‌سازی خطی اصلاح‌شده (ReLU) استفاده می‌شود [32]. این ساختار بهینه به‌خوبی در مرجع [11] نشان داده شده است.

۳-۲- مرحله دو: مدل کدگذار-کدگشا

ایده اصلی در این بخش مربوط به استفاده از مدل کدگذار-کدگشا در شبکه بازگشتی عصبی جهت استفاده در دنباله متن به‌صورت ورودی و بازنمایی آن عبارات کلیدی به‌صورت خروجی است. سپس مدل کدگذار-کدگشا برای استفاده از نگاشتی از توالی منبع و توالی هدف آماده می‌شود. به‌منظور ساده‌سازی، عبارت $(x; y)$ برای نشان دادن هر جفت داده در بقیه این بخش استفاده

که S_t وضعیت RNN در زمان t است. y_t نماد مقصد پیش‌بینی شده در زمان t توسط تابع $g(\cdot)$ است. مکانیسم توجه^{۲۲} در ابتدا برای دنباله-به-دنباله ارائه شد [31]، تا بارگذاری کل منابع را به یک بردار طول ثابت به‌عنوان محتوا بسپارد. در عوض، تمرکز بر روی استفاده از یک شیء پویا در تغییر روند رمزگشایی بود. یک گزینه طبیعی (به‌جای "توجه نرمی"^{۲۳})، نشان دادن c_t به‌عنوان مجموع وزنی از وضعیت‌های پنهان منبع است.

$$c_t = \sum_{\tau=1}^{T_s} \alpha_{t\tau} h_{\tau}; \quad \alpha_{t\tau} = \frac{e^{\eta(s_{t-1}, h_{\tau})}}{\sum_{\tau'} e^{\eta(s_{t-1}, h_{\tau'})}} \quad (۶)$$

که η تابع نشان‌دهنده میزان قوت مکانیسم توجه است.

۳-۳- روش پیشنهادی

در این بخش به تشریح ایده اصلی کار پرداخته شده است. ساختار این سیستم در دو مرحله مطرح شده است. در مرحله اول، خوشه‌بندی اسناد مشابه به کمک یک شبکه DCN انجام شده است. در این مرحله عبارات کلیدی ظاهر نشده، با استفاده از مشابهت‌های خوشه‌ای، به‌خوبی مشخص می‌شوند. در این مدل، دنباله‌ای از عبارات متنی اسناد هم‌خوشه به‌عنوان ورودی آموزش داده شده است. در واقع، عبارات ظاهر نشده در یک متن، با مفهوم مشابه در اسناد هم‌خوشه دیگر ظاهر شده‌اند. با این فرآیند، این عبارات ظاهر نشده وارد چرخه آموزش مدل شده‌اند. ساختار شبکه DCN در بخش ۳-۱ ارائه شده است. در مرحله دوم، مدل‌سازی تخمین عبارات کلیدی توسط RNN ارائه شده است. در این بخش، از یک تخمین‌گر عبارات کلیدی توسط آموزش یک مدل مبتنی بر RNN با آموزش دنباله-به-دنباله استفاده شده است. ورودی این بخش خوشه‌های تولید شده توسط شبکه DCN و خروجی آن عبارات کلیدی است. ساختار کلی RNN در بخش ۳-۲ ارائه شده است.

طبق بررسی‌های انجام شده، اسناد هم‌خوشه عبارات کلیدی ظاهر نشده مشابهی دارند. این میزان تشابه در جدول ۱ نشان داده شده است. نتایج جدول ۱ مربوط به ۳ دادگان متفاوت است. این دادگان در ادامه به‌طور کامل معرفی شده‌اند. نتایج جدول ۱ تشابه ۱۸ درصدی عبارات کلیدی ظاهر نشده در اسناد هم‌خوشه را نشان می‌دهد.

در نتیجه با یک خوشه‌بندی مناسب، اسناد مشابه شناسایی و عبارات کلیدی آن‌ها در مدل‌سازی نهایی دخیل می‌شوند. این رویکرد موجب تقویت مدل‌سازی در موارد عبارات کلیدی ظاهر نشده و در واقع مفاهیم معنایی متون می‌شود.

جدول ۱. تعداد عبارات کلیدی ظاهر نشده مشترک در خوشه‌ها

دادگان	تعداد عبارات کلیدی	% عبارات کلیدی ظاهر نشده در متن	% میانگین عبارات کلیدی ظاهر نشده مشترک در خوشه
--------	--------------------	---------------------------------	--

²⁴. Clustering-Friendly

²⁵. Reconstruction Sector

²⁶. Loss Function

²⁷. Mirrored Version

²². Attention Mechanism

²³. Soft Attention

می‌کند که ورودی‌ها حول موقعیت z و خروجی در موقعیت i تطابق دارند.

۴- نتایج آزمایش‌ها

در این بخش به بررسی نتایج حاصل از اجرای نهایی الگوریتم‌ها پرداخته شده است. در جدول ۲ مقایسه عملکرد روش‌های متداول قبلی و روش پیشنهادی نمایش داده شده است. در بین ۷ روش ارائه شده ۴ روش (TF-Idf و TextRank و KEA و Maui) جزو روش‌های مرسوم و سه روش دیگر دارای آموزش مبتنی بر شبکه‌های عمیق عصبی هستند. این روش‌ها به جز روش CopyRNN و روش پیشنهادی، رویکرد خاصی جهت شناسایی کلمات کلیدی ظاهر نشده ندارند.

جدول ۲. نتایج کیفیت الگوریتم‌ها بر اساس معیار f-measure

Method	Hulth2003	SemEval	Krapivin	NUS
TF-Idf	۰,۳۰۹	۰,۱۸۹	۰,۱۵۸	۰,۱۸۳
TextRank	۰,۲۸۱	۰,۱۸۷	۰,۱۶۶	۰,۱۹۶
KEA	۰,۱۲۵	۰,۰۲۶	۰,۱۵۲	۰,۰۸۴
Maui	۰,۰۵۷	۰,۰۳۹	۰,۲۱۶	۰,۲۶۸
RNN	۰,۰۶۴	۰,۱۲۴	۰,۰۸۸	۰,۱۲۷
CopyRNN	۰,۳۴۲	۰,۳۰۴	۰,۲۶۶	۰,۳۱۳
روش پیشنهادی	۰,۳۸۱	۰,۲۸۸	۰,۳۱۷	۰,۳۴۴

بهترین مقدار هر ستون در جدول ۲ پررنگ و دومین مقدار هر ستون خط زیر دار شده‌اند. نتایج جدول ۳ نشان‌دهنده بهبود مقادیر تشخیص عبارات کلیدی برای روش پیشنهادی است. با بررسی مقادیر دوم بهتر، روش CopyRNN مناسب‌تر بوده است. به‌طور کلی روش‌های یادگیری بانظارت بهتر عمل کرده است. خروجی اصلی این تحقیق، پیرامون عملکرد روش‌های مبتنی بر یادگیری عمیق و همچنین نحوه کشف موارد ظاهر نشده است. از این رو، نتایج بررسی موارد ظاهر نشده در جدول ۳ به‌طور خاص بررسی شده است.

جدول ۳. نتایج کیفیت الگوریتم‌ها بر اساس Recall روی موارد ظاهر نشده (روش‌های یادگیری عمیق)

Method	Hulth2003	SemEval	Krapivin	NUS
RNN	۰,۰۳۱	۰,۰۴۱	۰,۰۹۵	۰,۰۵۰
CopyRNN	۰,۰۴۷	۰,۰۴۳	۰,۱۱۳	۰,۰۵۸
روش پیشنهادی	۰,۰۵۱	۰,۰۴۱	۰,۱۲۸	۰,۰۹۳

نتایج جدول ۳، نمایانگر پیشرفت بسیار مناسب این روش نسبت به مدل پایه‌ای و همچنین روش CopyRNN است. روش CopyRNN به‌طور خاص روی عبارات کلیدی ظاهر نشده ارائه شده است که روش پیشنهادی از این روش بهتر عمل کرده است. نکته حائز اهمیت در این بخش، تعیین تعداد کاندیدا برای انتخاب عبارات کلیدی نهایی است. تعداد کاندیدای عبارات کلیدی در روش CopyRNN برای شناسایی عبارات صحیح، زیاد است. در این روش معمولاً باید ۳۰ کاندیدا لحاظ شوند. با افزایش تعداد کاندیدا،

می‌شود که x ترتیب کلمه یک متن منبع و y ترتیب کلمه از عبارت کلیدی آن است. در ساختار اصلی مدل تولید کلمات کلیدی، محتویات متن منبع با یک رمزگذار فشرده شده و فاکتورهای کلیدی مرتبط با رمزگشایی بر اساس آن بازسازی می‌شوند. هر دو بخش کد گزار و کدگشا با شبکه‌های عصبی بازگشتی (RNN) اجرا می‌شوند. کدگذار RNN دنباله متغیر ورودی $\mathbf{x} = (x_1, x_2, \dots, x_T)$ را به مجموعه‌ای از لایه بازنمایی مخفی $\mathbf{h} = (h_1, h_2, \dots, h_T)$ تبدیل می‌کند، با تکرار معادلات در طول زمان t داریم:

$$\mathbf{h}_t = f(x_t, \mathbf{h}_{t-1}) \quad (9)$$

که f یک تابع غیرخطی است؛ و بردار متن \mathbf{c}^{28} به‌عنوان لایه بازنمایی کل ورودی \mathbf{x} از طریق یک تابع غیرخطی q عمل می‌کند.

$$\mathbf{c} = q(h_1, h_2, \dots, h_T) \quad (10)$$

برای کدگشا از یک RNN دیگر استفاده می‌کنیم که بردار متن را بازسازی می‌کند و یک دنباله $\mathbf{y} = (y_1, y_2, \dots, y_T)$ لغت به لغت را از طریق یک مدل زبانی شرطی g تولید می‌کند:

$$\begin{aligned} \mathbf{s}_t &= f(y_{t-1}, \mathbf{s}_{t-1}, \mathbf{c}) \\ p(y_t | y_{1, \dots, t-1}, \mathbf{x}) &= g(y_{t-1}, \mathbf{s}_t, \mathbf{c}) \end{aligned} \quad (11)$$

که \mathbf{s}_t وضعیت پنهان RNN رمزگذار در زمان t است. تابع غیرخطی g یک طبقه‌بندی کننده softmax است که احتمال تمام کلمات را در واژگان را نمایش می‌دهد. y_t کلمه پیش‌بینی شده در زمان t است که با محاسبه کلمه‌ای با بزرگ‌ترین احتمال شرطی توسط $g(\cdot)$ به دست می‌آید. شبکه‌های کدگذار و کدگشا به‌طور مشترک آموزش می‌یابند تا حداکثر احتمال شرطی دنباله هدف را با توجه به توالی منبع مقایسه کنند. پس از آموزش، از جستجوی پرتوی 3^3 برای تولید عبارات استفاده شده و یک max-heap برای دستیابی به توالی‌های پیش‌بینی شده کلمه با بیشترین احتمال وجود دارد.

در بخش کدگذار از یک GRU استفاده شده است. با توجه به بهبود کارایی اعلام شده در [28] از یک forward GRU دیگر به‌عنوان کدگشا استفاده می‌شود. بردار متن \mathbf{c} به‌صورت مجموع وزنی بازنمایی پنهان محاسبه می‌شود:

$$\begin{aligned} \mathbf{c}_i &= \sum_{j=1}^T \alpha_{ij} h_j \\ \alpha_{ij} &= \frac{\exp(a(s_{i-1}, h_j))}{\sum_{k=1}^T \exp(a(s_{i-1}, h_k))} \end{aligned} \quad (12)$$

که $a(s_{i-1}, h_j)$ تابع هم‌ترازی نرم 3^1 است که شباهت بین s_{i-1} و h_j را اندازه‌گیری می‌کند. این تابع یعنی حداکثر درجه‌ای را تعیین

²⁸ Context Vector

²⁹ Conditional Language Model

³⁰ Beam Search

³¹ Soft Alignment Function

- [8] A. Ng, "Sparse autoencoder," *CS294A Lect. notes*, pp. 1–19, 2011.
- [9] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, "Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering," Oct. 2016.
- [10] V. M. Patel, H. Van Nguyen, R. R. Vidal, H. Van Nguyen, and R. R. Vidal, "Latent space sparse subspace clustering," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 225–232.
- [11] B. Yang, X. Fu, and N. D. Sidiropoulos, "Learning From Hidden Traits: Joint Factor Analysis and Latent Clustering," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 256–269, Jan. 2017.
- [12] Z. Liu, X. Chen, Y. Zheng, and M. Sun, "Automatic keyphrase extraction by bridging vocabulary gap," in *Proceedings of the Fifteenth Conference on Computational Natural Language Learning*, 2011, no. June, pp. 135–144.
- [13] Q. Zhang, Y. Wang, Y. Gong, and X. Huang, "Keyphrase extraction using deep recurrent neural networks on Twitter," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 836–845.
- [14] A. Hulth and B. B. Megyesi, "A study on automatically extracted keywords in text categorization," in *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL - ACL '06*, 2006, no. July, pp. 537–544.
- [15] O. Medelyan, E. Frank, and I. H. Witten, "Human-competitive tagging using automatic keyphrase extraction," in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3*, 2009, vol. 3, no. August, pp. 1318–1327.
- [16] I. Witten, G. Paynter, E. Frank, C. Gutwin, and C. Nevill-Manning, "KEA: Practical Automatic Keyphrase Extraction DL'99 *Proceedings of the fourth ACM conference on Digital libraries*," DOI, 1999.
- [17] S. Das Gollapalli and C. Caragea, "Extracting Keyphrases from Research Papers Using Citation Networks," in *AAAI*, 2014, pp. 1629–1635.
- [18] P. Lopez and L. Romary, "HUMB: Automatic key term extraction from scientific articles in GROBID," in *Proceedings of the 5th international workshop on semantic evaluation*, 2010, no. July, pp. 248–251.
- [19] R. Mihalcea, "Graph-based ranking algorithms for sentence extraction, applied to text summarization," in *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions -*, 2004, vol. 85, pp. 20-es.
- [20] M. Grineva, M. Grinev, and D. Lizorkin, "Extracting key terms from noisy and multitheme documents," in *Proceedings of the 18th international conference on World wide web - WWW '09*, 2009, p. 661.
- [21] Z. Liu, P. Li, Y. Zheng, and M. Sun, "Clustering to Find Exemplar Terms for Keyphrase Extraction," 2009.
- [22] Z. Liu, W. Huang, Y. Zheng, and M. Sun, "Automatic Keyphrase Extraction via Topic Decomposition," no. October, pp. 366–376, 2010.
- [23] F. Boudin, "Unsupervised Keyphrase Extraction with Multipartite Graphs," pp. 667–672, 2018.
- [24] T. Tomokiyo and M. Hurst, "A language model approach to keyphrase extraction," in *Proceedings of the ACL 2003 workshop on Multiword expressions analysis, acquisition and treatment -*, 2003, vol. 18, pp. 33–40.
- [25] R. Meng, S. Zhao, S. Han, D. He, P. Brusilovsky, and Y. Chi, "Deep Keyphrase Generation," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2017, pp. 582–592.
- [26] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *International conference on machine learning*, 2016, vol. 48, pp. 478–487.
- [27] J. Yang, D. Parikh, and D. Batra, "Joint Unsupervised Learning of Deep Representations and Image Clusters," in *2016 IEEE Conference on Computer Vision and Pattern*

روش‌های مختلف تقریباً خروجی مشابهی دارند و برتری نسبت به یکدیگر ندارند. زمانی شناسایی معنی می‌یابد که تعداد کاندیدا معقول باشد. افزایش تعداد کاندیدا، مسئله شناسایی را بی‌معنی می‌کند.

۵- نتیجه‌گیری

در این مطالعه، یک ساختار دویخشی جهت شناسایی عبارات کلیدی توسط شبکه‌های عصبی عمیق ارائه شد. در بخش اول به شناسایی مناسب موارد ظاهر نشده پرداخته شد. شناسایی عبارات کلیدی ظاهر نشده منجر به بالا رفتن کیفیت کلی شناسایی عبارات کلیدی شد. این مرحله در روش‌های گذشته به‌طور معمول صورت پذیرفته بود. در واقع با یک خوشه‌بندی عمیق، شناسایی عبارات ظاهر نشده مؤثر با استفاده از موارد هم‌خوشه انجام شد. سپس در بخش دوم، از اطلاعات به‌دست‌آمده در مدل‌سازی یک شبکه بازگشتی عمیق مبتنی بر دنباله-به-دنباله استفاده شد. ساختار متن و عبارت کلیدی به‌صورت دنباله‌های متنی به‌خوبی در یک شبکه عصبی عمیق بازگشتی مدل شد. این رویکرد توانست نتایج قابل قبولی برای مدل‌سازی موارد ظاهر نشده به نمایش گذارد. در نهایت نتایج نشان داد که این مدل‌سازی در شناسایی موارد ظاهر نشده به‌خوبی عمل کرد و همچنین در مقایسه با روش‌های قبلی، نتایج کلی موفق‌تری در تخمین عبارات کلیدی ارائه داد. نتایج به‌طور متوسط بهبود ۱۴ درصدی روی دادگان را نشان می‌دهد. رویکرد حاضر نشان‌دهنده تأثیر مدل‌سازی معنایی در کیفیت تخمین عبارات کلیدی است. در کارهای آینده عملیات پیدا کردن تعداد خوشه مناسب و نحوه ارجاع آن به RNN نهایی، موجب کیفیت مناسب‌تر الگوریتم خواهد شد.

مراجع

- [1] S. Song, H. Huang, and T. Ruan, "Abstractive text summarization using LSTM-CNN based deep learning," *Multimed. Tools Appl.*, vol. 78, no. 1, pp. 857–875, 2019.
- [2] J. Gu, Z. Lu, H. Li, and V. O. K. Li, "Incorporating Copying Mechanism in Sequence-to-Sequence Learning," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2016, pp. 1631–1640.
- [3] V. Abishek, C. Valliyammai, and B. Hariharan, *An Enhanced Deep Learning Model for Duplicate Question Pairs Recognition*, vol. 758. Springer Singapore, 2019.
- [4] O. Vinyals, L. Kaiser, T. Koo, S. Petrov, I. Sutskever, and G. Hinton, "Grammar as a foreign language," in *Advances in Neural Information Processing Systems*, 2015, pp. 2773–2781.
- [5] S. Shen et al., "Minimum Risk Training for Neural Machine Translation," *arXiv Prepr. arXiv1512.02433*, Dec. 2015.
- [6] K. S. Hasan and V. Ng, "Conundrums in unsupervised keyphrase extraction: making sense of the State-of-the-Art," in *Coling 2010 - 23rd International Conference on Computational Linguistics, Proceedings of the Conference*, 2010, no. August, pp. 365–373.
- [7] G. Andrew, R. Arora, J. Bilmes, and K. Livescu, "Deep canonical correlation analysis," in *International Conference on Machine Learning*, 2013, vol. 28, pp. 1247–1255.

- Recognition (CVPR)*, 2016, pp. 5147–5156.
- [28] B. Van Merri et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *arXiv Prepr. arXiv1406.1078*, 2014.
- [29] O. Vinyals and Q. Le, “A neural conversational model,” *arXiv Prepr. arXiv1506.05869*, 2015.
- [30] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [31] D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” pp. 1–15, 2014.
- [32] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, no. 3, pp. 807–814.



قطب علمی سامانه های پردازش دیجیتال

کنفرانس پردازش سیگنال و



سیستم های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸

موقعیت یابی داخلی با استفاده از بلوتوث کم انرژی و رویکرد مبتنی بر شبکه عصبی

نقیسه دهقان سلماسی^۱، مرضیه قاضی اسداللهی^۲، رضا عزمی^۳، سید محمدحسین هاشمی نژاد^۴

^۱ دانشجو، دانشکده فنی گروه مهندسی کامپیوتر، دانشگاه الزهرا، تهران
n.dehghan@student.alzahra.ac.ir

^۲ دانشجو، دانشکده فنی گروه مهندسی کامپیوتر، دانشگاه الزهرا، تهران
m.ghazi@student.alzahra.ac.ir

^۳ دانشیار گروه مهندسی کامپیوتر، دانشگاه الزهرا، تهران
azmi@alzahra.ac.ir

^۴ استادیار گروه مهندسی کامپیوتر، دانشگاه الزهرا، تهران
SMH.hasheminejad@Alzahra.ac.ir

چکیده

ارتقاء دستگاه های بی سیم، ایجاد زیرساخت های مخابراتی گسترده برای ارتباطات بی سیم و گسترش شبکه های ماهواره ای، امکان ایجاد سیستم های تعیین موقعیت بسیاری را برای کاربران سیار به خصوص در فضاهای باز فراهم نموده است. متأسفانه بسیاری از این سیستم ها در فضای بسته عملکرد خوبی ندارند و راهکارهای دیگری برای این محیطها لازم است. در این پژوهش به بررسی نمونه ای از تکنیک ها و قواعد موقعیت یابی داخلی با استفاده از سیگنال های رادیویی پرداخته شده، و تلاش دارد با استفاده از یک تکنولوژی جدید به نام بلوتوث کم انرژی، به بهبود نتایج حاصل از این نوع موقعیت یابی بپردازد. پیشرفت و توسعه این تکنولوژی، وابسته به این است که چه میزان از سوی دستگاه های هوشمند امروزی از قبیل تلفن های همراه هوشمند، تبلت ها و... مورد پشتیبانی قرار گیرد. در ادامه با تحلیل اطلاعات جمع آوری شده از محیط های سرپوشیده، از طریق پیاده سازی یک شبکه عصبی، اطلاعات نهفته در این مجموعه استخراج شده و عملیات موقعیت یابی مبتنی بر یادگیری و هوشمندی محقق گردیده است.

کلمات کلیدی

موقعیت یابی داخلی، شبکه های عصبی، الگوریتم نزدیک ترین همسایه، نشانه گذاری در موقعیت یابی داخلی، سیگنال های رادیویی

۱- مقدمه

فروشی ها، حضور پیدا کند، می توان با دانستن مکان وی پیام های تبلیغاتی و یا اعلان هایی را برایش ارسال نمود. حتی می توان نظر کاربران، میزان و درصد علاقه آنها را نسبت به محصول و کالای خاص بدست آورد. به این صورت که اگر کاربر در مکانی بیشتر توقف داشته باشد و یا اگر از مکانی تردد بیشتری داشته باشد، می توان حدس زد که درصد علاقه آنها نسبت به آن محصول بیشتر است. از دیگر مزایا و

بسیاری از سرویس ها و خدماتی که به کاربران داده می شود بر اساس موقعیت و مکان آنها می باشد، به طور مثال اگر کاربری در موقعیتی که در آنجا محصولی وجود دارد مانند فروشگاه ها و خرده

گیرنده وجود دارد). اما متأسفانه این رویکردها نتایج ضعیفی در داخل ساختمان‌ها دارند. چراکه متغیرهای اندازه‌گیری تا حد بسیار زیادی متأثر از محیط اندازه‌گیری هستند. در نتیجه سیگنال‌های رادیویی توسط بسیاری از موانع از جمله دیوارها، ممکن است جذب، بازتاب و تضعیف شوند. از این رو تخمین فاصله و موقعیت، پیچیده و پیچیده‌تر می‌گردد. [۵]

در پژوهش [1]، [6] از رویکرد نشانه‌گذاری استفاده شده است. این تکنیک کاربران حاضر در محیط داخلی را در یک شبکه بی‌سیم فرض می‌کند. هر یک از بخش‌های محیط دارای داده‌های قابل ثبت و اندازه‌گیری است. بعد از جمع‌آوری این داده‌ها (عملیات نشانه‌گذاری محیطی) از نقاط نمونه و معین شده در فاز اول، لازم است یک معیار و روش برای تطبیق این نتایج تعیین شود تا موقعیت نقطه ناشناخته را در فاز دوم تخمین بزند. اغلب از فاصله اقلیدسی برای پیدا کردن اختلاف بین مقادیر اندازه‌گیری شده در فاز اول و مقادیر تخمین زده شده در فاز دوم استفاده می‌شود.

۲- سیستم موقعیت‌یابی داخلی

۲-۱- مفاهیم اولیه

بیکن^{۱۴}: با توجه به محدودیت‌هایی که در ویژگی‌های بلوتوث کلاسیک وجود دارد، رویکرد موقعیت‌یابی براساس این تکنولوژی کمتر مورد استفاده قرار می‌گیرد. به علت طولانی بودن فرایند جستجوی دستگاه‌های بلوتوثی، زمان مورد نیاز برای فراهم کردن تعداد مناسبی از این دستگاه‌ها، چندان به صرفه نیست. به علت مسائل مربوط به انرژی مصرفی و هزینه، استفاده از این تکنولوژی و زیرساخت‌های لازم برای آن عموماً در مقایسه با زیرساخت‌های مبتنی بر وای‌فای قرار می‌گیرد. که می‌تواند به فرایندها و نیازهایی غیر از موقعیت‌یابی نیز سرویس‌هایی ارائه کند.

این وضعیت با معرفی فناوری بلوتوث ۴ در سال ۲۰۱۰، که شامل بلوتوث کم انرژی یا BLE/Bluetooth Smart می‌شود، تغییر کرد به علت مصرف کم انرژی و داشتن گزینه‌های مختلف برای پیکربندی (مثل ایجاد فاصله زمانی در ارسال تبلیغات و یا تنظیم میزان انرژی برای ارسال پیام و سیگنال) استفاده از این تکنولوژی، نه تنها در مقایسه با دیگر سرویس‌های پیشین مبتنی بر بلوتوث، بلکه در مقایسه با کلیه خدمات مربوط به موقعیت‌یابی امروزی که توسط وای‌فای انجام می‌گیرد، امیدوار کننده است. [1]

هر بیکن بسته اطلاعات سنسوری مربوط به خود را (متشکل از مقادیر major، minor، UUID) ارسال و گوشی هوشمند از هر بیکنی که می‌بیند این سیگنال را دریافت می‌کند و با آنالیز این اطلاعات به محاسبه فاصله مربوطه‌اش با بیکن می‌پردازد. این در حالی است که کلیه اطلاعات (شناسه بیکن، مقدار RSSI و فاصله محاسبه شده) از طریق یک رابط کاربری^{۱۵} ساده برای کاربر نیز قابل مشاهده است. [7]

ارتباط بیکن با گوشی‌های هوشمند و دیگر بیکن‌ها از طریق داشتن

اطلاعاتی به شرح زیر صورت می‌گیرد: [1]

• Mac Address

کاربرد این سیستم‌های مبتنی بر تشخیص موقعیت، می‌توان به استفاده در فرودگاه‌ها، هتل‌ها، مکان‌های آموزشی، بیمارستان‌ها و مراکز درمانی، موزه‌ها، مکان‌های توریستی و... اشاره نمود.

امروزه درحالی که سیستم‌های هدایت ماهواره‌ای مانند GPS^۲، GLONASS^۳ رای عموم افراد در دسترس است، باید گفت مشکلی برای پیدا کردن یک فرد یا یک دستگاه تلفن همراه در محیط خارجی وجود ندارد. اما با وجود اینکه تکنولوژی GPS نقش بسیار مهمی را در سیستم‌های موقعیت‌یابی خارجی ایفا می‌کند، میزان قدرت سیگنال در محیط داخلی ساختمان‌ها به شدت افت کرده و موقعیت‌یابی در محیط‌های سرپوشیده با نداشتن هیچ‌گونه خط دید و دسترسی از سوی ماهواره‌های تعیین موقعیت، به دشوارترین حالت خود می‌رسد از طرفی GPS موقعیت‌یابی را با دقتی بیش از ۵ متر انجام می‌دهد و این در حالی است که برای موقعیت‌یابی دقیق در محیط داخلی دقتی در حدود چند سانتی‌متر لازم است. از این رو لازم است از دیگر ابزارهای حسگر برای بهبود دقت موقعیت‌یابی در محیط‌های داخلی استفاده شود.

[1][2]

برای موقعیت‌یابی با دقت بالا از تکنولوژی‌هایی نظیر وای‌فای و بلوتوث کم انرژی^۴ و همچنین از روش‌های مبتنی بر مجاورت بر اساس آیدی^۵ روش‌های مبتنی بر قدرت و زمان دریافت سیگنال‌های دریافت شده از ایستگاه‌ها و آنتن‌ها^۶ روش‌های مبتنی بر زاویه دریافت و ارسال سیگنال‌ها^۷، روش مثلث‌بندی^۸ و روش مبتنی بر نشانه‌گذاری^۹ استفاده شده است. در روش نشانه‌گذاری از الگوریتم‌های یادگیری ماشین استفاده می‌شود. [3]

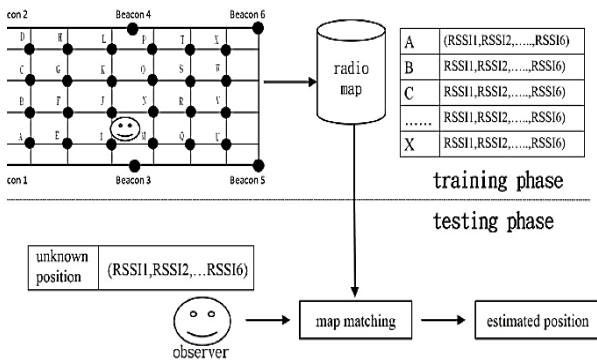
در این مقاله در صدد هستیم تا با بررسی مجموعه پژوهش‌های انجام شده در این حوزه، روشی برای ساخت سیستم موقعیت‌یابی داخلی با تکیه بر رویکرد نشانه‌گذاری و پیاده‌سازی روش‌هایی همچون نزدیکترین همسایه اقدام نماییم. در ادامه این روش را با بهره‌گیری از رویکردهای یادگیری ماشین و استفاده از شبکه‌های عصبی بهبود داده ایم تا بتوانیم به سرعت عملیات موقعیت‌یابی افزوده میزان خطای محاسباتی را به حداقل برسانیم. [2]

۱-۱- پژوهش‌های پیشین

در این بخش به طور مختصر به بیان روش‌هایی می‌پردازیم که در حیطه موقعیت‌یابی داخلی با استفاده از سیگنال‌های رادیویی انجام شده اند.

در این مقاله [1]، [4] متدهای مبتنی بر مثلث‌سازی به دو دسته تابشی^{۱۰} و زاویه‌ای^{۱۱} تقسیم شده است. به طور کلی این قبیل متدها، فاصله‌ای را تخمین می‌زنند که با قوی‌تر شدن سیگنال دریافتی از سوی فرستنده، میزان فاصله تخمین زده شده دقیق‌تر بیان می‌شود. ویژگی‌های مربوط به زمان انتشار سیگنال عبارتند از: زمان رسیدن سیگنال به گیرنده TOA^{۱۲}، اختلاف زمانی رسیدن سیگنال از فرستنده به گیرنده TDOA^{۱۳}، مسیر و جهت سیگنال دریافتی AOA (هنگامی که از آنتن‌های چند جهتی و یا مجموعه‌ای از آنتن‌ها استفاده می‌شود). تمامی متدهای ذکر شده، عملکرد خوبی در تعیین موقعیت در فضای باز دارند (فضایی که در آن خط دید انتشار مشخصی بین فرستنده و

جمع‌آوری می‌شوند. این مقادیر به همراه مختصات دقیق موقعیتی که جمع‌آوری شده‌اند در پایگاه داده مربوط به فاز اول نشانه‌گذاری ذخیره شده تا برای موقعیت‌یابی فاز بعدی، مورد استفاده قرار گیرد.



شکل ۱: فرآیند موقعیت‌یابی با روش نشانه‌گذاری [2]

در فاز دوم، موقعیت‌یابی انجام می‌شود. آن دستگامی که می‌بایست موقعیتش تعیین شود، مقادیر RSSI را از محیط دریافت و با متدی مشخص، با مجموعه داده جمع‌آوری شده در فاز پیش مقایسه می‌کند و موقعیت کاربر تخمین زده می‌شود. سپس می‌توان موقعیت تخمین زده شده توسط سیستم موقعیت‌یابی را با موقعیت حقیقی، مقایسه و نتیجه را به عنوان میزان دقت و کارایی روش مطرح نمود. [9][10]

بیشترین و پرکاربردترین الگوریتم‌ها و رویکردهای مقایسه‌ای و تخمین موقعیت استفاده شده در این مرحله عبارتند از [11][12]:

- روش‌های مبتنی بر احتمال (Probabilistic Methods)
- k-نزدیک ترین همسایه (K-Nearest Neighbors)
- شبکه‌های عصبی (Neural Networks)
- ماشین بردار پشتیبان (Support Vector Machine - SVM)
- کوچکترین چند ضلعی (Smallest M-Vertex Polygon)

۲-۲-۱- معماری و پیاده‌سازی سیستم

در پژوهش پیش رو به سراغ مقاله [2] رفته‌ایم در این مقاله پیاده‌سازی روش نشانه‌گذاری در یک محیط سرپوشیده به مساحت ۶ × ۴ مترمربع انجام شده است. برای فاز اول، ۶ بیکن مطابق شکل ۲ در محیط قرار داده شده است، در ۱۲ نقطه آموزشی مشخص شده به تعداد ۲۰ مرتبه سیگنال‌های دریافتی از بیکن‌ها جمع‌آوری و به همراه مختصات این نقاط در پایگاه‌داده ذخیره می‌گردد. با اتمام فاز اول و تکمیل مجموعه داده، تخمین مختصات ۶ موقعیت ناشناخته در محیط

• Universally Unique Identifier (UUID) - برای تمامی محصولات تولید یک شرکت یکسان است.

• Major Number - عددی که بیکن را به زیر مجموعه کوچکتری منتسب می‌کند.

• Minor Number - عددی که زیرمجموعه بیکن را به زیرمجموعه کوچکتری منتسب می‌کند.

ویژگی RSSI: همان‌طور که پیش‌تر گفته شد بسیاری از سرویس دهنده‌هایی که در محیط‌های داخلی^{۱۶} (IPS) خدماتی را ارائه می‌کنند، برای انجام متدهای موقعیت‌یابی از جمع‌آوری و بررسی ویژگی‌ای از سیگنال به نام میزان قدرت سیگنال دریافتی در گیرنده RSSI بهره می‌برند.

انتشار این سیگنال‌های غیرقابل پیش‌بینی در محیطی که به واسطه شرایط جغرافیایی با دیگر انواع سیگنال‌های نویزی و موانع فیزیکی احاطه شده است، توسعه سیستم ISP را با چالش‌هایی روبه‌رو می‌کند. اطلاعات انتشار یافته از طریق RSSI بیکن‌ها، به یک معیار ثابت از فاصله تبدیل می‌شود. این معیار بر اساس اختلاف بین قدرت سیگنال دریافتی و قدرت سیگنال مورد انتظار در فاصله ۱ متر، بدست می‌آید. اما همان‌طور که گفته شد، رسیدن به اطلاعات دقیق و بی‌نقص از طریق RSSI، به علت وجود عوامل ایجاد نویز، امری مشکل است. از طرفی، استفاده از قطعه سخت‌افزاری بیکن به تعداد زیاد، ریسک بروز خطا را بالاتر می‌برد. [7]

انتشار این سیگنال رادیویی در محیط از یک مدل لگاریتمی افت سیگنال پیروی می‌کند. رابطه بین قدرت سیگنال RSSI و فاصله بر اساس رابطه ۱ محاسبه می‌شود [8]:

$$RSSI(d) = RSSI(d_0) - 10n \log_{10}\left(\frac{d}{d_0}\right) \quad (1)$$

که در آن $RSSI(d)$ قدرت سیگنال دریافتی در فاصله d ، d_0 برابر فاصله ثابت است (معمولاً ۱ متر) و n بسته به محیط انتشار سیگنال، پارامتر افت سیگنال بر اثر انتشار در فضا^{۱۷} می‌باشد.

۲-۲- سیستم پیشنهادی

همان‌طور که گفته شد عملیات نشانه‌گذاری روشی مشتمل بر دو فاز است:

در فاز نخست یادگیری^{۱۸} انجام می‌شود. به این شکل که محیطی مجهز به تعداد مشخصی فرستنده و مولد سیگنال (بیکن) را ایجاد می‌شود. برای انجام عملیات موقعیت‌یابی لازم است در موقعیت‌های مشخص، که برای سهولت سعی می‌شود نقاط منظم در یک شبکه ماتریسی باشند، از پارامترهای مهم سیگنال که عمدتاً (و در این پژوهش) همان قدرت سیگنال دریافتی RSSI است، اندازه‌گیری به عمل آید. بدین شکل که در هر نقطه از فضای مختصاتی محیط مورد بررسی، به وسیله مقادیر RSSI که از هر بیکن در این محیط دریافت می‌کند، مشخص می‌گردد. مجموعه‌ای که در نهایت از نقاط و مقادیر سیگنال RSSI هر یک به دست می‌آید، مجموعه داده آموزشی سیستم خواهد بود. این داده‌ها توسط دستگاه‌های اندازه‌گیری و ثبت سیگنال (عموماً تلفن‌های همراه) در موقعیت‌هایی مشخص و معلوم،

¹ Retail Store

² Global Positioning System

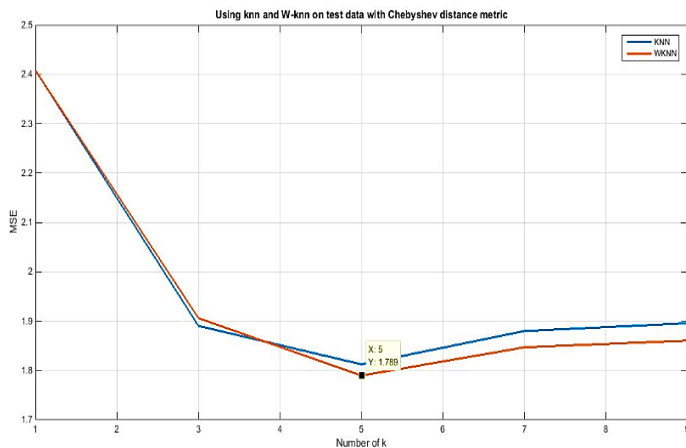
³ GLObal NAVigation Satellite System

- ¹² TOA: Time Of Arrival
- ¹³ Difference of Arrival TDOA: Time
- ¹⁴ beacon
- ¹⁵ Interface
- ¹⁶ Indoor Positioning Systems
- ¹⁷ Path loss
- ¹⁸ learning

- ⁴ Bluetooth Low Energy (BLE)
- ⁵ Cell identity (CID)-based methods
- ⁶ Range-based methods
- ⁷ Angle-based methods (AOA)
- ⁸ Trilateration
- ⁹ Fingerprinting-based methods
- ¹⁰ Lateration
- ¹¹ angulation

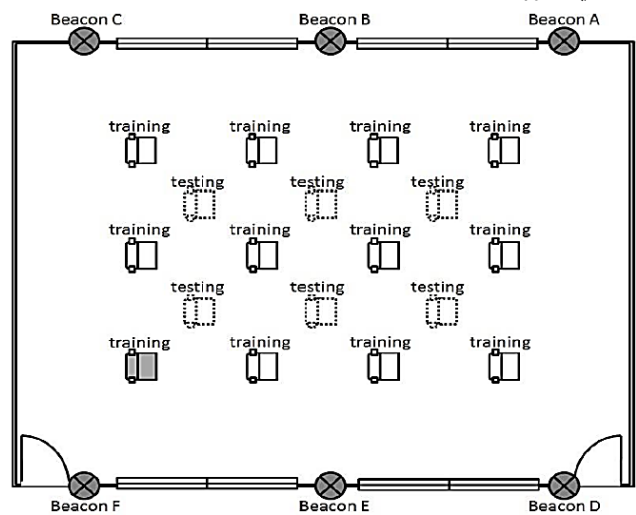
Metric+matchin	K=1	K=3	K=5	K=7	K=9
Euclidean+KNN	2.463	1.884	1.968	1.943	1.899
Chebyshev+KN	2.408	1.890	1.812	1.879	1.895
Euclidean+WKN	2.463	1.870	1.930	1.909	1.874
Chebyshev+WK	2.408	1.905	1.789	1.846	1.860
g					
N	0	1	3	8	6
NN	0	9	5	9	4

جدول ۱- میانگین خطای فاصله تخمین زده شده نسبت فاصله واقعی [2]



تاییدی است بر اینکه شاخص چبیشف تنها بیشترین اختلاف میان اعضای یک مجموعه را در نظر می‌گیرد و برای مجموعه داده‌ای که اغلب نقاط آن با یکدیگر اختلاف کوچکی دارند، ایده‌آل است. شرایط در نظر گرفته شده برای این پژوهش، فضای کلاسی با حداقل موانع محیطی است و همانطور که گفته شد تعداد بالایی (حدوداً ۲۰ مقدار) در هر نقطه آموزشی نمونه‌گیری انجام می‌شود. با چشم پوشی از این شرایط ایزوله و ایده‌آل، پیاده‌سازی روش k - نزدیک‌ترین همسایه که از جمله روش‌های کند بوده و نیاز به صرف زمان زیاد برای تعیین دقیق پارامترهای آن (مانند مقدار مناسب k و نوع شاخص فاصله مورد استفاده) دارد، از این رو لازم است تا در جستجوی روش سنجش‌ای با کارایی و سرعت بالاتر باشیم تا ضمن حفظ دقت مطلوب موقعیت‌یابی

با بکارگیری روش k - نزدیک‌ترین همسایه (KNN) و نوع وزن‌دار این روش (WKNN) انجام می‌شود. مقادیر خطای حاصل از پیاده‌سازی این دو روش با استفاده از دو شاخص اندازه‌گیری فاصله اقلیدسی و چبیشف و با تعیین مقادیر مختلف برای k ، در شکل ۳، ۴ و همچنین مطابق جدول ۱ به تفکیک آورده شده است.



شکل ۲- محیط و شرایط پیاده‌سازی تکنیک نشانه‌گذاری [2]

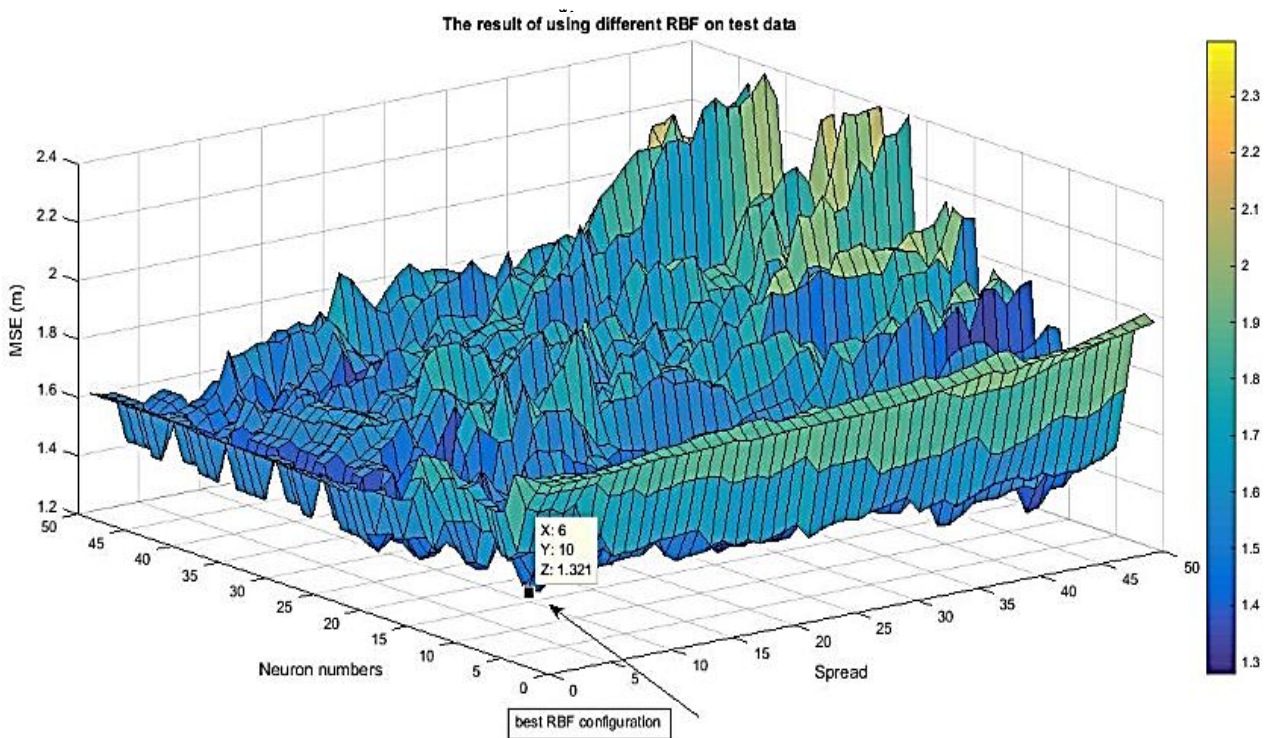
۲-۲-۲- ارزیابی سیستم

همان‌طور نتیجه می‌شود، کمترین اختلاف بین مقدار حقیقی و تخمین زده شده در حدود ۱,۷۸۹,۵۴ است که با تعیین $k=5$ و استفاده از روش wknn با شاخص فاصله چبیشف بدست می‌آید. این امر

آموزشی، در آزمایشات مربوط به تنظیم بهینه پارامترها یک نمونه از کل نقاط موجود در نقشه رادیویی (نقاط اندازه گیری شده) کنار گذاشته شده و مجموعه آموزشی با بقیه نقاط ساخته می‌شود. این مجموعه به شبکه RBF داده شده و شبکه فوق با در نظر گرفتن سیگنال‌های اندازه گیری شده آن به عنوان ورودی و بردار موقعیت متناظر به عنوان خروجی، به آموزش خود مطابق با روال ذکر شده می‌پردازد. سپس نمونه کنار گذاشته شده مورد سنجش قرار می‌گیرد. میزان خطا برابر با فاصله اقلیدسی خروجی شبکه (موقعیت محاسبه شده برای نمونه فوق) با مقدار واقعی موقعیت مورد سنجش است. این روال برای تک تک نمونه‌ها ادامه می‌یابد و در نهایت خطای کلی با میانگین‌گیری از مقدار خطاهای بدست آمده برای تمام نمونه‌ها حاصل می‌شود.

شکل ۳- نموداری مقایسه تغییرات خطا بر حسب تعداد k در دو روش KNN و WKNN با شاخص فاصله چیبشف

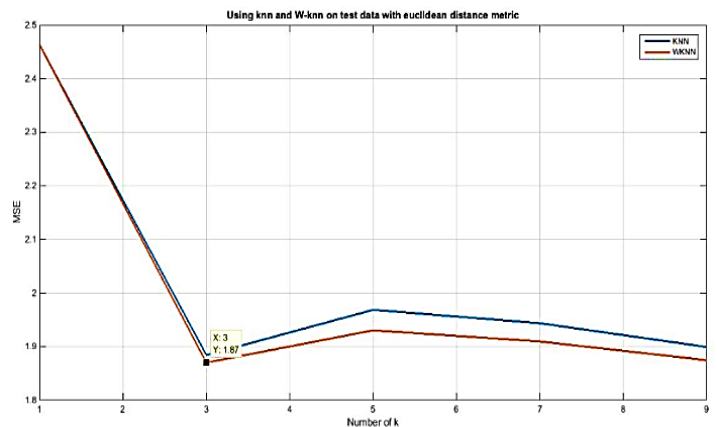
در فضای داخلی، برای پیاده‌سازی در شرایط واقعی‌تر نیز مناسب باشد. یکی از بهترین و کامل‌ترین روش‌های سنجشی که می‌توان در فاز دوم موقعیت‌یابی استفاده کرد، شبکه‌های عصبی هستند. شبکه‌های عصبی مبتنی بر توابع شعاعی RBF^۳، به علت سادگی پیاده‌سازی و زمان یادگیری کوتاه و سرعت در اجرا، از محبوبیت بالایی برخوردارند. [13] اما از آنجایی که به طور کلی رفتار خروجی هر شبکه عصبی تا حد زیادی بستگی به چگونگی تنظیم پارامترهای تعیین‌کننده آن دارد، لذا در وهله اول انتخاب پارامترهای بهینه باید در دستور کار قرار گیرد. معیار مورد بررسی برای عملکرد بهینه شبکه فوق، تنظیم مناسب پارامترها برای اندازه‌گیری خطا یا همان میزان فاصله اقلیدسی بین خروجی این شبکه و خروجی واقعی است. برای ساخت مجموعه



شکل ۵: نمودار میزان خطا بر حسب ترکیب بندی‌های مختلف شبکه عصبی RBF [2]

شکل ۴- نموداری مقایسه تغییرات خطا بر حسب تعداد k در دو روش KNN و WKNN با شاخص فاصله اقلیدسی

با بهره‌گیری از مجموعه داده جمع‌آوری در پژوهش [2]، پس از پیاده‌سازی روش k -نزدیک همسایه و دریافت نتایج یکسان با نتایج و خطای عنوان شده در آن پژوهش، از شبکه عصبی RBF به عنوان روش سنجش جایگزین برای تخمین موقعیت‌های ناشناخته استفاده نمودیم. با تنظیم پارامترهای spread و تعداد نورن‌ها (Neuron Number = ۱۰، Spread = ۶) مطابق شکل ۵ شبکه را به مدلی ایده‌آل رسانده و به دقت اندازه‌گیری یا میانگین



- [۵] راستگو چلمردی، محسن، عبداللہی، سعید، "بررسی روش‌ها و الگوریتم‌های موقعیت‌یابی برای عابر پیاده" نشریه علمی-ترویجی مهندسی نقشه‌برداری و اطلاعات مکانی، دوره نهم.
- R. Faragher and R. Harle, "Location Fingerprinting With Bluetooth Low Energy Beacons," IEEE J. Sel. Areas Commun., vol. 33, no. 11, pp. 2418–2428, Nov. 2015.
- A. Mackey and P. Spachos, "Performance evaluation of beacons for indoor localization in smart buildings," in 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2017, pp. 823–827.
- G. Li, E. Geng, Z. Ye, Y. Xu, J. Lin, and Y. Pang, "Indoor Positioning Algorithm Based on the Improved RSSI Distance Model," Sensors, vol. 18, no. 9, p. 2820, Aug. 2018.
- [۹] احسانی چیمه، حمیدرضا، کریمی، مصطفی، "سیستم موقعیت‌یابی مکانهای داخلی با استفاده از روش فینگرپرینت در سیستم‌های وای‌فای"، سومین کنفرانس بین‌المللی مهندسی برق.
- N. Vo, T. Ngo, and J. G. Kim, "Sequential learning for fingerprint based indoor localization Sequential learning for fingerprint based indoor localization," AEUE - Int. J. Electron. Commun., 2016.
- C. Sekhara and R. Annavarapu, "Machine learning algorithms for wireless sensor networks: A survey," Inf. Fusion, 2018.
- R. Hassan, S. M. Haque, and M. Imtiaz, "A novel cascaded deep neural network for analyzing smart phone data for indoor localization," Futur. Gener. Comput. Syst., vol. 101, pp. 760–769, 2019.
- M. Stella, M. Russo, and M. Saric, "RBF Network Design for Indoor Positioning Based on WLAN and GSM." 2014.
- "GitHub - ycpu/IPS: Indoor Positioning System Based on BLE Location Fingerprinting with Classification Approach." [Online]. Available: <https://github.com/ycpu/IPS>. [Accessed: 19-Aug-2019].

مربعات خطای^۱ برابر $MSE=1/321$ دست پیدا کردیم. این مقدار خطای شبکه در مقایسه با حداقل خطای گزارش شده در [2] دقیق‌تر و با صرف زمان کوتاه‌تری حاصل شده است. از این رو برای استفاده در عملیات موقعیت‌یابی در فضایی واقعی‌تر و غیر ایزوله شده مناسب‌تر است.

۳- نتیجه

تخمین موقعیت، یک تکنیک مهم برای سرویس‌های مبتنی بر تحلیل موقعیت است. امروزه موقعیت‌یابی در فضای خارجی با استفاده از زیرساخت ابزارهایی مانند GPS به سمت موقعیت‌یابی داخلی و در فضاهای بسته، سوق داده شده است. با وجود گستردگی وسیع مباحث قابل بررسی در این حوزه، باید گفت همچنان تکنیک‌های ارائه شده از فقدان سطح قابل قبولی از دقت و انعطاف‌پذیری برای اصلاح و توسعه برای کاربردهای تجاری موجود، رنج می‌برند. می‌توان علت اصلی این کاستی را به عدم توجه کافی و همه‌جانبه در انتخاب رویکردهای مناسب دانست. روش‌هایی که برای کاوش انبوه اطلاعات جمع‌آوری شده از یک فضای سرپوشیده، از بازدهی و قدرت آزمون کافی برخوردار باشند. در این پژوهش، تلاش شد تا ضمن بررسی یکی از روش‌های پیاده‌سازی شده (روش $wknn$ و knn) به منظور موقعیت‌یابی در فضای داخلی کلاس، از نقاط قوت این رویکرد استفاده و برای برطرف کردن کاستی‌ها و نارسایی‌های آن (دقت کمتر، سرعت و قابلیت تعمیم ضعیف‌تر، عدم توانایی کنترل حجم بالایی از داده‌ها و...) یک شبکه عصبی RBF با پیکربندی بهینه جایگزین شود. در ادامه با مقایسه خطای حاصل از موقعیت‌یابی با شبکه عصبی و خطای روش knn دریافتیم که شبکه عصبی با صرف حداقل زمان و با سرعت بالاتر قادر است با دقت بالاتری عملیات موقعیت‌یابی را در فضای مورد بحث انجام دهد.

آنچه در ادامه نیاز به بررسی عمیق‌تر دارد، مطالعه بر روی ارائه راهکاری است تا با صرف حداقل زمان بتواند برای هر موقعیت و فضای ارائه شده، یک پیکربندی بهینه برای شبکه عصبی یافته و متناسب با شرایط و موانع فیزیکی موجود در محیط عملیات آزمون و موقعیت‌یابی را پیش ببرد.

مراجع

- P. Kriz, F. Maly, and T. Kozel, "improving indoor localization using bluetooth low energy beacons," vol. 2016, 2016.
- Y. C. Pu and P. C. You, "Indoor positioning system based on BLE location fingerprinting with classification approach," Appl. Math. Model., vol. 62, pp. 654–663, 2018.
- P. Viswanath and M. Mathew, "Recent advances in indoor localization: A survey on theoretical approaches and," vol. 19, no. 2, pp. 1–9, " Ieeexplore.Ieee.Orgapplications 2017.
- F. Zafari, I. Papapanagiotou, M. Devetsikiotis, and T. Hacker, "An iBeacon based Proximity and Indoor Localization System," pp. 1–14, 2017.

¹ Mean Square Error - MSE



قطب علمی سامانه های پردازش دیجیتال

کنفرانس پردازش سیگنال و



سیستم های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸

بهبود آشکارسازی ناهنجاری در تصاویر ابرطیفی با بیان تنک و فیلتر گوسین

مرضیه بابائی^۱، مریم ایمانی^۲

^۱ کارشناسی ارشد، گروه مهندسی برق، دانشکده فنی مهندسی، دانشگاه خاتم، تهران

^۲ استاد یار، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران
maryam.imani@modares.ac.ir

چکیده

در این مقاله، روشی جهت بهبود آشکارسازی ناهنجاری در تصاویر ابرطیفی بر اساس بیان تنک و یک فیلتر نرم کننده پیشنهاد شده است. با استفاده از بیان تنک تصویر، تفاوت بین اهداف ناهنجار موجود در صحنه با پیش زمینه تصویر بهتر آشکار می شود. از آن جایی که اهداف ناهنجار، نادر بوده و درصد کمی از کل تصویر را تشکیل می دهند و از طرفی دارای طیف متفاوت نسبت به پیش زمینه هستند، به پیکسل های نویزی شبیه هستند. استفاده از فیلتر نرم کننده گوسین و محو کردن نویزهایی که مورد علاقه نیستند، قبل از به کارگیری بیان تنک، کمک شایانی به کاهش نرخ خطای آشکارساز خواهد کرد. مقایسه نتایج به دست آمده روش پیشنهادی با الگوریتمی بر مبنای پایه تنک به نام BJSR حاکی از بهبود آشکارسازی با میزان ۲/۲ درصد می باشد.

کلمات کلیدی

آشکارسازی ناهنجاری، بیان تنک، فیلتر گوسین، تصاویر ابرطیف

موجود نباشد به عنوان اهداف نظارت نشده (ناهنجاری) شناخته می شوند. در سال های اخیر افزایش تقاضا برای استفاده از این تصاویر با کاربرد آشکارسازی ناهنجاری در علوم مختلف از جمله کشاورزی و صنعت باعث توجه بیشتر دانشمندان به ارائه روش ها و الگوریتم های کاربردی تری در این زمینه گردیده است.

سال ۱۹۹۰ اولین الگوریتم آشکارسازی ناهنجاری موسوم به RX^1 توسط رید و یو معرفی گردید [۱]. این روش با محاسبه فاصله مایلانوبیس بین پیکسل تست و پس زمینه به آشکارسازی می پردازد:

(۱)

$$\delta_{RX}(r) = (r - \mu)^T C_{L \times L}^{-1} (r - \mu)$$

در این رابطه r پیکسل ورودی، μ و c به ترتیب بردار میانگین و ماتریس کواریانس پس زمینه و L تعداد باندهای مورد استفاده در تصویر

۱- مقدمه

تصاویر ابرطیفی بدست آمده از ثبت دقیق انرژی الکترومغناطیسی بازتابی از سطح اشیاء، داده هایی با ویژگی های طیفی متفاوت هستند که شامل اطلاعات بسیار زیادی از ناحیه تصویربرداری شده می باشند. از ویژگی های طیفی این داده ها می توان برای شناسایی و تفکیک مواد تشکیل دهنده تصویر استفاده کرد. کاربردهای بسیار زیادی تاکنون برای پردازش و استخراج اطلاعات از این داده ها پیشنهاد شده است. یکی از این کاربردهای موسوم به طبقه بندی، پیکسل هایی از تصویر را به عنوان اهداف مورد آزمایش در نظر می گیرد، چنانچه هدف مورد نظر دارای اطلاعات از پیش تعریف شده و در دسترس باشد به عنوان اهداف نظارت شده و اگر از اهداف مورد آزمایش هیچ گونه اطلاعات طیفی

¹ Reed-Xiaoli

می‌باشد. در صورتیکه δ_{RX} از یک حد آستانه مشخص، بزرگتر باشد پیکسل مورد بررسی ناهنجاری خواهد بود. الگوریتم RX به دو نوع کلی و محلی تقسیم می‌شود [۲،۳]. در الگوریتم‌های نوع اول، از کل داده‌های موجود در تصویر و تمامی باندهای طیفی به منظور محاسبه میانگین و ماتریس کواریانس نمونه استفاده می‌گردد، اما در الگوریتم‌های محلی ابتدا یک پنجره با ابعاد مشخص به مرکز پیکسل تحت آزمایش تعریف شده و با حرکت آن در تمام طول تصویر، میانگین نمونه را با استفاده از داده‌های درون پنجره محاسبه و در فرمول اصلی قرار می‌دهد.

$$S \in R^B \text{ تعریف می‌شود که در آن } B \text{ تعداد باندهای تصویربرداری است. در همین راستا دیکشنری مورد نظر و ضرایب } \alpha \text{ نیز شامل } B \text{ باند خواهند بود که حجم محاسبات را برای حل معادله شماره (۲) بسیار پیچیده و دشوار خواهد کرد. یکی از راه حل‌های ارائه شده برای کاهش حجم این محاسبات تقسیم کردن تصویر ابرطیف به زیرفضاهای کوچکتر است [۱۰]. اساس عملکرد این تقسیم بندی نیز بر پایه شباهت پیکسل‌های موجود در یک زیر فضا می‌باشد. بنابراین معادله بیان تنک در یک تصویر ابرطیفی را می‌توان به صورت رابطه زیر نمایش داد:}$$

$$S = [s_1, s_2, \dots, s_n] = [D\alpha_1 + \varepsilon_1, D\alpha_2 + \varepsilon_2, \dots, D\alpha_n + \varepsilon_n] = D[\alpha_1, \alpha_2, \dots, \alpha_n] + \varepsilon = D\Psi + \varepsilon \quad (۳)$$

بعد از سال ۱۹۹۰ تعداد زیادی الگوریتم در این راستا ارائه شدند، از جمله، در سال ۲۰۰۲ الگوریتم شبه آنی CRX^A توسط چنگ ارائه شد، که بر اساس RX کار می‌کرد ولی به جای ماتریس کواریانس نمونه طیفی از یک ماتریس همبستگی نمونه استفاده می‌کرد [۴]. سال ۲۰۱۷ مریم ایمانی روشی براساس معیارهای تمرکز میانگین و میانه را ارائه دادند که بهبود فزاینده‌ای در روش RX بوجود آورد [۵]. سال ۲۰۱۹ لینگ، لین و آن روشی بر پایه بیان تنک ارائه کردند که بر اساس آن هر پیکسل پس‌زمینه می‌تواند به صورت یک ترکیب خطی از پیکسل‌های همسایگی خود بیان شود، اما پیکسل‌های ناهنجار به دلیل اینکه در پس‌زمینه به صورت محدود وجود دارند هیچ فضای همسایگی نخواهند داشت در نتیجه نمی‌توان آن‌ها را به صورت ترکیب خطی با فضای همسایگی خود بیان کرد [۶].

$$\Psi = \operatorname{argmin} \|D\Psi - S\| \text{ s.t. } \|\Psi\|_0 \leq K_0 \quad (۴)$$

در رابطه (۴)، $\|\Psi\|_0$ تعداد سطرهای غیر صفر Ψ است که با حل معادله بالا در نهایت، Ψ یک ماتریس تنک با L سطر غیر صفر خواهد بود. برای حل معادله (۴) روش‌های مختلفی ارائه شده که برای مطالعه بیشتر می‌توان به مقاله‌های [۱۱،۱۲] مراجعه کرد.

۲-۱- بیان تنک در آشکارسازی ناهنجاری

همانگونه که بیان شد با استفاده از زیرفضاها و محدود کردن دیکشنری در تصاویر ابرطیفی از محاسبات پیچیده پردازش تنک کاسته می‌شود. چنانچه از پردازش تنک برای آشکارسازی ناهنجاری‌ها در یک تصویر ابرطیفی استفاده گردد، می‌بایست اختلاف طیفی تنک پیکسل‌ها با پیکسل‌های پس‌زمینه محاسبه شود. در همین راستا می‌توان با استفاده از پنجره دوگانه اطراف پیکسل تست و تخمین پس‌زمینه به صورت محلی و در قالب تعداد محدودی از زیرفضاهای استخراج شده از داده ابرطیفی، به آشکارسازی ناهنجاری در این تصاویر پرداخت. [۱۳]

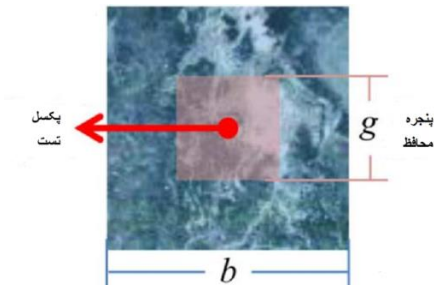
این روش که به $BJSR^2$ معروف است با استفاده از پنجره‌های دوگانه اطراف پیکسل تست می‌تواند به پردازش دقیق پیکسل‌هایی که بسیار نزدیک به هم و تقریباً در یک طیف و فرکانس قرار دارند بپردازد. براساس این فرضیه همانگونه که در شکل شماره (۱) نمایش داده شده است، پیکسل تست به رنگ قرمز و پنجره‌ی اطراف آن با رنگ نارنجی روشن با سایز g^2 قرار گرفته، این پنجره در پیشگیری از ترکیب پیکسل تست مورد بررسی با پیکسل‌های همسایگی آن مؤثر است. علاوه بر آن در خارج از این پنجره نیز یک پنجره همسایگی محلی دیگر با رنگ آبی با سایز b^2 خواهیم داشت. اساس عملکرد پنجره‌ی دوگانه در یک زیرفضا به این ترتیب خواهد بود که باید تمام حالت‌های احتمالی در مورد پیکسل‌های موجود در این پنجره مورد بررسی قرار گیرد:

۲- روش بیان تنک

یکی از دقیق‌ترین روش‌های تشخیص ناهنجاری و تشخیص اهداف که در سال‌های اخیر بسیار مورد توجه دانشمندان قرار گرفته، استفاده از بیان تنک یک تصویر است [۷،۸]. با توجه به اینکه سیگنال‌ها معمولاً در حوزه‌های متفاوت دارای نمایش تنک هستند می‌توان از بین تعداد انبوهی سیگنال پایه در یک حوزه از کمترین تعداد ضرایب برای نمایش سیگنال مورد نظر استفاده کرد. در الگوریتم بیان تنک به هر سیگنال پایه، اتم و به مجموعه سیگنال‌های پایه دیکشنری گفته می‌شود. در این روش سیگنال بسیار بزرگی به طول M را می‌توان تنها با N ضریب به طوریکه $N \ll M$ است نمایش داد [۹]. برای رسیدن به این منظور ابتدا باید دستگاه معادله خطی $X=D\alpha$ را تشکیل داد که در این رابطه بردار X بوسیله دیکشنری تشکیل شده از اتم‌های پایه D و تنک‌ترین ضریب α بازسازی می‌شود، اما از آنجاییکه این معادله دارای بیشمار جواب خواهد بود، برای حل آن و یافتن تنک‌ترین ضریب بازسازی نیاز است که معادله $X=D\alpha$ را به فرم معادله بهینه‌سازی (۲) تبدیل کنیم:

$$\alpha = \operatorname{argmin} \|D\alpha - S\| \text{ s.t. } \|\alpha\|_0 \leq K_0 \quad (۲)$$

حل معادله شماره (۲) بردار ضرایبی را در اختیار قرار می‌دهد که دارای تنک‌ترین ضریب هستند. در یک تصویر ابرطیفی که شامل صدها باند طیفی از ناحیه مورد نظر می‌باشد، بیان تنک محاسبات بسیار پیچیده‌تری خواهد داشت به گونه‌ای که هر سیگنال از تصویر به صورت



شکل ۱: پنجره دوگانه اطراف پیکسل تست مرکزی [۱۳]

مشکل این روش این است که مجموعه پیکسل همسایگی S ممکن است ترکیبی از چندین امضای طیفی باشد که تشخیص هنجار یا ناهنجار بودن پیکسل تست را مشکل می‌کند. روش بهتر استفاده از الگوریتمی است که می‌تواند به طور انحصاری اطلاعات پس‌زمینه محلی پیکسل تست را تخمین زند به این صورت که ابتدا یک دیکشنری فوق کامل از تصویر ابرطیفی به صورت بیان خطی مجموعه‌ی پیکسل‌های همسایه محلی $S = [s_1, s_2 \dots s_n] \in R^{B \times n}$ در نظر می‌گیرد، سپس پیکسل‌های کل تصویر ابرطیفی را به صورت یک ماتریس $H \in R^{B \times H}$ پشته می‌کند، (H تعداد پیکسل‌ها و B تعداد نوارهای طیفی است). اگر فرض کنیم C_{b_n} عناصر پس‌زمینه باشد و C_{t_n} اهداف ناهنجار، ستون‌های ماتریس H را می‌توان به صورت معادله (۸) نشان داد:

$$H = [BC_{b_1} \dots BC_{b_n} TC_{t_1} \dots TC_{t_n}] \quad (۸)$$

در این رابطه BC_{b_i} شامل N_{b_i} پیکسل پس‌زمینه متعلق به دسته‌ی N_{t_i} پیکسل هدف ناهنجار متعلق به دسته‌ی C_{t_i} ($t_i = 1 \dots t_n$) خواهد بود.

(۷)

$$H = \sum_{b_i=b_1}^{b_n} N_{b_i} + \sum_{t_i}^{t_n} N_{t_i}$$

به این معنا

که $C_{b_n} + C_{t_n}$ یک زیرفضا با بعد پائین همراه با انواع عناصر موجود در پوشش پس‌زمینه است که در فضای خیلی بزرگتر توسط H اندازه‌گیری می‌شود. برای بررسی حالت‌هایی که پیکسل‌های ناهنجاری اطراف همسایگی محلی S بوجود می‌آورند، نیاز به برآورد دقیق پس‌زمینه خواهیم داشت که با استفاده از بیان تنک می‌توان تخمین دقیقی از پس‌زمینه را بدست آورد. برای یک پیکسل تست فرض کنید یک ستون از S به نام S_j ($j = 1 \dots n$) متعلق به کلاس C_{b_i} باشد در این حالت می‌توان این ستون را به صورت یک ترکیب خطی از پیکسل‌های BC_{b_i} در H بیان کرد:

$$S_j = BC_{b_i} \beta_{C_{b_i,j}} + \varepsilon_j \quad (۸)$$

در رابطه (۸)، $\beta_{C_{b_i,j}}$ یک بردار ضرایب ناشناخته از N_{b_i} خواهد بود، ε_j نیز نویز تصادفی می‌باشد. بهتر است قبل از قضاوت برای تعیین

۱. حالت اول: تمام پیکسل‌های موجود در این مجموعه پیکسل همسایگی، پس‌زمینه هستند و چندین نوع پوشش زمینی را در بر می‌گیرند.

۲. حالت دوم: اکثر پیکسل‌های موجود در این پنجره به جز یک یا دو پیکسل ناهنجار، جزو پس‌زمینه هستند.

۳. حالت سوم: اکثریت و یا تمام پیکسل‌های موجود ناهنجار خواهند بود. که البته این حالت به دلیل احتمال پایین وجود ناهنجاری در واقعیت اتفاق نمی‌افتد، بنابراین تمرکز اصلی روی دو حالت اول خواهد بود.

بعد از مشخص کردن حالت‌های موجود در پنجره‌ی دوگانه به بیان آشکارسازی ناهنجاری با استفاده از پردازش تنک می‌توان پرداخت. برای پیکسل تست همانگونه که بیان شد ابتدا پنجره‌ی دوگانه اطراف آن ساخته می‌شود، سپس پیکسل تست به صورت ماتریس $S_c \in R^B$ بیان می‌گردد که در این رابطه B بیانگر تعداد باندها خواهد بود. همچنین پیکسل‌های همسایه پیکسل تست مرکزی به صورت مجموعه‌ای برابر با $S = [s_1, s_2 \dots s_n] \in R^{B \times n}$ در نظر گرفته می‌شوند. که n بیانگر تعداد پیکسل‌های موجود است. برای مشخص کردن اینکه پیکسل مورد نظر ناهنجار است یا نه، از آزمون فرضیه زیر استفاده می‌گردد:

$$\begin{aligned} H_0 : S_c &= Da_0 + e && \text{ناهنجار نیست} \\ H_1 : S_c &= Da_0 + t + e && \text{ناهنجار است} \end{aligned}$$

در این دو رابطه D در واقع ماتریس پایه (دیکشنری) پس‌زمینه برای پیکسل تست S_c خواهد بود و e نویز تصادفی با سطح انرژی پایین در نظر گرفته می‌شود. زمانیکه حالت H_1 اتفاق بیفتد به معنای این است که S_c (پیکسل تست مرکزی) دارای طیف ناهنجار می‌باشد بنابراین فرضیه H_1 را می‌توان به صورت رابطه (۵) بازنویسی کرد:

(۵)

$$S_c = D\alpha_0 + m$$

که m ترکیبی از نویز و طیف ناهنجار خواهد بود و بخش اعظمی از انرژی پیکسل تست مرکزی را اشغال کرده است. با توجه به رابطه‌ی (۵) برای بدست آوردن S_c :

ابتدا نیاز به یک دیکشنری فوق کامل داریم که شامل تمام اطلاعات موجود باشد. همچنین نیاز به یک معیار ارزیابی قوی برای قضاوت در مورد هنجار یا ناهنجار بودن پیکسل خواهیم داشت.

ساده‌ترین حالت این است که پیکسل تست را با پیکسل‌های اطراف آن در پنجره‌ی بیرونی مقایسه کنیم که در این صورت زمانی پیش‌بینی موفقیت‌آمیزی خواهیم داشت که تنها اطلاعات پس‌زمینه به هنجار یا ناهنجار بودن پیکسل تست کمک کند.

$$s_c = Aa + \varepsilon_c \quad (13)$$

که در این حالت انرژی سیگنال برابر است با :

$$\varepsilon_c = s_c - Aa = P \frac{1}{B} s_c \quad (14)$$

اما اگر پیکسل تست متعلق به پس‌زمینه نباشد و ناهنجار محسوب گردد، انرژی محاسبه شده‌ی پیکسل تست بسیار بزرگتر از میانگین انرژی Σ خواهد بود.

در حالت سوم اگر پیکسل تست ترکیبی از پس‌زمینه و ناهنجاری باشد، به دلیل اینکه در رابطه‌ی $s_c = t + n$ ، انرژی t بر کل پیکسل تسلط دارد، پیکسل تست ناهنجار محسوب می‌گردد و برای بدست آوردن انرژی آن از نسبت سیگنال به نویز استفاده می‌شود:

$$d = \frac{\|P \frac{1}{B} s_c\|_2^2}{\sum \frac{\|P \frac{1}{B} s_i\|_2^2}{n}} = \frac{s_c^T P \frac{1}{B} s_c}{\text{mean}(s_i^T P \frac{1}{B} s_i)} \cdot i = 1 \dots n \quad (15)$$

در رابطه (۱۵)، s_i پیکسل i ام از پنجره دوگانه خواهد بود و n تعداد پیکسل‌های موجود در ناحیه همسایگی می‌باشد.

۳- نرم کردن تصویر با استفاده از فیلتر گوسین

یکی از مهم‌ترین پردازش‌ها در تصاویر ابرطیفی حذف نویز یا اهدافی است که نیازی به حضور آنها در تصویر نیست، یکی از روش‌های دستیابی به این امر استفاده از فرایند فیلتر کردن در پردازش این تصاویر است.

فیلتر، فرآیندی است که برخی از اجزا یا ویژگی‌های ناخواسته را از سیگنال اصلی حذف می‌کند. ویژگی مشخص‌کننده فیلترها از بین بردن تمام یا بخشی از جنبه‌های سیگنال است. در واقع، این به معنای حذف برخی از فرکانس‌ها یا باندهای فرکانسی می‌باشد. با این حال، فیلترها تنها در دامنه فرکانس عمل نمی‌کنند، به ویژه در زمینه پردازش تصویر، که بسیاری از اهداف دیگر برای فیلتر کردن مد نظر است. تا به امروز فیلترهای بسیاری برای حذف نویزهای احتمالی در تصویر پیشنهاد شده، از جمله فیلتر گوسین که یک فیلتر میانگین وزن‌دار می‌باشد و از تابع گوسی اقتباس می‌شود. در ماسک فیلتر گوسی، بیشترین ارزش به پیکسل اصلی (مرکزی) داده می‌شود و پیکسل‌های همسایه، وزنی متناسب با فاصله‌شان تا پیکسل مرکزی به خود اختصاص می‌دهند. هرچه فاصله دورتر می‌شود مقدار وزن نیز کوچک‌تر می‌گردد. این نکته‌ای است که باعث می‌شود لبه‌ها و مرزها بهتر حفظ گردند.

این فیلتر بر اساس رابطه (۱۶) عمل می‌کند [۱۶]:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (16)$$

که σ یک عدد حقیقی و مثبت است. این تابع یک تابع پیوسته می‌باشد و در همه جا مثبت است تابع گوسی ماکزیمم خود را در $(0,0)$ اختیار

هنجار یا ناهنجار بودن پیکسل S_j از ماتریس H ، H را به عنوان یک دیکشنری فوق کامل برای بدست آوردن بردار ضرایب تنک b به کار ببریم. در همین راستا با توجه به دیکشنری مورد نظر مساله بهینه‌سازی توسط فرمول (۹) بازسازی می‌گردد:

$$B_j = \text{argmin} \|HB_j - S_j\| \quad s.t. \quad \|B_j\|_0 \leq L \quad (9)$$

که L سطح تنک خواهد بود. در نتیجه با استفاده از فرمول شماره (۹) می‌توان پیکسل تست را توسط دیکشنری فوق کاملی که از کل تصویر بدست آمده تخمین زد. مشابه پیکسل S_j ، تمام پیکسل‌های دیگر نیز که در پنجره‌ی همسایگی محلی قرار دارند، می‌توانند توسط دیکشنری فوق کامل H به صورت تنک بیان گردند:

$$S = [s_1 \dots s_n] = H[\alpha_1 \dots \alpha_n] + [\varepsilon_1 \dots \varepsilon_n] \quad (10)$$

از آنجایی که پنجره‌ی همسایگی محلی تنها شامل یک یا چند نوع از عناصر پس‌زمینه است، فرض بر این خواهد بود که تمام پیکسل‌ها در پنجره‌ی همسایگی محلی می‌توانند توسط یک زیر فضا با بعد پایین به صورت یک ترکیب خطی بیان گردند. در واقع این زیر فضا متعلق به دیکشنری پس‌زمینه یا ترکیبی از زیرفضای متعلق به دیکشنری پس‌زمینه خواهد بود. بنابراین مجموعه پیکسل همسایگی S را که ابتدا بیان شد می‌توان به صورت فرمول (۱۱) بازسازی کرد:

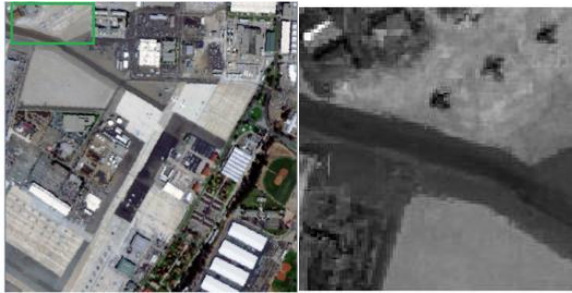
$$\Psi = \text{argmin} \|H\Psi - S_f\| \quad s.t. \quad \|\Psi\|_0 \leq l \quad (11)$$

در معادله (۱۱)، Ψ مجموعه‌ای از تمام بردارهای ضرایب تنک است.

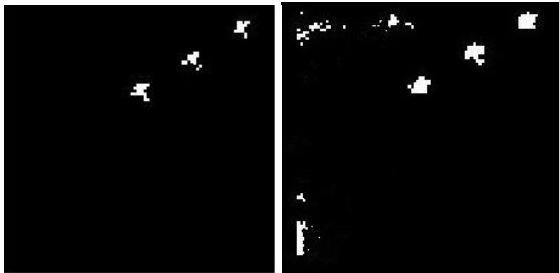
L بالاترین باندی است که بیشترین ضرایب غیر صفر Ψ را خواهد داشت. در این مقاله از روش SOMP برای کدگذاری تنک و مشاهده‌ی ماتریس ضرایب تنک Ψ و از روش K-SVD برای آپدیت دیکشنری استفاده می‌گردد [۱۴، ۱۵]. اگر مجموعه‌ای را به نام مجموعه‌ی A در نظر بگیریم که شامل دو زیر فضا از دیکشنری فوق کامل H باشد، (زیر فضای مربوط به ناهنجاری‌ها و زیر فضای مربوط به پس‌زمینه) به این ترتیب تمام پیکسل‌های تست S_c می‌توانند از طریق رابطه (۱۲) محاسبه گردند:

$$\Sigma = S - A\Psi = S - A((A)^T A)^{-1} (A)^T S = (I - A((A)^T A)^{-1} (A)^T) S = P \frac{1}{B} S \quad (12)$$

رابطه (۱۲) میانگین انرژی نویز پیکسل تست و ناحیه همسایگی آن (به طور کلی S) را بیان می‌کند. بنابراین اگر پیکسل تست متعلق به پس‌زمینه باشد و هیچ‌گونه ناهنجاری نداشته باشد، میانگین انرژی تقریباً برابر با مقدار نویز پیکسل تست خواهد بود و می‌توان پیکسل مورد نظر را با رابطه (۱۳) نمایش داد:

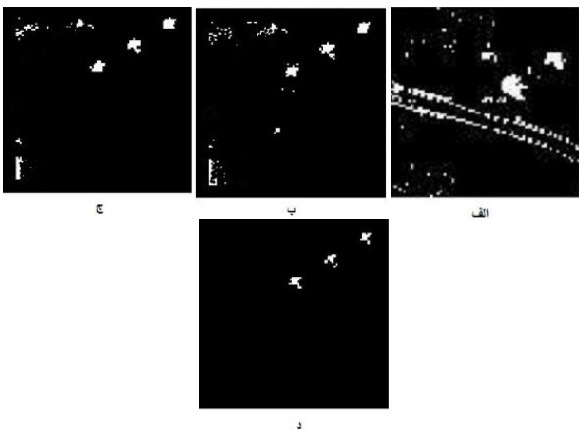


شکل ۳: داده (تصویر ابرطیف فرودگاه سن دیگو)



شکل ۴: تصویر (الف) اهداف منتخب از داده اصلی، تصویر (ب) خروجی الگوریتم GBJSR

شکل (۵) مقایسه‌ای چشمی بین روش‌های آشکارساز پیشنهادی ارائه شده. همانگونه که مشاهده می‌گردد روش GBJSR در آشکارسازی نسبت به دو روش RX, BJSR عملکرد بهتری داشته است.



شکل ۵: خروجی سه الگوریتم RX, BJSR, GBJSR تصویر (الف) الگوریتم RX، تصویر (ب) الگوریتم BJSR، تصویر (ج) الگوریتم GBJSR، تصویر (د) نقشه اصلی

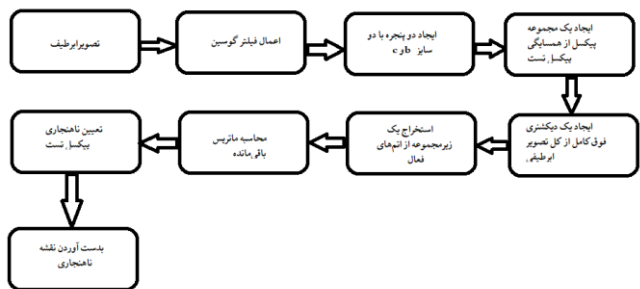
۲-۴ اعتبارسنجی

یکی از روش‌های قابل قبول برای ارزیابی و مقایسه روش‌های آشکارسازی در تصاویر ابرطیفی استفاده از منحنی ROC می‌باشد. در این منحنی هرچه سطح زیر نمودار بیشتر باشد به معنای عملکرد بهتر آشکارساز خواهد بود همانگونه که روی نمودار نیز مشاهده می‌گردد، روش ارائه شده عملکرد دقیق‌تری را نشان می‌دهد.

کرده و در بینهایت به صفر میل می‌کند. هرچه مقدار σ بزرگ‌تر باشد، آنگاه تابع با سرعت کمتری به صفر میل می‌کند.

در تصاویر ابرطیفی اهداف ناهنجار دارای دو ویژگی اساسی هستند: ۱- طیف این اهداف نسبت به پیش زمینه متفاوت است و ۲- تعداد اهداف ناهنجار در مقایسه با پیش زمینه بسیار کمتر است. همین موضوع باعث می‌شود که نویز ایجاد شده بر روی تصویر در بعضی از مکان‌های تصویر با اهداف ناهنجار مورد علاقه اشتباه گرفته شود. به عبارت دیگر نویز، به جای هدف ناهنجار به اشتباه آشکار شده و نرخ اعلام خطا افزایش می‌یابد. جهت جلوگیری از این امر پیشنهاد می‌گردد، در ابتدا، یک فیلتر گوسی به تصویر اعمال گردد تا نویزهای آن تا حد امکان مرتفع شود. سپس با استفاده از تئوری حسگری فشرده، بیان تنک تصویر ابرطیفی به دست آید. پس از آن با استفاده از محاسبه انرژی سیگنال و مقایسه آن با انرژی پیش زمینه، ناهنجار بودن آن آشکار خواهد شد.

راهکار پیشنهادی در این مقاله، بعد از اعمال فیلترگوسین روی تصویر و از بین بردن نویزهای اضافی، با استفاده از بیان تنک، می‌توان به آشکارسازی دقیق‌تری از ناهنجاری‌ها در تصویر ابرطیفی دست یافت.



شکل ۲: بلوک دیاگرام روش پیشنهادی

۴- داده مورد استفاده

داده استفاده شده مربوط به فرودگاه شهر سن دیگو واقع در آمریکا می‌باشد که از طریق سنجنده هوابرد AVIRIS در ۲۲۴ باند طیفی در محدوده‌ی طیفی ۳۷۰ تا ۲۵۱۰ نانومتر ثبت شده است. بعد از حذف باندهای جذبی و نویزی قسمتی از تصویر را با اندازه‌ی ۱۸۹ × ۱۰۰ × ۱۰۰ خواهیم داشت که شامل سه هواپیما به عنوان اهداف مورد نظر می‌باشد که الگوریتم‌های آشکارسازی RX, BJSR, GBJSR^۴ روی این بخش از تصویر اعمال می‌گردند.

۴-۱ نتایج پیاده‌سازی

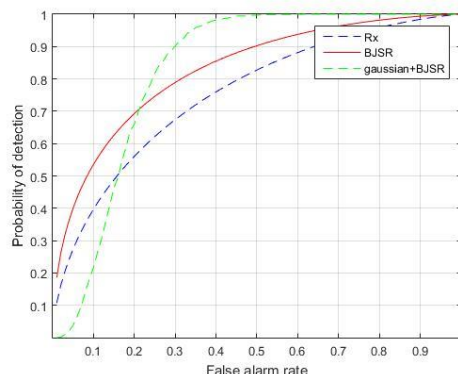
شکل (۳) خروجی الگوریتم GBJSR در کنار نقشه ناهنجاری اصلی را نشان می‌دهد. همان گونه که مشاهده می‌گردد این خروجی با توجه به تصویر اصلی آشکارسازی قابل قبولی را ارائه داده است و در نمایش ناهنجاری مخصوصاً لبه‌ها دقیق عمل کرده است.

⁴ Gaussian Background Joint Sparse Representatio

مراجع

- [1] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust, Speech Signal Process*, vol. 38, no.10, pp.1760–1770, Oct. 1990.
- [2] M.Theodore Eismann, "Hyperspectral remote sensing," SPIE, 2012.
- [3] E.Ashton and Alan Schaum, "Algorithms for the detection of sub-pixel targets in multispectral imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 64, no. 7, pp. 723–731, 1998.
- [4] Chein, Chang, "Anomaly Detection and Classification for Hyperspectral Imagery", *IEEE Transactions On Geoscience And Remote Sensing*, Vol. 40, No. 6, 2002.
- [5] M.Imani, "RX Anomaly Detector With Rectified Background", *IEEE Geoscience & Remote Sensing Letters*, vol.14,no.8,pp.1313–1317, Aug. 2017.
- [6] Q. Ling, W. Guo, Z. Lin, W. An, "A Constrained Sparse Representation Model for Hyperspectral Anomaly Detection", *IEEE Transactions on Geoscience & Remote Sensing*, vol.58,no.4,pp.2357–2371, April, 2019.
- [7] M.Imani, Attribute profile based target detection using collaborative and sparse representation, *Neurocomputing*, vol. 313, pp. 364–376, 2018.
- [8] M.Imani, Difference-Based Target Detection Using Mahalanobis Distance and Spectral Angle, *International Journal of Remote Sensing*, vol. 40, no. 3, pp. 811–831, 2019.
- [9] M.Aharon, M.Elad, and A.Bruckstein. The K-SVD: "An algorithm for designing of overcomplete dictionaries for sparse representation", *IEEE Trans. on Signal Processing*, vol.54,no. 11, pp. 4311–4322, 2006.
- [10] M.F. Duarte et al, "Distributed compressed sensing of jointly sparse signals", in *Proc. Conf. Rec. 39th Asilomar Conf. Signals Syst. Comput*, pp. 1537–1541, 2005.
- [11] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit", *Signal Process*, vol. 86, no. 3, pp. 572–588, 2006.
- [12] M.Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Berlin, Germany, Springer Verlag, 2010.
- [13] Li, H. zhang, L. zhang, L. Ma., "Hyperspectral Anomaly Detection by the use of Background Joint Sparse Representation," *IEEE Geoscience and Remote Sensing society*, vol.8,no.6;2523–2533, June. 2015.

جهت بهتر مشخص شدن تفاوت‌ها از شاخص AUC نیز کمک می‌گیریم که نشان دهنده سطح زیر نمودار ROC می‌باشد و هرچه عدد بزرگتری داشته باشد، نشان دهنده عملکرد بهتر الگوریتم مورد نظر می‌باشد. در جدول (۱) شاخص AUC هر سه الگوریتم مشاهده می‌شود.



شکل ۶: نمودار ROC توام سه الگوریتم RX, BJSR, GBJSR

جدول ۱- نتایج شاخص AUC برای سه الگوریتم RX, BJSR, GBJSR

الگوریتم	RX	BJSR	GBJSR
شاخص AUC	74.85	82.63	84.83

۵- نتیجه‌گیری

از آنجاییکه الگوریتم RX بر پایه محاسبه فاصله ماهالانوبیس بین پیکسل مورد آزمایش و پس زمینه استوار است، با محاسبه میانگین و ماتریس کواریانس، از کل داده‌های موجود در تصویر بدون در نظر گرفتن اقلیت یا اکثریت بودن پیکسل‌ها استفاده می‌کند. به همین دلیل عملکرد دقیقی برای آشکار کردن تعداد پیکسل اقلیت نخواهد داشت.

اما روش BJSR به دلیل تفکیک با دقت بسیار زیاد بین پیکسل‌های هدف و پس‌زمینه که با ایجاد پنجره دوگانه اطراف پیکسل مورد آزمایش میسر شده، عملکرد مطلوبی را در آشکار کردن پیکسل‌های اقلیت نشان داده است. با این حال این روش نیز به دلیل اینکه قابلیت تفکیک بین پیکسل‌های هدف و موارد نامطلوب (نویز) پس‌زمینه را ندارد، بسیاری از نویزهای موجود در پس‌زمینه را نیز به عنوان هدف ناهنجار در نظر می‌گیرد.

که برای رفع این مشکل در این مقاله از فیلتر گوسین بهره گرفتیم و همانگونه که آزمایش‌ها نشان داد روش GBJSR به دلیل حذف نویزهای اضافی عملکرد دقیق‌تری نسبت به دو روش قبلی دارد.

نتایج آزمایش‌ها نشان می‌دهد که الگوریتم BJSR نسبت به RX معمولی موفق به تشخیص ناهنجاری‌های بیشتری شده است و الگوریتم پیشنهادی یعنی GBJSR دارای عملکرد بهتری بوده است.

- [14] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit", *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.
- [15] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing over-complete dictionaries for sparse representation", *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [16] Access:
<https://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm>



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و



دانشگاه تبریز

سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸

پیاده سازی الگوریتم PCA برای بهبود اندازه گیری عمق عیوب زیرسطحی در آلیاژ مس به روش جریان گردابی پالسی

علی جهانی سه قلعه

کارشناس ارشد برق، سیستم‌های الکترونیک دیجیتال، موسسه آموزش عالی خراسان، مشهد، ایران

a.jahani@khorasan.ac.ir

چکیده

در میان عیوب ریخته گری، ترک‌ها و عیوب سطحی به وسیله چشم و یا با کمک روش‌های ساده‌ای مانند مایعات نافذ و مواد مغناطیسی، قابل تشخیص هستند، اما کشف عیوب زیرسطحی نیاز به روشهای پیچیده‌تری دارد. در این تحقیق از جریان‌های گردابی بمنظور بررسی غیر مخرب برای تعیین مکان نقص‌های زیرسطحی در قطعات ریخته گری و استخراج اطلاعات استفاده شده است. یک نمونه از آلیاژهای مس (UNS C83600)، که بر روی آن بصورت عمودی چهار عیب زیرسطحی ایجاد شده‌است، توسط یک کاوشگر ساخته شده از دو پیچه، مورد آزمون قرار گرفت. ویژگی‌های مقدار قله سیگنال، زمان رسیدن به قله، میانگین سیگنال، واریانس، توان سیگنال، شیب صعود، کشیدگی، انحراف استاندارد و زمان صعود از سیگنال دریافت شده توسط پیچه استخراج شد. در نهایت با استفاده از الگوریتم تحلیل مولفه اساسی (PCA) ویژگی‌های موثرتر با ابعاد کمتر مورد تحلیل قرار گرفت. برای مقایسه ویژگی‌ها، مقدار خطای جذر میانگین مربعات (RMSE) مورد ارزیابی قرار گرفت. ویژگی Y1 حاصل از اعمال PCA بین دو ویژگی (قله سیگنال) و ویژگی (زمان صعود) با کمترین مقدار خطای (۰,۱۷۲۵) و ویژگی میانگین با مقدار خطای (۰,۳۶۵۱) به عنوان بهترین ویژگی‌ها انتخاب می‌شوند.

کلمات کلیدی

عیوب زیرسطحی، آلیاژ مس (UNS C83600)، الگوریتم تجزیه به مولفه‌های اساسی، جریان گردابی

آزمون غیر مخرب، قراردادن جسم در معرض یک منبع انرژی است تا جریانی از انرژی به آن وارد گردد عیوب زیرسطحی باعث ایجاد تغییرات در جریان انرژی خواهد شد. با اندازه‌گیری انرژی خارج شده از جسم می‌توان این عیوب و ناهنجاری‌ها را تشخیص داد.

از انواع آزمون‌های غیرمخرب که متداول می‌باشد، می‌توان آزمون فراصوت، آزمون ترموگرافی و جریان گردابی را نام برد. این تحقیق به استفاده از جریان‌های گردابی برای کشف عیوب زیر سطحی و عمق وجود ضایعه می‌پردازد.

ژان برنارد فوکو، دانشمند فرانسوی، در سال ۱۸۳۰ ایجاد جریان گردابی در اجسام رسانا را کشف کرد. و در سال ۱۸۳۱ جوزف هنری در آمریکا و مایکل فارادی در انگلیس به صورت جداگانه القای الکترومغناطیسی را کشف کردند. اما سالها طول کشید که از این پدیده برای انجام آزمون‌های غیر مخرب در قطعات استفاده شود در حقیقت سالها بعد، یعنی در سال ۱۸۷۹ دیوید ادوارد هیوز دانشمند انگلیسی

۱- مقدمه

تبدیل فلزات خالص به محصولات قابل استفاده، به دو روش ریخته‌گری و ماشین‌کاری ممکن است. طراحان با انجام محاسبات و شبیه سازی، قطعات را طوری طراحی می‌کنند تا در معرض تنش‌های مکانیکی حین کار دوام مناسبی داشته باشند. وجود ترک، حفرات و ضایعه‌های سطحی و زیر سطحی موجب کاهش عمر مفید قطعات و افزایش هزینه‌های تعمیر نگهداری خواهد شد.

ترک‌ها و ضایعات سطحی با چشم و یا تکنیک‌هایی مانند مایعات نافذ و غیره براحتی قابل تشخیص هستند در صورتی که برای کشف عیوب زیرسطحی بصورتی که جسم قابلیت استفاده مجدد خود را از دست ندهد، نیاز به روش‌های پیچیده‌تری می‌باشد.

غیر فرو مغناطیسی ساخته شده از AISI type 316 stainless steel در [2] نشان داده شده است.

نمونه‌های دیگر، کشف و شناسایی عیوب شامل نقص در سطح داخلی لوله تحت عایق خوردگی بیرونی لوله‌های فولادی کربنی در صنعت نفت و گاز [3]، ترک‌های سطحی و زیرسطحی و تشخیص حفره در سازه‌ها [4]، خوردگی زیر سطح در یک ورق [5]، انواع مختلف نقص در سازه‌های پشتیبانی کننده توربین بخار [6] را می‌توان نشان داد.

A. Sophian و همکاران در [7] با استفاده از الگوریتم PCA و تحلیل داده‌ها توانستند روشی برای دسته بندی عیوب بین سه گروه عیوب سطحی، زیرسطحی و تغییر ضخامت ارائه دهند.

Tian و همکاران در [8] با معرفی ویژگی جدید به پردازش سیگنال برای کشف عیوب زیرسطحی پرداخته‌اند.

Xuan Bing Qiu و همکاران در [9] با استفاده از تبدیل موجک روشی برای دسته بندی عیوب را معرفی کردند.

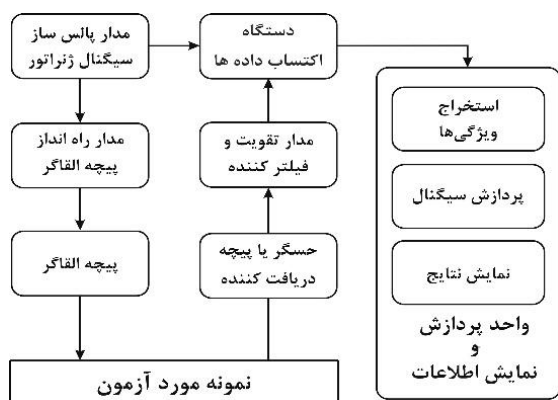
همچنین Yunze He و همکاران در [10] با پردازش سیگنالهای دریافتی به تصویری از عیوب زیر سطحی رسیدند.

Tiyan و همکاران در [11] به پردازش ویژگی‌هایی از سیگنال PEC در حوزه زمان پرداخته‌اند.

اما در این تحقیق سامانه‌ای پیاده سازی می‌شود که پالس‌های الکتریکی با فرکانسی مناسب را به پیچه القاگر تزریق نماید، سپس سیگنال دریافت شده، ابتدا فیلتر و آماده سازی شده و سپس مورد پردازش قرار می‌گیرد و ویژگی‌هایی از این سیگنال استخراج شده و مورد مطالعه قرار می‌گیرد. سپس با استفاده از الگوریتم تحلیل مولفه اصلی PCA، ویژگی جدیدی، به دست آید که بتواند نتایجی با خطای کمتر را ارائه نماید.

۲- ابزار و روش‌ها

برای انجام این آزمایش سامانه PEC مطابق ساختار شکل ۱ باید پیاده سازی گردد.



شکل ۱: شمای کلی یک سامانه PEC

مدار سیگنال ژنراتور پالسی با شکل موج مربعی با فرکانس و چرخه وظیفه مناسب تولید می‌کند. مدار راه انداز القاگر، این پالس را با ولتاژ

آمریکایی از این روش در بازرسی غیر مخرب قطعات ریخته گری شده، استفاده کرد و اکتشاف کاربردهای متفاوت جریان‌های گردابی آغاز شد.

در حقیقت جریان گردابی شارش الکترونها در رسانا می‌باشد. این جریان باعث ایجاد میدان مغناطیسی ثانویه‌ای می‌گردد که بعنوان میدان مغناطیسی جریان گردابی شناخته می‌شود و این میدان طبق قانون لنز در جهت معکوس میدان اولیه می‌باشد. [1]

دو تکنیک در بازرسی جریان گردابی وجود دارد: آزمون جریان گردابی عادی^۱ و جریان گردابی پالسی یا گذرا^۲. در آزمون جریان گردابی عادی، یک کاوشگر با یک سیگنال سینوسی هارمونیک در یک فرکانس یا چند فرکانس برای مطالعه سیگنال‌های جریانی گردابی تحریک می‌شود که در نهایت با تغییر امپدانس پیچه دریافت کننده، اثر خود را نمایان می‌کند. [1]

مسیر جریان‌های گردابی القاء شده در داخل قطعه به صورت حلقه‌های بسته در صفحه‌های عمود بر شار مغناطیسی است، به طوری که چگالی جریان گردابی به صورت نمایی از سطح به طرف درون قطعه کم می‌شود. تراکم جریان‌های گردابی در نقاطی از سطح قطعه در زیر سیم پیچ حداکثر بوده و با افزایش فاصله از سطح به سمت داخل قطعه کاهش می‌یابد. [1]

بنا بر تعریف عمق موثر استاندارد برای جریان گردابی عمقی است که تراکم جریان‌های گردابی ۳۷٪ یا $1/e$ تراکم آنها در سطح ماده باشد. رابطه (۱) معادله حاکم بر عمق نفوذ است. این رابطه به رابطه اثر پوستی جریان گردابی مشهور است. [1]

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} = \sqrt{\frac{2}{\omega \mu \sigma}} \quad (1)$$

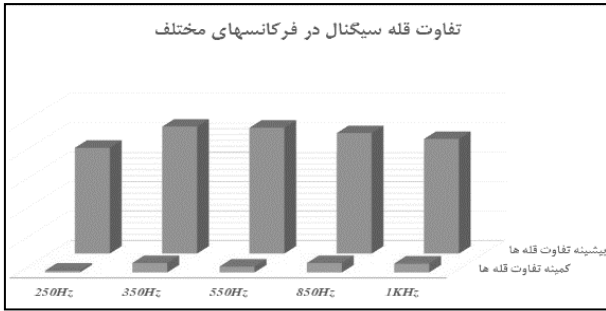
در رابطه فوق f نشان دهنده فرکانس بر حسب Hz، μ نفوذپذیری مغناطیسی بر حسب (H/m)، σ هدایت الکتریکی بر حسب (Simens/m) و ω سرعت زاویه‌ای بر حسب (rad/s) می‌باشد. همانطور که از این رابطه مشخص می‌شود از بین عوامل موثر بر عمق نفوذ، تنها فرکانس قابل تغییر و تنظیم است و دو عامل دیگر، رسانایی الکتریکی و نفوذپذیری مغناطیسی از خواص ذاتی ماده است.

در تکنیک جریان گردابی پالسی، تحریک پیچه القایی به وسیله شکل موج پالس انجام می‌شود. بدین معنی که یک پیچه القاگر که با پالس الکتریکی تغذیه می‌شود، از طریق میدان الکترومغناطیسی، جریان گردابی گذرا را در نمونه، بوجود می‌آورد. سپس سیگنال دریافتی در پیچه دریافت کننده توسط یک سامانه هوشمند پردازش می‌گردد و مورد تحلیل قرار می‌گیرد. مزیت اولیه و اصلی تکنیک PEC این است که پردازش سیگنال‌های آن نسبتاً ساده‌تر است، در حالی که تحلیل سیگنال‌های ECT نیاز به مهارت خاص آزمون‌گر دارد. [1]

بخش زیادی از فعالیت‌های محققان حوزه جریان‌های گردابی، در رابطه با موضوع کشف عیوب زیرسطحی بوده است که نشان می‌دهد این موضوع تمرکز اصلی پژوهشگران در بکارگیری PEC بوده است. به عنوان مثال، تشخیص و اندازه‌گیری نقص‌های زیرسطحی در یک نمونه

¹ Eddy Current Test

² Pulsed Eddy Current



شکل ۳: تفاوت قله‌های سیگنال در فرکانس‌های مختلف

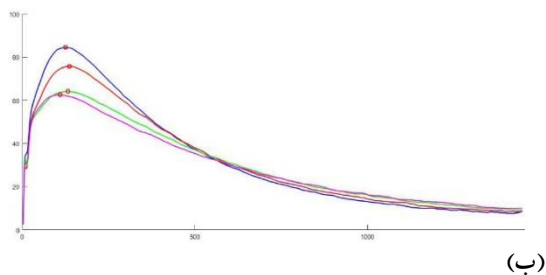
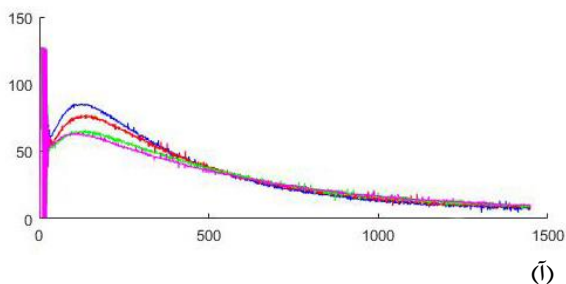
بنابر آنچه در نمودار شکل ۳ مشاهده می‌شود، بیشترین تفاوت‌ها در فرکانس ۳۵۰ هرتز رخ می‌دهد بنابراین فرکانس کارا برای اندازه‌گیری در این سامانه فرکانس ۳۵۰ هرتز می‌باشد. در این فرکانس با استفاده از رابطه ۱ و اطلاعات جدول ۱، عمق نفوذ برابر ۹,۱۲ میلی‌متر محاسبه می‌گردد.

۲-۳- پیش پردازش داده‌ها

برای استخراج ویژگی‌ها و پردازش سیگنال ابتدا باید بخشی از سیگنال دریافت شده که در بازه زمانی صفر بودن سطح ولتاژ پیچه القاگر می‌باشد، جداسازی گردد. زیرا دقیقاً در این بازه است که پیچه القا در وضعیت خاموش است و پیچه دریافت کننده تنها تحت تاثیر جریان گردابی باقیمانده در جسم می‌باشد.

همانطور که انتظار می‌رود سیگنال دریافت شده نویزدار است. با شناسایی منابع نویز و رفع آنها و استفاده از کابل‌های شیلددار میزان آن به حداقل رسانده شد اما همچنان مقداری نویز بر روی سیگنال حاصل باقی می‌ماند که باید توسط یک فیلتر پایین گذر، رفع گردد.

شکل ۴ نمونه‌ای از سیگنال نویزدار (آ) و فیلترشده (ب) را نشان می‌دهد.

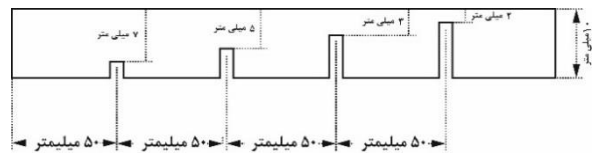


شکل ۴: نمونه سیگنال جدا شده: (آ) نویزدار (ب) فیلتر شده

و جریان مناسب برای پیچه القاگر آماده می‌نماید. میدان گذرا بوجود آمده در پیچه القاگر جریان‌های گردابی را داخل نمونه آزمون به وجود می‌آورد. سیگنال حاصل توسط حسگر یا پیچه دریافت کننده دریافت می‌شود. سپس در مرحله بعد، تقویت و فیلتر می‌شود. این سیگنال، در ادامه، به وسیله یک دستگاه اکتساب داده‌ها دریافت و ذخیره می‌شود تا در مرحله پایانی سیگنال حاصل، جداسازی، پردازش و تحلیل شود که عمق ضایعه تخمین زده شده و نمایش داده شود.

۲-۱- نمونه مورد آزمایش

برای انجام آزمایش نمونه‌ای از آلیاژی با پایه مس (UNS C83600) تهیه شده و مطابق ساختار شکل ۲ ماشینکاری می‌شود. این شیارها بر روی جسم ایجاد می‌گردند تا بتوان شرایط عیوب زیر سطحی را شبیه سازی کرد. [۱]



شکل ۲: ساختار نمونه ساخته شده جهت آزمایش

برخی خواص فیزیکی این آلیاژ که برای محاسبه عمق نفوذ لازم است در جدول ۱ آورده شده است.

جدول ۱: برخی مشخصات فیزیکی آلیاژ C83600 [۱۲]

مشخصه فیزیکی	مقدار
ضریب رسانش الکتریکی	$15\% IACS = 0.087 MS/cm$
تراوایی مغناطیسی	$\mu_r = 1$ $(\mu = 4\pi \times 10^{-7} H/m)$

۲-۲- نحوه انجام آزمایش

برای تعیین فرکانس کارا سه روش پیشنهاد شده است. [۱۳]

۱- تعیین فرکانس کارا بر اساس عمق نفوذ مورد نظر

۲- تعیین فرکانس کارا بر اساس حساسیت بیشینه

۳- استفاده از تحلیل رگرسیون برای تعیین فرکانس کارا

با استفاده از روش دوم، برای تعیین این فرکانس، قله‌های سیگنال در چهار ضایعه برای ۵ فرکانس ۲۵۰، ۳۵۰، ۵۵۰، ۸۵۰ و ۱ کیلوهرتز اندازه‌گیری شد. و بر اساس آن تفاوت بین دو قله سیگنال کمتر و همچنین تفاوت بین کمترین و بیشترین قله سیگنال، در هر فرکانس و در بین نتایج این پنج آزمایش اندازه‌گیری شد. این تفاوت‌ها در نمودار شکل ۳ قابل مشاهده است.

۲-۴- روش تحلیل مولفه اصلی^۳ (PCA)

محاسبه ماتریس کوواریانس برای پیدا کردن بیشترین تفکیک الزامی است.

گام سوم: محاسبه بردارها و مقادیر ویژه ماتریس کوواریانس
از آنجایی که ماتریس کوواریانس یک ماتریس مربعی است، ما می‌توانیم برای آن مقادیر ویژه و بردارهای ویژه را محاسبه کنیم. که ویژگی‌های مجموعه اطلاعات ما را بارزتر می‌کنند.

گام چهارم: مرتب کردن ماتریس مقادیر ویژه
در این گام لازم است ستون‌های ماتریس بردارهای ویژه، بر مبنای بزرگی مقادیر ویژه متناظر با آنها، از زیاد به کم، مرتب شوند و این به معنی مرتب شدن بردارهای ویژه بر اساس اهمیت و تاثیر آنها می‌باشد.

گام پنجم: انتخاب مولفه‌های اصلی و ایجاد بردار ویژگی
بردارهای ویژه متناظر با مقادیر ویژه بزرگتر، همان ویژگی‌های اصلی این اطلاعات هستند. اکنون می‌توان برای کاهش ابعاد اطلاعات، از بردارهای ویژه متناظر با مقادیر ویژه کمتر، صرف نظر کرد.

گام ششم: به دست آوردن مجموعه اطلاعات جدید
بعد از اینکه ویژگی‌های مورد نظر را انتخاب کردیم در حقیقت کفایت ترانهاده بردارهای ویژه بدست آمده را از سمت چپ در بردار اطلاعات اصلی، مانند رابطه ۲، ضرب کنیم. اطلاعات تازه با استفاده از ماتریس بردار ویژه به فضای مختصات جدید نگاشت می‌شوند.

$$Y = Q * X \quad (2)$$

در رابطه فوق ماتریس Q، همه مولفه‌های دستگاه اول را به دستگاه مختصات دوم نگاشت می‌کند.

۲-۵- استخراج ویژگی‌ها

تعداد ۹ ویژگی از سیگنال‌های فیلتر شده، استخراج شده‌اند. این ویژگی‌ها اغلب بر اساس سوابق تحقیقاتی انتخاب شده‌اند. در ادامه این ویژگی‌ها معرفی شده‌اند

۱) قله سیگنال

قله سیگنال یکی از متداول‌ترین ویژگی‌های بکار رفته در تحقیقات انجام شده، می‌باشد [1]. برای تعیین دقیقتر مقدار قله سیگنال از تابع $\max()$ در متلب همراه با تمهیداتی استفاده شد.

۲) زمان رسیدن به قله

زمان رسیدن به قله سیگنال دومین ویژگی است که در تحقیقات زیادی مورد تحلیل قرار گرفته است. [1]

۳) میانگین سیگنال

با توجه به اینکه مقدار میانگین منحنی با ارتفاع قله سیگنال تغییر می‌کند، به نظر می‌آید این ویژگی می‌تواند گزینه مناسبی برای تحقیق باشد. تابع $\text{mean}()$ در متلب مقدار میانگین سیگنال را محاسبه می‌کند.

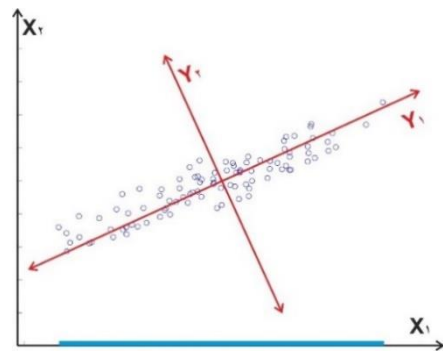
۴) واریانس سیگنال

تابع $\text{Var}()$ در متلب مقدار واریانس سیگنال ورودی را تعیین می‌کند. واریانس معیار مناسبی برای نشان دادن دامنه تغییرات یک متغیر نسبت به مقدار میانگین می‌باشد.

۵) توان سیگنال

این الگوریتم، از نتایج جبرخطی است که برای پیدا کردن الگوهای بهتر در اطلاعات، بکار می‌رود. موفقیت دیگری که با روش تحلیل مولفه اصلی به دست می‌آید، این است که می‌توان ابعاد اطلاعات را بدون آنکه اطلاعات زیادی از دست برود، کم کرد.

همانطور که در شکل ۵ قابل مشاهده است راستای پراکندگی اطلاعات با محور X_1 موازی نیست بنابراین تصویر نقاط بر روی محور X_1 دارای تداخل می‌باشد به زبان ریاضی واریانس کمتری دارد.



شکل ۵: دستگاه مختصات X_2/X_1 و نگاشت به Y_1/Y_2

حال دستگاه مختصات (Y_1, Y_2) را در نظر بگیرید که بر نقطه میانگین اطلاعات پایه گذاری شده و محور مولفه اول آن، Y_1 منطبق بر راستای پراکندگی اطلاعات می‌باشد.

در دستگاه مختصات جدید، بیشترین واریانس در تصویر نقاط اطلاعات بر روی محور Y_1 حاصل می‌شود. واضح است که انتقال اطلاعات به این دستگاه مختصات جدید، با استفاده از نگاشت خطی و ضرب ماتریسی امکان پذیر است. برای پیدا کردن ماتریس نگاشتی که بهترین تفکیک را نتیجه بدهد با استفاده از الگوریتم PCA، دو پارامتر بردار ویژه و مقدار ویژه تعریف می‌گردد هر بردار ویژه مولفه ای از دستگاه دوم را بر اساس مولفه‌های دستگاه اول تعریف می‌کند.

در دستگاه دوم (Y_1, Y_2) مولفه Y_1 دارای واریانس بیشتری نسبت به مولفه Y_2 می‌باشد. بزرگی پارامتر مقدار ویژه نشان دهنده این تفاوت و موثرتر بودن مولفه‌های دستگاه جدید می‌باشد بنابراین لازمست پس از محاسبه بردارهای ویژه این بردارها به ترتیب مقادیر ویژه متناظرشان مرتب شوند تا همیشه موثرترین مولفه‌ها در ابتدای فهرست قرار گیرند.

حال می‌توان بر اساس نتایج حاصله از یک یا چند مولفه حاصل که دامنه تغییرات کمتری دارند، در مقابل مولفه‌های موثرتر صرف نظر کرد به صورتی که به تحلیل داده‌ها صدمه چندانی وارد نشود.

روند گام به گامی برای پیاده سازی الگوریتم PCA وجود دارد:

گام اول: جمع آوری اطلاعات

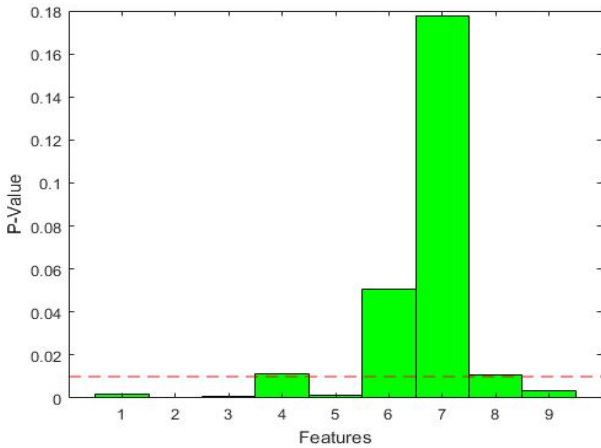
در گام نخست، اطلاعات جمع آوری شده و به صورت ماتریس ذخیره می‌گردد.

گام دوم: محاسبه ماتریس کوواریانس

³ Principal Component Analysis

بنابراین اگر مقدار محاسبه شده احتمال از یک سطح رواداری کمتر باشد، ویژگی مورد آزمون می‌تواند ویژگی خوبی باشد. این سطح رواداری معمولاً ۰٫۰۵ و یا ۰٫۰۱، در نظر گرفته می‌شود. در این تحقیق، سطح رواداری، سختگیرانه‌تر، ۰٫۰۱، در نظر گرفته شد.

شکل ۶ مقادیر احتمال حاصل از انجام آزمون T را برای هر ویژگی نمایش می‌دهد. با توجه به این نتایج، ویژگی‌های شماره ۴، ۶، ۷، ۸ رد شده و حذف می‌شوند.



شکل ۶: مقادیر احتمال حاصل از آزمون T ویژگی‌ها

۲-۷- مقایسه و ارزیابی ویژگی‌ها

برای مقایسه بهتر نتایج، پارامتر^۴ RMSE، برای نمایش تفاوت بین مقادیر تخمینی و مقادیر واقعی، محاسبه شد.

محاسبات انجام شده شامل دو مرحله می‌باشد. در مرحله اول ویژگی‌های استخراج شده به صورت جداگانه در مقایسه با مجموعه هدف که همان مقادیر عمق ضایعه‌ها می‌باشد، ترسیم می‌شوند. سپس خط رگرسیون بین نقاط اطلاعات رسم می‌شود و مقادیر RMSE در مقایسه با نقاط متناظر بر روی خط رگرسیون، مقادیر تخمینی، محاسبه می‌شوند.

در مرحله دوم، در ابتدا با پیاده سازی الگوریتم PCA، مولفه‌های Y1 و Y2 بدست می‌آیند. از مولفه Y2 در مقابل مولفه موثرتر Y1 صرفنظر می‌شود. و مولفه Y1 مانند هر کدام از ویژگی‌های سیگنال پردازش شده و مقادیر RMSE محاسبه می‌گردند. شکل ۷ نحوه انجام محاسبات را نشان می‌دهد.

توان سیگنال نیز از ویژگی‌هایی است که در سابقه مقالات پیشنهاد شده است. [11] این ویژگی با استفاده از رابطه ۶ محاسبه می‌شود.

$$P_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (6)$$

که در آن T زمان تناوب، t زمان، x مقادیر سیگنال گسسته است.

۶ شیب خط قسمت بالارونده سیگنال

این ویژگی در حقیقت شیب برداری است که از مبدا مختصات به نقطه قله متصل می‌گردد. واضح است که در این شرایط مولفه افقی این بردار زمان رسیدن به قله و مولفه عمودی آن مقدار قله است در نتیجه این ویژگی از تقسیم ویژگی شماره ۱، قله سیگنال، بر ویژگی شماره ۲، زمان رسیدن به قله، به دست می‌آید. این ویژگی در [11] نیز بکار گرفته شده است.

۷ کشیدگی

این ویژگی نیز در مقالات تحقیقاتی پیشنهاد شده است. [1] کشیدگی منحنی به عبارت ساده به همان ارتفاع توزیع داده‌ها اشاره می‌کند که همان ممان چهارم سیگنال و ضریبی برای تعیین کشیدگی سیگنال می‌باشد.

۸ پراکندگی استاندارد

پارامتر پراکندگی استاندارد با محاسبه میانگین فاصله اعضا تا مقدار میانگین محاسبه می‌شود تابع STD در متلب مقدار پراکندگی استاندارد منحنی سیگنال را محاسبه می‌کند.

۹ زمان محدوده بالا رفتن منحنی سیگنال

برداری که زمان انتقال از شکل موج ورودی X، به عبور از سطوح مرجع ۱۰٪ تا ۹۰٪ قله می‌رسد را بدست می‌دهد. این تابع برای تعیین زمان افزایش، سطوح حالت شکل موج ورودی را با روش هیستوگرام برآورد می‌کند. این ویژگی در [8] بکار گرفته شده است.

۲-۶- انتخاب ویژگی‌ها بوسیله آزمون T

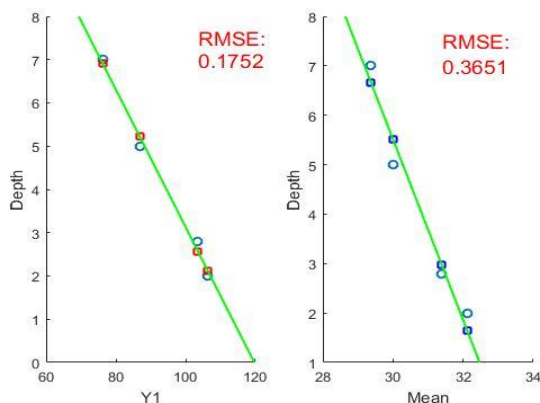
برای انتخاب ویژگی‌ها می‌توان از آزمون‌های استنباط آماری بهره برد. یکی از انواع این آزمون‌ها، آزمون T است که بر اساس میانگین جامعه‌های آماری، تصمیم‌گیری می‌کند. با توجه به تعداد کم اطلاعات، این آزمون در بین آزمون‌های مشابه، برای این تحقیق، مناسب‌تر به نظر می‌رسد.

روش محاسبه، به دست آوردن شواهدی از یک اختلاف معنی‌دار در میان دو جمعیت (بین مقادیر یک ویژگی و مجموعه هدف) می‌باشد. به عبارت دیگر، T از تفاوت میانگین دو مجموعه، تقسیم بر انحراف استاندارد است. انحراف استاندارد معیاری برای برآورد میزان نزدیکی میانگین نمونه به میانگین جمعیت آماری است. تابع TTest در نرم افزار متلب این آزمون را انجام می‌دهد و مقدار احتمال Pvalue را بر می‌گرداند.

مقدار P احتمال رد شدن فرضیه صفر می‌باشد و فرض صفر یعنی: ((اختلاف معنی‌داری بین دو مجموعه وجود ندارد))

به عبارت دیگر مقدار P میزان اتفافی بودن، اختلاف مشاهده شده از میانگین اطلاعات می‌باشد. هر چقدر این احتمال، عددی کوچک باشد اعتماد ما به واقعی بودن اختلاف مشاهده شده بیشتر است.

⁴ Root Mean Square Error



۳- نتایج

جدول ۲ مقادیر محاسبه شده RMSE را برای ویژگی‌ها و همچنین برای مولفه Y1 حاصل از پیاده سازی الگوریتم PCA، را نمایش می‌دهد.

جدول ۲: مقادیر RMSE

۹	۵	۳	۲	۱	
۰,۱۷۲۵	۰,۵۷۴۵	۰,۷۳۵۶	۰,۸۴۰۷	۰,۷۴۵۳	۱
۱,۳۱۰۳	۰,۵۷۱۶	۲,۰۴۵۳	۲,۰۶۵۵	----	۲
۰,۶۵۳۰	۰,۵۷۴۱	۰,۳۶۵۱	----	----	۳
۰,۵۷۱۸	۰,۵۷۴۱	----	----	----	۵
۰,۶۶۴۵	----	----	----	----	۹

شکل ۷: محاسبه RMSE برای ویژگی‌های سیگنال و مولفه Y1

با محاسبه RMSE، ویژگی‌های سیگنال استخراج شده، با یکدیگر مقایسه شدند. از اعمال الگوریتم PCA بین قله سیگنال و زمان صعود، مولفه جدیدی بوجود آمده که مقدار بسیار کم $RMSE=0,1725$ را نتیجه می‌دهد که نشان می‌دهد با این روش مقادیر تخمینی به مقادیر واقعی عمق عیوب بسیار نزدیکتر می‌شوند. همچنین ویژگی میانگین سیگنال معرفی گردید که نتیجه مناسب $RMSE=0,3651$ را نشان داد. میانگین می‌تواند ویژگی مناسبی برای پردازش این سیگنال باشد. با فرض ثابت بودن نوبز، به دست آوردن ویژگی میانگین، از استخراج ویژگی‌های دیگر بخصوص قله سیگنال ساده‌تر به نظر می‌رسد.

۴-۱- پیشنهادهاى ادامه تحقیق

در ادامه این تحقیق پیشنهاد می‌گردد:

- آلیاژهای دیگری نیز آزمایش شده و سیگنال‌ها با روش ارائه شده پردازش شده و نتایج مقایسه شوند.
- ویژگی‌های دیگری نیز، از هر دو حوزه زمان و فرکانس این سیگنال استخراج و با این روش مورد مطالعه قرار گیرند.
- برای تحلیل بهتر ویژگی سیگنال‌ها، از شبکه‌های عصبی و ماشین بردار پشتیبان استفاده گردد.

مراجع

- [1] Sophian • Guiyun Tian• Mengbao Fan, " Pulsed Eddy Current Non-destructive Testing and Evaluation: A Review" Chinese Mechanical Engineering Society and Springer- 2017
- [2] Arjun V, Sasi B, Rao B P C, et al. Optimisation of pulsed eddy current probe for detection of sub-surface defects in stainless steel plates. Sensors Actuators A Phys., 2015.
- [3] Majidnia S, Nilavalan R, Rudlin J. Investigation of an encircling Pulsed Eddy Current probe for corrosion detection. IEEE Sensors 2014 Proceedings, 2014.
- [4] Li Y, Yan B, Li D, et al. Pulse-modulation eddy current inspection of subsurface corrosion in conductive structures. NDT&E Int., 2016.
- [5] Qiu X, Zhang P, Wei J, et al. Defect classification by pulsed eddy current technique in con-casting slabs based on spectrum analysis and wavelet decomposition. Sensors Actuators A Phys., 2013.

مقادیر روی قطر جدول (تقاطع شماره‌های یکسان) مقادیر محاسبه شده از بکارگیری ویژگی‌های سیگنال می‌باشند مولفه Y1 مقادیر حاصل از بکارگیری PCA بین دو ویژگی سیگنال می‌باشد بهترین نتیجه از اعمال الگوریتم PCA بین دو ویژگی مقدار قله و زمان صعود به دست می‌آید. در بین ویژگی‌های اصلی، بکارگیری ویژگی شماره ۳ (میانگین) نتیجه ای با کمترین مقدار RMSE در بین ویژگی‌های سیگنال را نتیجه می‌دهد. جدول ۳ مقادیر متوسط خطای اندازه گیری در سه مورد از موارد فوق را نشان می‌دهد.

جدول ۳: محاسبه متوسط درصد خطا بین سه روش محاسبه

خطای متوسط حاصل از PCA بین دو ویژگی ۱ و ۹	۴,۳۹%
خطای متوسط با استفاده از قله سیگنال	۲۸,۹۶%
خطای متوسط با استفاده از ویژگی میانگین	۱۰,۴۶%

۴- نتیجه گیری

آزمایش‌ها و تحلیل‌های انجام شده نشان داد که کدام ویژگی‌ها برای این آلیاژ مس مناسب می‌باشند و فرکانس کارا در این آزمایش در محدوده ۳۰۰-۶۰۰ هرتز می‌باشد. در این تحقیق، با ساخت کاوشگر، نمونه آزمون مناسب، پیاده‌سازی آزمایشگاهی سامانه آزمون PEC، جمع‌آوری اطلاعات و پیاده‌سازی الگوریتم PCA در تحلیل ویژگی‌های سیگنال انجام شد.

- [6] Buck J A, Underhill P R, Mokros S G, et al. Pulsed eddy current inspection of support structures in steam generators. IEEE Sensors Journal, 2015.
- [7] A.sophian 2003
- [8] Tian G Y, Sophian A. Defect classification using a new feature for pulsed eddy current sensors. NDT & E International, 2005, 38(1): 77-82.
- [9] Qiu X B, Liu L L, Li C L, et al. Defect classification by pulsed eddy-current technique based on power spectral density analysis combined with wavelet transform. IEEE Trans. Magn., 2014, 50(9): 1-8.
- [10] He Y, Pan M, Luo F, et al. Pulsed eddy current imaging and frequency spectrum analysis for hidden defect nondestructive testing and evaluation. NDT E Int., 2011, 44(4): 344-352 .
- [11] Chen T, Tian G Y, Sophian A, et al. Feature extraction and selection for defect classification of pulsed eddy current NDT. NDT E Int., 2008, 41(6): 467-476.

[۱۲] م‌عطفی، ف: «کلید مس و آلیاژهای وابسته». جلد یک. چاپ اول. فدک ایساتیس، تهران. ص ۳۳۹، ۱۳۸۵.

[۱۳] سیف پور، م: «بررسی تغییرات ریز ساختاری در ورق فولادی تغییرشکل یافته حین فرآیند آنیل به روش غیر مخرب جریان گردابی»، پایان نامه کارشناسی ارشد، دانشگاه فردوسی مشهد، ۱۳۹۱.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و



سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸

آشکارسازی خودکار خونریزی‌های بسیار کوچک مغزی در تصاویر SWI مبتنی بر شبکه عصبی کانولوشن

زهره نادری اسرمی^۱، محمدرضا حسن زاده^۲، مهدی ازوجی^۳

^۱ دانشجوی کارشناسی ارشد، گروه الکترونیک، دانشگاه صنعتی نوشیروانی بابل، بابل

zohreh.ndr@gmail.com

^۲ استادیار، گروه الکترونیک، دانشگاه صنعتی نوشیروانی بابل، بابل

m.hasanzadeh@nit.ac.ir

^۳ استادیار، گروه الکترونیک، دانشگاه صنعتی نوشیروانی بابل، بابل

m.ezoji@nit.ac.ir

چکیده

در این مقاله، روشی مبتنی بر شبکه عصبی کانولوشن (CNN) برای آشکارسازی خودکار خونریزی‌های بسیار کوچک مغزی (CMBs) ارائه شده است. CMB تجمع‌های کوچکی از خون در مغز است که با افزایش سن، بیماری‌های مغزی-عروقی، فشار خون بالا و بیماری مزمن کلیوی در ارتباط است و همچنین خطر بروز سکته مغزی را افزایش می‌دهد. تشخیص CMB در روش بالینی فعلی به صورت دستی توسط متخصص انجام می‌شود که کاری دشوار، وقت‌گیر و مستعد خطا است. در این تحقیق، از ۲۰ تصویر SWI شامل ۳۲۷۶ پج با اندازه ۶۱×۶۱ استفاده شده است. طی آزمایش‌های گوناگون، معماری شبکه عصبی کانولوشن پیشنهادی به صورت سه لایه کانولوشن، دو لایه پولینگ و یک لایه تماماً متصل در نظر گرفته شده است. نتایج اعتبارسنجی متقابل ۱۰ پوشه‌ای بر روی پایگاه-داده‌ی مربوطه نشان می‌دهد که الگوریتم پیشنهادی در مقایسه با کارهای پیشین عملکرد بهتری دارد.

کلمات کلیدی

آشکارسازی خودکار، شبکه عصبی کانولوشن، خونریزی کوچک مغزی (CMB)، تصاویر تشدید مغناطیسی، یادگیری عمیق

شود. از میان روش‌های مختلف تصویربرداری تشدید مغناطیسی بیشتر

از دو روش T2-star و SWI^۲ استفاده می‌شود. در [3] نشان داده شد که

SWI تعداد CMBهای شناسایی شده را افزایش می‌دهد. در شکل (۱)

اسکن SWI از مجموعه نشان داده شده است.

پردازش تصاویر برای آشکارسازی خودکار CMBها با چالش‌هایی

روبه‌رو است: الف) اندازه متفاوت CMBها بین ۲ تا ۱۰ میلی‌متر، ب)

مکان‌های گوناگون حضور CMBها [1] و پ) شبه CMBهای بسیاری

وجود دارد که فرآیند تشخیص را دشوار می‌کند.

در روش بالینی فعلی، CMBها به صورت دستی برچسب‌گذاری

می‌شوند [4]. برچسب‌گذاری دستی دشوار، وقت‌گیر و مستعد خطا

۱- مقدمه

تجمع‌های بسیار کوچک از خون در مغز (CMB^۱) بیشتر در بیماران

مبتلا به سکته مغزی، زوال عقل، بیماری‌های مغزی-عروقی یافت می‌-

شود [1]. تشخیص وجود CMBها می‌تواند به پیش‌بینی بروز سکته

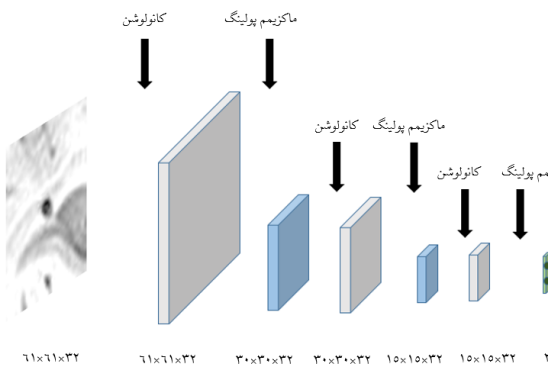
مغزی کمک کند. علاوه بر این، در سال‌های اخیر مطالعات بالینی نشان

می‌دهد که CMB می‌تواند منجر به اختلال شناختی مانند زوال عقل

شود [2]. تصویربرداری تشدید مغناطیسی به دلیل مزایایی که نسبت به

تصویربرداری‌های دیگر دارد، برای آشکارسازی CMBها استفاده می‌-

ارزیابی ساختار شبکه انتخاب شد. ساختار شبکه CNN پیشنهادی در شکل (۲) نشان داده شده است.



شکل ۲: ساختار شبکه CNN پیشنهادی

در ادامه، شبکه عصبی کانولوشن و پیش‌پردازش به عنوان پیش‌نیاز شرح داده شد.

۲-۱- شبکه عصبی کانولوشن

شبکه عصبی کانولوشن نوعی از شبکه‌های عصبی عمیق است. مدل یادگیری آن با ناظر و وزن‌ها با الگوریتم پس‌انتشار خطا^۷ اصلاح می‌شوند. شبکه عصبی کانولوشن از چندین لایه گوناگون تشکیل شده است: لایه کانولوشن، لایه پولینگ و لایه تماماً که در ادامه به صورت مختصر بررسی می‌شوند.

• لایه کانولوشن

لایه کانولوشن اصلی‌ترین لایه‌ی شبکه است که شامل مجموعه‌ای از فیلترهای قابل یادگیری است. هر فیلتر از لحاظ مکانی کوچک بوده اما در امتداد عمق توده ورودی ادامه پیدا می‌کند. به بیان ساده‌تر، فیلتر یک توده سه بعدی است که می‌توان به تعداد دلخواه فیلتر تعریف شود. اما باید توجه شود که تعداد زیاد فیلتر بار محاسباتی را افزایش می‌دهد. تابع فعال‌ساز که روی خروجی لایه کانولوشن اعمال می‌گردد می‌تواند خطی یا غیرخطی باشد. توابع غیرخطی برای جدا کردن داده‌هایی که به صورت خطی قابل جدا شدن نیستند استفاده می‌شود. در این مقاله توابع ReLU، Leaky ReLU و PReLU به عنوان تابع فعال‌ساز بررسی شده‌اند که به ترتیب با رابطه (۱)، (۲) و (۳) تعریف می‌شوند.

$$y = \max(0, x) \quad (1)$$

$$y = \begin{cases} 0.01x & x < 0 \\ x & x > 0 \end{cases} \quad (2)$$

$$y = \begin{cases} ax & x < 0 \\ x & x > 0 \end{cases} \quad (3)$$

• لایه پولینگ

لایه پولینگ معمولاً بعد از یک لایه کانولوشن قرار می‌گیرد و با بیشینه‌گیری^۸ یا میانگین‌گیری^۹، تعداد پیکسل‌های تصویر در عرض و ارتفاع

است. از این رو الگوریتم‌های خودکار و نیمه‌خودکار بسیاری ایجاد شدند. الگوریتم‌های موجود تعداد مثبت غلط (FP^۲) زیادی دارند که باعث کاهش ارزش آن‌ها و ادامه تحقیق در این زمینه می‌شود. در ادامه به بررسی چند نمونه از کارهای پیشین پرداخته شده است.

در مرجع [5] روشی نیمه خودکار برای آشکارسازی CMBها ارائه شده است. در این مقاله، ابتدا از یک الگوریتم آستانه گذاری برای جداسازی CMB از شبه CMBها و بعد از آن از طبقه‌بند ماشین بردار پشتیبان (SVM^۳) بهره گرفته و در انتها پس‌پردازش دستی انجام شد. حساسیت بعد از پس‌پردازش دستی به % ۸۱/۷ رسید. در مرجع [6] از طبقه‌بند جنگل تصادفی استفاده شد و به حساسیت % ۹۲/۰۴ رسید. مرجع [7] براساس مقدار شدت و طبقه‌بند جنگل تصادفی کاندیدهایی را به عنوان CMB در نظر گرفته است. سپس ویژگی‌های استخراجی با استفاده از شبکه ISA^۵ و طبقه‌بند ماشین بردار پشتیبان با حساسیت % ۸۹/۴۴ طبقه‌بندی شده است. در مرجع [8] ساختاری جدید از شبکه‌ی عصبی کانولوشن با لایه‌ی RBAP^۶ ارائه شد و به دقت % ۹۷/۱۸ دست یافت. مرجع [9] با به کارگیری شبکه عصبی کانولوشن بهینه به حساسیت % ۹۹/۷۴ دست یافت. در مرجع [10] یک روش مبتنی بر ResNet-50 پیشنهاد و حساسیت % ۹۵/۷۱ گزارش شد.

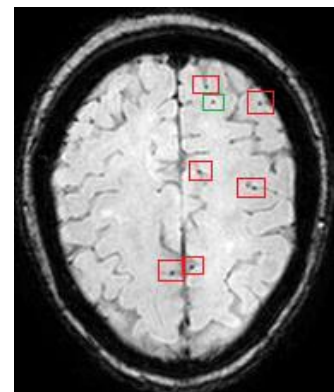
این مقاله، با آزمایش‌های مختلف بر روی پارامترهای قابل تنظیم شبکه عصبی کانولوشن، سعی در بهبود پارامترهای ارزیابی برای آشکارسازی CMBها دارد. سرانجام شبکه‌ای با سه لایه کانولوشن، دو لایه پولینگ و یک لایه تماماً متصل پیشنهاد شده است.

مقاله پیش‌رو در چهار بخش سازماندهی شده است: در بخش دوم، الگوریتم پیشنهادی مطرح شد. در بخش سوم، آزمایش‌ها و ارزیابی روش پیشنهادی ارائه و در بخش چهارم به جمع‌بندی کار

شکل ۱: اسکن SWI از جمجه

CMB و شبه CMBها به ترتیب با مستطیل سبز و قرمز نشان داده شده است.

پرداخته شده است.



۲- الگوریتم پیشنهادی

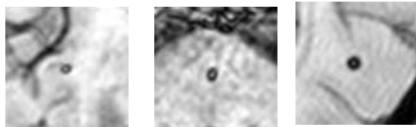
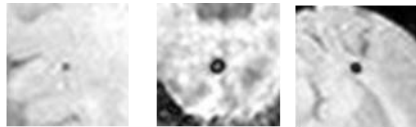
به منظور حل مساله آشکارسازی CMB مبتنی بر تصاویر SWI، استفاده از یک شبکه CNN پیشنهاد شد. الگوریتم پیشنهادی شامل یک گام پیش‌پردازش و یک شبکه عصبی کانولوشن است. برای بهبود عملکرد شبکه‌ی CNN آزمایش‌هایی انجام شد تا عواملی که می‌تواند عملکرد CNN را تحت تاثیر قرار دهد آزمایش شود و براساس نتایج پارامترهای

به عنوان داده‌های اعتبارسنجی برای آزمایش مدل حفظ می‌شود و بقیه داده‌ها به عنوان داده‌های آموزش استفاده می‌شوند. پس از آن از نتایج برای تولید یک برآورد واحد، میانگین گرفته می‌شود.

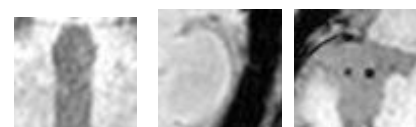
۳-۱- آماده‌سازی داده

در این مقاله، از ۲۰ حجم تصویر پایگاه داده‌ی SWI-CMB که در مرجع [11] در دسترس عموم قرار دارد استفاده شده است. اندازه اکثر تصاویر $150 \times 152 \times 512$ است. با جدا کردن و ذخیره اسلایس‌ها از حجم تصاویر، ۲۹۸۲ تصویر ایجاد شد که در ۶۳ تصویر از این مجموعه CMB وجود دارد. برای ایجاد پیچ‌های ورودی و برچسب‌ها از SNP^{11} [12] با پنجره لغزان 61×61 استفاده شد. دلیل انتخاب اندازه‌ی 61×61 برای پیچ این است که با این اندازه اطلاعات مفید برای آشکارسازی حفظ می‌شود [9]. به پیکسل‌های مرکزی پیچ که شامل CMB بود برچسب یک و غیر CMB برچسب صفر داده شد. شکل (۳) و شکل (۴) به ترتیب نمونه پیچ‌هایی با برچسب یک و صفر را نشان می‌دهد.

از آنجا که در بعضی از تصاویر CMB، چندین CMB وجود داشت، بعد از جدا کردن این CMB‌ها تعداد پیچ‌های CMB به ۷۴ پیچ و با الحاق داده‌های افزونه حاصل از قرینه کردن^{۱۱}، چرخش^{۱۲} و جابه‌جایی به ۱۷۷۶ پیچ رسید. از تصاویر غیر CMB هم، ۱۵۰۰ پیچ غیر CMB تولید شد. در نهایت تعداد کل پیچ‌ها برای آموزش و ارزیابی شبکه برابر ۳۲۷۶ پیچ شد که به صورت تصادفی ۹۰٪ آن برای آموزش و ۱۰٪ آن برای ارزیابی استفاده شد. یادآوری این نکته ضروری است که داده‌های افزونه مربوط به داده‌های آزمایش، از مجموعه آموزش حذف شده‌اند.



شکل ۳: نمونه پیچ‌های CMB



کاسته می‌شود، به دنبال آن تعداد پارامترها و پیچیدگی محاسبات کم می‌شود. عملیات پولینگ در هر عمق به صورت مستقل انجام می‌شود.

• لایه تماما متصل

همان‌گونه که از نام این لایه مشخص است، تمام نورون‌های این لایه به نورون‌های لایه قبل متصل است. لایه تماما متصل شبکه، کل ویژگی‌ها را ترکیب می‌کند تا تصویر ورودی را طبقه‌بندی کند.

۲-۲- پیش‌پردازش

در گام پیش‌پردازش مطابق با رابطه (۴) از نرمال سازی استفاده شده است، به گونه‌ای که اندازه بیشینه پیکسل به یک محدود شود. هدف از نرمال‌سازی این است که پیکسل‌های ورودی دارای توزیع مشابه شوند تا همزمان با آموزش شبکه، همگرایی سریع‌تر شود.

$$X = \frac{X0}{\max(X0)} \quad (4)$$

لایه‌ها	اندازه فیلتر	تعداد فیلتر	ابعاد خروجی
۱ پیچ ورودی پیش‌پردازش شده	-	-	(۶۱,۶۱,۱)
۲ کانولوشن+LeakyReLU	۱۱×۱۱	۳۲	(۶۱,۶۱,۳۲)
۳ ماکزیمم پولینگ	-	-	(۳۰,۳۰,۳۲)
۴ کانولوشن+LeakyReLU	۳×۳	۳۲	(۳۰,۳۰,۳۲)
۵ ماکزیمم پولینگ	-	-	(۱۵,۱۵,۳۲)
۶ کانولوشن+LeakyReLU	۳×۳	۳۲	(۱۵,۱۵,۳۲)
۷ تماما متصل	-	-	۲

X و $X0$ به ترتیب پیچ ورودی و پیچ نرمالیزه شده است.

۲-۳- ساختار شبکه پیشنهادی

شبکه‌ی CNN پیشنهادی از ۶ لایه تشکیل شده است. جدول (۱) لایه‌های شبکه عصبی کانولوشن پیشنهادی را نشان می‌دهد. آزمایش‌های انجام شده، نشان می‌دهد که تابع فعال‌ساز Leaky-ReLU لایه پولینگ مبتنی بر بیشینه به نتایج بهتری منجر می‌شود، از این رو در شبکه پیشنهادی قرار داده شدند. در پایان لایه تماما متصل قرار داده شد تا براساس الگوریتم آموزش، تصویر ورودی به عنوان CMB یا غیر CMB طبقه‌بندی شود.

جدول ۱: لایه‌های شبکه پیشنهادی

۳- آزمایش‌ها و ارزیابی

در این بخش از مقاله، پایگاه داده برای آموزش شبکه، معیارهای ارزیابی و نتایج بدست آمده معرفی می‌شوند. پیاده‌سازی‌ها در زبان برنامه‌نویسی پایتون بر روی Google Colab انجام شده است.

برای اعتبارسنجی در این مقاله از روش اعتبارسنجی متقابل ۱۰ پوشه‌ای استفاده شده است. در این روش، پایگاه داده به‌طور تصادفی به ۱۰ زیرمجموعه تقسیم می‌شود. در هر زیرمجموعه، تعدادی داده

شکل ۴: نمونه پچ‌های غیر CMB

(۲) پیدا است چپش لایه‌ها به ترتیب به صورت Conv/PI/Conv/PI/Conv انتخاب مناسبی خواهد بود.

جدول ۲: تاثیر چپش لایه‌ها بر کارایی شبکه CNN

آزمایش سه: تابع فعال‌سازی و پولینگ

ترکیب‌های مختلف تابع فعال‌سازی و پولینگ به عملکردهای مختلف شبکه منجر می‌شود، از این رو لازم است تا ترکیب‌های مختلف مورد

متوسط حساسیت (%)	متوسط ویژگی (%)	متوسط دقت (%)	لایه‌ها
۷۵	۹۷/۴۶	۹۶/۳۲	Conv
۷۷/۵	۹۸/۴	۹۷/۳۴	PI/Conv
۸۱/۲۵	۹۸/۲	۹۷/۳۴	Conv/PI/Conv
۹۵	۹۸/۱	۹۸/۱۰	PI/Conv/PI/Conv
۹۶/۲۵	۹۸/۲	۹۸/۳	Conv/PI/Conv/PI/Conv
۹۵	۹۸/۲	۹۸/۳۵	PI/Conv/PI/Conv/PI/Conv
۹۵	۹۸/۳۳	۹۸/۳	Conv/PI/Conv/PI/Conv/PI/Conv

آزمایش قرار گیرد. الگوریتم به ازای هر ترکیب از تابع فعال‌سازی و پولینگ ۱۰ بار اجرا شد و مقادیر متوسط معیارهای ارزیابی در جدول (۳) نشان داده شده است. با توجه به جدول (۳)، کارایی لایه پولینگ مبتنی بر میانگین کمتر از لایه پولینگ مبتنی بر بیشینه است.

جدول ۳: تاثیر ترکیب‌های مختلف تابع فعال‌سازی و پولینگ بر کارایی شبکه CNN

متوسط حساسیت (%)	متوسط ویژگی (%)	متوسط دقت (%)	پولینگ مبتنی بر	تابع فعال‌سازی
۹۶/۲۵	۹۸/۸	۹۸/۶۷	بیشینه	ReLU
۹۳/۷۵	۹۷/۶	۹۷/۴	بیشینه	PReLU
۹۸/۷۵	۹۹/۲۶	۹۹/۲۴	بیشینه	LeakyReLU
۹۲/۵	۹۸/۲	۹۷/۹	میانگین	ReLU
۸۶/۲۵	۹۶/۵۳	۹۶/۰۱	میانگین	PReLU
۹۸/۷۵	۹۷/۸	۹۷/۸۴	میانگین	LeakyReLU

آزمایش چهار: اندازه و تعداد فیلتر در لایه کانولوشن

در لایه کانولوشن اول، برای دستیابی به بهترین عملکرد تشخیص CMBها تعداد و اندازه‌های مختلفی از فیلتر متناظر در این لایه مورد بررسی قرار گرفت، الگوریتم به ازای هر حالت ۱۰ بار اجرا شد و مقادیر میانگین پارامترهای ارزیابی در جدول (۴) و (۵) آورده شده است.

متوسط حساسیت (%)	متوسط ویژگی (%)	متوسط دقت (%)	اندازه فیلتر
۹۸/۷۵	۹۹	۹۸/۹۸	۳
۹۹	۹۹/۴۶	۹۹/۴۳	۵
۹۹	۹۹	۹۹/۰۵	۷
۹۸/۷۵	۹۶/۰۶	۹۹/۰۵	۹
۹۷/۵	۹۶/۰۶	۹۸/۹۸	۱۱

بدین ترتیب همان‌گونه که از جدول (۴) و (۵) پیدا است، اندازهی فیلتر برابر با

۵ و تعداد فیلتر برابر با ۳۲ تنظیم شد.

جدول ۴: تاثیر اندازهی فیلتر لایه کانولوشن اول بر کارایی شبکه CNN

۲-۳- پارامترهای ارزیابی

برای ارزیابی نتایج طبقه‌بندی از سه شاخص یعنی حساسیت، ویژگی و دقت طبق رابطه (۵)، (۶) و (۷) استفاده شد. در این مطالعه، مثبت و منفی به ترتیب CMB و غیر CMB را نشان می‌دهد.

$$sensitivity = \frac{TP}{(TP + FN)} \quad (5)$$

$$specificity = \frac{TN}{(TN + FP)} \quad (6)$$

$$accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \quad (7)$$

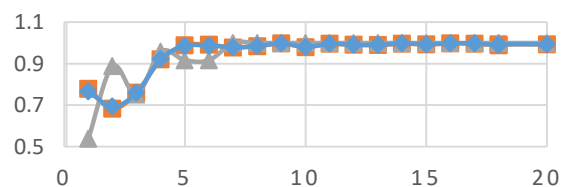
۳-۳- آزمایش‌ها

برای رسیدن به ساختار مطلوب و ارزیابی الگوریتم پیشنهادی آزمایش‌هایی انجام شد. در این آزمایش‌ها از بهینه‌ساز Adam که به حافظه کمی نیاز دارد و همگرایی آن سریع است استفاده شد. تابع هزینه در این مقاله تابع انتروپی متقابل در نظر گرفته شده است. وزن‌های اولیه CNN به صورت تصادفی و با اندازه‌ی بسته‌ای^{۱۳} برابر با ۱۲۸ تنظیم شد. در لایه تماماً متصل نیز از تابع softmax استفاده شده است.

آزمایش یک: تنظیم تعداد اپاک الگوریتم

در این آزمایش، الگوریتم به ازای تعداد اپاک‌های مختلف ۱۰ بار اجرا شد و میانگین معیارهای ارزیابی به ازای ۱۰ بار اجرا در شکل (۵) نشان داده شده است. همان‌گونه که از شکل (۵) پیدا است تعداد اپاک برابر با ۱۱ انتخاب مناسبی خواهد بود.

شکل ۵: نمودار نتایج حاصل از پارامترهای ارزیابی با تغییر تعداد اپاک



شکل ۵: نمودار نتایج حاصل از پارامترهای ارزیابی با تغییر تعداد اپاک

آزمایش دو: تنظیم تعداد لایه‌ها

تحقیقات نشان می‌دهد که تعداد لایه می‌تواند بر عملکرد CNN تاثیر بگذارد، به خصوص لایه‌های کانولوشن که برای استخراج ویژگی استفاده می‌شود. میانگین نتایج حاصل از ۱۰ بار اجرا برای چپش‌های مختلف لایه‌ها در جدول (۲) آورده شده است. همان‌گونه که از جدول

حساسیت (%)	۹۶/۹۴	۹۹/۷۴	۹۵/۷۱	۹۹/۹
------------	-------	-------	-------	------

۴- جمع‌بندی

هدف از این مقاله آشکارسازی خودکار خونریزی مغزی است. از آنجا که پردازش دستی این موضوع، دشوار، وقت‌گیر و مستعد خطا است، ارائه روشی برای خودکار کردن پردازش مربوطه ضروری است. استخراج ویژگی‌های تصویر، مرحله اصلی و مهم این پردازش است که در این مقاله از توانایی بالای شبکه عصبی کانولوشن استفاده شده است. به منظور کاهش پارامترهای شبکه و همچنین حفظ دقت کافی جهت طبقه‌بندی از شبکه عصبی کانولوشن ۶ لایه با ابعاد کوچک استفاده شده است. آزمایش‌های انجام شده بر روی پایگاه داده تصاویر SWI، نشان می‌دهد که روش پیشنهادی در مقایسه با روش‌های مرجع، نتایج بهتری را به دنبال دارد.

مراجع

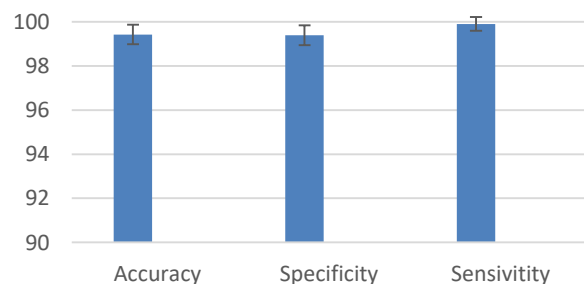
- [1] Wu Y, Chen T, "An Up-to-Date Review on Cerebral Microbleeds", *Journal of Stroke and Cerebrovascular Diseases*, Vol. 25, No. 6, pp.1301-1306, 2016.
- [2] Charidimou A, Krishnan A, Werring DJ, Jäger HR, "Cerebral microbleeds: a guide to detection and clinical relevance in different disease settings." *Neuroradiologie Scan*, Vol.4, No.02, pp 123-147, 2014.
- [3] Shams S, Martola J, Cavallin L, Granberg T, Shams M, Aspelin P, Wahlund LO, Kristoffersen-Wiberg M, "SWI or T2*: which MRI sequence to use in the detection of cerebral microbleeds? The Karolinska Imaging Dementia Study." *American Journal of Neuroradiology*, Vol. 36, No.6, pp 1089-109, 2015.
- [4] De Bresser J, Brundel M, Conijn MM, Van Dillen JJ, Geerlings MI, Viergever MA, Luijten PR, Biessels GJ, "Visual cerebral microbleed detection on 7T MR imaging: reliability and effects of image processing." *American Journal of Neuroradiology*, Vol.34, No.6, pp E61-E64, 2013.
- [5] Barnes SR, Haacke EM, Ayaz M, Boikov AS, Kirsch W, Kido D, "Semiautomated detection of cerebral microbleeds in magnetic resonance images." *Magnetic resonance imaging*, Vol.29, No.6, pp 844-852, 2011.
- [6] Fazlollahi A, Meriaudeau F, Villemagne VL, Rowe CC, Yates P, Salvado O, Bourgeat P, "Efficient machine learning framework for computer-aided detection of cerebral microbleeds using the radon transform." *2014 IEEE 11th international symposium on biomedical imaging (ISBI)*, 2014.
- [7] Dou Q, Chen H, Yu L, Shi L, Wang D, Mok VC, Heng PA, "Automatic cerebral microbleeds detection from MR images via independent subspace analysis based hierarchical features." *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 7933-7936, 2015.
- [8] Wang S, Jiang Y, Hou X, Cheng H, Du S, "Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling." *IEEE Access*, pp 16576-16583, 2017.
- [9] Hong J, Wang SH, Cheng H, Liu J, "Classification of cerebral microbleeds based on fully-optimized convolutional neural network." *Multimedia Tools and Applications*, pp.1-19, 2018.
- [10] Hong J, Cheng H, Zhang YD, Liu J, "Detecting cerebral microbleeds with transfer learning." *Machine Vision and Applications*, pp. 1-11, 2019.
- [11] Dou Q, Chen H, Yu L, Zhao L, Qin J, Wang D, Mok VC, Shi L, Heng PA, "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks." *IEEE transactions on medical imaging*, Vol.35, No.5, pp 1182-1195, 2016.
- [12] Zhang YD, Hou XX, Chen Y, Chen H, Yang M, Yang J, Wang SH, "Voxelwise detection of cerebral microbleed in CADASIL patients

جدول ۵: تاثیر تعداد فیلتر لایه کانولوشن اول با ابعاد ۵×۵ بر کارایی شبکه CNN

تعداد فیلتر	متوسط دقت (%)	متوسط ویژگی (%)	متوسط حساسیت (%)
۸	۹۸/۴۸	۹۸/۷۳	۹۳/۷۵
۱۶	۹۹/۱۷	۹۹/۱۳	۹۸/۹۱
۳۲	۹۹/۲۲	۹۹/۲	۹۹/۹
۶۴	۹۹/۲۴	۹۹/۲	۹۹/۸
۱۲۸	۹۸/۹۲	۹۸/۸۶	۹۹/۹۳

آزمایش پنج: کارایی شبکه پیشنهادی

به منظور بهبود عملکرد طبقه بندی CMBها، آزمایش‌هایی انجام شد تا تقریباً تمام عواملی را که می‌تواند عملکرد CNN را تحت تاثیر قرار دهد آزمایش شود. در نهایت، یک شبکه عصبی کانولوشن که شامل سه لایه کانولوشن و دو لایه پولینگ و یک لایه کاملاً متصل باشد به کارایی بهتری دست یافت. برای جلوگیری از تصادفی بودن نتایج ارزیابی ۱۰ بار اجرا شد. میانگین و انحراف از میانگین در نمودار شکل (۶) نشان داده شده است.



شکل ۶: نتایج ارزیابی در ۱۰ بار اجرای شبکه پیشنهادی

کارایی شبکه پیشنهادی با روش‌های دیگر در جدول (۶) نشان داده شده است. همان‌گونه که دیده می‌شود، روش پیشنهادی در مقایسه با روش‌های دیگر نتایج بهتری دارد.

جدول ۶: مقایسه روش پیشنهادی با مقالات پیشین

سال	[8]	[9]	[10]	شبکه پیشنهادی
سال	۲۰۱۷	۲۰۱۸	۲۰۱۹	۲۰۱۹
فرمت تصاویر	SWI	SWI	SWI	SWI
دقت (%)	۹۷/۱۸	۹۸/۳۲	۹۷/۴۶	۹۹/۴۱
ویژگی (%)	۹۷/۱۸	۹۶/۸۹	۹۹/۲۱	۹۹/۳۹

کنفرانس پردازش سیگنال و سیستم‌های هوشمند، ۲۷ الی ۲۸ آذرماه ۱۳۹۸، دانشگاه صنعتی شاهرود

by leaky rectified linear unit and early stopping." Multimedia Tools and Applications, Vol. 77, No. 17, pp 21825-21845, 2018.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و

سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه تبریز

جداسازی طیفی تصاویر ابرطیفی بر اساس ساختار تنک بلوکی

سید حسین موسوی آذرنگ^۱، روزبه رجبی^۲، هادی زبانی^۳

^۱ دانشجوی کارشناسی ارشد، گروه مخابرات، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی قم، قم،
h.mosaviazarang@yahoo.com

^۲ استادیار، گروه مخابرات، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی قم، قم،
rajabi@qut.ac.ir

^۳ استادیار، گروه مخابرات، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی قم، قم،
استادیار، بخش مخابرات و الکترونیک، دانشکده مهندسی برق و کامپیوتر، دانشگاه شیراز، شیراز،
zayyani@qut.ac.ir

چکیده

جداسازی طیفی داده‌های سنسجش از دور یکی از مسائل مهم، پیش از بهره برداری از این داده‌ها در زمینه‌های مختلف علمی و کاربردی است. با وجود قدرت تفکیک طیفی بالا در تصاویر ابرطیفی، به دلیل توان تفکیک مکانی پایین سنجنده‌ها، ممکن است مواد مختلفی درون یک پیکسل باشند. در این صورت، طیف پیکسل‌های ثبت شده توسط سنجنده، ترکیبی از طیف مولفه‌های اصلی موجود در آن پیکسل‌ها است. جداسازی طیفی، روشی است که مشخصه طیفی مولفه‌های اصلی پیکسل‌های مخلوط را استخراج کرده و سپس فراوانی‌های نسبی هر عضو پایانی را به دست می‌آورد. تجزیه به ماتریس‌های غیرمنفی یکی از روش‌های متداول در زمینه جداسازی طیفی است. این مقاله بر اساس ساختار تنک بلوکی توسعه داده شده است و سپس با الهام از راهبرد یادگیری بیزی تنک، الگوریتم پیشنهادی با استفاده از قید فراوانی مجموع به یک و قید فراوانی‌های نامنفی ارائه شده است. در انتها برای بررسی عملکرد، الگوریتم پیشنهادی بر روی داده‌های ساختگی و واقعی اجرا شده است برای مقایسه کمی نتایج از دو معیار AAD و MSE که در مطالعات جداسازی متداول هستند، استفاده شده است که معیار AAD با کاهش هشت درصدی و معیار MSE با کاهش چهار درصدی، نشان دهنده برتری الگوریتم پیشنهادی در مقایسه با سایر روش‌های پیشین می‌باشد.

کلمات کلیدی

تصاویر ابرطیفی، جداسازی طیفی، فراوانی نسبی، تنک بلوکی، یادگیری بیزی تنک.

۱- مقدمه

در رابطه با نوع مخلوط طیف‌های الکترومغناطیسی، فرض‌های متفاوتی وجود دارد بر مبنای این فرض‌ها، مدل‌های مخلوط متفاوتی مطرح شده‌اند، که در دودسته‌ی عمده‌ی خطی و غیرخطی جای می‌گیرند:

مدل مخلوط خطی: طیف‌های اندازه‌گیری شده را می‌توان به عنوان ترکیبی خطی از امضاهای طیفی مواد موجود در پیکسل مختلط بیان کرد.

مدل مخلوط غیرخطی: در این مدل، طیف ثبت شده در هر پیکسل یک تابع غیرخطی از طیف عناصر اصلی است که پرتوهای رسیده به سنجنده حاصل از بازتاب‌های چندگانه و تداخل‌های چندگانه بین مواد است [۵].

فرض دسته‌ای از روش‌های جداسازی طیفی وجود پیکسل خالص در تصویر برای هر یک از عناصر اصلی است به طوری که سیگنال ثبت شده در آن‌ها تنها شامل طیف فرکانسی آن عنصر است که روش VCA^۱ بر این فرض بیان شده است [۶]. تجزیه به ماتریس نامنفی برای حل مساله جداسازی منابع کور با الگوریتم تکراری عمل می‌کند که تعداد تکرار زیاد مانع از نتایج بهینه می‌گردد. برای بهبود عملکرد روش تجزیه به ماتریس نامنفی از قید تنگی استفاده می‌شود. به دلیل اینکه تعداد عناصر غیرصفر ماتریس فراوانی خیلی کم است و این به واقعیت نزدیک‌تر است [۷]. در جداسازی طیفی مدل نیمه نظارتی، فرض می‌شود که در تصویر مشاهده شده، می‌توان امضای طیفی عناصر را از کتابخانه ابرطیفی به دست آورد. این امر بیانگر مساله رگرسیون تنک می‌باشد. برای حل این مساله از روش SUNSAL می‌توان بهره جست [۸]. در ادامه چند نمونه از روش‌های مبتنی بر ساختار تنک بیان می‌شود. هدف الگوریتم ISMA^۲ پیدا کردن مجموعه اعضای پایانی بهینه برای تخمین فراوانی‌های نسبی است [۹]. الگوریتم SMP^۳ یکی دیگر از روش‌های تنک می‌باشد که برای حل مسائل جداسازی منابع کور به کار گرفته می‌شود [۱۰]. الگوریتم دیگر RSU^۴ است که از خطای میانگین مربعات برای جداسازی طیفی استفاده می‌کند [۱۱]. در روش‌های پیشین جداسازی طیفی نویز گوسی با مقدار متوسط برای هر باند بیان می‌گردید اما در روش SUBM^۵ از ایده تفاوت نویز گوسی موجود در هر باند برای جداسازی طیفی تصاویر ابرطیفی استفاده می‌شود [۱۲]. در روش SDW^۶ هدف استفاده از اطلاعات مکانی لبه‌ها در هر باند تصویر ابرطیفی و سپس وزن‌دار کردن این اطلاعات با استفاده از فیلتر سوبل برای جداسازی تنک می‌باشد [۱۳].

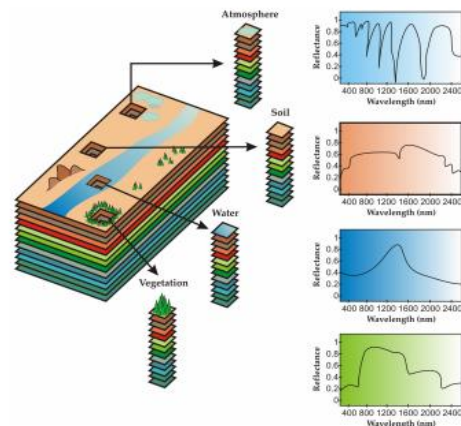
در این مقاله از راهبرد بیزی تنک و ساختار تنک بلوکی به منظور جداسازی طیفی تصاویر ابرطیفی استفاده شده است. اساس راهبرد یادگیری بیزی تنک استفاده از اطلاعات فوق پارامتر منتسب به هر پیکسل می‌باشد. در روش پیشنهادی با بهره‌گیری از ساختار تنک بلوکی موجود در داده‌های ابرطیفی علاوه بر اطلاعات فوق پارامتر منتسب به

یکی از مهم‌ترین کاربردهای فناوری سنجنش از دور تشخیص ویژگی از تصاویر تهیه شده توسط سنجنده‌های مستقر بر روی سکوها هوارد و فضایی می‌باشد [۱]. نتایج حاصله از این کار منجر به تولید اطلاعاتی با اعتماد پذیری بالا و اتخاذ تصمیماتی با ریسک کمتر توسط تصمیم گیرندگان می‌شود. یکی از جدیدترین فناوری‌های مطرح در سنجنش از دور، سنجنش از دور ابرطیفی است. سنجنده‌های ابرطیفی قابلیت ثبت انرژی الکترومغناطیسی بازتابی از سطح اشیاء را در باندهای طیفی با عرض‌های بسیار باریک دارا هستند.

فناوری سنجنش از دور در داده‌های ابرطیفی در دو دهه گذشته پیشرفت چشمگیری داشته است. حسگرهای فعلی روی سیستم عامل‌های موجود در هوافضا مناطق زیادی از سطح زمین را با وضوح طیفی، مکانی و زمانی بی‌سابقه‌ای پوشش می‌دهد. این ویژگی‌ها تعداد بی‌شماری از برنامه‌ها را که نیاز به شناسایی دقیق مواد یا برآورد پارامترهای فیزیکی دارند، قادر می‌سازد. خیلی اوقات، این برنامه‌ها به روش‌های پیچیده تجزیه و تحلیل داده‌ها متکی هستند [۲].

اطلاعات طیفی موجود در ابزارهای پیشرفته تصویربرداری ابرطیفی که در حال حاضر در حال بهره برداری هستند، دیدگاه‌های جدیدی را در بسیاری از زمینه‌های کاربردی همچون: نظارت بر فرایندهای زیست محیطی و شهری یا پیشگیری و رعایت عوامل مخرب، از جمله؛ تشخیص آب و هوا، تشخیص تهدیدات زیست محیطی، نظارت بر نفت و سایر انواع آلودگی شیمیایی پدید آورده‌اند.

ابزارهای پیشرفته ابرطیفی مانند هوافضا ناسا اکنون قادر به پوشش دادن منطقه‌ی طول موج از ۰٫۴ تا ۲٫۵ میکرومتر با استفاده از بیش از ۲۰۰ کانال طیفی با وضوح طیفی ۱۰ نانومتر است [۳]. مکعب داده اطلاعات ابرطیفی پشته‌ای از تصاویر است که در آن هر پیکسل (بردار) با یک امضای طیفی ارائه می‌شود که مواد تشکیل دهنده در آن را نمایش می‌دهد [شکل (۱) را ملاحظه کنید]. اگر قدرت تفکیک مکانی حسگر به اندازه کافی برای جدا کردن مواد مختلف مناسب نباشد، می‌تواند به طور مشترک یک پیکسل واحد را اشغال کند. به عنوان مثال، احتمال دارد که پیکسل در یک منطقه رویشی در واقع شامل ترکیبی از پوشش گیاهی و خاک باشد.



شکل ۱: مفهوم تصویر ابرطیفی [۴]

¹ Vertex Component Analysis

² Iterative Spectral Mixture Analysis

³ Subspace Matching Pursuit

⁴ Robust Sparse Unmixing

⁵ Sparse Unmixing Bandwise Model

⁶ Spatial Discontinuity Weighted

طول سیگنال می‌باشد) سیگنال را بازیابی کرد. این درحالی است که تئوری شانون-نایکوئیست بیان می‌کند برای بازسازی کامل یک سیگنال باند محدود، باید نرخ نمونه برداری حداقل به اندازه دوبرابر حداکثر فرکانس موجود در سیگنال باشد؛ در حالی که در حسگری فشرده به تعداد نمونه (یا اندازه گیری) خیلی کمتری نیاز است [۱۴].

در بسیاری از مسائل عملی ضرایب صفر به صورت خوشه‌ای ظاهر می‌شوند که به این ساختار تنک بلوکی گفته می‌شود. تعدادی از روش‌های بازیابی سیگنال تنک از این ساختار بهره می‌گیرند. روش BOMP یکی از این روش‌ها است که با استفاده تکراری از الگوریتم EM بر پایه بیزی سعی در بازیابی سیگنال تنک دارد [۱۵]. روش دیگر، روش PCSBL [۱۶] است که در بخش ۴ توضیح داده شده است. تصاویر ابرطیفی به دلیل همسایگی مواد مشابه از ساختار تنک بلوکی برخوردارند. در این مقاله بر اساس روش‌های بازیابی تنک بلوکی الگوریتم جدیدی برای جداسازی طیفی در بخش ۴ ارائه شده است.

۴- یادگیری بیزی تنک

مسئله بازیابی سیگنال تنک بلوکی $x \in R^n$ از بردار مشاهدات اندازه گیری شده از فرمول (۱) قابل بیان است. سیگنال x ساختار تنک بلوکی دارد اما الگوی دقیق بلوک مانند محل و اندازه هر بلوک برای ما قابل دسترس نیست. در چارچوب یادگیری بیزی تنک، برای تشویق تنکی سیگنال تخمین شده، x دارای توزیع پیشین گوسی است.

$$p(x|\alpha) = \prod_{i=1}^n p(x_i|\alpha_i) \quad (5)$$

که $\alpha \triangleq \{\alpha_i\}$ های پارامترهای غیرمنفی هستند که تنکی سیگنال x را کنترل می‌کند. واضح است که وقتی α_i به بی‌نهایت نزدیک می‌شود، هم ضریب مربوط x_i صفر می‌شود و هم ضرایب مربوط به $\{x_{i+1}, x_{i-1}\}$ صفر می‌شود. این اثر صفرشدن ضرایب، امکان بازیابی سیگنال‌های تنک بلوک را به روشی مطمئن‌تر فراهم می‌کند. از آنجا که احتمال وقوع خطاهای متوالی نسبت به ایجاد یک خطای واحد در هنگام بازسازی بسیار کمتر است، لذا صفر شدن ضرایب می‌تواند به بهبود قابلیت اطمینان بازیابی سیگنال تنک بلوک کمک کند. مدل یادگیری بیزی تنک پیشنهاد می‌دهد که در آن هر فوق‌پارامتر علاوه بر خود ضریب، به همسایگان خود نیز وابسته باشد. لذا اگر یکی از ضرایب شناسایی شود، این شناسایی در بهبود دو ضریب همسایه آن تأثیر خواهد گذاشت. از این نظر، به نظر می‌رسد در نظر گرفتن همسایگان، منعطف‌ترین چارچوب را برای توصیف سیگنال‌های تنک بلوکی ارائه می‌دهد [۱۶]. لازم به ذکر است یادگیری بیزی تنک از توزیع گاما به عنوان پیش فرض بر روی فوق‌پارامترهای $\{\alpha_i\}$ استفاده می‌کند.

۴-۱- الگوریتم پیشنهادی بیزی با واریانس نویز

شناخته

از روش یادگیری بیزی تنک در بازیابی سیگنال‌های تنک بلوکی می‌توان بهره جست. در این بخش برای سهولت فرض می‌شود که واریانس نویز شناخته شده و به صورت پیش فرض σ^2 است. لذا بر اساس یادگیری بیزی تنک، توزیع پسین x به صورت زیر بیان می‌شود:

هر پیکسل، از اطلاعات فوق پارامتر همسایگی‌های آن پیکسل نیز استفاده می‌شود که باعث بهبود عملکرد یادگیری بیزی تنک می‌گردد.

در بخش دوم مقاله به بیان ریاضی مساله پرداخته می‌شود. در بخش سوم ساختار سیگنال تنک که اساس پژوهش مقاله می‌باشد، بیان و بررسی می‌گردد. در بخش چهارم شرح الگوریتم پیشنهادی با استفاده از راهبرد یادگیری بیزی تنک بیان می‌شود. در بخش پنجم برای بررسی عملکرد، الگوریتم پیشنهادی بر روی داده ساختگی و واقعی با چند روش مقایسه می‌شود و نتایج آن بیان می‌شود. در بخش ششم نیز به جمع بندی و نتیجه گیری از ارائه روش پیشنهادی پرداخته می‌شود.

۲- بیان ریاضی مساله

در مدل مخلوط خطی فرض می‌شود که پاسخ طیفی یک پیکسل در هر باند طیفی، ترکیبی خطی از همه اعضای پایانی موجود در پیکسل در باند طیفی مربوطه است. برای هر پیکسل، مدل خطی را می‌توان بصورت زیر نوشت:

$$y_i = \sum_{j=1}^q a_{ij}x_j + n_i \quad (1)$$

که در آن y_i مقدار بازتاب اندازه گیری شده در باند طیفی i است، a_{ij} بازتاب j -امین عضو پایانی^۷ در باند طیفی i است، x_j فراوانی نسبی^۸ عضو پایانی j ام است، و n_i نشان دهنده خطا برای باند طیفی i است (نویز موثر بر فرایند اندازه گیری می‌باشد). اگر فرض شود که سنسور ابرطیفی جمع آوری داده‌ها در L باند طیفی قرار داشته باشد، می‌توان فرمول (۱) را به صورت ماتریسی فشرده به صورت زیر بازنویسی کرد:

$$y = Ax + n \quad (2)$$

که در آن y یک بردار ستونی $1 \times L$ (طیف اندازه گیری شده از پیکسل) است، A یک ماتریس $q \times L$ حاوی امضا طیفی (اعضای پایانی) است، x یک بردار $1 \times q$ که شامل فراوانی نسبی اعضای پایانی است، و n یک بردار $1 \times L$ جمع آوری اشتباهات موثر بر اندازه گیری‌ها در هر باند طیفی است. بردار فراوانی نسبی عناصر خالص در ترکیب هر پیکسل معمولاً با دو شرط زیر همراه است:

$$x \geq 0 \quad (3)$$

$$1^T x = 1 \quad (4)$$

که به ترتیب قید فراوانی غیر منفی^۹ (ANC) و قید فراوانی مجموع به یک^{۱۰} (ASC) هستند که به صورت فشرده نشان داده شده‌اند. 1^T بردار خطی می‌باشد که همه عناصر آن عدد یک می‌باشد. همچنین x_j بیانگر درصد حضور هر عضو پایانی در هر پیکسل را نمایش می‌دهد.

۳- نمایش تنک بلوکی سیگنال

حسگری یا نمونه‌برداری فشرده بر این اصل استوار است که اگر سیگنالی در یک پایه یا دیکشنری دارای نمایش تنک باشد، آن گاه می‌توان با تعدادی اندازه گیری از آن (که در حالت کلی خیلی کمتر از

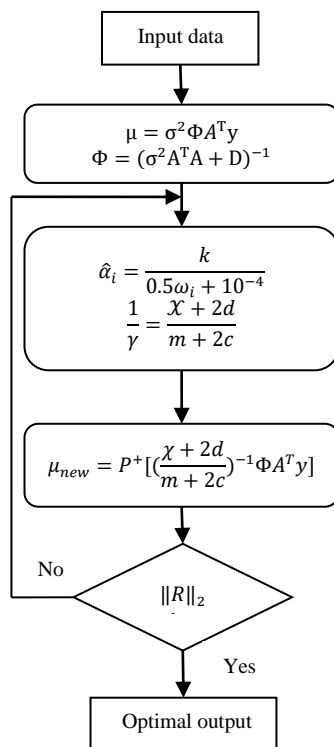
⁷ Endmember

⁸ Abundance fraction

⁹ Abundance Nonnegativity Constraint (ANC)

¹⁰ Abundance Sum Constraint (ASC)

به‌روزرسانی γ نیز همانند به‌روزرسانی دو پارامتر که در بخش قبل بیان شد، تا برآورده شدن معیار توقف بیان شده در فرمول (۱۳) ادامه می‌یابد. در شکل (۲) نمودار الگوریتم پیشنهادی ترسیم شده است.



شکل ۲: نمودار الگوریتم پیشنهادی

۵- داده‌ها

برای انجام آزمایش‌های ابر طیفی و الگوریتم پیشنهادی دو نوع داده استفاده می‌شود. داده‌های واقعی که با استفاده از طیف سنجهای شناخته شده ثبت شده و در دسترس هستند، و داده‌های ساختگی که با استفاده از کتابخانه‌های طیفی ساخته و آزمایش می‌شوند.

۵-۱- داده‌های ساختگی

تصاویر ساختگی، با استفاده از دسته‌ای از بازتاب‌های طیفی انتخاب شده از کتابخانه‌های طیفی، مانند کتابخانه‌ی طیفی دیجیتال USGS

$$p(x|\alpha, y) \propto p(x|\alpha)p(y|x) \quad (۶)$$

به راحتی می‌توان ثابت کرد که توزیع پسین $p(x|\alpha, y)$ دارای توزیع گوسی با میانگین و کواریانس به ترتیب صورت زیر می‌باشد:

$$\mu = \sigma^2 \Phi A^T y \quad (۷)$$

$$\Phi = (\sigma^2 A^T A + D)^{-1} \quad (۸)$$

که D ماتریس قطری است که عنصر قطری i -ام آن برابر است با

$$D \triangleq (\alpha_i + \beta \alpha_{i-1} + \beta \alpha_{i+1}) \quad (۹)$$

که β نشان‌دهنده این واقعیت است که وقتی $\beta_i > 0$ تنگی x_i نه تنها به فوق‌پارامتر α_i ، بلکه به همسایگی فوق‌پارامترهای $\{\alpha_{i+1}, \alpha_{i-1}\}$ نیز بستگی دارد. با توجه به مجموعه فوق‌پارامترهای تخمین زده شده $\{\alpha_i\}$ ، حداکثر پسین (MAP) از x به صورت زیر بیان می‌شود:

$$x_{MAP} = \mu = (\sigma^2 D + A^T A)^{-1} A^T y \quad (۱۰)$$

با توجه به \hat{x}_{MAP} ، فوق‌پارامترهای $\{\alpha_i\}$ با استفاده از الگوریتم حداکثر انتظار (EM) [۱۷] به صورت زیر تخمین زده می‌شود:

$$\hat{\alpha}_i = \frac{k}{0.5 \omega_i + 10^{-4}} \quad \forall i = 1, \dots, n \quad (۱۱)$$

که ω_i بیانگر میانگین و کواریانس وزنی به صورت زیر می‌باشد:

$$\omega_i = (\hat{\mu}_i^2 + \hat{\varphi}_{i,i}) + \beta (\hat{\mu}_{i+1}^2 + \hat{\varphi}_{i+1,i+1}) + \beta (\hat{\mu}_{i-1}^2 + \hat{\varphi}_{i-1,i-1}) \quad (۱۲)$$

همچنین k دارای محدوده‌ی $1 < k < 2.5$ می‌باشد.

به‌روز رسانی فوق‌پارامترهای α_i و همچنین \hat{x} ادامه می‌یابد تا زمانی که معیار توقف که در فرمول (۱۳) آمده است، محقق شود.

$$\|J_{new} - J_{old}\|_2 \leq \epsilon \quad (۱۳)$$

که J_{old} و J_{new} مقادیر تابع هزینه برای دو تکرار متوالی است و ϵ نیز در این مقاله برابر با 10^{-8} است.

۴-۲- الگوریتم پیشنهادی بیزی با واریانس نویز

ناشناخته

در قسمت قبلی فرض بر این بود که واریانس نویز شناخته شده می‌باشد. در این بخش به این حالت که عموماً در واقعیت نیز چنین است فرض بر این است که واریانس نویز شناخته شده نباشد. در این صورت فوق‌پارامتر جدیدی به نام γ معرفی می‌شود که بیانگر نویز واریانس می‌باشد:

$$p(\gamma) = \Gamma(c)^{-1} d^c \gamma^c e^{-d\gamma} \quad (۱۴)$$

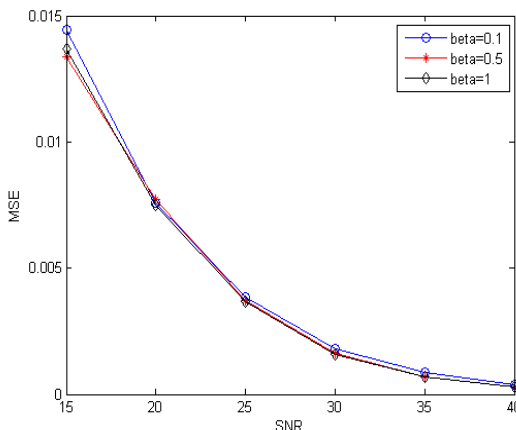
که $c = d = 10^{-4}$ می‌باشند. تنها تفاوت این بخش با حالت قبل، تخمین فوق‌پارامترهای α و واریانس نویز (یا معادل آن γ) است که باز هم الگوریتم (EM) بکارگرفته می‌شود. پس از بکارگیری الگوریتم (EM) تخمین α و تخمین γ باید صورت پذیرد که تخمین α همانند بخش قبلی است و تنها تفاوت در این بخش تخمین معادل نویز یعنی γ است که به صورت زیر انجام می‌شود:

$$\frac{1}{\gamma^{(\ell+1)}} = \frac{\|y - Ax\|_2^2 + (\gamma^{(\ell)})^{-1} \sum_{i=1}^n \rho_i + 2d}{m + 2c} \quad (۱۵)$$

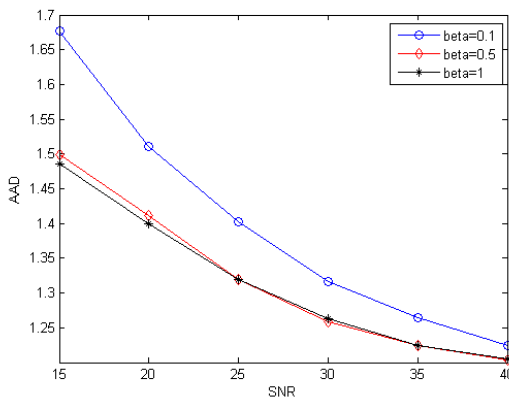
که $\rho_i \triangleq 1 - \Phi_{i,i}(\alpha_i^{(\ell)} + \beta \alpha_{i-1}^{(\ell)} + \beta \alpha_{i+1}^{(\ell)})$ می‌باشد و m نیز بیانگر تعداد باندهای طیفی می‌باشد.

داده تولید شده اضافه شده است. همچنین برای حالت های مختلف از پارامتر $\beta = [0.1, 0.5, 1]$ نیز الگوریتم پیشنهادی را اجرا نمودیم. ابتدا بر اساس پارامتر β نمودارهای ارزیابی براساس معیارهای MSE و AAD رسم شده‌اند.

همانطور که از هر دو شکل (۳) و (۴) پیداست با افزایش مقدار β هر دو نسبت میانگین خطای مربعات و فاصله بردار فروانی نسبی رو به کاهش سوق پیدا میکنند که نشان دهنده همگرایی الگوریتم پیشنهادی به حداقل خطای ممکن است. همچنین با افزایش مقدار SNR نیز هر دو نسبت میانگین خطای مربعات و فاصله بردار فروانی کاهش پیدا میکنند که با توجه به ذات SNR امری بدیهی است.



شکل ۳: میانگین خطای مربعات نسبت به β های متفاوت در محدوده SNR های مختلف در الگوریتم پیشنهادی



شکل ۴: فاصله بردار فروانی نسبی نسبت به β های متفاوت در محدوده SNR های مختلف در الگوریتم پیشنهادی

در شکل (۵) و (۶) نتایج الگوریتم پیشنهادی نسبت به دو معیار ارزیابی AAD و MSE با الگوریتمهای SUnSAL و RSU در SNR های مختلف مقایسه شده است.

تولید می‌شود انتخاب امضاهای اعضای پایانی که تا اندازه‌ای مستقل خطی باشند، دلخواه است. داده‌ی طیفی شامل ۲۲۴ باند طیفی است، که طول موج‌ها را از ۰.۳۸ تا ۲.۵ میکرومتر با دقت طیفی ۱۰ نانومتر پوشش می‌دهد. برای ساخت ترکیب‌های خطی، کل تصویر به واحدهای کوچک ۵×۵ تقسیم می‌شوند پیکسل‌های داخل هر بلوک خالص هستند و پوشش زمینی یکسانی دارند، که به صورت تصادفی از یک کلاس اعضای پایانی انتخاب شدند. سپس تصویر حاصل، برای شبیه سازی یک تصویر با پیکسل‌های آمیخته، از یک فیلتر پایین گذر عبور داده می‌شود. با این روش آمیختگی، یک صحنه‌ی ابرطیفی که در آن اعضای پایانی به صورت تکه‌های گسسته چیده شدند، با مدل آمیختگی خطی، مناسب است. فیلتر پایین گذر مورد استفاده فیلتر میانگین‌گیری ساده، $k \times k$ است، که مقدار k با درجه آمیختگی کنترل می‌شود. برای حذف مجدد پیکسل‌های خالص نیز، همه‌ی پیکسل‌هایی که فراوانی بیشتر از هشتاد درصد دارند، با یک مخلوط ساخته شده از همه‌ی پیکسل‌های با فراوانی یکسان، جایگزین شدند؛ یعنی هر عضو پایانی، فراوانی $1/c$ ترکیب را دارد، که منظور از $1/c$ ، تعداد اعضای پایانی است. برای شبیه سازی خطاها و نویزهای حسگر، نویز گوسی با میانگین صفر، به ترکیب اضافه می‌شود و فرض می‌شود که نویز از نظر فاصله‌ای و طیفی ناهمبسته است؛ بنابراین ماتریس کواریانس نویز برابر با $\sigma^2 I$ خواهد بود. پس نسبت سیگنال به نویز را می‌توان از $SNR = 10 \log_{10} (E[x^T x] / E[n^T n])$ محاسبه کرد.

برای مقایسه کمی نتایج، از دو معیار ارزیابی AAD^{۱۱} و MSE^{۱۲} استفاده شده است. این معیارها، دو معیار متداول در مطالعات جداسازی داده‌های ابرطیفی هستند که برای اندازه‌گیری میزان شباهت نتایج به دست آمده و مقادیر مرجع مورد استفاده قرار می‌گیرند. معیار AAD فاصله بردار فروانی می‌باشد که به صورت زیر بیان می‌شود [۱۸]:

$$AAD = \cos^{-1} \left(\frac{s^T \hat{s}}{\|s\| \|\hat{s}\|} \right) \quad (16)$$

s بردار فروانی واقعی یک پیکسل است و \hat{s} بردار فروانی تخمین شده می‌باشد.

معیار دیگر ارزیابی خطای میانگین مربعات است، که بر اساس معیار متداول خطای میانگین مربعات بین نتایج ثبت شده و تخمین زده شده به دست می‌آید [۱۹]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (17)$$

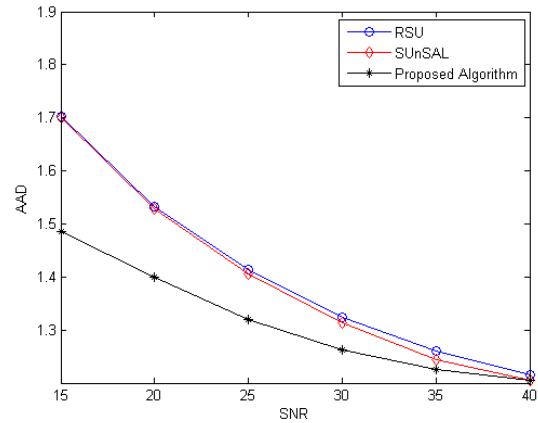
\hat{y}_i طیف تخمینی برای i امین پیکسل و y_i طیف دریافتی به وسیله حسگرهای سنجنده برای i امین پیکسل است.

الگوریتم پیشنهادی ابتدا بر روی داده‌های ساختگی اجرا شده است. برای تولید این مجموعه داده، دوازده امضای طیفی از کتابخانه USGS به صورت تصادفی با استفاده از فیلتر ۵×۵ و بدون پیکسل خالص انتخاب شده است. سپس نویز گوسی با میانگین صفر در ۶ مقدار سیگنال به نویز متفاوت یعنی $SNR = [15, 20, 25, 30, 35, 40]$ به

¹¹ Abundance Angel Distance (AAD)

¹² Mean Square Error (MSE)

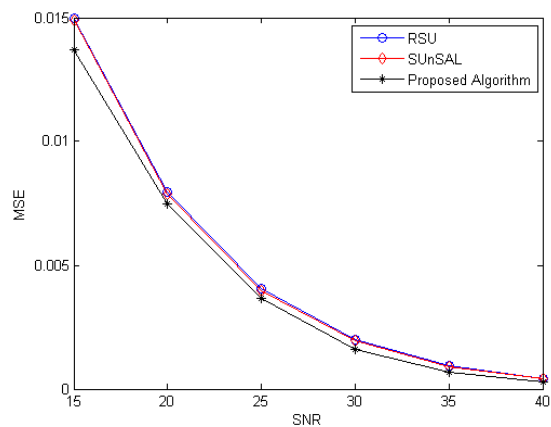
معدنی محفوظ نشده و آشکار است. برای بهبود عملکرد جداسازی، باندهای با نسبت سیگنال به نویز پایین، به دلیل تأثیرات اتمسفر، از ۲۲۴ باند مکعب داده حذف می‌شوند. بنابراین ۱۸۸ باند از این تصاویر مورد استفاده خواهند بود [۱۸]. در شکل (۷) نقشه‌های فراوانی نسبی تخمینی الگوریتم پیشنهادی دو ماده از داده‌ی کوپرایت با روش‌های RSU و SUnSAL مقایسه شده‌اند.



شکل ۵: مقایسه فاصله بردار فراوانی نسبی الگوریتم پیشنهادی با الگوریتم‌های RSU و SUnSAL

جدول ۱: مقایسه نقشه فراوانی نسبی دو عضو پایانی ساختگی با روش‌های الگوریتم پیشنهادی، RSU و SUnSAL نسبت به داده ساختگی واقعی

روش‌ها	نقشه فراوانی نسبی دو عضو پایانی ساختگی
داده ساختگی	
RSU	
SUnSAL	
الگوریتم پیشنهادی	



شکل ۶: مقایسه میانگین خطای مربعات الگوریتم پیشنهادی با الگوریتم‌های RSU و SUnSAL

در جدول (۱) نقشه فراوانی نسبی تخمین زده شده دو عضو پایانی ساختگی با استفاده از الگوریتم پیشنهادی با نسبت سیگنال به نویز 25dB و الگوریتم‌های RSU و SUnSAL مقایسه شده است. ارائه شده بیانگر این نکته است که روش پیشنهادی با دقت بیشتری نقشه فراوانی نسبی را نسبت به روش‌های دیگر تخمین می‌زند.

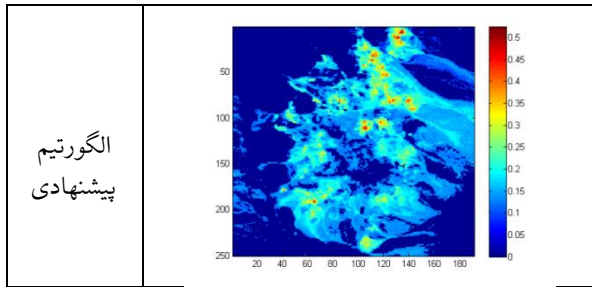
۵-۲- داده‌های واقعی

تا به امروز طیف سنج‌های متعددی در سال‌های مختلف تصاویری را ثبت کرده‌اند که برای آزمایش الگوریتم‌های مختلف استفاده می‌شود، که در ادامه به برخی از آن‌ها اشاره می‌شود داده‌های ابرطیفی ثبت شده با طیف سنج تصویربرداری مرئی-فروسرخ هوابرد در کوپرایت نوادا^{۱۳} حسگر AVIRIS یک طیف سنج تصویربرداری با ۲۲۴ کانال و دقت تقریباً ۱۰ نانومتر، محدوده‌ی طول موج ۰,۴ تا ۲,۵ میکرومتر را پوشش می‌دهد. دقت مکانی نیز ۲۰ متر است. استفاده از این داده مزایای زیادی دارد، از جمله اینکه، این مکان برای آزمایش‌های سنجش ازدور از سال ۱۹۸۰ مورد استفاده است و بسیاری از تحقیقات با دقت بالا روی این منطقه در دسترس است. ناحیه‌ی کوپرایت نیز یک ناحیه با مواد

۶- نتیجه‌گیری

در این مقاله با استفاده از راهبرد یادگیری بیزی تنک برای بازیابی سیگنال‌های تنک بلوکی که ساختار آن‌ها ناشناخته است، روش جدیدی به منظور جداسازی طیفی تصاویر ابرطیفی ارائه گردید. در مدل یادگیری بیزی تنک، میزان تنکی هر ضریب نه تنها به فوق پارامتر مربوطه بستگی دارد، بلکه به هایپرپارامترهای همسایه آن نیز بستگی دارد. هایپرپارامترها در سیگنال‌های تنک، می‌توانند با به حداکثر رساندن احتمال پسین آن‌ها، که در آن یک الگوریتم تکرار شونده با بهره‌گیری از

¹³ Cuprite, Nevada



شکل ۷: مقایسه نقشه فراوانی نسبی تخمینی الگوریتم

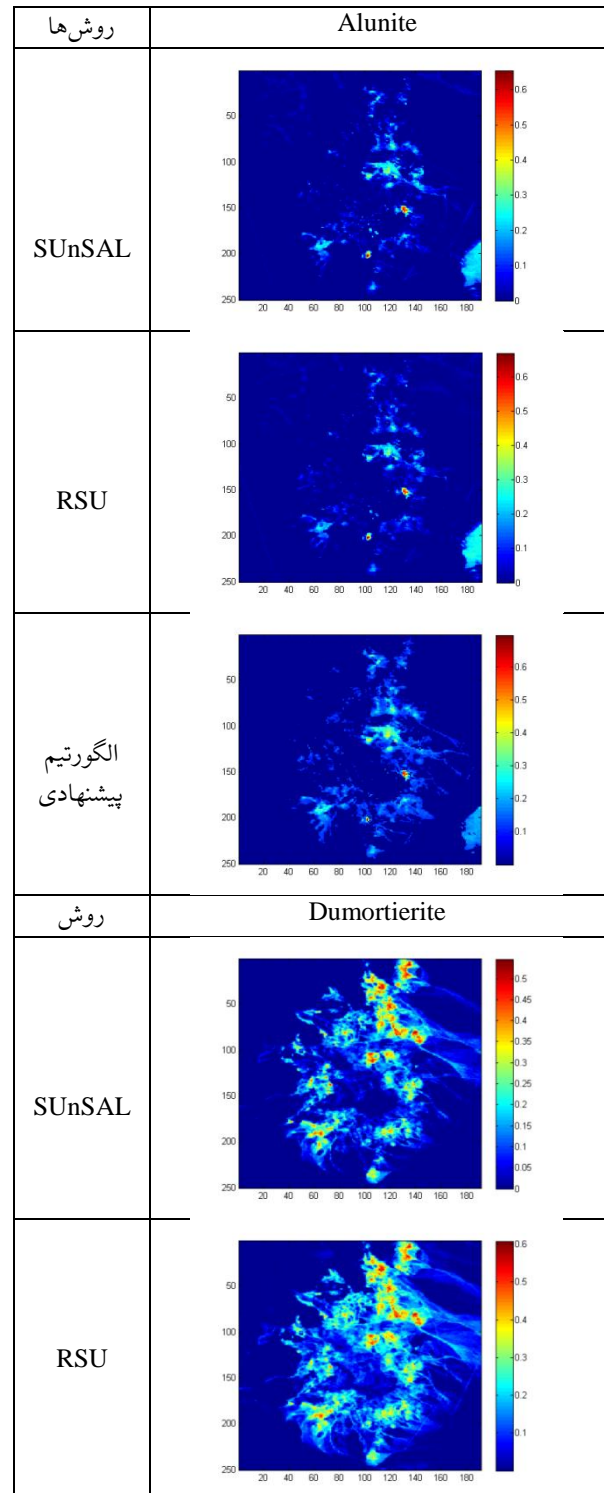
پیشنهادی با روش SUnSAL و RSU برای دو ماده نمونه

خطای میانگین مربعات نشان می‌دهد که الگوریتم پیشنهادی مقاله با بهره‌گیری از ساختار تنک بلوکی، حتی بدون نیاز به دانستن مکان دقیق و اندازه هر بلوک، کارایی بهتری در مقایسه با روش‌های پیشین به دست می‌آورد. برای پژوهش‌های آتی می‌توان از روش جداسازی تنک بلوکی مانند Block-IBA که نیازی به دانستن ساختار بلوکی سیگنال ندارند، بهره جست.

مراجع

- [۱] D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*. John Wiley & Sons, 2005.
- [۲] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on geoscience and remote sensing*, vol. 42, no. 8, pp. 1778-1790, 2004.
- [۳] R. O. Green *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote sensing of environment*, vol. 65, no. 3, pp. 227-248, 1998.
- [۴] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354-379, 2012.
- [۵] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014-2039, 2011.
- [۶] M. Nie, Z. Liu, X. He, Q. Qiu, Y. Zhang, and J. Chang, "End-member extraction based on segmented vertex component analysis in hyperspectral images," *Applied optics*, vol. 56, no. 9, pp. 2476-2482, 2017.
- [۷] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $L_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4282-4297, 2011.
- [۸] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2010: IEEE, pp. 1-4.
- [۹] D. M. Rogge, B. Rivard, J. Zhang, and J. Feng, "Iterative spectral unmixing for optimizing per-pixel endmember sets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 12, pp. 3725-3736, 2006.
- [۱۰] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3256-3274, 2014.
- [۱۱] Y. Ma, C. Li, X. Mei, C. Liu, and J. Ma, "Robust Sparse Hyperspectral Unmixing With $\ell_{2,1}$ Norm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1227-1239, 2016.
- [۱۲] C. Li *et al.*, "Sparse unmixing of hyperspectral data with bandwise model," *Information Sciences*, 2019.

فرمول حداکثر انتظار (EM) توسعه داده شده است، تخمین زده شوند. برای بررسی عملکرد الگوریتم پیشنهادی از دو دسته داده‌های ساختگی و داده‌های واقعی ثبت شده توسط سنجنده AVIRIS استفاده شده است. نتایج بر اساس معیارهای کمی فاصله زاویه طیفی و



- [۱۳] S. Zhang, J. Li, Z. Wu, and A. Plaza, "Spatial discontinuity-weighted sparse unmixing of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. ۵۷۷۹-۵۷۶۷, ۲۰۱۸.
- [۱۴] C. L. Farrow, M. Shaw, H. Kim, P. Juhás, and S. J. Billinge, "Nyquist-Shannon sampling theorem applied to refinements of the atomic pair distribution function," *Physical Review B*, vol. 84, no. 13, p. 134105, 2011.
- [۱۵] Y. Fu, H. Li, Q. Zhang, and J. Zou, "Block-sparse recovery via redundant block OMP," *Signal Processing*, vol. 97, pp. 162-171, 2014.
- [۱۶] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 360-372, 2014.
- [۱۷] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of machine learning research*, vol. 1, no. Jun, pp. 211-244, 2001.
- [۱۸] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765-777, 2007.
- [۱۹] A. Plaza, P. Martínez, R. Pérez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE transactions on geoscience and remote sensing*, vol. 42, no. 3, pp. 650-663, 2004.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه صنعتی شاهرود

۱۳۹۸ آذرماه ۲۸ الی ۲۷

کمی سازی استرس با بهره‌گیری از ویژگی‌های ارتباطات سیگنال‌های حجم خون

مویرگی و میزان تعریق پوست

عبدالعزیز عزیزی لاری^۱، علی مالکی^۲

^۱ پردیس علوم و فناوری‌های نوین، دانشگاه سمنان، سمنان،

a.azizi.lari@gmail.com

^۲ عضو هیئت علمی، گروه مهندسی پزشکی، دانشگاه سمنان، سمنان،

Amaleki@semnan.ac.ir

چکیده

امروزه تلاش‌های بسیاری جهت کمی‌سازی استرس افراد، مبتنی بر سیگنال‌های فیزیولوژیک انجام شده است. سیگنال‌های تعریق پوست و حجم خون مویرگی دو مورد پرکاربرد از این سیگنال‌ها هستند که ویژگی‌های مختلفی همچون زمانی، شکلی، فرکانسی و غیرخطی از آن‌ها استخراج شده است؛ با این حال در مطالعات گذشته، تعریف ویژگی روی هر یک از سیگنال‌ها به صورت جداگانه انجام شده است. هدف این پژوهش، استفاده از ویژگی‌های ارتباطات بین این دو سیگنال، جهت بهبود تشخیص استرس است. برای این هدف، از پایگاه داده‌گان DEAP استفاده شد و فرآیند تشخیص استرس شامل پیش‌پردازش سیگنال‌ها، استخراج ویژگی، کاهش ویژگی‌ها و طبقه‌بندی با هدف تفکیک استرس به دو سطح بالا و پایین انجام گردید. از ماتریس همبستگی و ماتریس شبکه به دست آمده از علیت گرنجر برای توصیف ارتباط دو سیگنال استفاده شد و ویژگی‌ها بر روی این ماتریس‌ها تعریف گردید. نتایج بر روی این ویژگی‌ها و با استفاده از دو طبقه‌بند بیزین و ماشین بردار پشتیبان (SVM) نشان داد که ویژگی‌های تعریف شده بر روی ماتریس شبکه (علیت گرنجر) با صحت طبقه‌بندی ۶۳ درصد نسبت به ویژگی‌های مختلف تعریف شده در مطالعات گذشته، و ویژگی‌های تعریف شده بر روی ماتریس همبستگی، عملکرد بهتری دارد. با ترکیب تمام ویژگی‌ها، صحت روش به بیش از ۶۵ درصد رسید.

کلمات کلیدی

کمی‌سازی استرس، ارتباطات، علیت گرنجر، همبستگی، طبقه‌بندی.

استرس و بازخورد آن به افراد، از اختلال‌ها پیش‌گیری کند. تلاش محققان در سال‌های اخیر جهت دقیق کردن، ساده‌سازی و مقاوم کردن سامانه‌های تشخیص استرس بوده است [۴]. در صورتی که با استفاده از دستگاه‌هایی به صورت غیرتهاجمی، از میزان استرس روزانه به افراد بازخورد داده شود، این می‌تواند به عنوان پیش‌آگهی استفاده شود. مطالعات متعددی نشان داده‌اند که با تحلیل سیگنال‌های فیزیولوژیک می‌توان به میزان استرس فرد پی برد [۵-۷].

تشخیص خودکار استرس کمک بزرگی به پیشرفت ارتباط بین انسان و ماشین می‌نماید و همین‌طور در بهبود دستگاه‌های بهداشتی

۱- مقدمه

استرس در جامعه بشری امروز، به طور گسترده‌ای مورد توجه قرار گرفته است. گویی هرچه پیشرفت فناوری در جوامع افزایش یافته و مسائل بهداشتی بیشتر حل شده است، مشکلات روحی و روانی خود را بیشتر نشان داده و بحث افسردگی، استرس و موضوعاتی از این دست، همچنان موضوع کار تحقیقات باقی مانده است [۱، ۲]. موقعیت‌های استرس‌زا در زندگی روزمره همه ما وجود دارند و می‌توانند منجر به اختلال‌هایی همچون اضطراب یا افسردگی شوند [۳]

در این میان، فناوری به کمک بشر آمده و سعی دارد با پایش پیوسته

و استفاده از ارتباطات بین آن‌ها در راستای بهبود تشخیص استرس است.

۲- روش کار

روش پیشنهادی که طرح‌واره‌ی کلی آن در شکل ۱ نشان داده شده است، شامل مراحل پیش‌پردازش سیگنال‌های PPG و GSR، توصیف ارتباط بین سیگنال‌ها، استخراج ویژگی، کاهش ویژگی و طبقه‌بندی است. در ادامه به توصیف دقیق این مراحل پرداخته شده است.

۲-۱- دادگان

دادگان DEAP در سال ۲۰۱۳ با همکاری چند دانشگاه اروپایی در آزمایشگاهی در شهر ژنو سوئیس و در فضایی کنترل شده تهیه شد. تهیه‌کنندگان این دادگان چند بعدی، به منظور پردازش حالت هیجانی فرد، سیگنال‌های مغزی و فیزیولوژیک پیرامونی را هم‌زمان با تحریک‌های ویدیویی ثبت نموده‌اند. برای ثبت سیگنال از دستگاه Biosemi ActiveTwo system استفاده شده است و از ویدیوهایی یک دقیقه‌ای به عنوان محرک هیجانی استفاده شده است.

۳۲ سوژه سالم که نیمی از آن‌ها مذکر بودند و سن آن‌ها بین ۱۹ تا ۳۷ سال بود در آزمایشگاه و شرایط کنترل‌شده در آزمایش شرکت کرده‌اند. پروتکل انجام آزمایش به این صورت بوده که هر فرد ۴۰ ویدیو یک دقیقه‌ای را تماشا می‌کرده است و پس از اتمام مشاهده هر ویدیو، ارزیابی خود را از شرایط. با پارامترهای انگیزتی، خوشایندی و تسلط به صورت عددی بین ۱ تا ۹ بیان نموده است. محدوده انگیزتی بین آرام یا خسته تا تحریک‌شده یا کاملاً هیجان‌زده تغییر می‌کند. محدوده خوشایندی از ناراحت و غمگین تا خوشحال و لذت‌بخش را در بر می‌گیرد. بازه تسلط یا کنترل‌پذیری اوضاع نیز از کاملاً بی‌اختیار و مطیع تا مسلط بر شرایط را در بر می‌گیرد.

برای دانلود دادگان به سایتی که برای این منظور در نظر گرفته شده است مراجعه شد. پس از امضای فایل قوانین و ارسال آن، نام کاربری و گذرواژه دریافت شد. دادگان با فرمت مناسب برای نرم‌افزار متلب با حجم حدود ۳ گیگابایت دریافت شد.^۷

۲-۲- پیش‌پردازش سیگنال‌ها

در قسمت پیش‌پردازش، ابتدا هر یک از سیگنال‌ها توسط میانگین واریانس سیگنال هر قطعه هنجار می‌شود. سپس یک فیلتر میانگین‌گذر بیضوی درجه ۷ با فرکانس قطع پایین ۰/۵ هرتز و فرکانس قطع بالای ۵۰ هرتز بر روی سیگنال PPG اعمال می‌شود تا سیگنال پالس از خط پایه جدا شود. سیگنال خط پایه از تفاضل

نقش بسزایی خواهد داشت. چندین مطالعه این موضوع را بررسی کرده‌اند که چگونه تشخیص خودکار استرس می‌تواند بر بهبود ایمنی رانندگان [۸، ۹] توانمندسازی واسطه‌های کاربری گفتار-پایه [۱۰، ۱۱] یا پیشگیری از آسیب‌های ناشی از استرس [۱۲] اثر داشته باشد. علاوه بر این مطالعه استرس از نگاه‌های مختلف باعث شناسایی بیشتر آثار آن بر تن و ذهن انسان می‌شود.

ورودی هر سامانه تشخیص استرس معمولاً سیگنال‌های فیزیولوژیک ثبت شده از بدن است. از جمله سیگنال‌های فیزیولوژیک شامل سیگنال تنفس، دمای پوست، فعالیت قلبی، فعالیت الکتریکی پوست^۱ (EDA) یا پاسخ گالوانیک پوست^۲ (GSR) در قالب مقاومت یا رسانایی پوست^۳ (SC)، پتانسیل الکتریکی پوست^۴ (SP)، فشار خون^۵ است. این سیگنال‌ها پس از اصلاح و پالایش وارد مرحله استخراج ویژگی می‌شوند. تاکنون ویژگی‌های مختلفی برای هر یک از سیگنال‌های فیزیولوژیک تعریف شده است. با وجود این، هنوز محققان به دنبال ویژگی‌های با اطلاعات بیشتر و قدرت تفکیک بالاتر در کاربردهای تشخیص استرس می‌باشند. بعداً این مرحله، مرحله انتخاب ویژگی و سپس ترکیب و طبقه‌بندی است. در این راستا از انواع روش‌های انتخاب یا کاهش بعد ویژگی و انواع مختلف طبقه‌بندها و رگرورها^۶ استفاده شده است. بهینه‌سازی و بهبود عملکرد این قسمت‌ها در مطالعات اخیر تشخیص استرس هدف محققان بوده است [۱۳-۱۸].

تاکنون در مطالعات مختلف از انواع سیگنال‌های فیزیولوژیک جهت تشخیص سطح استرس استفاده شده است. سیگنال‌های تعریق پوست (GSR) و حجم خون مویرگی که توسط حجم‌سنجی نوری (PPG) اندازه‌گیری می‌شود دو مورد پرکاربرد این سیگنال‌ها هستند که ثبت آن‌ها نیز ساده و با مزاحمت کم برای فرد انجام می‌شود [۱۹]. این سیگنال‌ها به عنوان ورودی به سامانه تشخیص استرس وارد می‌شوند. استخراج ویژگی‌های مناسب مبتنی بر ماهیت سیگنال‌های فیزیولوژیک نقش اساسی در فرآیند تشخیص دارد. تاکنون از ویژگی‌های مختلفی از این سیگنال‌ها همچون ویژگی‌های زمانی، شکلی، فرکانسی، خطی و غیرخطی جهت تشخیص استرس استفاده شده است. این حال در مطالعات گذشته، تعریف ویژگی به صورت جداگانه بر روی هر یک از سیگنال‌ها انجام شده است. حال آنکه وابستگی و ارتباط تنگاتنگی بین سیگنال‌های مختلف بدن وجود دارد و استفاده از ویژگی‌های این ارتباطات به توصیف پدیده استرس کمک شایانی می‌کند.

نقش ارتباطات و ارزیابی آنها در سیگنال‌هایی همچون سیگنال‌های مغزی نواحی مختلف سر به شدت مورد توجه و در حال بررسی است. با این حال، بررسی این ارتباطات در سیگنال‌های فیزیولوژیک پیرامونی کمتر مورد توجه قرار گرفته است. از این رو هدف این مقاله استخراج ویژگی‌های مشترک دو سیگنال GSR و PPG

1 Electrodermal Activity

2 Galvanic Skin Response

3 Skin Conductance

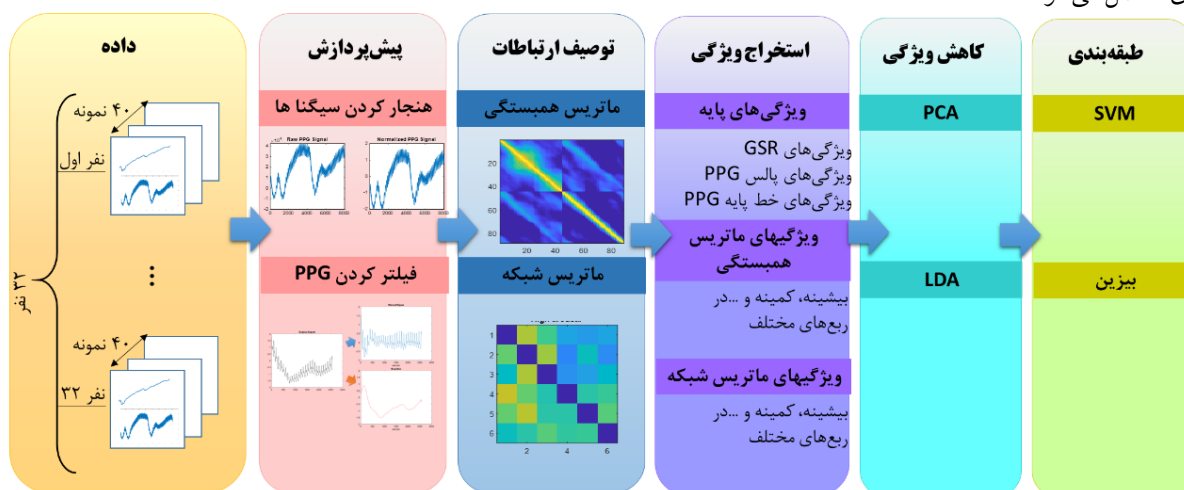
4 Skin Potential

5 Blood Pressure

6 Regressors

⁷ <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/>

سیگنال اولیه و سیگنال فیلتر شده پس از نرم‌سازی با پنجره ۱۰۰ نمونه‌ای حاصل می‌شود.



شکل ۳۶: طرحواره کلی روش پیشنهادی

۲-۳-۲- استفاده از روش علیت گرنجر جهت توصیف

ارتباطات

علیت گرنجر^۱ (GC) یکی از رایج‌ترین روش‌های مطالعه‌ی ارتباطات موثر^۲ و ارتباطات جهت‌دار از داده‌های مختلف همچون EEG است. مبنای این روش که ابتدا در اقتصاد مورد استفاده قرار گرفت، این است که علت، بر معلول خود در زمان پیشی می‌گیرد، بنابراین گفته می‌شود سری زمانی X دلیل گرنجر Y است، اگر X شامل اطلاعاتی در مورد آینده Y باشد که در گذشته‌ی Y (و گذشته‌ی سایر متغیرهای شرطی Z₁، Z₂، ...) موجود نباشد، یا به عبارت دیگر لحاظ کردن گذشته‌های Y، موجب کاهش چشمگیر خطای پیش‌بینی X شود. رایج‌ترین پیاده‌سازی GC از طریق مدل‌سازی خودبازگشتی چند متغیره^۳ (MVAR) سری‌های زمانی است، که هم امکان آزمون‌های چشمگیری آماری و هم محاسبه مقدار عددی GC را فراهم می‌کند. برای به دست آوردن علیت گرنجر از جعبه ابزار متلب^۴ GCCA استفاده شده است [۲۰]. این جعبه‌ابزار به رایگان قابل دانلود است.^۵ منابع مورد نظر قطعات سیگنال‌های GSR و خط پایه PPG است که در قسمت قبل برای محاسبه ماتریس همبستگی شرح داده شد. با این تفاوت که در اینجا به دلیل کم کردن بار محاسباتی تعداد قطعات با در نظر گرفتن گام بزرگتر یعنی ۲۰ ثانیه کمتر می‌شود. در واقع در این حالت هیچ همپوشانی بین قطعات وجود ندارد. ماتریس شبکه در این حالت، یک ماتریس ۶*۶ است که هر یک از گره‌های آن، قطعات سیگنال‌های GSR یا PPG است و هر یال، ارتباط مؤثر بین دو قطعه سیگنال را نمایش می‌دهد. برای به دست آوردن علیت گرنجر از مدل AR درجه ۱ استفاده شده است

۲-۳-۳- توصیف ارتباطات بین سیگنال‌های GSR و

PPG

برای بیان ارتباطات بین سیگنال‌های GSR و PPG از دو روش همبستگی و علیت گرنجر استفاده شده است. در ادامه نحوه استفاده هر یک از دو روش شرح داده شده است.

۲-۳-۱- استفاده از همبستگی جهت توصیف

ارتباطات

به منظور استفاده از روش همبستگی در هر یک از ۴۰ قسمت آزمایش، هر یک از این سیگنال‌های GSR و PPG، به قطعات زمانی ۲۰ ثانیه‌ای با گام مشخص یک ثانیه تقسیم می‌شود. سپس همبستگی بین هریک از قطعات محاسبه می‌شود و ماتریس همبستگی برای آن حاصل می‌شود. ابعاد ماتریس همبستگی برای هر قسمت از ۴۰ قسمت آزمایش که یک بازه ۶۳ ثانیه‌ای است برای گام یک ثانیه‌ای ۸۸*۸۸ خواهد شد زیرا در بازه ۶۳ ثانیه‌ای ۴۴ پنجره ۲۰ ثانیه‌ای با گام یک ثانیه‌ای قابل استخراج است و از آنجا که دو سیگنال GSR و PPG داریم بنابراین ۸۸ قطعه حاصل می‌شود.

همبستگی به صورت ضریب همبستگی پیرسون از رابطه (۱) به دست می‌آید؛

$$R(x, y) = \frac{cov(x, y)}{\sqrt{var(x)var(y)}} \quad (1)$$

در این رابطه x مقدار سیگنال اول و y مقدار سیگنال دوم است. Cov نشانگر کواریانس و Var نشانگر واریانس است.

¹ Granger Causality

² Effective connectivity

³ Multivariable Autoregressive model

⁴ Granger Causal Connectivity Analysis

⁵ <https://www.dcs.warwick.ac.uk/~feng/causality.html>

برای هر دو طبقه‌بند SVM و بیزین صادق است. این نتایج بر این موضوع که همبستگی به معنی علیت نیست صحه می‌گذارد.

جدول ۱۱: نتایج به ازای حالت‌های مختلف

طبقه بند بیزین	طبقه بند SVM	
۵۷،۰۸	۵۶،۱۳	ویژگی‌های برداشت شده از مطالعات گذشته
۵۸،۱۶	۵۷،۴۸	ویژگی‌های ماتریس همبستگی
۶۲،۸۹	۶۳،۳۲	ویژگی‌های ماتریس شبکه علیت گرنجر
۶۵،۱۲	۶۵،۷۸	ترکیب ویژگی

جدول (۳) نتایج کارهای مرتبط بر روی این دادگان را در مقایسه با پژوهش حاضر نشان می‌دهد. باید در نظر داشت که با اینکه در برخی پژوهش‌ها ممکن است با روش‌های طبقه‌بندی دیگری صحت بالاتری از این موارد به دست آمده باشد اما در این پژوهش درصد صحت بالا مد نظر نبوده و بیشتر نشان دادن اطلاعات موجود در ارتباطات بین سیگنال‌های PPG و GSR مد نظر قرار داشته است. بنابراین در جدول (۳) نیز برخی پژوهش‌هایی که از طبقه‌بندهای بیزین و SVM استفاده کرده‌اند لیست شده است.

جدول ۱۲: نتایج برخی پژوهش‌های مرتبط نسبت به پژوهش حاضر

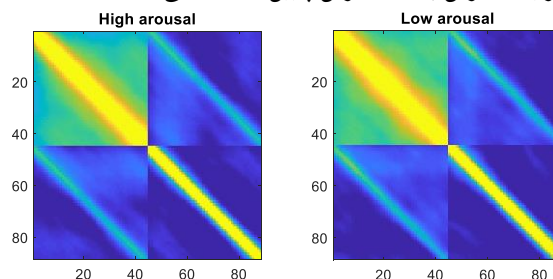
مرجع	طبقه‌بند	سیگنال مورد استفاده	درصد صحت
[۱۷]	Gaussian Naive Bayes	EEG	۶۲٪
[۱۷]	Gaussian Naive Bayes	PPG, GSR, Respiratory, EMG, Temperature	۵۷٪
[۲۱]	RNN and SVM	PPG	۶۱٪
[۲۲]	Bayesian SVM with causality features	EEG	۶۵٪
تحقیق حاضر	GSR, PPG		۶۵٪

۴- جمع‌بندی و پیشنهاد کارهای آتی

هدف این مقاله، استخراج ویژگی‌های مناسب ارتباطی بین سیگنال‌های PPG و GSR جهت بهبود تشخیص استرس است. بحث ارتباط بین سیگنال‌های مختلف بدن موضوعی است که در سال‌های اخیر مورد توجه قرار گرفته است و کارایی فراوانی در مسائل مختلف داشته است. عمده بحث ارتباطات در مطالعات گذشته بر روی سیگنال EEG انجام شده است. مطالعات کمی ارتباطات را در سیگنال‌های فیزیولوژیک پیرامونی بررسی کرده‌اند. براساس دانش ما، در هیچ‌یک از پژوهش‌های انجام شده در زمینه ارتباطات در سیگنال‌های فیزیولوژیک پیرامونی، تشخیص استرس یا هیجان بررسی نشده است.

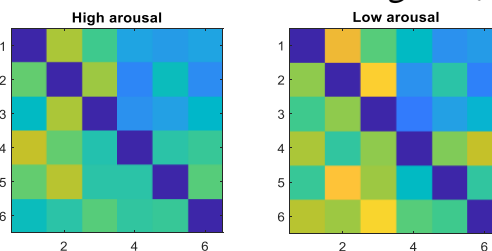
برای توصیف ارتباط بین سیگنال PPG و GSR دو روش پیشنهاد شد. در روش اول، ارتباطات در قالب ماتریس همبستگی توصیف شد. برای استخراج این ماتریس هر یک از این سیگنال‌های PPG و GSR، به قطعات زمانی ۲۰ ثانیه‌ای با گام یک ثانیه تقسیم می‌شود سپس

دوم و سوم که مبین ارتباطات GSR با قطعات PPG است قطر اصلی ایجاد شده است که نشان می‌دهد قطعات GSR و PPG هم‌زمان همبستگی بالاتری نسبت به یکدیگر دارند. در مجموع می‌توان گفت در شرایط استرس پایین مقدار همبستگی‌ها نسبت به شرایط استرس بالا افزایش می‌یابد. شکل (۲) میانگین ماتریس همبستگی برای تمام افراد در شرایط استرس بالا و استرس پایین را نشان می‌دهد.



شکل ۳۷: میانگین ماتریس همبستگی برای تمام افراد در شرایط استرس بالا و استرس پایین

برای مشاهده اثر استرس در ماتریس شبکه (علیت گرنجر) مشابه با روند طی شده برای ماتریس همبستگی، میانگین ماتریس بر روی همه افراد در شرایط استرس بالا و پایین محاسبه شده است. شکل (۳) این نمودار را نشان می‌دهد.



شکل ۳۸: میانگین ماتریس شبکه برای تمام افراد در شرایط استرس بالا و پایین

همان‌طور که دیده می‌شود در شرایط استرس پایین ارتباطات بیشتر می‌شود. این موضوع می‌تواند به دلیل هماهنگی ریمتی اتفاق افتاده بین دستگاه‌های مختلف بدن در شرایط استرس پایین باشد. این نتیجه با نتیجه مشاهده شده در ماتریس همبستگی هم مطابقت دارد.

۳-۱- نتایج طبقه‌بندی

جدول (۲) خروجی نتایج به ازای ویژگی‌های مختلف را نشان می‌دهد. سطر اول مربوط به شرایطی است که فقط از ویژگی‌های پایه استفاده شده است؛ در سطرهای دوم و سوم به ترتیب فقط ویژگی‌های ماتریس همبستگی و فقط ویژگی‌های ماتریس شبکه استفاده شده است. در سطر چهارم همه ویژگی‌ها با هم ترکیب شده‌اند.

نتایج نشان می‌دهد ویژگی‌های تعریف شده بر روی ماتریس شبکه علیت گرنجر بهترین نتیجه را در خروجی داشته است. این ویژگی‌ها با صحت طبقه‌بندی ۶۳،۳۲ درصد هم نسبت به ویژگی‌هایی آماری، فرکانسی و زمانی تعریف شده در مطالعات گذشته عملکرد بهتری داشته است و هم نسبت به ویژگی‌های تعریف شده بر روی ماتریس همبستگی. ترکیب تمام ویژگی‌ها با یکدیگر باعث می‌شود که صحت روش به ۶۵،۷۸ درصد صحت تنها با دو سیگنال برسد. این نتیجه

- driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156-166, 2005.
- [8] H. Gao, A. Yüce, and J.-P. Thiran, "Detecting emotional stress from facial expressions for driving safety," in *Image Processing (ICIP), 2014 IEEE International Conference on*, 2014, pp. 5961-5965: IEEE.
- [9] J. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156-166, 2005.
- [10] R. Fernandez and R. W. Picard, "Modeling drivers' speech under stress," *Speech communication*, vol. 40, no. 1-2, pp. 145-159, 2003.
- [11] G. Zhou, J. H. Hansen, and J. F. Kaiser, "Nonlinear feature based classification of speech under stress," *IEEE Transactions on speech and audio processing*, vol. 9, no. 3, pp. 201-216, 2001.
- [12] J. Wijsman, B. Grundlehner, H. Liu, H. Hermens, and J. Penders, "Towards mental stress detection using wearable physiological sensors," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 1798-1801: IEEE.
- [13] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," in *International Conference on Mobile Computing, Applications, and Services*, 2010, pp. 211-230: Springer.
- [14] Y. Deng, C.-H. Chu, H. Si, Q. Zhang, and Z. Wu, "AN INVESTIGATION OF DECISION ANALYTIC METHODOLOGIES FOR STRESS IDENTIFICATION," *International Journal on Smart Sensing & Intelligent Systems*, vol. 6, no. 4, 2013.
- [15] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, "Emotion assessment: همبستگی بین دو به دوی قطعات محاسبه می‌شود و ماتریس همبستگی برای آن حاصل می‌شود. در روش دوم برای توصیف ارتباطات از علیت گرنجر استفاده شده است. منابع در این روش مشابه روش همبستگی قطعات ۲۰ ثانیه‌ای از سیگنال‌ها هستند با این تفاوت که گام ۲۰ ثانیه‌ای برای آن در نظر گرفته شده است تا بار و زمان محاسبات کاهش یابد.
- نتایج این پژوهش نشان می‌دهد که ویژگی‌های تعریف شده بر روی ماتریس شبکه (علیت گرنجر) بهترین نتیجه را در خروجی دارند. این ویژگی‌ها هم نسبت به ویژگی‌هایی آماری، فرکانسی و زمانی تعریف شده در مطالعات گذشته عملکرد بهتری دارند و هم نسبت به ویژگی‌های تعریف شده بر روی ماتریس همبستگی. این نتیجه برای هر دو طبقه‌بند SVM و بیزین صادق بود. ترکیب تمام ویژگی‌ها با یکدیگر باعث شد که صحت روش در مقایسه با مطالعات گذشته بیشتر شود.
- ### مراجع
- [1] O. M. Mozos *et al.*, "Stress detection using wearable physiological and sociometric sensors," *International journal of neural systems*, vol. 27, no. 02, p. 1650041, 2017.
- [2] C. Sandi, "Stress and cognition," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 4, no. 3, pp. 245-261, 2013.
- [3] S. S. Dickerson and M. E. J. P. b. Kemeny, "Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research," *Psychological bulletin*, vol. 130, no. 3, p. 355, 2004.
- [4] K. Mohanavelu, R. Lamshe, S. Poonguzhali, K. Adalarasu, and M. Jagannath, "Assessment of Human Fatigue during Physical Performance using Physiological Signals: A Review," *Biomedical and Pharmacology Journal*, vol. 10, no. 4, pp. 1887-1896, 2017.
- [5] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Computer methods programs in biomedicine*, vol. 108, no. 3, pp. 1287-1301, 2012.
- [6] A. Riera *et al.*, "Electro-physiological data fusion for stress detection," ed, 2012.
- [7] J. A. Healey and R. W. Picard, "Detecting stress during real-world

- Arousal evaluation using EEG's and peripheral physiological signals," *Multimedia content representation, classification and security*, pp. 530-537, 2006.
- [16] A. De Santos, C. Sánchez-Avila, J. Guerra-Casanova, and G. Bailador-Del Pozo, "Real-time stress detection by means of physiological signals," in *Recent Application in Biometrics: InTech*, 2011.
- [17] S. Koelstra *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, 2012.
- [18] U. Zalabarria, E. Irigoyen, R. Martínez, and A. Salazar-Ramirez, "Detection of Stress Level and Phases by Advanced Physiological Signal Processing Based on Fuzzy Logic," in *International Conference on European Transnational Education*, 2016, pp. 301-312: Springer.
- [19] M. Saidi, H. Hassanpoor, and A. Azizi Lari, "Proposed new signal for real-time stress monitoring: Combination of physiological measures," *AUT Journal of Electrical Engineering*, vol. 49, no. 1, pp. 11-18, 2017.
- [20] Q. Luo *et al.*, "Spatio-temporal Granger causality: A new framework," *NeuroImage*, vol. 79, pp. 241-263, 2013.
- [21] L. A. B. Chacon, A. Fedoskin, E. Shcheglakova, S. Neamsup, and A. Rashed, "Emotion Analysis Using Heart Rate Data," in *International Conference on Database and Expert Systems Applications*, 2019, pp. 147-154: Springer.
- [22] S. Y. Chung and H. J. Yoon, "Affective classification using Bayesian classifier and supervised learning," in *12th International Conference on Control, Automation and Systems*, 2012, pp. 1768-1771: IEEE.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود

پیش بینی مدت زمان باقی مانده چرخه زمانی فرآیندهای سازمانی براساس جنبه‌های کنترلی و جریان داده

ایمان فیروزیان^۱، مرتضی زاهدی^۲، حمید حسن پور^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

Iman.firoozian@shahroodut.ac.ir

^۲ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

zahedi@shahroodut.ac.ir

^۳ استاد، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود

h.hassanpour@shahroodut.ac.ir

چکیده

پیش بینی دقیق زمان تکمیل اجرای نمونه‌فرآیندها در شرایطی که مهلت های زمانی تعیین می‌شود، بسیار حیاتی محسوب می‌شود. پیش بینی در فرآیندهای سازمانی بسیار چالش برانگیز می‌باشد، زیرا عوامل متعددی بر روی روند نمونه فرآیندها تأثیرگذار می‌باشد. کارهای پیشین در حوزه پیش بینی، تنها از اطلاعات آماری مربوط به پرونده های پیشین استفاده می‌کردند که به تنهایی منجر به پیش بینی دقیق نمی‌شد. در این مقاله، یک رویکرد جدید برای افزایش دقت پیش بینی ارائه می‌دهیم که از جنبه های کنترلی و جریان داده همزمان بهره می‌برد. رویکرد این مقاله، یک مدل فرآیندی را می‌سازد که با اطلاعات داده ای و زمانی غنی سازی شده است تا بتواند پیش بینی زمانی را انجام دهد. بنابر عوامل و نیازهایی که مطرح شد، پیش بینی را به صورتی پیاده سازی نمودیم تا بتواند از داده های گذشته برای تعیین احتمال فعالیت ها و وظایف آینده استفاده شود. نتایج ارائه شده بهبود پیش بینی را نسبت به روش های برپایه جریان کنترلی نشان می‌دهد.

کلمات کلیدی

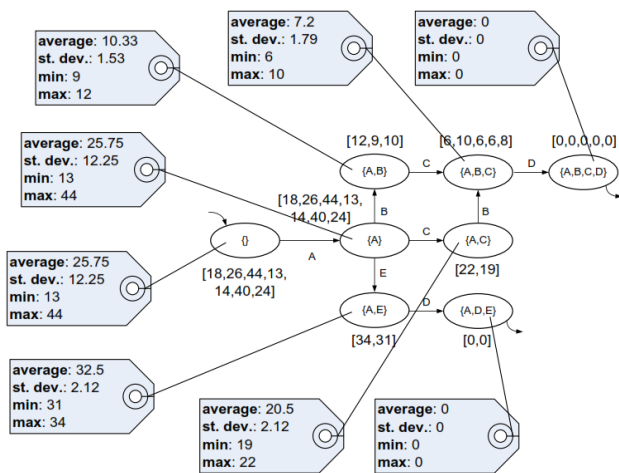
چرخه زمانی فرآیندهای سازمانی، سیستم مدیریت فرآیندهای کسب و کار، پیش بینی زمانی.

سازمانی مربوطه ثبت و آنالیز می‌شوند. فرآیندکاوی، از جهتی دیگر، کل فرآیند سازمانی را در نظر می‌گیرد و بنابراین آنالیزهای کلی را بر روی داده هایی انجام می‌دهد که از فرآیند مربوطه بدست می‌آید. یک فرآیند سازمانی را می‌توان به صورت یک توالی از عملیات های انجام شده برای رسیدن به یک هدف خاص در نظر گرفت. ایده اصلی فرآیندکاوی، کشف مدل های فرآیندی، نظارت (برای مثال، نظارت انحرافات با مقایسه مدل و سابقه رویداد) و بهبود فرآیندهای سازمانی (همانند داده کاوی ساختار سازمانی، داده کاوی شبکه اجتماعی و سیستم پیشنهادگر

۱. مقدمه

تعداد شرکت هایی که از سیستم های اطلاعاتی فرآیند-آگاه برای مدیریت فرآیندهای سازمانی خود بهره می‌برند، در حال گسترش می‌باشد. تمام این سیستم ها وقایع مربوط به هر فعالیت را در قالب سابقه رویداد ذخیره می‌کنند. این سابقه رویدادها تحت جوانب مختلفی تحلیل و آنالیز می‌شوند. دو محور مهم فرآیندکاوی که تحلیل و آنالیز را انجام می‌دهد، شامل داده‌کاوی و فرآیندکاوی می‌باشد. داده کاوی از جنبه داده آنالیز را انجام می‌دهد: مجموعه ای از مقادیر که مستقل از فرآیند

است. در پژوهش دیگری، Van der Aalst و همکاران [۱۴] از اطلاعات موجود در سابقه رویداد برای پیش‌بینی طول مدت زمان باقی‌مانده انجام پرونده بر اساس مدت‌های زمانی مشاهده‌شده در گذشته استفاده می‌کند. آن‌ها با طراحی یک سیستم گذر حالت، مدت زمان باقی‌مانده برای هر حالت ملاقات شده را از دنباله‌های موجود در سابقه رویداد جمع‌آوری می‌کنند. نهایتاً هنگامی که یک پرونده انجام می‌شود، زمان باقی‌مانده در هر حالت، با میانگین‌گیری از طول مدت باقی‌مانده از پرونده‌های پیشین بر اساس سابقه رویداد پیش‌بینی می‌شود. رویکرد ارائه‌شده توسط Rogge-Solti [۱۵] نیز مشابه آن است با این تفاوت که داده‌ها و منابع را در نظر نگرفته است اما از مدل‌های GDT_SPN به جای سیستم‌های گذر استفاده می‌کند، که باعث دقت بیشتری در هنگام استفاده از کنترل موازی می‌شود. از کارهای ارائه شده در [۱۴] و [۱۶] می‌توان برای محک روش پیشنهادی مقاله استفاده نمود.



شکل 39: سیستم‌گذار حاشیه‌نویسی شده ارائه شده توسط Aalst و همکاران [۱۴]

Folino و همکاران [۱۷]، بهبودی را با توسعه [۱۴] برپایه خوشه‌بندی مبتنی بر پیش‌بینی ارائه دادند. آن‌ها از اطلاعات متنی اضافی یک دنباله (برای مثال، بار کاری جاری سیستم) برای خوشه‌بندی استفاده می‌کنند. ایده پیشنهادی برپایه گروه‌بندی دنباله‌های مشابه می‌باشد و پیش‌بینی را برای دنباله‌های جدید که مشابه با دنباله‌های پیشین می‌باشد، انجام می‌دهد. آن‌ها از پیش‌بینی برای اعلام هشدار در صورت تجاوز از مهلت تعیین‌شده استفاده می‌کنند. ترکیب این روش با روش [۱۵] نیز می‌تواند در قالب یک پروپوزال مطرح شود.

و سیستم پیش‌بینی) با استخراج دانش از روی سابقه رویدادها می‌باشد. سابقه رویدادها، فایل‌هایی حاوی داده برای ارجاع به اجرای یک فرآیند خاص می‌باشد. اطلاعات مرسوم که از سابقه رویداد استخراج می‌شود، در حقیقت جریان کنترلی کار را نشان می‌دهد. علاوه بر اطلاعات مربوط به جریان کار، سابقه رویدادها اغلب حاوی داده اضافی (همانند منابع،

برچسب زمانی) می‌باشد که انواع مختلف آنالیز را از جنبه‌های مختلف انجام می‌دهد. آنالیز داده که شامل بیش از یک جنبه می‌باشد به عنوان آنالیز چندجنبه‌ای شناخته می‌شود.

امروزه، چندین تکنیک جدید در رویکردهای چندجنبه‌ای ارائه می‌شود. ایده اصلی در این موارد، تعریف الگوریتمی که قادر به استخراج و کسب سود از ترکیب تکنیک‌های حوزه فرآیندکاوی و داده‌کاوی می‌باشد. فرآیندکاوی را می‌توان به دو دسته بر اساس زمان انجام آنالیز تقسیم بندی کرد: آنالیز پسین و آنالیز زمان-اجرا. در آنالیز پسین، یک بخش متناهی از داده‌های پیشین را به صورت ورودی دریافت می‌کند و تلاش دارد دانش درون آن را استخراج نماید. رویکرد دوم، اطلاعات را تا زمانی که فرآیند سازمانی در حال اجرا می‌باشد، در نظر می‌گیرد. از دید سازمانی، دستیابی سریع به اطلاعات جدید و دانش جدید بسیار ارزشمند می‌باشد. برای مثال، در یک مؤسسه مالی با مدیریت پرونده‌ها، تشخیص سریع کلاهبرداری به جهت پرهیز از ازدست دادن منابع مهم می‌باشد.

در این مقاله، تکنیکی را ارائه می‌دهیم که در شرایط عملیاتی استفاده شود. رویکردی که قادر به انجام پیش‌بینی زمان اتمام نمونه فرآیند برای حالت‌های در حال اجرا می‌باشد. پیش‌بینی برای اساس جنبه‌های مختلف انجام می‌شود؛ پیش‌بینی بر اساس جریان فعالیت‌های پرونده جاری و همچنین بر روی داده نمونه فرآیند جاری انجام می‌شود.

ادامه این مقاله به صورت زیر ساختار بندی شده است: بخش دوم کارهای مرتبط در حوزه پیش‌بینی را در چارچوب فرآیندکاوی را مرور می‌کند. مدل پیشنهادی در بخش چهارم ارائه داده می‌شود. آزمایشات بر روی پایگاه داده واقعی انجام داده می‌شود و نتایج آن گزارش می‌شود.

۲. پیشینه تحقیق

بسیاری از کارهای مرتبط، اساس پیش‌بینی را برپایه مشاهدات پیشین و یا براساس اعتقادات اولیه قراردادند. کارهای مرتبط با پیش‌بینی براساس سری‌های زمانی نیز به صورت گسترده وجود دارد. سری‌های زمانی سنسج‌های پریودیکی از داده‌ها را لحاظ می‌کنند. برای مثال، مقدار نهایی روزانه سهام یک شرکت در بازار بورس و یا ارقام فروش هفتگی یک محصول یک سری زمانی می‌باشد. متدهای بسیاری برای ایجاد مدل‌های مناسب برای فرآیند تولیدکننده چنین داده‌هایی وجود دارد. مروری از این روش‌ها در [۱۰] آمده است. آنالیز روندها و نقاط تغییر در سری‌های زمانی در [۱۱] یافت می‌شود اما به حوزه این پروپوزال مربوط نمی‌شود.

یکی از روش‌های مدیریت فرآیندها، مدل‌سازی فرآیندها توسط مدل‌های SPN می‌باشد که آنالیز زمان-گسسته این مدل‌ها توسط van der Aalst [۱۲] و همچنین رساله دکتری Reijers [۱۳] بررسی شده

است ابتدا برخی مفاهیم موجود در فرآیند سازمانی را به صورت فرمال تعریف کنیم.

می‌توان یک فرآیند سازمانی را به صورت یک چندتایی از $\langle T, R_s, R_o, W, start, end \rangle$ در نظر گرفت که در آن:

T مجموعه‌ای از کارهای قابل تعریف در فرآیند سازمانی است که به صورت $T = \{t_1, t_2, t_3, \dots, t_n\}$ نشان می‌دهیم.

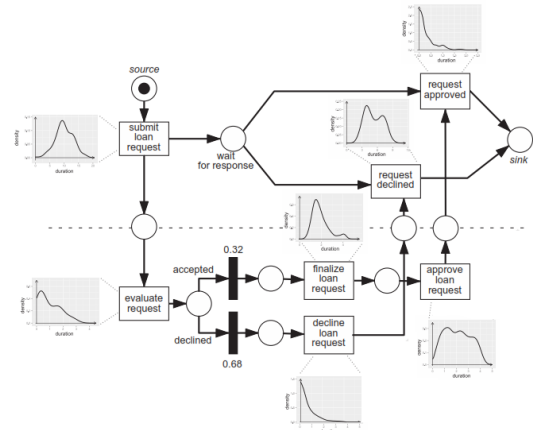
R_s مجموعه منابع انسانی

R_o مجموعه از نقش‌ها است و برای هر نقش $o \in R_o$ یک زیر مجموعه از منابع $o.resources \subseteq R_s$ به آن نسبت داده شده است که به معنای منابعی است که دارای توانایی‌های تعریف شده در نقش o هستند.

W جریان‌کاری تعریف شده روی مجموعه T است. در واقع W یک گراف جهت دار با مجموعه راس‌های $V = T \cup \{start, end\}$ و مجموعه یال‌های $E \subseteq T \times T$ مشروط به اینکه، به راس $start$ هیچ یال ورودی وجود نداشته باشد و از راس end هیچ یال خروجی وجود نداشته باشد و برای هر راس $t \in T$ در گراف W حداقل یک مسیر با شروع از راس $start$ و پایان راس end و شامل t وجود داشته باشد. به بیان بهتر $E = T \times T \cup T \times \{end\} \cup \{start\} \times T$ به هر راس $t \in T$ یک زیر مجموعه نقش‌های مناسب شده است که مجموعه نقش‌های مجاز به انجام کار t هستند که به صورت $t.roles \subseteq R_o$ نشان می‌دهیم.

در ضمن می‌توان به تعاریف مشابه از یک فرآیند سازمانی با WF-net نیز اشاره داشت [۲]. که در این تعریف مجموعه V به صورت $V = T \cup Split \cup Join$ است که $Split = Join$ می‌باشد. در این مقاله ما فرض می‌کنیم منبع تخصیص یافته به کار $t \in T$ تصمیم می‌گیرد کدام یک از فعالیت‌های مجموعه $Join$ یا $Split$ اتفاق بیافتد و در ادامه شرح مفصل‌تری از آن بیان خواهیم کرد.

وقتی یک پرونده به عنوان یک درخواست در سازمان ایجاد می‌شود می‌بایست مسیرهای از پیش تعریف شده در یکی از فرآیندهای سازمان را طی کند و منابع تعریف شده در آن فرآیند روی آن کار انجام دهند. بنابراین با توجه به نوع پرونده ایجاد شده فرآیند سازمانی مناسب انتخاب می‌شود و یک نمونه فرآیند^۳ از فرآیند سازمانی ایجاد می‌شود. برای نمایش و مدل کردن هر پرونده ایجاد شده در نمونه فرآیند یک توکن^۴ در رأس شروع فرآیند سازمانی قرار می‌گیرد و به صورت خودکار یک نمونه کاری در رأس بعدی $next_W(t) = t^{next} | (t, t^{next}) \in E, W = (E, V)$ الگوریتم تخصیص منبع در رأس t^{next} فعال می‌شود. این الگوریتم یکی از منابع مجموعه $R = \{r | \forall o \in t^{next}.roles, r \in o.resources\}$ انتخاب می‌کند و نمونه کاری ایجاد شده به آن تخصیص داده می‌شود. منبع انتخاب شده فعالیت تعریف شده در نمونه کاری را تکمیل می‌کند و برای نشان دادن این وضعیت توکن به راس t^{next} منتقل می‌شود تا سرانجام با طی فعالیت‌های تعریف شده در



شکل 40: مدل ارائه شده توسط Andreas Rogge-soltie برای حل احتمالاتی فرآیند [۱۵]

Hwang و همکاران [۱۸] و همچنین Zheng و همکاران [۱۸] نیز در بحث پیش‌بینی عملکرد، نتایج آزمایشات خود را ارائه دادند. آن‌ها از فرمول‌هایی برای محاسبه کیفیت سرویس‌دهی همانند مدت زمان مورد انتظار در ترکیب سرویس‌ها استفاده می‌کنند. فرض رایج در این کارها آنست که سرویس‌ها از بلوک‌های ساختاری (بلوک‌های تک ورودی-تک خروجی) استفاده می‌کند. این کار را می‌توان با کار ارائه شده توسط Kiepuszewski و همکاران [۱۹] مقایسه نمود. روش‌های پیشنهادی را می‌توان برای فرآیندهای سازمانی نیز به کار برد. در این پروپوزال با حذف فرض ساختار بلوکی در این کار، مدل‌های پیچیده‌تری را پوشش می‌دهیم و نمونه فرآیندهای در حال اجرا را نیز در نظر می‌گیریم.

۳. مدل پیشنهادی

۳-۱- تحلیل اثر رفتاری منابع در فرآیند

مؤلفه اصلی در تغییر وضعیت پرونده‌ها در فرآیندهای سازمانی، منابع انسانی می‌باشد. تحلیل اثر رفتاری منابع که منجر به تحلیل دقیق‌تر عملکرد فرآیند سازمانی می‌شود که نهایتاً به افزایش دقت عملیات پیش‌بینی منجر می‌شود. منابع در فرآیندهای سازمانی به دو دسته منابع انسانی و منابع غیر انسانی دسته‌بندی می‌شوند. هنگامی که یک پرونده در یک فرآیند سازمانی ایجاد می‌شود، مجموعه از آیت‌های کاری با توجه به جریان کاری تعریف شده در فرآیند در طول زمان روی آن پرونده تعریف می‌شود و آیت‌های کاری^۱ ایجاد شده در قالب یک فعالیت^۲ به مجموعه‌ای از منابع مجاز در فرآیند سازمانی تخصیص داده می‌شود. پس از مطالعه و بررسی الگوهای منابع موجود در مقاله [۱] که توسط آلست و همکارانش ارائه شد، می‌توان تأثیر منابع را در فرآیند سازمانی بر اطلاعات موجود در پرونده‌ها و مسیر حرکت پرونده‌ها در فرآیند سازمانی خلاصه کرد.

۳-۲- مدل سازی رفتار منابع

برای ارائه یک تعریف فرمال از مدل رفتاری منابع در این بخش لازم

³ Process Instance

⁴ Token

¹ Work Item

² Activity

فرآیند سازمانی به راس پایانی end برسد.

اجرای فعالیت بر روی نمونه کاری ایجاد شده است. خروجی این توابع علاوه بر بردار اطلاعات پرونده به منبع انتخاب شده نیز وابسته است. بنابراین برای تغییر در جهت و انداز بردار اطلاعات پرونده در راس t توسط منبع r از بردار $U(t,r)$ استفاده می‌کنیم و برای تغییر بعد آن از ماتریس $C(t,r)$ استفاده می‌کنیم. تابع $P(c.dataValues,t,r)$ و مجموعه حالت بعدی را برمی‌گرداند و خروجی تابع $F(c.dataValues,t,r)$ حاوی اطلاعات توزیع آماری است که مدت زمان فعالیت منبع از آن تبعیت می‌کند.

افزودن مکانیزم یادگیری به مدل رفتاری

در بخش قبل اثر رفتاری منبع بر روی پرونده در فرآیند سازمانی به ترتیب با یک بردار تغییر و اندازه، یک ماتریس تغییر بعد، تابع انتخاب حالت بعدی و تابع توزیع آماری مدل شد. برای پیش‌بینی عملکرد یک منبع میبایست از یک عامل یادگیر استفاده کنیم به نحوی که این عامل یادگیر با بررسی سابقه فعالیت های منبع بر روی پرونده های پیشین و بررسی سابقه روخدادهای گذشته سیستم آموزش ببیند. و برای پیش‌بینی عملکرد منبع قبل از فرا رسیدن زمان تخصیص به آن از عامل یادگیر استفاده شود و عملکرد عامل یادگیر به عنوان نسخه پیش بینی شده منبع اصلی در نظر گرفته شود. برای این منظور در این مقاله چهار شبکه عصبی چند لایه برای اثرهای رفتاری چهارگانه منبع در نظر می‌گیریم. بنابراین به ازای هر منبع که در فرآیند سازمانی تعریف شده چهار شبکه عصبی چند لایه استفاده می‌شود.

۳-۵- آموزش مدل رفتاری

برای آموزش عامل های یادگیری از سابقه رویدادهای سیستم استفاده می‌کنیم. در سابقه روخدادها به هر پرونده که در فرآیند وارد شده است دنباله ای از رویداد ها تحت عنوان trace وجود دارد. به ازای هر نمونه کاری ایجاد شده در فرآیند حداقل سه رویداد تخصیص (Allocation) ، شروع (Start) و تکمیل (Complete) وجود دارد. در رویداد تخصیص منبع انجام کار مشخص می‌شود و در رویداد شروع، منبع کار را روی پرونده مربوطه شروع می‌کند و در رویداد تکمیل، منبع زمان تکمیل را به سیستم اعلام می‌کند. برای اینکه امکان آموزش عامل های یادگیر وجود داشته باشد می‌بایست قبل از رویداد شروع و بعد از رویداد تکمیل اطلاعات موجود در پرونده در سیستم ذخیره شود یا به عبارت بهتر سابقه روخدادها علاوه بر trace ها با اطلاعات پرونده غنی شود. در شبه کد ذیل با استفاده از این اطلاعات غنی شده برای ایجاد Dataset و آموزش عامل های یادگیر استفاده می‌کنیم.

Learning Algorithm:

برای مدل کردن فعالیتی که منبع روی نمونه کاری ایجاد شده در هر راس از فرآیند سازمانی میتوان در ادامه تعریف ذیل را اضافه کرد. همانطور که در بخش قبل بیان کردیم هر منبع در نمونه کاری ایجاد شده بر روی پرونده جاری فعالیت های ذیل را انجام می‌دهد. الف) تغییر یا تکمیل اطلاعات ب) تعیین مسیر حرکت پرونده در نمونه فرآیند. در ادامه برای مدل سازی رفتاری منبع لازم است ابتدا اطلاعات ذخیره شده در پرونده را مدل کنیم.

۳-۳- مدل سازی اطلاعات پرونده

هر پرونده که در یک سازمان ایجاد میشود حاوی اطلاعاتی درباره درخواست، نوع درخواست و درخواست دهنده است و وقتی در یک فرآیند سازمانی قرار میگیرد و قرار است روی آن مجموعه مشخص از فعالیت ها انجام شود. هر فعالیتی که روی آن انجام میشود و مدت زمانی روی آن کار انجام میشود از ابتدا تا انتها، همگی وابسته به اطلاعات و داده هایی است که در قالب فیلدها و مقادیر آنها در پرونده قرار گرفته است. برای مدل سازی این مجموعه از اطلاعات میتوان با کمی سازی مقادیر فیلدهای موثر در فرآیند به یک بردار از مقادیر عددی رسید که در ادامه این مقاله آن را به صورت $case.dataValues = [v_1, v_2, v_3, \dots, v_k]$ نامگذاری می‌کنیم. مقادیر اولیه این بردار در زمان ایجاد پرونده مشخص می‌شود و تا زمانی که پرونده در نمونه فرآیند از یک حالت به حالت دیگر می‌رود مقادیر آن ممکن است عوض شود و در برخی موارد ممکن است اندازه این بردار تغییر کند.

۳-۴- مدل سازی اثر رفتاری منابع

منابعی که به صورت غیر مستقیم بر اساس نقش های تعریف شده در هر رأس از فرآیند سازمانی قرار گرفته اند، توانایی انجام فعالیت هایی را دارند که در فرآیند سازمانی بر روی پرونده ها انجام می‌شود. بعد از تخصیص نمونه کاری ایجاد شده در رأس t روی پرونده c به منبع $\{a \in \{r \mid \forall o \in t.roles, r \in o.resources\}$ ، اثر رفتاری منبع a یکی از موارد ذیل است.

تغییر در جهت و اندازه بردار $c.dataValues$

تغییر در بُعد بردار $c.dataValues$

تعیین مقدار t^{next} در صورتی که اندازه مجموعه $\{(t, t^{next}) \mid (t, t^{next}) \in E\}$ بزرگتر از یک باشد.

مدت زمانی که لازم است منبع روی نمونه کاری ایجاد شد فعالیت کند. برای مدل کردن چهار اثر رفتاری منابع که ذکر شد، می‌توان از برخی توابع و عملیات جبری روی بردارها استفاده کرد. به عنوان مثال، برای تغییر در اندازه و جهت بردار $c.dataValues$ کافی است، از عملگر جمع بردارها استفاده کنیم و بدین معناست که ویرایش اطلاعات موجود در پرونده توسط منبع مانند برداری است که با بردار اطلاعات پرونده جمع می‌شود. به همین صورت می‌توان تغییر بُعد بردار اطلاعات را بر ضرب برداری بردار اطلاعات در ماتریس اثر رفتاری منابع در هر فعالیت مدل کرد. برای مدل سازی دو مورد باقیمانده نیز از یک تابع استفاده می‌کنیم که بردار اطلاعات پرونده ورودی آن است و خروجی آن به ترتیب حالت یا حالت بعدی و پارامتر های توزیع آماری مدت زمان

شبیه ساز تعریف کردیم و به کمک آن چهار نوع فایل سابقه رخداد با ۱۵۰۰، ۳۰۰۰، ۴۵۰۰ و ۶۰۰۰ ایجاد کردیم. علاوه بر الگوریتم پیشنهادی الگوریتم های معرفی شده در مقالات را نیز رو پایگاه داده ساخته شده اجرا شد و نتایج را در جدول ۴ و شکل ۲ آمده است. همانطور که در شکل ۲ نشان داده شده است نتایج برای الگوریتم های مقالات ... و الگوریتم پیشنهادی برای دو حالت مجموعه و لیست انتزاعی از سیستم انتقال (transition system abstraction) اجرا گرفته شده است. برای مقایسه نتایج نیز از دو معیار معرفی شده در مقاله ... به نامهای MAPE و RMSPE استفاده شد نحوه محاسبه ای دو معیار فاصله بین مقادیر پیشبینی شده و مقادیر واقعی در فرمولهای ۱ و ۲ آمده است.

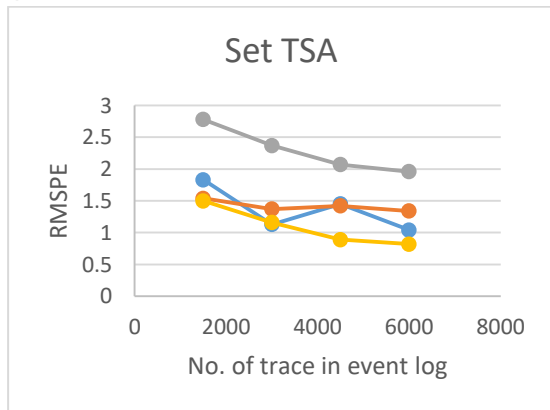
فرمول ۱:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|$$

فرمول ۲:

$$RMSPE = 100\% \sqrt{\frac{\sum_{i=1}^n (A_i - F_i)^2}{n}}$$

Polynomial Kernel RBF Kernel Aalst et al. NNR



Polynomial Kernel RBF Kernel Aalst et al. NNR

```

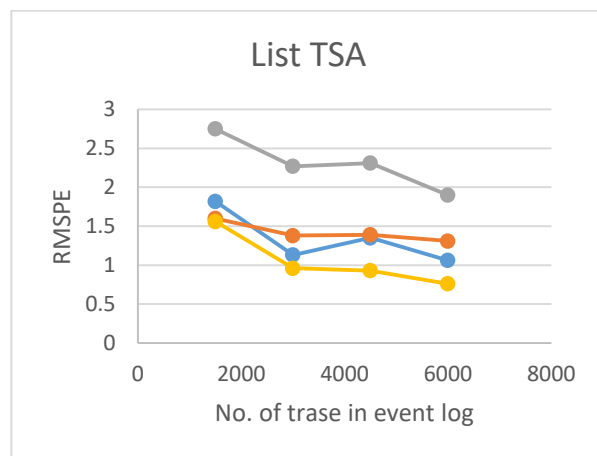
1 Workflow ← (T,Rs,Ro,W,start-end)
2   foreach r ∈ Rs do
3     foreach
4       t | (t,t') ∈ V and W = (V,E) and r ∈ t.rollse.resources do
5         U_Training_Set[t,r] = Empty_Set()
6         C_Training_Set[t,r] = Empty_Set()
7         P_Training_Set[t,r] = Empty_Set()
8         F_Training_Set[t,r] = Empty_Set()
9         U_NetWork[t,r] = new Architect.Perceptron(IN_SIZE,
10            IN_SIZE*2, OUT_SIZE)
11        C_NetWork[t,r] = new Architect.Perceptron(IN_SIZE,
12            IN_SIZE*2, OUT_SIZE)
13        P_NetWork[t,r] = new Architect.Perceptron(IN_SIZE,
14            IN_SIZE*2, OUT_SIZE)
15        F_NetWork[t,r] = new Architect.Perceptron(IN_SIZE,
16            IN_SIZE*2, OUT_SIZE)
17      end
18    foreach trace ∈ eventLog do
19      foreach event ∈ trace do
20        if (event.type is 'Start') then
21          input = trace.getCase.dataValues
22          startTime = event.time
23        end
24        if (event.type is 'Complete') then
25          output = trace.getCase.dataValues
26          completeTime = event.Time
27        end
28        U_Training_Set[event.task,event.resource].Add({input,output})
29
30        C_Training_Set[event.task,event.resource].Add({input,output.size(
31            ))
32
33        P_Training_Set[event.task,event.resource].Add({input,event.getNe
34            xtEvent.task})
35
36        F_Training_Set[event.task,event.resource].Add({input,completeTim
37            e-startTime })
38      end//Event
39    end//Trace
40  end
41  foreach
42    t | (t,t') ∈ V and W = (V,E) and r ∈ t.rollse.resources do
43    U_NetWork[t,r].trainer.train(U_Training_Set[t,r])
44    C_NetWork[t,r].trainer.train(C_Training_Set[t,r])
45    P_NetWork[t,r].trainer.train(P_Training_Set[t,r])
46    F_NetWork[t,r].trainer.train(F_Training_Set[t,r])
47  end //task
48 end //Resources

```

۴. نتایج آزمایشات

برای آزمایش روش پیشنهادی از یک شبیه ساز برای ایجاد یک پایگاه داده استفاده میکنیم در این شبیه ساز امکان ایجاد پایگاه داده ای وجود دارد که مسیر حرکت پرونده و مدت زمان انجام کارها توسط منبع به داده های ذخیره شده در پرونده وابسته است. این شبیه ساز برای هر پرونده تعدادی داده تصادفی ایجاد میکند که این داده ها بر مسیر حرکت پرونده و زمان انجام کار تا تاثیر گذار هستند. در شکل ذیل یک فرایند سازمانی فرضی با ۸ کار و ۱۵ منبع نشان داده شده است. این فرایند در

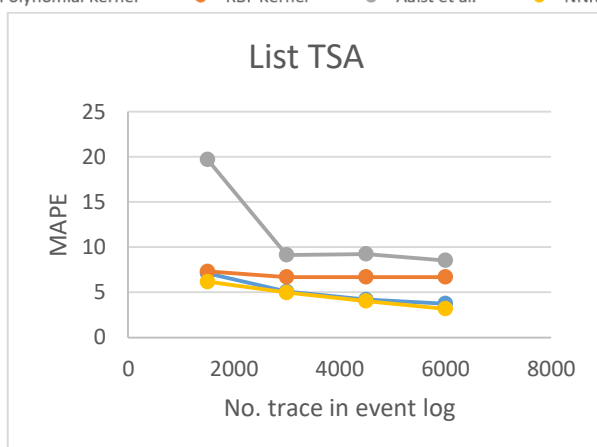
No. of Traces	Polynomial Kernel		RBF Kernel		Aalst et al.		NNR	
	PEVM	RMSPE	PEVM	RMSPE	PEVM	RMSPE	PEVM	RMSPE
Using set transition system abstraction with no limitation								
1500	6.56	1.83	7.79	1.54	20.53	2.78	6.58	1.5
3000	5.07	1.13	6.68	1.37	9.82	2.37	4.42	1.16
4500	4.7	1.45	6.68	1.42	8.02	2.07	3.48	0.89
6000	3.39	1.04	6.68	1.34	8.16	1.96	3.09	0.82
Using list transition system abstraction with no limitation								
1500	7.12	1.82	7.31	1.6	19.71	2.75	6.17	1.56
3000	5.08	1.13	6.69	1.38	9.13	2.27	4.98	0.96
4500	4.19	1.35	6.69	1.39	9.24	2.31	4.04	0.93
6000	3.75	1.06	6.69	1.31	8.54	1.9	3.18	0.76



Polynomial Kernel RBF Kernel Aalst et al. NNR



Polynomial Kernel RBF Kernel Aalst et al. NNR



همانطور که در شکل مشخص است در کلیه الگوریتم‌های مورد مقایسه با افزایش تعداد trace‌های موجود در فایل سابقه رخدادها نتیجه پیشبینی بهتر شده است. همانطور که انتظار میرفت عامل یادگیر در مدل کردن رفتار منابع تاثیر خوبی بر عملکرد الگوریتم پیشبینی داشته است و

نتایج بهتری نسبت به سایر الگوریتم‌ها داشته است.

۵. نتیجه‌گیری

در این مقاله، تکنیکی را ارائه دادیم که در شرایط عملیاتی استفاده شود. رویکردی که قادر به انجام پیش‌بینی زمان اتمام نمونه فرآیند برای حالت‌های در حال اجرا می‌باشد. پیش‌بینی برای اساس جنبه‌های مختلف انجام می‌شود؛ پیش‌بینی بر اساس جریان فعالیت‌های پرونده جاری و همچنین بر روی داده نمونه فرآیند جاری انجام می‌شود.

مراجع

- [13] Van Der Aalst, W.M., K.M. Van Hee, and H.A. Reijers, *Analysis of discrete-time stochastic petri nets*. Statistica Neerlandica, 2000. 54(2): p. 237-255.
- [14] Reijers, H.A., Design and control of workflow processes: business process management for the service industry. 2003: Springer-Verlag.
- [15] Van der Aalst, W.M., M.H. Schonenberg, and M. Song, *Time prediction based on process mining*. Information Systems, 2011. 36(2): p. 450-475.
- [16] Rogge-Solti, A. and M. Weske, Prediction of business process durations using non-Markovian stochastic Petri nets. Information Systems, 2015.
- [17] Rogge-Solti, A., Probabilistic Estimation of Unobserved Process Events. 2014.
- [18] Folino, F., M. Guarascio, and L. Pontieri, Discovering context-aware models for predicting business process performances, in On the Move to Meaningful Internet Systems: OTM 2012. 2012, Springer. p. 287-304.
- [19] Hwang, S.-Y., et al., A probabilistic approach to modeling and estimating the QoS of web-services-based workflows. Information Sciences, 2007. 177(23): p. 5484-5503.
- [20] Kiepuszewski, B., A.H.M. ter Hofstede, and C.J. Bussler. *On structured workflow modelling*. in *Advanced Information Systems Engineering*. 2000. Springer.
- [21] Wombacher, A. and M. Iacob. Estimating the Processing Time of Process Instances in Semi-structured Processes--A Case Study. in Services Computing (SCC), 2012 IEEE Ninth International Conference on. 2012. IEEE.
- [22] Anklesaria, K.P. and Z. Drezner, *A multivariate approach to estimating the completion time for PERT networks*. Journal of the Operational Research Society, 1986: p. 811-815.
- [1] Van Der Aalst, W., Process mining: discovery, conformance and enhancement of business processes. 2011: Springer Science & Business Media.
- [2] Van Der Aalst, W., Process mining: discovery, conformance and enhancement of business processes. 2011: Springer Science & Business Media.
- [3] Hammer, M. and J. Champy, Reengineering the Corporation: Manifesto for Business Revolution, A. 2009: Zondervan.
- [4] Weske, M., Business process management: concepts, languages, architectures. 2012: Springer Science & Business Media.
- [5] Van Der Aalst, W.M., A.H. Ter Hofstede, and M. Weske, *Business process management: A survey*, in *Business process management*. 2003, Springer. p. 1-12.
- [6] Reijers, H.A. and S.L. Mansar, Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega, 2005. 33(4): p. 283-306.
- [7] Van der Aalst, W.M., Verification of workflow nets, in Application and Theory of Petri Nets 1997. 1997, Springer. p. 407-426.
- [8] Jansen-Vullers, M. and M. Netjes. Business process simulation—a tool survey. in Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools, Aarhus, Denmark. 2006.
- [9] Lenz, R. and M. Reichert, *IT support for healthcare processes—premises, challenges, perspectives*. Data & Knowledge Engineering, 2007. 61(1): p. 39-58.
- [10] Rozinat, A., et al., *Discovering simulation models*. Information Systems, 2009. 34(3): p. 305-327.
- [11] De Gooijer, J.G. and R.J. Hyndman, *25 years of time series forecasting*. International journal of forecasting, 2006. 22(3): p. 443-473.
- [12] Zeng, L., et al., Event-driven quality of service prediction, in Service-Oriented Computing—ICSOC 2008. 2008, Springer. p. 147-161.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه صنعتی شاهرود

۲۷ الی ۲۸ آذرماه ۱۳۹۸

تحلیل مرجع مشترک مبتنی بر منابع دانش و شبکه عصبی

سمیرا حورعلی^۱، مرتضی زاهدی^۲، منصور فاتح^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
s.hourali@shahroodut.ac.ir

^۲ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
zahedi@shahroodut.ac.ir

^۳ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود
mansour_fateh@shahroodut.ac.ir

چکیده

تحلیل مرجع مشترک به‌عنوان یکی از اصلی‌ترین وظایف در حوزه پردازش زبان طبیعی و استخراج اطلاعات، اهمیت فراوانی دارد. هرچه این مسئله با کیفیت بالاتری انجام شود، سامانه‌هایی که از آن استفاده می‌کنند، عملکرد مطلوب‌تری خواهند داشت. هرچند طی چند دهه گذشته تحلیل هم‌ارجاعی همواره یک موضوع تحقیقاتی فعال بوده است؛ اما هنوز هم دقت آن در حد قابل قبول برای استفاده در کاربردهای درک متن نیست. یکی از دلایل دشواری این مسئله نیاز آن به منابع دانش مختلف از جمله دانش لغوی، دانش نحوی، دانش جهان، ساختار گفتمان و دانش معنایی است. به‌گونه‌ای که در یک متن نمونه ممکن است، هر یک از موارد هم‌ارجاعی نیازمند یک یا چند منبع دانش برای تحلیل باشد؛ در واقع هر نوع از هم‌ارجاعی نیاز به دانش و اطلاعات مخصوص به خود برای حل شدن دارد. در این مقاله سعی شده است که با استفاده از منابع دانش، بازنمایش دانش و شبکه عصبی این مشکل رفع شود. جهت بررسی عملکرد روش پیشنهادی شبیه‌سازی گسترده‌ای صورت گرفته است. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی توانسته به نتایج بهتری نسبت به سایر روش‌های جدید تحلیل هم‌ارجاعی دست یابد و مقدار F1 روی پایگاه داده CONLL-2012 به میزان ۷/۱ درصد بهبود داده شده است.

کلمات کلیدی

تحلیل مرجع مشترک، منبع دانش، نامبری، شبکه عصبی، بازنمایش دانش.

موجودیت استفاده می‌شوند نامبری^۱ گوئیم. بنابراین می‌توان گفت همه نامبری‌هایی که به یک موجودیت یکسان اشاره می‌کنند با یکدیگر هم مرجع^۲ هستند. منظور از تحلیل مرجع مشترک نامبری‌هایی در متن است که به یک موجودیت اشاره می‌کنند. اگر این فرآیند به‌درستی انجام شود، می‌توان از آن برای بهبود عملکرد سیستم‌های پردازش متن نظیر استخراج اطلاعات، ترجمه ماشینی، تحلیل احساسات، سرقت ادبی و غیره استفاده کرد.

علیرغم تلاش‌های زیادی که طی چهار دهه اخیر روی مسئله تحلیل مرجع انجام شده است کارایی روش‌های حل این مسئله هنوز به حد قابل قبولی نرسیده است به‌گونه‌ای که این مسئله هنوز یک موضوع تحقیقاتی فعال است. یکی از دلایل آن متنوع بودن دلایل هم‌ارجاعی است به بیان دیگر ممکن است هرکدام از چندین نامبری درون یک سند با قوانین متفاوتی نظیر نزدیکی فاصله، تطبیق کامل رشته‌ای، تطبیق کلمه

۱- مقدمه

معمولاً حین نوشتن یک متن، برای اشاره به یک موجودیت (یک شخص، سازمان، محل و غیره) تنها از نام آن استفاده نمی‌کنیم؛ بلکه بسته به شرایط، به دلیل جلوگیری از تکرار و بیان اطلاعات بیشتری در مورد آن موجودیت یا تأکید بر یک ویژگی خاص، از عبارات توصیفی مختلف نظیر عبارات اسمی و گاهی ضمائر برای اشاره به آن موجودیت استفاده می‌کنیم. مثلاً ممکن است برای اشاره یک شخص نام کاملش بیان شود (عادل فردوسی پور)، یا تنها نام خانوادگی (فردوسی پور) او و در موارد غیر رسمی‌تر تنها از نام کوچک (عادل) استفاده می‌کنیم. حتی ممکن است شخص یا اشیا را با ویژگی‌ها و یا کاربردهایشان توصیف کنیم (مجری برنامه نود یا گزارشگر خوب فوتبال) و یا با استفاده از یک ضمیر (او، ش و غیره). به‌چنین عباراتی که برای اشاره به یک

یک شبکه عصبی عمیق از پیش آموزش داده شده است که توسط مدل-های زبانی آموزش دیده و ویژگی‌های استخراج شده توسط این روش جهت بالا بردن کارایی روش پیشنهادی و شناسایی درست موجودیت-های نامدار و نامبری‌ها مؤثر است. در تعبیه واژگان، هر واژه‌ای به بردار بازنمایشی^۴ نگاشت داده می‌شود. بنابراین هر واژه با بردار بازنمایشی که حاوی یکسری اطلاعات معنایی و ارتباطی از واژه است، نمایش داده می‌شود و بازنمایشی از واژگان با توجه به زمینه‌ی متنی که واژگان در آن قرار دارند، بدست می‌آید. پس از تعبیه واژگان آن‌ها را به‌عنوان ورودی به یک شبکه Bi-GRU داده تا دنباله کلمات استخراج شود.

هدف تشخیص مجموعه نامبری‌های درست C برای نامبری m از مجموعه دنباله کلمات^۵ S با داشتن مجموعه دانش G (که شامل سه‌تایی‌هایی است که در بخش ۲-۲ توضیح داده شده‌اند) در سند D طبق رابطه (۱) است.

$$G = \frac{\sum_{c \in C} e^{F(c,m,G,D)}}{\sum_{s \in S} e^{F(s,m,G,D)}} \quad (1)$$

با فرض اینکه بردار کلمات استخراج شده توسط روش تعبیه واژه RoBERTa به‌صورت X_1, X_2, \dots, X_T باشد و خروجی شبکه Bi-GRU نیز به‌صورت $X_1^*, X_2^*, \dots, X_T^*$ باشد، نمایش برداری توسط روابط (۲) و (۳) برای دنباله کلمات موجود در متن استخراج می‌گردد.

$$\hat{X}_i = \sum_{t=1}^T a_t X_t \quad (2)$$

$$a_t = \frac{e^{\alpha_t}}{\sum_{z=1}^T e^{\alpha_z}} \quad (3)$$

$$\alpha_t = NN_{\alpha}(X_t^*) \quad (4)$$

در نهایت برای دنباله کلمات i بازنمایش رابطه (۵) استخراج می‌شود که در آن X_{START}^* و X_{END}^* ابتدا و انتهای تعبیه دنباله کلمات، $\emptyset(i)$ بردار ویژگی و NN شبکه عصبی روبه‌جلو است.

$$S_i = [X_{START}^*, X_{END}^*, \hat{X}_i, \emptyset(i)] \quad (5)$$

۲-۲- منابع دانش

همان‌طور که گفته شد، رویکرد اصلی این مقاله استفاده از منابع دانش است، برای استفاده از این منابع همه آن‌ها رو به یک فرمت تبدیل کردیم، فرمت در نظر گرفته شده سه‌تایی ("ابتدا" (لیستی از کلمات)، رابطه بین ابتدا و انتها، "انتها" (لیستی از کلمات)) است. پس از تبدیل کردن اطلاعات منابع به این فرمت همه آن‌ها را با هم ادغام کرده تا منبع دانش بدست آید. منابع دانش استفاده شده برای آموزش سیستم عبارتند از:

منبع دانش^۱ OMCS [۱۴]: این منبع شامل ۶۰۰ هزار سه‌تایی است، به‌عنوان مثال (غذا، استفاده می‌شود برای، خوردن) و (باد، توانایی دارد، به شوق وزیدن). تمام روابط در این منبع به‌صورت دستی مشخص شده است. در این مقاله سه‌تایی‌های با امتیاز بالا (بالتر از ۳) انتخاب شده است که تعداد آن‌ها ۶۱۶۷۳ سه‌تایی است.

سر، توازی نحوی، شباهت معنایی به مرجع خود تحلیل شوند. همچنین با در نظر گرفتن انواع مختلف موارد هم‌ارجاعی متوجه می‌شویم که هر نوع از هم‌ارجاعی نیاز به دانش و اطلاعات مخصوص به خود برای حل شدن دارد. روش‌های تحلیل مرجع مشترک که تا به حال ارائه شده‌اند عموماً به چهار دسته قطعی (مبتنی بر قاعده یا قانون)، آماری، مبتنی بر یادگیری تقویتی و مبتنی بر یادگیری عمیق تقسیم می‌شوند. مرجع [۱] رویکردی مبتنی بر قاعده برای تحلیل مرجع مشترک ارائه داده است که در آن قواعد به‌صورت دست‌نویس استخراج و به سیستم آموزش داده شده است. از مزایای این روش می‌توان به دقت بالا و سادگی طراحی اشاره کرد، اما قابلیت انعطاف این روش پایین است و لازم است برای هر زبان طبیعی مجزا، سیستم مجدداً از ابتدا توسط افراد خیره طراحی شود.

مرجع [۲] رویکردی آماری و با ناظر برای حل مسئله مرجع مشترک ارائه داده است که ویژگی‌های سطح موجودیت در خوشه‌ها و نامبری‌ها استخراج شده و به آن آموزش داده می‌شود. ضعف این‌گونه روش‌ها نیاز به داده‌های فراوان برچسب‌گذاری شده است که باید به‌صورت دستی یا خودکار برچسب‌گذاری شوند. مراجع [۳] و [۴] رویکردهایی مبتنی بر یادگیری تقویتی برای حل مسئله ارائه داده‌اند که در آن‌ها مجموعه‌ای از عمل‌ها برای یادگیری اتصال درست بین حالت‌ها یا نامبری‌ها انجام می‌شود، سپس عامل تحت پاداش‌ها و خطاهایی که دریافت می‌کند عمل درست یا اتصال نامبری‌های هم‌مرجع را آموزش می‌بیند. مراجع [۵] - [۹] رویکردهایی مبتنی بر یادگیری عمیق برای تحلیل مرجع مشترک ارائه داده‌اند که در آن‌ها شبکه‌های عصبی عمیق توسط توابع خطای تعریف شده، جهت یافتن زنجیره‌های هم‌ارجاعی آموزش می‌بینند. از معایب این روش‌ها می‌توان به پیچیده بودن مدل و ساختار محاسباتی پیچیده اشاره کرد.

رویکرد این مقاله برای حل مسئله تحلیل مرجع مشترک استفاده منابع دانش برای یادگیری بهتر سیستم و شناسایی درست موجودیت-های نامدار است. در ادامه این مقاله، در بخش دوم مدل در نظر گرفته شده برای حل مسئله را تشریح می‌کنیم. در بخش سوم نتایج شبیه‌سازی را ارائه می‌کنیم و در پایان، نتایج را شرح می‌دهیم.

۲- روش پیشنهادی

در این بخش ابتدا در فاز پیش‌پردازش چگونگی استخراج نامبری‌های موجود در متن شرح داده می‌شود؛ سپس با استفاده از منابع دانش و طبق روش تصمیم‌گیری چند شاخصه ارائه شده و روش وزن‌دهی فازی نامبری‌های هم‌مرجع شناسایی و استخراج می‌شوند.

۲-۱- بازنمایش دنباله کلمات^۲

روش RoBERTa: در مدل پیشنهادی برای استخراج نامبری‌ها ابتدا توسط روش RoBERTa [۱۰] تعبیه واژگان استخراج می‌شود. روش RoBERTa جدیدترین روش پیش‌آموزش شبکه است و در سایر زمینه‌های پردازش به نتایج ایده‌آلی دست یافته است. طبق نتایج شبیه‌سازی این روش نسبت به سایر روش‌های تعبیه واژگان نظیر BERT [۱۱]، ELMo [۱۲] و GloVe [۱۳] بهتر عمل می‌کند. RoBERTa در واقع

سالم") با جستجو در منبع دانش یافت شود، این رابطه را توسط میانگین تعبیه کلمات "سالم" و "خوراکی" نمایش داده می‌دهیم.

در نتیجه، برای دنباله کلمات s مجموعه دانش بازیابی شده را با \mathcal{K}_s نشان می‌دهیم، که حاوی m_s بازنمایش دانش مرتبط به- صورت $k_{1,s}, k_{2,s}, \dots, k_{m_s,s}$ می‌باشد. برای ادغام دانش فوق‌الذکر در مدل پیشنهادی، با این چالش روبرو هستیم که تعداد زیادی تعبیه وجود دارد، درحالی‌که بیشتر آن‌ها در زمینه‌های خاص بی‌فایده هستند. برای حل این مشکل، روابط (۷، ۸ و ۹) برای انتخاب دانش مناسب پیشنهاد شده است.

۲-۴- یافتن زنجیره‌های هم‌مرجع

برای هر زوج دنباله کلمات (جفت) (s_d, s_i) که s_d دنباله موردنظر و s_i دنباله کلمات کاندید شده برای هم‌مرجع بودن یا نبوده با دنباله s_d است، ابتدا تعبیه‌های این جفت (e_{s_d}, e_{s_i}) به هم الحاق شده و تعبیه کلی $e_{d,i}$ حاصل می‌شود که جهت انتخاب دانش برای هر دو دنباله استفاده می‌شود. برای مثال برای دنباله s ، وزن هر $k_i \in \mathcal{K}_s$ توسط رابطه (۷) و (۸) محاسبه می‌شود.

$$w_i = \frac{e^{\beta k_i}}{\sum_{k_j \in \mathcal{K}_s} e^{\beta k_j}} \quad (7)$$

$$\beta_k = NN_B([e_{s_d}, k]) \quad (8)$$

دانش کلی نیز به صورت رابطه (۹) بدست می‌آید.

$$o_s = \sum_{k_i \in \mathcal{K}_s} w_i \cdot k_i \quad (9)$$

امتیاز نهایی برای جفت (s_i, s_j) توسط رابطه (۱۰) محاسبه می‌شود.

$$F(s_i, s_j) = f_m(s_i) + f_m(s_j) + f_c(s_i, s_j) \quad (10)$$

در رابطه (۱۰) $f_m(s_i)$ تابع امتیازدهی به دنباله کلمات s_i است که مشخص شود یک نامبری معتبر است یا خیر. $f_c(s_i, s_j)$ نیز تابع امتیازدهی به دو دنباله s_i و s_j است که مشخص شود آیا دو دنباله هم-مرجع هستند یا خیر. در رابطه (۱۲) منظور از \odot ضرب عنصر به عنصر^{۱۴} است. این دو تابع به صورت روابط (۱۱) و (۱۲) تعریف می‌شوند. جهت بررسی اعتبار دنباله کلمات از جهت نامبری بودن، تعبیه دنباله به همراه دانش مربوط به آن را که توسط رابطه (۹) استخراج شد به یک شبکه عصبی روبه‌جلو می‌دهیم تا امتیاز آن دنباله مشخص شود.

$$f_m(s_i) = NN_m([e_{s_i}, o_{s_i}]) \quad (11)$$

$$f_c(s_i, s_j) = NN_c([e_{s_i}, o_{s_i}, e_{s_j}, o_{s_j}, e_{s_i} \odot e_{s_j}, o_{s_i} \odot o_{s_j}]) \quad (12)$$

بعد از محاسبه امتیازات هم‌مرجعی برای تمام دنباله کلماتی که نامبری هستند، تابع softmax طبق رابطه (۱۳) روی محتمل‌ترین کاندیدها برای پیش‌بینی نهایی اعمال می‌شود.

$$\hat{F}(s_i, s_j) = \frac{e^{F(s_i, s_j)}}{\sum_{s_i \in S} e^{F(s_i, s_j)}} \quad (13)$$

در رابطه (۱۳) کاندیدهای دارای \hat{F} بالاتر از مقدار آستانه t انتخاب می‌شوند.

منبع دانش پزشکی [۱۵]: این منبع شامل اطلاعات پزشکی است به‌عنوان مثال سه‌تایی (ام آر آی، است، یک آزمایش) و (تزریق، است، درمان) که در مجموع شامل ۲۲۲۳۴ سه‌تایی است.

ویژگی‌های زبان‌شناسی: علاوه بر منابع دانش دست‌نویس که به آن‌ها اشاره شد، ویژگی‌های زبان‌شناسی نظیر جاندار بودن، جنسیت نیز در نظر گرفته شده است. تجزیه‌کننده^{۱۵} استنفورد برای تولید تعداد، جاندار و جنسیت برای همه عبارات اسمی به کار می‌رود، که می‌تواند به‌طور خودکار دانش زبانی (به‌صورت سه‌تایی) برای داده‌های ما تولید کند. ویژگی تعداد شامل مقدار s برای مفرد و p مقدار برای جمع است. ویژگی‌های جاندار نشان‌دهنده جاندار که s یا p است که در صورت جاندار، مقدار مرد، زن و خنثی بودن برای ویژگی جنسیت استفاده می‌شود. برای مثال نامبری "پسران" به‌عنوان جمع و مرد برچسب‌گذاری می‌شود که برای نمایش از سه‌تایی‌های ("پسران"، تعداد، جمع) و ("پسران"، جنسیت و جاندار، مرد) استفاده می‌شود. در نتیجه ۴۰۱۴۹ سه‌تایی برای جمع بودن و ۴۰۴۶۲ سه‌تایی برای جنسیت و جاندار در نظر گرفته می‌شود.

دانش معنایی [۱۶]: برای جمع‌آوری دانش معنایی، ابتدا با استفاده از تجزیه‌کننده استنفورد اطلاعات پایگاه ویکی‌پدیا انگلیسی^{۱۶} را تجزیه نمودیم و تمام لبه‌های وابستگی را با فرمت (گزاره^{۱۷}، آرگومان^{۱۸}، رابطه، شماره) استخراج شد. در تجزیه‌کننده استنفورد، هنگامی که فعل یک فعل ارتباطی (مانند هستم (am)، است (is)) است، بین پیش‌بینی کننده (گزاره) و موضوع، لبه "nsubj" ایجاد می‌شود. هر جفت در این منبع دانش با یک احتمال خلفی رابطه (۶) اندازه‌گیری می‌شود.

$$P_r(a|p) = \frac{Count_r(p, a)}{Count_r(p)} \quad (6)$$

در رابطه (۶)، $Count_r(p)$ و $Count_r(p, a)$ به ترتیب نشان‌دهنده این هستند که چند بار p و جفت گزاره-آرگومان (p, a) در رابطه r وجود داشته‌اند. در آزمایشات زمانی که $P_r(a|p) > 0.1$ و $Count_r(p, a) > 10$ بود، سه‌تایی (p, r, a) را در نظر گرفتیم. به‌عنوان مثال، ("سگ"، nsubj، "پارس") یک رابطه معنایی معتبر است. سرانجام، دو رابطه معنایی nsubj و dobz برای منبع دانش معنایی انتخاب شدند که به ترتیب شامل ۱۷۰۷۴ و ۴۵۳۶ جفت گزاره-آرگومان برای nsubj و dobz می‌باشد.

۲-۳- بازنمایش دانش^{۱۹}

برای هر دنباله کاندید شده جهت نامبری بودن یا هم‌مرجع بودن با نامبری دیگر، دانش‌های مختلفی از منبع دانش و با روش‌های مختلف می‌توان استخراج کرد. به‌عنوان مثال جهت سادگی و تعمیم در مدل پیشنهادی از تطابق ساختار رشته‌ای^{۱۱} برای استخراج دانش استفاده نمودیم. به‌طور خاص، برای هر سه‌تایی $t \in \mathcal{G}$ که ابتدا^{۱۲} و انتهای^{۱۳} آن لیستی از کلمات است، اگر ابتدای آن مشابه با رشته‌ای در دنباله کلمات s باشد، آن را یک سه‌تایی مرتبط می‌دانیم. بنابراین، اطلاعات t را با میانگین‌گیری تعبیه تمام کلمات انتهای آن، رمزنگاری می‌کنیم. برای مثال، اگر s "سیب" باشد و سه‌تایی ("سیب"، هست یک، "خوراکی

۳- نتایج شبیه‌سازی

در این بخش روش پیشنهادی از جنبه‌های مختلف مورد بررسی قرار گرفته و با سایر روش‌ها مقایسه شده است. جهت بررسی کارایی روش پیشنهادی در این مقاله از سه پایگاه داده [۱۷] CONLL-2012، [۱۸] English Gigaword و MEDCo-A [۱۹] استفاده شده است. در جدول (۱) روش پیشنهادی طبق معیارهای MUC [۲۰]، B^3 [۲۹]، $CEAF_{\phi 4}$ [۲۲]، مقادیر دقت، بازخوانی و $Avg.F1$ (میانگین خروجی سه معیار B^3 ، $CEAF_{\phi 4}$ و MUC)، روی دو پایگاه [۱۷] CONLL-2012 با روش‌های قطعی [۱]، آماری [۲]، مبتنی بر یادگیری تقویتی عمیق [۴] و مبتنی بر یادگیری عمیق [۷] مقایسه شده است. همان‌طور که جدول (۱) مشاهده می‌شود، روش پیشنهادی طبق همه معیارها بر روش‌های قبلی برتری دارد و مقدار $F1$ به میزان $7/1$ درصد بهبود یافته است.

جدول ۱: نتایج ارزیابی روش پیشنهادی روی پیکره CoNLL-2012

روش	MUC			B3			CEAF_φ4			Avg.F1
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	
راگوناتان و سایرین [۱]	۲۵/۵	۵۸/۹	۳۵/۶	۲۲/۹	۶۴/۳	۳۳/۸	۳/۴	۵/۷	۴/۲	۶۵/۷
کلارک و مینگ [۲]	۷۶/۱	۶۹/۴	۷۲/۶	۶۵/۶	۵۶/۰	۶۰/۴	۵۹/۴	۵۳/۰	۵۶/۰	۶۳/۰
سایرین [۷]	۸۱/۴	۷۹/۵	۸۰/۴	۷۲/۲	۶۹/۵	۷۰/۸	۶۸/۲	۶۷/۱	۶۷/۶	۷۳/۰
سایرین [۴]	۸۵/۴	۷۷/۹	۸۱/۴	۷۷/۹	۶۶/۴	۷۱/۷	۷۰/۶	۶۶/۳	۶۸/۴	۷۳/۸
پیشنهادی	۸۴/۲	۸۹/۱	۸۶/۵	۷۸/۲	۸۱/۲	۷۹/۷	۷۴/۱	۷۹/۳	۷۶/۶	۸۰/۹

دلیل برتری روش پیشنهادی در نظر گرفتن اطلاعات معنایی در شبکه از پیش آموزش دیده RoBERTa است. با در نظر گرفتن تعبیه واژگان توسط این روش و استفاده از شبکه Bi-GRU جهت استخراج دنباله کلمات اطلاعات بهتری در اختیار سیستم قرار گرفته و سیستم توانایی بالاتری در تشخیص درست نامبری‌ها پیدا می‌کند. همچنین از آنجا که مهم‌ترین مسئله در تحلیل مرجع مشترک فقدان دانش است و در روش پیشنهادی سعی شده است، سیستم از این جهت تأمین شود، در نتیجه اطلاعات مفید معنایی و زبان‌شناسی بهتری در اختیار سیستم قرار می‌گیرد، بنابراین سیستم به‌خوبی آموزش دیده و باعث می‌شود در استخراج نامبری‌های درست و زنجیره‌های هم‌مرجع بهتر عمل کند. همان‌طور که در بخش ۲-۱ اشاره شد، جهت استخراج دنباله کلمات از تعبیه واژگان ورودی از شبکه GRU دوجهته استفاده شده است، جهت

بررسی کارایی این شبکه عملکرد شبکه‌های LSTM و LSTM، دوجهته نیز در این زمینه بررسی شد. در جدول (۲) قدرت این شبکه‌های بازگشتی در استخراج دنباله کلمات نشان داده شده است. همان‌طور که مشخص است شبکه GRU دوجهته دارای عملکرد بهتر است و توانایی بالاتری در استخراج وابستگی‌های طولانی دارد. در بعضی زمینه‌های پردازش متن نیز شبکه‌های GRU توانسته‌اند نسبت به شبکه‌های LSTM نتایج بهتری تولید کنند، دلیل این برتری این است که در این شبکه‌ها گیت‌های فراموشی و ورودی در گیت به‌روزرسانی ادغام شده‌اند که باعث بالاتر رفتن سرعت عملکرد آن‌ها می‌شود. همچنین استفاده از تعبیه واژگان RoBERTa باعث شده است که اطلاعات معنایی بهتری در اختیار این شبکه قرار گیرد و به دلیل داشتن سرعت بالای این شبکه‌ها دنباله کلمات با سرعت بالاتری استخراج شود؛ در نتیجه نامبری‌ها با دقت بالاتر استخراج شده و عملکرد سیستم بهبود داده می‌شود.

جدول ۲: توانایی استخراج دنباله کلمات توسط شبکه‌های عصبی عمیق بازگشتی

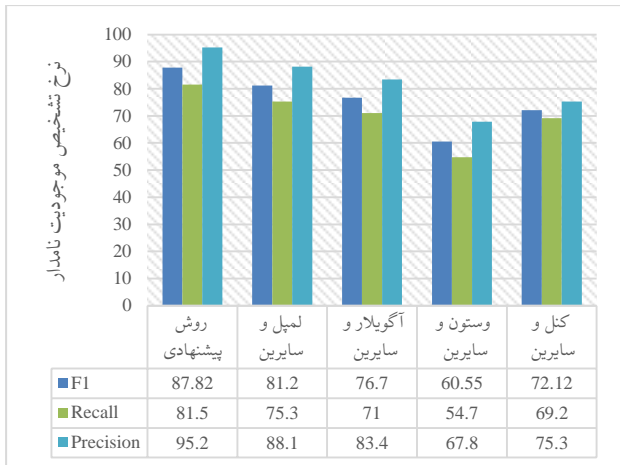
شبکه بازگشتی	Prec.	Rec.	F1
RNN	۷۹/۲۲	۸۳/۱۳	۸۱/۱۲
LSTM	۸۱/۴۹	۸۵/۹۹	۸۳/۶۷
Bidirectional LSTM	۸۴/۱۹	۹۰/۸۲	۸۷/۳۷
GRU	۸۵/۳۱	۹۳/۹۰	۸۹/۳۹
Bidirectional GRU	۸۹/۹۳	۹۷/۹۵	۹۳/۷۶

برای نمایش تأثیر نحوه استخراج تعبیه واژگان بر روی عملکرد روش پیشنهادی و مقدار $Avg.F1$ ، روش‌های مختلف تعبیه واژگان نظیر [۲۵] Word2Vec، [۱۳] GloVe، [۱۲] ELMo، BERT و RoBERTa در جدول (۳) روی مجموعه توسعه پیکره CoNLL-2012 با یکدیگر مقایسه شده‌اند. همان‌طور که مشخص است، کارکرد روش GloVe از روش Word2Vec اندکی بهتر است؛ زیرا در روش GloVe بردار اولیه بر اساس طول همجواری^{۱۵} با شمارش تعداد ظاهر شدن دو کلمه کنار هم در این بازه، تشکیل می‌شود و در مرحله بعد، بر اساس نزدیکی یا دوری دو کلمه به یکدیگر، این اعداد، وزن دهی می‌شوند. روش ELMo از روش GloVe بهتر عمل می‌کند؛ زیرا در ساختار خود از مدل‌های زبانی دوجهته و شبکه BiLSTM استفاده کرده و ویژگی‌های نحوی و معنایی و زبان‌شناسی بهتری در اختیار سیستم قرار می‌دهد. در نهایت روش RoBERTa بهتر از تمام روش‌های تعبیه واژگان عمل می‌کند زیرا این روش معماری روش BERT را بهبود داده و توانسته است ویژگی‌های بهتری نسبت به این روش ارائه دهد و در اکثر زمینه‌های پردازش متن نتایج قابل قبولی ارائه کند.

نوع موجودیت (مانند اشخاص، مکان‌ها و سازمان‌ها) به‌طور جداگانه محاسبه شود.

$$(14) \quad \text{دقت} = \frac{\text{تعداد موجودیت‌های درست استخراج شده}}{\text{تعداد کل موجودیت‌های استخراج شده}}$$

$$(15) \quad \text{بازخوانی} = \frac{\text{تعداد موجودیت‌های درست استخراج شده}}{\text{تعداد کل موجودیت‌های درست}}$$



شکل ۱: دقت و بازخوانی در تشخیص موجودیت‌های نامدار روی پیکره English Gigaword

۴- نتیجه‌گیری

در این مقاله، یک روش تحلیل مرجع مشترک در متن بر پایه منبع دانش و شبکه عصبی ارائه شد. مهم‌ترین مسئله در پایین بودن دقت سایر روش‌ها در این زمینه، فقدان دانش است. از آنجا که مسئله تحلیل مرجع مشترک دارای ابهامات زیادی است و هر مورد هم‌ارجاعی با دانشی خاص قابل حل است، در روش پیشنهادی سعی شد تا با منابع دانش مختلف نظیر دانش معنایی، زبان‌شناسی و پزشکی دقت این مسئله بهبود داده شود. جهت پیش‌آموزش شبکه از روش RoBERTa که جدیدترین روش آموزش شبکه تحت مدل‌های زبانی و یادگیری عمیق است، استفاده شد. استفاده از این روش و ساختار شبکه بازگشتی عمیق باعث شد نامبری‌ها با دقت بالاتری استخراج شوند. در این روش، انتخاب هم‌مرجع مناسب برای نامبری موردنظر بر اساس تمام معیارهای مؤثر در تحلیل هم‌ارجاعی انجام شد. شبکه عصبی عمیق عملکرد مدل را بهینه کرده و نیازی به صرف وقت اضافی برای تعیین وزن ویژگی‌ها ندارد؛ زیرا ویژگی‌های مفید تحت این ساختار استخراج می‌شوند. استفاده از فرآیند بازنمایش دانش نیز باعث همگرایی سریع‌تر روش پیشنهادی شد. در نهایت ترکیب همه این راهکارها باعث دقیق‌تر شدن تحلیل هم‌مرجعی در مقایسه با روش‌های پیشین شد. در این مقاله، مقایسه‌ی مناسبی بین روش پیشنهادی و روش‌های جدید تحلیل هم‌ارجاعی صورت گرفت. همچنین نشان داده شد که روش پیشنهادی با کمترین میزان خطا، مسئله تحلیل هم‌ارجاعی را به‌خوبی مدیریت می‌کند. سیستم پیشنهادی عملکردی مناسب در انواع مختلف دادگان دارد و معیار Avg.F1 در مجموعه دادگان CoNLL-2012 ۷/۱ درصد افزایش

جدول ۳: نتایج ارزیابی تعبیه واژگان روی پیکره توسعه CoNLL-2012

	Avg. F1	Δ
مدل پیشنهادی (تعبیه واژه RoBERTa)	۸۱/۷	-
تعبیه واژه Word2Vec	۷۴/۶	۷/۱
تعبیه واژه GloVe	۷۷/۳	۴/۴
تعبیه واژه ELMo	۷۸/۲	۳/۵
تعبیه واژه BERT	۷۹/۲	۲/۵

جهت بررسی میزان تأثیر هریک از منابع دانش بر عملکرد مدل پیشنهادی، تک‌تک این منابع را از مدل حذف کردیم و بررسی که کردیم که با حذف هریک از آن‌ها دقت روش پیشنهادی روی دو پایگاه داده [۱۷] CoNLL-2012 و [۱۸] MEDCo-A چه تغییری می‌کند. همان‌طور که در جدول (۴) مشخص است، هر نوع دانش روی پیکره‌ای خاص عملکرد متغیری دارد. برای مثال، وجود دانش زبان‌شناسی در پیکره CoNLL-2012 بالاترین تأثیر را نسبت به بقیه دانش‌ها دارد. درحالی‌که در پیکره پزشکی [۱۸] MEDCo-A حذف دانش پزشکی بیشترین تأثیر منفی را بر عملکرد مدل پیشنهادی می‌گذارد.

جدول ۴: نتایج حذف منابع دانش روی پیکره MEDCo-A و CoNLL-2012

پایگاه داده	CoNLL-2012		MEDCo-A	
	F1	ΔF1	F1	ΔF1
مدل به‌طور کامل	۸۰/۹	-	۸۸/۹	-
منبع دانش OMCS	۸۰/۱	-۰/۸	۸۸/۷	-۰/۲
منبع دانش پزشکی	۷۹/۱	-۱/۸	۸۸/۱	-۰/۸
منبع دانش زبان‌شناسی	۷۸/۳	-۲/۶	۸۸/۶	-۰/۳
منبع دانش معنایی	۷۸/۸	-۲/۱	۸۸/۴	-۰/۵

جهت ارزیابی عملکرد روش پیشنهادی در شناسایی موجودیت‌های نامدار، آن را با روش‌های جدید در این زمینه [۲۴-۲۷] مقایسه کرده‌ایم. مقدار دقت و بازخوانی در شکل (۱) توسط روابط (۱۴) و (۱۵) محاسبه شده است. از آنجا که در روش پیشنهادی از ساختار شبکه عصبی از پیش آموزش دیده در شناسایی نامبری‌ها استفاده شده و این نوع شبکه قابلیت استخراج خودکار ویژگی‌ها از موجودیت‌ها را دارد؛ این روش عملکرد بهتری در مقایسه با سایر روش‌ها دارد.

همان‌گونه که در شکل (۱) نشان داده شده است، مقدار معیار Avg.F1 برای روش پیشنهادی در مقایسه با روش‌های قبلی تا ۶/۶ درصد بهبود یافته است؛ زیرا استفاده از مدل‌های زبانی توسط روش پیش‌آموزش استفاده شده و تعبیه درست واژگان، اطلاعات معنایی خوبی رو در اختیار سیستم قرار داده است. روش‌های لمپل و آگویلار بعد از روش پیشنهادی به ترتیب بالاترین مقادیر مربوط به F1 را دارند. ترکیب این دو روش نیز می‌تواند، روش مناسبی برای تشخیص موجودیت‌ها باشد که لازم است در آن مقادیر دقت و بازخوانی برای هر

- [19] C. Napoles, M. Gormley, and B. Van Durme, "Annotated gigaword," in *Proceedings of the Joint Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction*, 2012, pp. 95–100.
- [20] M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman, "A model-theoretic coreference scoring scheme," in *Proceedings of the 6th conference on Message understanding*, 1995, pp. 45–52.
- [21] A. Bagga and B. Baldwin, "Algorithms for scoring coreference chains," in *The first international conference on language resources and evaluation workshop on linguistics coreference*, 1998, vol. 1, pp. 563–566.
- [22] X. Luo, "On coreference resolution performance metrics," in *Proceedings of the conference on human language technology and empirical methods in natural language processing*, 2005, pp. 25–32.
- [23] X. Rong, "word2vec parameter learning explained," *arXiv preprint arXiv:1411.2738*, 2014.
- [24] G. Aguilar, S. Maharjan, A. P. López-Monroy, and T. Solorio, "A multi-task approach for named entity recognition in social media data," *arXiv preprint arXiv:1906.04135*, 2019.
- [25] L. Canale, P. Lisena, and R. Troncy, "A Novel Ensemble Method for Named Entity Recognition and Disambiguation Based on Neural Network," in *International Semantic Web Conference*, 2018, pp. 91–107.
- [26] L. Weston *et al.*, "Named Entity Recognition and Normalization Applied to Large-Scale Information Extraction from the Materials Science Literature," *Journal of chemical information and modeling*, 2019.
- [27] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," *arXiv preprint arXiv:1603.01360*, 2016.

زیرنویس‌ها

- 1 Mention
- 2 Co-referent
- 3 Span representation
- 4 Vector representation
- 5 Span
- 6 Commonsense knowledge graph
- 7 Parser
- 8 Predicate
- 9 Argument
- 10 Knowledge representation
- 11 String match
- 12 Head
- 13 Tail
- 14 Element-wise multiplication
- 15 Window Size

یافته است. در آینده قصد داریم، مدل ارائه‌شده را با افزودن منابع دانش بیشتر، جهت افزایش دقت تشخیص زنجیره‌های هم‌مرجع بهبود دهیم.

مراجع

- [1] K. Raghunathan *et al.*, "A multi-pass sieve for coreference resolution," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 2010, pp. 492–501.
- [2] K. Clark and C. D. Manning, "Entity-centric coreference resolution with model stacking," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2015, vol. 1, pp. 1405–1415.
- [3] K. Clark and C. D. Manning, "Deep Reinforcement Learning for Mention-Ranking Coreference Models," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 2256–2262.
- [4] H. Fei, X. Li, D. Li, and P. Li, "End-to-end Deep Reinforcement Learning Based Coreference Resolution," in *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 2019, pp. 660–665.
- [5] R. Zhang, C. N. dos Santos, M. Yasunaga, B. Xiang, and D. Radev, "Neural coreference resolution with deep biaffine attention by joint mention detection and mention clustering," *arXiv preprint arXiv:1805.04893*, 2018.
- [6] K. Lee, L. He, M. Lewis, and L. Zettlemoyer, "End-to-end Neural Coreference Resolution," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 188–197.
- [7] K. Lee, L. He, and L. Zettlemoyer, "Higher-order coreference resolution with coarse-to-fine inference," *arXiv preprint arXiv:1804.05392*, 2018.
- [8] M. Joshi, O. Levy, D. S. Weld, and L. Zettlemoyer, "BERT for Coreference Resolution: Baselines and Analysis," *arXiv preprint arXiv:1908.09091*, 2019.
- [9] B. Kantor and A. Globerson, "Coreference Resolution with Entity Equalization," in *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 2019, pp. 673–677.
- [10] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," *arXiv preprint arXiv:1907.11692*, 2019.
- [11] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [12] M. Peters *et al.*, "Deep Contextualized Word Representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 2018, vol. 1, pp. 2227–2237.
- [13] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.
- [14] P. Singh, T. Lin, E. T. Mueller, G. Lim, T. Perkins, and W. L. Zhu, "Open Mind Common Sense: Knowledge acquisition from the general public," in *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, 2002, pp. 1223–1237.
- [15] O. Uzuner, A. Bodnari, S. Shen, T. Forbush, J. Pestian, and B. R. South, "Evaluating the state of the art in coreference resolution for electronic medical records," *Journal of the American Medical Informatics Association*, vol. 19, no. 5, pp. 786–791, 2012.
- [16] J. R. Hobbs, "Resolving pronoun references," *Lingua*, vol. 44, no. 4, pp. 311–338, 1978.
- [17] S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, and Y. Zhang, "CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes," in *Joint Conference on EMNLP and CoNLL-Shared Task*, 2012, pp. 1–40.
- [18] J. L. Raisaro *et al.*, "MedCo: Enabling Privacy-Conscious Exploration of Distributed Clinical and Genomic Data," in *4th International Workshop on Genome Privacy and Security (GenoPri'17)*, 2017.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه صنعتی شاهرود

۲۷ الی ۲۸ آذرماه ۱۳۹۸

ارائه یک معماری یادگیری عمیق برای شناسایی اعمال انسان در ویدئو

مهدی سوادى حسینی^۱، فواد قادری^۲*

^۱ دانشجوی کارشناسی ارشد، دانشگاه تربیت مدرس، mehdi.savadihosseini@modares.ac.ir

^۲ استادیار دانشکده مهندسی برق و کامپیوتر، آزمایشگاه تعامل انسان و کامپیوتر، دانشگاه تربیت مدرس، fghaderi@modares.ac.ir

چکیده

شناسایی اعمال انسان در ویدئو با توجه به کاربردهای بسیاری که در زمینه‌های مختلف از جمله امنیت، سلامت، مدیریت هوشمند شهر و ساختمان و ... دارد، طی چند سال اخیر توجه زیادی را به خود جلب کرده است. از سوی دیگر، رویکردهای یادگیری عمیق مانند شبکه‌های عصبی پیچشی (CNN) و شبکه‌های عصبی بازگشتی (RNN) در زمینه‌های مختلف توانسته‌اند نتایج بسیار خوبی را کسب کنند. با این وجود، علی‌رغم تغییرات زمانی قابل توجه در محتوای ویدئو، تاکنون شبکه‌های عصبی بازگشتی در شناسایی اعمال انسان ضعیف‌تر از روش‌های مبتنی بر شبکه‌های پیچشی عمل کرده‌اند. در این پژوهش با معرفی یک رویکرد سلسله‌مراتبی زمانی در دو سطح محلی و سراسری، با استفاده از شبکه‌های پیچشی متورم دو جریانی که از شبکه‌های عصبی پیچشی دوبعدی معروف برای دسته‌بندی تصاویر ایجاد شده‌اند و همچنین استفاده از پشته‌ای از لایه‌های واحدهای بازگشتی گیت‌دار (GRU) توانسته‌ایم رویکردی جدید برای شناسایی اعمال انسان در ویدئو ارائه دهیم. در این رویکرد ابتدا ویژگی‌های فضایی-زمانی را برای هر دو جریان فضایی و زمانی به صورت محلی با استفاده از شبکه‌های پیچشی سه‌بعدی متورم (I3D) فضایی و زمانی استخراج می‌کنیم که این کار باعث تبدیل دنباله طولانی قاب‌های یک ویدئو به یک دنباله کوتاه‌تر و حاوی اطلاعات مفیدتر می‌شود. این دنباله کوتاه‌شده را به شبکه‌ای از GRU ها می‌دهیم و در نهایت نیز نتایج دو جریان را تجمیع می‌کنیم. برای تجمیع نتایج، یک لایه جدید با نام میانگین وزن‌دهی شده را معرفی می‌کنیم که اهمیت هر جریان را در فرایند آموزش به صورت خودکار فرا می‌گیرد. ارزیابی‌ها نشان‌دهنده نتایجی قابل قبول برای دو مجموعه داده HMDB51 و UCF101 هستند. روش پیشنهادی موجب ۱,۶ درصد بهبود در صحت دسته‌بندی نمونه‌های مجموعه داده پرچالش HMDB51 نسبت به نتایج گزارش شده بهترین روش موجود گردیده است.

کلمات کلیدی

شبکه‌های عصبی پیچشی متورم، واحد بازگشتی گیت‌دار، شناسایی اعمال، معماری دو جریانی.

۱- مقدمه

بسیاری از جمله نظارت و امنیت، راهبری خودروهای خودران، بینایی ربات، مدیریت ساختمان و شهر هوشمند و ... دارد. تلاش‌های بسیاری برای شناسایی اعمال انسان در ویدئو صورت گرفته است. در این تلاش‌ها از دو رویکرد اصلی استفاده شده است: رویکردهای مبتنی بر ویژگی‌های دست‌ساز^۲ و رویکردهای مبتنی بر یادگیری عمیق. Bobick و همکارانش [۲] از تصاویر انرژی حرکت^۳ و تصاویر تاریخچه حرکت^۴ برای شناسایی اعمال استفاده می‌کردند. Ho و همکارانش [۳] با تفکیک پیش‌زمینه تصاویر تاریخچه حرکت،

پژوهش‌های آماری نشان می‌دهد که در هر دقیقه بیش از ۵۰۰ میلیون ویدئو در یوتیوب بارگذاری می‌شود [۱]. افزایش روز افزون محتوای ویدئویی، لزوم ایجاد سیستم‌های هوشمندی که قادر به فهم محتوای ویدئوها باشند را بیش از پیش ایجاد می‌کند. شناسایی اعمال انسان^۱ در ویدئو یکی از مهمترین بخش‌های چنین سیستمی است. با وجود تلاش‌های بسیار در این زمینه، شناسایی اعمال انسان همچنان یک کار پرچالش محسوب می‌شود. شناسایی اعمال انسان در ویدئو کاربردهای

² Hand-crafted

³ Motion energy image

⁴ Motion history image

¹ Human action recognition

نمونه سه‌بعدی آن‌ها کرد. Carreira و Zisserman [۲۳] با متورم‌سازی معماری Inception-V1 [۲۴] به نوع سه‌بعدی آن و با اعمال اندکی تغییرات، توانستند به نتایج خوبی برسند.

هر چند شبکه‌های عصبی پیچشی سه‌بعدی ویژگی‌های فضایی-زمانی را در کنار هم فرا می‌گیرند، اما به ترتیب زمانی پیشامدها حساس نیستند. اعمال در ویدئوهای مختلف بدون توجه به اینکه با سرعت‌های متفاوتی انجام می‌گیرند، معمولاً از ترتیب مشخصی پیروی می‌کنند. برای همین شبکه‌های عصبی بازگشتی^{۱۱} می‌توانند بسیار مفید واقع شوند. یکی از انواع این شبکه‌ها حافظه ماندگار کوتاه مدت (LSTM) [۲۵] است. Yong و همکارانش [۲۶] یک رویکرد سلسله مراتبی مبتنی بر شبکه‌های عصبی بازگشتی ارائه کردند. با وجود اینکه LSTM ها در سایر زمینه‌ها مانند پردازش صوت و متن بسیار خوب عمل کرده‌اند [۲۷-۳۱]، در شناسایی اعمال انسان نتوانسته‌اند نتایج خوبی داشته باشند. Joe و همکارانش [۳۲] از یک لایه LSTM در انتهای شبکه CNN استفاده کردند. آن‌ها همچنین از یک لایه ادغام زمانی نیز در شبکه‌ای مشابه استفاده کردند و به این نتیجه رسیدند که لایه ادغام زمانی از LSTM بهتر عمل می‌کند. ما در این پژوهش با بررسی مشکلات شبکه‌های عصبی بازگشتی، یک رویکرد با کارایی بالا را ارائه می‌دهیم.

در ادامه در بخش دوم به توصیف روش پیشنهادی پرداخته می‌شود. بخش سوم را به ارزیابی اختصاص داده‌ایم که با معرفی مجموعه داده‌ها آغاز شده و با مقایسه نتایج با سایر پژوهش‌ها به اتمام می‌رسد. در پایان در بخش چهارم جمع‌بندی ارائه می‌شود.

۲- روش پیشنهادی

روش پیشنهادی در شکل ۱ نشان داده شده است. در این روش ابتدا با استفاده از شبکه‌های عصبی پیچشی سه‌بعدی متورم^{۱۲} (I3D) ویژگی‌های فضایی-زمانی محلی استخراج می‌شود. سپس دنباله‌ای بردارهای یادگرفته شده خروجی از I3D به شبکه‌ای چند لایه از سلول‌های GRU داده می‌شود و در نهایت نتایج دو جریان فضایی و زمانی با هم در لایه‌ی میانگین وزن‌دهی شده^{۱۳} تجمیع می‌شود. معماری پیشنهادی به صورت انتها به انتها قابل یادگیری می‌باشد.

زیردنباله‌های فضایی-زمانی

همانطور که در شکل ۱ مشاهده می‌شود، ویدئو با تعداد قاب‌های N به T بخش مساوی (احتمالاً به جز بخش آخر) تقسیم می‌شود. به دنباله‌ای که از این بخش‌ها ایجاد می‌شود، زیردنباله‌های فضایی-زمانی می‌گوییم. تقسیم دنباله‌ی طولانی ویدئو به دنباله‌ای از زیر دنباله‌های کوتاه فضایی-زمانی می‌تواند پیچیدگی ویژگی‌های فضایی-زمانی را کاهش دهد و به فرایند یادگیری کمک کند. استخراج زیر دنباله‌های فضایی-زمانی به دو صورت انبوه و خلوت می‌تواند انجام گیرد. در حالت خلوت از قاب t تا $t+L$ (تعداد قاب‌هایی است که همزمان به صورت یک تنسور

هیستوگرام جهت‌دار گرادیان^۵ را برای شناسایی اعمال استخراج می‌کردند. Roh و همکارانش [۴] تصاویر تاریخیچه حرکت را از دوبعدی به سه‌بعدی گسترش دادند. مسیرهای متراکم^۶ [۵] نقطه عطفی در روش‌های مبتنی بر ویژگی‌های دست‌ساز بود. این مسیرها با استفاده از نمونه‌گیری متراکم و ضبط ویژگی‌های حرکتی محلی استخراج می‌شوند. Wang و همکارانش [۶] با تخمین حرکات دوربین مسیرهای بهبود یافته را ارائه دادند. مسیرهای متراکم برای محاسبه نیاز به محاسبات سنگین دارد، Eleonora و همکارانش [۷] با شناسایی نقاط ساکن، مسیرهای مربوط به آنجا را حذف کردند. محبوب‌ترین رویکرد مبتنی بر ویژگی‌های دست‌ساز، کیف کلمات بصری^۷ [۸، ۹] یا انواع برگرفته از آن است [۱۰، ۱۱]. کیف کلمات بصری از ویژگی‌های ارزشمندی همچون مسیرهای متراکم و همچنین روش‌های کدگذاری قدرتمند مانند بردارهای فیشر^۸ [۱۱] استفاده می‌کند. Peng و همکارانش [۱۲] به جای یک لایه استفاده از بردارهای فیشر از پشت‌های از بردارهای فیشر استفاده می‌کنند. Fernando و همکارانش [۱۳] رویکردی تابعی برای استخراج تکامل زمانی حرکت ارائه کردند. گسترش این رویکرد به صورت سلسله مراتبی انجام شد [۱۴].

در حالی که رویکردهای مبتنی بر ویژگی‌های دست‌ساز رو به رشد بودند، یادگیری عمیق در زمینه‌هایی همچون پردازش تصاویر و صدا به موفقیت‌های بزرگی دست یافته بود. پژوهشگران حوزه شناسایی اعمال با الهام از رویکردهای دسته‌بندی تصاویر، معماری‌های مختلفی برای شناسایی اعمال ارائه کردند. Simonyan و همکارانش [۱۵] معماری دو جریانی که از یک قاب ویدئو برای جریان فضایی و پشت‌های از جریان نوری^۹ برای جریان زمانی استفاده می‌کرد، ارائه کردند. جریان فضایی با استفاده از مجموعه داده بسیار بزرگ ImageNet [۱۶] آموزش داده می‌شود. شبکه‌های تقسیم زمانی [۱۷، ۱۸] با عمیق‌تر کردن معماری دو جریانی آن را گسترش دادند. با وجود اینکه این معماری از تعداد قاب‌های بیشتری (۳ یا ۷) استفاده می‌کند، اما باز هم این تعداد برای شناسایی اعمال پیچیده و کلاس‌های نزدیک به هم کافی نیستند. رویکرد رمزگذاری خطی زمانی [۱۹] برای رفع این مشکل ارائه شد. Zhu و همکارانش [۲۰] با ارائه لایه ادغام هرم‌های زمانی، برای استخراج ویژگی‌های زمانی، قاب‌های بیشتری را استفاده کردند و به نتایج بهتری رسیدند.

اگر ویدئوها را به صورت مکعب‌های سه‌بعدی در نظر بگیریم^{۱۰} شبکه‌های عصبی پیچشی^{۱۰} (CNN) سه‌بعدی به خوبی می‌توانند ویژگی‌های فضایی-زمانی را فرا بگیرند. Ji و همکارانش [۲۱] یک معماری مبتنی بر شبکه‌های پیچشی سه‌بعدی ارائه کردند که به علت فیلترهای بزرگ هزینه‌ی محاسباتی زیادی را در پی داشت. پژوهشگران با استفاده از فیلترهای $3 \times 3 \times 3$ در کلیه لایه‌ها هزینه محاسباتی را کاهش دادند [۲۲]. موفقیت شبکه‌های عمیق در دسته‌بندی تصاویر دوبعدی، پژوهشگران را تشویق به تبدیل معماری‌های موفق دوبعدی به

⁵ Histogram of oriented gradient

⁶ Dense trajectories

⁷ Bag of visual words

⁸ Fisher vectors

⁹ Optical flow

¹⁰ Convolutional Neural Networks

¹¹ Recurrent Neural Networks

¹² Inflated 3D CNN

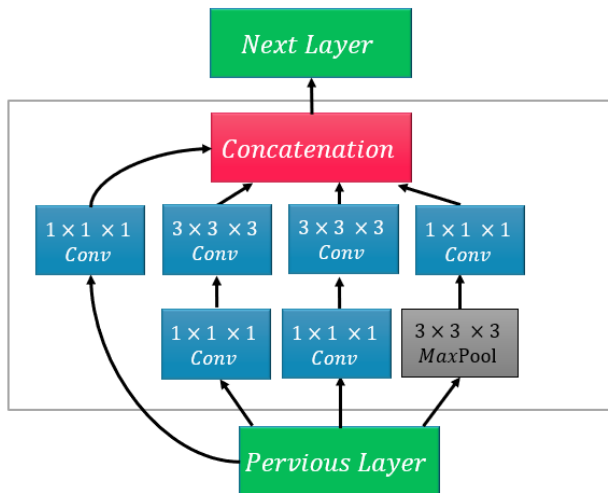
¹³ Weighted average

۲-۴- لایه تجمیع

همانطور که اشاره شد، از دو جریان فضایی و زمانی در شبکه استفاده میکنیم. در بسیاری از رویکردهای مطرح اخیر، پیشبینی‌های تولید شده در هر دو جریان در نهایت و بیرون از شبکه تجمیع می‌شوند، که تقریباً تمام این رویکردها از میانگین برای این کار استفاده می‌کنند. در حالت دو جریانی میانگین به معنای ضرب نتیجه‌ی هر جریان در عدد ۰,۵ و سپس جمع هر دو جریان است. در این حالت تاثیر هر دو جریان فضایی و زمانی به یک اندازه است اما آیا واقعا این چنین است. می‌توان از میانگین وزندهی شده برای این کار استفاده کرد به این صورت که به صورت یکسان، دو جریان ارزشگذاری نشوند. میانگین وزندهی شده به صورت زیر تعریف می‌شود:

$$WeightedAverage = \alpha \times A + \beta \times B \quad \# \quad (1)$$

باید توجه داشت که مجموع پارامترهای α و β برابر یک است، در واقع می‌توان گفت که $\alpha = 1 - \beta$ و به یافتن پارامتر α پرداخت. می‌توان این ضریب را به صورت دستی مقداردهی کرد اما ما این کار را به خود شبکه سپرده‌ایم و یک لایه جدید در شبکه برای اینکار تعریف کرده‌ایم. تنها پارامتر این لایه α است که در فرایند یادگیری فراگرفته می‌شود. پارامترهای α و β با ضرب نقطه‌ای در بردارهای A و B که خروجی دو جریان فضایی و زمانی هستند اعمال می‌شوند.

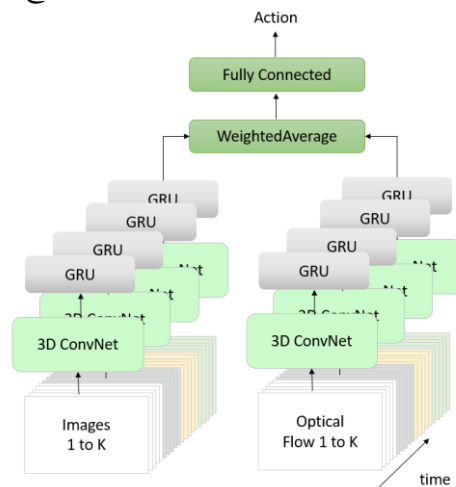


شکل 42. ماژول Inception متورم. شکل بالا ماژول اصلی استفاده شده در معماری I3D روش پیشنهادی ما را تشکیل می‌دهد.

۳- ارزیابی

در این بخش به ارزیابی روش پیشنهادی می‌پردازیم. ابتدا مجموعه داده‌های مهم معرفی می‌گردند. در ادامه به تاثیر نوآوری‌های روش پیشنهادی می‌پردازیم و در انتهای این بخش نیز کارایی روش پیشنهادی با روش‌های دیگر مورد بررسی قرار می‌گیرد.

سه‌بعدی به I3D داده می‌شوند) و سپس از $t+L+1$ تا $t+2L+1$ و ... به صورت غیر تکراری به شبکه داده می‌شود و در حالت انبوه از قاب t تا $t+L$ و سپس از $t+1$ تا $t+L+1$ و ... دنباله‌های محلی استخراج می‌شوند.



شکل 41. روش پیشنهادی. در ابتدا برای هر دو قاب متوالی جریان نوری محاسبه می‌شود، سپس برای هر دو دنباله، زیر دنباله‌های فضایی-زمانی محاسبه می‌شود، ویژگی‌های فضایی-زمانی هر زیر دنباله به طور مجزا توسط I3D استخراج می‌شود. دنباله کوتاه جدید به شبکه‌ای از سلول‌های GRU داده می‌شود و در نهایت نتایج دو جریان در لایه میانگین وزن‌دار تجمیع می‌شوند.

۲-۲- I3D

در طی سالیان اخیر معماری‌های بسیار موفق در زمینه‌ی دسته‌بندی تصاویر ارائه شده‌اند که بخش زیادی از طراحی آن‌ها با آزمایش و خطا صورت گرفته است. ما به جای تکرار این روند در دسر ساز برای مدل‌های فضایی-زمانی به منظور استخراج ویژگی‌های فضایی-زمانی محلی، به سادگی مدل‌های موفق دسته‌بندی تصاویر را از دوبعدی به سه بعدی تبدیل می‌کنیم. این کار با تکرار فیلترهای دوبعدی در راستای بعد زمان صورت می‌پذیرد. به این صورت که فیلترهای مربعی به مکعب تبدیل می‌شوند. ما در این پژوهش از معماری I3D ارائه شده توسط Inception و Carreira و Zisserman [۲۳] استفاده می‌کنیم، ماژول Inception استفاده شده در این معماری در شکل ۲ نشان داده شده است.

۲-۳- لایه‌های بازگشتی

بعد از استخراج ویژگی‌های زیردنباله‌ها توسط I3D، دنباله‌ای از بردارهای ویژگی بدست می‌آید که ترتیب زمانی حرکات محلی را در خود دارند. این دنباله را به شبکه‌ای از سلول‌های واحد بازگشتی گیت‌دار^{۱۴} (GRU) می‌دهیم. علت استفاده از GRU به جای LSTM تعداد پارامترهای کمتری است که GRU دارد. باید توجه داشت که GRU با تعداد پارامتر کمتر کارایی نزدیک به LSTM دارد.

¹⁴ Gated Recurrent Unit

۱-۳- مجموعه داده‌ها

زیردنباله‌های فضایی-زمانی. در ارزیابی روش پیشنهادی از روش خلوت برای استخراج زیردنباله‌ها استفاده کرده‌ایم، علت این کار کاهش بار محاسباتی است. برای ارزیابی تمام مقادیر برای L بررسی نگردیده است و تنها به چند حالت پرداخته شده است. در اینصورت نتیجه زیر بهینه خواهد بود.

تاثیر استفاده از GRU. در روش پیشنهادی ما از GRU به جای LSTM استفاده شده است که علت این کار تعداد پارامترهای کمتر GRU است. همانطور که در جدول ۳ نشان داده شده است، دو لایه GRU در مقایسه با LSTM نتایج بهتری را خواهد داشت. در مورد مجموعه داده‌هایی که استفاده کرده‌ایم، به علت اینکه تعداد نمونه‌ها کم است، تعداد پارامترهای زیاد باعث بیش برآزش خواهد شد.

جدول ۱. ارزیابی تاثیر طول زیر دنباله‌های فضایی-زمانی بر مجموعه داده HMDB51.

L=20	L=15	L=10	
۷۶,۰۷	۸۳,۸۹	۸۴,۵۹	جداسازی اول
۷۷,۰۳	۸۳,۵۹	۸۲,۹۱	جداسازی دوم
۷۸	۸۴,۰۸	۸۳,۰۱	جداسازی سوم

جدول ۲. ارزیابی تاثیر طول زیردنباله‌های فضایی-زمانی بر مجموعه داده UCF101.

L=20	L=15	L=10	
۹۱,۰۲	۹۵,۶۵	۹۴,۹۵	جداسازی اول
۹۲	۹۶,۸۷	۹۶,۴۸	جداسازی دوم
۹۱,۲۶	۹۵,۷۶	۹۴,۴۹	جداسازی سوم

جدول ۳. نتایج لایه‌های مختلف بازگشتی بر روی جداسازی اول مجموعه داده HMDB51.

صحت	پارامتر	معماری
۸۴,۵۹	۲,۷۶۸,۶۹۲	۲ لایه GRU به عرض ۲۵۶
۷۸,۹۱	۳,۶۸۷,۲۲۰	۲ لایه LSTM به عرض ۲۵۶
۸۲,۳۲	۶,۳۲۱,۷۱۶	۱ لایه LSTM به عرض ۵۱۲

۳-۴- مقایسه با سایر پژوهش‌ها

در این بخش بهترین نتایج بدست آمده را با سایر پژوهش‌ها مقایسه می‌کنیم. ابتدا بهترین نتایجی که از ارزیابی‌های پیشین بدست آوردیم را گزارش می‌دهیم. میانگین بهترین نتایج بر روی سه جداسازی استاندارد برای هر مجموعه داده محاسبه می‌شود. پژوهش‌های مورد مقایسه، پژوهش‌هایی هستند که اخیراً نتایج خوبی را بر روی مجموعه داده‌های معرفی شده داشته‌اند. بسیاری از این روش‌ها از یادگیری عمیق استفاده کرده‌اند. در جدول ۶ نتایج روش پیشنهادی با سایر پژوهش‌ها بر روی مجموعه داده HMDB51 مقایسه گردیده است، همانطور که مشاهده می‌شود،

برای ارزیابی روش پیشنهادی از دو مجموعه داده HMDB51 [۳۳] و UCF101 [۳۴] استفاده می‌کنیم. مجموعه داده HMDB51 از فیلم‌های واقعی و ویدئوهای یوتیوب ایجاد شده است و شامل ۶۷۶۶ ویدئو در ۵۱ کلاس است. مجموعه داده UCF101 از ۱۳۳۲۰ ویدئو در ۱۰۱ کلاس مختلف تشکیل شده است. برای هر کدام از این مجموعه‌های داده‌ای سه جداسازی آموزش/آزمایش مجزا ارائه گردیده است. نتایج بر روی این سه مجموعه داده با معیار صحت^{۱۵} گزارش می‌گردد تا با نتایج پژوهش‌های دیگر قابل مقایسه باشد.

۳-۲- جزئیات پیاده‌سازی

برای پیاده‌سازی روش پیشنهادی از کتابخانه Keras بر روی بستر متن‌باز Tensorflow استفاده کرده‌ایم. کراس یک کتابخانه متن‌باز است که امکان پیاده‌سازی سریع مدل‌های یادگیری عمیق را فراهم می‌کند. همچنین از کتابخانه‌های Numpy برای کار با داده‌ها و OpenCV برای پردازش ویدئوها در سطح پایین و همچنین از یک کارت گرافیکی NVIDIA GeForce GTX 1080 Ti استفاده کرده‌ایم. برای استخراج جریان نوری از الگوریتم TV-L1 [۳۵] که از تمام پیکسل‌ها استفاده می‌کند و برای بهینه‌سازی از الگوریتم RMSProp استفاده می‌کنیم. نرخ یادگیری این الگوریتم ۰,۰۰۱ و نرخ کاهش نیز ۰,۰۱ است. همچنین از دو لایه GRU به عرض ۲۵۶ نورون برای ارزیابی اولیه استفاده شده است. خروجی ماژول‌های I3D با اعمال میانگین سراسری^{۱۶} تبدیل به بردار خواهد شد و به لایه‌های GRU داده می‌شود.

۳-۳- بررسی تاثیر نوآوری‌ها

در این بخش به بررسی نقش نوآوری‌ها در کارایی مدل می‌پردازیم. ابتدا نقش لایه میانگین وزن‌دار بر کارایی مدل مورد ارزیابی قرار می‌گیرد. در ادامه زیردنباله‌های فضایی-زمانی و سپس تاثیر استفاده از GRU به جای LSTM بررسی می‌شود.

لایه تجمیع. استفاده از میانگین برای تجمیع کاری رایج است. شاید این پرسش مطرح شود که چرا از بیشه یا کمینه و ... استفاده نمی‌شود؟ استفاده از میانگین به روش تجربی و آزمون و خطا در تقریباً همه معماری‌های دوجریانی به اثبات رسیده است. وقتی میانگین می‌گیریم یعنی تاثیر دو جریان فضایی و زمانی را یکسان فرض کرده‌ایم، در حالیکه آیا واقعا این چنین است؟ ما میزان تاثیر هر یک از جریان‌ها را به شبکه سپرده‌ایم که در فرایند یادگیری بسته به نوع مجموعه داده خود شبکه بیاموزد که چه ضریبی را برای هر جریان در نظر بگیرد. با اعمال این روش بر روی جداسازی اول از مجموعه داده HMDB51 به صحت ۸۴,۵۹ رسیدیم که افزایش محسوسی نسبت به حالتی که از میانگین معمولی استفاده می‌کنیم داشته است. در حالت میانگین معمولی صحت ۸۳,۳۲ است.

¹⁵ Accuracy

¹⁶ Global average pooling

به علت اینکه مجموعه داده UCF101 کلاس‌های بیشتری دارد و برخی کلاس‌ها بسیار به هم شبیه هستند کل شبکه باید به صورت انتها به انتها آموزش داده شود. با تنظیم دقیق حداقل لایه‌های انتهایی I3D می‌توان بر این مشکل غلبه کرد و کارایی بهتری را بدست آورد.

۴- نتیجه‌گیری

در این پژوهش با توسعه یک معماری جدید عمیق که از معماری‌های پایه‌ای دسته‌بندی تصاویر استفاده می‌کند، توانستیم با چالش‌های موجود بر سر راه شبکه‌های عصبی بازگشتی در شناسایی اعمال انسان در ویدئو مقابله کنیم. همچنین موفق شدیم با استفاده از معماری ترکیبی از شبکه‌های عصبی پیچشی سه بعدی که از نمونه موفق دو بعدی خود ایجاد شده‌اند، ویژگی‌های محلی و با استفاده از پشت‌های از شبکه‌های عصبی بازگشتی ویژگی‌های سراسری را استخراج کنیم. شبکه‌های عصبی پیچشی متورم که در سطح پایین قرار دارند، با استخراج ویژگی‌های مفید فضایی-زمانی به صورت محلی، دنباله‌ی قاب‌های ویدئو را به دنباله‌ای کوتاه‌تر تبدیل می‌کنند که این امر سبب می‌شود شبکه‌های بازگشتی بتوانند بهتر ویژگی‌های سرتاسری را فرا بگیرند. برای کاهش تعداد پارامترهای شبکه، به جای LSTM از GRU که در اکثر موارد با کارایی مشابه با LSTM و در عین حال تعداد پارامترهای کمتر عمل می‌کند استفاده کرده‌ایم. رویکرد پیشنهادی ما باعث بهبود ۱,۶ درصدی صحت روی مجموعه داده پر چالش HMDB51 در مقایسه با بهترین روش موجود گردید. در مورد مجموعه داده UCF101 نیز می‌توان با یادگیری انتها به انتها کل شبکه به نتایج بهتری دست یافت.

مراجع

- [1] J. Loke Hale 2019, More Than 500 Hours Of Content Are Now Being Uploaded To YouTube Every Minute, TubeFilter, viewed 7 May 2019, <https://www.tubefilter.com/2019/05/07/number-hours-video-uploaded-to-youtube-per-minute/>.
- [2] A. F. Bobick, and J. W. Davis, "The Recognition of Human Movement Using Temporal Templates," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 3, pp. 257-267, 2001.
- [3] Hu, Y., Cao, L., Lv, F., Yan, S., Gong, Y., & Huang, T. S. (2009, September). Action detection in complex scenes with spatial and temporal ambiguities. In 2009 IEEE 12th International Conference on Computer Vision (pp. 128-135). IEEE.
- [4] M.-C. Roh, H.-K. Shin, and S.-W. Lee, "View-independent human action recognition with Volume Motion Template on single stereo camera," Pattern Recogn. Lett., vol. 31, no. 7, pp. 639-647, 2010.
- [5] H. Wang, A. Kläser, C. Schmid, and L. Cheng-Lin, "Action Recognition by Dense Trajectories," IEEE Conference on Computer Vision & Pattern Recognition, 06/20, 2011.
- [6] Wang, H., & Schmid, C. (2013). Action recognition with improved trajectories. In Proceedings of the IEEE international conference on computer vision (pp. 3551-3558).
- [7] E. Vig, M. Dorr, and D. Cox, "Space-Variant Descriptor Sampling for Action Recognition Based on Saliency and Eye Movements," Computer Vision – ECCV 2012, pp. 84-97.
- [8] X. Peng, L. Wang, X. Wang, and Y. Qiao, "Bag of visual words and fusion methods for action recognition," Comput. Vis. Image Underst., vol. 150, no. C, pp. 109-125, 2016.
- [9] J. Sivic, and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," in Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2, 2003, pp. 1470.

نتیجه روش پیشنهادی ما با ۱,۶ درصد اختلاف بالاتر از بهترین روش موجود است.

جدول 13. بهترین نتایج بر روی دو مجموعه داده HMDB51 و UCF101.

UCF101	HMDB51	
۹۵,۶۵	۸۴,۵۹	جداسازی اول
۹۶,۸۷	۸۳,۵۹	جداسازی دوم
۹۵,۷۶	۸۴,۰۸	جداسازی سوم
۹۶,۰۹	۸۴,۰۸	میانگین

جدول ۶. مقایسه نتایج برای مجموعه داده HMDB51.

روش	توضیحات	سال	صحت
روش پیشنهادی	I3D + GRU	۲۰۱۹	۸۴,۰۸
لی وانگ و همکاران [۳۶]	HAF+BOW/FV	۲۰۱۹	۸۲,۴۸
ژو و همکاران [۲۰]	DTPP	۲۰۱۸	۸۲,۱
ژو وانگ و همکاران [۳۷]	SVMP + I3D	۲۰۱۸	۸۱,۳
روی و همکاران [۳۸]	-	۲۰۱۸	۸۱,۱
کریرا و زیسرمن [۲۳]	Two-Stream I3D + Kinetics pre-training	۲۰۱۷	۸۰,۷
فیزا و همکاران [۳۹]	iDT+DA-VLAD	۲۰۱۸	۸۰,۱
تران دو و همکاران [۴۰]	-	۲۰۱۸	۷۸,۷
ژویی و همکاران [۴۱]	-	۲۰۱۸	۷۸,۷
شوانگ و همکاران [۴۲]	-	۲۰۱۸	۷۴,۲

جدول ۷. مقایسه نتایج برای مجموعه داده UCF101

روش	توضیحات	سال ارائه	صحت
روش پیشنهادی	-	۲۰۱۹	۹۶,۰۹
چوتاس و همکاران [۴۳]	I3D + PoTion	۲۰۱۸	۹۸,۲
کریرا و زیسرمن [۲۳]	Two-Stream I3D + Kinetics pre-training	۲۰۱۷	۹۸
ژو و همکاران [۲۰]	DTPP	۲۰۱۸	۹۸
علی دیا و همکاران [۴۴]	HATNet (32 frames)	۲۰۱۹	۹۷,۷
تران دو و همکاران [۴۰]	-	۲۰۱۸	۹۷,۳
ژویی و همکاران [۴۱]	Hidden Two-Stream (I3D)	۲۰۱۸	۹۷,۱
یانگ و همکاران [۴۵]	Full IF-TNN	۲۰۱۹	۹۶,۲
شویانگ و همکاران [۴۲]	-	۲۰۱۸	۹۶
ژینگ و همکاران [۴۶]	S-TPNet + iDT	۲۰۱۹	۹۶

در جدول ۷ نتیجه روش ما با روش‌های دیگر بر روی مجموعه داده‌ی UCF101 مقایسه گردیده است. در مورد این مجموعه داده رویکرد ما نتوانسته است بهترین کارایی را داشته باشد، که علت اصلی این است که شبکه I3D استفاده شده در رویکرد ما، تنظیم دقیق^{۱۷} نشده است. این نتیجه می‌توانست در مورد مجموعه داده HMDB51 هم بدست آید اما

¹⁷ Fine tune

- [32] J. Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, "Beyond short snippets: Deep networks for video classification." pp. 4694-4702.
- [33] Kuehne, H., Jhuang, H., Garrote, E., Poggio, T., & Serre, T. (2011, November). HMDB: a large video database for human motion recognition. In 2011 International Conference on Computer Vision (pp. 2556-2563). IEEE.
- [34] K. Soomro, A. Roshan Zamir, and M. Shah, "UCF101: A Dataset of 101 Human Actions Classes From Videos in The Wild," CoRR, 12/03, 2012.
- [35] J. S. Pérez, E. Meinhardt-Llopis, and G. Facciolo, "TV-L1 optical flow estimation," *Image Processing On Line*, vol. 2013, pp. 137-150, 2013.
- [36] L. Wang, P. Koniusz, and D. Q. Huynh, "Hallucinating Bag-of-Words and Fisher Vector IDT terms for CNN-based Action Recognition," arXiv preprint arXiv:1906.05910, 2019.
- [37] Wang, J., Cherian, A., Porikli, F., & Gould, S. (2018). Video representation learning using discriminative pooling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1149-1158).
- [38] D. Roy, and C. K. Mohan, "Unsupervised universal attribute modelling for action recognition," *IEEE Transactions on Multimedia*, 2018.
- [39] Murtaza, F., HaroonYousaf, M., & Velastin, S. A. (2018, October). DA-VLAD: Discriminative Action Vector of Locally Aggregated Descriptors for Action Recognition. In 2018 25th IEEE International Conference on Image Processing (ICIP) (pp. 3993-3997). IEEE.
- [40] Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., & Paluri, M. (2018). A closer look at spatiotemporal convolutions for action recognition. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 6450-6459).
- [41] Zhu, Y., Lan, Z., Newsam, S., & Hauptmann, A. (2018, December). Hidden two-stream convolutional networks for action recognition. In Asian Conference on Computer Vision (pp. 363-378). Springer, Cham.
- [42] Sun, S., Kuang, Z., Sheng, L., Ouyang, W., & Zhang, W. (2018). Optical flow guided feature: A fast and robust motion representation for video action recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1390-1399).
- [43] Choutas, V., Weinzaepfel, P., Revaud, J., & Schmid, C. (2018). Potion: Pose motion representation for action recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7024-7033).
- [44] A. Diba, M. Fayyaz, V. Sharma, M. Paluri, J. Gall, R. Stiefelhagen, and L. Van Gool, "Holistic Large Scale Video Understanding," arXiv preprint arXiv:1904.11451, 2019.
- [45] K. Yang, J. Fu, X. Guo, Y. Lu, P. Qiao, D. Li, and Y. Dou, "IF-TTN: Information Fused Temporal Transformation Network for Video Action Recognition," arXiv preprint arXiv:1902.09928, 2019.
- [46] Z. Zheng, G. An, D. Wu, and Q. Ruan, "Spatial-temporal pyramid based Convolutional Neural Network for action recognition," *Neurocomputing*, vol. 358, pp. 446-455, 2019.
- [10] L. Liu, L. Wang, and X. Liu, "In defense of soft-assignment coding," in Proceedings of the 2011 International Conference on Computer Vision, 2011, pp. 2486-2493.
- [11] F. Perronnin, J. S. S., nchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in Proceedings of the 11th European conference on Computer vision: Part IV, Heraklion, Crete, Greece, 2010, pp. 143-156.
- [12] X. Peng, C. Zou, Y. Qiao, and Q. Peng, "Action recognition with stacked fisher vectors." pp. 581-595.
- [13] B. Fernando, E. Gavves, J. Oramas, A. Ghodrati, and T. Tuytelaars, "Rank pooling for action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 4, pp. 773-787, 2016.
- [14] B. Fernando, P. Anderson, M. Hutter, and S. Gould, "Discriminative hierarchical rank pooling for activity recognition." pp. 1924-1932.
- [15] K. Simonyan, and A. Zisserman, "Two-stream convolutional networks for action recognition in videos." pp. 568-576.
- [16] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). IEEE.
- [17] L. Wang, Y. Xiong, Z. Wang, Y. Qiao, D. Lin, X. Tang, and L. Van Gool, "Temporal segment networks for action recognition in videos," *IEEE transactions on pattern analysis and machine intelligence*, 2018.
- [18] Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., & Van Gool, L. (2016, October). Temporal segment networks: Towards good practices for deep action recognition. In European conference on computer vision (pp. 20-36). Springer, Cham.
- [19] Diba, A., Sharma, V., & Van Gool, L. (2017). Deep temporal linear encoding networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 2329-2338).
- [20] Zhu, J., Zhu, Z., & Zou, W. (2018, August). End-to-end video-level representation learning for action recognition. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 645-650). IEEE.
- [21] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 221-231, 2012.
- [22] Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks. In Proceedings of the IEEE international conference on computer vision (pp. 4489-4497).
- [23] Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset. In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6299-6308).
- [24] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
- [25] S. Hochreiter, and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [26] Du, Y., Wang, W., & Wang, L. (2015). Hierarchical recurrent neural network for skeleton based action recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1110-1118).
- [27] Yao, L., Torabi, A., Cho, K., Ballas, N., Pal, C., Larochelle, H., & Courville, A. (2015). Describing videos by exploiting temporal structure. In Proceedings of the IEEE international conference on computer vision (pp. 4507-4515).
- [28] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention." pp. 2048-2057.
- [29] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator." pp. 3156-3164.
- [30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," *Advances in Neural Information Processing Systems*, vol. 4, 09/10, 2014.
- [31] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional LSTM." pp. 273-278.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه تبریز

۲۷ الی ۲۸ آذرماه ۱۳۹۸

تحلیل ارزش و بخش بندی مشتریان با استفاده از تکنیک های داده کاوی و RFM (مطالعه موردی شرکت پتروشیمی بجنورد)

نفیسه روحانی^۱، احسان جعفری^۲، شادی لنگری^۳

^۱ دانشجوی کارشناسی ارشد مهندسی نرم افزار، موسسه غیرانتفاعی اشراق، بجنورد
n.rohany61@gmail.com

^۲ عضو هیئت علمی گروه کامپیوتر، موسسه غیرانتفاعی اشراق، بجنورد
ehsannjafari@gmail.com

^۳ عضو هیئت علمی گروه کامپیوتر، موسسه غیرانتفاعی اشراق، بجنورد
shadilangari@gmail.com

چکیده

بخش بندی بهینه و مؤثر مشتریان بر مبنای ویژگی های مناسب و مرتبط، می تواند به نفوذ بیشتر کسب و کار در بازار کمک کند. روش های متعددی برای بخش بندی مشتریان بر اساس تحلیل رفتار آنها وجود دارد که RFM یکی از روش های مرسوم و پرکاربرد در این زمینه است. متغیرهای RFM پیش بینی کننده های خوبی برای رفتار آتی مشتریان هستند. هدف اصلی این پژوهش بررسی تکنیک های داده کاوی و RFM در تحلیل مؤثر مشتریان است. در ادامه با توجه به نوع داده و نیاز سازمان از الگوریتم های خوشه بندی و تکنیک RFM برای ارائه راه حل های استراتژیک استفاده شده است و در پایان با تحلیل نتایج بر پایه روش بوستون، اطمینان خوشه بندی نیز محاسبه گردیده است. با بررسی مقالات و تحقیقاتی در این زمینه، پیشنهاداتی مطابق با نتایج حاصل از این بررسی برای تحقیقات آتی ارائه شده است.

کلمات کلیدی

تکنیک RFM، مدیریت ارتباط با مشتریان، داده کاوی، برآورد فاصله ای میانگین.

۱- مقدمه

کند [۲]. بخش بندی مشتریان یکی از روش های داده کاوی است که هدف آن گروه بندی مشتریان با نیازها و رفتار خرید مشابه برای تعیین راهبردهای مختلف به منظور حداکثر نمودن پاسخ به برنامه های بازاریابی و کاهش هزینه های سازمان است؛ بخش بندی بهینه و مؤثر مشتریان بر مبنای ویژگی های مناسب و مرتبط، می تواند به نفوذ بیشتر کسب و کار در بازار کمک کند. روش های متعددی برای بخش بندی مشتریان بر اساس تحلیل رفتار آنها وجود دارد که RFM یکی از روش های مرسوم و پرکاربرد در این زمینه است.

متغیرهای RFM پیش بینی کننده های خوبی برای رفتار آتی مشتریان هستند. هدف اصلی این پژوهش بررسی تکنیک های داده کاوی

یکی از مهمترین مواردی که در سازمان های امروز به خوبی قابل رویت است، فراوانی و حجم بالای داده ها و اطلاعاتی است که در سازمان تولید می شود. این حجم وسیع از داده و اطلاعات در عین اینکه سازمان را با مسائلی در رابطه چگونگی نگهداری، جمع آوری و طبقه بندی داده ها روبه رو می سازد، دانشی را در درون خود نهفته دارند که یک سازمان یادگیرنده و هوشمند از آن به مانند فرصتی جهت ارتقای توانمندی های خود استفاده می کند [۱]. مدیریت ارتباط با مشتری بر پایه سازی، حفظ و افزایش روابط بلند مدت با مشتریان تمرکز می

شاخص وفاداری مشتری است را شناسایی نمودند و بررسی کردند که کدام شاخص‌ها در پیش بینی عملکرد کسب و کار ارزشمندتر هستند [۵].

۲-۲- RFM

یکی از مدل‌های مطرح شده در تحلیل ارزش مشتری، مدل است که برای اولین بار توسط هاگس (Hughes) در سال ۱۹۹۴ ارائه شده است و به بیان تفاوت میان مشتریان با استفاده از سه متغیر تازگی، تکرار و ارزش پولی می‌پردازد. پارامترهای این مدل به شرح زیر است:

تازگی آخرین خرید (R): نشان دهنده تازگی است که مدت زمان بین آخرین تعامل تجاری با حال را نشان می‌دهد، هرچه این مدت کمتر باشد، R بیشتر است.

تکرار خرید (F): نشان‌دهنده تکرار است و تعداد تراکنش‌ها را در یک بازه زمانی مشخص نشان می‌دهد. هرچه تعداد تکرار بیشتر باشد F بزرگتر است.

ارزش پولی خرید (M): ارزش پولی معاملات را در بازه ای خاص نشان می‌دهد. هرچه ارزش پولی بیشتر باشد M بزرگتر است [۶]. اکثر محققان بر این عقیده هستند که با توجه به صنایع مختلف می‌بایست برای دسته‌بندی مشتریان، وزن‌های مختلفی را به متغیرهای RFM اختصاص داد. به عنوان مثال برای تحلیل ارزش مشتریان بانک‌ها، بهتر است که بیشترین وزن به متغیر F، سپس R و کمترین وزن برای M تخصیص یابد [۷]. در پژوهش دیگر داده‌ها بر اساس وزن دهی هر یک از فاکتورهای R، F و M در چهار خوشه جای گرفتند که هر یک مشخصه یکی از انواع "مشتریان وفادار"، "مشتریان از دست رفته"، "مشتریان بالقوه"، و "مشتریان جدید" می‌باشد [۸].

۳-۳- پیشینه تجربی

برادران و بیگلری برای بخش بندی مشتریان صنایع تولید و پخش کالاهای پر گردش از مدل بهبود یافته RFM استفاده کردند. آنها با جایگزینی متغیر توالی خرید (C) (توالی خرید مشتری طی یک دوره خاص و برابر تعداد ماه‌هایی از سال است که مشتری طی این دوره خرید کرده است) با متغیر تأخر خرید در مدل RFM کیفیت بخش بندی را بهبود دادند. نتایج نشان داد بخش بندی مشتریان بر اساس مدل CFM در مقایسه با RFM دقت بیشتری دارد [۹].

کرامتی و خالقی از مدلی به نام RFM-PCB، برای بخش بندی مشتریان خرده فروشی برخط کالاهای دیجیتالی کامپیوتری بهره بردند. آنها در این مدل به استثنای متغیر R مقادیر متغیرهای FM را در سطح دسته کالاهای به صورت جداگانه محاسبه کردند، سپس با اجرای روشهای قوانین انجمنی در هر خوشه پیشنهادی متناسبی با ویژگیهای آن خوشه ارائه داده اند. نتایج نشان داد مدل پیشنهادی نسبت به مدل RFM ساده در پیش بینی رفتار آتی مشتریان دقت بیشتری دارد [۱۰].

آخوندزاده نوقایی و همکارانش در تحقیقی به شناسایی گروه‌های رفتاری مشتریان و ویژگی‌های هر یک در صنعت مخابرات پرداختند. بدین منظور آنان ابتدا مشتریان را بر اساس بخش بندی کردند، سپس با استفاده از قوانین انجمنی به K-means و روش RFM متغیرهای

RFM در تحلیل موثر مشتریان است. در ادامه با توجه به نوع داده و نیاز سازمان از الگوریتم‌های خوشه‌بندی و تکنیک RFM برای ارائه راه‌حل‌های استراتژیک استفاده شده است و در پایان با تحلیل نتایج بر پایه روش بوستون، اطمینان خوشه بندی نیز محاسبه گردیده است. با بررسی مقالات و تحقیقاتی در این زمینه، پیشنهادهای مطابق با نتایج حاصل از این بررسی برای تحقیقات آتی ارائه شده است.

۲- بیان مساله

یکی از مسائل مهمی که سازمان‌ها باید به سرعت در ارتباط با آن تصمیم‌گیری نمایند، مشتری می‌باشد. اما دیگر داشتن سیستم‌های مدیریت ارتباط با مشتری برای مدیریت موثر ارتباط با مشتری و پشتیبانی از تصمیم‌گیری با توجه به شرایط کسب و کار کنونی دیگر کفایت نمی‌کند.

در واقع مدیران بخش‌های فروش و بازاریابی نیاز دارند تا به شناخت دقیق تری از مشتریان دست یابند. در این شرایط می‌توان با توجه به تحلیل‌های به دست آمده از تکنیک‌های داده‌کاوی به اتخاذ تصمیمات مناسب استراتژیک جهت افزایش میزان فروش و هم چنین حفظ و وفادار کردن مشتریان اقدام نمود.

۳- پیشینه پژوهش

۳-۱- بخش بندی و تحلیل رفتار مشتری

با استفاده از داده‌های دربارۀ میلیون‌ها تراکنش مشتری به همراه اطلاعات جمعیت شناختی، جغرافیایی، و روان شناختی، سازمان‌ها می‌توانند مشتریان را براساس رفتار خرید بخش بندی نمایند. ابزارهای هوش تجاری جایگزین بخش بندی‌های سنتی نیستند. آن‌ها ابزارهای جدید قدرتمندی را که با ابزارهای موجود به خوبی کار کنند را فراهم می‌کنند تا به سازمان‌ها در ایجاد بخش‌های باریک از مشتریان، فهمیدن نیازها و ارزش‌های هر یک از بخش‌ها، ایجاد محصولات و خدماتی که نیاز آن‌ها را بهتر پاسخ دهد، کمک می‌کند [۳].

تجزیه و تحلیل رفتارها برای شناسایی انواع خاصی از رفتارها برای پیش بینی گرایش‌ها یا سرمایه‌ها استفاده می‌گردد. به طور کلی، هر الگوی رفتاری که رویدادهای مهمی از کسب و کار را نشان دهد، ارزشمند است و دنبال می‌گردد. این نوع از فرآیندهای تحلیلی، از داده‌های گذشته برای جستجوی الگوهای رفتاری که قبل از رویدادهای مهم به وقوع می‌پیوندد، استفاده می‌نمایند و سپس سعی در شناسایی این الگوهای رفتاری دارند. این تجزیه و تحلیل برای تحلیل گرایش‌های خرید، فرسایش مشتریان، فعالیت‌های تحت وب، تجزیه و تحلیل شبکه‌های اجتماعی و تجزیه و تحلیل تمایلات امکان پذیر است [۴]. مورگان و رگو (۲۰۰۶) با اشاره به اهمیت داده‌های بازخورد مشتری برای تعیین اهداف و نظارت بر عملکرد براساس معیارها (متریک‌ها)، ۶ معیار بازخورد مشتری را که شامل ۳ شاخص رضایت مشتری و ۳

^۱ Customer attrition

^۲ Sentiment analysis

^۳ Morgan & Rego

شرکت های خارجی می باشند. مشتریان خارجی این شرکت شامل کشورهای عراق، افغانستان، ترکیه، ایتالیا و آذربایجان می باشند. محصولات قابل فروش شامل کریستال ملامین، پودر اوره، ازت، گاز آمونیاک و آب صنعتی می باشد. داده‌های واقعی برای کار داده‌کاوی نیاز به پیش پردازش و آماده سازی دارد تا نتایج خروجی دقیق تر باشد، که در فاز بعدی به این کار خواهیم پرداخت.

فاز دوم: آماده سازی و پیش پردازش داده ها

این مرحله از متد CRISP، شامل انتخاب داده‌ها، پاکسازی داده‌ها، آماده کردن داده جهت داده‌کاوی است. پاکسازی داده در واقع مرحله کنترل کیفی قبل از تحلیل داده است و یکی از وظایف آن پرکردن یا حذف داده‌های مفقوده است. داده‌های مفقوده را در صورتی می‌توان حذف نمود که تعداد ویژگی‌های فاقد مقدار در یک نمونه زیاد نباشد و همچنین مقادیر مفقوده را نتوان به صورت دستی پرکرد، در داده‌های موجود نیز با توجه به اینکه مقادیر مفقوده از فیلدهای مهمی به شمار می‌آید که پر کردن آن به صورت دستی امکان‌پذیر نبود لذا تعدادی از آن حذف شد و تعداد داده‌های مورد بررسی به 12385 رکورد کاهش یافت. هم چنین با در نظر گرفتن هدف پژوهش، تعداد ۷ فیلد که شامل مشخصات توزین بوده به دلیل کم اهمیت بودن در نظر گرفته نشد. در این مرحله داده‌ها آماده مدل‌سازی می‌باشند. یکی دیگر از روش‌های رایج برای تشخیص داده‌های پرت یا نویز و همچنین توصیف داده، روش خوشه بندی است [۱]، بنابراین در فاز بعدی به خوشه‌بندی داده‌ها با روش کای میانگین می‌پردازیم.

فاز سوم: خوشه بندی داده ها با روش کای میانگین

مرحله اول: خوشه بندی داده‌ها با الگوریتم کای میانگین

در این مرحله به منظور یافتن داده‌های پرت و همچنین توصیف داده‌ها، از الگوریتم کای میانگین (K-means) برای خوشه‌بندی داده‌ها استفاده می‌شود. الگوریتم کای میانگین یکی از روش‌های خوشه‌بندی داده‌ها می‌باشد که به دلیل سرعت و سادگی این الگوریتم به طور گسترده مورد استفاده قرار می‌گیرد. آخوندزاده نوقابی روش‌های مختلفی به منظور خوشه بندی استفاده می‌شود اما دلیل استفاده از این روش در این پژوهش اینست که تعداد خوشه‌های بهینه با استفاده از شاخص ریاضی قابل محاسبه است و این مساله اعتبار خروجی خوشه‌بندی را در پی دارد [۱۶]. الگوریتم k-means، پارامتر k را به عنوان ورودی گرفته و مجموعه n شیء را به k خوشه افراز می‌کند به طوریکه سطح شباهت داخلی خوشه‌ها بالا بوده و سطح شباهت اشیاء بیرون خوشه‌ها پائین باشد. شباهت هر خوشه نسبت به متوسط اشیاء آن خوشه سنجیده شده که متوسط مرکز خوشه نیز نامیده می‌شود. در این روش فاصله‌ها بر اساس فاصله اقلیدسی تعیین می‌شود [۱۷]. به منظور خوشه‌بندی از نرم‌افزار Spss Clementine 12.0 استفاده شده است. با توجه به اینکه تعداد خوشه بهینه با روش کای میانگین بر اساس شاخص محاسبه می‌شود، لذا در مدل‌سازی تعداد ۲ الی ۸ خوشه را با نرم افزار مدل‌سازی کرده و پس از آن، با استفاده از شاخص دیویس- بولدین به محاسبه خوشه بهینه می‌پردازیم.

شناسایی الگوهای رفتاری مشتریان در هر بخش پرداختن بر اساس نتایج، هفت گروه رفتاری مختلف از مشتریان شناسایی شد که این امر به دید و شناخت بهتر نسبت به الگوهای رفتاری مشتریان و بهبود راهبردهای بازاریابی انجامید [۱۱].

سوئینی و فتح‌علیزاده در پژوهشی برای بخش بندی مشتریان صنعت بیمه، مدل RFM را به LRFMC (طول ارتباط با مشتری، C میانگین خسارت وارده توسط مشتری) توسعه دادند که نتایج بخش بندی حاکی از بهبود کیفیت خوشه‌ها بود [۱۲].

وی، لین، ونگ و وو پژوهشی با عنوان کاربرد مدل LRFM در بخش بندی بازار کلینیک دندانپزشکی انجام دادند و در آن بر اساس مدل LRFM به شناسایی مشتریان وفادار کلینیک پرداختند. طبق نتایج، بیماران به چهار دسته مشتریان وفادار، فعال، جدید و نامعلوم بخش بندی شدند و راهبردهای مناسب هر بخش تعیین شد [۱۳]. سایر نتایج: استفاده از خوشه بندی های فازی موجب می‌شود تا یک داده تنها محدود به یک خوشه نباشد. به عبارت دیگر بتواند عضو دو یا بیشتر خوشه باشد که موجب تحلیل های متفاوت میشود [۱۴]. استفاده از سری زمانی برای بخش بندی با تحلیل های جدیدی از RFM صورت گرفته به طوری که تغییر مروری رفتار کاربر را در نظر گرفته و این تحلیل ها روی داده های حجمی نتایج جالب تر و تحلیل های موثرتری به همراه داشته است [۱۵]. با مرور پژوهش‌های پیشین در می‌یابیم که تکنیک RFM روشی قدرتمند برای شناخت ویژگی‌های حیاتی هر سازمان می‌باشد. ترکیب این تکنیک با سایر روش‌ها از جمله داده‌کاوی می‌تواند نتایج جالبی را برای سازمان به همراه داشته باشد. بر این اساس در قسمت بعد از ترکیب الگوریتم‌های داده‌کاوی و تکنیک RFM استفاده خواهیم نمود.

۴- یافته های پژوهش

فاز اول: شناخت کسب و کار و آشنایی با نوع داده ها

مورد مطالعاتی این پژوهش شرکت پتروشیمی بجنورد می باشد که تحقیق روی بخشی از داده های فروش انجام شده است. تعداد رکورد مورد بررسی حدود 12389 رکورد داده (۱۲,۳ مگابایت) است. این داده ها شامل داده های ۸ ماه کاری فروش شرکت پتروشیمی از فروردین تا آبان ماه سال ۱۳۹۵ است. فیلد های اطلاعاتی شامل مشخصات بارگیری محصولات، نوع محصولات، اطلاعات شرکت حمل و نقل و مشتریان می باشد که در شکل 1 مشاهده می شود.

ردیف	کد مشتری	نام مشتری	آدرس مشتری	تاریخ	مقدار	نوع	وضعیت
1	10001	شرکت پتروشیمی	شهر بجنورد	1395/01/01	1000	گاز	موجود
2	10002	شرکت پتروشیمی	شهر بجنورد	1395/01/05	2000	گاز	موجود
3	10003	شرکت پتروشیمی	شهر بجنورد	1395/01/10	3000	گاز	موجود
4	10004	شرکت پتروشیمی	شهر بجنورد	1395/01/15	4000	گاز	موجود
5	10005	شرکت پتروشیمی	شهر بجنورد	1395/01/20	5000	گاز	موجود

شکل ۱: داده‌های اولیه مربوط به واحد فروش در پتروشیمی

بررسی اجمالی داده ها نشان می‌دهد که برخی مشتریان این شرکت مشتری مستقیم و برخی مشتریان واسطه برای فروش محصولات به

محصولاتی است که برای کشور ایران تولید و به فروش می‌رسد دچار سوگیری شده است. خوشه بندی به ما نشان می‌دهد که پیاده سازی تکنیک RFM روی این توزیع داده نمی‌تواند نتیجه خوبی به همراه داشته باشد زیرا متغیر فرکانس (f) تحت تاثیر داده‌های فروش ایران می‌شود چون قطعا میزان فرکانس یعنی دفعات خرید در ایران بیش از خارج کشور است بدین ترتیب برای بررسی بیشتر در مرحله دوم فرآیند خوشه بندی را به نحوی انجام می‌دهیم که داده‌ها را به دو بخش مجزا تقسیم کرده، یعنی یک بخش شامل داده‌های فروش به کشور ایران و دسته دوم داده‌های فروش به سایر کشورها. با این کار می‌توانیم تحلیل RFM را روی داده‌های فروش خارج از کشور برای مقایسه فروش بین کشورهای خریدار انجام دهیم.

مرحله دوم: محاسبه فاکتورهای RFM

داده‌های فروش به کشورهای خارجی شامل ۲۴۱۱ رکورد شامل پنج کشور افغانستان، عراق، ایتالیا، ترکیه و آذربایجان می‌باشد. در این مرحله به منظور آماده سازی داده‌ها برای مدل‌سازی با روش RFM، باید سه فاکتور فرکانس یا تکرار خرید، تازگی خرید و ارزش پولی هر خرید را محاسبه کنیم. در پژوهش دیگری داده‌ها بر اساس وزن دهی هر یک از فاکتورهای R، F و M در چهار خوشه جای گرفتند که هر یک مشخصه یکی از انواع "مشتریان وفادار"، "مشتریان از دست رفته"، "مشتریان بالقوه"، و "مشتریان جدید" می‌باشد [۸]. نحوه وزن دهی در این پژوهش به صورت جدول ۲ در زیر می‌باشد.

جدول ۲: محاسبه فاکتورهای RFM

نوع مشتریان	نحوه وزن دهی به فاکتورها
مشتریان وفادار	$R \uparrow F \uparrow M \uparrow$
مشتریان از دست رفته	$R \downarrow F \downarrow M \downarrow$
مشتریان جدید	$R \uparrow F \downarrow M \downarrow$
مشتریان بالقوه	$R \uparrow F \uparrow M \downarrow$

با توجه به این‌که در این پژوهش قصد داریم روی مشتریان سودآور تمرکز کنیم بنابراین بیشترین وزن را به ترتیب به فاکتور M، F و R اختصاص می‌دهیم.

مرحله سوم: خوشه بندی کشورهای خریدار بر مبنای سه فاکتور RFM پس از وزن دهی و محاسبه فاکتورهای RFM، به خوشه بندی‌های داده‌های مشتریان بر مبنای این سه فاکتور می‌پردازیم. در قسمت قبل ادعا کردیم که استفاده از شاخص ریاضی دیویس- بولدین اعتبار خروجی خوشه‌بندی را در پی دارد، بار دیگر از روش خوشه بندی کای میانگین (K-means) برای خوشه بندی استفاده کرده و سپس با شاخص دیویس بولدین تعداد خوشه بهینه محاسبه می‌گردد. نتایج محاسبات، سه خوشه را به عنوان خوشه بهینه انتخاب می‌کند. نتایج خوشه بندی در جدول ۳ نمایش داده شده است. فاز چهارم: تحلیل خوشه‌ها

همانطور که گفته شد برای محاسبه خوشه بهینه از شاخص دیویس بولدین استفاده می‌شود. که شکل این محاسبات به صورت زیر است (دیویس و بولدین، ۱۹۷۹):

شاخص دیویس بولدین از معیار شباهت بین دو خوشه (Rij) استفاده میکند که بر اساس پراکندگی یک خوشه (Si) و عدم شباهت بین دو خوشه (dij) تعریف میشود. شباهت بین دو خوشه را میتوان به صورت‌های مختلفی تعریف کرد ولی بایستی شرایط زیر را دارا باشد:

$$R_{ij} \geq 0 \quad (1-4)$$

$$R_{ij} = R_{ji} \quad (2-4)$$

اگر S_i و S_j هر دو برابر صفر باشند آنگاه R_{ij} نیز برابر صفر باشد.

$$R_{ik} > R_{kj} \text{ آنگاه } d_{ik} = d_{kj} \text{ و } SK < S_j \quad (3-4)$$

$$R_{ij} > R_{ik} \text{ آنگاه } d_{ik} > d_{ij} \text{ و } SK = S_j \quad (4-4)$$

معمولا شباهت بین دو خوشه به صورت زیر تعریف میشود:

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \quad (5-4)$$

که در آن d_{ij} و s_i بر اساس فرمولهای زیر محاسبه میشوند:

$$d_{ij} = d(v_i, v_j) \quad (6-4)$$

$$s_i = \frac{1}{|c_i|} \sum_{x \in c_i} d(x, v_i) \quad (7-4)$$

به این ترتیب شاخص دیویس بولدین به صورت زیر تعریف میشود:

$$DB = \frac{1}{nc} \sum_{i=1}^{nc} R_i \quad (8-4)$$

که در آن R_i به صورت زیر محاسبه میشود:

$$i=1 \dots nc \quad (R_{ij}), R_i = \max \quad (9-4)$$

این شاخص در واقع میانگین شباهت بین هر خوشه با شبیه‌ترین خوشه به آن را محاسبه میکند. میتوان دریافت که هرچه مقدار این شاخص کمتر باشد، خوشه‌های بهتری تولید شده است [۱۸].

پس از محاسبات انجام شده، شاخص دیویس- بولدین تعداد ۴ خوشه را به عنوان تعداد خوشه بهینه تعیین کرد. حال می‌توانیم به تحلیل خوشه‌ها بپردازیم. جدول ۱ مشخصات خوشه‌ها را نشان می‌دهد.

جدول ۱: مشخصات خوشه‌های بهینه

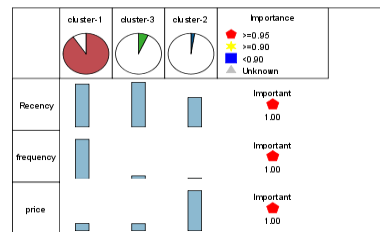
خوشه‌ها	تعداد رکوردها	ماه‌های فروش	نوع محصول	کشور خریدار
خوشه ۱	۵۹۷۶	مرداد، شهریور، مهر (۶۷,۵٪)	اوره حمایتی (۹۵,۷۵٪)	ایران (۱۰۰٪)
خوشه ۲	۲۳۴۸	مرداد ماه بیشترین فروش (۲۲,۵۳٪)	اوره صادراتی (۹۹,۹۶٪)	افغانستان (۹۲,۹۷٪)
خوشه ۳	۳۱۰۷	مهرماه بیشترین فروش (۲۳,۷۵٪)	آمونیاک مایع (۲۸,۸۷٪)	ایران (۹۷,۹۷٪)
خوشه ۴	۹۵۴	تیر ماه (۱۰۰٪)	اوره حمایتی (۹۶,۴۴٪)	ایران (۱۰۰٪)

یکی از مهمترین مزایای تکنیک خوشه بندی، توصیف داده‌ها می‌باشد. همانطور که در جدول بالا مشاهده می‌شود خوشه بندی توانسته داده‌ها را بر اساس شباهت ویژگی‌ها دسته بندی کند که در ۴ دسته مجزا تقسیم بندی شده است. نتایج مرحله اول خوشه بندی نشان می‌دهد توزیع داده‌ها در خوشه‌ها به دلیل اینکه حجم بیشتر داده‌ها شامل

جدول ۳: مشخصات خوشه‌های بهینه با الگوریتم کای میانگین

خوشه ها	تعداد رکوردها	کشورهای خریدار	میانگین فاکتور F	میانگین فاکتور R	میانگین فاکتور M
خوشه ۱	۲۱۷۹	افغانستان	۲۱۸۳	۱۳۲,۵۳۹	۶۳۴۳۷۰ ۱,۲۳۹
خوشه ۲	۶۳	ترکیه، آذربایجان، ایتالیا	۳۵,۳۸۱	۹۱,۰	۳۵۴۵۰۲ ۶۹,۸۴۱
خوشه ۳	۱۶۵	عراق	۱۶۵,۰	۱۳۸,۰	۶۳۰,۸۶۶ ۶,۶۶۷

با توجه به این‌که خوشه‌بندی یکی از مهمترین روش‌های توصیفی می‌باشد بنابراین ابتدا به مقایسه و توصیف خوشه‌ها می‌پردازیم. شکل ۲ به صورت نمودار به توصیف هر یک از فاکتورهای RFM و وضعیت آن‌ها در خوشه‌ها می‌پردازد.



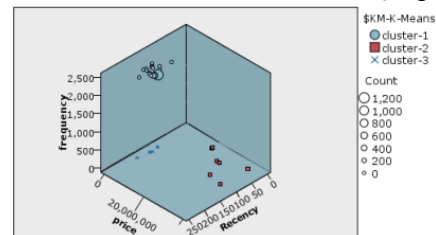
شکل ۲: توصیف نموداری خوشه‌های بهینه با الگوریتم کای میانگین

با مقایسه شکل ۲ و جدول ۲ می‌توان خوشه‌ها را به صورت جدول زیر توصیف نمود. خوشه دوم، شکل جدید از مشتریان را ایجاد کرده که آن‌ها را تحت عنوان مشتریان جدید پرسود در جدول ۳ نامگذاری می‌کنیم.

جدول ۴: نام‌گذاری خوشه‌های بهینه بر مبنای وضعیت سه فاکتور RFM

خوشه ها	نام‌گذاری مشتریان خوشه‌ها	M	F	R
خوشه ۱	مشتریان بالقوه	↓	↑	↑
خوشه ۲	مشتریان جدید پرسود	↑	↓	↑
خوشه ۳	مشتریان جدید کم سود	↓	↓	↑

شکل ۳ به طور واضح وضعیت خوشه ای مشتریان بر مبنای سه فاکتور RFM را نمایش می‌دهد که این شکل منظم خوشه‌ای بیانگر صحت محاسبه تعیین خوشه بهینه است.



شکل ۳: وضعیت خوشه‌های بهینه با روش خوشه بندی کای میانگین (k-means)

تحلیل خوشه اول

مشتریان خوشه اول تحت عنوان مشتریان بالقوه نامگذاری شده‌اند. این گروه از مشتریان با توجه به اینکه از ما زیاد خرید می‌کنند (فاکتور F زیاد) و همچنین ارتباط خود را با ما ادامه داده و به سمت رقیب نرفتند (فاکتور R زیاد) در دسته مشتریان خوب قرار می‌گیرند که نیاز به توجه ویژه دارند.

شکل ۴: درصد محصولات خریداری شده توسط کشور افغانستان

شکل ۴ محصولات خرید شده توسط مشتریان این خوشه یعنی کشور افغانستان را نمایش می‌دهد که بیشترین درصد خرید (۹۹,۹۵٪) مربوط به محصول اوره صادراتی می‌باشد.

تحلیل خوشه دوم

مشتریان خوشه دوم بر مبنای وضعیت سه فاکتور RFM تحت عنوان مشتریان جدید پرسود نامگذاری شده‌اند. با توجه به اینکه این گروه از مشتریان اخیراً از ما خرید داشته‌اند (فاکتور R بالا) و در گذشته زیاد خرید نکرده‌اند (فاکتور F پایین) بنابراین جدید هستند. و از سوی دیگر با وجود خرید کم (فاکتور M پایین) سودآوری خوبی برای شرکت داشته‌اند، بنابراین در دسته مشتریان پرسود قرار می‌گیرند.

شکل ۵: درصد محصولات خریداری شده توسط مشتریان خوشه دوم

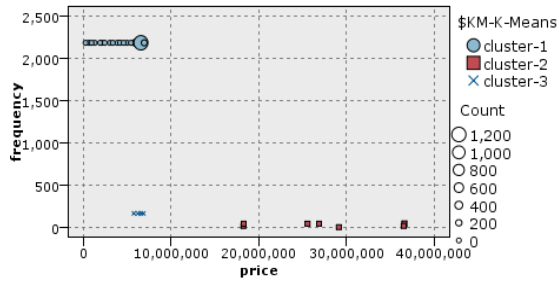
محصولات خریداری شده توسط مشتریان خوشه دوم

شکل ۵ محصولات خرید شده توسط مشتریان کشورهای ترکیه، ایتالیا و آذربایجان را نمایش می‌دهد که بیشترین درصد خرید مربوط به محصول کریستال ملامین BB می‌باشد.

تحلیل خوشه سوم

مشتریان خوشه سوم (کشور عراق) بر مبنای وضعیت سه فاکتور RFM تحت عنوان مشتریان جدید کم سود نامگذاری شده‌اند. با توجه به اینکه این گروه از مشتریان اخیراً از ما خرید داشته‌اند (فاکتور R بالا) و در گذشته زیاد خرید نکرده‌اند (فاکتور F پایین) بنابراین جدید هستند. و از سوی دیگر میزان سودآوری این گروه نسبت به گروه مشتریان خوشه دوم کمتر است، بنابراین در دسته مشتریان کم سود قرار می‌گیرند.

شکل ۶: درصد محصولات خریداری شده توسط کشور عراق

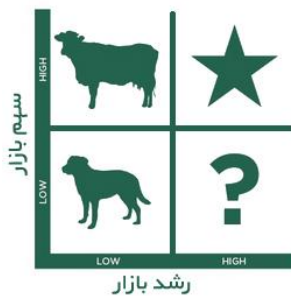


شکل ۶ محصولات خرید شده توسط مشتریان عراق را نمایش می‌دهد که بیشترین درصد خرید (۱۰۰٪) مربوط به محصول اوره صادراتی می‌باشد.

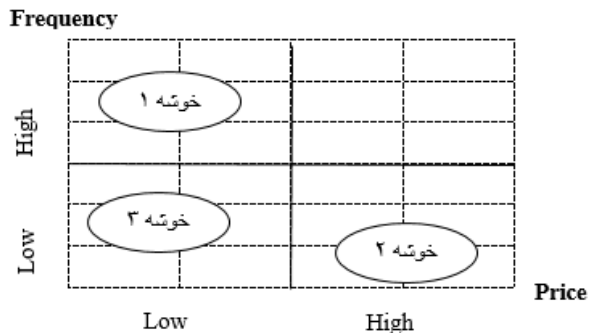
مقایسه نتایج

بر مبنای نتایج به دست آمده از فرآیند خوشه‌بندی، مشتریان در سه دسته مجزا خوشه‌بندی و یک به یک توصیف شده‌اند. در شرایط بازار، به منظور اتخاذ تصمیمات استراتژیک نیازمند مقایسه و پیش بینی مشتریان از نظر رفتار خرید و تحلیل رفتار آن‌ها می‌باشیم. در شکل‌های ۱۴ و ۱۵ مشتریان هر یک از خوشه‌ها را از نظر تکرار خرید (frequency) و ارزش پولی خرید (monetry) یا (price) مقایسه می‌کنیم. همانطور که در تصویر مشاهده می‌شود بالاترین حجم خرید به ترتیب متعلق به خوشه‌های یک، سه و دو می‌باشد، از طرفی بالاترین ارزش پولی خرید مشتریان به ترتیب به خوشه‌های دو (ترکیه، ایتالیا، آذربایجان)، یک (افغانستان) و سه (عراق) تعلق دارد.

شکل ۸: وضعیت خرید مشتریان بر مبنای دو بعد ارزش پولی خرید (price) و تعداد خریدها (frequency)

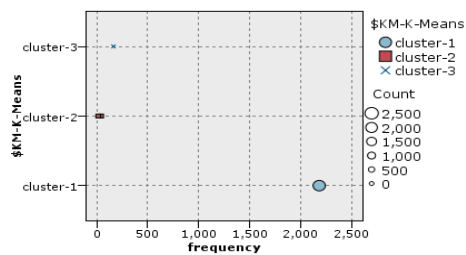
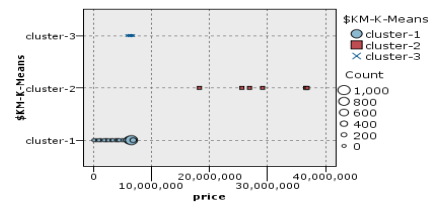


شکل ۹: ماتریس بوستون (BCG)، ابزاری برای برنامه ریزی استراتژیک فعالیت‌ها.



شکل ۱۰: وضعیت خوشه‌ها بر مبنای تعداد خرید و سهم خرید

وضعیت خوشه ۱ (کشور افغانستان) نشانه ویژگی مشتریان وفادار هستند که بیشتر با توجه به نتایج قبلی تحت عنوان مشتریان بالقوه نامگذاری شدند. مشتریان بالقوه مشتریانی هستند که فاصله چندانی تا مشتریان ایده‌آل ندارند، از سوی دیگر با توجه به این‌که بیشترین تکرار خرید را داشته‌اند بنابراین کشور افغانستان مشتری وفادار و مستعدی است که به مشتری ایده‌آل تبدیل شود. اگر نمودار رشد- سهم (نمودار بوستون) را بتوانیم مبنای قرار دهیم، پیشنهاد میشود شرکت بیشترین توجه را به این گروه از مشتریان معطوف نماید و تلاش کند تا وضعیت این مشتری از جایگاه ربع دوم به ربع اول یعنی مشتریان ایده‌آل تغییر کند، بنابراین شرکت باید به دنبال استراتژی‌هایی باشد که سود بیشتری از معاملات با این گروه مشتریان کسب کند، مثلاً با اقدام به مبادلات پایاپای، افزایش حجم خرید، ارائه مزایای ویژه، حمل و نقل رایگان و غیره....



شکل ۷: مقایسه خوشه‌ها را از نظر تکرار خرید (frequency)

به منظور ارائه تصمیمات استراتژیک نیاز به تحلیل‌های بهتری از رفتار مشتریان داریم. شکل ۸ وضعیت خرید مشتریان را بر مبنای دو بعد ارزش پولی خرید (price) و تعداد خریدها نمایش می‌دهد. بر مبنای شکل ۸ و ماتریس BCG در شکل ۹، نمودار شکل ۱۰ را طراحی و تحلیل می‌کنیم. ماتریس BCG یا ماتریس بوستون ابزاری برای بررسی فعالیت‌های لازم برای برنامه‌ریزی استراتژیک در شرایط رقابتی بازار است که در شکل ۹ نمایش داده شده است [۱۹].

با توجه به اینکه بیشترین تغییر در فاکتور F رخ می‌دهد لذا شرایط سه خوشه را بر مبنای این فاکتور مورد ارزیابی قرار می‌دهیم. برای این منظور ابتدا فاکتور فرکانس را بر مبنای زمان به صورت بازه‌ای تبدیل می‌کنیم، سپس در نرم افزار آماری Minitab 16، با درجه اطمینان ۹۵٪ به محاسبه بازه اطمینان می‌پردازیم. نتایج به صورت جدول زیر مشاهده می‌شود.

جدول ۵: نمایش بازه‌ای اطلاعات هر خوشه به شکل گزارش شده اولیه و با اعمال درجه اطمینان ۹۵٪

	خوشه ۱	خوشه ۲	خوشه ۳
تبدیل کردن فاکتور فرکانس به صورت بازه‌ای	(۲۲, ۵۱۰)	(۰, ۱۸)	(۰, ۵۹)
محاسبه بازه اطمینان با درجه اطمینان ۹۵٪	(۴۱۷, ۹)	(۴, ۸۱)	(۳۸, ۵۸)
	۱۲۲, ۶	(۰, ۴۴)	(۲, ۹۲)

قبل از تبدیل کردن بازه‌ای داده‌ها، جایگاه خوشه‌ها به ترتیب بر مبنای بزرگی فرکانس عبارتند از خوشه اول < خوشه سوم < خوشه دوم. حال قصد داریم تا این نتیجه را در شرایط اطمینان بار دیگر محاسبه کنیم. در سال ۲۰۰۹ دکتر جهان‌شاهلو در مقاله خود به منظور رتبه بندی واحدهای مجزا که ورودی‌های داده به صورت فاصله‌ای هستند از روش سنگوپتا استفاده کرده که به صورت زیر می‌باشد: [۲۱].

تعریف: به ازای مقادیر کران بالا و کران پایین هر واحد، مقدار E را به این صورت نمایش می‌دهیم:

$$E = \langle m(E), w(E) \rangle$$

به طوریکه $m(E)$ مقدار میانگین و $w(E)$ مقدار پهنای بازه هستند که به این صورت رابطه (۴-۱۱) محاسبه می‌شوند:

$$m(E) = \frac{1}{2}(el + eu) \quad (4-11)$$

$$w(E) = \frac{1}{2}(eu - el)$$

پس از این تعریف سنگوپتا و پال تابعی به منظور مقایسه دو مقدار فاصله‌ای واحدهای D و E معرفی کردند:

$$A(<) = \frac{m(D) - m(E)}{w(D) + w(E)} \quad (4-12)$$

$A(<)$ را می‌توان به عنوان «درجه مقبولیت اینکه بازه اول نسبت به بازه دوم پایین تر باشد» تفسیر کرد. درجه مقبولیت $A < B$ را می‌توان بر اساس موقعیت نسبی میانگین و پهنای بازه B نسبت به میانگین و پهنای بازه A به صورت زیر تقسیم بندی کرد:

وضعیت خوشه دوم (کشور ایتالیا، ترکیه، آذربایجان) نماد مشتریان پرسودی است که بیشتر با توجه به نتایج گذشته، تحت عنوان مشتریان جدید پر سود نامگذاری شده بود. با توجه به اینکه فاکتور R زیاد است بنابراین این گروه از مشتریان جدید هستند، برای تبدیل کردن این گروه مشتریان به مشتریان ایده‌آل (حرکت از ربع چهارم به ربع اول) نیاز است تا تعداد خرید (فاکتور F) افزایش یابد. با توجه به اینکه این گروه مشتریان بسیار سودآور هستند، به سازمان پیشنهاد می‌گردد تا با تسهیل کردن فرآیند معاملات و مدیریت بهتر ارتباط با مشتری به سودآوری دست یابد. وضعیت خوشه سوم نماد مشتریان کم سود برای سازمان است. بر مبنای نمودار رشد- سهم معمولاً برای مشتریانی که در این گروه (ربع سوم) قرار می‌گیرند پیشنهاد می‌شود که وقت و انرژی زیادی قائل نشوند حتی گاهی برخی سازمان‌ها به دلیل هزینه‌های بالای مدیریت ارتباط با مشتری این گروه از مشتریان را حذف می‌کنند اما با توجه به اینکه شرایط در همه سازمان‌ها و معاملات یکسان نیست و همچنین با توجه اینکه تعداد دفعات خرید این گروه نسبت به گروه دوم بیشتر هم هست، پیشنهاد می‌گردد بر مبنای شرایط و امکانات سازمان تصمیمات بینابینی در خصوص این گروه اتخاذ نمایند.

ارزیابی نتایج با اطمینان بالاتر

نوسان پذیری و عدم قطعیت جز جدایی ناپذیر دنیای واقعی است. این تغییرات در میزان ورودی و خروجی واحدها باعث می‌شود تا نتوانیم اعتماد کافی به نتایج حاصل از محاسبه را داشته باشیم. بنابراین در بیشتر موارد با داده‌های نا دقیق و نا مشخص مواجه هستیم. برای غلبه بر این مشکل و نیز افزایش میزان اطمینان به نتایج حاصل از نتایج، در این پایان نامه پیشنهاد می‌شود تا داده‌ها را به صورت فاصله‌ای در نظر گرفته و سپس به ارزیابی نتایج پردازیم.

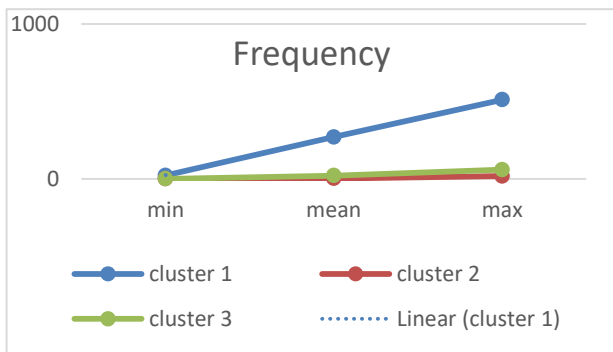
استفاده از برآورد فاصله‌ای میانگین به منظور افزایش اطمینان در محاسبه نتایج می‌باشد، علاوه بر این استفاده از برآورد فاصله‌ای میانگین، دارای مزایای زیر است:

۱. ضریب اطمینان تحت کنترل مدیر یا تحلیلگر است، ۲. پراکندگی داده‌ها را در نظر گرفته و سپس فاصله را ایجاد می‌کند، ۳. با افزایش اندازه نمونه، بازه‌ها کوتاهتر و قابلیت اطمینان آن بالاتر می‌رود.

به منظور کاهش خطای برآورد، یک برآورد فاصله‌ای با اطمینان ۹۵٪ محاسبه می‌شود. اگر توزیع جامعه مجهول باشد ولی اندازه نمونه بزرگ باشد ($n \geq 30$) با استفاده از رابطه (۴-۱۰) می‌توان یک برآورد فاصله‌ای با قطعیت ۱۰۰٪ $(1-\alpha)$ برای پارامتر μ انجام داد. S مقدار انحراف معیار نمونه‌ای، α مقدار سطح معناداری را نشان می‌دهد. [۲۰].

$$\bar{X} - z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \quad (4-10)$$

در این پژوهش قصد داریم تا با استفاده از برآورد فاصله‌ای میانگین از درستی داده‌ها و نتایج، اطمینان حاصل کنیم. سه فاکتور مهمی که برای خوشه‌بندی از آنها استفاده کردیم شامل فاکتورهای R، F و M می‌باشد.



شکل ۱۲: وضعیت خوشه‌ها در شرایط معمول

۵- نتیجه

ترکیب ماتریس بوستون با تکنیک RFM منجر به تحلیل‌های دقیق‌تری شد تا بر آن اساس بتوان در حل مسائل استراتژیکی سازمان و تصمیم‌گیری استفاده کرد.

در این پژوهش، با استفاده از برآورد فاصله‌ای میانگین از درستی داده‌ها و نتایج، اطمینان حاصل شد. سه فاکتور مهمی که برای خوشه‌بندی از آنها استفاده کردیم شامل فاکتورهای $F \cdot R$ و M می‌باشد. با توجه به اینکه بیشترین تغییر در فاکتور F رخ می‌دهد لذا شرایط سه خوشه را بر مبنای این فاکتور مورد ارزیابی قرار می‌دهیم. برای این منظور ابتدا فاکتور فرکانس را بر مبنای زمان به صورت بازه‌ای تبدیل می‌کنیم، سپس در نرم افزار آماری minitab16، با درجه اطمینان ۹۵٪ به محاسبه بازه اطمینان می‌پردازیم. پس از محاسبه نتایج با استفاده از برآورد فاصله‌ای میانگین و رتبه‌بندی آن‌ها با استفاده از رابطه سنگوپتا، تایید می‌شود که نتایج پایدار است. با توجه به نتایج مقایسه‌ها، ترتیب رتبه‌بندی خوشه‌ها به صورت خوشه اول < خوشه سوم < خوشه دوم می‌باشد که این نتیجه پایداری و اطمینان نتایج حاصل از خوشه‌بندی را تایید می‌کند.

آنچه که در این پژوهش حائز اهمیت است، نحوه تحلیل داده‌ها و چگونگی استفاده از الگوریتم‌های داده‌کاوی به شکلی است که دقیق‌ترین و بهترین نتیجه را به همراه داشته باشد. در ادامه به ارائه پیشنهاداتی برای پژوهش‌های آتی می‌پردازیم:

۱. ترکیب روش‌های فازی با الگوریتم‌های خوشه‌بندی در شرایطی که نوع فیلد داده‌ها کیفی باشند.
۲. استفاده از الگوریتم‌های پیش‌بینی‌کننده مانند شبکه عصبی یا درخت تصمیم برای پیش‌بینی‌های فروش.
۳. تحلیل مشتریان داخل کشور و بررسی و پیش‌بینی نتایج با استفاده از روش داده‌کاوی.

مراجع

- [۱] فدایی نژاد، م.، صادقی شریف، ج و بنایان، طراحی سیستم پشتیبان تصمیم‌گیری جهت مدیریت بانکی از منظر تجهیز منابع (مورد پژوهشی بانک کشاورزی). مدیریت فناوری اطلاعات، ۳ (۶): ۱۰۸۰، ۱۳۹۰-۸۹.
- [۲] Jayachandran, S. Sharma, S. Kaufman, P. Raman, P, The role of relational information process and

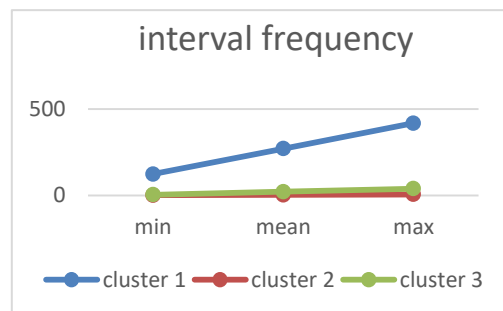
$$A (A < B) \begin{cases} = 0 & m(A) = m(B) \\ > 0, < 1 & aR > bL \\ & m(A) < m(B) \\ \geq 1 & aR \leq bL \\ & m(A) < m(B) \end{cases} \quad (۴-۱۳)$$

اگر $A (A < B) = 0$ ، آنگاه فرض «A پایین تر از B است» پذیرفته نمی‌شود. اگر $0 < A (A < B) < 1$ ، آنگاه مفسر فرض $A < B$ را با درجات مختلف رضایت از صفر تا یک (بدون احتساب صفر و یک) می‌پذیرد. اگر $A (A < B) \geq 1$ آنگاه مفسر مطلقاً از فرض $A < B$ راضی می‌شود، یعنی به عبارت دیگر می‌پذیرد که $A < B$ صحیح است. در واقع این رابطه روشی برای رتبه‌بندی واحدها یا مشتریان بر مبنای ارزش آن‌هاست که نتایج رتبه‌بندی می‌تواند در شرایط اطمینان با نتایج اولیه متفاوت باشد بنابراین در این قسمت با اعمال بازه اطمینان ۹۵٪ بررسی می‌کنیم تا ببینیم آیا نتایج پایدار است یا خیر. اگر نتایج منجر به رتبه‌بندی متفاوتی شود این نشان می‌دهد که انحراف معیار داده‌های اولیه منجر به گمراهی نتایج خوشه‌بندی شده است. پس از محاسبه نتایج با استفاده از برآورد فاصله‌ای میانگین و رتبه‌بندی آن‌ها با استفاده از رابطه سنگوپتا، تایید می‌شود که نتایج پایدار است نتایج نهایی محاسبات در جدول ۶ نمایش داده شده است.

جدول ۶: نتایج مقایسه خوشه‌ها در فرکانس خرید بر مبنای رابطه سنگوپتا

نتیجه بر مبنای درجه مقبولیت	نتیجه محاسبه فرمول سنگوپتا	مقایسه خوشه‌ها
تایید	۱,۰۱	خوشه اول < خوشه دوم
تایید	۰,۵۳	خوشه سوم < خوشه دوم
تایید	۰,۸۶۴	خوشه اول < خوشه سوم

با توجه به نتایج مقایسه‌های جدول ۶ ترتیب رتبه‌بندی خوشه‌ها به صورت خوشه اول < خوشه سوم < خوشه دوم می‌باشد که این نتیجه پایداری و اطمینان نتایج حاصل از خوشه‌بندی را تایید می‌کند. مقایسه شکل ۱۱ و شکل ۱۲ نیز این ادعا را اثبات می‌کند. با توجه به این دو شکل ملاحظه می‌شود نتایج پایدار است.



شکل ۱۱: وضعیت خوشه‌ها با محاسبه برآورد فاصله میانگین با درجه اطمینان ۹۵٪

- Based Systems, doi: 10.1016/j.knosys.2017.05.027.2017.
- [۱۶] محمدی، ش. و علیزاده، س. تحلیل مشکلات شعب بانک آینده در سراسر کشور با استفاده از روش داده‌کاوی، نشریه مدیریت فناوری اطلاعات 333-350. 1393.
- [۱۷] Singh, Hardeep "Clustering of text documents by implementation of K-means algorithms". Streamed *Info-Ocean*.62-12.2016.
- [۱۸] Davies D.L, Bouldin D.W.; "A cluster separation measure"; IEEE Transactions on Pattern Analysis and Machine, Vol. 24 No. 2, 1979.
- [۱۹] Iona, a, Vasile Mirea and Cezar Iescu, analysis of service management in the materials industry using the BCG matrix method, *Amfiteatru Economic*, Vol XI • Nr. 26 • June 2009.
- [۲۰] زینال زاده، ا. حسینی دولاما، جعفر. باقرزاده خواجه، مجید. ۱۳۸۹. برآورد فاصله‌ای زمان تکمیل پروژه بر اساس رویکرد شبیه سازی. فراسوی مدیریت. سال چهارم. ص ۲۳-۱۱.
- [۲۱] Jahanshahloo.G.H, Hosseinzadeh Lotfi.F, Davoodi.A.R. Extension of TOPSIS for decision-making problems with interval data: interval efficiency. *Mathematical and Computer Modeling* 49, pp. 1137-1142.2009.
- technology use in customer relationship management. *Journal of Marketing* 69(4):177-192 ,2005.
- [۳] Shu-Hsien Liao, Pei-Hui Chu, Pei-Yuan Hsiao, Data mining techniques and applications ,A decade review from 2000 to 2011, *Expert Systems with Application*, 2012.
- [۴] Loshin, David. Business intelligence : the savvy manager's guide. 2nd edition. Elsevier Inc. USA, 2013.
- [۵] Morgan, N.A., Rego, L. L. The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance. *Marketing Science Journal* , October, pp. 426-439. 2006.
- [۶] Ghazanfari, M. Malek Mohamadi, S. Alizadeh, S. Data mining application for country segmentation based on the RFM model. *Int.J. Data Analysis Techniques and Strategies*, Vol. 1, No. 2, 2008.
- [۷] رزمی، ج و قنبری، آ.ارائه مدلی نوین جهت محاسبه ارزش دوره عمر مشتری. مدیریت فناوری اطلاعات، ۱ (۲) :۵۰-۳۵. ۱۳۸۸.
- [۸] Dursun A , MeltemCaber, Using datamining techniques for profiling profitable hotel customers: An application of RFM analysis, *Tourism Management Perspectives* 18 (2016) 153-160,2016.
- [۹] برداران، و. و بیگلری، م. بخش بندی مشتریان صنایع تولید و پخش کالاهای پرگردش بر اساس مدل بهبود یافته RFM مطالعه موردی: شرکت گلستان. نشریه مدیریت بازرگانی 4-23. ۱۳۹۳.
- [۱۰] کرامتی، ع. و خالقی، ر. توسعه یک سیستم پیشنهاددهنده محصول طراحی مدل ترکیبی با بهره گیری از روش‌های فیلترینگ مشارکت محور، کشف قوانین انجمنی و بخش بندی مشتریان. نشریه تخصصی مهندسی صنایع، ۴۸ (۲) 280-257. ۱۳۹۳.
- [۱۱] [آخوندزاده نوقایی، ا.؛ البدوی، ا. و اقدسی، م. کاوش پویایی مشتری در طراحی بخش بندی با استفاده از روش‌های داده‌کاوی. نشریه مدیریت فناوری اطلاعات، 30-1-6. ۱۳۹۳.
- [۱۲] Soeini, R. A. & Fathalizade, E. Customer Segmentation RFM Model in the Insurance Based on Modified Industry. *Proceedings of 2012 4th International Conference on Machine Learning and Computing*, Singapore, 101-104, 2012.
- [۱۳] Wei J , Ming-Chun L , Hsuan-Kai Ch, H-Hung Wu, Customer relationship management in the hairdressing industry: An application of data mining techniques, *Expert Systems with Applications* 40, 7513-7518,2013.
- [۱۴] A. Joy Christy , A. Umamakeswari, L. Priyatharsini, A. Neyaa, RFM ranking – An effective approach to customer segmentation, *Journal of King Saud University – Computer and Information Sciences*, 2018.
- [۱۵] Meina Song, Xuejun Zhao, Haihong E, Zhonghong Ou, Statistics-based CRM approach via time series segmenting RFM on large scale data, *Knowledge-*



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و



دانشگاه صنعتی شاهرود

سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸

تحلیل رفتار گراف‌های تصادفی قدرت در مسأله جانمایی بهینه واحدهای اندازه‌گیری فاز در شبکه‌های هوشمند الکتریکی

سید محمد شهرآئینی^۱، احد الوندی^۲، و شهلا خرمالی^۲

^۱ استادیار گروه برق، دانشگاه گلستان، گرگان،

m.shahr@gu.ac.ir

^۲ دانشجوی کارشناسی مهندسی برق، دانشگاه گلستان، گرگان

shahla.khormali3011@gmail.com و ahadalvandi@yahoo.com

چکیده

در این مقاله، دو روش مختلف برای ایجاد گراف‌های تصادفی قدرت پیشنهاد شده است. دو روش فوق مبتنی بر دو مدل گراف‌های تصادفی یعنی روش‌های ادگار گیلبرت و اردوس رینی هستند. برای اثبات کارایی گراف‌های تصادفی در تحلیل سیستم‌های قدرت و مشابهت گراف‌های ایجاد شده با این روش‌ها به سیستم قدرت، مسأله جانمایی واحدهای اندازه‌گیری فاز روی گراف‌های تصادفی ایجاد شده پیاده و حل شده است. نتایج شبیه‌سازی حاکی از آن است که دو مدل گراف‌های تصادفی قدرت پیشنهادی قابلیت خوبی در ایجاد شبکه‌هایی مشابه با شبکه قدرت دارند. این مشابهت در حل مسأله جانمایی واحدهای اندازه‌گیری فاز، که یکی از شناخته شده‌ترین مسائل مرتبط با ساختار شبکه قدرت است؛ به اثبات رسیده است.

کلمات کلیدی

گراف تصادفی قدرت، جانمایی بهینه واحدهای اندازه‌گیری فاز، تخمین حالت، مشاهده‌پذیری

OPP اطلاق خواهیم کرد؛ قصد دارد با جانمایی تعداد کمینه واحد اندازه‌گیری فاز کل شبکه قدرت را مشاهده‌پذیر کند [۲]. در حالت کلی دو گونه مختلف OPP وجود دارد [۳]: نوع اول: OPP با اهداف اقتصادی و نوع دوم: OPP با اهداف کارایی.

نوع اول OPP قصد دارد با کمینه‌ترین تعداد واحد اندازه‌گیری فاز، یک یا چند کارایی بهره‌برداری را بهبود بخشد. غالباً این هدف کارایی مشاهده‌پذیری کل سیستم قدرت است اما اهداف دیگری مانند پایداری ولتاژ و یا حذف نوسانات فرکانس کوچک نیز می‌توانند مدنظر باشند.

۱- مقدمه

«واحد اندازه‌گیری فاز» یک تجهیز جدید اندازه‌گیری است که در اواخر قرن ۲۰ میلادی به سیستم قدرت معرفی شد. این تجهیز با استفاده از ساعت سیستم موقعیت یاب جهانی (GPS) و پردازنده دیجیتال سیگنال قابلیت اندازه‌گیری فاز و فرکانس شکل موج‌های ولتاژ و جریان را در گستره‌ی وسیع جغرافیائی داراست [۱]. از ابتدای معرفی واحد اندازه‌گیری فاز (PMU)، مسأله جانمایی این واحدها همواره مورد توجه پژوهشگران بوده است [۴-۸]. مسأله جانمایی بهینه PMU (Optimal PMU Placement) که پس از این به اختصار به آن

گراف‌های تصادفی اختصاص دارد. همچنین در این بخش گراف تصادفی قدرت معرفی می‌شود و پارامترهای دو مدل ارائه شده به گونه‌ای تنظیم می‌شود که گراف تصادفی حاصل مشابه شبکه قدرت باشد. در بخش چهارم یازده کیس نمونه IEEE بررسی و هر دو مدل ارائه شده در بخش سوم برای بازسازی یک گراف مشابه برای هر کیس پیاده خواهد شد. همچنین در این بخش برای اثبات کارایی مدل‌های پیشنهادی، یک گراف تصادفی قدرت ۱۰۰۰ باس تولید و مسأله جانمایی PMU برای آن حل خواهد شد. این مقاله با نتیجه‌گیری در بخش پنجم پایان می‌یابد.

۲- تخمین حالت و مشاهده‌پذیری

در حالت کلی، تخمین حالت عبارت است از محاسبه حالت‌های دقیق سیستم با استفاده از اندازه‌گیری‌های نادقیق و دارای خطا [۲۵]. حالت در سیستم قدرت اندازه ولتاژ باس‌ها ($|V|$) و فاز ولتاژ باس‌ها (δ) به جز باس شماره ۱ (δ_1) است. برای یک سیستم N باس، بردار حالت (x) دارای بعد $(n=2N-1)$ است. در حالتی که تعداد پارامترهای اندازه‌گیری شده در شبکه m باشد، معادلات تخمین حالت به فرم زیر هستند:

$$z = h(x) + e \quad (1)$$

که در آن، z ماتریس اندازه‌گیری و e خطاهای اندازه‌گیری هستند و هر دو دارای ابعاد $(m \times 1)$ هستند؛ و $h(x)$ معادلات شبکه است که اندازه‌گیری‌ها را به حالت‌های شبکه مرتبط می‌کند و دارای ابعاد $(m \times n)$ است.

در حالت کلی بسته به نوع داده‌های اندازه‌گیری شده، معادلات تخمین حالت به فرم خطی و یا خیر خطی خواهند بود. در شرایط استفاده از داده‌های ولتاژ، جریان و توان اکتیو و راکتیو که توسط سیستم SCADA اندازه‌گیری شده‌اند، معادلات بدست آمده غیرخطی است و به این سیستم «تخمین‌گر سنتی»^۱ گفته می‌شود. در حالت استفاده از داده‌های هم‌زمان فاز PMUها، معادلات بدست آمده خطی است و به این نوع «تخمین‌گر بر پایه PMU^۲» شناخته می‌شود. نهایتاً در حالت استفاده از هر دو نوع داده (PMU و SCADA) معادلات غیرخطی است و این تخمین‌گر بنام هیبرید^۳ شناخته می‌شود [۹].

در معادلات تخمین حالت مرتبه ماتریس h دارای اهمیت بالایی است و مشاهده‌پذیری سیستم را تعیین می‌کند [۲۵]: اگر $\text{Rank}(h) = n$ سیستم «کاملاً مشخص»^۴ است؛ در حالت $\text{Rank}(h) > n$ سیستم «فرا مشخص»^۵ است؛ و اگر $\text{Rank}(h) < n$ سیستم «فرو مشخص»^۶ خواهد بود. روش دیگری نیز برای محاسبه مشاهده‌پذیری سیستم وجود دارد که به آن «مشاهده‌پذیری توپولوژیکی»^۷ گویند. مشاهده‌پذیری توپولوژیکی با استفاده از مبانی

هدف در نوع دوم OPP بدست آوردن حداکثر یک کارایی با تعداد مشخص PMU است. مانند نوع اول غالباً این کارایی مشاهده‌پذیری شبکه قدرت انتخاب می‌شود.

همانگونه که در بالا ذکر شد غالباً هدف کارایی در OPP مشاهده‌پذیری سیستم قدرت است [۳]. در حالت کلی دو روش کلی برای محاسبه مشاهده‌پذیری سیستم قدرت وجود دارد [۹]: مشاهده‌پذیری جبری و مشاهده‌پذیری ساختاری.

مشاهده‌پذیری ساختاری با استفاده از نظریه گراف‌ها و روابطی که در شبکه قدرت حاکم است تعیین می‌کند که یک توزیع خاص از تعداد مشخص PMU سیستم فوق را مشاهده‌پذیر می‌کند یا خیر. در این روش سیستم قدرت توسط یک گراف $G(V,E)$ بیان می‌شود که باس‌های سیستم گره‌های این گراف (V) و خطوط انتقال یال‌های آن (E) هستند.

تحلیل شبکه‌ی قدرت به عنوان یک شبکه پیچیده (Complex Network) در سال‌های اخیر مورد توجه پژوهش‌گران بوده است [۱۰]. در این پژوهش‌ها سیستم قدرت به عنوان یک شبکه پیچیده و به صورت گراف مدل می‌شود و با ترکیب تئوری گراف‌ها و روابط حاکم در سیستم قدرت، روش‌های جدید تحلیل سیستم قدرت ارائه شده است. تعریف معیارهای میانگی [۱۱] و همچنین تحلیل قابلیت اطمینان [۱۲] نمونه‌هایی از این روش‌ها هستند.

مدل‌ها و تحلیل‌های مرتبط با سیستم قدرت غالباً روی سیستم‌های تست IEEE اجرا می‌شود که در طی سالیان مختلف از سیستم‌های قدرت حقیقی الهام گرفته شده‌اند [۱۵]. تعداد و تنوع این سیستم‌های تست زیاد نیست و همین امر تحلیل سیستم‌های قدرت به عنوان شبکه‌های پیچیده را با محدودیت روبرو کرده است. برای رفع این نقیصه پژوهش‌گران روش‌هایی را پیشنهاد کرده‌اند که گراف‌های تصادفی (Random Graph) پیشنهاد می‌کنند که دارای خواص مشابه با سیستم‌های تست IEEE است [۱۳-۱۴].

گراف تصادفی گرافی است که خواص آن از قبیل تعداد رأس‌ها (Vertices)، تعداد یال‌ها (Edges)، و یا نحوه اتصال رأس‌ها به یک دیگر به روشی تصادفی تعیین شده است [۱۶]. انواع مختلف گراف تصادفی معرفی شده است که گراف‌های تصادفی «اردوس رینی» و «ادگار گیلبرت» نمونه‌هایی از آن‌ها هستند.

با در نظر گرفتن مواردی که تا کنون مطرح شد چنین می‌توان نتیجه گرفت که بسیاری از تحلیل‌های مرتبط با شبکه قدرت (مانند تحلیل مشاهده‌پذیری ساختاری) بر روی گراف استخراج شده از این شبکه‌ها قابل انجام است. از سوی دیگر تعداد و تنوع شبکه‌های تست IEEE زیاد نیست و این امر سبب می‌شود که امکان تحلیل و بررسی مدل‌ها در سیستم‌های قدرت کاهش یابد. به همین دلیل انگیزه‌ی این تحقیق ارائه دو مدل جهت ایجاد گراف تصادفی قدرت و بررسی رفتار این گراف‌ها در حل مسأله جانمایی واحدهای اندازه‌گیری فاز (OPP) است.

در بخش دوم مقاله تخمین حالت، مشاهده‌پذیری سیستم قدرت و مسأله جانمایی واحدهای اندازه‌گیری فاز معرفی می‌شوند. بخش سوم این مقاله به معرفی گراف‌های تصادفی و دو مدل پرکاربرد در تولید

¹Traditional State Estimator

²PMU based State Estimator

³Hybrid State Estimator

⁴Fully Determined

⁵Over Determined

⁶Under Determined

⁷Topological Observability

که در آن، A^+ ماتریس مجاورت تعمیم یافته است؛ PMU^+ برداری n بعدی است که مکان PMUها را مشخص می‌کند؛ و $\hat{1}$ بردار n بعدی است که همه درایه‌های آن ۱ است. توضیح اینکه در رابطه (۴) شرط بزرگ‌تری به معنی بزرگ‌تر بودن تک تک درایه‌ها است.

۲-۱- جانمایی بهینه واحدهای اندازه‌گیری فاز

همانگونه که در مقدمه مطرح شد، جانمایی بهینه واحدهای اندازه‌گیری فاز (OPP) یک مسأله شناخته شده است که سعی دارد یک کارایی مشخص در سیستم قدرت را به صورت بهینه پیاده‌سازی کند [۴-۸]. همچنین مطرح شد که دو نوع مختلف OPP وجود دارد: OPP با اهداف اقتصادی و OPP با اهداف کارایی [۳].

در این پژوهش OPP نوع اول (یعنی هدف اقتصادی) مدنظر است و کارایی مورد نظر مشاهده‌پذیری کل سیستم تعریف می‌شود. با توجه به موارد طرح شده مسأله جانمایی بهینه واحدهای اندازه‌گیری فاز به فرم یک مسأله بهینه‌سازی و به صورت زیر قابل بیان است:

$$\begin{cases} \text{Min Cost}_{PMU} \\ \text{s.t. } A^+ \cdot PMU > \hat{1} \end{cases} \quad (5)$$

روش‌های مختلفی برای حل مسأله بهینه‌سازی مانند برنامه‌ریزی خطی [۴ و ۶]، الگوریتم ژنتیک [۵]، آبکاری فلزات [۸] استفاده شده است. در این پژوهش، الگوریتم ژنتیک برای حل مسأله بهینه‌سازی فوق استفاده شده است.

۳- گراف‌های تصادفی

همانطور که پیشتر مطرح شد، گراف تصادفی $G(V, E)$ گرافی است که خواص آن از قبیل تعداد رأس‌ها (V)، تعداد یال‌ها (E)، و یا نحوه اتصال رأس‌ها به یکدیگر به روشی تصادفی تعیین می‌شود [۱۶]. همچنین مطرح شد که تا کنون انواع مختلف گراف تصادفی معرفی شده است که هرکدام از آن‌ها دارای خواص مشخصی است.

مدل‌های گراف تصادفی متفاوت، توزیع‌های احتمال متفاوتی را در نمودارها تولید می‌کنند. این مدل‌ها بوسیله شروع با مجموعه‌ای از تعداد n رأس‌های جدا از هم و افزودن گره‌های پایایی بین رأس‌های تصادفی بدست می‌آیند. دو مدل گراف تصادفی شناخته شده، مدل «اردوس رینی» و مدل «ادگار گیلبرت» می‌باشند [۱۶].

۳-۱- مدل ادگار گیلبرت

بیشترین مورد مطالعه، مدلی است که توسط ادگار گیلبرت، به صورت $G(n, p)$ ارائه شده است، که در آن هر گره ممکن به طور مستقل با احتمال $1-p$ رخ می‌دهد. احتمال به دست آوردن هر

تئوری گراف‌ها و همچنین نوع و محل دستگاه‌های اندازه‌گیری، مشاهده‌پذیری بودن سیستم را مشخص می‌کند. همانگونه که قبلاً مطرح شد، PMU قابلیت اندازه‌گیری فاز هم‌زمان ولتاژ و جریان را داراست و این باعث می‌شود که اگر در یک باس PMU با تعداد کافی کانال اندازه‌گیری ولتاژ و جریان نصب شود، آن باس و تمامی باس‌های مجاور خود را مشاهده‌پذیر کند [۹]. توضیح اینکه اندازه‌گیری ولتاژ توسط «ترانس اندازه‌گیری ولتاژ» (PT^A) انجام می‌شود و سبب می‌شود باس مجهز به PMU مشاهده‌پذیر شود؛ و اندازه‌گیری جریان عموماً توسط «ترانس اندازه‌گیری ولتاژ» (CT^A) انجام می‌شود و باعث می‌شود شود باس مجاور مشاهده‌پذیر شود. با در نظر گرفتن این قابلیت ذکر شده‌ی PMU و در شرایطی که تعداد معینی از این واحدها و با توزیع مناسب در کل شبکه قدرت نصب شده باشد، می‌توان مشاهده‌پذیری کل سیستم را با استفاده از «ماتریس مجاورت»^۸ بدست آورد. در زیربخش بعدی به معرفی این ماتریس و نحوه محاسبه مشاهده‌پذیری توپولوژیکی می‌پردازیم.

فرض کنید که کل شبکه الکتریکی با گراف پیوسته و بدون جهت $G_E(V_E, E_E)$ نمایش داده شود؛ بطوری که V_E رأس‌های گراف باس‌های شبکه الکتریکی و E_E یال‌ها در این گراف خطوط انتقال هستند. این گراف می‌تواند با استفاده از ماتریس مجاورت بیان شود. بنا به تعریف برای یک گراف با n رأس، ماتریس مجاورت (A) یک ماتریس $(n \times n)$ است که درایه‌های آن به صورت زیر تعیین می‌شوند:

$$A = [a_{ij}]_{N \times N} \quad (2)$$

$$a_{ij} = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ bus connect to } j^{\text{th}} \text{ bus} \\ 0 & \text{otherwise} \end{cases}$$

از روی ماتریس مجاورت می‌توان «ماتریس مجاورت تعمیم یافته»^۹ را تشکیل داد که به صورت زیر تعریف می‌شود:

$$M = xI_{N \times N} + yJ_{N \times N} + zA_{N \times N} \quad (3)$$

$$x, y, z \in \mathbb{R}, z \neq 0$$

که در آن، I ماتریس همانی با سایز N و J ماتریس مربعی N بعدی است که تمام درایه‌های آن مساوی با ۱ است.

حالت خاص ($y=0$ و $x=z=1$) ماتریسی مجاورت تعمیم یافته‌ی

خاصی (A^+) را ایجاد می‌کند که در تحلیل مشاهده‌پذیری توپولوژیکی گراف با استفاده از PMU کاربرد دارد. هر سطر و یا ستون این ماتریس به یک رأس و تمامی رأس‌های مجاور آن اشاره می‌کند، در نتیجه اگر مکان PMUها در یک گراف n رأسی با یک بردار بیان شود، یک سیستم مشاهده‌پذیر توپولوژیکی است اگر شرط زیر برقرار باشد:

$$A^+ \cdot PMU > \hat{1} \quad (4)$$

⁸Potential Transformer

⁹Current Transformer

¹⁰Adjacency Matrix

¹¹Generalized Adjacency Matrix

عناصر است و هر عنصر با احتمال $\binom{N}{M}^{-1}$ رخ می‌دهد. متغیر

تصادفی G^M گرافی را ایجاد می‌کند که از این طریق ایجاد می‌شود [۲۴-۲۲].

در مدل اردوس رینی که به صورت $G(n, M)$ مدل می‌شود، M نشان دهنده تعداد گره هاست که احتمال برابر را به همه نمودارها با گره‌هایی به تعداد دقیقاً M گره اختصاص می‌دهد. هر عنصر با احتمال $1/\binom{N}{M}$ رخ می‌دهد. یک نمودار به طور یکتا، در

حالت تصادفی، از مجموعه کلیه نمودارهایی که n تعداد گره و M تعداد لبه (حاشیه) دارند انتخاب می‌شود [۱۹-۲۱].

۳-۳- تشابه‌ها و تفاوت‌های دو مدل

در رابطه با مقایسه مدل‌های اصلی، مدل $G(n, p)$ (مدل ادگار گیلبرت) رایج ترین مدل است. مدل $G(n, p)$ از لحاظ ریاضیاتی به راحتی حل و فصل نمی‌شود. مدل سازی $G(n, M)$ و $G(n, p)$ با استفاده از رابطه $M = \binom{n}{2} * P$ انجام می‌شود. با نزدیک شدن n

به بی‌نهایت، $G(n, p)$ باید مشابه $G(n, M)$ رفتار کند. قانون بزرگی اعداد می‌گوید $G(n, p)$ تقریباً شامل همان تعداد گره‌ها همانند $G(n, M)$ خواهد بود [۱۶-۱۸].

در بحث پیرامون تکامل $G(n, p)$ ، اگر $1 < np$ باشد، $G(n, p)$ به احتمال زیاد هیچیک از اجزای متصل بزرگتر از اندازه $O(\log(n))$ را ندارد و اجزاء از درختان مربوط به گراف تشکیل شده‌اند. اگر $np = c$ و به طوریکه $c > 1$ باشد، چرخه‌ها ظاهر می‌شوند و تقریباً همه راس‌ها در درخت‌ها، متصل می‌شوند. اگر $np = 1$ باشد، $G(n, p)$ به احتمال زیاد بزرگترین مؤلفه سفارشی $n^{2/3}$ را دارد و اگر $np < 1$ باشد، $G(n, p)$ به احتمال زیاد یک جزء غول پیکر منحصر به فرد را دارا می‌باشد که حاوی بخشی از رئوس است. هیچ مؤلفه دیگری حاوی رئوس بیشتر از $O(\log(n))$ نخواهد بود [۱۶-۱۸].

۳-۴- گراف‌های تصادفی قدرت

پس از بررسی دو مدل شناخته شده گراف تصادفی، در این بخش به معرفی گراف تصادفی قدرت می‌پردازیم. گراف تصادفی قدرت بیشتر توسط پژوهشگران مطرح شده است [۱۳-۱۴].

به صورت کلی یک شبکه قدرت را می‌توان در قالب یک گراف $G(V, E)$ نمایش داد؛ که در آن مجموعه رئوس $\{V\}$ مجموعه کلیه باس‌های آن شبکه و مجموعه $\{E\}$ مجموعه خطوط انتقال شبکه قدرت مذکور می‌باشند. در شبکه قدرت مورد بحث تعداد رئوس (باس‌ها) و تعداد خطوط انتقال را به ترتیب با n و M نمایش می‌دهیم. از مدت‌ها پیش و با الهام گرفتن از شبکه‌های واقعی انتقال (و گاهاً توزیع)،

نمودار تصادفی خاص با M گره به صورت $p^m(1-p)^{N-m}$ است که نماد $N = \binom{n}{2}$ [۱۶-۱۸].

در مدل ادگار گیلبرت که به صورت $G(n, p)$ مدل می‌شود، n تعداد راس‌های جدا از هم، p احتمال رخ دادن هر گره ممکن به صورت مستقل نسبت به یکدیگر و M تعداد گره‌ها می‌باشد. همچنین احتمال به دست آوردن یک نمودار خاص به صورت $p^m(1-p)^{N-m}$ می‌باشد که مقدار $N = \binom{n}{2}$ است. نمودار با اتصال گره‌ها به طور تصادفی ساخته می‌شود. با افزایش p از ۰ به ۱، احتمال زیادی وجود دارد که نمودارهایی با تعداد گره‌های بیشتر تولید شود [۱۶-۱۸].

در مدل ادگار گیلبرت $G(n, P(\text{edge}) = p)$ شامل تمام نمودارهایی است که دارای مجموعه راس $V = \{1, 2, \dots, n\}$ می‌باشد که در آن گره‌ها به طور مستقل و با احتمال p انتخاب می‌شوند. به عبارت دیگر، اگر G_0 گرافی با مجموعه V راس و تعداد M گره باشد، سپس رابطه زیر برقرار است:

$$P(\{G_0\}) = P(G = G_0) = p^m(1-p)^{N-m} \quad (۶)$$

متغیر تصادفی G_p گرافیکی را ایجاد می‌کند که از این طریق نشان داده می‌شود. برای $M \cong pN$ ، این دو مدل تقریباً قابل تعویض هستند [۲۴-۲۲].

۳-۲- مدل اردوس رینی

یک مدل مشابه نزدیک به مدل پیشین مطرح شده، مدل اردوس رینی است که به صورت $G(n, M)$ ارائه شده است و احتمال برابر را به تمام نمودارها با گره‌هایی دقیقاً به اندازه M اختصاص می‌دهد. با $0 < M < N$ ، عناصر $G(n, M)$ دارای $\binom{N}{M}$ عناصر است و هر عنصر با احتمال $1/\binom{N}{M}$ رخ می‌دهد. این مدل را می‌توان به عنوان یکتصویر

لحظه‌ای در یک زمان خاص (M) از روند نمودار تصادفی G_n مشاهده کرد، که یک فرآیند تصادفی است که با n راس و بدون گره شروع می‌شود و در هر مرحله یک گره جدید اضافه می‌کند که به طور یکتا از مجموعه گره‌های از دست رفته انتخاب شده است [۱۶-۱۸].

مدل اردوس رینی $G(n, M)$ شامل تمام نمودارهایی است که دارای مجموعه راس $V = \{1, 2, \dots, n\}$ که M تعداد گره دارد، می‌باشد، که در آن نمودارها احتمال یکسانی دارند. بنابراین با مقادیر $0 \leq M \leq N$ و $N = \binom{n}{2}$ ، گراف $G(n, M)$ دارای $\binom{N}{M}$

IEEE شبکه‌های تست مختلف با اندازه‌های مختلف ارائه کرده است که اطلاعات کامل آنها در [۱۵] قابل دسترس می‌باشد.

جدول ۱۴: نتایج مربوط به کیس‌های نمونه IEEE و دو گراف تصادفی قدرت ایجاد شده برای هر کیس

Calculation Method	#Node	#Links	P	%PMU	Degree Mean	Degree Variance
1.Test Case	14	20	-	28.571	2.8571	1.3626
2.Edros		20	-	28.571	2.8571	2.7473
3.Gillbert		19	0.2	35.714	2.7143	2.8352
1.Test Case	24	34	-	37.5	2.8333	1.0145
2.Edros		34	-	37.5	2.8333	2.5797
3.Gillbert		32	0.13	37.5	2.6667	1.5362
1.Test Case	30	41	-	36.667	2.7333	1.9954
2.Edros		41	-	36.667	2.7333	1.7885
3.Gillbert		41	0.11	36.667	2.7333	1.9264
1.Test Case	39	46	-	41.026	2.3590	1.0256
2.Edros		46	-	41.026	2.3590	2.4993
3.Gillbert		43	0.06	41.026	2.2051	1.4305
1.Test Case	57	78	-	35.088	2.7368	1.2331
2.Edros		78	-	38.596	2.7368	2.4474
3.Gillbert		82	0.05	38.596	2.8772	2.8596
1.Test Case	118	179	-	32.203	3.0339	2.4775
2.Edros		179	-	38.983	3.0339	2.8023
3.Gillbert		172	0.025	41.525	2.9153	3.0868
1.Test Case	145	422	-	24.828	5.8207	22.1204
2.Edros		422	-	26.207	5.8207	5.5093
3.Gillbert		411	0.040	25.517	5.6690	5.0147
1.Test Case	200	245	-	35	2.4500	2.9020
2.Edros		245	-	42.5	2.4500	2.2588
3.Gillbert		243	0.012	44	2.4300	2.5177
1.Test Case	300	409	-	36	2.7267	2.3799
2.Edros		409	-	41	2.7267	2.6675
3.Gillbert		408	0.009	42.667	2.7200	2.9715
1.Test Case	1354	1710	-	36.632	2.5258	3.7129
2.Edros		1710	-	41.728	2.5258	2.3663
3.Gillbert		1710	0.0026	41.95	2.5258	2.7048
1.Test Case	1888	2308	-	38.771	2.4449	3.6313
2.Edros		2308	-	43.962	2.4449	2.4093
3.Gillbert		2308	0.0013	43.273	2.4417	2.4778

گرفته می‌شود. سپس در گام نخست تعداد گره‌ها و یال‌ها، میانگین و انحراف معیار در «معیار میانگی درجه»، و نهایتاً درصد باس‌های پوشش داده شده با PMU برای مشاهده‌پذیری کامل این شبکه‌ها بررسی و محاسبه و تحلیل می‌شوند.

در گام بعدی، برای هر کیس و با توجه به مشخص بودن مقدار n و M که به ترتیب تعداد رأس‌ها و یال‌ها را بیان می‌کنند؛ با استفاده از مدل اردوس رینی گراف تصادفی ایجاد می‌شود و پارامترهای میانگین و واریانس میانگی درجه، و درصد باس‌های تحت پوشش PMU برای مشاهده‌پذیری کامل محاسبه می‌شوند.

در آخرین گام و با استفاده از مدل ادگار گیلبرت، تعداد گره‌ها برابر هر کدام از کیس‌ها فرض می‌شود و با تغییر احتمال P در این مدل، گراف تصادفی ایجاد می‌شود که دارای تعداد لینک‌های تقریباً مساوی با شبکه تست IEEE مدنظر داشته باشد. در این حالت پارامترهای تعداد یال‌ها، احتمال P ، میانگین و واریانس میانگی درجه،

هر دو مدل تصادفی معرفی شده در بخش‌های پیشین قابلیت مدل‌سازی شبکه‌های قدرت را دارند. به همین دلیل به گراف تصادفی که یک شبکه قدرت را مدل می‌کند «گراف تصادفی قدرت» خواهیم گفت. در بخش بعدی به این مدل‌سازی خواهیم پرداخت.

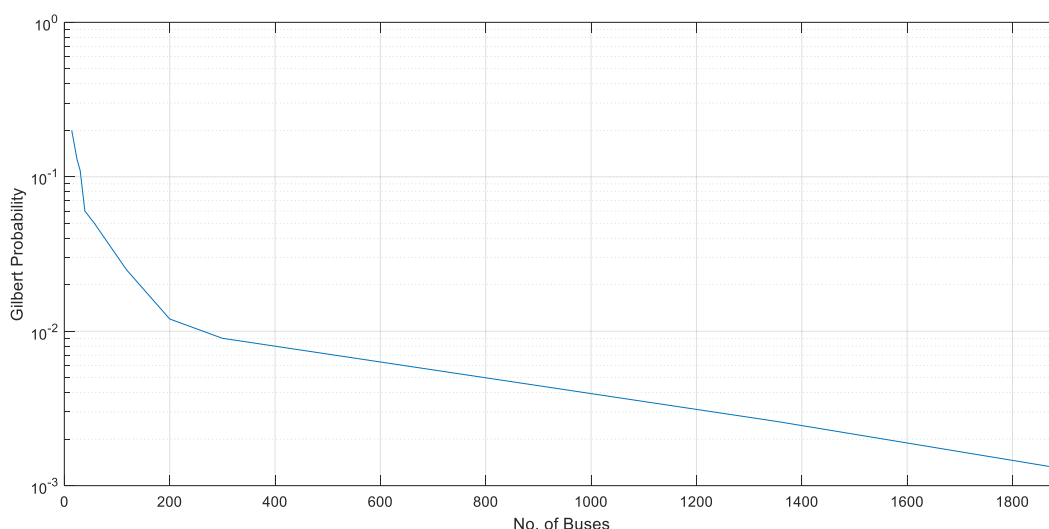
۴- مدل‌سازی شبکه‌های تست IEEE بوسیله گراف‌های تصادفی

پیشتر مطرح شد که IEEE شبکه‌های تست مختلف با سایزهای مختلف ارائه کرده است [۱۵]. همچنین مطرح شد که هر دو مدل گراف تصادفی ارائه شده (یعنی اردوس رینی و ادگار گیلبرت) قابلیت ایجاد گراف تصادفی قدرت را دارا هستند.

در این بخش ابتدا گراف‌های متناظر با یازده شبکه تست ۱۴، ۲۴، ۳۰، ۳۹، ۵۷، ۱۱۸، ۱۴۵، ۲۰۰، ۳۰۰، ۱۳۵۴، و ۱۸۸۸ در نظر

و درصد باس‌های تحت پوشش PMU به جهت مشاهده‌پذیری کامل برای هر کیس محاسبه می‌شود.

نتایج هر سه گام بالا برای یازده کیس مورد بررسی در جدول ۱ ارائه شده است. برای هر کیس سه ردیف مجزا وجود دارد که به معنی تحلیل سه گراف متفاوت برای هر کیس است: گراف اول گراف متناظر با کیس IEEE و دو گراف تصادفی با مشخصات مشابه با آن کیس.



شکل 43: تغییرات احتمال مدل تصادفی گیلبرت (P) بر حسب تعداد گره-های گراف

انتقال (M) حدوداً برابر با کیس واقعی باشد. در سطر سوم مرتبط با هر کیس این مقدار احتمال علاوه بر اطلاعات پیشین درج شده است. بررسی مقدار P نشان می‌دهد که با رشد سایز شبکه این احتمال کاهش می‌یابد و بین تعداد گره‌های گراف تصادفی گیلبرت قدرت و احتمال P یک رابطه نمائی وجود دارد. شکل ۱ این رابطه نمائی را نشان می‌دهد. همچنین درصد پوشش PMU بدست آمده برای گراف‌های تصادفی قدرت گیلبرت و مقایسه آنها با کیس واقعی حاکی از این حقیقت است که گراف گیلبرت نیز قابلیت خوبی در شبیه‌سازی شبکه‌های قدرت دارد و مانند مدل تصادفی اردوس، مشابهت در سایزهای پائین‌تر بهتر است. در انتها و با دقت روی گراف ارائه شده در شکل ۱ می‌توان وجود رابطه نمائی بین تعداد گره‌ها و احتمال P در کیس‌های ۳۰۰ به بالا را مشاهده نمود. ضمناً پیشتر و در رابطه (۷) مطرح شد که میانگین معیار درجه در شبکه‌های قدرت بین ۲ تا ۳ است و لذا می‌توان محدوده تغییرات تعداد لینک‌ها در یک شبکه با تعداد گره مشخص بدست آورد.

جدول ۲: نتایج مربوط به گراف تصادفی قدرت ۱۰۰۰ باس با دو مدل

Method	# N	# L	P	% PMU	Degree Mean	Degree Variance
Edros	1000	1500	-	39.90	3	3.0731
Gillbert	1000	1688	.0033	36.00	3.3760	3.3500

در نتیجه، برای اثبات کارایی مدل گراف تصادفی قدرت ارائه شده در این تحقیق کیس نمونه جدید ۱۰۰۰ باس به صورت تصادفی و با دو مدل اردوس و گیلبرت به صورت زیر ایجاد شده‌اند: در مدل اردوس میانگین میانگین درجه عدد سه انتخاب شده است و برای مدل گیلبرت با توجه به نمودار شکل ۱ احتمال 0.0033 انتخاب شده است. در جدول ۲ مقادیر ورودی در هر مدل به صورت برجسته و زیرخط‌دار مشخص شده‌اند. پارامترهای گراف تصادفی حاصل در جدول ۲ آمده است. نتایج حاصل حاکی از کارایی دو مدل ارائه شده است.

بررسی اطلاعات ارائه شده در سطر اول هر کیس صحت این ادعا را تایید می‌کند که با پوشش حدود یک سوم باس‌ها کل شبکه‌های قدرت (به استثنای یک مورد) مشاهده‌پذیر می‌شوند [۲]. همچنین اطلاعات مربوط به میانگین معیار درجه حاکی از آن است که در همگی کیس‌ها (به استثنای کیس ۱۴۵ باس) این مقدار بین ۲ تا ۳ است و با افزایش این مقدار درصد باس‌های تحت پوشش PMU افزایش می‌یابد (کیس ۱۱۸ و ۱۴۵ را ببینید). توجه این موضوع این است که مقدار بالاتر میانگین میانگین درجه به معنی تعداد بیشتر متوسط اتصالات رأس‌ها به یکدیگر است. طبق خواصی که برای PMU تعریف شد، اتصالات بیشتر در شبکه به معنی تعداد کمتر PMU است. در واقع میانگین معیار میانگین درجه در یک شبکه‌های قدرت که دارای n باس و M خط انتقال است را می‌توان به این صورت محاسبه کرد: هر خط انتقال دو باس را به یکدیگر متصل کرده؛ در نتیجه شبکه دارای $2M$ اتصال است، پس برای محاسبه میانگین معیار میانگین داریم:

$$\text{Mean(degree)} = \frac{2M}{n} \quad (7)$$

همچنین در کلیه کیس‌ها (به استثنای کیس خاص ۱۴۵ باس) انحراف معیار میانگین درجه پائین بوده و بین ۱ تا ۴ هستند.

سطر دوم کیس‌ها گراف‌های تصادفی هستند که با مدل اردوس رینی و با ورودی‌های $G(n, M)$ ایجاد شده‌اند. برای هر یک از گراف‌های اردوس ایجاد شده برای هر کیس پارامترهای ذکر شده در بالا اندازه‌گیری شده‌اند. بررسی این پارامترها نشان می‌دهد که مدل اردوس رینی قابلیت خوبی در ایجاد گراف‌های تصادفی قدرت دارد و درصد پوشش PMU در سطرهای دوم کیس‌ها تاییدی بر این نکته می‌باشد. همچنین مشابهت گراف تصادفی قدرت پیشنهاد شده با کیس‌های حقیقی در شبکه‌ها با سایز کوچک‌تر بیشتر است.

نهایتاً و برای گراف تصادفی ادگار گیلبرت برای هر کیس با روش سعی و خطا مقدار احتمال P به گونه‌ای انتخاب شده که تعداد خطوط

۵- نتیجه‌گیری

- در این مقاله، ابتدا مسأله جانمایی واحدهای اندازه‌گیری فاز معرفی و نحوه حصول مشاهده‌پذیری کامل در سیستم قدرت با توجه به گراف این شبکه مطرح شد. سپس مفهوم گراف‌های تصادفی قدرت معرفی شد و دو روش شناخته شده در تولید گراف‌های تصادفی (یعنی روش‌های ادگار گیلبرت و اردوس رینی) معرفی شدند و به گونه‌ای تعمیم یافتند که قابلیت ایجاد گراف تصادفی را داشته باشند. یازده کیس مختلف IEEE بررسی شد، و برای هر کیس دو گراف تصادفی قدرت ایجاد شد و هر سه گراف (گراف اصلی و دو گراف تصادفی) از نقطه نظر پارامترهای: تعداد گره، تعداد لینک، درصد پوشش PMU برای حصول مشاهده-پذیری، و میانگین و واریانس معیار میانگی درجه بررسی شدند. نتایج حاصل از بررسی کیس‌ها و گراف‌های تصادفی قدرت مشابه آنها به قرار زیر هستند:
- شبکه‌های قدرت (فارغ از اندازه‌ی آنها) دارای متوسط معیار درجه بین ۲ تا ۳ هستند؛ یعنی در شبکه‌های قدرت به طور میانگین هر باس به دو یا سه باس دیگر از طریق خطوط انتقال متصل است. در نتیجه در این شبکه‌ها با جانمایی PMU روی حدود یک سوم باس‌ها مشاهده-پذیری کامل حاصل می‌شود. بین متوسط معیار میانگی درجه و درصد PMU مورد نیاز رابطه معکوس برقرار است یعنی با افزایش این مقدار متوسط درصد فوق کم می‌شود. احتمال اتصال گره در مدل تصادفی ادگار گیلبرت با افزایش سایز شبکه کاهش می‌یابد که کاملاً طبیعی است؛ زیرا باید مقدار متوسط دو تا سه اتصال برای هر گره حفظ شود و با افزایش سایز گراف کاهش این احتمال این مهم را محقق می‌کند. همچنین احتمال گیلبرت و سایز شبکه دارای یک رابطه نمائی کاهشی هستند. نهایتاً نتایج حاصل نشان می‌دهد که مشابهت گراف‌های تصادفی ارائه شده در سایزهای بالا کاهش می‌یابد، اما همچنان در شرایط عدم وجود تنوع در کیس‌های نمونه، گراف‌های تصادفی قدرت پیشنهادی توانائی قابل قبولی در شبیه‌سازی‌های شبکه قدرت دارند. این قابلیت در مسأله جانمایی واحدهای اندازه‌گیری فاز به خوبی نشان داده شده است.
- ### مراجع
- [۱] M. Shahraeini and M. H. Javidi, *Wide Area Measurement Systems*. INTECH Open Access Publisher, 2012.
 - [۲] D. Brueni and L. Heath, "The PMU Placement Problem", *SIAM Journal on Discrete Mathematics*, vol. 19, no. 3, pp. 744-761, 2005.
 - [۳] R. Kumar, M. Giesselmann, S. Bayne and M. He, "GENERALIZED METHOD FOR FORMULATION OF OPTIMAL PMU PLACEMENT PROBLEM," *International Journal of Power and Energy Systems*, vol. 35, no. 1, 2015.
 - [۴] B. Gou, "Generalized integer linear programming formulation for optimal PMU placement," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1099-1104, Aug. 2008
 - [۵] C. Peng and X. Xu, "A hybrid algorithm based on BPSO and immune mechanism for PMU optimization placement," in *2008 World Cong.on Intelligent Control and Automation*, pp.7036-7040.
 - [۶] B. Gou, "Optimal placement of PMUs by integer linear programming," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1525-1526, Sep. 2008
 - [۷] F. J. Marín, F. García-Lagos, G. Joya, and F. Sandoval, "Optimal phasor measurement unit placement using genetic algorithms," *Computational Methods in Neural Modeling*, vol. 2686, pp. 486-493, 2003
 - [۸] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Trans. Power Del.*, vol. 20, no. 4, pp. 2381-2388, Oct. 2005
 - [۹] M. Shahraeini and M. H. Javidi, "A Survey on Power System Topological Observability," *IEEE Power Engineering and Automation Conference (PEAM 2011)*, Wuhan, China, Sep. 8-9, 2011.
 - [۱۰] G. Pagani and M. Aiello, "The Power Grid as a complex network: A survey", *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 11, pp. 2688-2700, 2013.
 - [۱۱] Z. Wang, A. Scaglione and R. Thomas, "Electrical Centrality Measures for Power Grids", in *Control and Optimization Methods for Electric Smart Grids*, Springer New York, pp. 239-255, 2012.
 - [۱۲] E. Zio and L. Golea, "Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements", *Reliability Engineering & System Safety*, vol. 101, pp. 67-74, 2012.
 - [۱۳] Zhifang Wang, A. Scaglione and R. Thomas, "Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks", *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 28-39, 2010.
 - [۱۴] Z. Wang, R. Thomas and A. Scaglione, "Generating Random Topology Power Grids", *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, 2008.
 - [۱۵] "System test cases archive" [Online]. Available: <http://www.ee.washington.edu/research/pstca>
 - [۱۶] Bollobás, Béla (2001). *Random Graphs* (2nd ed.). Cambridge University Press.
 - [۱۷] Frieze, Alan; Karonski, Michal (2015). *Introduction to Random Graphs*. Cambridge University Press.
 - [۱۸] Béla Bollobás, *Random Graphs*, 1985, Academic Press Inc., London Ltd.
 - [۱۹] Rick Durrett, *Random Graph Dynamics* (Cambridge Series in Statistical and Probabilistic Mathematics), Cambridge University Press, New York, NY, 2006
 - [۲۰] D. B. West. *Introduction to Graph Theory* (2nd Edition). Edited by Prentice Hall. Prentice Hall, 2001.
 - [۲۱] D. J. Watts, S. H. Strogatz (1998). "Collective dynamics of 'small-world' networks". *Nature* 393 (6684): 440-442.
 - [۲۲] Rick Durrett, *Random Graph Dynamics* (Cambridge Series in Statistical and Probabilistic Mathematics), Cambridge University Press, New York, NY, 2006.
 - [۲۳] D. B. West. *Introduction to Graph Theory* (2nd Edition). Edited by Prentice Hall. Prentice Hall, 2001
 - [۲۴] D. J. Watts, S. H. Strogatz (1998). "Collective dynamics of 'small-world' networks". *Nature* 393 (6684): 440-442
 - [۲۵] A. Abur and A. Expósito, *Power System State Estimation*. Hoboken: Taylor and Francis, 2004.



بهبود امنیت لایه فیزیکی شبکه مشارکتی با استفاده از انتخاب رله‌ی تقویت -و- ارسال در حضور چندین شنودگر

سحر پرکوک^۱، محمد ترابی^۲، سعید شکرالهی^۳

^۱ دانشجوی کارشناسی ارشد مهندسی برق، دانشگاه شهید بهشتی، تهران،

s.parkook@mail.sbu.ac.ir

^۲ عضو هیئت علمی دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران

m_torabi@sbu.ac.ir

^۳ عضو هیئت علمی پژوهشکده فضای مجازی، دانشگاه شهید بهشتی، تهران

s_shokrollahi@sbu.ac.ir

چکیده

یکی از چالش‌های مهم شبکه‌های بی‌سیم، آسیب‌پذیری در برابر حملات شنودگر است. از راهکارهای مطرح شده برای کاهش این آسیب‌پذیری برقراری امنیت در لایه فیزیکی با استفاده از روش ارتباط مشارکتی است. در این مقاله، بهبود امنیت و محرمانگی شبکه مشارکتی با چندین رله‌ی تقویت -و- ارسال با استفاده از روش انتخاب رله‌ی جزئی بر روی کانال‌های محوشدگی رایلی در حضور چندین شنودگر مورد بررسی قرار می‌گیرد. بدین منظور با استفاده از روابط ریاضی برای معیارهای ارزیابی محرمانگی شبکه‌های مخابراتی نظیر ظرفیت محرمانگی غیر صفر، احتمال قطع محرمانگی و میانگین ظرفیت محرمانگی که در این مقاله بدست آورده می‌شود، به تحلیل و ارزیابی کارایی امنیت ارسال و دریافت اطلاعات در شبکه مشارکتی با انتخاب رله پرداخته می‌شود. صحت روابط بدست آمده با استفاده از شبیه‌سازی مونت- کارلو مورد راستی‌آزمایی قرار می‌گیرد.

کلمات کلیدی

امنیت لایه فیزیکی، ظرفیت محرمانگی، انتخاب رله، شبکه مشارکتی

و روش‌های چندگانگی (انتخاب رله) و سیستم‌های چند ورودی چند خروجی (MIMO) و مدنظر قرار گرفته است [1-6].

مبنای تئوری امنیت لایه فیزیکی، تئوری اطلاعاتی است که بر مبنای مفاهیم امنیت کامل شانون است [7]. امنیت لایه فیزیکی ابتدا با بررسی کانال‌های شنودگر بدون حافظه توسط واینر مطرح شد [8]. در کانال شنودگری که توسط واینر پیشنهاد شده دو کاربر مجاز از طریق یک کانال اصلی در ارتباط هستند و یک شنودگر به نسخه کاهش یافته‌ی خروجی کانالی که به گیرنده مجاز می‌رسد دسترسی دارد. اثبات شده‌است که امنیت کامل داده‌های ارسالی در صورتی به دست می‌آید که ظرفیت کانال اصلی (از منبع به مقصد) از کانال شنودگر (از منبع به شنودگر) بزرگتر باشد. سپس نتایج تحقیقات واینر از کانال شنودگر بدون حافظه به کانال شنودگر گوسی گسترش یافت و مفهوم ظرفیت محرمانگی مطرح گردید [9]. ظرفیت محرمانگی (حداکثر نرخ ارسال که شنودگر قادر به دی‌کد اطلاعات نمی‌باشد) برابر اختلاف بین ظرفیت کانال اصلی و کانال شنودگر است. در نتیجه ارتباط امن امکان

۱- مقدمه

در شبکه‌های بی‌سیم، ارسال بین کاربران مجاز به راحتی به دلیل ماهیت پخش مودم بی‌سیم به وسیله یک شنودگر شنود می‌شود و ارسال بی‌سیم در برابر حملات شنودگر به شدت آسیب‌پذیر است. به منظور افزایش امنیت ارسال و دریافت، سیستم‌های ارتباطی موجود معمولاً از تکنیک‌های رمزنگاری برای مانع شدن از حملات یک شنودگر به داده‌ی ارسالی بین کاربران مجاز استفاده می‌کنند. این روش‌ها بر پایه‌ی کلید محرمانه بوده و ممکن است متن رمز شده با یک جست و جوی کامل کلید رمز گشایی شود.

امنیت لایه فیزیکی به عنوان یک روش جایگزین و مکمل برای ایجاد امنیت در ارتباطات بی‌سیم در برابر حملات شنودگر با استفاده از خواص فیزیکی کانال بی‌سیم، ارائه شده‌است. در کارهای تئوری و عملی انجام شده نشان داده شده‌است که برای بهبود امنیت سیستم‌های ارتباطی دیجیتال، ایده‌های امنیت لایه فیزیکی نظیر مدولاسیون، کدینگ

ارتباط دارد. کانال اصلی متشکل از کانال‌های $S \rightarrow R_i \rightarrow D$ و $S \rightarrow D$ است.

$$\gamma_{SR_i} = |h_{SR_i}|^2 P / N_M \quad , \quad \gamma_{SD} = |h_{SD}|^2 P / N_M$$

و $\gamma_{R_i D} = |h_{R_i D}|^2 P / N_M$ به ترتیب SNR مربوط به کانال‌های $S \rightarrow R_i$ ، $S \rightarrow D$ و $R_i \rightarrow D$ است و h_{SD} و h_{SR_i} و $h_{R_i D}$ به ترتیب ضریب محوشدگی کانال‌ها می‌باشند و P میانگین توان سیگنال ارسالی می‌باشد.

الگوریتم انتخاب رله، بهترین رله b را انتخاب می‌کند به طوری که:

$$b = \arg \max_{i \in \mathcal{R}} \{\gamma_{R_i D}\} \quad (1)$$

که $\mathcal{R} = \{1, \dots, N\}$ است. می‌توان نشان داد که SNR کانال اصلی برابر است با:

$$\gamma_M = \gamma_{SD} + \frac{\gamma_{SR_b} \gamma_{R_b D}}{1 + \gamma_{SR_b} + \gamma_{R_b D}} \quad (2)$$

شنودگر k ام (E_k) هم می‌تواند سیگنال ارسالی از منبع و رله را شنود کند. یک سیستم ارتباط مشارکتی متشکل از دو کانال $S \rightarrow E_k$ و $S \rightarrow R_i \rightarrow E_k$ است. γ_{SE_k} و $\gamma_{R_i E_k}$ به ترتیب SNR مربوط به کانال‌های $S \rightarrow E_k$ و $R_i \rightarrow E_k$ می‌باشند.

مشابه رابطه (۲)، SNR در شنودگر k ام برابر است با:

$$\gamma_{W_k} = \gamma_{SE_k} + \frac{\gamma_{SR_i} \gamma_{R_i E_k}}{1 + \gamma_{SR_i} + \gamma_{R_i E_k}} \quad (3)$$

و $\gamma_{R_i E_k} = |h_{R_i E_k}|^2 P / N_W$ و $\gamma_{SE_k} = |h_{SE_k}|^2 P / N_W$ است و h_{SE_k} و $h_{R_i E_k}$ به ترتیب ضریب محوشدگی کانال‌های $S \rightarrow E_k$ و $R_i \rightarrow E_k$ هستند.

مشابه آنچه برای سیستم ارسال مستقیم (بدون مشارکتی) بیان شده است، برای سیستم مشارکتی، SNR کانال شنودگر برای چندین شنودگر برابر است با [11]:

$$\gamma_W = \max_{k \in \mathcal{E}} \{\gamma_{W_k}\} \quad (4)$$

که $\mathcal{E} = \{1, \dots, K\}$ است.

۲-۱- تابع چگالی احتمال SNRهای مربوط به کانال‌های اصلی و شنودگر

با فرض $\gamma_{SR_b D} \square \frac{\gamma_{SR_b} \gamma_{R_b D}}{1 + \gamma_{SR_b} + \gamma_{R_b D}}$ ، تابع چگالی احتمال SNR کلی $\gamma_M = \gamma_{SD} + \gamma_{SR_b D}$ برابر است با [9]:

$$f_{\gamma_M}(x) = f_{\gamma_{SD}}(x) * f_{\gamma_{SR_b D}}(x) \quad (5)$$

از آنجاییکه γ_{SD} یک متغیر تصادفی رابلی است با فرض $\bar{\gamma}_{SD} = \bar{\gamma}$ ، تابع چگالی احتمال آن برابر است با:

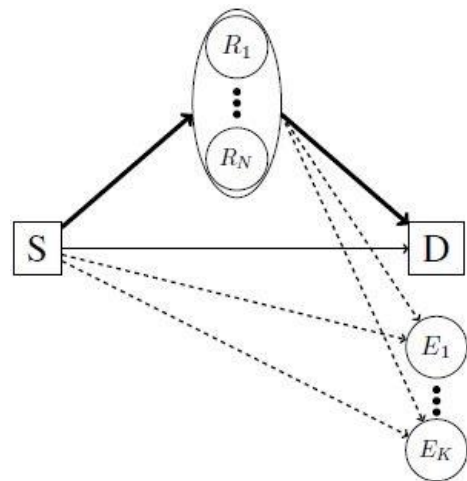
$$f_{\gamma_{SD}}(x) = \frac{1}{\bar{\gamma}} \exp\left(-\frac{x}{\bar{\gamma}}\right) \quad (6)$$

پذیر نیست مگر اینکه کانال اصلی نسبت سیگنال به نویز (SNR) بهتری نسبت به کانال شنودگر داشته باشد. در مراجع [10,11] به بررسی و تحلیل کارایی امنیت سیستم ارتباطی مستقیم پرداخته شده است، که در آن یک فرستنده اطلاعات را به گیرنده مورد نظر ارسال می‌کند در حالی که یک یا چندین شنودگر قابلیت دریافت اطلاعات ارسالی را دارند. از طرف دیگر نشان داده شده است [12,13]، شبکه‌های مشارکتی با رله‌ی انتخابی تقویت -و- ارسال عملکرد بهتری نسبت به شبکه‌های مخابراتی ارسال مستقیم (بدون مشارکتی) دارند.

در این مقاله، به بررسی و بهبود امنیت و محرمانگی شبکه مشارکتی با چندین رله‌ی تقویت -و- ارسال و بهبود امنیت لایه فیزیکی با استفاده از روش انتخاب رله‌ی جزئی بر روی کانال‌های محوشدگی رایلی در حضور چندین شنودگر می‌پردازیم. روابط ریاضی برای محاسبه احتمال ظرفیت محرمانگی غیر صفر، احتمال قطع محرمانگی و میانگین ظرفیت محرمانگی را بدست می‌آوریم. با استفاده از روابط بدست آمده به ارزیابی کارایی و امنیت ارسال اطلاعات در شبکه مشارکتی می‌پردازیم. همچنین با استفاده از شبیه‌سازی مونت-کارلو صحت روابط بدست آمده را راستی‌آزمایی می‌کنیم.

ساختار مطالب این مقاله به قرار زیر است. در بخش ۲، به توضیح مدل شبکه مشارکتی پرداخته می‌شود. در بخش ۳، ظرفیت محرمانگی شبکه مشارکتی با انتخاب رله‌ی تقویت -و- ارسال بررسی شده و روابط ریاضی احتمال ظرفیت محرمانگی غیر صفر، احتمال قطع محرمانگی و میانگین ظرفیت محرمانگی بدست می‌آید. در بخش ۴، با استفاده از نتایج عددی بدست آمده از روابط ریاضی و شبیه‌سازی مونت-کارلو، به بررسی کارایی سیستم می‌پردازیم و در نهایت در بخش ۵، نتیجه‌گیری مقاله ارائه می‌شود.

۲- مدل سیستم



شکل ۱: مدل سیستم

شکل (۱) یک سیستم ارتباطی مشارکتی را نشان می‌دهد که فرستنده منبع (S) در حضور چندین شنودگر E_1 و E_2 و ... و E_K با گیرنده مجاز در مقصد (D) از طریق رله‌های تقویت و ارسال R_1 و R_2 و ... و R_N

۳- ظرفیت محرمانگی

اگر فرستنده هیچ اطلاعاتی درباره‌ی کانال‌ها نداشته باشد، استراتژی ارسال پیام با سطح توان محدود توصیه می‌شود. بنابراین ظرفیت کانال اصلی برابر است با:

$$C_M = \frac{1}{2} \log_2(1 + \gamma_M) \quad (16)$$

که ضریب 1/2 به این دلیل است که ارسال در دو مرحله انجام می‌شود [10]. به همین ترتیب، ظرفیت کانال شنودگر برابر است با:

$$C_W = \frac{1}{2} \log_2(1 + \gamma_W) \quad (17)$$

با توجه به مدل واینر، ظرفیت محرمانگی برابر است با:

$$C_s = C_M - C_W \quad (18)$$

در حالتیکه $\gamma_M \leq \gamma_W$ باشد، پیام ارسال نمی‌شود مگر اینکه حفاظت اطلاعات در برابر شنودگر مهم نباشد. بنابراین ظرفیت محرمانگی به صورت زیر بیان می‌شود:

$$C_s = \begin{cases} \frac{1}{2} [\log_2(1 + \gamma_M) - \log_2(1 + \gamma_W)] & \gamma_M > \gamma_W \\ 0 & \gamma_M \leq \gamma_W \end{cases} \quad (19)$$

۳-۱- احتمال ظرفیت محرمانگی غیر صفر

با توجه به رابطه (۱۹) در حالتی که $\gamma_M > \gamma_W$ باشد، ظرفیت محرمانگی مثبت خواهد شد و در حالتی که $\gamma_M \leq \gamma_W$ باشد، ظرفیت محرمانگی صفر می‌شود. بنابراین احتمال غیر صفر بودن ظرفیت محرمانگی به صورت زیر تعریف می‌شود:

$$\Pr(C_s > 0) = \Pr(\gamma_M > \gamma_W) = \Pr(x > x') \\ = \int_{x=0}^{\infty} \int_{x'=0}^x f_{\gamma_M, \gamma_W}(x, x') dx dx' \quad (20)$$

با فرض اینکه γ_M و γ_W غیر همبسته هستند (همبستگی آن‌ها قابل اغماض است) می‌توان نوشت:

$$\Pr(C_s > 0) = \int_{x=0}^{\infty} \int_{x'=0}^x f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' \quad (21)$$

با جایگذاری روابط (۱۲) و (۱۵) در رابطه‌ی فوق و با حل انتگرال‌های فوق رابطه‌ی زیر را بدست می‌آوریم:

$$\Pr(C_s > 0) = \frac{1}{\bar{\gamma}} \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{i} \sum_{j=0}^K \binom{K}{j} \left(\frac{1}{\bar{\gamma}_C - \bar{\gamma}'} \right)^j \sum_{m=0}^j \binom{j}{m} \\ \times \bar{\gamma}_C^{j-m} (\bar{\gamma}')^m \left(\frac{1}{\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}} + \frac{1}{\bar{\gamma}}} - \frac{1}{\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}} + \frac{j+1}{\bar{\gamma}}} \right) \quad (22)$$

۳-۲- احتمال قطع محرمانگی

یکی از معیارهای ارزیابی کارایی امنیت لایه فیزیکی احتمال قطع محرمانگی ($P_{out}(R_s)$) است. احتمال قطع محرمانگی احتمالی است که ظرفیت محرمانگی از حد آستانه‌ی R_s کمتر باشد. بنابراین:

$$P_{out}(R_s) = \Pr(C_s < R_s) = \Pr(C_s < R_s | \gamma_M > \gamma_W) \\ \times \Pr(\gamma_M > \gamma_W) + \Pr(C_s < R_s | \gamma_M \leq \gamma_W) \Pr(\gamma_M \leq \gamma_W)$$

برای سهولت در تحلیل ریاضی، می‌توان از تعریف

$$\gamma_{SR_b D} = \frac{\gamma_{SR_b} \gamma_{R_b D}}{1 + \gamma_{SR_b} + \gamma_{R_b D}} \square \min(\gamma_{R_b D}, \gamma_{SR_b D})$$

بنابراین:

$$F_{\gamma_{SR_b D}} = 1 - (1 - F_{\gamma_{SR_b}})(1 - F_{\gamma_{R_b D}}) \quad (7)$$

با فرض $\gamma_{SD} = \gamma_{SR_i} = \gamma_{R_i D} = \bar{\gamma}$ داریم:

$$F_{\gamma_{SR_b}} = 1 - \exp\left(-\frac{x}{\bar{\gamma}}\right) \quad (8)$$

$$F_{\gamma_{R_b D}} = \left(1 - \exp\left(-\frac{x}{\bar{\gamma}}\right)\right)^N \quad (9)$$

بنابراین:

$$F_{\gamma_{SR_b D}} = \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{\bar{\gamma}} \exp\left(-\frac{x}{\bar{\gamma}}(i+1)\right) \quad (10)$$

برای تعیین تابع چگالی احتمال از رابطه‌ی (۵) (مشابه [12])،

تبدیل لاپلاس می‌گیریم:

$$MGF_{\gamma_M}(s) = MGF_{\gamma_{SD}}(s) MGF_{\gamma_{SR_b D}}(s) \quad (11)$$

که

$$MGF_{\gamma_{SD}}(s) = \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{1}{1 + \frac{\bar{\gamma}}{s}} \frac{1}{i+1} \quad \text{و} \quad MGF_{\gamma_{SR_b D}}(s) = \frac{1}{1 + s\bar{\gamma}}$$

است و با عکس تبدیل لاپلاس از رابطه‌ی (۱۱) تابع چگالی احتمال کانال اصلی برابر خواهد شد با:

$$f_{\gamma_M} = \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{\bar{\gamma}} \left(\exp\left(-\frac{x}{\bar{\gamma}}\right) - \exp\left(-\frac{x}{\bar{\gamma}}(i+1)\right) \right) \quad (12)$$

به همین روش تابع چگالی احتمال γ_W برای یک شنودگر برابر خواهد شد با:

$$f_{\gamma_W} = \frac{1}{\bar{\gamma}_C - \bar{\gamma}'} \left(\exp\left(-\frac{x'}{\bar{\gamma}_C}\right) - \exp\left(-\frac{x'}{\bar{\gamma}'}\right) \right) \quad (13)$$

با توجه به رابطه (۴) تابع توزیع احتمال کانال شنودگر به صورت

زیر خواهد بود:

$$F_{\gamma_W} = \left(1 - \frac{1}{\bar{\gamma}_C - \bar{\gamma}'} \left(\bar{\gamma}_C \exp\left(-\frac{x'}{\bar{\gamma}_C}\right) - \bar{\gamma}' \exp\left(-\frac{x'}{\bar{\gamma}'}\right) \right) \right)^K \quad (14)$$

که $\bar{\gamma}_C = \frac{\bar{\gamma}_{SR_i} \bar{\gamma}_{R_i E_k}}{\bar{\gamma}_{SR_i} + \bar{\gamma}_{R_i E_k}}$ و $\bar{\gamma}' = \bar{\gamma}_{SE_k} = \bar{\gamma}_{R_i E_k}$ با استفاده از

مشق گیری از رابطه (۱۴)، رابطه‌ی تابع چگالی احتمال را به صورت زیر بدست می‌آوریم:

$$f_{\gamma_W}(x') = \frac{N}{\bar{\gamma}_C - \bar{\gamma}'} \left(\exp\left(-\frac{x'}{\bar{\gamma}_C}\right) - \exp\left(-\frac{x'}{\bar{\gamma}'}\right) \right) \\ \times \left(1 - \frac{1}{\bar{\gamma}_C - \bar{\gamma}'} \left(\bar{\gamma}_C \exp\left(-\frac{x'}{\bar{\gamma}_C}\right) - \bar{\gamma}' \exp\left(-\frac{x'}{\bar{\gamma}'}\right) \right) \right)^{K-1} \quad (15)$$

(۲۳)

با استفاده از $\int_0^\infty \ln(1+x) \exp(-\mu x) dx = \frac{1}{\mu} \exp(\mu) E_1(\mu)$ و

روابط (۱۲) و (۱۵) خواهیم داشت:

$$I_1 = \int_{x=0}^\infty \int_{x'=0}^x \log_2(1+x) f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' \quad (29)$$

با جایگذاری روابط (۱۲) و (۱۵) و با حل انتگرال مربوط به γ_W ،

خواهیم داشت:

$$I_1 = \frac{1}{\bar{\gamma}} \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{i} \sum_{j=0}^K \binom{K}{j} \left(\frac{1}{\bar{\gamma}_C - \bar{\gamma}} \right)^j \\ \times \sum_{m=0}^j \binom{j}{m} \bar{\gamma}_C^{j-m} (\bar{\gamma}')^m \int_{x=0}^\infty \log_2(1+x) \\ \times \left(\exp \left(- \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}} \right) x \right) \right. \\ \left. - \exp \left(- \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}} \right) x \right) \right) dx \quad (30)$$

با حل انتگرال فوق رابطه زیر را بدست می‌آوریم:

$$I_1 = \frac{1}{\ln(2)\bar{\gamma}} \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{i} \sum_{j=0}^K \binom{K}{j} \left(\frac{1}{\bar{\gamma}_C - \bar{\gamma}} \right)^j \sum_{m=0}^j \binom{j}{m} \\ \times \bar{\gamma}_C^{j-m} (\bar{\gamma}')^m \left[\frac{1}{\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}}} \exp \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}} \right) \right. \\ \times E_1 \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}} \right) - \frac{1}{\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}}} \\ \left. \times \exp \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}} \right) E_1 \left(\frac{j-m}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}} \right) \right] \quad (31)$$

و همچنین بطور مشابه خواهیم داشت:

$$I_2 = \int_{x'=0}^\infty \int_{x=x'}^\infty \log_2(1+x') f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' \quad (32)$$

با جایگذاری روابط (۱۲) و (۱۵) و با حل انتگرال مربوط به γ_M ،

خواهیم داشت:

$$I_2 = \frac{K}{\bar{\gamma}_C - \bar{\gamma}} \sum_{n=1}^N \binom{N}{n} (-1)^{n+1} \frac{n+1}{n} \sum_{u=0}^{K-1} \binom{K-1}{u} \left(\frac{1}{\bar{\gamma}_C - \bar{\gamma}} \right)^u \\ \times \sum_{s=0}^u \binom{u}{s} \bar{\gamma}_C^{u-s} (\bar{\gamma}')^s \int_{x'=0}^\infty \log_2(1+x') \\ \times \left(\exp \left(- \frac{x'}{\bar{\gamma}} \right) - \frac{1}{n+1} \exp \left(- \frac{n+1}{\bar{\gamma}} x' \right) \right) \left(\exp \left(- \frac{x'}{\bar{\gamma}_C} \right) \right. \\ \left. - \exp \left(- \frac{x'}{\bar{\gamma}'} \right) \right) \exp \left(- \left(\frac{u-s}{\bar{\gamma}_C} + \frac{s}{\bar{\gamma}'} \right) x' \right) dx' \quad (33)$$

با حل انتگرال فوق رابطه زیر را بدست می‌آوریم:

واضح است که $\Pr(C_s < R_s | \gamma_M \leq \gamma_W) = 1$ چون $R_s > 0$ با

استفاده از رابطه (۱۹)، رابطه (۲۳) به صورت زیر در می‌آید:

$$P_{out} = \Pr \left(\frac{1}{2} \log_2 \left(\frac{1+\gamma_M}{1+\gamma_W} \right) < R_s | \gamma_M > \gamma_W \right) \Pr(\gamma_M > \gamma_W) \\ + 1 - \Pr(\gamma_M > \gamma_W) = \Pr(x < 2^{2R_s} (1+x') - 1 | x > x') \\ \times \Pr(x > x') + 1 - \Pr(x > x') \quad (24)$$

با فرض $\alpha = 1 - \Pr(x < 2^{2R_s} (1+x') - 1 | x > x')$ و γ_M و γ_W

همبستگی آن‌ها قابل اغماض است) آنگاه خواهیم داشت:

$$P_{out}(R_s) = 1 - \Pr(x > \alpha) = 1 - \int_{x'=0}^\infty \int_{x=\alpha}^\infty f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' \quad (25)$$

با جایگذاری روابط (۱۲) و (۱۵) در رابطه‌ی فوق و با حل

انتگرال‌های فوق رابطه احتمال قطع محرمانگی را به صورت زیر بدست

می‌آوریم:

$$P_{out}(R_s) = 1 - \frac{K}{\bar{\gamma}_C - \bar{\gamma}} \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{i} \sum_{j=0}^{K-1} \binom{K-1}{j} \\ \times \left(\frac{1}{\bar{\gamma}_C - \bar{\gamma}} \right)^j \sum_{m=0}^j \binom{j}{m} \bar{\gamma}_C^{j-m} (\bar{\gamma}')^m \left(\exp \left(- \frac{2^{2R_s} - 1}{\bar{\gamma}} \right) \right. \\ \times \left[\frac{1}{\frac{2^{2R_s}}{\bar{\gamma}} + \frac{j-m+1}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'}} - \frac{1}{\frac{2^{2R_s}}{\bar{\gamma}} + \frac{j-m}{\bar{\gamma}_C} + \frac{m+1}{\bar{\gamma}'}} \right] \\ \left. - \frac{1}{i+1} \exp \left(- \frac{(i+1)(2^{2R_s} - 1)}{\bar{\gamma}} \right) \right) \\ \times \left[\frac{1}{\frac{2^{2R_s}}{\bar{\gamma}} + \frac{j-m+1}{\bar{\gamma}_C} + \frac{m}{\bar{\gamma}'}} - \frac{1}{\frac{2^{2R_s}}{\bar{\gamma}} + \frac{j-m}{\bar{\gamma}_C} + \frac{m+1}{\bar{\gamma}'}} \right] \quad (26)$$

۳-۳- میانگین ظرفیت محرمانگی

با استفاده از میانگین ظرفیت محرمانگی (b/s/Hz) هم می‌توان امنیت

لایه فیزیکی را تحلیل کرد. میانگین ظرفیت محرمانگی در واقع، با

استفاده از متوسط‌گیری از روی مقادیر لحظه‌ای و با فرض اینکه γ_M و

γ_W غیر همبسته هستند (همبستگی آن‌ها قابل اغماض است)، ظرفیت

محرمانگی به صورت زیر بدست می‌آید:

$$\bar{C}_s = E(C_s) = \int_{-\infty}^{+\infty} C_s f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' \quad (27)$$

با جایگذاری رابطه (۱۹) در رابطه‌ی فوق خواهیم داشت:

$$\bar{C}_s = \frac{1}{2} \int_{x=0}^\infty \int_{x'=0}^x (\log_2(1+x) - \log_2(1+x')) \\ \times f_{\gamma_M}(x) f_{\gamma_W}(x') dx dx' = \frac{1}{2} (I_1 - I_2) \quad (28)$$

$$\begin{aligned}
\bar{C}_s &= \frac{0.5}{\ln(2)\bar{\gamma}} \sum_{i=1}^N \binom{N}{i} (-1)^{i+1} \frac{i+1}{i} \sum_{j=0}^K \binom{K}{j} \left(\frac{1}{\bar{\gamma}_c - \bar{\gamma}'} \right)^j \\
&\times \sum_{m=0}^j \binom{j}{m} \bar{\gamma}_c^{j-m} (\bar{\gamma}')^m \left(\frac{1}{\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}}} \right) \\
&\times \exp\left(\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}} \right) E_1\left(\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{1}{\bar{\gamma}} \right) \\
&- \frac{1}{\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}}} \exp\left(\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}} \right) \\
&\times E_1\left(\frac{j-m}{\bar{\gamma}_c} + \frac{m}{\bar{\gamma}'} + \frac{i+1}{\bar{\gamma}} \right) - \frac{0.5K}{\ln(2)(\bar{\gamma}_c - \bar{\gamma}')} \sum_{n=1}^N \binom{N}{n} \\
&\times (-1)^{n+1} \frac{n+1}{n} \sum_{u=0}^{K-1} \binom{K-1}{u} \left(\frac{1}{\bar{\gamma}_c - \bar{\gamma}'} \right)^u \sum_{s=0}^u \binom{u}{s} \bar{\gamma}_c^{u-s} (\bar{\gamma}')^s \\
&\times \left(\frac{1}{\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}}} \exp\left(\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) \right) \\
&\times E_1\left(\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) - \frac{1}{\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}}} \\
&\times \exp\left(\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) E_1\left(\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) \\
&- \frac{1}{n+1} \left(\frac{1}{\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}}} \exp\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) \right) \\
&\times E_1\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) - \frac{1}{\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}}} \\
&\times \exp\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) E_1\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) \Big) \\
\end{aligned} \tag{۳۵}$$

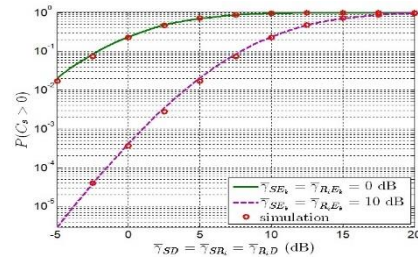
$$\begin{aligned}
I_2 &= \frac{K}{\ln(2)(\bar{\gamma}_c - \bar{\gamma}')} \sum_{n=1}^N \binom{N}{n} (-1)^{n+1} \frac{n+1}{n} \sum_{u=0}^{K-1} \binom{K-1}{u} \left(\frac{1}{\bar{\gamma}_c - \bar{\gamma}'} \right)^u \\
&\times \sum_{s=0}^u \binom{u}{s} \bar{\gamma}_c^{u-s} (\bar{\gamma}')^s \left(\frac{1}{\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}}} \exp\left(\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) \right) \\
&\times E_1\left(\frac{1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) \frac{1}{\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}}} \\
&\times \exp\left(\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) E_1\left(\frac{1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) \\
&- \frac{1}{n+1} \left(\frac{1}{\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}}} \exp\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) \right) \\
&\times E_1\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s+1}{\bar{\gamma}_c} + \frac{s}{\bar{\gamma}} \right) - \frac{1}{\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}}} \\
&\times \exp\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) E_1\left(\frac{n+1}{\bar{\gamma}} + \frac{u-s}{\bar{\gamma}_c} + \frac{s+1}{\bar{\gamma}} \right) \Big) \\
\end{aligned} \tag{۳۴}$$

با جایگذاری روابط I_1 و I_2 در رابطه ۲۷، میانگین ظرفیت محرمانگی به صورت زیر بدست می‌آید:

۴- نتایج عددی و شبیه‌سازی

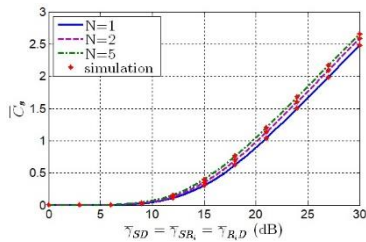
در این بخش، با استفاده از دو نتایج عددی حاصل از روابط (۲۲)، (۲۶) و (۳۵) و همچنین نتایج حاصل از شبیه‌سازی مونت-کارلو، نتایج عددی و شبیه‌سازی احتمال ظرفیت محرمانگی غیر صفر، احتمال قطع محرمانگی و میانگین ظرفیت محرمانگی سیستم، به ازای تعداد N رله و K شنودگر مورد بررسی قرار می‌گیرد.

شکل ۵: احتمال قطع محرمانگی با $\bar{\gamma}_{SE_k} = \bar{\gamma}_{R_i E_k} = 10$ dB و $\bar{\gamma}_{SD}$ بر حسب $R_s = 0.1$



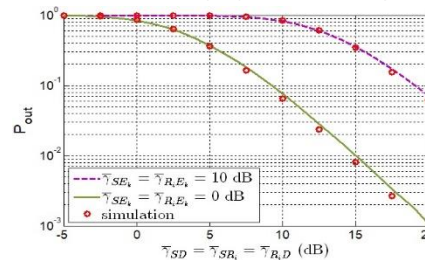
شکل ۲: احتمال ظرفیت محرمانگی غیر صفر با $N = K = 5$ بر حسب $\bar{\gamma}_{SD}$

در شکل (۵) نتایج مربوط به احتمال قطع محرمانگی با $\bar{\gamma}_{SE_k} = \bar{\gamma}_{R_i E_k} = 10$ dB و $R_s = 0.1$ بر حسب $\bar{\gamma}_{SD}$ آورده شده است. مشاهده می‌شود که با افزایش $\bar{\gamma}_{SD}$ احتمال قطع محرمانگی کاهش می‌یابد. همچنین با افزایش تعداد رله‌ها N احتمال قطع محرمانگی کاهش می‌یابد.



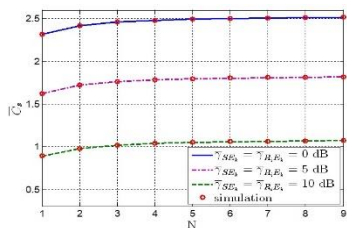
شکل ۶: میانگین ظرفیت محرمانگی با $\bar{\gamma}_{SE_k} = \bar{\gamma}_{R_i E_k} = 10$ dB بر حسب $\bar{\gamma}_{SD}$

در شکل (۲)، نتایج مربوط به احتمال ظرفیت محرمانگی غیر صفر برای تعداد رله و شنودگر $N = K = 5$ بر حسب $\bar{\gamma}_{SD}$ آورده شده است. مشاهده می‌شود که با افزایش $\bar{\gamma}_{SD}$ احتمال ظرفیت محرمانگی غیر صفر افزایش می‌یابد. همچنین با افزایش $\bar{\gamma}_{SE_k}$ احتمال ظرفیت محرمانگی غیر صفر کاهش می‌یابد.



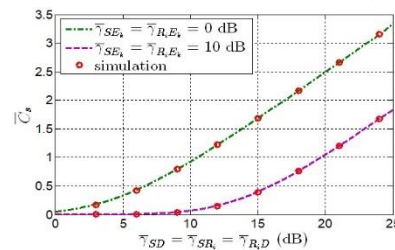
شکل ۳: احتمال قطع محرمانگی با $N = K = 5$ و $R_s = 0.1$ بر حسب $\bar{\gamma}_{SD}$

در شکل (۶) نتایج مربوط به میانگین ظرفیت محرمانگی با $\bar{\gamma}_{SE_k} = \bar{\gamma}_{R_i E_k} = 10$ dB بر حسب $\bar{\gamma}_{SD}$ آورده شده است. مشاهده می‌شود که با افزایش $\bar{\gamma}_{SD}$ میانگین ظرفیت محرمانگی افزایش می‌یابد. همچنین با افزایش تعداد رله‌ها N میانگین ظرفیت محرمانگی افزایش می‌یابد.



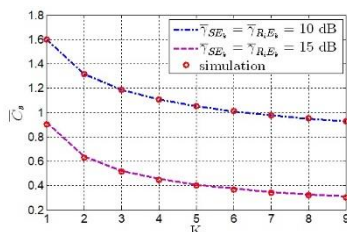
شکل ۷: میانگین ظرفیت محرمانگی با $\bar{\gamma}_{SD} = \bar{\gamma}_{SR_i} = \bar{\gamma}_{R_i D} = 20$ dB بر حسب تعداد رله‌ها N

در شکل (۳) نتایج مربوط به احتمال قطع محرمانگی با $N = K = 5$ و $R_s = 0.1$ بر حسب $\bar{\gamma}_{SD}$ آورده شده است. مشاهده می‌شود که با افزایش $\bar{\gamma}_{SD}$ احتمال قطع محرمانگی کاهش می‌یابد. همچنین با افزایش $\bar{\gamma}_{SE_k}$ احتمال قطع محرمانگی افزایش می‌یابد.



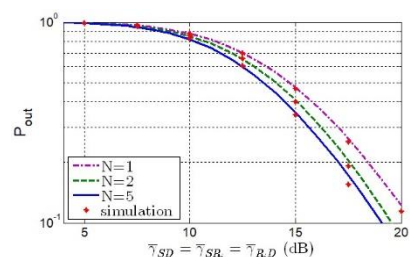
شکل ۴: میانگین ظرفیت محرمانگی با $N = K = 5$ بر حسب $\bar{\gamma}_{SD}$

در شکل (۷) نتایج مربوط به میانگین ظرفیت محرمانگی با $\bar{\gamma}_{SD} = \bar{\gamma}_{SR_i} = \bar{\gamma}_{R_i D} = 20$ dB بر حسب N آورده شده است. مشاهده می‌شود که با افزایش تعداد رله‌ها N میانگین ظرفیت محرمانگی افزایش می‌یابد. همچنین با افزایش $\bar{\gamma}_{SE_k}$ میانگین ظرفیت محرمانگی کاهش می‌یابد.



شکل ۸: میانگین ظرفیت محرمانگی با $\bar{\gamma}_{SD} = \bar{\gamma}_{SR_i} = \bar{\gamma}_{R_i D} = 20$ dB بر حسب تعداد شنودگرها K

در شکل (۴) نتایج مربوط به میانگین ظرفیت محرمانگی با $N = K = 5$ بر حسب $\bar{\gamma}_{SD}$ آورده شده است. مشاهده می‌شود که با افزایش $\bar{\gamma}_{SD}$ میانگین ظرفیت محرمانگی افزایش می‌یابد. همچنین با افزایش $\bar{\gamma}_{SE_k}$ میانگین ظرفیت محرمانگی کاهش می‌یابد.



در شکل (۸) نتایج مربوط به میانگین ظرفیت محرمانگی با مشاهده می‌شود که با افزایش تعداد شنودگرها K میانگین ظرفیت محرمانگی کاهش می‌یابد. همچنین با افزایش $\bar{\gamma}_{SE_k}$ میانگین ظرفیت محرمانگی کاهش می‌یابد.

در شکل‌های (۲)–(۸) با مقایسه‌ی نتایج عددی بدست آمده با شبیه‌سازی مشاهده می‌شود که با هم مطابقت دارند که نشانگر صحت روابط بدست آمده است.

۵- نتیجه‌گیری

در این مقاله، به بررسی و بهبود امنیت و محرمانگی شبکه مشارکتی با چندین رله‌ی تقویت -و- ارسال با استفاده از روش انتخاب رله‌ی جزئی بر روی کانال‌های محوشدگی رایلی در حضور چندین شنودگر پرداختیم. روابط ریاضی برای محاسبه احتمال ظرفیت محرمانگی غیر صفر، احتمال قطع محرمانگی و میانگین ظرفیت محرمانگی را بدست آوردیم. با استفاده از روابط بدست آمده به ارزیابی کارایی و امنیت ارسال اطلاعات در شبکه مشارکتی با رله‌ی انتخابی تقویت -و- ارسال پرداختیم. همچنین با استفاده از شبیه‌سازی مونت-کارلو صحت روابط ریاضی بدست آورده شده را بررسی کردیم.

مراجع

- [1] Y. Zou, J. Zhu, X. Wang, and V. Leung, "Improving diversity for physical-layer security in wireless communications," *IEEE Network*, vol. 29, no. 1, pp. 42–48, Jan. 2015.
- [2] Y. Zou, X. Wang, L. Hanzo. "A survey on wireless security: technical challenges, recent advances and future trends". *Proceedings of the IEEE*, vol.104, no. 9, Sept 2016.
- [3] B. He, Q. Ni, J. Chen, L. Yang, and L. Lv, "User-pair selection in multiuser cooperative networks with an untrusted relay," *IEEE Trans. Veh. Technol.*, pp. 1–12, 2018.
- [4] D. B. Rawat, T. White, M. S. Parwez, C. Bajracharya, and M. Song, "Evaluating secrecy outage of physical layer security in large-scale MIMO wireless communications for cyber-physical systems," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1987–1993, 2017.
- [5] M. K. Shukla, S. Member, S. Yadav, N. Purohit, "Secure transmission in cellular multiuser two-way amplify-and-forward relay networks," *IEEE Trans. Veh. Technol.*, pp. 1-10, 2018.
- [6] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secrecy cooperative networks with outdated relay selection over correlated fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7599–7603, Aug. 2017.
- [7] C. Shannon, "Communication theory of secrecy systems" *Bell Syst. Tech. Journal*, vol. 29, pp. 656–715, 1949.
- [8] D. Wyner, "The Wire-Tap Channel," *Bell Syst. Tech. Journal*, vol. 54, pp. 1355–1387, 1975.
- [9] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. on Inform. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.
- [10] J. Barros, M. Rodrigues, "Secrecy capacity of wireless channels" *IEEE Intl. Sym. Info. Theory*, pp.356–60, 2006.
- [11] P. Wang, G. Yu, Z. Zhang, "On the secrecy capacity of fading wireless channel with multiple eavesdroppers," *IEEE Int. Symp. Inf. Theory*, June 2007, pp. 1301–1305.
- [12] M. Torabi, W. Ajib and D. Haccoun, "Performance analysis of amplify-and-forward cooperative networks with relay selection over rayleigh fading channels," *IEEE Vehicular Tech. Conference*, April 2009, pp. 1-6.
- [13] M. Torabi and D. Haccoun, "Capacity of amplify-and-forward selective relaying with adaptive transmission under outdated channel information," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2416–2422, June 2011.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



۲۷ الی ۲۸ آذرماه ۱۳۹۸

پیش‌بینی دمای اتاق با استفاده از یادگیری تقویتی

زهرا مرتضایی^۱، مهسا قاسمی^۲، منصور فاتح^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود
zm.mortezaie@gmail.com

^۲ دانشجوی کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود
mahsa.ghasemi@shahroodut.ac.ir

^۳ استادیار، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود
mansoor_fateh@shahroodut.ac.ir

چکیده

عوامل بسیاری مانند دما، روشنایی مناسب و هوای مطبوع در آسایش ساکنین یک ساختمان تأثیرگذارند. در ساختمان‌های معمولی، این عوامل توسط ساکنین تنظیم می‌شود. برای افزایش سطح آسایش ساکنان، می‌توان سیستمی برای تنظیم خودکار میزان مطلوبیت، با استفاده از قابلیت یادگیری و پیش‌بینی عوامل مؤثر، طراحی کرد. این ایده منجر به پیاده‌سازی ساختمان‌های هوشمند شده است. هدف از این مقاله پیش‌بینی و یادگیری دمای مطلوب روزهای متوالی است. این پیش‌بینی بر اساس دمای فعلی اتاق، رطوبت کنونی اتاق، میزان تابش خورشید و وضعیت آب و هوای منطقه انجام شده است. در این مقاله پیش‌بینی دمای روزهای متوالی توسط روش ϵ -greedy انجام شده است. دقت پیش‌بینی روش یادگیری تقویتی ۸۱٫۵۶٪ است که بالاتر از دقت دیگر روش‌های طبقه‌بندی است.

کلمات کلیدی

یادگیری تقویتی، پیش‌بینی دما، آسایش حرارتی، ساختمان‌های هوشمند.

۱- مقدمه

در مرجع [۳] برای پیش‌بینی دما از یک شبکه عصبی چند لایه پرسپترون استفاده شده است که در آن دما و رطوبت هر ساعت، به‌عنوان پارامترهای ورودی در نظر گرفته شده‌اند. در مرجع [۴] از شبکه عصبی بازگشتی برای پیش‌بینی دما استفاده شده است. در مرجع [۵] با توجه به اطلاعات ۳۰ ایستگاه آب و هوایی ایران، با استفاده از سیستم تطبیقی عصبی فازی و سیستم شبکه عصبی به پیش‌بینی درجه حرارت هوا پرداخته شده است. در نهایت این تحقیق، نتایج حاصل از این دو روش پیش‌بینی را با هم مقایسه کرده است. همچنین در مرجع [۶] برای پیش‌بینی درجه هوا از اطلاعات ۵۰ ایستگاه آب و هوایی ایران استفاده شده است. در این پیش‌بینی از روش‌های سیستم تطبیقی عصبی فازی با بخش بندی شبکه، سیستم تطبیقی عصبی فازی با خوشه‌کاشی، شبکه عصبی و رگرسیون بردار پشتیبان استفاده شده است. علاوه بر این، در مرجع [۷] از دو شبکه عصبی چند لایه پرسپترون و راف برای پیش‌بینی

آسایش حرارتی، بصری و تهویه، سه عامل مؤثر در آسایش ساکنین ساختمان هستند [۱]. هدف از ارائه سیستم ساختمان‌های هوشمند تامین خودکار آسایش ساکنین به وسیله تعدادی عوامل ورودی تاثیرگذار است. آسایش حرارتی وابسته به داده‌های محیطی و عملکرد ساکنین است [۲]. داده‌های محیطی به دو دسته داده‌های داخل (ابعاد و دمای ساختمان و اتاق‌ها) و خارج ساختمان (پیش‌بینی وضع هوا و دمای فعلی محیط) تقسیم می‌شوند. تغییرات دمای هوا روی دمای داخلی ساختمان و در نتیجه آسایش حرارتی ساکنان، تاثیرگذار است. با پیش‌بینی این تغییرات می‌توان دمای ساختمان را به‌صورت خودکار در وضعیت مطلوب قرار داد.

مطرح می‌شود که ضمن رعایت راحتی و آسایش ساکنان سعی در ایجاد تعادل بین نیاز آن‌ها و صرفه جویی در مصرف انرژی دارد.

نیاز ساکنان در ساعات مختلف شبانه روز متفاوت است و این امر موجب تغییر مقدار انرژی مصرفی می‌شود. به‌عنوان نمونه در ساختمان‌های اداری عموماً در بازه زمانی ۹ شب تا ۶ صبح نیاز چندانی به مصرف انرژی وجود ندارد. در مقابل در ساعت ۷ صبح تا ۸ شب مصرف انرژی به اوج خود می‌رسد. بنابراین در ساعات خاصی از شبانه‌روز، مصرف انرژی افزایش می‌یابد. همچنین در ساختمان‌های مسکونی، بیمارستان‌ها و آسایشگاه‌ها با احتمال بالایی ساکنان به‌صورت تمام وقت حضور دارند. از طرفی با افزایش تعداد افراد موجود در یک ساختمان، میزان مصرف انرژی افزایش می‌یابد؛ در نتیجه نیاز کاربران از ابعاد مختلف با کاهش مصرف انرژی، در تناقض است.

با توجه به محدودیت منابع انرژی، بهتر است مصرف انرژی در این محیط‌ها مدیریت شود [۱۳]. یکی از منابع انرژی نیازمند به مدیریت، منابع گرمایشی/سرمایشی است. عوامل زیادی در مدیریت سیستم گرمایشی/سرمایشی یک ساختمان هوشمند تاثیر دارند. این عوامل شامل داده‌های زیادی هستند که به دو دسته عمده عوامل محیطی و ساکنان تقسیم می‌شوند. داده‌های محیطی می‌توانند، دمای داخل و خارج ساختمان، ابعاد ساختمان، میزان تابش خورشید، سرعت باد، رطوبت و میزان بارش باران باشند [۲ و ۱۴]. داده‌های ساکنان شامل تعداد افراد و حضور یا عدم حضور آن‌ها هستند. مسئله مهم در این بخش، محدودیت داده‌های در دسترس است. به‌عنوان مثال در یک ساختمان اداری احتمالاً داده‌های مربوط به یک فرد، در طول زمان تغییر می‌کند و یا بدلیل انتقال، اخراج یا ارتقا وی از دست می‌رود و نیاز به جایگزینی داده‌های جدید است. علاوه بر این، داده‌های محیطی مانند دما احتمالاً در طول ساعات‌های شبانه‌روز و در طول فصل‌های مختلف سال تغییر می‌کنند لذا نیاز به سیستم یادگیری وجود دارد. این سیستم یادگیری باید توانایی کار با داده‌های محدود را داشته باشد و بتواند به‌طور مداوم خود را با شرایط جدید، تطبیق دهد. بنابراین می‌توانیم از روش‌های یادگیری تقویتی برای این منظور، استفاده کنیم.

۳- یادگیری تقویتی

در یکی از موارد بسیار مهم در تصمیم‌گیری دمای مطلوب، پیش‌بینی دمای محیط در آینده نزدیک است؛ همان‌طور که در بخش قبلی اشاره شد، با توجه به تغییر دمای محیط در ساعت‌ها و فصل‌های مختلف، نیاز به سیستم یادگیری وجود دارد که بتواند خود را به‌طور مداوم با شرایط جدید تطبیق دهد. روش‌های یادگیری تقویتی برای چنین سیستمی مناسب هستند. یادگیری تقویتی از این رو مورد توجه است که عامل‌های هوشمند با توجه به بازخورد محیط، برای انجام یک هدف خاص آموزش می‌بینند. این بازخورد متناسب با عمل عامل روی محیط است و توسط سیگنالی به‌صورت پاداش یا تنبیه به عامل منتقل می‌شود. می‌توان مسئله پیش‌بینی دما در آینده نزدیک را با استفاده از یک تابع مقدار، مطابق با مرجع [۱۵]، به‌صورت رابطه (۱) بیان کرد. بدین صورت ارزش عمل a ، زمانی که در حالت s قرار دارد و از سیاست π استفاده می‌کند، توسط رابطه‌ی زیر بیان می‌شود:

دما استفاده شده است و کارایی این دو شبکه با یکدیگر مقایسه شده است. در مرجع [۸] روشی برای پیش‌بینی دمای هوا با استفاده از شبکه عصبی- فازی بازگشتی مدل ممدانی ارائه شده است. در مرجع [۹] با استفاده از روش رگرسیون بردار پشتیبان حداقل مربعات و با پارامترهای ورودی میانگین رطوبت نسبی، میانگین سرعت باد، میانگین حداقل دما و میانگین حداکثر دمای ماهانه، میانگین حداکثر دمای ماهانه ایستگاه سینوپتیک شهرکرد پیش‌بینی شده است. همچنین در مرجع [۱۰] با استفاده از داده‌های دمای سالانه شهر بوشهر و روش تحلیل‌های رگرسیون و شبکه عصبی مصنوعی، پیش‌بینی دما انجام شده است. برای بهبود دقت شبکه عصبی جهت پیش‌بینی میانگین دمای هوا، در مرجع [۱۱] الگوریتمی شامل ترکیب شبکه عصبی پرسپترون چندلایه، الگوریتم ژنتیک و PCA به کار گرفته شده است. در مرجع [۱۲] مدلی ترکیبی از پیش‌پردازش داده و شبکه عصبی ماژولار برای پیش‌بینی دمای هوا در شهر تهران ارائه شده است. در این مدل ابتدا با استفاده از روش گام به گام ویژگی‌های مناسب استخراج می‌شود و سپس با شبکه بهبود یافته نگاشت خود سازمانده به‌وسیله بهینه‌سازی ازدحام پرندگان، داده‌های مناسب برای هر شبکه تولید می‌شود. مدل ماژولار مورد نظر از چهار شبکه عصبی تابع پایه شعاعی، چندلایه پرسپترون، رگرسیون عمومی و زمان تأخیر توزیع‌شده ساخته شده است. داده‌های مورد استفاده شامل شش ویژگی نقطه شبنم، رطوبت، فشار سطح دریا، دید، سرعت باد و بارندگی هستند و دما برای روز بعد پیش‌بینی می‌شود. در مقاله‌های ذکر شده، برای پیش‌بینی دما از روش یادگیری تقویتی استفاده نشده است. همچنین بیشتر پیش‌بینی‌های انجام شده بر پایه شبکه‌های عصبی هستند. در ضمن برای پیش‌بینی دما، سیستم یادگیری مورد نیاز است تا بتواند خود را با تغییرات دمایی، به‌طور مداوم تطبیق دهد. بنابراین در این مقاله برای پیش‌بینی دمای داخلی ساختمان در روزهای متوالی از سیستم یادگیری تقویتی استفاده شده است.

در این مقاله سعی شده با استفاده از یادگیری تقویتی و با در نظر گرفتن دمای کنونی اتاق، رطوبت داخلی اتاق، میزان تابش خورشید و پیش‌بینی وضع هوا، دمای داخلی ساختمان در روزهای متوالی پیش‌بینی شود. در بخش دوم این مقاله، ساختمان‌های هوشمند معرفی می‌شوند؛ در بخش سوم درباره یادگیری تقویتی بحث می‌شود؛ در بخش چهارم روش پیشنهادی بیان می‌شود؛ در بخش پنجم شبیه‌سازی و نتایج حاصل نشان داده می‌شود؛ در بخش ششم به بررسی کلی این مقاله خواهیم پرداخت و در نهایت در بخش هفتم چند نمونه از اقدامات مؤثر برای توسعه روش پیشنهادی، ارائه خواهد شد.

۲- ساختمان هوشمند

در اغلب ساختمان‌ها، مقدار انرژی الکتریکی زیادی برای روشنایی داخلی، سیستم خنک‌کننده هوا و تهویه هوا مصرف می‌شود [۱]. همچنین برای گرم کردن محیط، آب و پخت و پز از میزان قابل توجهی گاز طبیعی استفاده می‌شود. از آنجا که بسیاری از منابع انرژی مصرفی تجدیدنپذیر هستند، لزوم مدیریت مصرف انرژی در ساختمان‌ها احساس می‌شود [۱۳]. در راستای این امر بحث ساختمان‌های هوشمند

۲-۳- فرآیند تصمیم‌گیری مارکوف

به محیطی که در آن وضعیت عامل در حالت بعدی فقط به وضعیت عامل در حالت فعلی مرتبط است، محیط مارکوف می‌گویند؛ به عبارت دیگر وضعیت کنونی عامل حاوی خلاصه‌ای از اطلاعات گذشته است و برای تغییر وضعیت آن به اطلاعات دیگری نیاز نیست [۱۵]. با توجه به این توضیحات، محیط مورد بررسی در این مقاله خاصیت مارکوف دارد. با استفاده از این خاصیت می‌توان با توجه به حالت و عمل فعلی، حالت بعدی و میزان پاداش مورد انتظار را پیش‌بینی کرد. برای رابطه‌ی این مسئله، اگر برای حالت فعلی s عمل a انتخاب شود، حالت بعدی برابر $\mathbf{s}' = \mathbf{s} + \mathbf{a}$ خواهد شد. چون عمل، پیش‌بینی دمای اتاق در روزهای متوالی است، برای هر حالت s و عمل a مدل انتقال توسط رابطه‌ی (۲) بیان می‌شود:

$$\rho_{ss'}^a = P_r \{s_{t+1} = s' | s_t = s, a_t = a\} \quad (2)$$

$\rho_{ss'}^a$ برای $\mathbf{s}' = \mathbf{s} + \mathbf{a}$ ، برابر یک می‌شود. برای باقی حالت‌ها، احتمال انتقال برابر صفر است. به‌طور مشابه، ارزش مورد انتظار برای حالت بعدی s' ، با توجه به حالت s و عمل a برابر با رابطه‌ی (۳):

$$R_{ss'}^a = E \{r_{t+1} | s_t = s, a_t = a\} \quad (3)$$

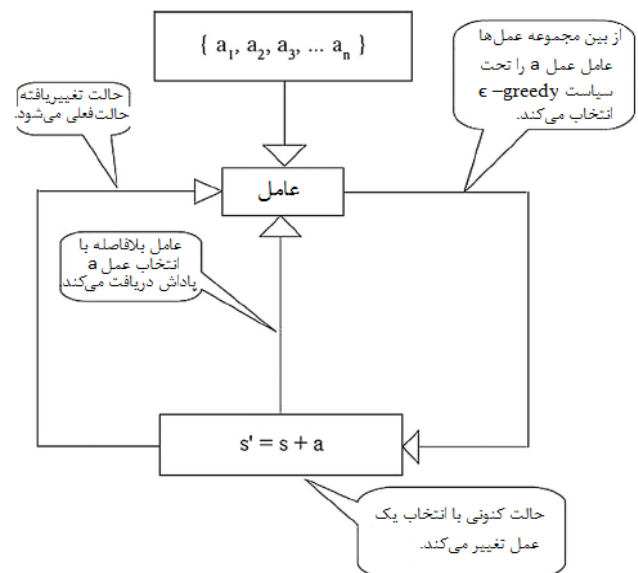
۴- روش پیشنهادی

برای پیش‌بینی دمای اتاق می‌توان از روش‌های مختلفی استفاده کرد که در این مقاله از سیاست ϵ -greedy استفاده می‌شود؛ در این روش، احتمال انتخاب عمل منجر به بیشترین پاداش، $1 - \epsilon$ و شانس انتخاب اعمال با پاداش‌های کم‌تر، متناسب با مقدار ϵ تغییر می‌کند. در این مقاله، ابتدا مقدار اولیه ϵ ، 0.9 در نظر گرفته می‌شود. در طول آموزش با افزایش دانش عامل، این مقدار کاهش داده می‌شود. آزمایش‌ها بر روی داده‌های مربوط به دمای اتاق (میانگین دمای مطلوب ساکنین)، میزان رطوبت اتاق، میزان تابش خورشید و پیش‌بینی وضع دمای بیرون در ۲۸ روز از ماه مارس سال ۲۰۱۲ انجام شده است. در هر روز ۲۵ درصد داده‌ها برای آزمون و ۷۵ درصد داده‌ها برای آموزش استفاده شده‌اند؛ همچنین داده‌های آزمون و آموزش به صورت تصادفی انتخاب می‌شوند.

در این داده‌ها دمای اتاق بین بازه ۱۱ تا ۲۶ درجه سانتی‌گراد و پیش‌بینی وضع هوا بین بازه ۴ تا ۲۶ درجه سانتی‌گراد در حال تغییر است؛ از طرفی رطوبت کنونی اتاق در بازه ۲۷ تا ۶۱ درصد و میزان تابش خورشید در بازه ۴- تا ۱۰۲۸ وات بر متر مربع، هستند. برای آموزش و پیش‌بینی بهتر، هر ۵ واحد تغییر رطوبت و هر ۵۰ واحد تغییر در میزان تابش خورشید یک واحد لحاظ می‌شوند. بنابراین حالت‌های مورد نظر، دمای کنونی، رطوبت کنونی اتاق، میزان تابش خورشید و پیش‌بینی وضع هوای بیرون هستند. همچنین عمل، پیش‌بینی دمای اتاق در روزهای متوالی در نظر گرفته می‌شوند. همان‌طور که در قسمت ۳-۱ توضیح داده شد، برای هر عملی در هر حالت، پاداشی به اندازه یک در نظر گرفته می‌شود. در نتیجه تکرار یک عمل مشخص روی یک حالت خاص و منجر شدن به حالت بعدی یکسان، پاداش بیشتری برای عمل مورد نظر در طول زمان ایجاد می‌کند.

$$Q^\pi(s, a) = E_\pi \{R_t | s_t = s, a_t = a\} \\ = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a \right\} \quad (1)$$

در رابطه‌ی (۱)، E_π برای عاملی با سیاست π ، امیدریاضی پاداش را نشان می‌دهد، R_t ارزش مورد انتظار است که به‌عنوان تابعی از مجموعه پاداش‌ها تعریف می‌شود، r_t پاداش عددی عامل در گام t است. Q^π تابع مقدار برای سیاست π است. γ ضریبی بین صفر و یک است که اهمیت پاداش‌های آینده را تعیین می‌کند. هر چه این مقدار به یک نزدیک‌تر باشد میزان تاثیر پاداش‌های آینده بیشتر است. شکل (۱) چهارچوب یادگیری تقویتی را نشان می‌دهد [۱۵]. در این شکل، عامل می‌تواند یک عمل از فضای اعمال خود را انتخاب کند و آن عمل را در یک زمان مشخص انجام دهد. این عمل باعث تغییر حالت کنونی عامل می‌شود. از سویی دیگر، عامل بلافاصله یک پاداش برای انتخاب عمل دریافت می‌کند که این به نوبه خود منجر به بروز رسانی سیاست برای انتخاب گام بعدی می‌شود.



شکل ۱: چهارچوب یادگیری تقویتی [۱۵]

۳-۱- محیط، حالت، عمل‌ها و پاداش

در این مقاله، محیط مورد نظر یک اتاق است. حالت، دمای کنونی اتاق، رطوبت کنونی اتاق و میزان تابش خورشید، پیش‌بینی وضع هوای بیرون و عمل، پیش‌بینی دمای اتاق در روزهای متوالی است. همچنین در هر حالت برای هر عمل، پاداشی به اندازه یک در نظر گرفته می‌شود. تکرار یک عمل مشخص روی یک حالت خاص و ایجاد حالت بعدی یکسان، به تدریج پاداش بیشتری برای عمل مورد نظر ایجاد می‌کند. در این مقاله تلاش می‌شود تا دمای اتاق در روزهای متوالی، به‌وسیله یادگیری تقویتی پیش‌بینی شود.

۵- شبیه‌سازی و نتایج

خطای بین ۰ تا ۱، بهترین عملکرد را نسبت به سایر روش‌های یادگیری مورد مقایسه داشته است.

۷- کارهای آینده

برای بهبود روش پیش‌بینی دمای اتاق، می‌توان پارامترهای بیشتری را دخیل نمود. این پارامترها می‌توانند شامل سرعت وزش باد و میزان رطوبت بیرونی باشند. همچنین برای پیش‌بینی دما می‌توان علاوه بر سیاست ϵ -greedy از روش‌های دیگر یادگیری تقویتی مثل الگوریتم یادگیری Q یا الگوریتم سارسا استفاده کرد.

مراجع

- [1] Yang, R. and Wang, L., Development of multi-agent system for building energy and comfort management based on occupant behaviors. *Energy and Buildings*, vol. 56, pp.1-7, 2013
- [2] Ekwevugbe, T., Brown, N. and Fan, D., A design model for building occupancy detection using sensor fusion". In IEEE International Conference on Digital Ecosystems Technologies (DEST), pp. 1-6, 2012.
- [3] Mba, L., Meukam, P. and Kemajou, A., Application of artificial neural network for predicting hourly indoor air temperature and relative humidity in modern building in humid region". *Energy and Buildings*, vol. 121, pp.32-42, 2016.
- [4] Baboo, D.S.S. and Shereef, I.K., An Efficient Temperature Prediction System using BPN Neural Network. *International Journal of Environmental Science and Development*, vol. 2, no. 1, pp.2010-0264, 2011.
- [5] Kisi, O. and Shiri, J., Prediction of long-term monthly air temperature using geographical inputs. *International Journal of Climatology*, vol. 34, no. 1, pp.179-186. 2014.
- [6] Kisi, O. and Sanikhani, H., Prediction of long-term monthly precipitation using several soft computing methods without climatic data. *International Journal of Climatology*, vol. 35, no. 14, pp.4139-4150, 2015.

[۷] رنجبر خدیوی، م. و خان‌محمدی، س.، "پیش‌بینی دما با استفاده از شبکه عصبی راف". اولین کنفرانس ملی ریاضیات صنعتی، تبریز، دانشگاه تبریز، ۱۳۹۳.

[۸] موید صفاری، م.ص.، موسوی نصر، س.م. و حور علی، م.، "تخمین دمای هوا با استفاده از شبکه عصبی فازی بازگشتی مبتنی بر آموزش با الگوریتم فرگشتی پرورشی". کنفرانس بین‌المللی یافته‌های نوین پژوهشی درمهندسی برق و علوم کامپیوتر، تهران، موسسه آموزش عالی نیکان، ۱۳۹۴.

[۹] رضائی، م. و رضائیان‌دلویی، م.، "پیش‌بینی دمای بیشینه با استفاده از روش رگرسیون بردار پشتیبان حداقل مربعات (مطالعه موردی: شهرکرد)". اولین همایش ملی گیاهان دارویی، طب سنتی و کشاورزی

هدف از این تحقیق، پیش‌بینی دمای روزهای متوالی در یک اتاق است. باید توجه داشت که تغییرات ناگهانی دما در اثر پیش‌بینی نادرست برای ساکنان نامطلوب است. لذا روشی برای پیش‌بینی مناسب است که در آن علاوه بر درصد تشخیص درست و بالا، میزان اختلاف بین دمای پیش‌بینی شده و دمای واقعی، کم باشد. بسیاری از مطالعات و تحقیقات علمی که برای پیش‌بینی یک هدف مطلوب، انجام شده‌اند، از روش‌های متفاوتی مانند شبکه‌های عصبی [۱۶ و ۱۷] و رگرسیون [۱۸] و الگوریتم‌های یادگیری ماشین [۱۹] استفاده کرده‌اند. در این مقاله برای بررسی عملکرد روش پیشنهادی در زمینه پیش‌بینی دما، نتایج حاصل از روش پیشنهادی را با نتایج حاصل از روش‌های ماشین بردار پشتیبان، شبکه عصبی سه لایه پرسپترون، مدل ترکیبی گوسین و رگرسیون خطی مورد مقایسه قرار داده‌ایم. در جدول (۱) نتایج حاصل از پیش‌بینی دمای اتاق توسط روش پیشنهادی و سایر روش‌های مورد مقایسه نشان داده شده است. همان‌طور که در جدول (۱) مشاهده می‌شود، ستون چهارم نشان‌دهنده میزان اختلاف بین دمای پیش‌بینی شده و دمای واقعی است. این اختلاف در یادگیری تقویتی نسبت به روش‌های دیگر کمتر است. همچنین درصد پیش‌بینی درست دما با استفاده از روش یادگیری تقویتی نسبت به سایر روش‌های یادگیری مورد استفاده، بالاتر است. در واقع این روش تعداد دماهای بیشتری را به صورت دقیق پیش‌بینی می‌کند و مجموع مربعات خطای آن از سایر روش‌ها کمتر است. با در نظر گرفتن این توضیحات، روش یادگیری تقویتی نسبت به سایر روش‌های مورد استفاده در این مقاله، عملکرد بهتری دارد.

جدول ۱: نتایج حاصل از آزمایش روش‌های مختلف با استفاده از ۲۵ درصد داده‌های هر روز برای آزمون و ۷۵ درصد داده‌ها برای آموزش

روش	درصد تشخیص درست	مجموع مربعات خطا	بازه خطا
یادگیری تقویتی	۸۱٫۵۶	۳۷	۱ تا ۰
ماشین بردار پشتیبان	۷۹٫۲۵	۵۶	۲ تا ۰
سه لایه پرسپترون	۷۷٫۲۶	۸۲	۲ تا ۰
مدل ترکیبی گوسین	۷۵٫۵۷	۱۷۸	۳ تا ۰
رگرسیون خطی	۷۴٫۲۵	۱۸۴	۳ تا ۰

۶- نتیجه

در دمای یک محیط در طول زمان‌ها و فصل‌های مختلف به‌طور مداوم در حال تغییر است. از این رو، در این مقاله از روش یادگیری تقویتی ϵ -greedy استفاده شده تا بتواند با چنین شرایطی، پیش‌بینی درستی از دما انجام دهد. روش یادگیری پیشنهادی با روش‌های یادگیری با نظارت مانند شبکه عصبی سه لایه پرسپترون، ماشین بردار پشتیبان و روش یادگیری بدون نظارت مدل ترکیبی گوسین و رگرسیون خطی مقایسه شد. آزمایش‌ها بر روی داده‌های مربوط به دمای یک اتاق، میزان رطوبت اتاق، میزان تابش خورشید و پیش‌بینی وضع دمای بیرون در ۲۸ روز از ماه مارس سال ۲۰۱۲ انجام شد. در هر روز ۲۵ درصد داده‌ها برای آزمون و ۷۵ درصد داده‌ها برای آموزش استفاده شدند. نتایج حاصل نشان داد که روش یادگیری تقویتی با درصد تشخیص ۸۱٫۵۶ و بازه

ارگانیک، همدان، انجمن ارزیابان محیط زیست هگمتانه، مرکز توسعه همایش‌های آریا هگمتان، ۱۳۹۳.

[۱۰] سلیقه، م.، بلیانی، ی.، بیگلی، ز.، و خسروی، ی.، مقایسه و پیش‌بینی دمای سالانه شهر بوشهر به روش تحلیل‌های رگرسیونی و شبکه عصبی مصنوعی". دومین کنفرانس بین‌المللی مخاطرات محیطی، تهران، دانشگاه خوارزمی، ۱۳۹۲.

[۱۱] ابطی، ا.، قویدل، ا. و لاهوتی، ا.، "پیش‌بینی میانگین دمای هوا با رویکرد ترکیبی شبکه عصبی، ژنتیک و PCA". دهمین کنفرانس بین‌المللی مهندسی صنایع، تهران، انجمن مهندسی صنایع ایران، دانشگاه صنعتی امیرکبیر، ۱۳۹۲.

[۱۲] احمدی، ع.، زرگران، ز. و تقوی، ف.، پیش‌بینی دمای هوا با استفاده از شبکه‌های عصبی ترکیبی ماژولار". یازدهمین کنفرانس سراسری سیستم‌های هوشمند، انجمن سیستم‌های هوشمند ایران، دانشگاه خوارزمی، ۱۳۹۱.

[13] Wang, Z., Wang, L., Dounis, A.I. and Yang, R., Multi-agent control system with information fusion-based comfort model for smart buildings. *Applied Energy*, vol. 99, pp.247-254, 2012.

[14] Zhang, F., de Dear, R. and Candido, C., Thermal comfort during temperature cycles induced by direct load control strategies of peak electricity demand management. *Building and Environment*, vol. 103, pp.9-20, 2016.

[15] Sutton, R.S., Barto, A.G. and Bach, F., Reinforcement learning: An introduction". MIT press, 1998.

[۱۶] نوذری، ح. و آزادی، س.، پیش‌بینی شوری آب زیرزمینی زیر لوله‌های زهکش با استفاده از شبکه عصبی". *مجله مدل‌سازی در مهندسی، دوره ۱۶، شماره ۵۲، ۱۳۹۷*.

[۱۷] افرند، م. و اسفه، م. ه.، مدل‌سازی با استفاده از شبکه عصبی مصنوعی جهت پیش‌بینی هدایت حرارتی نانوسیال نانولوله کربنی چند جداره عامل دار - آب و ارائه رابطه تجربی جدید". *مجله مدل‌سازی در مهندسی، دوره ۱۶، شماره ۵۳، ۱۳۹۷*.

[۱۸] محمودی، م. و نادری بختیاری، ع.، پیش‌بینی خواص مکانیکی آلومینیوم حاصل از فرآیند نورد در کانال همسان زاویه‌دار با استفاده از شبکه عصبی مصنوعی و رگرسیون غیرخطی". *مجله مدل‌سازی در مهندسی، دوره ۱۵، شماره ۵۱، ۱۳۹۶*.

[۱۹] شفیعی نیک آبادی، م. و عظیمی، س.ع.، پیش‌بینی تقاضا در زنجیره تامین با استفاده از الگوریتم‌های یادگیری ماشین (مورد مطالعه: زنجیره تامین شرکت ایران خودرو)". *مجله مدل‌سازی در مهندسی، دوره ۱۳، شماره ۴۱، ۱۳۹۴*.



طبقه‌بندی تصاویر ابرطیفی با استفاده از پروفایل‌های توصیفی و گشتاورهای تغییر ناپذیر

زهرا رستمی^۱، حسن قاسمیان^۲، مریم ایمانی^۳

^۱ دانشجوی کارشناسی ارشد، گروه آموزشی مهندسی برق مخابرات سیستم، دانشگاه تربیت مدرس، تهران
z-rostami@modares.ac.ir

^۲ استاد گروه مخابرات سیستم، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران
ghassemi@modares.ac.ir

^۳ استادیار گروه مخابرات سیستم، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران
maryam.imani@modares.ac.ir

چکیده

پروفایل‌های توصیفی (APs) از طریق اعمال متوالی فیلترهای توصیفی به تصویر، حاصل می‌شوند. ثابت شده است که APs در سنجش از دور، برای توصیف ویژگی‌های مکانی اشیا موجود در صحنه تصویر، روش قدرتمندی است؛ اما هنگامی که بافت پیچیده‌ای در تصویر وجود داشته باشد، استفاده‌ی مستقیم از APs کافی نیست. در این مقاله، از گشتاورهای تصویر در کنار پروفایل‌های توصیفی، استفاده شده است. گشتاورها، ویژگی‌های بافت را از تصویر استخراج می‌کنند. مراحل انجام کار به این صورت است که ابتدا پروفایل‌های توصیفی و گشتاورهای تصویر، ایجاد شده و سپس این ویژگی‌ها ترکیب شده‌اند و در نهایت طبقه‌بندی با SVM انجام شده است. نتایج حاصل، روی دو تصویر ابرطیفی دانشگاه پاویا و تصویر ابرطیفی ایندیانا نشان می‌دهد که روش به کار رفته، نسبت به حالتی که تنها پروفایل‌های توصیفی استفاده شود، صحت طبقه‌بندی را افزایش داده است.

کلمات کلیدی

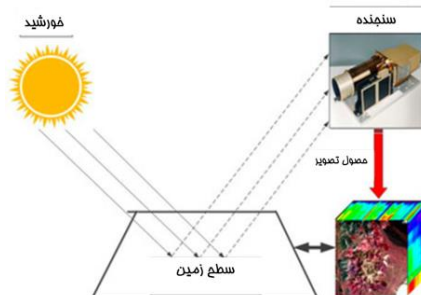
سنجش از دور، ابرطیفی، طبقه‌بندی، پروفایل توصیفی، گشتاور

اطلاعات بسیار غنی، تفسیر خودکار این تصاویر به صورت یک چالش باقی مانده است.

۱- مقدمه

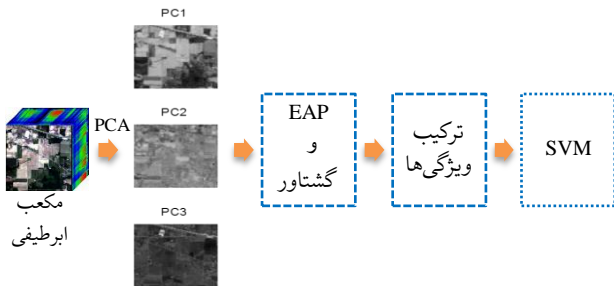
سنجش از دور ابرطیفی روشی برای استخراج اطلاعات اشیا موجود بر سطح زمین بر اساس میزان بازتاب دریافتی (در باندهای مختلف طیف الکترومغناطیس) توسط سنجنده‌های مستقر بر روی ماهواره است. این موضوع در شکل (۱) نشان داده شده است.

زمینه‌های تحقیقاتی مختلفی در پردازش تصاویر ابرطیفی وجود دارد. طبقه‌بندی این تصاویر، یعنی تخصیص برچسب به هر پیکسل موجود در تصویر، به منظور تولید نقشه‌های پوشش زمین، حوزه‌ی تحقیقاتی مهمی در سنجش از دور به شمار می‌رود. این تصاویر، دقت طیفی، مکانی و زمانی بسیار بالایی دارند. علیرغم رزولوشن بالا و



شکل ۱: نمایش فرآیند سنجش از دور و بدست آوردن تصویر ابرطیفی

یک مجموعه از گشتاورها ایجاد شده، با دنبال هم قرار دادن APها و ویژگی‌های گشتاور، این ویژگی‌ها ترکیب شده و در نهایت طبقه‌بندی با ماشین بردار پشتیبان انجام شده است.



شکل ۲: مراحل انجام کار در این مقاله

در ادامه، ساختار مقاله به این صورت است که در بخش دوم روش‌های به کار رفته جهت استخراج ویژگی‌ها توضیح داده شده، در بخش سوم نحوه ارزیابی و نتایج آورده شده و در بخش چهارم جمع‌بندی صورت گرفته است.

۲- استخراج اطلاعات مکانی

همانطور که در بخش مقدمه گفته شد، جهت استخراج اطلاعات مکانی از هر مولفه‌ی حاصل از تبدیل PCA، از فیلترهای توصیفی برای ایجاد APها و از گشتاورهای تصویر استفاده شده است. در این بخش هر یک از این روش‌ها شرح داده شده است.

۲-۱- پروفایل‌های توصیفی

بر خلاف فیلترهای رایجی که در پردازش تصویر به کار می‌روند و مبتنی بر پیکسل هستند، فیلتر توصیفی مبتنی بر ناحیه است. این فیلترها تصویر را با ادغام اجزای به هم پیوسته‌ی آن در سطوح خاکستری مختلف، مطابق با یک معیار، پردازش می‌کنند. اجزای به هم پیوسته، مجموعه‌ای از پیکسل‌های متصل و دارای شدت یکسان است.

عملیات فیلتر کردن تصویر باینری به کمک فیلترهای توصیفی به صورت زیر انجام می‌شود: اگر اجزای به هم پیوسته‌ی موجود در تصویر را با C نشان دهیم، ابتدا یک ویژگی توصیفی، برای مثال مساحت هر یک از اجزای به هم پیوسته، $A(C_i)$ ، در نظر گرفته شده و یک سطح آستانه (λ) برای این ویژگی توصیفی تعیین می‌شود. برای هر یک از اجزای به هم پیوسته‌ی موجود در تصویر، ویژگی توصیفی آن محاسبه شده و با سطح آستانه‌ی مورد نظر مقایسه می‌شود [6]. برای مثال، مقایسه می‌تواند طبق معیار P به صورت رابطه‌ی (۱) انجام شود:

$$P(C) = A(C_i) > \lambda \quad (1)$$

اگر برای یک جز پیوسته، عبارت فوق صحیح باشد، آن جز بدون تغییر باقی می‌ماند؛ در غیر این صورت با جز پیوسته‌ی مجاور ادغام می‌شود. این موضوع در شکل (۳) نشان داده شده است.

در مورد تصاویر سطح خاکستری، ابتدا یک تصویر با دسته‌ای از تصاویر باینری که با گذاشتن آستانه در هریک از سطوح خاکستری‌اش به دست می‌آید، بیان شده، سپس فیلترهای توصیفی بر هر یک از تصاویر باینری به صورتی که شرح داده شد اعمال می‌شود، تصویر

طبقه‌بندی دقیق صحنه‌های پیچیده، مانند مناطق شهری، در داده‌ی خام سنجنش از دور با مشکلاتی روبه‌رو است. برای مثال در تصویر یک منطقه‌ی شهری، مشخصات طیفی کلاس‌هایی مانند سطح جاده‌ها و محوطه‌ی پارکینگ‌ها بسیار مشابه است به گونه‌ای که تنها با استفاده از اطلاعات طیفی تصویر خام، نمی‌توان آن‌ها را از یکدیگر تشخیص داد؛ بنابراین الگوریتم‌های پردازش تصویر جهت به کارگیری اطلاعات مکانی برای طبقه‌بندی این تصاویر، توسعه داده شده است. اطلاعات مکانی شامل مشخصات بافت و مشخصات هندسی اشیاء موجود در تصویر است. در مقالات سنجنش از دور روش‌های مختلفی برای استخراج اطلاعات مکانی و بافت از تصاویر به کار گرفته شده است.

از میان روش‌های استخراج اطلاعات مکانی، در سال‌های اخیر، استفاده از پروفایل‌های توصیفی (AP) توسط محققین مورد توجه قرار گرفته است [1]. پروفایل‌های توصیفی با اعمال دنباله‌ای از فیلترهای توصیفی به تصویر سطح خاکستری، بدست می‌آیند [2]. فیلترهای توصیفی عملگرهایی هستند که در چارچوب مورفولوژی ریاضی تعریف شده‌اند و مولفه‌های متصل تصویر را در سطوح مختلف با توجه به برخی اندازه‌گیری‌های روی مولفه‌ها ادغام می‌کنند. با استفاده از APها اطلاعات مربوط به اندازه و شکل اشیاء موجود در تصویر حاصل می‌شود. در این روش ویژگی‌های مکانی مهم نواحی و اشیاء داخل تصویر مانند لبه‌ها به خوبی حفظ می‌شود.

پس از موفقیت APها در مدل‌سازی اطلاعات مکانی، نسخه‌ی بهبود یافته‌ای از آن با عنوان پروفایل توصیفی خود دوگانه (SDAP) مطرح شد [3] و ثابت شده است که به لحاظ صحت طبقه‌بندی و هزینه‌ی محاسباتی، نسبت به APها بهتر عمل می‌کند. در این روش نواحی تیره و روشن تصویر به صورت همزمان پردازش می‌شود، در نتیجه ابعاد ویژگی‌ها کاهش یافته است.

استفاده از APها و SDAPها به صورت مستقیم برای توصیف کاملی از اطلاعات ساختاری و بافت جهت طبقه‌بندی کافی نبود، بنابراین پروفایل‌های توصیفی مبتنی بر هیستوگرام (HAP) ارائه شد [4]. در این روش هیستوگرام محلی هر پیکسل، در هر یک از تصاویر پروفایل توصیفی، محاسبه شده و سپس در یک دنباله قرار گرفته است. ثابت شده است که این روش برای طبقه‌بندی تصاویر با رزولوشن بالا بسیار بهینه‌تر است. اما در این روش ابعاد فضای ویژگی بسیار زیاد است.

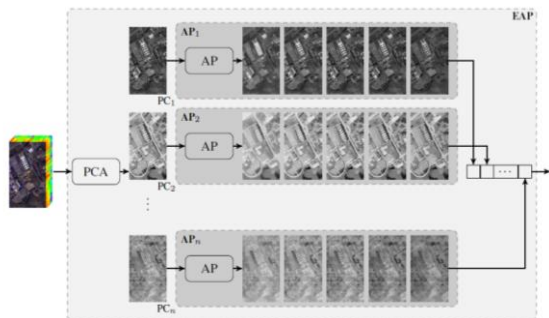
پس از HAPها، پروفایل‌های توصیفی مبتنی بر ویژگی‌های محلی (LFAP) معرفی شد [5]. در این روش به جای ساختن هیستوگرام‌های محلی، برخی ویژگی‌های آماری پیکسل‌ها استخراج شده است. مشکلات موجود در روش قبلی در این روش وجود ندارد، ضمن اینکه توصیف بهتری از بافت‌های محلی در تصاویر با رزولوشن بالا، نسبت به APها، فراهم می‌کند.

در این مقاله، جهت طبقه‌بندی تصاویر ابرطیفی، از APها به همراه گشتاورها برای استخراج اطلاعات مکانی و بافت استفاده شده و مراحل انجام کار در شکل (۲) نشان داده شده است.

ابتدا ابعاد تصویر ابرطیفی به کمک تبدیل PCA به سه مولفه کاهش یافته است. سپس برای هر یک از مولفه‌های حاصل، یک AP و

$$AP = \{\varphi_n, \dots, \varphi_1, f, \rho_1, \dots, \rho_n\} \quad (2)$$

در رابطه‌ی (۲)، φ_i و ρ_i به ترتیب نشان دهنده‌ی نازک کردن و ضخیم کردن توصیفی با حد آستانه‌ی λ_i است. در مورد تصاویر ابرطیفی که تعداد ابعاد تصویر زیاد است، ابتدا با استفاده از یک الگوریتم کاهش بعد مانند PCA، تعداد ابعاد تصویر کاهش داده شده، سپس برای هر مولفه‌ی حاصل، یک AP ایجاد می‌شود. با در کنار هم قرار دادن این APها یک پروفایل توصیفی گسترش یافته (EAP) حاصل می‌شود. این موضوع در شکل (۵) نشان داده شده است.



شکل ۵: نحوه‌ی تشکیل EAP برای تصویر ابرطیفی [8]

ویژگی‌های توصیفی که برای تولید APها، بسیار در مقالات استفاده شده است، شامل موارد زیر است:

- ۱- مساحت ناحیه‌ها (مربوط به اندازه‌ی نواحی)
 - ۲- انحراف معیار (شاخصی برای نشان دادن همگنی نواحی)
 - ۳- گشتاور اینرسی (شاخصی برای اندازه‌گیری کشیدگی نواحی)
- با در نظر گرفتن چند ویژگی توصیفی متفاوت و ایجاد یک EAP به ازای هر ویژگی توصیفی و در کنار هم قرار دادن آن‌ها یک پروفایل توصیفی گسترش یافته با چند ویژگی توصیفی (EMAP) ایجاد می‌شود.

۲-۲- گشتاورها

از نقطه نظر ریاضی، گشتاورها، تصویرهای یک تابع به پایه‌های چند جمله‌ای‌ها است. اگر $f(x, y)$ یک تابع و $\{h_{pq}(x, y) | p, q\}$ اعداد صحیح: p, q پایه‌های چند جمله‌ای باشند، گشتاور مرتبه‌ی $p+q$ به صورت رابطه‌ی (۳) محاسبه می‌شود:

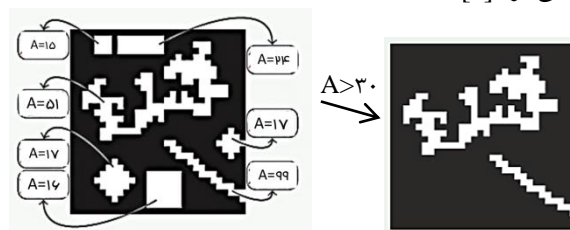
$$m_{pq} = \sum_{(x,y) \in \Omega} h_{pq}(x, y) f(x, y), \quad (3)$$

در معادله‌ی (۳) Ω ناحیه‌ی مورد نظر برای محاسبه‌ی گشتاور است [9]. در این مقاله، Ω یک پنجره‌ی مربعی به مرکز پیکسل مورد نظر و اندازه‌ی دلخواه است.

بسته به پایه‌های چند جمله‌ای استفاده شده، سیستم‌های گوناگونی از گشتاورها تعریف شده است. گشتاور هندسی ساده‌ترین نوع گشتاورها است که عبارت پایه‌ی چند جمله‌ای در آن $x^p y^q$ است: p و q اعداد صحیح غیر منفی هستند.

مجموعه‌ای از گشتاورها شامل همه‌ی m_{pq} هایی که در آن $p+q \leq n$ باشد، دارای $\frac{1}{2}(n+1)(n+2)$ عضو است. برای مثال گشتاورهای مرتبه‌ی صفر دارای یک عضو m_{00} ، گشتاورهای مرتبه‌ی

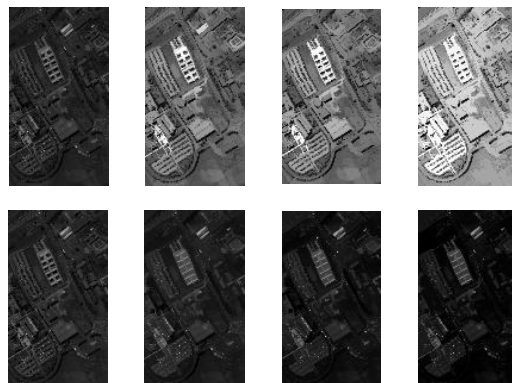
سطح خاکستری فیلتر شده با استفاده از تصاویر باینری فیلتر شده ساخته می‌شود. [7]



شکل ۳: نمایش نحوه‌ی فیلتر شدن تصویر به کمک فیلترهای توصیفی؛ (سمت چپ تصویر باینری با اجزای به هم پیوسته و ویژگی توصیفی هر کدام (A) و سمت راست تصویر فیلتر شده) [8]

اگر C جز پیوسته‌ای باشد که معیاری را برآورده کند، آنگاه همه‌ی مجموعه‌هایی که C عضوی از آنهاست نیز آن معیار را برآورده می‌کنند؛ در این صورت آن معیار، یک معیار افزایشی است. برای مثال عبارت «باید مساحتی برابر λ یا بیشتر داشته باشد» یک نمونه معیار افزایشی است.

فیلترهای توصیفی، شامل باز کردن و بستن توصیفی است. عملگر باز کردن روی اشیا روشن (نسبت به محیط اطراف) عمل می‌کند و اشیا را که از حد آستانه کوچکتر باشد را با پس زمینه ادغام می‌کند در حالی که اشیا تیره بدون تغییر باقی می‌ماند. در مقابل، عملگر بستن روی اشیا تیره عمل می‌کند و اشیا کوچک و تیره را حذف کرده و اشیا روشن را بدون تغییر باقی می‌گذارد. اگر در تعریف باز کردن (بستن) توصیفی از یک معیار غیر افزایشی استفاده شود، فیلتر حاصل نازک کردن (ضخیم کردن) توصیفی نامیده می‌شود. شکل (۴) تصاویر حاصل از این فیلترها را روی یک تصویر سنجنش از دور نشان می‌دهد.



شکل ۴: تصویر اصلی (اولین شکل از سمت چپ) و تصاویر حاصل از فیلترهای توصیفی با ویژگی توصیفی مساحت (بالا): تصاویر حاصل از نازک کردن، پایین: ضخیم کردن توصیفی با مقادیر آستانه، به ترتیب از چپ به راست: ۵۰۰، ۱۰۰۰، ۵۰۰۰

فرض می‌شود مجموعه‌ی $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ شامل مقادیر حد آستانه باشد، به گونه‌ای که اگر $i < j$ آنگاه $\lambda_i \leq \lambda_j$ است. پروفایل توصیفی با اعمال دنباله‌ای از عملگرهای نازک کردن و ضخیم کردن، به تصویر سطح خاکستری تک باند f به صورت رابطه‌ی (۲) حاصل می‌شود.

است. جهت استخراج گشتاورها برای تصویر ابرطیفی، هر ۷ گشتاور، برای هر مولفه، محاسبه شده و سپس همه‌ی آنها، در یک مجموعه قرار داده شده‌اند. سه اندازه‌ی مختلف ۳، ۹ و ۱۵ نیز برای اندازه‌ی پنجره در نظر گرفته شده است. پس از استخراج اطلاعات مکانی و بافت با استفاده از فیلترهای توصیفی و گشتاورها، آنها را در یک مجموعه، به دنبال هم، قرار داده تا ورودی سیستم طبقه‌بند باشند.

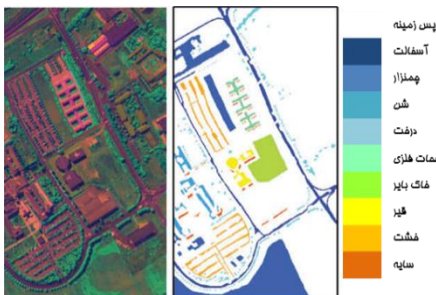
جهت طبقه‌بندی از SVM با هسته‌ی چندجمله‌ای، بدون تنظیم پارامترها استفاده شده است.

۳-۲- نرم افزار پیاده سازی و داده‌ها

پیاده سازی همه‌ی روش‌ها با نرم‌افزار MATLAB و پیاده‌سازی طبقه‌بند SVM با کتابخانه Libsvm [10] انجام شده است. داده‌های ابرطیفی استفاده شده، شامل دو تصویر دانشگاه پاولیا و تصویر ایندیانا است.

۱- تصویر ابرطیفی دانشگاه پاولیا

این تصویر، قسمتی از منطقه‌ی شهری شهر پاولیا (دانشگاه پاولیا) واقع در کشور ایتالیا است که به کمک سنجنده ROSIS جمع آوری شده و دارای ۱۱۵ باند طیفی در محدوده‌ی طول موج ۰/۴۳ تا ۰/۸۶ میکرومتر است که بعد از حذف باندهای نویزی شامل ۱۰۳ باند طیفی شده است. ابعاد آن ۳۴۰ × ۶۱۰ پیکسل است. این منطقه، دارای ۹ گونه پوشش شهری است؛ که در شکل (۶) نشان داده شده است [11].



شکل ۶: داده‌ی ابرطیفی دانشگاه پاولیا (سمت چپ) و نقشه‌ی مرجع آن (سمت راست)

۲- تصویر ابرطیفی ایندیانا

این تصویر به کمک سنجنده AVIRIS جمع آوری شده و دارای ابعاد ۱۴۵ × ۱۴۵ پیکسل و ۲۲۴ باند طیفی در محدوده‌ی طول موج ۰/۴ تا ۲/۵ میکرومتر است. بعد از حذف باندهای نویزی، تصویری شامل ۲۰۰ باند حاصل شده است. این داده شامل ۱۶ پوشش گیاهی مختلف است و در شکل (۷) نمایش داده شده است [11].



شکل ۷: داده ابرطیفی ایندیانا (سمت چپ) و نقشه‌ی مرجع آن (سمت راست)

اول شامل سه عضو m_{00} ، m_{01} و m_{10} است و به همین ترتیب مرتبه‌ی های بالاتر نیز تعریف می‌شود. m_{00} جرم تصویر، m_{01} و m_{10} مرکز ثقل و m_{02} ، m_{11} ، m_{20} (گشتاورهای مرتبه‌ی دوم) جهت تصویر را نشان می‌دهند.

اگر در تعریف گشتاور هندسی عبارت پایه‌ی چند جمله‌ای به صورت $(x - \bar{x})^p (y - \bar{y})^q$ تغییر کند، که در آن \bar{x} و \bar{y} به صورت رابطه‌ی (۴) تعریف می‌شوند:

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{و} \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (4)$$

گشتاور حاصل، گشتاور مرکزی نامیده می‌شود. گشتاور مرکزی نسبت به حرکت انتقالی محورهای مختصات تغییر ناپذیر است.

مجموعه‌ای از گشتاورها که نسبت به مقیاس، انتقال و جهت تغییر ناپذیر هستند، شامل ۷ تابع است که با استفاده از گشتاورهای مرکزی تا مرتبه‌ی سوم، ایجاد شده‌اند. چهار معادله‌ی اول از این هفت تابع، در رابطه‌ی (۵) ذکر شده است.

$$\begin{aligned} \phi_1 &= m_{20} + m_{02} \\ \phi_2 &= (m_{20} - m_{02})^2 + 4m_{11}^2 \\ \phi_3 &= (m_{30} - 3m_{12})^2 + (3m_{21} - m_{03})^2 \\ \phi_4 &= (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2 \\ \phi_5 &= \dots \\ \phi_6 &= \dots \\ \phi_7 &= \dots \end{aligned} \quad (5)$$

برای محاسبه‌ی گشتاورهای تصویر، یک پنجره با اندازه مناسب انتخاب شده، به گونه‌ای که مرکز پنجره روی پیکسل مورد نظر قرار گرفته و گشتاور در آن پنجره محاسبه شده و مقدار حاصل، به آن پیکسل نسبت داده می‌شود.

۳- ارزیابی

در این قسمت نحوه‌ی ارزیابی روش پیشنهادی شرح داده شده است.

۳-۱- تنظیم پارامترها

در این مقاله از تصاویر سنجنش از دور ابرطیفی استفاده شده است. ابتدا ابعاد تصویر، با استفاده از تبدیل PCA به سه بعد کاهش یافته، دلیل استفاده از این روش سادگی و بهینه بودن آن است. برای ایجاد AP از هر مولفه، دو ویژگی توصیفی مساحت (a) و گشتاور اینرسی (mi) انتخاب شده و مقادیر حد آستانه برای هر کدام از ویژگی‌های فوق به صورت زیر در نظر گرفته شده است:

$$\lambda_{a_i} = \{1.00, 5.00, 10.00, 20.00, 30.00, 40.00, 50.00, 60.00, 70.00, 80.00\}$$

$$\lambda_{mi_i} = \{0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55\}$$

به این ترتیب به ازای هر یک از ویژگی‌های توصیفی، با استفاده از فیلترهای توصیفی و مقادیر آستانه، یک EAP برای سه مولفه ایجاد شده و با در کنار هم قرار دادن هر دو EAP یک EMAP ساخته شده است. سپس از گشتاورهای تغییر ناپذیر در کنار APها استفاده شده

۳-۳- معیارهای ارزیابی

عملکرد هر روش به صورت کمی به کمک معیارهای زیر ارزیابی شده است:

۱- صحت کلی (OA): نسبت تعداد نمونه‌هایی که به درستی طبقه بندی شده‌اند به تعداد همه‌ی نمونه‌های آزمایشی.

۲- ضریب توافق کاپا (K): این ضریب در واقع نوعی صحت کلی است با این تفاوت که سعی می‌کند قسمتی از صحت طبقه‌بندی را که ممکن است در اثر تصادف (شانس) باشد در نظر نگیرد. محاسبه‌ی این ضریب به کمک فرمول (۶) انجام می‌شود:

$$\hat{k} = \frac{OA - P_e}{1 - P_e} \quad (6)$$

$$P_e = \frac{\sum_{c=1}^M m_c \cdot N_c}{N^2} \quad (7)$$

در فرمول (۷)، N تعداد کل نقاط آزمایشی مربوط به همه دسته‌ها؛ N_c تعداد کل نقاط آزمایشی دسته c ؛ و m_c تعداد کل نقاطی است که در خروجی طبقه‌بند برچسب دسته c به آن‌ها داده شده است.

۳-۴- طبقه‌بندی و نتایج

تعداد نمونه‌هایی که برای آموزش و آزمایش طبقه‌بند از تصویر ابرطیفی دانشگاه پاویا انتخاب شده، در جدول (۱) آورده شده (نقاط، به صورت تصادفی از روی نقشه‌ی مرجع انتخاب شده است).

جدول ۱: تعداد نمونه‌های آموزشی و آزمایشی داده‌ی ابرطیفی دانشگاه پاویا

کلاس‌ها	تعداد نمونه‌های آموزشی	تعداد نمونه‌های آزمایشی
آسفالت	۵۴۸	۶۶۳۱
چمنزار	۵۴۰	۱۸۱۶۹
شن	۳۹۲	۲۰۹۹
درختان	۵۲۴	۳۰۶۴
ورق فلزی	۲۶۵	۱۳۴۵
خاک	۵۳۲	۵۰۲۹
قیر	۳۷۵	۱۳۳۰
آجر	۵۱۴	۳۶۸۲
سایه	۲۳۱	۹۴۷

نتایج طبقه‌بندی تصویر ابرطیفی دانشگاه پاویا در جدول (۲) و شکل (۸) آورده شده است. در هر مورد طبقه‌بندی ۵ مرتبه انجام شده و میانگین آن‌ها گزارش شده است.

با توجه به نتایج طبقه‌بندی تصویر ابرطیفی دانشگاه پاویا، مشاهده می‌شود که کمترین دقت طبقه‌بندی مربوط به داده‌ی خام ابرطیفی است. استخراج اطلاعات مکانی توسط APها دقت طبقه‌بندی را افزایش داده است. در میان APها بیشترین دقت، مربوط به ویژگی‌های پروفایل توصیفی مساحت است که دقت طبقه‌بندی آن ۹ درصد بهتر از طبقه‌بندی داده‌ی خام ابرطیفی است. طبقه‌بندی با EAPa نسبت به طبقه‌بندی با EMAP بهتر است زیرا هم تعداد ویژگی‌ها کمتر است و هم دقت آن بالاتر است.

افزودن ویژگی‌های گشتاور به هر یک از EAPها باعث افزایش دقت طبقه‌بندی، نسبت به حالتی که تنها پروفایل‌های توصیفی استفاده شده، شده است؛ که در مورد EAPmi این افزایش دقت، مشهودتر است (افزایش دقت به اندازه‌ی تقریباً ۵ درصد) اما ترکیب گشتاورها با EMAP تقریباً تاثیری در نتایج طبقه‌بندی نداشته است.

تعداد نمونه‌های آموزشی تصویر ابرطیفی ایندیانا، ۱۰ درصد از کل تعداد پیکسل‌های مربوط به هر کلاس است؛ که به صورت تصادفی از روی نقشه‌ی مرجع انتخاب شده است. مابقی پیکسل‌های هر کلاس، برای آزمایش طبقه‌بند استفاده شده است. همانند قبل، در مورد هر دسته از ویژگی‌ها، طبقه‌بندی ۵ مرتبه انجام شده و میانگین نتایج در جدول (۲) و شکل (۹) گزارش شده است.

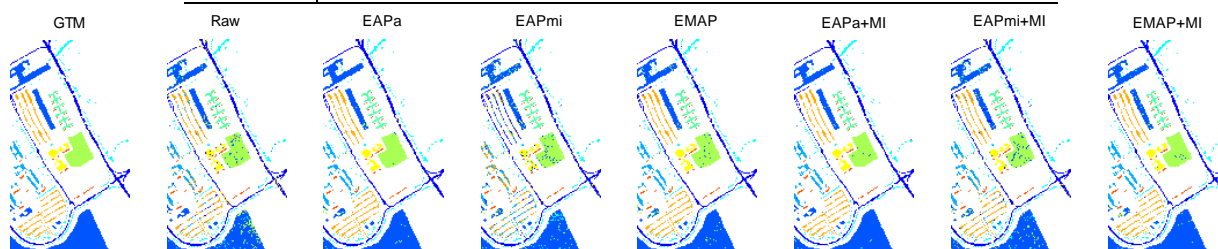
با توجه به جدول (۲) و نتایج طبقه‌بندی داده‌ی ابرطیفی ایندیانا مشاهده می‌شود که APها دقت طبقه‌بندی را نسبت به داده‌ی خام افزایش داده‌اند. نتایج، تقریباً مشابه داده‌ی ابرطیفی دانشگاه پاویا است.

افزودن گشتاورها به APها باعث افزایش دقت در مورد پروفایل توصیفی مساحت شده است، اما تاثیری در دقت EAPmi و EMAP نداشته است.

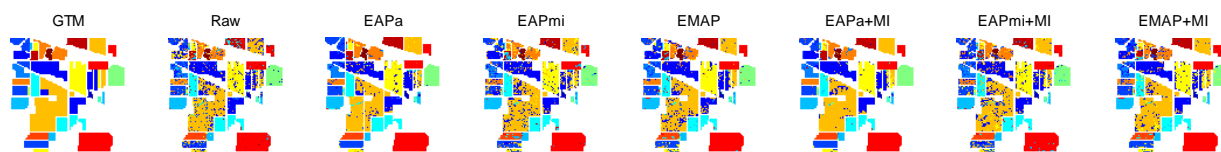
در حالت کلی می‌توان نتیجه گرفت که ترکیب گشتاورهای تغییر ناپذیر با پروفایل‌های توصیفی، برای داده‌های ابرطیفی مناطق شهری مناسب تر است. در این حالت دقت طبقه‌بندی نسبت به حالتی که تنها از پروفایل‌های توصیفی استفاده شده است، بیشتر است.

جدول ۲: نتایج طبقه‌بندی داده‌ی ابرطیفی دانشگاه پاویا و تصویر ابرطیفی ایندیانا

داده‌ی خام	EAP		EMAP	EAP+گشتاور		EMAP+گشتاور	
	EAPa	EAPmi		EAPa+گشتاور	EAPmi+گشتاور		
دانشگاه پابویا							
تعداد ویژگی	۱۰۳	۶۳	۱۲۶	۱۲۶	۱۲۶	۱۸۹	
OA	۹۰/۳۴	۹۸/۹۵	۹۷/۴۱	۹۹/۰۲	۹۰/۱۶	۹۷/۴۱	
K	۰/۸۷	۰/۹۸	۰/۹۶	۰/۹۹	۰/۸۷	۰/۹۶	
ایندیانا							
تعداد ویژگی	۲۰۰	۶۳	۱۲۶	۱۲۶	۱۲۶	۱۸۹	
OA	۷۷/۳۶	۹۱/۰۹	۸۵/۶۴	۹۲/۳۶	۷۹/۵۴	۸۵/۶۴	
K	۰/۷۴	۰/۹۰	۰/۸۴	۰/۹۱	۰/۷۷	۰/۸۴	



شکل ۸: نقشه‌ی مرجع و نقشه‌های طبقه‌بندی با پروفایل‌های توصیفی و طبقه‌بندی با ترکیب پروفایل‌های توصیفی و گشتاورها (تصویر ابرطیفی دانشگاه پابویا) (MI: Moment Invariant)



شکل ۹: نقشه‌ی مرجع و نقشه‌های طبقه‌بندی با پروفایل‌های توصیفی و طبقه‌بندی با ترکیب پروفایل‌های توصیفی و گشتاورها (تصویر ابرطیفی ایندیانا)

۴- جمع‌بندی

را مدل می‌کند؛ که جهت بهبود نتایج طبقه‌بندی، با ویژگی‌های بافتی حاصل از گشتاورهای تغییر ناپذیر، ترکیب شده است.

فیلترهای توصیفی تصویر را بدون از بین بردن یا ایجاد لبه‌های جدید، بلکه فقط با ادغام اجزای به هم پیوسته، پردازش می‌کنند.

روش کار به این صورت است که ابتدا با استفاده از تبدیل PCA تعداد ابعاد تصویر به سه بعد کاهش یافته، سپس APها و گشتاورهای هر مولفه استخراج شده و در یک مجموعه به دنبال هم قرار گرفته‌اند و در نهایت طبقه‌بندی با SVM انجام شده است.

پیاده‌سازی روش پیشنهادی به کمک دو داده‌ی ابرطیفی دانشگاه پابویا (تصویری از یک منطقه‌ی شهری) و ایندیانا (تصویری از یک منطقه‌ی کشاورزی) نشان می‌دهد که روش پیشنهادی، برای داده‌های ابرطیفی مناطق شهری مناسب‌تر است. در این حالت، دقت طبقه‌بندی، نسبت به حالتی که تنها از پروفایل‌های توصیفی استفاده شده، بیشتر است.

سنجش از دور شامل بدست آوردن اطلاعات از یک شی یا صحنه بدون هیچ گونه تماس فیزیکی مستقیم است. داده‌های سنجش از دور به دلیل دقت بالا، تنوع زیاد، دسترسی آسان و هزینه نسبتاً کم، نخستین عامل در مطالعه سطح زمین محسوب می‌شوند. طبقه‌بندی این تصاویر، یعنی تخصیص برجسب به هر پیکسل موجود در تصویر، به منظور تولید نقشه‌های پوشش زمین، حوزه‌ی تحقیقاتی مهمی در سنجش از دور به شمار می‌رود.

گاهی برای آنالیز تصاویر سنجش از دور ابرطیفی، تنها اطلاعات طیفی کافی و یا مفید نیست بنابراین الگوریتم‌های پردازش تصویر جهت به کارگیری اطلاعات مکانی، توسعه داده شده است.

در این مقاله، طبقه‌بندی تصاویر سنجش از دور ابرطیفی، با استفاده از ترکیب پروفایل‌های توصیفی با گشتاورهای تغییر ناپذیر، انجام شده است. پروفایل‌های توصیفی، ویژگی‌های مکانی اشیا تصویر

مراجع

- [1] M.-T. Pham, S. Lefèvre, E. Aptoula, and L. Bruzzone, "Recent developments from attribute profiles for remote sensing image classification," arXiv Prepr. arXiv1803.10036, 2018.
- [2] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological Attribute Profiles for the Analysis of Very High Resolution Images," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 10, pp. 3747–3762, 2010.
- [3] M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Self-dual attribute profiles for the analysis of remote sensing images," in International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing, 2011, pp. 320–330.
- [4] B. Demir and L. Bruzzone, "Histogram-based attribute profiles for classification of very high resolution remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 4, pp. 2096–2107, 2015.
- [5] M.-T. Pham, S. Lefèvre, and E. Aptoula, "Local feature-based attribute profiles for optical remote sensing image classification," IEEE Trans. Geosci. Remote Sens., vol. 56, no. 2, pp. 1199–1212, 2017.
- [6] W. Liao et al., "Taking Optimal Advantage of Fine Spatial Resolution: Promoting partial image reconstruction for the morphological analysis of very-high-resolution images," IEEE Geosci. Remote Sens. Mag., vol. 5, no. 2, pp. 8–28, 2017.
- [7] C. Vision, I. Understanding, and A. No, "Attribute Openings, Thinnings, and Granulometries," vol. 64, no. 3, pp. 377–389, 1996.
- [8] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectral-spatial classification techniques based on attribute profiles," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 5, pp. 2335–2353, 2014.
- [9] F. Mirzapour and H. Ghassemian, "Moment-based feature extraction from high spatial resolution hyperspectral images," Int. J. Remote Sens., vol. 37, no. 6, pp. 1349–1361, 2016.
- [10] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, p. 27, 2011.
- [11] Universidad-del-País-Vasco. 2014. Hyperspectral Remote Sensing Scenes. www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes



افزایش دایره واژگان و بهبود عملکرد واژه‌نامه احساسی Senti-N-Gram با استفاده از پسوندها و پیشوندهای منفی‌ساز

حامد زرگری^۱، دانشجوی دکتری، مرتضی زاهدی^۱، استادیار

۱- دانشکده مهندسی برق و کامپیوتر - دانشگاه صنعتی شاهرود - شاهرود

{h.zargari, zahedi}@shahroodut.ac.ir

چکیده

با رشد روز افزون رسانه‌های دیجیتال، آنالیز احساسات به یکی از موضوعات تحقیقاتی پر کاربرد تبدیل شده است. کاربردهای گسترده تجزیه و تحلیل احساسات در حوزه‌های متنوع، تجاری، سیاسی، اجتماعی، بهداشتی و امنیتی محققان را به سمت طراحی و پیاده‌سازی سیستم‌های تجزیه و تحلیل احساسات، که دارای دقت بالایی بوده و همچنین قادر به تجزیه و تحلیل برخط اطلاعات باشند سوق داده است. واژه‌نامه‌های احساسی یکی از مهمترین و کاربردترین ابزارها در حوزه، تجزیه و تحلیل احساسات محسوب می‌شوند. یکی از مشکلات واژه‌نامه‌های احساسی، عدم دقت این قبیل از واژه‌نامه‌ها در شناسایی و تشخیص کلمات و عبارات منفی در مقایسه با کلمات و عبارات مثبت است. با توجه به این امر در این مقاله با ارائه روشی ابتکاری اقدام به توسعه واژه‌نامه احساسی با استفاده از پیشوندها و پسوندهای منفی‌ساز زبان انگلیسی شده است. در روش پیشنهادی در این مقاله پسوندها و پیشوندهای منفی‌ساز به واژه‌نامه احساسی Senti-N-Gram اضافه و این کلمات جدید امتیازدهی می‌شوند. در نهایت نتایج شبیه سازی نشان داد که این واژه‌نامه توسعه یافته نسبت به واژه‌نامه موجود در شرایط یکسان در پایگاه داده‌های متداول در حدود ۲٪ موجب افزایش دقت تشخیص گردیده است.

کلمات کلیدی

تشخیص احساس، واژه‌نامه احساسی، منفی‌سازها، Senti-N-Gram

۱- مقدمه

همواره در فرآیند تصمیم‌سازی مهم بوده است [5]، لذا نظراتی که کاربران محصولات در مورد یک محصول ارائه می‌دهند از دیدگاه تولیدکنندگان برای ارزیابی محصول و پیش‌بینی عکس‌العمل‌های مشتریان نسبت به ورود محصولات جدید، بسیار پراهمیت و مفید است. از دیدگاه تجاری تجزیه و تحلیل احساسات امکان ارائه مشاوره و توصیه‌های برخط را برای هر دو گروه مشتری و بازرگان فراهم می‌نماید [6]. علاوه بر این امروزه تحلیل احساسات سیاسی، روشی متداول برای تحلیل عقاید مردم در قبال کارزار سیاسی است [7]. طبقه بندی احساسات یک فرایند وابسته به دامنه و حساس به متن محسوب می‌شود [5,8,9]. این وابستگی در موضوعاتی همچون دامنه، موضوع، دوره زمانی و سبک زبان نشان داده شده است [10]. تجزیه و تحلیل احساسات در حوزه‌های وسیعی از دامنه‌ها همچون بررسی فیلم [11-13]، بررسی تدریس [14-17]، بررسی محصول [18,19]، یادگیری الکترونیکی [20,21]، بررسی هتل [22-24] و بسیاری دیگر از موارد مورد استفاده قرار گرفته است.

اساس تجزیه و تحلیل احساسات بر مبنای تعیین قطبیت یک متن بنا شده است [25,26]. تجزیه و تحلیل احساسات عمدتاً در سطح

یکی از آرزوهای بشر ساخت رباتی بوده که بتواند بفهمد سایر انسانها چه فکری می‌نمایند و برای رسیدن به این آرزو تلاشهای بسیاری صورت پذیرفته است. تحلیل احساسات به عنوان "وظیفه یافتن نظرات نویسندگان درباره اشخاص خاص" تعریف شده است [1]. ظهور آنالیز احساسات بشکل مدرن کنونی به اواسط سال ۲۰۰۰ و از زمان پیدایش بررسی‌های آنلاین و رسانه‌های اجتماعی بر می‌گردد [2]. اهمیت آنالیز احساسات با معرفی ۲,۰ web که کاربران را قادر به بیان نظرات و دیدگاه‌هایشان درباره موضوعات مختلف از طریق روش‌هایی همچون انجمن‌های مصرف‌کنندگان^۱، رسانه‌های اجتماعی^۲ و درگاه‌های خدمات الکترونیک^۳ می‌نماید به مراتب نمایان‌تر گردیده است [3,4]. با توجه به این نکته که آنچه دیگران فکر می‌کنند

¹ Consumer forums

² Social media

³ E-governance

در این تحقیق یک روش جدید با هدف توسعه واژه‌نامه احساسی Senti-N-Gram پیشنهاد شده است. در روش پیشنهادی با استفاده از تعدادی از الگوهای پیشوندی و پسوندی منفی‌کننده، اقدام به توسعه مجموعه لغات واژه‌نامه پایه شده و کارایی واژه‌نامه جدید با استفاده از یک مجموعه از متنهای طبقه بندی شده بررسی شده است.

در ادامه ساختار مقاله به شرح زیر است. در بخش دوم روش پیشنهادی جهت توسعه واژه‌نامه احساسی Senti-N-Gram بیان شده و بخش سوم به ارائه نتایج حاصل از ارزیابی پرداخته است. در نهایت قسمت آخر مقاله نیز به نتیجه گیری و خاتمه اختصاص یافته است.

۲- روش پیشنهادی

رویکرد پیشنهادی در توسعه این واژه‌نامه یک رویکرد اتوماتیک محسوب می‌شود. در این پژوهش از واژه‌نامه Senti-N-Gram به عنوان واژه‌نامه احساسی پایه استفاده شده است. فرآیند توسعه و امتیازدهی واژه‌نامه توسعه یافته بر مبنای قواعد نحوی و با توجه به کلمات احساسی و امتیازات موجود در خود واژه‌نامه صورت پذیرفته است. با توجه به توضیحات فوق این بخش به سه زیربخش مجزا تقسیم شده است. ابتدا تعدادی از مهمترین واژه‌نامه‌های موجود معرفی شده و نقاط قوت و ضعف هرکدام بیان خواهد شد. سپس پایگاه داده‌های موجود مورد مطالعه قرار گرفته و در نهایت روش پیشنهادی بیان می‌شود.

۲-۱- واژه‌نامه‌ها

اولین واژه‌نامه را می‌توان واژه‌نامه [58] General Inquire دانست که در اواخر دهه ۱۹۶۰ توسط فلیپ استون و همکاران در هاروارد به صورت دستی و باهدف کشف و ردیابی موضوعات روانشناسی تولید شده است. این واژه‌نامه یکی از قدیمی ترین واژه‌نامه‌های تولید شده بصورت دستی است که هنوز هم برای تولید و توسعه واژه‌نامه‌های جدید مورد استفاده قرار می‌گیرد. این واژه‌نامه حاوی ۸۶۴۱ کلمه احساسی است که برخی از کلمات جنبه‌های احساسی چندگانه دارند. مجموع کل کلمات احساسی این واژه‌نامه ۱۱۷۸۹ کلمه بوده که در ۱۸۲ دسته‌ی مرتبط با حوزه مورد استفاده، قطبیت، طبقات اجتماعی و غیره گروه‌بندی شده‌اند.

واژه‌نامه بعدی واژه‌نامه [59] LIWC است که برای مطالعه اجزای هیجانی، شناختی، ساختاری در نمونه‌های مختلف متون طراحی شده است. ایده‌ی اولیه‌ی طراحی این واژه‌نامه، شناسایی و گردآوری گروهی از کلمات دارای پایه احساسی و ابعاد شناختی است که غالباً در روانشناسی اجتماعی، بهداشتی و شخصیتی مورد بررسی قرار می‌گیرد. این واژه‌نامه شامل تقریباً ۴۵۰۰ کلمه و ریشه‌ی کلمه است که در ۱۰۰ دسته طبقه‌بندی شده‌اند

سند، جمله و جنبه مورد بررسی قرار گرفته است [27,2829]. در تجزیه و تحلیل احساسات در سطح سند جهت گیری احساسی کلی سند مورد توجه قرار می‌گیرد. این روش تنها در هنگامیکه تمامی سند در رابطه با یک ویژگی باشد مطلوب است [27]. در تجزیه و تحلیل احساسات در سطح جمله تنها قطبیت جملات ذهنی مورد بررسی قرار می‌گیرد. در نهایت تجزیه و تحلیل احساسات در سطح ویژگی، به دنبال استخراج جفت‌های هدف و احساس مرتبط با آن هدف است [30,31].

یکی از مهمترین موضوعات در حوزه تجزیه و تحلیل احساسات تعیین مثبت بودن یا منفی بودن یک جمله یا متن است [32]. دسته متداولی از تحقیقات از روشهای مبتنی بر واژه‌نامه‌های احساسی برای مشخص نمودن جهت گیری قطبی مثبت و یا منفی کلمات و عبارات استفاده می‌نمایند [33-35].

در جمع بندی کارهای انجام شده می‌توان کارهای انجام شده در ساخت واژه‌نامه را به ۵ زیر گروه تقسیم کرد.

- واژه‌نامه‌های ساخته شده بصورت کاملاً دستی [36-38]
- واژه‌نامه‌های ساخته شده با استفاده از روش بوت استرپ و کلمات بذر [30,39,40]
- واژه‌نامه‌های ساخته شده با روش انتقال از یک حوزه به حوزه جدید [41-43]
- واژه‌نامه‌های ساخته شده به صورت نیمه اتوماتیک [44]
- واژه‌نامه‌های ساخته شده با استفاده از الگوریتم های مبتنی بر یادگیری ماشین [43,45,46]

بعضاً در مواردی نیز از قابلیت‌های حوزه‌های مختلف به صورت همزمان استفاده شده است. نتایج حاصل از آزمایش برخی از این واژه‌نامه‌ها [44,47-49] نشان دهنده عملکرد مطلوبتر آنها در تشخیص عبارات احساسی مثبت و نتایج ضعیفتر آنها در حوزه تشخیص عبارات منفی است. با توجه به این موضوع که، منفی‌کننده‌ها یکی از مهمترین پدیده‌های زبانی موثر در تغییر قطبیت کلمات احساسی محسوب می‌شوند. تعداد بسیار زیادی از محققان [50-54] نشان داده‌اند که توجه به ساختارهای زبانی همچون منفی‌کننده‌ها باعث افزایش کارایی سیستم‌های تجزیه و تحلیل احساسات می‌گردند.

برخی تحقیقات با استفاده از روشهای متفاوتی همچون توسعه دامنه لغات منفی موجود در در واژه‌نامه [49,55] افزایش ضریب تاثیر کلمات منفی [47] اعمال تابع جریمه در نتایج حاصل [44] و همچنین توجه به عبارات طعنه و بسیاری دیگر از موارد سعی در بهبود تشخیص جملات منفی نموده‌اند.

از آنجاییکه فرآیند تولید واژه‌نامه‌های دستی فرآیندی بسیار وقت گیر و پرهزینه محسوب می‌شود لذا بسیاری از واژه‌نامه‌های پیشنهادی [39,44,47,48,56,57] با استفاده از واژه‌نامه‌های تولید شده پیشین و همچنین بهره گیری از روابط معنایی و نحوی سعی در تولید واژه‌نامه‌های جدید نموده‌اند.

این پایگاه داده توسط دانشگاه استنفورد در سال ۲۰۰۴ برای زبان انگلیسی گردآوری و برچسب گذاری شده است [63]. در ادامه یک متن نمونه از حوزه کتاب با برچسب منفی از این پایگاه داده در شکل ۱

I have read all of Grisham's books and this was by far the most boring one! The story is about a lawyer who gives up his high paying job to help the homeless in an office staffed by two other people. There is no mystery to the story. Grisham's books are normally mysterious, can't wait to finish it, a "who done it" style. This book does not have any surprises, its dull. I expected a book like other books (f. i. Pelican Brief etc) he has written. It definitely was not! I would definitely not recommend it to anybody

آورده شده است.

۲-۳- واژه‌نامه توسعه یافته

شکل 44 متن نمونه از حوزه کتاب با برچسب منفی از پایگاه داده taboda در این روش واژه‌نامه Senti-N-Gram بعنوان واژه‌نامه پایه مورد استفاده قرار گرفت. این واژه‌نامه یک واژه‌نامه نیمه دستی بوده که بر مبنای واژه‌نامه [48] VADER بنا شده و از عبارات احساسی Unigram، Bigram و Trigram پشتیبانی می‌نماید. عبارات Bigram و Trigram این واژه‌نامه ترکیبی از کلمات احساسی موجود در واژه‌نامه VADER و تشدید کننده‌های موجود در [64] و منفی-سازها بر اساس استاندارد دانشگاه کمبریج⁴ است.

جدول 1: پیشوندها و پسوندهای منفی ساز

الگوی پیشوندها و پسوندها		تعداد	مثال
W1	W2		
X	disX	۳۹۹	honest-dishonest
X	imX	۱۵۹	possible-impossible
X	inX	۷۱۲	consistent-inconsistent
X	malX	۷۴	adroit-maladroit
X	misX	۱۸۲	fortune-misfortune
X	nonX	۱۵۶۳	sense-nonsense
X	unX	۲۱۲۲	happy-unhappy
X	Xless	۳۳۳	gut-gutless
lX	illX	۲۵	legal-illegal
rX	irX	۳۰	responsible-irresponsible
Xless	Xful	۲۷	harmless-harmfu

لازم به ذکر است منفی‌سازهای استفاده شده برای ایجاد واژه‌نامه Senti-N-Gram پسوند و پیشوند نیستند. بر اساس توضیح فوق، در این تحقیق از مجموعه‌ای از ۱۱ الگوی پسوندی و پیشوندی

در واژه‌نامه بعدی [60] با طراحی یک بازی گروهی و گرفتن نظریات افراد درباره کلمات احساسی، نسبت به تعیین قطبیت کلمات احساسی با اجماع گروه، دسته بندی انجام شده است. بازی با استفاده از برج‌های بابل طراحی شده و سعی کرده یک پایگاه داده مناسب از لغات احساسی نیز جمع‌آوری کند. واژه‌نامه بعدی [48] VADER می‌باشد که در سال ۲۰۱۴ ایجاد شده است. این واژه‌نامه کاملاً دستی است و با استفاده از روش جمع‌سپاری سعی کرده نظرات افراد را راجع به بار احساسی کلمات جمع‌آوری کند. واژه‌نامه دیگر که در سال ۲۰۱۸ طراحی شده، [61] WKWSC است که با استفاده از گردآوری لغات و گسترش آنها یک واژه‌نامه مبتنی بر امتیاز برای بار احساسی کلمات جمع‌آوری کرده و در نهایت درصد درستی واژه‌نامه خود را بر روی پایگاه داده‌های مختلف با ۶ واژه‌نامه مهم در این حوزه مقایسه کرده است.

یکی از پرکاربردترین واژه‌نامه‌ها واژه‌نامه SentiWordnet است که در سال ۲۰۰۶ پیشنهاد شده [39] و در طراحی آن از کلمات بذر و روابط معنایی استفاده شده است. نسخه‌ی ۱ این واژه‌نامه با استفاده از یک الگوریتم یادگیرنده ناظر در چهار مرحله تشکیل شده که به‌طور مختصر شامل توسعه مجموعه بذر، آموزش طبقه‌بند، طبقه‌بندی Synsetها و در نهایت ترکیب طبقه‌بندها است. SentiWordnet در چند مرحله به روز رسانی نیز شده است که بروز رسانی‌های مختلف آن با ورژن‌های ۲ و ۳ در [34,62] آورده شده است.

بر اساس مرور انجام شده می‌توان دید، واژه‌نامه‌های مختلف با رویکردهای متفاوتی، هم به صورت کاملاً دستی، هم به صورت نیمه دستی و هم به صورت تمام اتوماتیک برای تشخیص احساس بر اساس امتیازدهی‌های مختلف ساخته شده که در این میان واژه‌نامه‌های عمومی به خاطر گستره بسیار بالا و عدم معطوف بودن به یک حوزه خاص معمولاً در مجموع درصدهای درستی کمتری نسبت به واژه‌نامه‌های تخصصی در حوزه مربوطه دارند. علاوه بر آن اکثر واژه‌نامه‌های موجود، غالباً توسعه یافته چند واژه‌نامه پایه بوده و یا بر اساس الگوریتم‌های یادگیری ماشین، در پایگاه داده‌های مختلف و یا با توجه به روابط معنایی و نحوی طراحی و توسعه داده شده‌اند که در روش پیشنهادی در این مقاله نیز ما با بهره‌گیری از روابط معنایی و نحوی، سعی در ارتقاء دایره واژگان و دقت تشخیص یکی از بهترین واژه‌نامه‌های موجود در این حوزه داشته‌ایم.

۲-۲- پایگاه داده

برای پیاده‌سازی روش پیشنهادی باید کارایی آن بر روی یک پایگاه داده استاندارد مورد بررسی و با مقالات مرجع بررسی شده بر روی همان پایگاه داده، مقایسه گردد. پایگاه داده‌ای که در این تحقیق مورد استناد قرار گرفته پایگاه داده taboda است که به صورت رایگان برای موارد علمی قابل دانلود است. پایگاه داده taboda در خود مجموعه‌ای شامل ۴۰۰ متن از ۸ حوزه مختلف را جای داده است. این ۴۰۰ متن، به صورت دستی در دو گروه منفی و مثبت برچسب‌گذاری شده‌اند. این پایگاه داده به عنوان یک معیار مقایسه اولیه برای اکثر روش‌ها در حیطه واژه‌نامه‌های عمومی مورد استناد قرار می‌گیرد.

⁴ <https://dictionary.cambridge.org/grammar/british-grammar/questions-and-negative-sentences/negation>

کارایی داشته‌ایم اما در اکثر موارد واژه‌نامه توسعه داده شده نسبت به واژه‌نامه پایه کارایی بهتری از خود نشان می‌دهد.

جدول ۲: تصمیم‌گیری بر اساس کل سند، امتیازات مربوط به واژه‌نامه Senti-N-Gram

	TP	FP	TN	FN	P	R	F	A
BOOKS	23	14	11	2	0.622	0.92	0.742	0.68
CARS	25	19	6	0	0.568	1	0.725	0.62
COMPUTERS	25	16	9	0	0.61	1	0.758	0.68
COOKWARE	25	22	3	0	0.532	1	0.694	0.56
HOTELS	25	19	6	0	0.568	1	0.725	0.62
MOVIES	24	18	7	1	0.571	0.96	0.716	0.62
MUSIC	23	21	4	2	0.523	0.92	0.667	0.54
PHONES	22	13	12	3	0.629	0.88	0.733	0.68

جدول ۳: تصمیم‌گیری بر اساس کل سند، امتیازات مربوط به روش پیشنهادی

	TP	FP	TN	FN	P	R	F	A
BOOKS	23	14	11	2	0.622	0.92	0.742	0.68
CARS	25	17	8	0	0.595	1	0.746	0.66
COMPUTERS	25	13	12	0	0.658	1	0.794	0.74
COOKWARE	25	22	3	0	0.532	1	0.694	0.56
HOTELS	25	19	6	0	0.568	1	0.725	0.62
MOVIES	23	18	7	2	0.561	0.92	0.697	0.6
MUSIC	23	17	8	2	0.575	0.92	0.708	0.62
PHONES	23	15	10	2	0.605	0.92	0.73	0.66

با توجه به ستون صحت (A)، واژه‌نامه توسعه یافته در هر یک از ۸ حوزه به طور متوسط چیزی در حدود ۱,۷٪ نسبت به واژه‌نامه پایه کارایی مناسب‌تری از خود نشان داده است. علاوه بر آن در دقت و معیار F نیز روش پیشنهادی از خود بهبود نشان داده است. در مجموع روش پیشنهادی توانسته نسبت به روش قبلی در حدود ۱,۷ درصد صحت را بهبود ببخشد که با توجه به درصدهای متداول در حوزه واژه‌نامه‌های عمومی مناسب به نظر می‌رسد.

برای مقایسه بهتر در شکل ۲ جوابهای روش پیشنهادی پایه بر روی پایگاه داده taboda در هر یک از هشت حوزه برای معیار صحت به تفکیک آورده شده تا امکان مقایسه دقیق حوزه به حوزه واژه‌نامه توسعه داده شده و واژه‌نامه پایه با یکدیگر فراهم شود. همانطور که در شکل ۲ می‌توان دید در تمامی حوزه‌ها غیر از دو حوزه سینمایی و تلفن روش پیشنهادی معادل یا بهتر از روش پایه بوده و فقط در این دو حوزه افت کارایی اندکی داشتیم که ممکن است بدلیل لغات تخصصی خاص موجود در این دو حوزه باشد. با توجه به ارتقاء انجام شده در سایر حوزه‌ها در مجموع کارایی واژه‌نامه توسعه داده شده نسبت به واژه‌نامه پایه بهتر بوده است.

منفی‌ساز [32,33] برای توسعه واژه‌نامه Senti-N-Gram و بالا بردن کارایی آن استفاده شده است. جدول ۱ لیست پسوندها و پیشوندهای منفی‌کننده مورد استفاده در این پیاده‌سازی را به همراه مثالی از هر یک نشان می‌دهد. تعداد کل کلمات ایجاد شده با پسوند یا پیشوند نیز در ستون مربوطه درج شده است. با این کار و حذف ترکیبات نامصطلح بر اساس مجموعه لغات Google Ngram در مجموع ۵۶۲۶ ترکیب جدید به واژه‌نامه Senti-N-Gram اضافه شده است.

به طور کلی ۴ روش برای برخورد با منفی‌کننده‌ها وجود دارد.

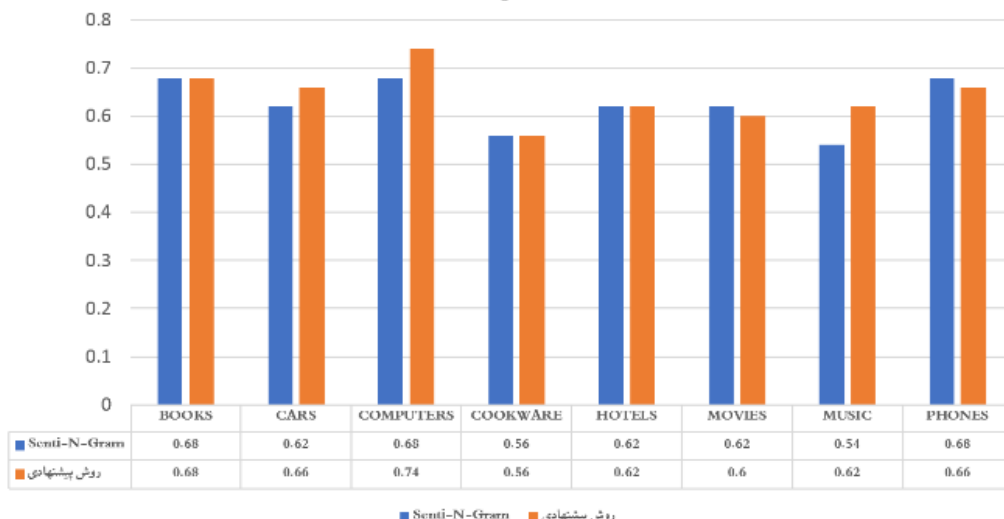
- معکوس کردن امتیاز کلمه احساسی پس از اعمال منفی‌کننده [33,38,48,65,66]
- شیفت ثابت امتیاز کلمه احساسی پس از اعمال منفی‌کننده [47,67,68]
- محاسبه مستقل امتیاز برای هر کلمه احساسی پس از اعمال منفی‌کننده [44,69,70]
- روشهایی مبتنی بر یادگیری ماشینی [71]

در این تحقیق برای تخصیص امتیاز به لغات اضافه شده از رویکرد اول استفاده شد. برای نمونه اگر امتیاز کلمه احساسی عدد ۲/۵ باشد با اضافه شدن منفی‌ساز به ابتدا یا انتهای آن، امتیاز آن برابر ۲/۵- لحاظ شد. به این ترتیب در حقیقت برای اختصاص امتیاز از امتیازات دستی داده شده در واژه‌نامه پایه کمک گرفته شد و هیچ نیازی به پروسه اولیه آموزش در روش پیشنهادی برای اختصاص امتیاز به لغات اضافه شده نبود.

۳- نتایج

در این بخش نتایج اعمال واژه‌نامه توسعه یافته و واژه‌نامه پایه بر روی پایگاه داده taboda با یکدیگر مقایسه شده تا مشخص شود که کلمات احساسی اضافه شده به واژه‌نامه پایه تا چه حد می‌تواند در بهبود کارایی آن موثر باشند. بر اساس نتایج جداول ۱ و ۲ می‌توان دید که در اکثر حوزه‌ها روش پیشنهادی بر اساس واژه‌نامه توسعه داده شده در شرایط یکسان توانسته کارایی مناسب‌تری نسبت به واژه‌نامه Senti-N-Gram از خود نشان دهد. با توجه به عمومی بودن هر دو واژه‌نامه و عدم انجام پروسه مجزا برای امتیازدهی به کلمات احساسی اضافه شده، می‌توان مشاهده کرد اضافه شدن کلمات احساسی منفی حاصل از ترکیب پسوندها و پیشوندهای منفی‌ساز به واژه‌نامه توانسته در غالب موارد، کارایی واژه‌نامه را به صورت معنی داری بهبود دهد. با توجه به اینکه روش پیشنهادی مستقل از واژه‌نامه و بر اساس پسوند و پیشوند منفی‌ساز طراحی شده، می‌توان به راحتی از این روش برای توسعه کلیه واژه‌نامه‌های عمومی موجود استفاده نمود.

با مقایسه جداول ۲ و ۳ می‌توان دید ستون‌های مربوط به TP و FN در مجموع ثابت است و لذا واژه‌نامه توسعه داده شده در تشخیص متون مثبت مشابه واژه‌نامه پایه عمل کرده است. در مورد FP و TN نیز روش پیشنهادی در مجموع نتایج بهتری داشته که بیانگر بالا رفتن کارایی واژه‌نامه توسعه داده شده در تشخیص متون منفی است. اگرچه در برخی حوزه‌ها مانند حوزه PHONES کمی افت



شکل 45: میزان درصد درستی بر حسب حوزه‌های مختلف در واژه‌نامه توسعه داده شده و پایه

۴- نتیجه گیری و خاتمه

در این تحقیق روشی برای ارتقاء کارایی واژه‌نامه‌های عمومی موجود با اضافه کردن پسوند و پیشوندهای منفی ساز به کلمات احساسی موجود در واژه‌نامه پیشنهاد شد. بر اساس روش پیشنهادی با اضافه کردن پسوندها و پیشوندهای منفی ساز به کلمات احساسی موجود و منفی کردن امتیاز موجود در واژه‌نامه برای کلمه احساسی اولیه می‌توان یک واژه‌نامه توسعه یافته مبتنی بر واژه‌نامه پایه ساخت.

روش پیشنهادی با استفاده از ۱۱ پسوند و پیشوند منفی ساز مرسوم انگلیسی بر روی واژه‌نامه Senti-N-Gram پیاده و در نهایت ارتقاء کارایی مشهود آن نسبت به واژه‌نامه اولیه بر روی یک پایگاه داده استاندارد عمومی نشان داده شد.

از نقاط قوت روش پیشنهادی به سهولت ارتقا و گسترش لغات و تعمیم پذیری آن بر روی کلیه واژه‌نامه‌های عمومی و حوزه محور می‌توان اشاره کرد. عدم نیاز به آموزش و محاسبه امتیاز برای کلمات احساسی جدید تولیدی نیز می‌تواند از جمله امتیازات روش پیشنهادی محسوب گردد. به عنوان پیشنهاد برای ادامه کار، ارتقاء دایره واژگان بر اساس پسوندها و پیشوندهای منفی ساز دیگر موجود در زبان انگلیسی و استفاده از پسوندها و پیشوندهای تشدید کننده و تضعیف کننده را می‌توان مدنظر قرار داد.

مراجع

- doi:10. 1007/s10115-018-1236-4
- [7]. Bose, R. , Dey, R. K. , Roy, S. & Sarddar, D. Analyzing Political Sentiment Using Twitter Data. in 427–436 (Springer, Singapore, 2019). doi:10. 1007/978-981-13-1747-7_41
- [8]. Turney, P. D. & D. , P. Thumbs up or thumbs down?: sentiment classification using machine learning techniques. *Proc. 40th Annu. Meet. Assoc. Comput. Linguist. - ACL '02* 417 (2001). doi:10. 3115/1073083. 1073153
- [9]. Aue, A. & Gamon, M. Customizing Sentiment Classifiers to New Domains: a Case Study. (2005).
- [10]. Read, J. Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification. 43–48 (2005).
- [11]. Singh, V. K. , Piryani, R. & Uddin, A. Sentiment Analysis of Movie Reviews A new Feature-based Heuristic for Aspect-level Sentiment Classification.
- [12]. Chakraborty, K. , Bhattacharyya, S. , Bag, R. & Hassanien, A. A. *Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Techniques. Social Network Analytics* (Elsevier Inc. , 2019). doi:10. 1016/b978-0-12-815458-8. 00007-4
- [13]. Tripathy, A. , Agrawal, A. & Rath, S. K. Classification of sentiment reviews using n-gram machine learning approach. *Expert Syst. Appl.* **57**, 117–126 (2016).
- [14]. Aung, K. Z. & Myo, N. N. Sentiment analysis of students' comment using lexicon based approach. *Proc. - 16th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2017* 149–154 (2017). doi:10. 1109/ICIS. 2017. 7959985
- [15]. Asghar, M. Z. , Ullah, I. , Shamshirband, S. , Kundi, F. M. & Habib, A. Fuzzy-Based Sentiment Analysis System for Analyzing Student Feedback and Satisfaction. (2019). doi:10. 20944/PREPRINTS201907. 0006. V1
- [16]. Sagum, R. A. , Vera, J. G. M. De, Lansang, P. J. S. , Narciso, D. S. R. & Respeto, J. K. Application of Language Modelling in Sentiment Analysis for Faculty Comment Evaluation. **1**, (2015).
- [17]. Rani, S. & Kumar, P. A Sentiment Analysis System to Improve Teaching and Learning. *Computer (Long. Beach. Calif.)* **50**, 36–43 (2017).
- [18]. Mirtalaie, M. A. , Hussain, O. K. , Chang, E. & Hussain, F. K. Extracting sentiment knowledge from pros/cons product reviews: Discovering features along with the polarity strength of their associated opinions. *Expert Syst. Appl.* **114**, 267–288 (2018).
- [19]. Avanc, L. V & Nunes, M. G. V. Lexicon-based Sentiment Analysis for Reviews of Products in Brazilian Portuguese. 10–14 (2014). doi:10. 1109/BRACIS. 2014. 57
- [20]. Ortigosa, A. , Martín, J. M. & Carro, R. M. Sentiment analysis in Facebook and its application to e-learning. *Comput. Human Behav.* **31**, 527–541 (2014).
- [21]. Kechaou, Z. , Ben Ammar, M. & Alimi, A. Improving e-learning with sentiment analysis of users' opinions. in *2011 IEEE Global Engineering Education Conference (EDUCON)* 1032–1038 (IEEE, 2011). doi:10. 1109/EDUCON. 2011. 5773275
- [22]. Hu, Y. -H. H. , Chen, Y. -L. L. & Chou, H. -L. L. Opinion mining from online hotel reviews – A text summarization approach. *Inf.*
- [1]. Feldman, R. Techniques and applications for sentiment analysis. *Commun. ACM* **56**, 82 (2013).
- [2]. Mäntylä, M. V *et al.* The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Comput. Sci. Rev.* **27**, 16–32 (2018).
- [3]. Speriou, M. , Upadhyay, S. & Baldrige, J. Twitter Polarity Classification with Label Propagation. *dl. acm. org* 53–63 (2011).
- [4]. Tang, D. , Wei, F. , Qin, B. , Zhou, M. & Liu, T. Building large-scale twitter-specific sentiment lexicon : A representation learning approach. *COLING 2014 - 25th Int. Conf. Comput. Linguist. Proc. COLING 2014 Tech. Pap.* 172–182 (2014).
- [5]. Pang, B. & Lee, L. *Opinion Mining and Sentiment Analysis: Foundations and Trends in Information Retrieval.* **2**, (2008).
- [6]. Yue, L. , Chen, W. , Li, X. , Zuo, W. & Yin, M. A survey of sentiment analysis in social media. *Knowl. Inf. Syst.* 1–47 (2018).

- [47].Taboada, M. , Brooke, J. , Tofloski, M. , Voll, K. & Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **37**, 267–307 (2011).
- [48].Hutto, C. J. C. , And, E. G. -E. international A. conference on weblogs, 2014, U. & Gilbert, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. in *ICWSM 216–225* (2014).
- [49].Thelwall, M. , Buckley, K. , Paltoglou, G. , Cai, D. & Kappas, A. Sentiment strength detection in short informal text. *J. Am. Soc. Inf. Sci. Technol.* **61**, 2544–2558 (2010).
- [50].Benamara, F. , Chardon, B. , Mathieu, Y. , Popescu, V. & Asher, N. How do Negation and Modality Impact on Opinions? - W12-3802. pdf. *ExProm 2012 -- Proc. Work. Extra-Propositional Asp. Mean. Comput. Linguist.* 10–18 (2012).
- [51].Wiegand, M. , Balahur, A. , Roth, B. , Klakow, D. & Montoyo, A. A Survey on the Role of Negation in Sentiment Analysis. *Proc. Work. Negation Specul. Nat. Lang. Process.* 60–68 (2010).
- [52].Councill, I. G. , McDonald, R. & Velikovich, L. What's great and what's not: learning to classify the scope of negation for improved sentiment analysis. *Proc. ACL Work. Negation Specul. Nat. Lang. Process. Uppsala Sweden* 51–59 (2010).
- [53].Jia, L. , Yu, C. & Meng, W. The effect of negation on sentiment analysis and retrieval effectiveness. 1827 (2009). doi:10.1145/1645953.1646241
- [54].Barnes, J. , Veldal, E. & Øvrelid, L. Improving Sentiment Analysis with Multi-task Learning of Negation. (2019).
- [55].Neviarouskaya, A. , Prendinger, H. & Ishizuka, M. SentiFul : A Lexicon for Sentiment Analysis. **2**, 22–36 (2011).
- [56].Nielsen, F. Å. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *CEUR Workshop Proc.* **718**, 93–98 (2011).
- [57].Dey, A. , Jenamani, M. & Thakkar, J. J. Lexical TF-IDF: An n-gram Feature Space for Cross-Domain Classification of Sentiment Reviews. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **10597 LNCS**, 380–386 (2017).
- [58].Hartman, J. J. , Stone, P. J. , Dunphy, D. C. , Smith, M. S. & Ogilvia, D. M. The General Inquirer: A Computer Approach to Content Analysis. *Am. Sociol. Rev.* **32**, 859 (1967).
- [59].Pennebaker, J. W. , Francis, M. E. & Booth, R. J. Linguistic Inquiry and Word Count LIWC 2001. (2001).
- [60].Hong, Y. , Kwak, H. , Baek, Y. & Moon, S. Tower of babel: A crowdsourcing game building sentiment lexicons for resource-scarce languages. *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web* 549–556 (2013). doi:10.1145/2487788.2487993
- [61].Khoo, C. S. & Johnkhan, S. B. Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *J. Inf. Sci.* **44**, 491–511 (2018).
- [62].Esuli, A. Automatic generation of lexical resources for opinion mining. *ACM SIGIR Forum* **42**, 105 (2010).
- [63].Taboada, M. & Grieve, J. Analyzing Appraisal Automatically. *AAAI Spring Symp. - Tech. Rep.* **2004**, 158–161 (2004).
- [64].Brooke, J. A SEMANTIC APPROACH TO AUTOMATED TEXT SENTIMENT ANALYSIS. *Analysis* **26**, (2009).
- [65].Polanyi, L. & Zaenen, A. Contextual valence shifters. in *AAAI Spring Symposium - Technical Report SS-04-07*, 106–111 (Springer-Verlag, 2005).
- [66].Benamara, F. , Cesarano, C. , Picariello, A. , Reforgiato, D. & Subrahmanian, V. S. Sentiment analysis: Adjectives and adverbs are better than adjectives alone. *ICWSM 2007 - Int. Conf. Weblogs Soc. Media* (2007).
- [67].Satthar, F. S. Modelling SO-CAL in an inheritance-based sentiment analysis framework. *OpenAccess Ser. Informatics* **49**, 46–53 (2015).
- [68].Liu, J. & Seneff, S. Review sentiment scoring via a parse-and-paraphrase paradigm. 161–169 (2009). doi:10.3115/1699510.1699532
- [69].Kiritchenko, S. , Zhu, X. & Mohammad, S. M. Sentiment analysis of short informal texts. *J. Artif. Intell. Res.* **50**, 723–762 (2014).
- [70].Jurek, A. *et al.* Improved lexicon-based sentiment analysis for social media analytics. *Secur. Inform.* **4**, 9 (2015).
- [71].Cruz, N. P. , Taboada, M. & Mitkov, R. A Machine-Learning Approach to Negation and Speculation Detection for Sentiment Analysis. *J. Assoc. Inf. Sci. Technol.* **67**, 2118–2136 (2015).
- Process. Manag.* **53**, 436–449 (2017).
- [23].Akhtar, N. , Zubair, N. , Kumar, A. & Ahmad, T. ScienceDirect ScienceDirect Aspect based Sentiment Oriented Summarization of Hotel Reviews. *Procedia Comput. Sci.* **115**, 563–571 (2017).
- [24].Al-Smadi, M. , Al-Ayyoub, M. , Jararweh, Y. & Qawasmeh, O. Enhancing Aspect-Based Sentiment Analysis of Arabic Hotels' reviews using morphological, syntactic and semantic features. *Inf. Process. Manag.* **56**, 308–319 (2019).
- [25].Heerschop, B. , Van Iterson, P. , Hogenboom, A. , Frascar, F. & Kaymak, U. Analyzing sentiment in a large set of Web data while accounting for negation. *Adv. Intell. Soft Comput.* **86**, 195–205 (2011).
- [26].Hogenboom, A. *et al.* Exploiting emoticons in polarity classification of text. *J. Web Eng.* **14**, 022–040 (2015).
- [27].Liu, B. *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies* **5**, (Morgan & Claypool Publishers , 2012).
- [28].Tubishat, M. , Idris, N. & Abushariah, M. A. M. Implicit aspect extraction in sentiment analysis: Review, taxonomy, opportunities, and open challenges. *Inf. Process. Manag.* **54**, 545–563 (2018).
- [29].Rana, T. A. & Cheah, Y. N. Aspect extraction in sentiment analysis: comparative analysis and survey. *Artif. Intell. Rev.* **46**, 459–483 (2016).
- [30].Hu, M. & Liu, B. Mining and summarizing customer reviews. *Proc. 2004 ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '04* 168 (2004). doi:10.1145/1014052.1014073
- [31].Tripathy, A. , Agrawal, A. & Rath, S. K. Classification of sentiment reviews using n-gram machine learning approach. *Expert Syst. Appl.* **57**, 117–126 (2016).
- [32].Mohammad, S. , Dunne, C. & Dorr, B. Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. *EMNLP 2009 - Proc. 2009 Conf. Empir. Methods Nat. Lang. Process. A Meet. SIGDAT, a Spec. Interes. Gr. ACL, Held Conjunction with ACL-IJCNLP 2009* **2**, 599–608 (2009).
- [33].Neviarouskaya, A. , Prendinger, H. & Ishizuka, M. SentiFul: A lexicon for sentiment analysis. *IEEE Trans. Affect. Comput.* **2**, 22–36 (2011).
- [34].Baccianella, S. , Esuli, A. & Sebastiani, F. SENTIWORDNET 3. 0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proc. 7th Int. Conf. Lang. Resour. Eval. Lr. 2010* 2200–2204 (2010).
- [35].Bradley, M. & Lang, P. Affective norms for English words (ANEW): Instruction manual and affective ratings. (1999).
- [36].General, T. The General Inquirer : Useful or Not ? 1 By George Psathas. 163–164 (1968).
- [37].Elbagir, S. & Yang, J. Language Toolkit and VADER Sentiment. **0958**, (2019).
- [38].Khoo, C. S. G. & Johnkhan, S. B. Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *J. Inf. Sci.* **44**, 491–511 (2018).
- [39].Esuli, A. & Sebastiani, F. SENTIWORDNET: A high-coverage lexical resource for opinion mining. *Evaluation* 1–26 (2007). doi:10.1.1.61.7217
- [40].Strapparava, C. & Valitutti, A. WordNet-Affect: an affective extension of WordNet BT - Proceedings of LREC. **4**, 1083–1086 (2004).
- [41].Dong, X. & De Melo, G. A helping hand: Transfer learning for deep sentiment analysis. *ACL 2018 - 56th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.* **1**, 2524–2534 (2018).
- [42].Wu, F. & Huang, Y. Sentiment domain adaptation with multiple sources. *54th Annu. Meet. Assoc. Comput. Linguist. ACL 2016 - Long Pap.* **1**, 301–310 (2016).
- [43].Xing, F. Z. , Pallucchini, F. & Cambria, E. Cognitive-inspired domain adaptation of sentiment lexicons. *Inf. Process. Manag.* **56**, 554–564 (2019).
- [44].Dey, A. , Jenamani, M. & Thakkar, J. J. Senti-N-Gram: An n-gram lexicon for sentiment analysis. *Expert Syst. Appl.* **103**, 92–105 (2018).
- [45].Yates, A. , Goharian, N. & Yee, W. G. Semi-supervised probabilistic sentiment analysis: Merging labeled sentences with unlabeled reviews to identify sentiment. *Proc. Am. Soc. Inf. Sci. Technol.* **50**, 1–10 (2013).
- [46].Araque, O. , Guerini, M. , Strapparava, C. & Iglesias, C. A. Neural domain adaptation of sentiment lexicons. *2017 7th Int. Conf. Affect. Comput. Intell. Interact. Work. Demos, ACIIW 2017 2018-Janua*, 105–110 (2018).



تشخیص نوع نویز در تصویر با استفاده از هیستوگرام و آنتروپی

زینب خدابخشی^۱، حمید حسن پور^۲، سکینه اسدی امیری^۳

^۱ دانشگاه صنعتی شاهرود، شاهرود،
khodabakhshi_zeinab@yahoo.com
^۲ استاد، دانشگاه صنعتی شاهرود، شاهرود
h_hassanpour@yahoo.com
^۳ استادیار، دانشگاه مازندران، بابلسر
asadi_amiri@yahoo.com

چکیده

نویز مقادیر ناخواسته‌ای است که به دلایل مختلف روی تصویر ایجاد می‌شود. در بسیاری از تکنیک‌های پردازش تصویر اغلب نیاز هست تا پیش‌پردازش‌هایی روی تصویر انجام شود. یکی از پیش‌پردازش‌های مهم حذف نویز تصویر است. برای حذف بهتر نویز باید نوع نویز در تصویر تشخیص داده شود. در این مقاله روشی برای تشخیص سه نوع نویز گوسی، نمک و لفل و تصادفی با استفاده از ویژگی‌های هیستوگرام و آنتروپی پیشنهاد شده است. استفاده از دو ویژگی هیستوگرام و آنتروپی، نوع نویز موجود در تصویر را به خوبی توصیف می‌کنند. از آنجایی که ویژگی‌های محلی تصویر اطلاعات بهتر و دقیق‌تری را به ما می‌دهند، در روش پیشنهادی از هیستوگرام محلی تصویر استفاده شده است. بدین منظور هیستوگرام ناحیه هموار تصویر محاسبه شده است که بیانگر اطلاعات قابل توجهی از نوع نویز تصویر است. در انتها از دسته‌بند ماشین بردار پشتیبان برای دسته‌بندی نوع نویز تصویر استفاده شده است. روش پیشنهادی روی تصاویر پایگاه داده CSIQ مورد ارزیابی قرار گرفت، که دقت آن ۹۷.۹۲ درصد است.

کلمات کلیدی

نویز، هیستوگرام، آنتروپی، ماشین بردار پشتیبان

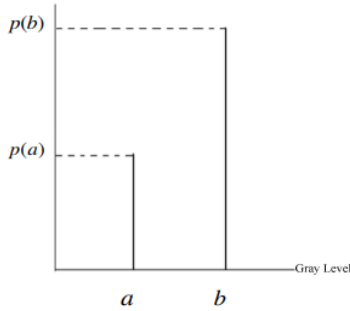
تابع چگالی احتمال^۴ یا هیستوگرام متفاوتی می‌باشند [۴]. بنابراین برای تشخیص نوع نویز در تصویر می‌توان از تابع چگالی احتمال و یا هیستوگرام بهره برد. استفاده از هیستوگرام برای تشخیص نوع نویز تصویر در [۲] پیشنهاد شده است. اما این نکته را باید در نظر گرفت که هیستوگرام کل تصویر ممکن است اطلاعات دقیقی در شناسایی نوع نویز به ما ندهد. بنابراین نیاز هست تا از اطلاعات محلی تصویر استفاده کرد. مرجع [۵] برای تشخیص و کلاسه‌بندی انواع نویز در تصویر از شبکه عصبی کانولوشنی با الگوریتم back propagation و تکنیک‌های بهینه‌سازی گرادیان نزولی بهره می‌برد.

بعضی از نویزها مانند نویز متناوب^۵ دارای تابع چگالی احتمال خاصی نمی‌باشند. این نوع نویز مدلی از یک یا چند موج سینوسی است که با تصویر اصلی جمع می‌شوند. از آنجایی که این نوع نویز دارای نظم خاص و فرکانس مشخصی است، از طریق تبدیل فوریه می‌توان نویز متناوب را در تصویر مشاهده کرد. جهت تشخیص این نوع نویز در تصویر می‌توان از معیار همبستگی متقابل^۱ بین سطرها و ستون‌های

۱- مقدمه

اغلب تصاویر دیجیتال در زمان عکس برداری و انتقال الکترونیکی دچار خرابی‌هایی می‌شوند. یکی از خرابی‌هایی که تصاویر را در طول فرآیندهای فوق دچار خرابی می‌کند، نویز است. نویز مقادیر ناخواسته‌ای است که به دلایل مختلف بر روی تصویر ایجاد می‌شود. بسته به نوع خرابی‌ها مانند دستگاه‌های ثبت تصاویر و یا عملیات صورت گرفته، یک تصویر ممکن است به نویزهای متفاوتی آغشته شود [۱]. جهت رفع نویزهای ایجاد شده در تصویر، فیلترهایی متناسب با آن‌ها ارائه شده است. اما بیشتر روش‌های رفع نویز با این فرض که نوع نویز را می‌دانند اقدام به رفع نویز تصویر می‌کنند.

برای کاهش اثرات نویز در تصویر باید نوع و مدل نویز شناسایی شود. مطالعه دقیق نوع و مدل‌های نویز یکی از عناصر اصلی در سیستم‌های حذف نویز می‌باشد [۲،۳]. نویزهایی مانند، نویز ضربه^۱ (نمک و لفل^۲ و تصادفی^۳)، گوسی^۴، یکنواخت^۵، گاما^۶ و ریلی^۷ دارای



شکل ۱: نمودار تابع چگالی احتمال نویز نمک و فلفل [۴]

همان‌طور که از شکل (۱) پیداست، دو مقدار نمک و فلفل نسبت

به سایر پیکسل‌های تصویر دارای فراوانی بیشتری می‌باشند.

در سال‌های اخیر برای شناسایی نویز ضربه از تکنیک‌های هوش

مصنوعی مانند شبکه عصبی مصنوعی (ANN^۱)، شبکه عصبی فازی

(NFN^۲) و ماشین بردار پشتیبان (SVM^۳) استفاده شده است [۱۰].

۲-۲- نویز ضربه با مقدار تصادفی

در مدل نویز ضربه‌ای با مقدار تصادفی پیکسل نویزی باعث تغییر بعضی از پیکسل‌ها با احتمال p می‌شود. مقدار سایر پیکسل‌های تصویر با احتمال $1-p$ سالم باقی می‌ماند [۱۱]. رابطه (۳) مدل این نوع نویز را بیان می‌کند:

$$n(i, j) = \begin{cases} m_{i, j} & \text{with probability } p \\ o(i, j) & \text{with probability } (1-p) \end{cases} \quad (3)$$

در این رابطه مانند معادله (۲)، p احتمال سطح نویز ضربه

می‌باشد، $m_{i, j}$ مقداری تصادفی با توزیع یکنواخت بین ۰ و ۲۵۵ است.

چالش بزرگ در این نوع نویز پیدا کردن پیکسل نویزی از میان تعداد زیادی سطح خاکستری می‌باشد [۱۲].

۲-۳- نویز گوسی

نویز گوسی دارای توزیع نرمال (یا توزیع گوسی) می‌باشد [۱۳]. در این نوع نویز فراوانی یک سری از مقادیر پیکسل‌ها بیشتر از سایر مقادیر می‌باشد. این مدل نویز بیشتر در تصاویر طبیعی مشاهده می‌شود. تابع چگالی احتمال نویز گوسی با شدت نویز x به صورت معادله (۴) می‌باشد:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad (4)$$

در این معادله μ مقدار میانگین، σ انحراف معیار استاندارد و σ^2

واریانس توزیع گوسی می‌باشد [۱۴]. شکل (۲) نمودار تابع چگالی

احتمال نویز گوسی را نشان می‌دهد.

تصویر در حوزه مکان استفاده کرد. در [۶] برای رفع نویز متناوب، نواحی فرکانسی مشابه نویز با استفاده از یک الگوریتم قطعه‌بندی مبتنی بر باند طیفی تصویر نویزی به کمک آستانه وقتی شناسایی می‌شود. سپس این نواحی با توجه به تعیین پهنای باند هر مؤلفه نویز متناوب به صورت جداگانه توسط الگوریتم رشد می‌کند. در ادامه نواحی فرکانسی مشخص شده رفع نویز می‌شود.

در بخش بعدی نویز و مدل‌های آن مرور می‌شود. در بخش

سوم، روش پیشنهادی جهت تعیین نوع نویز معرفی خواهد شد. در ادامه

در بخش‌های ۴ و ۵ نتایج تجربی و نتیجه‌گیری بیان می‌شود.

۲- نویز و مدل آن

نویز یکی از انواع خرابی است که باعث کاهش کیفیت تصویر می‌شود. وجود نویز در تصویر علاوه بر کاهش کیفیت تصویر باعث از دست رفتن اطلاعات و جزئیات تصویر شده و اثرات نامطلوبی مانند لبه‌های غیر واقعی، ایجاد مصنوعات، گوشه‌ها و تاروی روی تصویر می‌گردد [۷]. رابطه (۱) مدل خرابی تصویر را بیان می‌کند:

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \quad (1)$$

در این رابطه $g(x, y)$ تصویر خراب است. به زبان ریاضی، پیچش

تصویر سالم $f(x, y)$ با تابع خرابی $h(x, y)$ منجر به ایجاد تصویر خراب اولیه می‌شود. با جمع تصویر خراب اولیه با تابع نویز $n(x, y)$ ، تصویر خراب نهایی بوجود می‌آید [۸]. تابع نویز می‌تواند هر یک از نویزهای ضربه نمک و فلفل، ضربه تصادفی، گوسی و غیره باشد.

به دلیل تصادفی بودن مکان و مقدار نویز به غیر از نویز نمک و

فلفل که دارای دو مقدار ثابت می‌باشد، هر مدل نویز دارای تابع چگالی احتمال متفاوتی می‌باشد [۹]. در ادامه سه نوع نویز ضربه نمک و فلفل، ضربه با مقدار تصادفی و نویز گوسی به صورت اجمالی بررسی می‌شود.

۲-۱- نویز ضربه نمک و فلفل

در نویز ضربه‌ای مقدار نویز نمک یا فلفل جایگزین مقدار پیکسل در تصویر اصلی می‌شود. نویز ضربه‌ای در دو نوع ثابت و تصادفی دسته‌بندی می‌شود. نویز ضربه با مقدار ثابت، نویز نمک و فلفل خوانده می‌شود. برای مثال نویز ضربه با مقدار ثابت می‌تواند دو مقدار صفر (فلفل) و ۲۵۵ (نمک) را داشته باشد. تابع چگالی احتمال نویز نمک و فلفل با مقدار ثابت ۰ و ۲۵۵ در معادله (۲) نشان داده شده است:

$$n(i, j) = \begin{cases} b = 255 & \text{with probability } \frac{p}{2} \\ a = 0 & \text{with probability } \frac{p}{2} \\ o(i, j) & \text{with probability } 1-p \end{cases} \quad (2)$$

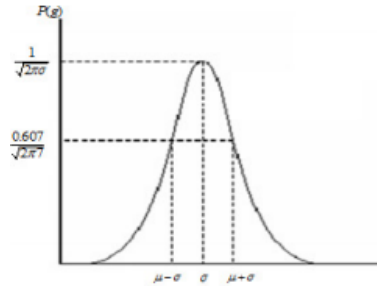
در این رابطه $o(i, j)$ پیکسل‌های غیر نویزی تصویر است. p

احتمال سطح نویز و $n(i, j)$ نمایش تصویر خراب شده توسط نویز

نمک و فلفل است [۹]. در شکل (۱) نمودار تابع چگالی احتمال این

نوع نویز نمایش داده شده است.

دارای دو مقدار ۰ و ۲۵۵ نیست. بنابراین نمودار هیستوگرام این نوع نویز، دارای دو قله با ارتفاع زیاد می‌باشد. نویز گوسی ماهیت متفاوتی با نویز نمک و فلفل دارد، نمودار هیستوگرام این نوع نویز دارای توزیع گوسی می‌باشد. توزیع هیستوگرام نویز تصادفی تقریباً شبیه به توزیع نویز گوسی است، اما توزیع هیستوگرام نویز گوسی دارای نظم بهتری نسبت به نویز تصادفی است. در نویز تصادفی مقدار نویز به صورت یکنواخت به تمام پیکسل‌های تصویر اعمال می‌شود. شکل (۴) نمودار هیستوگرام تصویر سالم و تصاویر آغشته به نویز گوسی، نمک و فلفل و تصادفی را نشان می‌دهد.

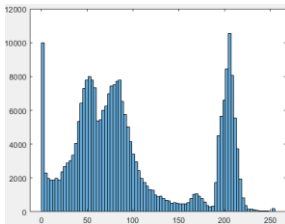


شکل ۲: نمودار تابع چگالی احتمال توزیع گوسی [۴]

همان‌طور که در شکل (۲) مشخص است فراوانی پیکسل‌ها در نویز گوسی دارای توزیع نرمال می‌باشد.

شکل (۳) نمونه‌ای از تصویر آغشته به سه نوع نویز، نمک و فلفل،

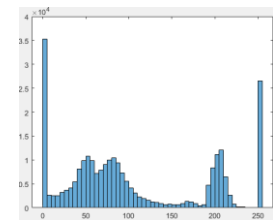
تصادفی و گوسی را نشان می‌دهد.



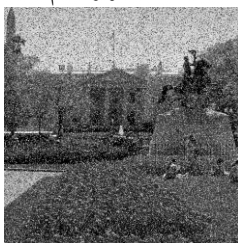
(ب) هیستوگرام تصویر الف



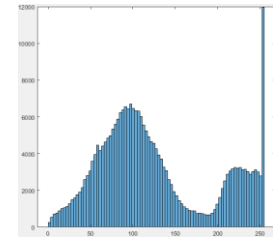
(الف) تصویر سالم



(ت) هیستوگرام تصویر پ



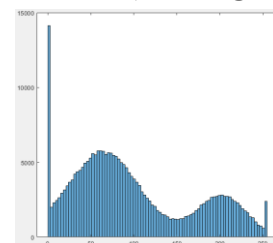
(ب) تصویر آغشته به نویز نمک و فلفل



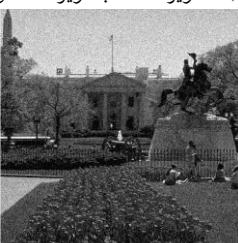
(ح) هیستوگرام تصویر ث



(ث) تصویر آغشته به نویز تصادفی



(ج) هیستوگرام تصویر خ



(خ) تصویر آغشته به نویز گوسی

شکل ۴: ترسیم هیستوگرام تصاویر آغشته به نویز

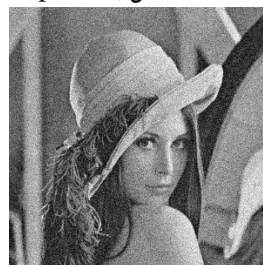
استفاده از ویژگی‌های محلی تصویر اطلاعات دقیق‌تری از آن را دارا می‌باشد. به دلیل اینکه هیستوگرام محلی نتایج دقیق‌تری را نسبت به هیستوگرام سراسری تصویر می‌دهد، در روش پیشنهادی از هیستوگرام محلی تصویر استفاده می‌شود. ناحیه محلی در نظر گرفته شده شامل ناحیه هموار تصویر می‌باشد. در ناحیه هموار تغییرات ارزش پیکسل‌ها کم است. تغییرات ناگهانی در این ناحیه نشان از وجود نویز در تصویر می‌باشد. بنابراین در روش پیشنهادی از هیستوگرام ناحیه هموار تصویر استفاده می‌شود. برای بدست آوردن هیستوگرام محلی تصویر، تصویر



(ب) تصویر آغشته به نویز نمک و فلفل با $p=0.1$



(الف) تصویر سالم



(د) تصویر آغشته به نویز تصادفی با $p=0.3$



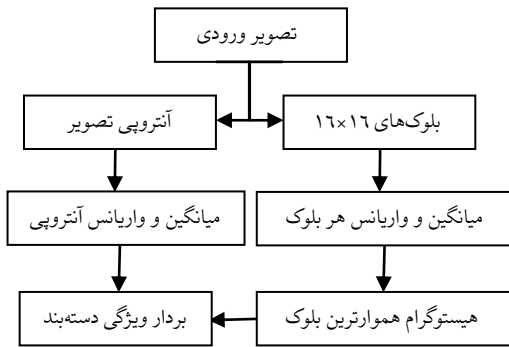
(ج) تصویر آغشته به نویز گوسی با $p=0.1$

شکل ۳: تصاویر آغشته به نویز نمک و فلفل، تصادفی و گوسی

۳- روش پیشنهادی

تشخیص نوع نویز تصویر قبل از عملیات بهسازی آن به انتخاب فیلتر مناسب برای رفع نویز کمک شایانی می‌کند. برای فیلتر کردن تصویر بهتر است در ابتدا نوع نویز در آن مشخص شود. در روش پیشنهادی استفاده از هیستوگرام و آنتروپی تصویر به ما کمک می‌کند تا نوع نویز با دقت بالایی شناسایی شود. مقادیر پیکسل‌ها در یک تصویر دیجیتال، ویژگی‌هایی مانند روشنایی و وضوح تصویر را بیان می‌کند. هیستوگرام تصویر توزیع فراوانی ارزش پیکسل‌های آن را نشان می‌دهد. محور افقی در نمودار هیستوگرام شامل مقادیر روشنایی در تصویر n بیتی است، و محور عمودی بیان‌کننده میزان فراوانی پیکسل‌های تصویر می‌باشد.

در روش پیشنهادی از هیستوگرام تصویر عاری از نویز و سه نوع نویز گوسی، نمک و فلفل و تصادفی استفاده شده است. نویز نمک و فلفل دو مقدار بیشینه و کمینه، که در اینجا به ترتیب دو مقدار ۲۵۵ و ۰ را با فراوانی زیاد، روی تصویر قرار می‌دهد، اما الزاماً این نوع نویز



شکل ۶: دیاگرام روش پیشنهادی

۴- نتایج تجربی

برای ارزیابی روش پیشنهادی از پایگاه داده CSIQ استفاده شده است. نمونه تصاویر موجود در این پایگاه فقط آغشته به نویز گوسی هست، بنابراین در روش پیشنهادی تصاویر مورد استفاده به انواع نویز با پنج سطح شدت متفاوت آغشته شده‌اند.

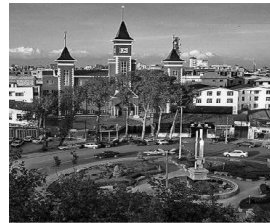
برای تشخیص نوع نویز در تصویر از دسته‌بند ماشین بردار پشتیبان استفاده شده است [۱۵]. ویژگی که برای دسته‌بند استفاده می‌شود، در نظر گرفتن هیستوگرام؛ میانگین و واریانس آنتروپی تصویر است. همانطور که در بخش روش پیشنهادی اشاره شد، از هیستوگرام هموارترین ناحیه از تصویر استفاده می‌شود، چرا که سیستم بینایی انسان به نویز موجود در نواحی هموار تصویر تمرکز بیشتری نسبت به نواحی غیرهموار تصویر دارد. بنابراین تصویر به بلوک‌هایی تقسیم‌بندی می‌شود، و با استفاده از میانگین و واریانس، هموارترین بلوک بدست می‌آید. هموارترین بلوک بدست آمده برای هیستوگرام تصویر استفاده می‌شود. پس از تعیین هیستوگرام، میانگین و واریانس آنتروپی تصویر نیز در نظر گرفته می‌شود.

همان‌طور که اشاره شد، آنتروپی و توزیع مقادیر پیکسل‌ها در هیستوگرام تصویر عاری از نویز و تصاویر آغشته به نویز باهم فرق دارد. در روش پیشنهادی جهت تشخیص نوع نویز، دسته‌بندی در چند مرحله انجام می‌شود. در مرحله اول استفاده از دسته‌بند، با در نظر گرفتن دو معیار هیستوگرام و آنتروپی، تصویر نویزی از غیر نویزی تشخیص داده می‌شود. در مرحله دوم به دلیل متفاوت بودن ماهیت نویز نمک و فلغل از نویز گوسی و تصادفی، تصاویر آغشته به نویز نمک و فلغل از دو نوع نویز دیگر قابل تشخیص است. در نهایت نویز گوسی و تصادفی از هم تشخیص داده می‌شوند.

پایگاه داده‌ای که برای تشخیص نوع نویز ساخته شده شامل: بردار ویژگی با مقادیر هیستوگرام ناحیه هموار تصویر سالم و نویزی به همراه میانگین و واریانس آنتروپی آنها می‌باشد. برای این پایگاه از ۳۰ تصویر مرجع پایگاه CSIQ استفاده شده است. مقادیر هیستوگرام ناحیه هموار و میانگین و واریانس این ۳۰ تصویر به همراه ۴۵۰ تصویر آغشته به نویز نمک و فلغل، گوسی و تصادفی با ۵ سطح شدت نویز مختلف برای هر نوع نویز در این پایگاه به کار برده شد. از پایگاه ساخته شده با ۴۸۰ تصویر، ۸۰ درصد برای آموزش دسته‌بند و ۲۰ درصد برای آزمون عملکرد دسته‌بند استفاده شده است. در روش پیشنهادی برای آموزش

آغشته به نویز را به بلوک‌های 16×16 تقسیم‌بندی می‌کنیم. پس از آن واریانس بلوک‌های بدست آمده محاسبه می‌شود. از میان بلوک‌های بدست آمده، بلوکی به عنوان هموارترین بلوک شناخته می‌شود که مقدار واریانس آن نسبت به بقیه بلوک‌ها کمتر باشد. انتخاب هموارترین بلوک، به صورت واضحی پیکسل‌های نویزی را از غیر نویزی نشان می‌دهد. بنابراین هیستوگرام هموارترین بلوک انتخاب شده برای تشخیص نوع نویز، محاسبه می‌گردد.

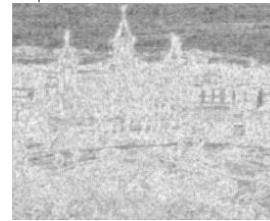
بعد از بدست آوردن هیستوگرام تصویر، آنتروپی آن را بدست می‌آوریم. میزان آنتروپی هر پیکسل با پیکسل همسایه آن فرق می‌کند. این معیار میزان بی‌نظمی که در پیکسل‌های تصویر وجود دارد را بیان می‌کند. بی‌نظمی که در تصویر نویزی وجود دارد به مراتب متفاوت از خود تصویر سالم است. بنابراین میزان آنتروپی تصویر عاری از نویز با تصویر آغشته با نویز نمک و فلغل، تصادفی و گوسی با هم فرق دارد. همان‌طور که در شکل (۵) مشاهده می‌شود، نقشه آنتروپی تصاویر آغشته به نویز با هم متفاوت است.



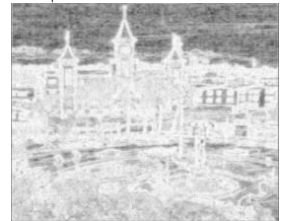
ب) تصویر خاکستری سالم



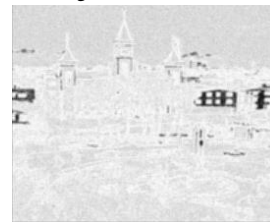
الف) تصویر رنگی سالم



ت) نقشه آنتروپی تصویر آغشته به نویز نمک و فلغل



پ) نقشه آنتروپی تصویر سالم



ج) نقشه آنتروپی تصویر آغشته به نویز تصادفی



ث) نقشه آنتروپی تصویر آغشته به نویز گوسی

شکل (۵): نقشه آنتروپی تصویر سالم و تصویر آغشته به نویز

با توجه به تفاوت نقشه آنتروپی در شکل (۵) استفاده از این معیار به همراه هیستوگرام تصویر کمک می‌کند تا تشخیص نوع نویز تصویر با دقت بالاتری صورت گیرد. جهت استفاده از آنتروپی تصاویر در روش پیشنهادی، از مقدار آنتروپی هر تصویر میانگین و واریانس گرفته شد. در نهایت هیستوگرام هموارترین بلوک از تصویر به همراه میانگین و واریانس آنتروپی جهت تعیین نوع نویز استفاده می‌شود. شکل (۶) دیاگرام این روند را نشان می‌دهد.

- [13] Agarwal R., "Bit plane average filtering to remove gaussian noise from high contrast images." International Conference on Computer Communication and Informatics (ICCCI). pp. 1-5, 2012.
- [14] Ahmed H., Amin M., and Abd El-Samie F. E., "Simultaneous denoising and compression of multispectral images." *Journal of Applied Remote Sensing* 7, no. 1, pp. 1-17, 2013.
- [15] Cortes C., Vapnik V., "Support-vector networks", *Machine learning*, 20(3), 273-297, 1995.
- [16] Rodriguez J. D., Perez A., & Lozano J. A., "Sensitivity analysis of k-fold cross validation in prediction error estimation". *IEEE transactions on pattern analysis and machine intelligence*, 32(3), pp. 569-575, 2010.
- دسته‌بند از روش اعتبارسنجی k -دسته‌ای^{۱۴} با ۱۰ دسته استفاده می‌شود [۱۶]. این عمل بدین صورت انجام می‌شود که مجموعه آموزش به ۱۰ دسته برابر تقسیم و هر بار نه دسته برای آموزش و یک دسته باقیمانده برای آزمون در نظر گرفته خواهد شد. این روند ده بار تکرار می‌شود، دقت دسته‌بند، میانگین دقت این ده بار تکرار است. در نهایت دقت روش پیشنهادی برای تشخیص سه نوع نویز نمک و فلفل، تصادفی و گوسی، روی پایگاه داده CSIQ با استفاده از ماشین بردار پشتیبان به میزان ۹۷/۹۲ درصد می‌باشد.

۵- نتیجه

زیرنویس‌ها

- 1 Impulse Noise
- 2 Salt and Pepper Impulse Noise (SPIN)
- 3 Random Value Impulse Noise (RIVN)
- 4 Gaussian
- 5 Uniform
- 6 Gamma
- 7 Rayleigh
- 8 Probability Density Function (PDF)
- 9 Periodic Noise
- 10 Cross Correlation
- 11 Artificial Neural Network
- 12 Neural Fuzzy Network
- 13 Support Vector Machine
- 14 K-Fold Cross Validation

تشخیص نوع نویز در تصویر کمک می‌کند تا با انتخاب فیلتر متناسب با نوع نویز، عملیات رفع نویز با نتایج بهتری انجام شود. در روش پیشنهادی برای تشخیص نوع نویز در تصویر از میانگین و واریانس هیستوگرام هموارترین ناحیه از تصویر به همراه آنتروپی استفاده شده است. نویزهای موجود در روش پیشنهادی شامل سه نوع نویز ضربه نمک و فلفل، تصادفی و گوسی می‌باشد. نتایج استفاده از دسته‌بند SVM روی پایگاه داده CSIQ نشان دهنده این است که روش پیشنهادی با دقت ۹۷/۹۲٪ نوع نویز در تصویر را درست تشخیص می‌دهد.

مراجع

- [1] Frankel, David S., *Model Driven Architecture: Applying MDA to Enterprise Computing*, OMG Press, Wiley Publishing, 2003.
- [2] Chan, R. H., Ho, C. W., Nikolova, M., "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization". *IEEE Transactions on image processing*, 14(10), pp. 1479-1485, 2005.
- [3] Gonzalez R. C., Woods R. E., "Digital Image Processing", second ed., Prentice Hall, Englewood, Cliffs, NJ, 2002.
- [4] Patil, J. & Jadhav S., "A Comparative Study of Image Denoising Techniques", *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 2, No. 3. 2013.
- [5] Khaw, H. Y., Soon, F. C., Chuah, J. H., & Chow, C. O., "Image noise types recognition using convolutional neural network with principal components analysis". *IET Image Processing*, 11(12), pp. 1238-1245. 2017.
- [6] Moallem, P., Masoumzadeh, M., & Habibi, M., "A novel adaptive Gaussian restoration filter for reducing periodic noises in digital image". *Signal, Image and Video Processing*, 9(5), pp. 1179-1191, 2015.
- [7] Boyat, A. K., Joshi, B. K., "A review paper: noise models in digital image processing", *arXiv preprint arXiv:1505.03489*, pp. 63-75, 2015.
- [8] Ajay Kumar B., Brijendra K. J., "A review paper: noise models in digital image processing." *Signal & Image Processing: An International Journal (SIPIJ)* Vol.6, No.2, pp. 63-75, 2015.
- [9] Tiwari S., Shukla V.P. Biradar S. and Singh A., "Texture features based blur classification in barcode images". *International Journal of Information Engineering and Electronic Business*, 5(5), p.34, 2013.
- [10] Deng, H. Y., Zhu, Q. X., Song, X. L., "A nonlinear diffusion for salt and pepper noise removal". *13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 231-234. IEEE, 2016
- [11] Lehtinen, Jaakko, et al. "Noise2noise: Learning image restoration without clean data." *arXiv preprint arXiv:1803.04189*, 2018.
- [12] Thangaraj V., "Context model based edge preservation filter for impulse noise removal." *Expert Systems with Applications* Vol. 88, pp. 29-44, 2017.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



دانشگاه تبریز

۲۷ الی ۲۸ آذرماه ۱۳۹۸

ارزیابی دو روش استخراج ویژگی مبتنی بر پردازش زمان-فرکانس سیگنال رادار با هدف تشخیص زمین خوردن در سالمندان

فاطمه تاجیک^۱، ندا فرجی^۲

^۱ دانشجوی کارشناسی ارشد مخابرات سیستم، دانشکده فنی مهندسی-گروه مهندسی برق، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران
Fafa.tajik@gmail.com

^۲ استادیار، دانشکده فنی مهندسی-گروه مهندسی برق، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران
nfaraji@eng.ikiu.ac.ir

چکیده

در این مقاله، دو الگوریتم پردازش سیگنال زمان-فرکانس برای استخراج ویژگی از سیگنال رادار مورد ارزیابی قرار می‌گیرد که با کمک آنها تشخیص اتوماتیک زمین خوردن سالمندان با دقت بالا قابل تحقق خواهد بود. نقطه شروع در تشخیص زمین خوردن مبتنی بر رادار استخراج ویژگی از سیگنال راداری است. بردار ویژگی استخراج شده در قسمت طبقه‌بندی برای تشخیص فعالیت زمین خوردن از غیرزمین خوردن مورد استفاده قرار می‌گیرد. در اولین روش، تبدیل موجک به عنوان یک رویکرد متداول پردازش زمان-فرکانس جهت استخراج ویژگی از سیگنال رادار به‌کاررفته و کارایی انواع موجک‌های مادر با هم مقایسه می‌شوند. همچنین آنالیز کپسترال فرکانس مل به عنوان یک روش پردازش زمان-فرکانس مطرح در پردازش گفتار به عنوان روش دوم در استخراج ویژگی مطرح می‌شود. سپس هر دو روش استخراج ویژگی در سه طبقه‌بندی‌کننده K-نزدیکترین همسایه، ماشین بردار پشتیبان و درخت تصمیم‌گیری به‌جهت کارایی طبقه‌بندی‌کننده‌ها مورد ارزیابی جامع قرار گرفته و مقایسه می‌شوند. پیاده‌سازیهای انجام شده، نشان دهنده عملکرد بهتر طبقه‌بندی‌کننده‌ها در روش استخراج ویژگی مبتنی بر آنالیز کپسترال فرکانس مل در مقایسه با روش تبدیل موجک است.

کلمات کلیدی

استخراج ویژگی، تبدیل موجک، ضرایب کپسترال فرکانس مل، تشخیص زمین خوردن مبتنی بر رادار

۱- مقدمه

هر ساله از سه نفر سالمند بالای ۶۵ سال یک نفر زمین می‌خورد. بنابراین بهتر است افراد سالخورده‌ای که در معرض خطر زمین خوردن هستند، به مراکز مراقبت از سالمندان منتقل شوند که خود هزینه‌ای گزاف در پی دارد. تشخیص سریع زمین خوردن^۱ منجر به کمک‌های فوری توسط مراقبین، درمان موثر و کاهش هزینه‌های پزشکی و درمانی می‌شود و در سال‌های اخیر تبدیل به یک زمینه پژوهشی مهم شده با این امید که امکان زندگی مستقل را برای افراد سالخورده فراهم کند [۱].

دستگاه‌های خانگی تشخیص زمین خوردن با قیمت پایین‌تر نسبت به دستگاه‌های بالینی در بازار در دسترس هستند. این دستگاه‌ها می‌توانند پوشیدنی [۲] و غیرپوشیدنی [۳] باشند. دستگاه‌های پوشیدنی مانند شتاب‌سنج و دکمه‌های فشار که بیشتر رایج هستند، به‌راحتی

زمین خوردن یکی از دلایل اصلی مرگ و میر در بین افراد مسن (۶۵ سال و بالاتر) است. پیش‌بینی می‌شود جمعیت بیش از ۶۵ سال در سراسر جهان تا سال ۲۰۳۰ به یک میلیارد افزایش یابد. این پیری جمعیت باعث می‌شود که مدل ارائه مراقبت‌های بهداشتی نسبت به گذشته متفاوت گردد و از رویکرد بیمارستان‌محور متداول که در آن بیماری تشخیص داده شده و بیمار باید برای درمان در شرایط حاد در بیمارستان‌های تخصصی بستری شود، به یک مدل خانه‌محور که در آن مراقبت از بیمار تا حد امکان با استفاده از فن‌آوری‌های جدید در خانه خود انجام می‌شود، تغییر کند [۱].

ارزیابی قرار می‌دهیم. اولی، روش استخراج ویژگی مبتنی بر موجک است که در آن انواع موجک‌های مادر به لحاظ کارایی تشخیص با هم مقایسه می‌شوند. در دومین روش، از ضرایب کپستراتل فرکانس مل^۲ به عنوان بردار ویژگی استخراج شده از سیگنال رادار بهره می‌جوییم. هر دو روش استخراج ویژگی با استفاده از سه خانواده طبقه‌بندی کننده متداول در حوزه بازشناسی الگو مورد بررسی قرار گرفتند. طبقه‌بندی کننده‌ها عبارتند از: ماشین بردار پشتیبان^۴، K نزدیکترین همسایه^۵ و درخت تصمیم‌گیری^۶. نتایج آزمایشات حاکی از عملکرد بسیار بهتر ویژگی‌های MFCC در مقایسه با ویژگی‌های مبتنی بر موجک است.

این مقاله به شرح زیر سازماندهی شده است. بخش دوم پیش زمینه مختصری در مورد رادار داپلر ارائه می‌دهد. در بخش سوم، استخراج ویژگی از سیگنال رادار باهدف تشخیص زمین‌خوردن با دو روش مختلف شرح داده شده است. بخش چهارم آزمایشات و نتایج را شرح می‌دهد. سرانجام، ما در بخش پنجم مقاله نتیجه‌گیری می‌کنیم.

۲- رادار داپلر

رادار داپلر^۷ وسیله‌ای است که با استفاده از اثر داپلر، قادر به اندازه‌گیری سرعت اجسام می‌باشد [۱۴]. این اندازه‌گیری با فرستادن سیگنال مایکروویو به سمت هدف موردنظر و دریافت انعکاس آن، انجام می‌شود. رادار داپلر تغییرات فرکانس سیگنال منعکس شده را نسبت به سیگنال اصلی آنالیز کرده و براساس آن سرعت هدف مورد نظر نسبت به منبع رادار و جهت پرتو مایکروویو به دست می‌آید.

فرکانس داپلر (F_d)، که برابر با اختلاف فرکانس بین سیگنال ارسالی و دریافتی است، از رابطه (۱) به دست می‌آید [۱۵]

$$F_d \approx 2v \frac{F_t}{c} \quad (1)$$

که در آن c سرعت نور، v سرعت هدف و F_t فرکانس اصلی فرستنده است. بادر نظر گرفتن این مورد که سرعت نور $c = 3 \times 10^8$ متر بر ثانیه است و برای فرکانس ارسالی برابر با $24/125$ گیگاهرتز و با این فرض که سرعت زمین‌خوردن انسان قبل از برخورد به زمین نوعاً در حدود پنج متر بر ثانیه است، فرکانس داپلر حدود 800 هرتز خواهد بود.

در شکل (۱) برای یک رادار $24/125$ گیگاهرتز که در تشخیص زمین‌خوردن مورد استفاده قرار گرفته یک نمونه سیگنال خروجی در حوزه زمان به همراه اسپکتروگرام آن نشان داده شده است. در این شکل که زمین‌خوردن در آن حول ثانیه دو اتفاق افتاده، سیگنال رادار حول این زمان محدوده فرکانسی صفر تا 700 هرتز را پوشش داده است که منطبق با عدد به دست آمده در حالت تئوری می‌باشد.

شکسته می‌شوند و از آنجایی که باید پوشیده و یا حمل شوند به نحوی مزاحم هستند. علاوه بر این، دستگاه‌های فشار دکمه‌ای برای کاربرانی که دارای اختلال شناختی هستند مناسب نیست. بیشتر این سیستم‌های نظارتی پوشیدنی فقط بر روی تشخیص یک رویداد زمین‌خوردن تمرکز می‌کنند. این درحالی است که برای جلوگیری از خطر زمین‌خوردن، آگاهی از علت رویداد زمین‌خوردن نیز ضروری است. برای رفع این چالش، اخیراً یک سیستم نظارت مداوم مبتنی بر سنسور پوشیدنی ارائه شده است که قادر به تشخیص زمین‌خوردن و نیز شناسایی الگوی زمین‌خوردن و فعالیت‌های مرتبط با رویداد زمین‌خوردن می‌باشد [۴]. دستگاه‌های غیرپوشیدنی اغلب بدون مزاحم و قابل قبول‌تر برای کاربران هستند. یک وسیله غیرپوشیدنی متداول برای تشخیص زمین‌خوردن دوربین ویدیویی است که در آن از تکنیک پردازش تصویر در تشخیص زمین‌خوردن استفاده می‌شود [۳]. با این حال، سیستم‌های مبتنی بر دید نمی‌توانند در محیط‌های کم‌نور کار کنند. روش دیگر، آرایه میکروفون است که از سیگنال صوتی تولید شده از برخورد با کف برای تشخیص زمین‌خوردن استفاده می‌کند [۵]. البته دستگاه‌های صوتی نیاز به محیط‌های نسبتاً آرام دارند. اگر چه هر دو سیستم بصری و صوتی می‌توانند با یکدیگر برای تشخیص زمین‌خوردن در خانه کار کنند و مکمل یکدیگر شوند تا نتایج قابل اطمینان‌تری ارائه دهند، اما استفاده از هر دو در محیط حمام وجود ندارد. آمار نشان می‌دهد که این مکان برای شناسایی زمین‌خوردن چالش‌برانگیز است. فیلمی که توسط دوربین فیلمبرداری در حمام گرفته می‌شود واضح نیست. همچنین ابعاد کوچک حمام موجب ایجاد انعکاس می‌شود که عملکرد سیستم صوتی را دچار مشکل می‌کند [۶].

سیستم تشخیص زمین‌خوردن مبتنی بر رادار از این جهت که می‌تواند در محیط‌های کم‌نور و یا پرسر و صدا کار کند، از سیستم مبتنی بر تصویر و سیستم مبتنی بر آکوستیک متمایز می‌شود. علاوه بر این، چنین سیستمی می‌تواند نگرانی‌های مربوط به حریم خصوصی را، به‌ویژه در حمام یا اتاق خواب، رفع کند [۱]. تکنیک‌های پیشرفته تشخیص و طبقه‌بندی سیگنال‌های رادار برای تشخیص زمین‌خوردن از فعالیت‌هایی مانند ایستادن، نشستن، زانو زدن و سایر فعالیت‌ها در مراجع [۶] و [۷] مطرح شده است. این طبقه‌بندی در [۷-۱۰] با استفاده از ویژگی‌های استخراج شده از نمایش زمان-فرکانس سیگنال رادار به دست آمد. سیگنال‌های زمین‌خوردن مبتنی بر رادار با استفاده از تبدیل موجک^۸ (WT) در [۱۱] و [۶] مورد تجزیه و تحلیل قرار گرفت و ویژگی‌های استخراج شده برای تشخیص زمین‌خوردن به کار رفت.

اگر چه دامنه تحقیقات در این حوزه هنوز با چالش‌هایی روبه‌رو است، در آینده‌ای نزدیک، راه حل‌هایی جدید بر اساس اینترنت پیشرفته اشیا برای جلوگیری از آسیب زمین‌خوردن به وجود خواهد آمد [۱۲]. اخیراً، یک روش تشخیص زمین‌خوردن با استفاده از یادگیری عمیق و ادغام چند حسگر ارائه شده که در آن یک رادار موج پیوسته و دوربین‌های نوری به طور همزمان اطلاعات مربوط به حرکات انسانی را ثبت می‌کنند. با ادغام اطلاعات چند حسگر، عملکرد کلی سیستم تشخیص زمین‌خوردن به طرز چشمگیری بهبود می‌یابد [۱۳].

در این مقاله، دو روش پردازش زمان-فرکانس را برای استخراج ویژگی از سیگنال رادار داپلر با هدف تشخیص زمین‌خوردن مورد

نیست، زیرا سیگنال زمین‌خوردن رفتاری غیرثابت دارد. چنین سیگنال نایبستانی را می‌توان به خوبی با استفاده از نمایش‌های زمان - فرکانس بازنمایی کرد [۶].

استخراج ویژگی از سیگنال رادار، با استفاده از تبدیل موجک شامل دو مرحله است. در مرحله اول با استفاده از ضرایب تجزیه موجک در یک سطح دلخواه مانند سطح دو، بازه محتمل برای رویداد زمین‌خوردن شناسایی می‌شود. در مرحله دوم، تنها مقدار داده در بازه محتمل برای استخراج ویژگی‌ها مورد استفاده قرار می‌گیرد.

پس از اعمال تبدیل موجک ایستان^۱ بر روی داده، عملیات فریم‌بندی روی ضرایب موجک در یک سطح مانند سطح دو D_2 انجام می‌شود. پنجره همینگ $w(n)$ با طول ۵۰۰ میلی ثانیه معادل با N نمونه روی ضرایب موجک اعمال می‌شود و مقدار در نظر گرفته شده برای شیفت پنجره ۲۵۰ میلی‌ثانیه است. سپس انرژی هر فریم محاسبه شده و فریمی که انرژی بیشتری در سطح دو تجزیه نسبت به فریم‌های دیگر دارد، به عنوان فریم محتمل (با اندیس j_p) که زمین‌خوردن در آن اتفاق افتاده در نظر گرفته می‌شود. در مرحله بعد، انرژی ضرایب موجک در آن فریم و چهار فریم قبل و بعد را در شش مقیاس محاسبه و نرمالایز کرده و به عنوان بردار ویژگی استفاده می‌کنیم. بدین ترتیب بردار ویژگی ما شامل ۵۴ مولفه می‌شود. استخراج ویژگی در فرمول‌های (۲) تا (۶) نشان داده شده است.

$$C(j) = \sum_{l=0}^N (w(n)D_2(n + jN/2))^2 \quad (2)$$

$$E_i(j) = \sum_{l=0}^N (w(n)D_i(n + jN/2))^2 \quad (3)$$

$$\tilde{E}_i(j) = \frac{E_i(j)}{\sum_{m=-M}^M E_i(j_p + m)}, \quad M = 4 \quad (4)$$

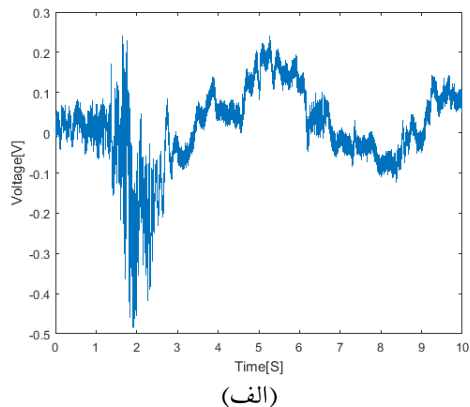
$$V(j) = [\tilde{E}_1(j), \tilde{E}_2(j), \dots, \tilde{E}_6(j)] \quad (5)$$

$$V = [V_{j_p-M}, V_{j_p-M+1}, \dots, V_{j_p+M}] \quad (6)$$

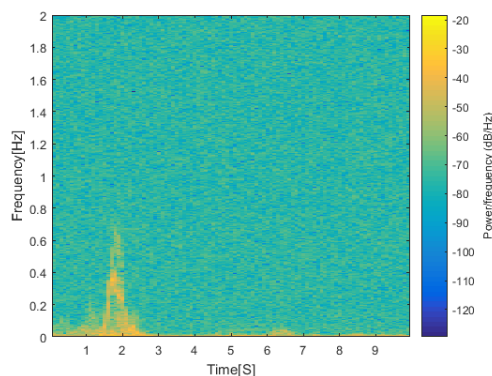
که در این روابط j اندیس فریم، i اندیس سطح تجزیه در مقیاس 2^i ، C انرژی کوتاه-مدت ضرایب ویولت در مرحله پیش‌آزمون، E انرژی کوتاه-مدت ضرایب ویولت، \tilde{E} انرژی‌های نرمالایز شده و V نمایانگر بردار ویژگی است. مرحله نهایی استفاده از ویژگی‌های استخراج شده به عنوان ورودی به یک طبقه‌بندی‌کننده است، که در آن بخشی از داده‌ها برای آموزش طبقه‌بندی‌کننده و بقیه برای آزمایش و ارزیابی عملکرد استفاده می‌شود [۶].

۳-۲- استخراج ویژگی بر اساس ضرایب کپسترال فرکانس مل

یکی از روش‌های استخراج ویژگی که در سیستم‌های تشخیص زمین‌خوردن استفاده شده، محاسبه ضرایب کپسترال فرکانس مل است. اساس این روش بر مبنای سیستم شنوایی انسان طراحی شده و به‌طور گسترده برای استخراج ویژگی از سیگنال گفتار به‌کار می‌رود. یکی از دلایل کارایی بالای این ضرایب، درجه وضوح بالای آن می‌باشد. به این



(الف)



(ب)

شکل ۱: شکل موج خروجی رادار داپلر موج پیوسته با فرکانس حامل ۲۴/۱۲۵ گیگاهرتز (الف) سیگنال در حوزه زمان (ب) نمایش اسپکتروگرام

۳- استخراج ویژگی از سیگنال رادار باهدف تشخیص زمین‌خوردن

روش پردازش کلی برای تشخیص زمین‌خوردن مبتنی بر رادار به این صورت است که ابتدا مرحله استخراج ویژگی روی سیگنال رادار اجرا می‌شود که در آن خصوصیات خاصی از داده راداری به منظور کاهش آلاطم‌های کاذب و افزایش نرخ تشخیص استخراج می‌شود. مرحله دوم، استفاده از ویژگی‌های استخراج شده به عنوان ورودی به یک طبقه‌بندی‌کننده است، که در آن بخشی از داده‌ها برای آموزش و بقیه برای آزمایش و ارزیابی عملکرد استفاده می‌شود. در ادامه دو روش استخراج ویژگی به‌کار رفته در این مقاله شرح داده می‌شود.

۳-۱- استخراج ویژگی بر اساس تبدیل موجک

معمولاً از تبدیل فوریه برای تجزیه و تحلیل داده‌ها در حوزه فرکانس استفاده می‌شود. سیگنال زمین‌خوردن مدت کوتاهی دارای فرکانس بالا و مدت طولانی دارای فرکانس پایین است [۶]. فرض بر این است که داده‌ها در تبدیل فوریه مشخصات آماری نامتغیر با زمان دارند، از این جهت تبدیل فوریه در تشخیص زمین‌خوردن مبتنی بر رادار مناسب

جمعا بردار ویژگی ۳۹ بعدی نهایی را برای هر فریم تشکیل می‌دهد. بردار ویژگی نهایی برای یک سیگنال رادار از پشت سرهم قرار دادن بردارهای ویژگی فریم‌ها تشکیل می‌شود.

۴- آزمایشات و نتایج

در این مقاله، طبقه‌بندی کننده‌های k -نزدیکترین همسایه، ماشین بردار پشتیبان و درخت تصمیم‌گیری برای ارزیابی دو روش استخراج ویژگی به کار رفته است. در ابتدا به شرح داده راداری مورد ارزیابی می‌پردازیم.

۴-۱- داده راداری مورد ارزیابی

داده مورد استفاده در این مقاله، توسط یک گروه تحقیقاتی ژاپنی و مشابه روالی که در مرجع [۱۷] معرفی شده جمع‌آوری و برای هدف تحقیقاتی ما در اختیار قرار گرفت. برای ثبت داده از یک رادار داپلر مایکروویو استفاده شده که رادار داپلر از نوع موج پیوسته است. رادار داپلر مایکروویو (InnoSenT: IPS154) در ۲۴.۱۲۵ گیگاهرتز و در ارتفاع ۲.۵ متر از کف متصل شده و اشعه مایکروویو را به کف پخش می‌کند. محدوده مورد نظر، یک دایره با شعاع ۱.۵ متر است و سنسور داپلر بر روی یک برد قرار گرفته تا سیگنال خروجی را به میزان ۴۰+ دسیبل تقویت کند. سیگنال تقویت شده در کامپیوتر، با استفاده از مبدل آنالوگ به دیجیتال با فرکانس نمونه‌برداری ۴۰۰۰ هرتز ذخیره می‌شود. در آزمایش‌ها، از افراد خواسته شد تا تلاش کنند سه نوع حرکت به عنوان رویدادهای زمین‌خوردن: (A1) زمین‌خوردن در حین قدم‌زدن، (A2) لیزخوردن و (A3) پشت‌پا خوردن؛ و سه نوع حرکت به عنوان رویدادهای غیرزمین‌خوردن: (B1) راه‌رفتن در محدوده مورد بررسی، (B2) نشستن و (B3) بلندشدن انجام دهند. هر حرکت توسط رادار داپلر به مدت ۱۰ ثانیه اندازه‌گیری شد [۱۷].

۴-۲- ارزیابی طبقه‌بندی کننده‌ها

در این بخش به ازای هر روش استخراج ویژگی، انواع طبقه‌بندی کننده‌ها را ارزیابی کرده و در انتها در مورد نتایج به بحث و بررسی می‌پردازیم. برای اعمال طبقه‌بندی کننده‌ها در نرم‌افزار MATLAB 2017b از ابزار Classification Learner App روی سیستمی با مشخصات CPU Corei7 و چهار گیگابایت حافظه RAM استفاده کردیم. طبقه‌بندی کننده نزدیکترین همسایه به ازای معیارهای فاصله k های مختلف، طبقه‌بندی کننده ماشین بردار پشتیبان به ازای کرنل‌های مختلف و طبقه‌بندی کننده درخت تصمیم‌گیری نیز به ازای معیارهای تقسیم مختلف در روش اعتبارسنجی متقابل ۵ لایه^۱ بر روی داده‌ها مورد بررسی قرار گرفت.

موارد زیر در ارزیابی تشخیص زمین‌خوردن مبتنی بر رادار انجام شد:

۱- در بررسی روش استخراج ویژگی مبتنی بر موجک، تعداد زیادی از خانواده موجک‌ها که برای تابع $\text{swt}(\cdot)$ در نرم‌افزار MATLAB تعریف شده و به ازای مرتبه‌های مختلف از موجک‌ها مورد ارزیابی جامع قرار گرفت.

معنی که تغییرات جزئی با استفاده از مقیاس مل اثر خود را به خوبی نشان می‌دهند. نقطه قوت دیگر این روش در استفاده از تبدیل کسینوسی گسسته^۲ می‌باشد که علاوه بر حذف جزئیات ساختار طیفی همبستگی بین ویژگی‌ها را از بین برده و عملیات طبقه‌بندی را بهبود می‌بخشد [۱۶].

برای استخراج ضرایب کپسترال به عنوان ویژگی ابتدا پنجره همینگ، رابطه (۷)، با طول N روی سیگنال راداری $y(n)$ اعمال می‌شود. پس از فریم‌بندی و اعمال پنجره از سیگنال تبدیل فوریه گسسته گرفته می‌شود. طیف سیگنال از تعدادی فیلتر با عرض باند مقیاس مل عبور داده می‌شود و عمل لگاریتم روی آن طبق رابطه (۹) اجرا می‌شود. به منظور کاهش تعداد مؤلفه‌های بردار ویژگی، مقادیر لگاریتمی خروجی فیلترهای مل در تبدیل کسینوسی گسسته ضرب شده و نتیجه به دست آمده برابر با تعداد ضرایب MFCC مورد نظر می‌باشد [۱۶].

$$w[n] = \begin{cases} 0.46 + 0.54 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n < N-1, \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

$$Y_{l,k} = \sum_{n=0}^{N-1} y(n+l.n_s)w(n)e^{-\frac{j2\pi kn}{N}}, \quad 0 \leq k \leq N-1 \quad (8)$$

$$\tilde{Y}_{l,m} = \log\left(\sum_{k=0}^{N-1} |Y_{l,k}| \cdot F_{k,m}\right), \quad 0 \leq m \leq M-1 \quad (9)$$

$$c_l[n] = \sum_{m=0}^{M-1} \tilde{Y}_{l,m} \cos\left(\frac{\pi n \left(m + \frac{1}{2}\right)}{M}\right) \quad 0 \leq m < M, \quad (10)$$

که در رابطه (۸) l شمارنده فریم، k اندیس ضرایب فوریه و n_s میزان شیفت پنجره است. در رابطه (۹)، F معرف فیلتر مل، m شمارنده فیلترها و M تعداد کل فیلترهای مل است. خروجی این مرحله برای هر فریم یک بردار است. در رابطه (۱۰) $\tilde{Y}_{l,m}$ لگاریتم خروجی m امین فیلتر و $c_l[n]$ معادل با n امین ضریب MFCC برای فریم l ام می‌باشد.

برای بالاتر بردن دقت تشخیص، می‌توان از مشتقات ضرایب کپسترال نسبت به زمان استفاده کرد. در حقیقت ضرایب کپسترال اطلاعات استاتیک سیگنال را مدل می‌کنند در حالی که مشتقات ضرایب کپسترال حاوی اطلاعات دینامیکی هستند. تلفیق ضرایب کپسترال و مشتقات آن می‌تواند ویژگی‌های بهتری از سیگنال را بیان نماید [۱۶]. روابط مربوط به محاسبه مشتقات در (۱۱) تا (۱۳) آمده است.

$$\Delta \bar{c}_l = \bar{c}_{l+2} - \bar{c}_{l-2} \quad (11)$$

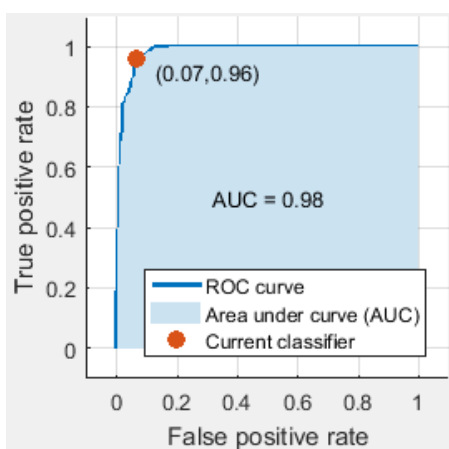
$$\Delta \Delta \bar{c}_l = \Delta \bar{c}_{l+1} - \Delta \bar{c}_{l-1} \quad (12)$$

$$\bar{x}_l = [\bar{c}_l \quad \Delta \bar{c}_l \quad \Delta \Delta \bar{c}_l] \quad (13)$$

در رابطه (۱۳)، \bar{x}_l معرف بردار ویژگی نهایی است. در این مقاله تعداد فیلترهای مل ۲۳ در نظر گرفته شده و ۱۳ ضریب کپسترال فرکانس مل به اضافه ۱۳ ضریب دلتا کپسترال و ۱۳ ضریب دلتا کپسترال

جدول ۲: سطح زیر منحنی مشخصه عملکرد سیستم (AUC)، دقت، حساسیت، صحت و امتیاز F1 برای انواع طبقه‌بندی کننده‌ها و دو روش استخراج ویژگی

		SVM	KNN	DT
WT (sym6)	AUC	0.91	0.83	0.78
	حساسیت	0.81	0.85	0.76
	دقت	0.9	0.81	0.77
	صحت	0.86	0.83	0.77
	امتیاز F1	0.85	0.83	0.77
MFCC	AUC	0.98	0.94	0.85
	حساسیت	0.93	0.93	0.84
	دقت	0.96	0.95	0.81
	صحت	0.95	0.94	0.82
	امتیاز F1	0.95	0.94	0.83



(ب)

شکل ۲: منحنی مشخصه عملکرد بهترین طبقه‌بندی کننده ماشین بردار پشتیبان در تشخیص زمین‌خوردن مبتنی بر ویژگی‌های MFCC

روش‌های تجزیه و تحلیل زمان-فرکانس از جمله تبدیل موجک، روی سیگنال‌های ثبت شده توسط رادار داپلر نقشی اساسی در استخراج ویژگی‌های زمین‌خوردن ایفا می‌کند. در این مقاله، علاوه بر مطالعه جامع و ارزیابی انواع موجک‌های مادر، ویژگی‌های ضرایب کپسترال فرکانس مل نیز برای تشخیص زمین‌خوردن با استفاده از سه خانواده طبقه‌بندی کننده SVM، KNN، و Decision Tree مورد بررسی قرار گرفت. بر اساس آزمایشات انجام شده، ویژگی‌های MFCC با طبقه‌بندی کننده ماشین بردار پشتیبان با کرنل گوسی بالاترین عملکرد آشکارسازی را نتیجه داد و عملکرد بسیار بالاتری نسبت به روش استخراج ویژگی مبتنی بر موجک نشان داد.

سپاسگزاری

از گروه تحقیقاتی ژاپنی در دانشگاه Aoyama Gakuin و پروفیسور Takashi Kaburagi که داده‌های راداری را در اختیار ما گذاشتند، کمال تشکر را داریم.

جدول ۱: مقایسه امتیاز F1، حساسیت، دقت و صحت طبقه‌بندی کننده ماشین بردار پشتیبان روی ویژگی‌های مبتنی بر موجک

موجک مادر	صحت	دقت	حساسیت	امتیاز F1
sym6	0.86	0.90	0.81	0.85
db8	0.86	0.90	0.81	0.85
dmey	0.84	0.91	0.77	0.83
coif4	0.85	0.91	0.77	0.84
bior4-4	0.83	0.86	0.79	0.83

در جدول (۱)، مقادیر چند معیار ارزیابی آشکارسازی (شامل حساسیت^{۱۱}، دقت^{۱۲}، صحت^{۱۳} و امتیاز F1^{۱۴}) برای چند نمونه موجک از جمله db8، sym6، coif4، bior4.4، dmey و برای بهترین طبقه‌بندی کننده SVM (با کرنل گوسی) نشان داده شده است. مقایسه انواع موجک‌ها تفاوت فاحشی را به لحاظ معیار F1-score نشان نمی‌دهد و انتخاب db8 و sym6 بهترین نتیجه را به لحاظ معیار F1-score با مقداری برابر با ۰٫۸۵، در پی داشته است.

۲- سه خانواده طبقه‌بندی کننده یعنی SVM، KNN و DT در پارامترهای مختلف از آنها برای دو روش استخراج ویژگی یعنی موجک با موجک مادر sym6 و MFCC مورد ارزیابی قرار گرفت. در هر خانواده طبقه‌بندی کننده بهترین نتیجه آشکارسازی به لحاظ معیار F1-score در جدول (۲) نمایش داده شده است.

با توجه به نتایج در جدول (۲)، درخت تصمیم‌گیری درمقایسه با دو روش طبقه‌بندی K-نزدیکترین همسایه و ماشین بردار پشتیبان عملکرد ضعیفتری را در هر دو روش استخراج ویژگی نشان می‌دهد. همچنین برای هر دو روش استخراج ویژگی، رتبه طبقه‌بندی کننده‌ها با مقایسه F1-score ها به این ترتیب است: ۱- ماشین بردار پشتیبان ۲- K-نزدیکترین همسایه ۳- درخت تصمیم‌گیری.

در مجموع، ویژگی‌های زمان-فرکانس مبتنی بر ضرایب کپسترال فرکانس مل و طبقه‌بندی کننده ماشین بردار پشتیبان با مقدار معیار F1-score برابر با ۰٫۹۵، بهترین نتیجه آشکارسازی زمین‌خوردن را در پی داشته و میزان بهبود نسبی آشکارسازی در این حالت نسبت به روش مبتنی بر موجک حدود ۱۰ درصد است. در شکل (۲) منحنی مشخصه عملکرد آشکارساز (گیرنده)^{۱۵} برای بهترین نتیجه آشکارسازی با ضرایب MFCC نمایش داده شده است. این آشکارساز، طبقه‌بندی کننده ماشین بردار پشتیبان با کرنل گوسی با میزان محدودیت جعبه^{۱۶} برابر با سه است که مقیاس کرنل^{۱۷} در آن بصورت اتوماتیک تنظیم شده است. طبق جدول (۲)، طبقه‌بندی کننده قابل رقابت از خانواده KNN است که بهترین نتیجه آن به ازای K برابر با یک، معیار فاصله همبستگی و در حالت استاندارد کردن ویژگی‌ها به دست آمد.

۵- نتیجه‌گیری

تشخیص سریع زمان واقعی زمین‌خوردن می‌تواند باعث کمک‌های پزشکی سریع و به این ترتیب، نجات جان، به حداقل رساندن آسیب و کاهش اضطراب سالمندانی شود که تنها زندگی می‌کنند. یک روش نوین، استفاده از فن‌آوری رادار برای تشخیص زمین‌خوردن سالمندان می‌باشد.

زیرنویس‌ها

مراجع

- ¹ Fall Detection
- ² Wavelet Transform (WT)
- ³ Mel Frequency Cepstral Coefficients (MFCC)
- ⁴ Support Vector Machine (SVM)
- ⁵ K Nearest Neighbor (KNN)
- ⁶ Decision Tree (DT)
- ⁷ Doppler Radar
- ⁸ Stationary Wavelet Transform (SWT)
- ⁹ Discrete Cosine Transform (DCT)
- ¹⁰ 5-fold Cross Validation
- ¹¹ Sensitivity
- ¹² Precision
- ¹³ Accuracy
- ¹⁴ F1-score
- ¹⁵ Receiver Operating Characteristic (ROC)
- ¹⁶ Box constraint level
- ¹⁷ Kernel scale

- [1] B. Jekanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in Proc. IEEE Radar Conf., May 2016, pp. 1–6.
- [2] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," IEEE Sensors J., vol. 15, no. 3, Mar. 2015.
- [3] V. Vaidehi, K. Ganapathy, K. Mohan, A. Aldrin, and K. Nirmal, "Video based automatic fall detection in indoor environment," in Proc. Int. IEEE ICRTIT Conf., Chennai, India, Jun. 2011, pp. 1016-1020.
- [4] F. Hussain, F. Hussain, M. Ehatisham-ul-Haq, M. A. Azam, "Aware Fall Detection and Recognition Based on Wearable Sensors," IEEE Sensors Journal. vol. 19, Issue: 12, 2019.
- [5] Y. Li, Z. Zeng, M. Popescu, and K. C. Ho, "Acoustic fall detection using a circular microphone array," in Proc. 32th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf., Buenos Aires, Argentina, Sep. 2010, pp. 2242-2245.
- [6] B. Y. Su, K. C. Ho, M. J. Rantz, and M. Skubic, "Doppler radar fall activity detection using the wavelet transform," IEEE Trans. Biomed. Eng., vol. 62, no. 3, pp. 865–875, 2015.
- [7] S. Tomii and T. Ohtsuki, "Falling detection using multiple Doppler sensors," in Proc. IEEE Int. Conf. e-Health Networking, Applications and Services, Beijing, China, Oct. 2012, pp. 196–201.
- [8] L. Liu, M. Popescu, M. Skubic, M. Rantz, T. Yardibi, and P. Cuddihy, "Automatic fall detection based on Doppler radar motion," in Proc. 5th Int. Conf. Pervasive Computing Technologies for Healthcare, Dublin, Ireland, May 2011, pp. 222-225.
- [9] M. Wu, X. Dai, Y. D. Zhang, B. Davidson, J. Zhang, and M. G. Amin, "Fall detection based on sequential modeling of radar signal time-frequency features," in Proc. IEEE Int. Conf. Healthcare Informatics, Philadelphia, PA, Sept. 2013, pp. 169–174.
- [10] F. Wang, M. Skubic, M. Rantz, and P. E. Cuddihy, "Quantitative gait measurement with pulse-Doppler radar for passive in-home gait assessment," IEEE Trans. Biomed. Eng., vol. 61, no. 9, pp. 2434–2443, Sept. 2014.
- [11] Q. Wu, Y. D. Zhang, W. Tao, and M. G. Amin, "Radar-based fall detection based on Doppler time-frequency signatures for assisted living," IET Radar Sonar Navig., vol. 9, no. 2, pp. 164–172, Feb. 2015.
- [12] T. Xu, Y. Zhou, J. Zhu, "New Advances and Challenges of Fall Detection Systems: A Survey," Appl. Sci., vol. 8, no. 3, p. 418, 2018.
- [13] X. Zhou, L.-C. Qian, P.-J. You, Z.-G. Ding, Y.-Q. Han, "Fall detection using convolutional neural network with multi-sensor fusion", Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW), Jul. 2018, pp. 1-5.
- [14] F. Pfanner, T. Allmendinger, T. Flohr, and M. KachelrieB, "Monitoring respiratory motion using continuous wave Doppler radar in a near field multi antenna approach," in Proc. Int. IEEE NSS/MIC Conf., Anaheim, CA, USA, Oct. 2012, pp. 3575-3581.
- [15] M. G. Amin, Y. D. Zhang, F. Ahmad, K. C. Ho, "Radar signal processing for elderly fall detection: The future for in-home monitoring", IEEE Signal Processing Magazine, vol. 33, no. 2, pp. 71-80, 2016.
- [16] U. Zafar Khan, A. Wahid, Dr. U. Akram, "Detection of Acoustic Events by using MFCC and Spectro-Temporal Gabor Filterbank Features," Proceedings of the 8th International Conference on Signal Processing Systems, Auckland, New Zealand, November. 2016, pp. 158-164.
- [17] K. Shiba, T. Kaburagi, and Y. Kurihara, "Fall Detection utilizing Frequency Distribution Trajectory by Microwave Doppler Sensor," IEEE Sensors Journal, vol. 17, pp. 7561–7568, Nov. 2017.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



۲۷ الی ۲۸ آذرماه ۱۳۹۸

افزایش کیفیت تصاویر درماتوسکوپی ملانوما با استفاده مناسب از ضرایب موجک در ساختار یادگیری عمیق

هاله فاتح^۱، محمد تشنه لب^۲، منصور فاتح^۳

^۱ دانشجوی کارشناسی ارشد هوش مصنوعی، دانشکده برق و کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران

h.fateh@email.kntu.ac.ir

^۲ استاد تمام، دانشکده برق و کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران

teshnehlab@eetd.kntu.ac.ir

^۳ استادیار، دانشکده برق و کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود

mansoor_fateh@shahroodut.ac.ir

چکیده

در تشخیص مبتنی بر تصاویر بیماری‌ها، کیفیت مناسب تصاویر دقت تشخیص بیماری را افزایش می‌دهد. سرطان‌های پوست از شایع‌ترین سرطان‌های بدن انسان به‌شمار می‌رود و ملانوما یکی از خطرناک‌ترین انواع آن است. درماتوسکوپ وسیله‌ای غیرتهاجمی برای عکس‌برداری از آسیب‌های پوستی است. در این مقاله با استفاده از روش یادگیری عمیق به افزایش کیفیت تصاویر درماتوسکوپی پرداخته شده است. در روش پیشنهادی، تخمین ضرایب جزئیات موجک یک تصویر با رزولوشن بالا، با رزولوشن پایین انجام می‌شود. این تخمین با استفاده از شبکه SRCNN انجام می‌شود. استفاده از ضرایب موجک، امکان آموزش شبکه با تعداد نمونه‌های کمتر در زمان کمتر را میسر می‌کند. استفاده مناسب از این ضرایب و اعمال تغییراتی مناسب بر روی آن، ارتقای کیفیت تصویر نهایی را نیز به ارمغان آورده است که این کیفیت را با معیار PSNR می‌سنجیم. خروجی PSNR روش پیشنهادی برای چندین تصویر استاندارد ۱۸۰,۹۹ است درحالی‌که بهترین خروجی روش‌های پیشین بر روی این تصاویر ۴۴,۰۲ است.

کلمات کلیدی

ملانوما، یادگیری عمیق، موجک، درماتوسکوپ، تصویربرداری با وضوح بالا

ضایعات پوستی استفاده می‌کنند. اما دقت این دستگاه هنوز برای این امر خطیر کافی نیست.

امروزه حجم دادگان موجود در تصاویر و ویدیوها ما را برآن می‌دارد تا ابزارهای مفیدی جهت استخراج اطلاعات از تصاویر و دنباله‌های ویدیویی خلق کنیم تا بتوانیم بدون دخالت انسان به طبقه‌بندی یا تحلیل محتوا پردازیم.

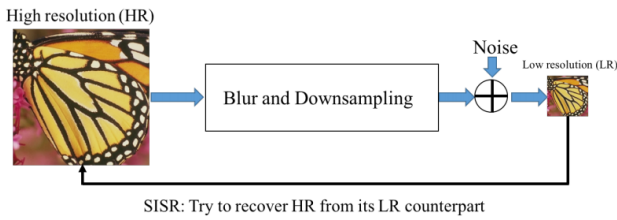
تهیه تصاویر با کیفیت بالا، از جمله علوم مورد پژوهش و پر اهمیت در عصر حاضر است. عدم وضوح تصاویر در بسیاری از موارد، مانند تشخیص‌های پزشکی، متضمن هزینه‌های سنگین و بروز مخاطرات جانی برای افراد است [۳].

گاه به دلیل مناسب نبودن کیفیت یا وجود نویز در تصاویر، امکان تفسیر دقیق تصویر برای پزشک مقدور نبوده و نیاز به تجدید عکس برداری است. عکس برداری مجدد نیز دارای عوارضی برای بیماران و صرف وقت و هزینه است. از این‌رو نیازمند افزایش کیفیت تصاویر هستیم. از

۱- مقدمه

ملانوما یکی از شایع‌ترین انواع سرطان پوست است که تشخیص آن در مراحل نخست باعث کاهش مرگ‌ومیر ناشی از این سرطان مهلک می‌شود. لذا تشخیص خوش‌خیم یا بدخیم بودن و موقعیت تومور آن در پزشکی از جایگاه ویژه‌ای برخوردار است [۱]. برای پیشگیری و تشخیص به هنگام در طول ۲۰ سال گذشته، چندین روش عکس برداری برای این منظور پیشنهاد شده است. با این حال عوارض جانبی این روش‌ها سبب می‌شود تا تنها در مواقع ضروری از آن‌ها استفاده شود.

تصاویر پزشکی بزرگترین منبع اطلاعات در تشخیص و مراقبت‌های پزشکی و در عین حال یکی از سخت‌ترین منابع برای تجزیه و تحلیل است. امروزه پزشکان باید تجزیه و تحلیل تصویر پزشکی انجام شده توسط متخصصان رادیولوژیست را جهت تشخیص به صورت دقیق بررسی نمایند [۲]. متخصصان از دوربین درماتوسکوپ برای مشاهده



شکل ۲: معماری روش SISR [۶]

رزولوشن بالای تصویر واحد، یک مشکل کلاسیک در حیطه وظایف رایانه‌ای است [۷]. اخیراً روش‌های SISR مبتنی بر یادگیری عمیق، به ویژه شبکه‌های عصبی پیچشی، عملکرد بسیار پیشرفته‌ای را در زمینه‌های مختلفی از قبیل تشخیص چهره، تصویربرداری پزشکی و سیستم‌های نظارتی و تصاویر طبیعی بدست آورده‌اند [۸]. در این مقاله یک روش SISR مبتنی بر نمونه خارجی ارائه شده است. در این روش ضرایب جزئیات معکوس موجکی یک تصویر دلخواه HR، بوسیله یک شبکه عصبی کانولوشن به عنوان ماشین یادگیری، بر اساس تصویر LR تخمین زده می‌شود. این روش می‌تواند جایگزین ارتقای پرهزینه سخت‌افزاری در برنامه‌های کاربردی مانند بزرگنمایی دوربین‌ها، تصویربرداری پزشکی و تصویربرداری ماهواره‌ای باشد.

فرض متداول این است که تصویر LR، با روش درون‌یابی Bicubic از تصویر با رزولوشن بالا ساخته شده است. SISR به روش‌های درون‌یابی، مبتنی بر بازسازی، مبتنی بر یادگیری [۹] و مبتنی بر نمونه‌های خارجی [۱۰] دسته‌بندی می‌شوند.

در روش‌های مبتنی بر یادگیری، تصویر اصلی به کمک تصویر نمونه‌برداری شده مبتنی بر روش‌های آماری، تخمین زده می‌شود. هنگامی که مجموعه داده‌های یادگیری شبکه (دریافت شده از داده‌های آموزش)، با داده‌های آزمون در یک راستا باشند، این روش وضوح فوق العاده مبتنی بر مثال خارجی شناخته می‌شود [۱۱]. از جمله مزیت‌های SISR مبتنی بر نمونه خارجی، زمان اجرای سریع آن بر روی تصاویر آزمون است که امکان اجرای آن بر روی سخت افزار منابعی نظیر تلفن‌های هوشمند را فراهم می‌کند. نوآوری روش SISR معرفی شده، استفاده از ضرایب معکوس موجکی می‌باشد.

در ادامه این مقاله در بخش دوم، مروری بر پژوهش‌های پیشین در این زمینه ارائه شده است. در بخش سوم روش پیشنهادی تشریح گشته است. بخش چهارم شامل نتایج حاصل از معماری شبکه پیشنهادی می‌باشد. به نتیجه‌گیری و کارهای آینده در بخش پنجم پرداخته شده است.

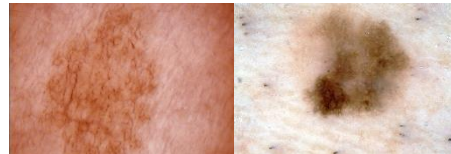
۲- پژوهش‌های پیشین

از آنجایی که یادگیری عمیق امروزه کاربرد و تاثیر بسیاری در پردازش تصویر دارد، دونگ و همکاران [۱۲] برای اولین بار یک شبکه عصبی پیچشی به نام SRCNN^{۱۱} و در ادامه کیم و همکاران [۱۳] شبکه‌ای عمیق‌تر با ۲۰ لایه کانولوشن به نام VDSR^{۱۲} با روند آموزش آسان‌تر معرفی کردند و همچنین یک شبکه عمیق کاملاً بازگشتی [۱۴] با لایه پیچشی و نظارتی بازگشتی پیشنهاد کردند که به کنترل مقدار پارامترهای

راهکارهای بهبود این امر می‌توان به تجهیز دستگاه‌های تصویر برداری پزشکی اشاره نمود که هزینه بسیاری دارد. در واقع بدست آوردن تصاویر با رزولوشن بالاتر نیازمند اسکن چندگانه توسط دستگاه‌های پیشرفته‌تر و تصویربرداری در مدت زمان بیشتر است.

فناوری هوش مصنوعی در سال‌های اخیر رشد چشمگیری داشته و ابزار مورد استفاده در این فناوری توانایی افزایش کیفیت تصاویر با وضوح پایین را دارد.

با ارتقای کیفیت و بهبود مصورسازی بافت‌های کوچک در تصاویر درماتوسکوپی، تشخیص نهایی در زمینه خوش‌خیم یا بدخیم بودن توده دقیق‌تر خواهد شد [۴]. امروزه روش‌های یادگیری عمیق برای تجزیه و تحلیل تصاویر پزشکی، امکان توسعه سیستم‌های تشخیص هوشمند مبتنی بر تصویربرداری پزشکی را فراهم آورده است که می‌تواند به متخصص انسانی در تصمیم‌گیری‌های بهتر درباره سلامتی بیماران کمک کنند [۵]. در شکل (۱) نمونه‌ای از تصویر درماتوسکوپی از ضایعه ملانوما و نرمال نشان داده شده است.



شکل ۱: نمایی از تصاویر درماتوسکوپی

یادگیری عمیق^۱ و زیر مجموعه‌های آن از جمله شبکه عصبی پیچشی عمیق^۲ کاربرد بسیاری در زمینه‌های بینایی ماشین از جمله بهبود تصویر، حذف نویز^۳ و حذف مصنوعات^۴ را دارد.

یکی از عوامل موثر در کیفیت تصاویر، تصویر برداری در مدت زمان بیشتر است که اجرای آن محدودیت‌های بسیاری از نظر تجهیزات برای کادر درمانی و از نظر وضعیت پایدار جسمی برای بیماران به همراه دارد. به عنوان مثال در بیماری‌های ریوی به دلیل آنکه بیماران قادر به حبس طولانی مدت نفس خود نیستند، بدست آوردن تصاویر با کیفیت بالا در اسکن^۵ بسیار دشوار است [۴].

بازیابی تصاویر با وضوح بالا^۶ از یک یا چند تصویر تخریب شده با وضوح پایین^۷، فرآیند وضوح فوق العاده^۸ نام دارد. با توجه به تعداد تصاویر LR^۹ ورودی، SR^{۱۰} را می‌توان به وضوح فوق العاده تصویر واحد [SISR]^۹ و وضوح فوق العاده چند تصویر [MISR]^{۱۰} طبقه‌بندی کرد [۷]. در شکل (۲) معماری روش SISR نشان داده شده است.

¹ Deep Learning

² Deep Convolutional Neural Network [DCNN]

³ Denoising

⁴ Deblurring

⁵ HRCT [High-Resolution CT]

⁶ High Resolution [HR]

⁷ Low Resolution [LR]

⁸ Super Resolution [SR]

⁹ single image super-resolution [SISR]

¹⁰ multi-image super-resolution [MISR]

¹¹ Super-Resolution Convolutional Neural Network

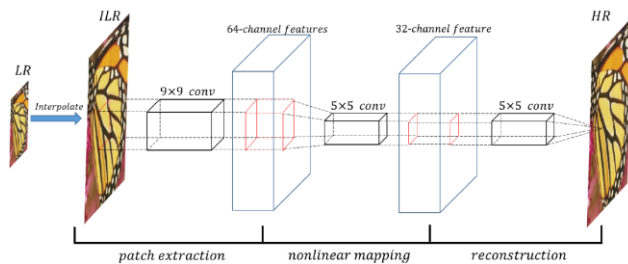
¹² Very deep convolutional network

جهت تبدیل یک تصویر با LR (زیرگروه تقریب^{۲۲}) به HR (زیرگروه جزئیات^{۲۳}) بود، استفاده می‌شود.

۳- روش پیشنهادی

یادگیری عمیق، یک دسته از الگوریتم‌های یادگیری ماشین است. این نوع یادگیری، به معنی استفاده از شبکه‌های عصبی داده‌محور با تعداد لایه‌های مخفی زیاد است. در یادگیری عمیق، به صورت خودکار، مهندسی ویژگی انجام می‌شود. بدین معنا که شبکه، به صورت خودکار، به استخراج ویژگی و الگوهای مناسب از داده‌های آموزشی می‌پردازد. این ویژگی اساسی یادگیری عمیق، از دخالت انسان در استخراج ویژگی و وابستگی نتایج به این مرحله جلوگیری می‌کند [۱۰].

اخیرا از زمینه بینایی ماشین شبکه SRCNN^{۲۴} به عنوان یک روش مبتنی بر یادگیری عمیق برای تصویربرداری با وضوح بالا ارائه شده است. شبکه SRCNN می‌تواند به صورت مستقیم یک نگاشت بین تصاویر با کیفیت پایین و بالا ایجاد نماید [۱۲]. مطالعات اخیر نشان داده است که استفاده از SRCNN برای تصویربرداری در پزشکی و سایر زمینه‌ها، از نظر کیفیت تصویر و سرعت پردازش، دارای عملکرد برتر نسبت به روش‌های پیشین است. این روش در تصویربرداری پزشکی در مقایسه با روش‌های معمول درون‌یابی خطی^{۲۵}، به صورت قابل توجهی کیفیت تصویر را بهبود می‌بخشد [۳۲] و به عنوان روشی موثر برای بهبود تصاویر در رادیوگرافی، سی تی اسکن قفسه سینه [۴] و تصاویر ماموگرام سینه [۳۳] استفاده شده است. در شکل (۳) نمایی از معماری SRCNN را مشاهده می‌کنید.



شکل ۳: معماری SRCNN [۶]

تبدیل موجک^{۲۶} یک تصویر، تجزیه فضایی - فرکانسی است که توسط آن یک فیلتر بر روی تصویر اعمال می‌شود.

تبدیل موجک گسسته^{۲۷} به دلیل داشتن ویژگی‌های منحصر به فرد نسبت به سایر تبدیل‌های حوزه فرکانس، کارایی بهتری دارد و همین موضوع سبب شده است که در پردازش تصویر بیشتر مورد استفاده قرار گیرد [۳۴]. از جمله این ویژگی‌ها می‌توان به نمایش چندریزی^{۲۸} آن اشاره کرد. در واقع قدرت چندریزی، توانایی تجزیه یک سیگنال

مدل و بهبود عملکرد شبکه کمک می‌کند. شبکه DRCN^{۱۳} بر اساس معماری باقیمانده از هی و همکاران [۱۶] طراحی شده است. از آنجا که روش پیشنهادی در گروه مبتنی بر نمونه و با استفاده از تبدیل موجک است، در ادامه روش‌های SISR را در این دو دسته مرور می‌شود.

۱-۲- روش‌های SISR مبتنی بر نمونه

از اولین پژوهش‌ها در SISR مبتنی بر نمونه، نگاشت متناظر قطعه به قطعه یک تصویر HR به قطعه متناظر در تصویر LR، با قاعده مخصوص همان قطعه است. فریمن و همکاران [۱۷] گلاسز و همکاران [۱۸] و زونتاک و ایرانی [۱۹] با استفاده از مدل‌های تصادفی مارکوف به بهبود این امر پرداختند.

تیموفته و همکاران [۲۰] با استفاده از رگرسیون محلی قاب‌ها^{۱۴}، برای نگاشت ساختار محلی تصویر از قطعات تصویر با LR به تصویر با HR پرداختند. یانگ و همکاران [۲۱]، زید و همکاران [۲۲] برای غلبه بر مشکلات حافظه زیاد و الزامات محاسباتی، رویکردهای مبتنی بر کدگذاری تنک را^{۱۵} پیشنهاد کردند. وانگ و همکاران [۲۳]، پله و الاد [۲۴] روش‌های نگاشت غیرخطی بین تصاویر LR و HR با استفاده از بردارهای پایه تنک استفاده کردند.

بر اساس مطالعه ستی و همکاران [۲۵]، دونگ و همکاران [۲۶] روش‌های غیرخطی ارائه شده توانستند دقت بازسازی را در نگاشت بین تصاویر LR به HR در مقایسه با روش‌های خطی، بهبود ببخشند. همچنین دونگ و همکاران به طور خاص، یک رویکرد مبتنی بر شبکه‌های عصبی کانولوشنی برای یادگیری نگاشت قطعه‌های مقیاس ناپذیر تصویر با LR با استفاده از درون‌یابی به قطعه‌های تصویر با HR ارائه دادند.

۲-۲- روش‌های SISR مبتنی بر دامنه موجک

تبدیل موجک^{۱۶} یک تصویر، تجزیه فضایی - فرکانسی است. در تبدیل موجک، یک فیلتر بر روی تصویر اعمال می‌شود که این فیلتر موجب می‌شود. خروجی موجک تصویر تبدیل شده در حالت Low-pass و High-pass باشد [۲۷]. این روش از مدل‌های چند مقیاسی و ویژگی‌های تنک موجک در تصاویر استفاده می‌کند.

کروس و همکاران [۲۸]، جاستین کی رومیگ و همکاران [۲۹]، کافوریو و همکاران [۳۱]. به بررسی ویژگی‌های تبدیل موجک در زمینه‌های وسیعی از جمله فشرده سازی تصویر^{۱۷}، حذف نویز^{۱۸}، بهبود کیفیت تصویر^{۱۹} و تقویت لبه^{۲۰} پرداختند. در این تحقیق، از بحث مورد توجه در پژوهش‌های گذشته که در رابطه با استفاده از تبدیل موجک^{۲۱}

¹³ Deeply-recursive convolutional network (DRCN)

¹⁴ Anchored neighborhood regression

¹⁵ sparse coding

¹⁶ wavelet transform

¹⁷ image compression

¹⁸ noise elimination

¹⁹ Image enhancement

²⁰ edge enhancement

²¹ Wavelet Tranfrom

²² Approximate Band

²³ Detail Band

²⁴ Super-Resolution Convolutional Neural Network

²⁵ Conventional Linear Interpolation

²⁶ wavelet transform

²⁷ Discrete Wavelet Transform

²⁸ Multi-resolution

همانطور که در شکل ۴ نشان داده شده است، اولین لایه از شبکه، تصویر LR را از ورودی استخراج می‌کند و در لایه دوم تصویر سطح پایین را به موجک نگاشت می‌دهد. در نهایت، با ترکیب در یک همسایگی فضایی ضرایب موجک برای سه زیر باند دقیق تولید می‌شوند.

در روش پیشنهادی این مقاله مشابه مرجع [۱۰] یک روش SISR ارائه شده است با این تفاوت که در این روش، تخمین ضرایب معکوس موجک یک تصویر با رزولوشن بالا، بر اساس تصویر با رزولوشن پایین، محاسبه می‌شود. بدین منظور از شبکه عصبی کانولوشن به عنوان یک ماشین یادگیری استفاده شده است.

۱-۳- معیار ارزیابی کیفیت تصاویر

ارزیابی میزان تغییرات کیفیت تصاویر توسط دو روش کلی مقایسه توسط انسان براساس سیستم بینایی و روش‌های ارزیابی کمی انجام می‌پذیرد. روش‌های ارزیابی کمی، براساس نوع معیار به دو دسته اصلی متکی بر تصویر مرجع و مستقل از آن تقسیم می‌شوند.

PSNR یکی از متداولترین روش‌های متکی بر تصویر مرجع برای اندازه‌گیری کیفیت تصاویر است. هرچه میزان PSNR بیشتر باشد، بازسازی تصویر بهتر صورت گرفته است.

ساده ترین معیار ارزیابی که به طور گسترده مورد استفاده قرار می‌گیرد، MSE است که در زمینه بهینه‌سازی استفاده می‌شود.

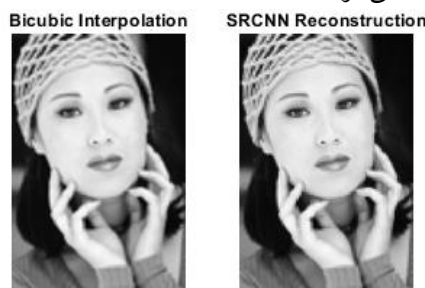
$$PSNR = 10 * \log_{10} \frac{MAX^2}{MSE} \quad (1)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (2)$$

MAX بیشترین مقدار پیکسل موجود در تصویر است. i و j سطر و ستون تصویر را نشان می‌دهد. I و K به ترتیب تصویر رزولوشن بالای اصلی و تصویر بازسازی شده است. برای مقایسه بهتر ابتدا روش پیشنهادی را بر روی همان دادگان مرجع [۱۰] اجرا نموده و سپس به بررسی آن در تصاویر درماتوسکوپی ملانوما پرداخته می‌شود.

۴- نتایج

در این مقاله ابتدا از روش درون یابی Bicubic رزولوشن تصاویر افزایش می‌یابد سپس روش SRCNN پایه برای افزایش کیفیت تصاویر اعمال می‌شود.



شکل ۵: خروجی حاصل از روش SRCNN و Bicubic

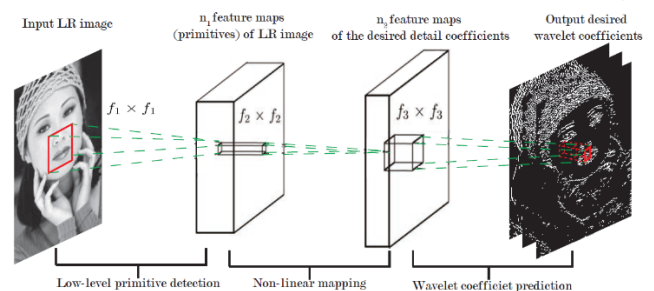
پس از محاسبه معیار PSNR مطابق جدول (۲) برای دو روش فوق به برتری SRCNN با $PSNR = 37.887617$ dB نسبت به روش Bicubic با $PSNR = 37.690386$ dB مشخص می‌گردد.

به جزئیات کوچک است. از آنجا که سیستم بینایی انسان^{۲۹} حساسیت‌های مختلفی نسبت به باندهای فرکانسی گوناگون دارا می‌باشد، به تجزیه چندریزیگی تبدیل موجک بسیار شبیه است، بنابراین می‌تواند به خوبی سیستم بینایی انسان را مدل‌سازی کند [۳۵]. همچنین این تبدیل برای شناسایی نواحی فرکانسی سیگنال تصویر بسیار مناسب است [۳۶]. با توجه به دوبعدی بودن تصاویر، اگر یک تصویر توسط تبدیل موجک گسسته مورد تجزیه قرار گیرد، چهار تصویر بدست می‌آید. این تصاویر شامل یک تصویر مربوط به کلیات و سه تصویر مربوط به جزئیات [جزئیات افقی، عمودی و قطری] است.

پس از معرفی تبدیل موجک و عملکرد خوب آن، محققان تبدیل‌های مختلفی را بر این اساس ایجاد کردند. این تبدیل‌ها توسعه‌ای بر موجک کلاسیک جهت بهبود آن هستند [۳۷]. در زمینه کاربرد تبدیل موجک در SISR کومار و همکاران [۱۱] به تخمین ضرایب جزئیات موجک یک تصویر HR با استفاده از تصویر LR پرداختند. این تخمین با استفاده از شبکه SRCNN انجام می‌شود. استفاده از ضرایب موجک، امکان آموزش شبکه با تعداد نمونه‌های کمتر در زمان کوتاه‌تر را میسر می‌کند.

روش مطرح شده در مرجع [۱۱] بر خلاف مدل‌های پیشین از یک رویکرد متمایز برای تخمین ضرایب جزئیات موجک استفاده می‌کند. این روش با تخمین یک خروجی تک^{۳۰} توسط شبکه، دارای دقت بازسازی بالاتر و سرعت آزمون بهینه در مقایسه با روش‌های پیشین است. در این روش، از یک شبکه پیچشی برای تخمین ضرایب جزئیات قطعه‌های یک تصویر LR (زیر باند تقریبی)^{۳۱} استفاده می‌شود که آن را از روش پیشنهادی دونگ و همکاران [۲۶] برای تخمین پیکسل‌های تصویر HR به صورت مستقیم، متمایز می‌سازد.

در شکل (۴) شمایی از روش پیشنهادی در مرجع [۱۰] را مشاهده می‌کنید:

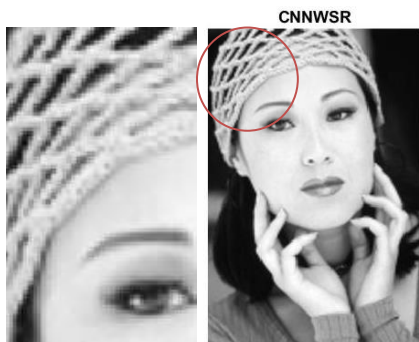


شکل ۴: شمایی از روش پایه [۱۰]

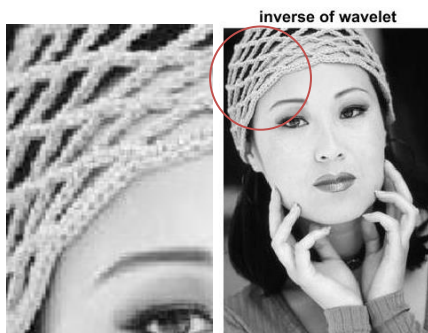
²⁹ Human Visual System [HVS]

³⁰ sparse

³¹ approximation sub-band



شکل ۸: خروجی حاصل از روش CNNWSR



شکل ۹: خروجی حاصل از روش پیشنهادی

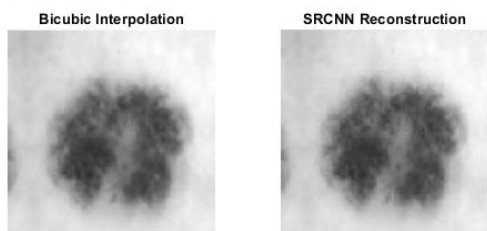
بر اساس جدول (۲) مشاهده می‌شود بهترین نتایج بر اساس معیار PSNR در روش پیشنهادی مقاله بدست می‌آید و همچنین مشاهده می‌شود که بافت و لبه در این روش دارای تفکیک پذیری و وضوح بالاتری است.

جدول ۲: مقایسه معیار PSNR در روش های موجود

روش	PSNR
Bicubic	37.690386 dB
SRCNN	37.887617 dB
CNNWSR	38.219048 dB
روش پیشنهادی	181.704783 dB

جدول (۲) به بررسی معیار PSNR در روش‌های مطرح شده، می‌پردازد که بهترین نتایج بر اساس معیار PSNR در روش پیشنهادی مقاله بدست می‌آید.

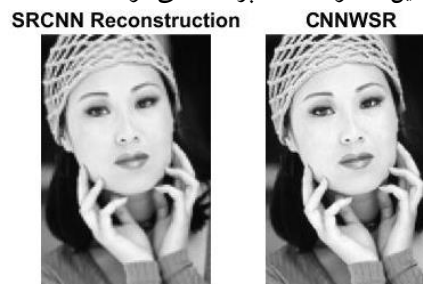
در ادامه معماری‌های فوق بر روی دادگان ملانوما پیاده‌سازی می‌گردد.



شکل ۱۰: خروجی حاصل از روش SRCNN و Bicubic در تصاویر درماتوسکوپی ملانوما

پس از محاسبه معیار PSNR مطابق جدول (۲) برای دو روش فوق برتری SRCNN با $PSNR = 43.604867$ dB نسبت به روش Bicubic با $PSNR = 43.511114$ dB مشخص می‌گردد.

در ادامه مشابه روش CNNWSR [۱۰]، با استفاده از تخمین ضرایب جزئیات موجک یک تصویر با رزولوشن بالا، از روی تصویر با رزولوشن پایین به افزایش کیفیت تصاویر در مقایسه با SRCNN که روش برتر پیشین تا کنون است، پرداخته می‌شود.



شکل ۶: خروجی حاصل از روش SRCNN و CNNWSR

پس از محاسبه معیار PSNR مطابق جدول (۲) در روش CNNWSR با $PSNR = 38.219048$ dB برتری این روش نسبت به دو روش پیشین مشخص می‌گردد.

یکی از نوآوری‌های مقاله اینست که در روش CNNWSR تبدیل موجک را تا سه سطح ادامه داده و سپس به بررسی تاثیرات و نتایج آن می‌پردازد.



شکل ۷: خروجی حاصل از تبدیل موجک در سه سطح

بر اساس جدول (۱) مشاهده می‌شود بهترین نتایج شبکه و بالاترین میزان PSNR در صورتی است که ضرایب سطح اول موجک را تخمین بزنیم.

جدول ۱: مقایسه معیار PSNR در تبدیل موجک در سه سطح

METHOD	PSNR
1level wavelet CNNWSR	38.219048 dB
2level wavelet	33.875647 dB
3level wavelet	30.198194 dB

نوآوری دیگر این مقاله تخمین ضرایب جزئیات تبدیل موجک معکوس یک تصویر با رزولوشن بالا، از روی تصویر با رزولوشن پایین برای افزایش کیفیت تصاویر است. در شکل (۸) و (۹) روش پیشنهادی مقاله با CNNWSR مقایسه می‌گردد.

۵- نتیجه‌گیری و پیشنهاد برای کارهای آتی

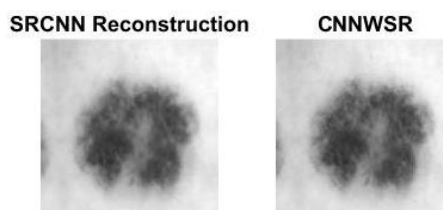
در این مقاله به بررسی روش‌های افزایش وضوح تصاویر همچون SRCNN، Bicubic و CNNWSR پرداخته شد. در ادامه یک روش SISR مبتنی بر مثال خارجی ارائه شد که ضرایب معکوس موجک یک تصویر با رزولوشن بالا دلخواه را بر اساس تصویر با رزولوشن پایین داده شده، تخمین می‌زند. با بررسی معیار PSNR به مقایسه روش پیشنهادی با سایر روش‌های مطرح پیشین پرداخته و مشاهده شد که در روش پیشنهادی تصاویر با رزولوشن بالا با کیفیت بالاتری از تصاویر با رزولوشن پایین بازسازی می‌گردد.

به منظور بهبود نتایج موجود پیشنهاد می‌گردد در پژوهش‌های آتی علاوه بر تبدیل موجک پایه، از سایر تبدیل‌های توسعه یافته موجک از جمله موجک‌های چند جهت برای شناسایی لبه‌ها، تبدیل موجک گسسته افزونه^{۳۳} و مرکزک‌ها^{۳۲} استفاده شود.

۶- منابع

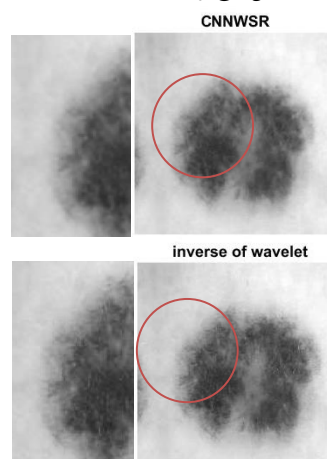
- [1] Zhang XJCAS. Melanoma segmentation based on deep learning. 2017;22[sup1]:267-77.
- [2] Tran GS, Nghiem TP, Nguyen VT, Luong CM, Burie J-C. Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss. Journal of Healthcare Engineering. 2019;2019.
- [3] Geras KJ, Wolfson S, Shen Y, Wu N, Kim S, Kim E, et al. High-resolution breast cancer screening with multi-view deep convolutional neural networks. arXiv preprint arXiv:170307047. 2017.
- [4] Umehara K, Ota J, Ishida T. Application of super-resolution convolutional neural network for enhancing image resolution in chest CT. Journal of digital imaging. 2018;31[4]:441-50.
- [5] Lopez AR, Giro-i-Nieto X, Burdick J, Marques O, editors. Skin lesion classification from dermoscopic images using deep learning techniques. 2017 13th IASTED International Conference on Biomedical Engineering [BioMed]; 2017: IEEE.
- [6] Yang W, Zhang X, Tian Y, Wang W, Xue J-H, Liao QJIToM. Deep learning for single image super-resolution: A brief review. 2019.
- [7] Ziwei L, Chengdong W, Dongyue C, Yuanchen Q, Chunping W, editors. Overview on image super resolution reconstruction. The 26th Chinese Control and Decision Conference [2014 CCDC]; 2014: IEEE.
- [8] Zhao X, Zhang Y, Zhang T, Zou XJIToIP. Channel splitting network for single MR image super-resolution. 2019;28[11]:5649-62.

در ادامه مشابه روش CNNWSR [۱۰]، با استفاده از تخمین ضرایب جزئیات موجک یک تصویر با رزولوشن بالا، از روی تصویر با رزولوشن پایین به افزایش کیفیت تصاویر در مقایسه با SRCNN که روش برتر پیشین تا کنون است، پرداخته می‌شود.



شکل ۱۱: خروجی حاصل از روش SRCNN و CNNWSR در تصاویر درماتوسکوپی ملانوما

پس از محاسبه معیار PSNR مطابق جدول (۳) در روش SRCNN با $PSNR=44.025949$ dB برتری این روش نسبت به دو روش پیشین مشخص می‌گردد.



شکل ۱۲: خروجی حاصل از روش CNNWSR و روش پیشنهادی در تصاویر درماتوسکوپی ملانوما

بر اساس جدول (۳) مشاهده می‌شود در تصاویر درماتوسکوپی ملانوما بهترین نتایج بر اساس معیار PSNR در روش پیشنهادی مقاله به دست می‌آید. همچنین بافت و لبه در روش پیشنهادی دارای تفکیک پذیری و وضوح بالاتری است.

جدول ۱۵: مقایسه معیار PSNR در روش‌های موجود بر روی دادگان ملانوما

روش	PSNR
Bicubic	43.511114 dB
SRCNN	43.604867 dB
CNNWSR	44.025949 dB
روش پیشنهادی	180.998495 dB

بر اساس جدول (۳) مشاهده می‌گردد روش پیشنهادی مقاله در تصاویر درماتوسکوپی ملانوما پوستی نیز برای افزایش رزولوشن، خروجی PSNR بهتری نسبت به روش‌های پیشین دارد.

³² RDWT

³³ Contourlet Transform

- [24] Peleg T, Elad MJ. A statistical prediction model based on sparse representations for single image super-resolution. 2014;23[6]:2569-82.
- [25] Sethi N, Sharma DJ. A novel method of image encryption using logistic mapping. 2012;1[2]:115-9.
- [26] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. 2015;38[2]:295-307.
- [27] Mallat SJA, Burlington, Mass. *A Wavelet Tour of Signal Processing: The Sparse Way*, 832 pp. 2008.
- [28] Crouse M, Nowak RD, Baraniuk RG. Wavelet-based statistical signal processing using hidden Markov models. 1998;46[4]:886-902.
- [29] Romberg JK, Choi H, Baraniuk RG. Bayesian tree-structured image modeling using wavelet-domain hidden Markov models. 2001;10[7]:105-114.
- [30] Romberg JK, Choi H, Baraniuk RG, editors. *Multiscale edge grammars for complex wavelet transforms*. Proceedings 2001 International Conference on Image Processing [Cat No 01CH37205]; 2001: IEEE.
- [31] Di Sciascio E, Guaragnella C. Edge enhancement for subband coded images. 2001;40[5]:729-39.
- [32] Ren H, El-Khamy M, Lee J, editors. *CT-SRCNN: cascade trained and trimmed deep convolutional neural networks for image super resolution*. 2018 IEEE Winter Conference on Applications of Computer Vision [WACV]; 2018: IEEE.
- [33] Umehara K, Ota J, Ishida T. Super-resolution imaging of mammograms based on the super-resolution convolutional neural network. *Open J Med Imaging*. 2017;180.
- [34] Kaviani HR, Karimi N, Samavi S, editors. *Robust watermarking in singular values of contourlet coefficients*. 2011 7th Iranian Conference on Machine Vision and Image Processing; 2011: IEEE.
- [35] Dinghui Z, Haixia D, Chao Z, editors. *Researches on Digital Image Watermarking*. 2007 8th International Conference on Electronic Measurement and Instruments; 2007: IEEE.
- [36] Chalamala SR, Kakkirala KR, Mallikarjuna RGB, editors. *Analysis of wavelet and contourlet transform based image watermarking techniques*. 2014 IEEE International Advance Computing Conference [IACC]; 2014: IEEE.
- [37] Easley G, Labate D, Lim W-Q, Analysis CH. *Sparse directional image representations using the discrete shearlet transform*. 2008;25[1]:25-46.
- [9] Xu W, Chen R, Huang B, Zhang X, Liu CJS. *Single Image Super-Resolution Based on Global Dense Feature Fusion Convolutional Network*. 2019;19[2]:313.
- [10] Patterson J, Gibson A. *Deep learning: A practitioner's approach*: O'Reilly Media, Inc.; 2017.
- [11] Kumar N, Verma R, Sethi A. *Convolutional neural networks for wavelet domain super resolution*. 2017;90:65-71.
- [12] Dong C, Loy CC, He K, Tang X, editors. *Learning a deep convolutional network for image super-resolution*. European conference on computer vision; 2014: Springer.
- [13] Kim J, Kwon Lee J, Mu Lee K, editors. *Accurate image super-resolution using very deep convolutional networks*. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
- [14] Kim J, Kwon Lee J, Mu Lee K, editors. *Deeply-recursive convolutional network for image super-resolution*. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
- [15] Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, et al., editors. *Photo-realistic single image super-resolution using a generative adversarial network*. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.
- [16] He K, Zhang X, Ren S, Sun J, editors. *Deep residual learning for image recognition*. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
- [17] Freeman WT, Jones TR, Pasztor EC. *Applications. Example-based super-resolution*. 2002[2]:56-65.
- [18] Glasner D, Bagon S, Irani M, editors. *Super-resolution from a single image*. 2009 IEEE 12th International Conference on Computer Vision; 2009: IEEE.
- [19] Zontak M, Irani M, editors. *Internal statistics of a single natural image*. CVPR 2011; 2011: IEEE.
- [20] Timofte R, De Smet V, Van Gool L, editors. *A+: Adjusted anchored neighborhood regression for fast super-resolution*. Asian conference on computer vision; 2014: Springer.
- [21] Yang M, Zhang L, Yang J, Zhang D, editors. *Metaface learning for sparse representation based face recognition*. 2010 IEEE International Conference on Image Processing; 2010: IEEE.
- [22] Zeyde R, Elad M, Protter M, editors. *Low-complexity SingleImage super-resolution based on nonnegative neighbor embedding*. International Conference on Curves and Surfaces; 2012.
- [23] Wang D, Lu H, Yang M-H. *Online object tracking with sparse prototypes*. 2012;22[1]:314-25.



بهبود الگوریتم ترمیم تصویر بر اساس اطلاعات ساختار و بافت با استفاده از مفهوم Quad-Tree

کیمیا پیوندی^۱، فرزین یغمایی^۲

^۱ دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، kpeyvandi@semnan.ac.ir

^۲ دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، f_yaghmaee@semnan.ac.ir

چکیده

در این مقاله، یک روش جدید برای ترمیم تصویر با استفاده از ساختار و بافت و بر اساس Quad-Tree آورده شده است. به طور کلی یک تصویر قابل تجزیه به دو تصویر ساختار و بافت می‌باشد که هر کدام از این اجزاء می‌توانند ویژگی‌های خاصی از تصویر را نشان دهند. در نتیجه در این مقاله ابتدا تجزیه تصویر انجام شده سپس هر یک از اجزاء به طور جداگانه ترمیم می‌شوند. در بخش ترمیم از مفهوم Quad-Tree استفاده شده است که یک داده ساختار درختی می‌باشد و می‌تواند روی یک تصویر به کار گرفته شود. با اعمال Quad-Tree تصویر به نواحی متمایزی افزای می‌شود به طوری که پیکسل‌های هر ناحیه بر اساس یک حد آستانه‌ای به هم شباهت دارند و نسبت به ناحیه دیگر متفاوت هستند. اندازه هر ناحیه بدست آمده پس از اعمال Quad-Tree می‌تواند اهمیت آن ناحیه را نشان دهد و ما از این امر در ترمیم ساختار و بافت استفاده می‌نماییم. پس از ترمیم هر یک از اجزاء، تصویر نهایی از ترکیب دو جزء ترمیم شده بدست می‌آید. نتایج روش پیشنهادی با استفاده از معیارهای کمی چون PSNR و SSIM نشان می‌دهد که استفاده از ساختار و بافت و نیز به کارگیری Quad-Tree می‌تواند به کیفیت ترمیم تصویر کمک شایانی نماید.

کلمات کلیدی

ترمیم تصویر، PCA، ساختار، بافت، Quad-Tree

۱- مقدمه

مطرح گردید که با عنوان پوشاندن خطا نامیده شده است. پس از آن برای حذف خراش از تصویر، حذف یک شیء خاص در تصویر که از نظر کاربر نامطلوب است، حذف لوگو، حذف متن و ... مورد توجه قرار گرفت. در یک مسأله ترمیم، تصویر داده شده با نماد I ، ناحیه هدف (حفره) با نماد Ω ، ناحیه منبع با نماد Φ و مرز ناحیه هدف با نماد $\delta\Omega$ نمایش داده می‌شود که در جدول ۱ نمادهای مورد استفاده در یک مسأله ترمیم تصویر قابل مشاهده می‌باشد.

روش‌های مختلفی برای ترمیم تصویر مطرح شده است که به طور کلی در چهار دسته قابل طبقه‌بندی می‌باشد: روش‌های مبتنی بر انتشار^۱، روش‌های مبتنی بر نمونه^۲، روش‌های کمینه کردن مرتبه^۳ و روش‌های Sparsity به طوری که در هر دسته روش‌های مختلفی ارائه شده است و در بخش ۲ به این روش‌ها اشاره شده است.

ترمیم تصویر یکی از حوزه‌های پردازش تصویر می‌باشد که در چند سال اخیر مورد توجه محققان قرار گرفته است. در ترمیم تصویر هدف حذف یک یا چند ناحیه از تصویر است که به کمک یک ماسک مشخص شده است. این نواحی می‌توانند لوگو، متن، خراش^۱ یا هر شیء ای از تصویر باشد که از نظر کاربر نامطلوب است. پس از حذف ناحیه مورد نظر، لازم است این ناحیه از تصویر بر اساس سایر نواحی پر شود به طوری که تصویر نهایی از نظر سیستم بینایی انسان مطلوب به نظر بیاید.

به طور کلی در یک مسأله ترمیم، لازم است کاربر نواحی که نیاز به ترمیم دارند را مشخص نماید که این نواحی به عنوان ماسک^۲ نامیده می‌شود. سپس لازم است این بخش بر اساس نواحی سالم تصویر که ناحیه منبع^۳ به حساب می‌آیند، پر گردد که به این فرآیند بازسازی تصویر یا ترمیم تصویر گفته می‌شود. بحث بازسازی تصویر ابتدا در برای بازسازی بلوک‌های از دست رفته در انتقال اطلاعات از راه دور

جدول ۱- نمادهای مورد استفاده در مسأله ترمیم تصویر

نماد	تعریف
I	تصویر اصلی
Ω	ناحیه هدف

Φ	ناحیه منبع
Ψ_p	یک وصله مرزی به مرکز P
φ	ناحیه منبع از وصله منتخب

۲- کارهای دیگران

روش‌های مبتنی بر انتشار یکی از قدیمی‌ترین روش‌های ترمیم تصویر محسوب می‌شود که نخستین بار توسط Bertalimo در سال ۲۰۰۰ مطرح شده است [۱]. پس از آن روش‌های مختلفی برای ترمیم تصویر در این حوزه مطرح گردید که به عنوان نمونه استفاده از مدل CCD^A است که بر اساس الگوریتم کاهش تغییرات کل^۹ تصویر به منظور ترمیم آن بیان شده است [۲].

روش‌های مبتنی بر نمونه یکی از متداول‌ترین روش‌های ترمیم تصویر محسوب می‌شوند [۳، ۴] و نخستین بار توسط Criminisi در سال ۲۰۰۴ مطرح شده‌اند [۵]. در این روش برای وصله‌ها در ناحیه هدف، اولویت تعیین شده است و این وصله‌ها بر اساس اولویت تعیین شده پر می‌شوند. در روش دیگری مفهوم اولویت ساختاری و سازگاری بافتی برای یک وصله مطرح شده است [۶]. یک روش ترمیم تصویر و ویدیو مبتنی بر وصله مطرح شده است که برای یافتن وصله‌های مشابه، ابتدا تصویر به فضای ویژگی با ابعاد کمتر تبدیل می‌شود سپس در فضای جدید جستجو برای یافتن وصله مشابه انجام می‌شود [۷].

روش‌های مبتنی بر کمینه کردن مرتبه در چند سال اخیر مورد توجه قرار گرفته است، به طوری که در این روش‌ها تصویر به صورت یک ماتریس یا یک تنسور در نظر گرفته می‌شود [۸، ۱۶]. سپس لازم است مقادیر پیکسل‌های ناحیه هدف به گونه‌ای تخمین زده شود که مرتبه ماتریس یا تنسور کمینه گردد. ذکر این نکته ضروری است که مسأله کمینه کردن مرتبه ماتریس، به عنوان مسأله NP شناخته می‌شود و به همین دلیل گفته می‌شود استفاده از نرم اتمی، بهترین جایگزین به منظور کمینه کردن مرتبه ماتریس یا تنسور می‌باشد [۹]. به طور کلی نرم اتمی ماتریس یا تنسور بر اساس تجزیه SVD بدست می‌آید اما اخیراً روش‌های مطرح شده است که تعاریف دیگری برای نرم اتمی ارائه داده‌اند [۱۰].

در اکثر روش‌های ارائه شده در حوزه ترمیم که از مفهوم کمینه کردن مرتبه ماتریس یا تنسور استفاده می‌کنند، ناحیه هدف کوچک بوده یا پیکسل‌های ناحیه هدف در کل تصویر پخش شده‌اند. اما اخیراً الگوریتم‌هایی هم مطرح شده‌اند که از ترکیب روش‌های مبتنی بر کمینه کردن مرتبه و مبتنی بر نمونه استفاده می‌کنند به طوری که می‌توانند ترمیم را در یک مسأله با ناحیه هدف بزرگتر انجام دهند [۱۱].

در این روش برای یافتن یک پیکسل هدف (که بخشی از پیکسل‌های آن معلوم و برخی مجهول هستند) تعداد n وصله مشابه یافته می‌شود و پیکسل‌های هر وصله به صورت یک آرایه ستونی قرار می‌گیرند [۱۱]. از کنار هم قرار دادن این آرایه‌های ستونی یک ماتریس بدست می‌آید که مرتبه پایینی دارد، به طوری که برای یافتن مقادیر مجهول این ماتریس، لازم است از روش‌های کمینه کردن مرتبه ماتریس استفاده شود. همچنین روش دیگری هم ابتدا ترمیم را بر اساس وصله انجام می‌دهد سپس سعی می‌نماید مرتبه ماتریس ترمیم شده را کمینه نماید

[۱۵]. روش دیگری برای اولویت وصله‌ها بر اساس Structure sparsity مطرح شده است که در دسته روش‌های مبتنی بر Sparsity قرار می‌گیرد. به طوری که ترکیب استفاده از وصله و نیز مفهوم Sparsity می‌تواند نتایج مطلوبی برای ترمیم تصویر ایجاد کند [۱۲، ۱۳].

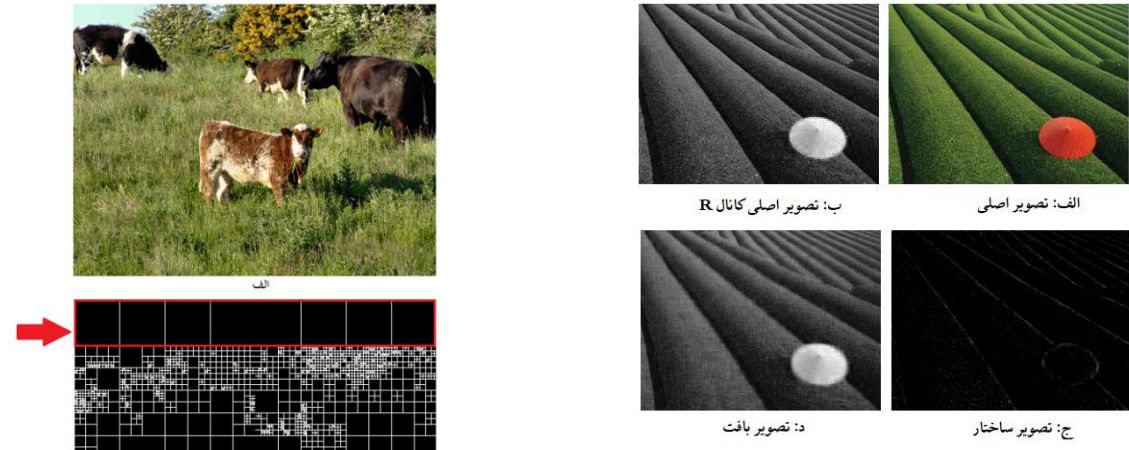
۳- روش پیشنهادی

به طور کلی یک تصویر قابل تجزیه به دو بخش ساختار و بافت می‌باشد که هر یک از این اجزاء می‌تواند ویژگی‌های متمایزی از تصویر را نشان دهد. اکثر روش‌های ترمیم تصویر که تاکنون مطرح شده‌اند یا مناسب برای ترمیم ساختار هستند و یا در ترمیم بافت کارایی دارند، در نتیجه به کارگیری آنها در ترمیم بسیاری از تصاویر سبب ایجاد اثرات مصنوعی می‌گردد. در این مقاله ابتدا روش پیشنهادی ما که جداسازی ساختار و بافت بر اساس روش PCA است توضیح داده می‌شود سپس هر یک از اجزای ساختار و بافت به صورت جداگانه بر اساس روش ما ترمیم شده و تصویر نتیجه بدست می‌آید. در راستای روش پیشنهادی و به منظور ترمیم اجزای تصویر از مفهوم Quad-tree استفاده می‌کنیم که این مفهوم در بخش ۳-۲ آورده شده است. همچنین نمادهای مورد استفاده در این مقاله در جدول ۱ قابل مشاهده است.

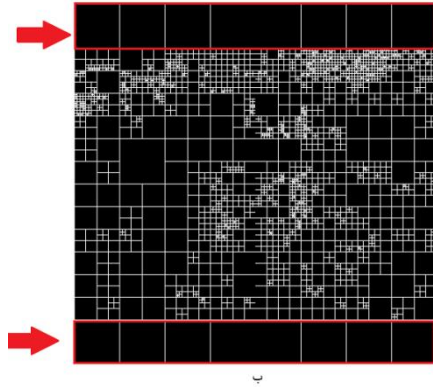
۳-۱- تجزیه تصویر به دو مولفه ساختار و بافت بر اساس PCA

به طور کلی روش PCA در تحلیل داده‌ها به کار می‌رود به طوری که از کاهش ابعاد داده‌ها برای شناسایی و یافتن الگو استفاده می‌نماید. یکی از کاربردهای PCA در حوزه پردازش تصویر می‌باشد، به طوری که با اعمال PCA بر روی یک تصویر می‌توان آن را به دو تصویر تجزیه نمود که یکی از این اجزاء دارای مرتبه کمینه بوده و جزء دیگر Sparse می‌باشد که در شکل ۱ نمونه‌ای از تجزیه تصویر بر اساس تکنیک PCA قابل مشاهده است. در واقع روش PCA تصویر را به دو بخش تصویر با مرتبه پایین^{۱۰} و تصویر Sparse تجزیه می‌کند که در روش پیشنهادی ما این اجزاء به عنوان بافت (تصویر مرتبه پایین) و ساختار (تصویر Sparse) در نظر گرفته شده‌اند.

به طور کلی اکثر روش‌های تجزیه تصویر به ساختار و بافت، اگرچه این دو جزء را تجزیه می‌کند اما تصویر اصلی نمی‌تواند از اجزاء بدست آید. در واقع هدف اکثر روش‌های تجزیه تصویر، استخراج ویژگی‌های مختلف می‌باشد در حالی که با استفاده از روش PCA جمع دو جزء بدست آمده می‌تواند تصویر اصلی را به ما بدهد. به طور کلی در تجزیه تصویر داریم: $I = U \oplus V$. به طوری که I تصویر اصلی، V جزء بافت و U ساختار به حساب می‌آید و عملگر \oplus در روش‌های مختلف تجزیه، متفاوت است که در شیوه PCA این عملگر + می‌باشد (جمع دو تصویر).



شکل ۱- تجزیه تصویر بر اساس PCA



شکل ۲- اعمال Quad-Tree روی یک نمونه تصویر

۲-۲- مفهوم Quad-Tree

مفهوم Quad-Tree یک داده ساختار درختی می‌باشد که در آن هر گره حداکثر می‌تواند چهار فرزند داشته باشد. به طور کلی Quad-Tree یک ساختار داده است که برای افراز یک ناحیه دوبعدی به صورت بازگشتی مورد استفاده قرار داده می‌شود. وقتی از یک Quad-Tree برای نمایش یک تصویر استفاده می‌شود، می‌توان تشخیص داد که کدام بخش‌های تصویر اهمیت بیشتری دارد و کدام نواحی کم‌اهمیت‌تر هستند. به طوری که هر چه مربع‌های یافته شده کوچکتر باشد یعنی این ناحیه اهمیت بیشتری دارد و برعکس. در شکل ۲ به عنوان نمونه شکل Quad-Tree از یک تصویر آورده شده است. به طوری که الف تصویر اصلی و ب تصویر Quad-Tree آن می‌باشد.

ذکر این نکته ضروری است که Quad-Tree فقط می‌تواند برای تصاویر مربعی به کار گرفته شود، در این حالت اگر اندازه تصویر $n \times n$ باشد اندازه تصویر حاصل از Quad-Tree هم $n \times n$ خواهد شد. اما همان‌طور که در شکل ۲ دیده می‌شود اندازه تصویر اصلی (الف) و تصویر حاصل از Quad-Tree با هم برابر نیست. دلیل این امر این است که تصویر اصلی غیرمربعی است و ما می‌خواهیم Quad-Tree را برای تصاویر غیرمربعی نیز به کار بگیریم. برای این منظور از Padding پیکسل‌های سیاه به اطراف تصویر استفاده نموده‌ایم تا بتوانیم یک تصویر غیرمربعی را به مربعی تبدیل کنیم. در نتیجه همان‌طور که در شکل ۲-ب دیده می‌شود مستطیل بالا و پایین به دلیل Padding ایجاد شده است.

در استفاده از Quad-Tree برای یک تصویر لازم است یک حد آستانه هم ذکر شود که این حد آستانه عددی بین صفر و یک می‌باشد. این حد آستانه نشان می‌دهد لازم است پیکسل‌های یک ناحیه تا چه حد به همدیگر مشابه باشند. حد پایین استفاده از Quad-Tree به این صورت است که هر پیکسل یک ناحیه می‌باشد و حد بالای آن این است که کل تصویر یک ناحیه است که به صورت تجربی عدد $0,6$ به عنوان حد آستانه در نظر گرفته شده است.

۳-۳- ترمیم ساختار بر اساس Quad-Tree

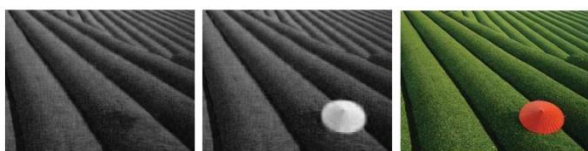
در یک مسأله ترمیم تصویر، تصویر اصلی به همراه ماسک تصویر داده شده است و قرار است ما ناحیه هدف را بر اساس ناحیه منبع پر کنیم به طوری که نتیجه از نظر سیستم بینایی انسان طبیعی به نظر برسد. برای این منظور ابتدا Quad-Tree را روی تصویر اصلی به همراه ماسک اعمال می‌کنیم. در این صورت اندازه مربع‌های ایجاد شده در اطراف تصویر به ما نشان می‌دهد کدام بخش ناحیه هدف مهم‌تر است و اولویت بالاتری برای پر کردن دارد. در نتیجه جهت پر کردن ناحیه هدف در تصویر ساختار بر اساس ساختار درختی ایجاد شده توسط Quad-Tree بدست می‌آید. به عبارت دیگر در روش پیشنهادی ما ترمیم جزء ساختار در جهتی انجام می‌شود که این جهت بر اساس ساختار درختی بدست آمده از Quad-Tree می‌باشد.

هدف ما در ترمیم ساختار تصویر، ارائه یک روش جدید است که از Quad-tree استفاده می‌نماید. برای این منظور فرض کنید Ψ_p یک وصله مرزی به مرکز p می‌باشد به طوری که ϕ ناحیه منبع از یک وصله محسوب می‌شود. هدف ما این است که جهت انتشار درست برای یک وصله را بیابیم. برای این منظور از نقطه p شروع می‌کنیم و پیکسل‌هایی که به عنوان همسایه ۸ گانه پیکسل مورد نظر است را بررسی می‌کنیم. هر پیکسل به یکی از نواحی که توسط Quad-Tree افراز شده است تعلق دارد. پیکسلی را انتخاب می‌کنیم که در همسایگی ۸ گانه پیکسل p بوده و متعلق به کوچکترین ناحیه باشد. زیرا هر چه یک پیکسل متعلق به ناحیه کوچکتر باشد نشان می‌دهد که اطلاعات بیشتری در آن است در حالی که وقتی یک پیکسل در ناحیه بزرگی قرار گرفته است احتمالاً متعلق به یک ناحیه هموار^{۱۱} می‌باشد. این روند را در کل وصله مورد نظر ادامه می‌دهیم یعنی پیکسل بعدی را می‌یابیم که در همسایگی ۸ گانه قرار داشته باشد و در کوچکترین ناحیه نسبت به

و در مورد تصویر ساختار از یافتن جهت درست ترمیم استفاده می‌نماید در نتیجه بیشترین کارایی را خواهد داشت.

۳-۴- ترمیم بافت

مرحله دیگری از ترمیم تصویر به روش پیشنهادی، ترمیم بافت می‌باشد. در ترمیم بافت از یافتن وصله‌های مشابه با وصله هدف استفاده می‌کنیم با این تفاوت که وقتی چند وصله کاندید با وصله هدف یافته می‌شود، وصله‌ای را انتخاب می‌کنیم که از نظر شکل بدست آمده با Quad-Tree به وصله هدف شبیه‌تر باشد. این امر سبب می‌شود ترمیم بافت نسبت به روش‌های موجود بهتر انجام گیرد. ترمیم تصویر بافت که در بخش ۳-۱ در شکل ۱-د آورده شده است با روش پیشنهادی انجام شده و در شکل ۵ قابل مشاهده است. شکل ۵ مربوط به ترمیم تصویر بافت کانال R از تصویر اصلی (تصویر اصلی: شکل ۱-الف) می‌باشد و تصویر ساختار کانال G و B نیز به همین شکل انجام می‌شود.



الف: تصویر اصلی ب: تصویر بافت از کانال R ج: تصویر بافت ترمیم شده

شکل ۵- یک نمونه از ترمیم بافت

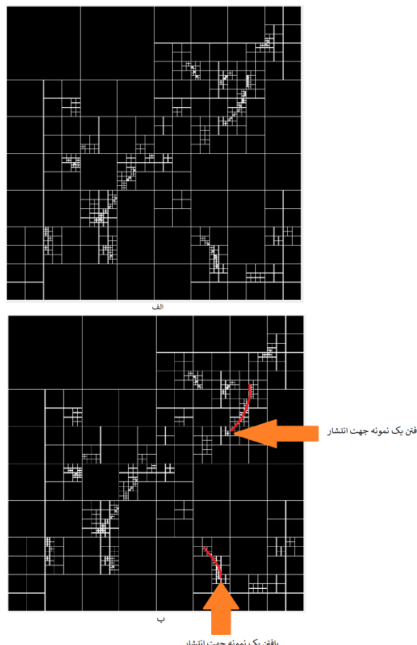
۳-۵- ترمیم تصویر اصلی

در این مرحله لازم است ترمیم تصویر اصلی بر اساس مجموع نتیجه حاصل از ترمیم اجزای آن بدست آید. یکی از مزایای استفاده از PCA به منظور تجزیه تصویر این است که تصویر نتیجه، حاصل جمع دو تصویر تجزیه شده می‌باشد. در نتیجه به راحتی می‌توان تصویر نتیجه را بر اساس جمع دو مولفه ساختار و بافت ترمیم شده بدست آوریم. لازم به ذکر است که تجزیه و ترمیم برای هر یک از کانال‌های رنگی R، G و B به صورت مجزا انجام می‌گردد. در بخش ۴ نتایج بدست آمده و ارزیابی آن آورده شده است.

۴- نتایج بدست آمده

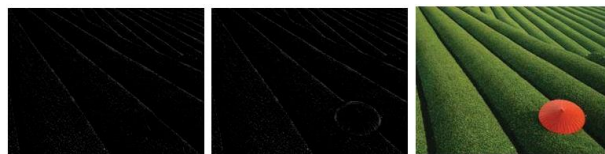
در روش پیشنهادی ابتدا هر یک از کانال‌های تصویر بر اساس روش PCA به دو جزء ساختار و بافت تجزیه می‌شود سپس بخش ساختار و بافت هر کدام با یک روش مجزا بر اساس Quad-Tree ترمیم می‌گردد که در شکل ۶ بخشی از نتایج روش پیشنهادی ما قابل مشاهده است. شکل ۶ (الف-۱، الف-۲، الف-۳) تصویر اصلی و شکل ۶ (ب-۱، ب-۲، ب-۳) تصویر بافتی است که ماسک تصویر رانشان می‌دهد. در شکل (ج-۱، ج-۲، ج-۳) خروجی ترمیم بر اساس روش ماتریس تقریب [۱۵] آورده شده است، همچنین شکل (د-۱، د-۲، د-۳) نتیجه بر اساس مرجع [۱۶] می‌باشد. شکل (ه-۱، ه-۲، ه-۳) خروجی ترمیم بر اساس الگوریتم Patch Sparsity [۱۳] را نشان می‌دهد و در نهایت نتایج روش پیشنهادی ما در (ی-۱، ی-۲، ی-۳) آورده شده است.

سایر پیکسل‌های ۸ گانه باشد و از این طریق یک جهت انتشار در وصله می‌یابیم. یک نمونه از یافتن جهت درست ترمیم ساختار بر اساس Quad-Tree که روش پیشنهادی ما مطرح می‌کند در شکل ۳ قابل مشاهده است.



شکل ۳- نمونه‌ای از یافتن مسیر انتشار درست بر اساس Quad-Tree

پس از تعیین جهت انتشار لازم است از پیکسل‌هایی برای ترمیم استفاده کنیم که در جهت انتشار هستند. این امر کمک می‌کند ترمیم ساختار به خوبی انجام شود. ترمیم تصویر ساختار که در بخش ۳-۱ در شکل ۱-ج آورده شده است با روش پیشنهادی انجام شده و در شکل ۴ قابل مشاهده است. شکل ۴ مربوط به ترمیم تصویر ساختار کانال R از تصویر اصلی (تصویر اصلی: شکل ۱-الف) می‌باشد و تصویر ساختار کانال G و B نیز به همین شکل انجام می‌شود.



الف: تصویر اصلی ب: تصویر ساختار از کانال R ج: ترمیم شده تصویر ساختار

شکل ۴- یک نمونه از ترمیم ساختار

در واقع یافتن جهت درست برای ترمیم، یک مسأله مهم در این حوزه محسوب می‌شود به طوری که در [۱۷] از هیستوگرام گرادیان برای یافتن جهت درست استفاده شده است. در حالی که نوآوری روش پیشنهادی ما استفاده از Quad-Tree برای یافتن جهت درست ترمیم می‌باشد. ذکر این نکته ضروری است که یافتن جهت درست در فرآیند ترمیم اگرچه می‌تواند کمک شایانی نماید اما در مورد تصاویر غیر بافتی قابلیت کاربرد دارد [۱۷]، پس در مورد هر تصویری کارا نمی‌باشد. این در حالی است که روش پیشنهادی ما جداسازی را انجام داده

پیشنهادی ما می‌تواند در ترمیم انواع مختلف تصاویر کارایی داشته باشد. این روش از مفهوم PCA در جداسازی تصویر استفاده می‌نماید که جزء مرتبه پایین با عنوان بافت و جزء Sparse با عنوان ساختار در نظر گرفته می‌شود. نوآوری دیگر روش پیشنهادی، استفاده از Quad-Tree برای یافتن جهت درست ترمیم در تصویر ساختار می‌باشد. همچنین در تصویر بافت برای یافتن وصله‌های مشابه از Quad-Tree استفاده شده است. نتایج روش پیشنهادی چه از نظر کمی و چه از نظر کیفی قدرت روش ما را نشان می‌دهد به طوری که به منظور ارزیابی کمی در تصاویری که مرجع موجود باشد معیارهای SSIM و PSNR به کار گرفته شده است.

همان‌طور که در شکل ۶ مشاهده می‌شود روش پیشنهادی نسبت به سایر روش‌ها عملکرد بهتری داشته و کیفیت تصویر پس از ترمیم بهتر است. برای ارزیابی کیفیت تصاویر از معیار PSNR و SSIM استفاده نموده‌ایم که در جدول ۲ مقایسه بین روش‌های مختلف برای ترمیم تصویر ۶-الف-۲ قابل مشاهده است. لازم به ذکر است که استفاده از معیارهای PSNR و SSIM زمانی امکان‌پذیر است که تصویر مرجع موجود باشد در نتیجه ارزیابی کمی نتیجه ترمیم تصویر ۶-الف-۱ و ۶-الف-۳ به دلیل نبودن تصویر مرجع امکان‌پذیر نمی‌باشد.

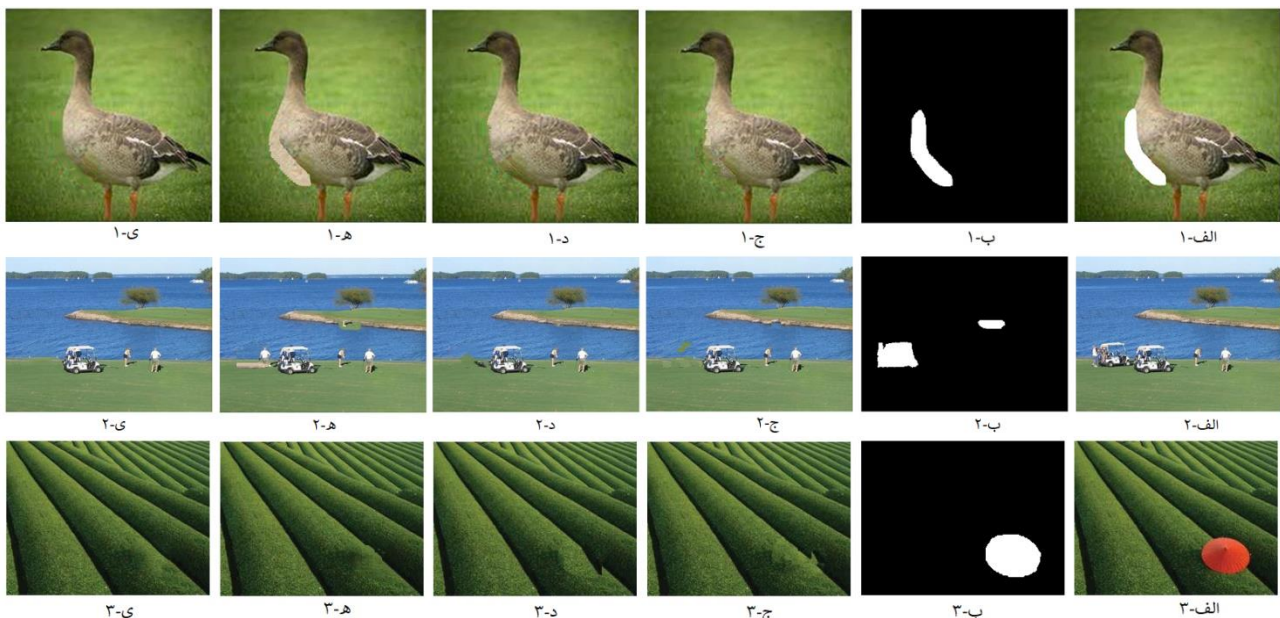
جدول ۲- مقایسه روش‌های مختلف ترمیم با معیارهای کمی مربوط به ترمیم تصویر در شکل ۶-الف-۲

روش ترمیم استفاده شده	معیار PSNR	معیار SSIM
روش [۱۵]، (شکل ۶-ج-۲)	۲۷,۰۱۴	۰,۹۶۰
روش [۱۶]، (شکل ۶-د-۲)	۲۷,۴۸۶	۰,۹۶۴
روش [۱۳]، (شکل ۶-ه-۲)	۲۵,۴۳۵	۰,۹۵۸
روش پیشنهادی، (شکل ۶-ی-۲)	۲۸,۰۲۱	۰,۹۷۱

۵- نتیجه گیری

در این مقاله یک روش جدید برای ترمیم تصویر پیشنهاد شده است به طوری که از اطلاعات ساختار و بافت برای ترمیم استفاده می‌کند. همچنین مفهوم Quad-Tree در ترمیم اجزاء تصویر به کار گرفته می‌شود. به عبارت دیگر بر اساس ساختار درختی ایجاد شده توسط Quad-Tree به خوبی می‌توان ترمیم را انجام دهیم به طوری که روش

- 1 Scratch
- 2 Mask
- 3 Source region
- 4 Error concealment
- 5 Diffusion based inpainting
- 6 Patch based inpainting
- 7 Low rank inpainting
- 8 Curvature Driven Diffusion
- 9 Total Variation (TV)
- 10 Low rank
- 11 Smooth



شکل ۶- مقایسه روش پیشنهادی با روش‌های مختلف ترمیم، (الف-۱، الف-۲، الف-۳) تصویر اصلی، (ب-۱، ب-۲، ب-۳): تصویر باینری که ماسک تصویر را نشان می‌دهد، (ج-۱، ج-۲، ج-۳): ترمیم به روش ماتریس تقریب [۱۵]، (د-۱، د-۲، د-۳): ترمیم به روش HOSVD [۱۶]، (ه-۱، ه-۲، ه-۳): ترمیم به روش Sparsity [۱۳]، (ی-۱، ی-۲، ی-۳): ترمیم به روش پیشنهادی ما

مراجع

- [1] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, "Image inpainting," in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 417-424, 2000.
- [2] T. F. Chan, J. Shen, "Nontexture Inpainting by Curvature-Driven Diffusions," Journal of Visual Communication and Image Representation, Vol. 12, Issue. 4, PP. 436-449, December 2001.
- [3] T. Ruzic, A. Pizurica, "Context-aware patch-based image inpainting using Markov random field modeling," IEEE Trans. Image Processing, Vol. 24, No. 1, PP. 444-456, 2015
- [4] H. Wang, L. Jiang, R. Liang, X. X. Li, "exemplar based image inpainting using structure consistent patch matching," Neuro computing, vol. 269, pp. 90-96, 2017.
- [5] A. Criminisi, P. Pérez, K. Toyama, "Region filling and object removal by exemplar-based image inpainting," IEEE Trans. Image Processing, Vol. 13, PP. 1200-1212, 2004.
- [6] Z. Chen et al, "Structure aware image inpainting using patch scale optimization," Journal of Visual Communication and Image Representation, vol. 40. pp. 312-323, 2016.
- [7] M. Isogawa, D. Mikami, K. Takahashi, A. Kojima, "Image and Video completion via feature reduction and compensation," Multimedia Tools and Applications, vol. 76, pp. 9443-9462, 2017.
- [8] J. Liu, P. Musialski, P. Wonka, J. Ye, "Tensor completion for estimating missing values in visual data," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 208 - 220, 2013.
- [9] M. Fazel, "Matrix rank minimization with applications," PhD thesis, Stanford University, 2002.
- [10] J. A. Bengua, H. N. Phien, H. D. Tuan, M. N. Do, "Efficient Tensor Completion for Color Image and Video Recovery: Low-Rank Tensor Train," IEEE Transactions on Image Processing, vol. 26, pp. 2466 - 2479, 2017.
- [11] Q. Guo, S. Gao, X. Zhang, Y. Yin, C. Zhang, "Patch-Based Image Inpainting via Two-Stage Low Rank Approximation", IEEE Transactions on Visualization and Computer Graphics, 2017.
- [12] D. Jin, X. Bai, "Patch-Sparsity-Based Image Inpainting Through a Facet Deduced Directional Derivative," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, pp. 1310-1324, 2019.
- [13] Image Inpainting by Patch Propagation Using Patch Sparsity," IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 5, MAY 2010.
- [14] V. Kumar, J. Mukherjee, S. K. Das Mandal, "Image Inpainting through Metric Labelling Via Guided Patch Mixing," IEEE Trans. on Image Processing, Vol. 25, No. 11, PP. 5212-5226, 2016.
- [15] V. K. Alilou, F. Yaghmaee, "Exemplar-based image inpainting using svd-based approximation matrix and multi-scale analysis," Multimedia Tools and Application, vol. 76, Issue. 5, pp. 7213-7234, 2017.
- [16] M. Ghorai, S. Mandal, B. Chanda, "A two-step image inpainting algorithm using tensor svd," In: Computer Vision-ACCV 2014 Workshops, pp. 63-77, 2014.
- [17] V. K. Alilou, F. Yaghmaee, "Non-texture image inpainting using histogram of oriented gradients," Journal of Visual Communication and Image Representation, vol. 48, pp. 43-53, 2017.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



۲۷ الی ۲۸ آذرماه ۱۳۹۸

یک روش سریع برای انتخاب بهترین گراف‌ها جهت ترکیب به منظور افزایش دقت انتشار برچسب

^۱ امیرنامجوی، علیرضا بساقزاده^۲

^۱ دانشجوی کارشناسی ارشد کامپیوتر- هوش مصنوعی و رباتیک، دانشگاه تربیت دبیر شهید رجایی، تهران
namjoy@sru.ac.ir

^۲ استادیار، گروه هوش مصنوعی و رباتیک، دانشکده مهندسی کامپیوتر دانشگاه تربیت دبیر شهید رجایی، تهران
a.bosaghzadeh@sru.ac.ir

چکیده

بسیاری از اوقات بررسی یک پدیده تنها از یک منظر (ویژگی) به تنهایی نمی‌تواند به تشخیصی کامل و دقیق از آن منجر شود. همچنین اینکه کدام ویژگی بهترین توصیف از پدیده مورد بررسی را ارائه می‌دهد از پیش برای ما مشخص نیست. لذا انتخاب بهترین ویژگی‌ها و ترکیب آنها می‌تواند منجر به نتایج مطلوب شود. می‌توان برای توصیف روابط بین نمونه داده‌ها از منظر هر کدام از ویژگی‌ها یک گراف ساخت. گرافی که در آن وزن بین نمونه‌های متعلق به یک کلاس زیاد و بین کلاس‌های مختلف کم باشد مطلوب‌تر است و دقت انتشار برچسب روی آن بالا خواهد بود، لذا بر این اساس و بر مبنای داده‌های برچسب‌دار، می‌توان بهترین گراف‌ها را انتخاب و سپس ترکیب کرد. در این مقاله یک روش رتبه‌بندی برای انتخاب گراف مناسب بر مبنای جمع مقادیر وزن‌های بین نمونه‌های هر کلاس و یک روش برای بهبود نتایج ترکیب و انتشار برچسب پیشنهاد شده است. نتایج بدست آمده نشان می‌دهد که انجام دو مرحله انتخاب و ترکیب در کنار یکدیگر می‌تواند منجر به نتیجه‌ای مشابه دقت بهترین گراف و یا حتی بالاتر از آن شود.

کلمات کلیدی

انتخاب بهترین گراف، ترکیب اطلاعات، ماتریس مجاورت، ویژگی، انتشار برچسب

از کاربردهای عملی استفاده شده است. موفق‌ترین رویکرد در یادگیری نیمه نظارتی، رویکردهای مبتنی بر گراف هستند [11]. هر گراف را با سه مولفه به صورت $G = (V, E, W)$ نمایش می‌دهند. که در آن V یک مجموعه از رأس‌ها بر اساس مجموعه داده $X = \{x_i, i = 1, \dots, n\}$ و E یک مجموعه از یال‌ها با ابعاد $n * n$ و W یک تابع وزن متقارن غیرمنفی برای نمایش وزن یال‌ها می‌باشد که می‌تواند نشان دهنده شباهت بین نمونه‌ها باشد. در صورتی که وزن یک یال بزرگتر از صفر باشد، نشان دهنده این است که یالی بین دو رأس وجود دارد، در غیر این صورت بین دو رأس یالی وجود نخواهد داشت [12].

در یادگیری نیمه نظارتی بر پایه گراف، ایده کلیدی فرض خمینه^۲ می‌باشد که در آن نمونه‌های متصل شده با وزن بزرگ در گراف، برچسب‌های مشابهی دارند. یکی از راه‌های بیان کردن ساختار خمینه در فضای با بعد بالا، استفاده از گراف‌های همسایگی است [11].

۱- مقدمه

فناورهای اخیر جمع‌آوری گسترده انواع مختلف داده‌ها را مقرون به صرفه کرده است. برای ایجاد یک دید جامع از داده‌های مختلف (انواع مختلف و از منابع مختلف) در مورد یک موضوع خاص (مثلاً، بیماری، یک فرآیند بیولوژیکی و یا تشخیص چهره) لازم است این داده‌های مختلف با یکدیگر ترکیب شوند. می‌توان با ساخت گراف‌هایی بصورت شبکه شباهت برای هر کدام از انواع داده در دسترس (موجود) و سپس ترکیب (همجوشی^۱) آن‌ها بطور موثر در یک شبکه که طیف کامل داده‌های مهم را بازنمایی کند، این کار را انجام داد [1].

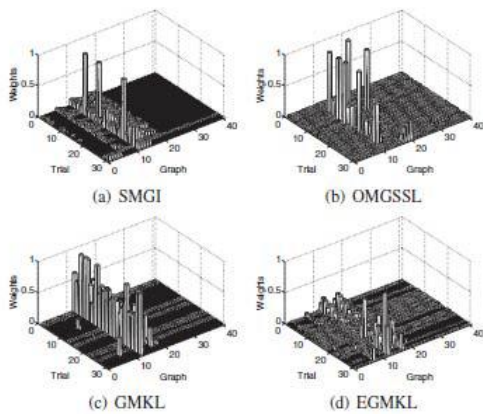
در سال‌های اخیر به روش‌های یادگیری نیمه نظارتی بر پایه گراف [10]-[2]، در یادگیری ماشین توجه زیادی شده است و به خاطر انعطاف‌پذیری و سادگی پیاده‌سازی، به طور گسترده‌ای در خیلی

ارتباطات قوی و ارتباطات ضعیفی که در تعداد زیادی از گراف‌ها وجود دارند، تقویت شده و سایر ارتباطات تضعیف می‌شوند.

موسایوکی کاراسایوما و همکاران [11] یک روش جدید برای ترکیب گراف‌ها با قید تنکی (در ضرایب ترکیب گراف‌ها) به نام SMGI^۴ پیشنهاد داده‌اند که بر اساس ادعای آنها، این تنکی در ضرایب با هیچ روش موجود دیگری قابل دستیابی نیست. ایده اصلی بر این اساس است که در میان گراف‌هایی که بر اساس ویژگی‌ها و یا پارامترهای مختلف ساخته شده‌اند ممکن است برخی گراف‌ها حاوی اطلاعات بی-ربط یا نویزی باشند و همچنین اهمیت برخی گراف‌ها نسبت به سایرین بیشتر باشد که این موارد با مشخص کردن ضرایب گراف‌ها مشخص می‌شود. این ویژگی منحصر بفرد دو مزیت را ایجاد می‌کند:

- ۱- بهبود عملکرد پیش بینی با حذف گراف‌های بی ربط و نویزی
- ۲- تفسیرپذیری نتایج: یعنی به سادگی می‌توان گراف‌های حاوی اطلاعات مفید را شناسایی نمود.

نتایج آزمایشات تجربی نشان داده است که مزیت روش پیشنهادی نه تنها در عملکرد پیش‌بینی بلکه همچنین در توانایی انتخاب گراف است. نتایج آزمایش و مقایسه انجام شده (شکل ۱) با چند روش مشابه به خوبی نشان می‌دهد که، ویژگی منحصر بفرد این روش، تنکی وزن گراف‌ها است. به این معنی که فقط یک بخش از وزن‌ها مقادیر غیر صفر دارند و باقی دقیقاً برابر صفر هستند.



شکل ۴۶: مقایسه ضرایب بدست آمده برای گراف‌ها [11]

عملکرد الگوریتم انتشار برچسب به شدت به روش ایجاد گراف ورودی وابسته است. به این معنا که باید معیار شباهت یا تابع فاصله که یال‌های گراف را مشخص می‌کند انتخاب شده و تابع فاصله نیز خود، پارامترهایی دارد که باید تنظیم شود و همچنین براساس ویژگی‌های مختلف می‌توان گراف‌های مختلفی ساخت. مهم ترین مسئله این است که در خیلی از کاربردهای عملی، ما تعدادی گراف مختلف از منابع داده داریم، ما به هر حال نمی‌توانیم متوجه شویم که کدام گراف برای پیش بینی، مهم تر است. لذا انتخاب گراف‌های مناسب که حاوی اطلاعات بیشتری هستند، اهمیت ویژه‌ای دارد. در این مقاله یک روش سریع برای رتبه‌بندی، جهت انتخاب بهترین گراف‌های ساخته شده با پارامترها و ویژگی‌های مختلف، که توصیف قویتری از ارتباطات بین داده‌های هرکلاس دارند و یک روش برای بهبود نتایج ترکیب و انتشار برچسب، با استفاده از داده‌های برچسب دار پیشنهاد شده است.

یکی از رویکردهای خوب پذیرفته شده در یادگیری نیمه نظارتی، انتشار برچسب است که برچسب هر گره در گراف داده شده را به گره‌های همسایه آن بر اساس شباهت آن‌ها منتشر می‌کند و بصورت مسئله حداقل مربعات فرموله می‌شود [5], [7], [13].

با توجه به اینکه هر کدام از روش‌های ساخت گراف پارامترهایی دارند که تعیین مقدار بهینه آنها می‌تواند مشکل باشد و همچنین براساس ویژگی‌های مختلف می‌توان گراف‌های مختلفی ایجاد کرد که هر کدام مزایا و معایبی دارند، لذا در بسیاری از موارد نیاز به استفاده از ترکیب گراف‌هایی که براساس ویژگی‌های مختلف و یا از یک ویژگی با پارامترهای مختلف ساخته شده‌اند می‌باشد.

ساده ترین راه ترکیب اطلاعات، الحاق کردن یا کنار هم گذاشتن ویژگی‌های نرمال شده می‌باشد که معمولاً منجر به نتایج خوبی نمی‌شود [1]. راه دیگر تحلیل مستقیم هر نوع داده قبل از ترکیب است [17]-[14] که چنین تحلیل‌های مستقلی منجر به نتایج ناسازگار می‌شود که یکپارچه سازی آن‌ها مشکل است [1]. رویکرد دیگر، انتخاب اولیه مجموعه‌های مشخصی از منظرهای مختلف اطلاعات و متریک‌های مشخص برای ترکیب می‌باشد [16]. این رویکرد این خطر را ایجاد می‌کند مجموعه‌های مناسبی که حاوی اطلاعات خوبی برای حل مسئله می‌باشند انتخاب نشود. همچنین می‌تواند منجر به بایاس شدن و از دست رفتن اطلاعات مکمل ارزشمند شود. برای حل این مشکلات روش‌هایی ارائه شده‌اند که قادر به یادگیری متریک مناسب و انتخاب گراف‌های حاوی اطلاعات و حذف گراف‌های بی ربط و نویزی می‌باشند [11].

مساله انتخاب ویژگی، یکی از مسائلی است که در محث یادگیری ماشین و همچنین شناسایی آماری الگو مطرح است [18]. این مساله در بسیاری از کاربردها (مانند طبقه بندی) اهمیت به سزایی دارد، زیرا در این کاربردها تعداد زیادی ویژگی وجود دارد، که بسیاری از آنها یا بلااستفاده هستند و یا اینکه بار اطلاعاتی چندانی ندارند. حذف نکردن این ویژگی‌ها حتی اگر مشکلی از لحاظ اطلاعاتی ایجاد نکند ولی باعث محاسباتی را برای کاربرد مورد نظر بالا می‌برد و علاوه بر این باعث میشود که اطلاعات غیر مفید زیادی را به همراه داده‌های مفید ذخیره کنیم. برای مساله انتخاب ویژگی، راه حل‌ها و الگوریتم‌های فراوانی ارائه شده است که بعضی از آنها قدمت سی یا چهل ساله دارند. مشکل بعضی از الگوریتم‌ها در زمان ارائه، بار محاسباتی زیاد آنها بود، اگرچه امروزه با ظهور کامپیوترهای سریع و منابع ذخیره سازی بزرگ این مشکل، به چشم نمی‌آید ولی از طرف دیگر، مجموعه داده‌های بسیار بزرگ برای مسائل جدید باعث شده است که همچنان پیدا کردن یک الگوریتم سریع برای این کار مهم باشد.

بو وانگ و همکاران [1] روشی تحت عنوان SNF^۳ ارائه نموده‌اند که برای ترکیب داده‌های مربوط به mRNA expression, DNA methylation و microRNA (miRNA) expression برای پنج دیاناست سرطان مورد استفاده قرار گرفته و نتایج آزمایشات آنها نشان داده است که ترکیب اطلاعات بطور قابل ملاحظه‌ای از آنالیز یک نوع داده منفرد بهتر عمل می‌کند. در روش SNF کار انتخاب ویژگی و ترکیب همزمان انجام می‌شود. به این صورت که در فرآیند ترکیب، پس از هر تکرار

۲- روش پیشنهادی

فرض کنید که تعداد n تصویر برای کلاسیابی داریم که L عدد از آنها برچسب‌دار و باقیمانده $u = n - L$ بدون برچسب هستند. برای هر تصویر v بردار ویژگی استخراج و بصورت $\{x_1^{(i)}, \dots, x_n^{(i)}\}_{i=1}^v$ نمایش داده می‌شود. ماتریس $Y \in \mathbb{R}^{n \times c}$ شامل برچسب n نمونه برای c کلاس است و هر تصویر متعلق به یکی از c کلاس می‌باشد [19]. به اینصورت که برای برچسب هر نمونه یک ماتریس $I^* C$ داریم که عدد متناظر کلاس آن 1 و باقی صفر می‌باشد. ترتیب قرارگیری داده‌ها به صورت $X = [x_1^T \dots x_u^T]^T$ می‌باشد. به این معنی که ابتدا داده‌های برچسب‌دار و سپس بدون برچسب قرار می‌گیرند، لذا ماتریس مجاورت نیز بدین صورت خواهد بود:

$$W = \begin{bmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{bmatrix}$$

در روش پیشنهادی ما از قسمت برچسب‌دار (W_{ll}) برای تخمین کیفیت گراف مجاورت استفاده می‌کنیم. ایده روش پیشنهادی این است مقدار مجموع وزن‌های درون کلاسی ($SWICW^l$) یعنی وزن بین نمونه‌های متعلق به یک کلاس برای گراف‌هایی که ارتباطات بین داده‌های هر کلاس با یکدیگر را بهتر بیان می‌کنند در مقایسه با گراف‌های ضعیف‌تر، بیشتر است. عبارت دیگر گرافی که در آن وزن بین نمونه‌های متعلق به یک کلاس زیاد و بین کلاس‌های مختلف کم باشد مطلوب‌تر است و دقت انتشار برچسب روی آن بالا خواهد بود، لذا بر این اساس و بر مبنای داده‌های برچسب‌دار، می‌توان بهترین گراف‌ها را انتخاب و سپس ترکیب کرد.

همان‌طور که بیان شد برای ارزیابی گراف مجاورت با این روش نیاز به داده‌های برچسب‌دار می‌باشد، لذا از قسمت برچسب‌دار گراف بعنوان یک تعمیم برای ارزیابی کل گراف استفاده می‌شود. الگوریتم روش پیشنهادی بدین صورت است: ابتدا براساس ویژگی‌های مختلف استخراج شده از تصاویر با استفاده از روش L1-R [20] گراف مجاورت با ویژگی‌ها و پارامترهای مختلف ساخته می‌شود. در مرحله بعد قسمت برچسب‌دار هرکدام از گراف‌ها جدا شده و مجموع وزن بین نمونه‌های متعلق به یک کلاس ($SWICW$) توسط فرمول (۱) محاسبه می‌شود. سپس گراف‌ها بر اساس مقدار $SWICW$ بصورت نزولی مرتب و بر این اساس رتبه بندی می‌شوند.

$$SWICW = \sum_{k \in C_n} \sum_{i=1}^{nk} \sum_{j=1}^{nk} W_{ij}^k \quad (1)$$

در فرمول (۱) $C1$ کلاس یک و Cn کلاس n ام داده‌ها، nk تعداد نمونه‌های کلاس k ام و W_{ij}^k وزن (شباهت) بین نمونه i و j از کلاس k می‌باشد.

پس از رتبه‌بندی گراف‌ها می‌توان تعدادی مشخص را که بعنوان پارامتر ورودی دریافت می‌شود، از ابتدای لیست انتخاب کرد و یا از بزرگترین گپ بین مقادیر $SWICW$ نرمال شده برای تعیین تعداد گراف‌های انتخابی استفاده کرد.

الگوریتم ۱: محاسبه مقدار $SWICW$ برای گراف‌ها

ورودی: مجموعه گراف‌های ساخته شده براساس ویژگی‌ها و پارامترهای مختلف W ، برچسب داده‌های برچسب‌دار L خروجی: گراف‌های مرتب شده W ، بردار مقادیر $SWICW$ متناظر گراف‌های ورودی

۱. مقدار اولیه $SWICW^i, i = 1, \dots, n$ برابر صفر شود.
۲. برای هریک از گراف‌های ورودی $W^i, i = 1, \dots, n$ مراحل زیر انجام شود:

۲,۱. بخش برچسب‌دار گراف W_{ll}^i جدا شود.
۲,۲. برای هریک از کلاس‌های $C^k, k = 1, \dots, m$ در گراف W_{ll}^i مراحل زیر انجام شود:

۲,۲,۱. مجموع مقادیر وزن بین نمونه‌های متعلق به کلاس C^k محاسبه و به مقدار $SWICW^i$ افزوده شود.
۳. مجموعه گراف‌های W براساس بردار مقادیر $SWICW$ متناظر، بصورت نزولی مرتب شوند.

علاوه بر رتبه بندی گراف‌ها با این روش، می‌توان با نرمالسازی مقادیر $SWICW$ بصورتی که مجموع آن یک شود، از آن برای تعیین وزن متناظر گراف‌ها نیز استفاده کرد. در روش ترکیب میانگین گیری مرتب وزن‌دار ($OWAF^l$) از مقادیر نرمال شده $SWICW$ متناظر گراف‌ها بعنوان وزن استفاده شده است.

برای بهبود دقت شناسایی، ارتباطات اشتباه بین داده‌های کلاس‌های مختلف را باید حذف کرد. به این معنی که دو گره که متعلق به دو کلاس مختلف هستند نباید بین آنها لبه‌ای وجود داشته باشد. حذف لبه‌های اشتباه باعث میشود که در فرآیند انتشار برچسب، برچسب‌ها بخاطر وجود لبه‌های غلط، به اشتباه منتشر نشوند. از آنجاکه انجام این کار برای بخش برچسب‌دار داده‌ها به سادگی قابل انجام است، لذا پیش از ترکیب گراف‌های انتخاب شده، در قسمت داده‌های برچسب‌دار گراف W_{ll} در صورتی که بین دو گره از دو کلاس مختلف یالی وجود داشته باشد، این یال حذف میشود (با صفر کردن وزن آن یال) تا مانع انتشار برچسب اشتباه بشود. نتایج آزمایش‌ها نشان می‌دهد که انجام این مرحله قبل از ترکیب اطلاعات، می‌تواند منجر به بهبود دقت انتشار برچسب شود.

۳- نتایج

در این بخش به بررسی و تحلیل نتایج آزمایش‌های انجام شده پرداخته می‌شود. در بخش ۱-۳ مجموعه داده‌های مورد استفاده به همراه ویژگی‌های استخراج شده از هر یک و پس از آن در بخش ۲-۳ روش استفاده شده برای ساخت گراف مجاورت و انتشار برچسب ذکر شده است.

در بخش ۳-۳ نتایج پیاده‌سازی روش پیشنهادی برای انتخاب بهترین گراف‌ها مورد بررسی قرار می‌گیرد و در نهایت در بخش ۳-۴ نتایج ترکیب گراف‌های انتخاب شده توسط روش پیشنهادی با نتایج ترکیب کل گراف‌ها مقایسه خواهند شد.

در همه آزمایش‌ها ۳۰ درصد داده‌ها برچسب‌دار در نظر گرفته شده‌اند. کلیه پیاده‌سازی‌ها با استفاده از نرم افزار متلب انجام شده است.

۳-۱ - مجموعه داده تصاویر چهره

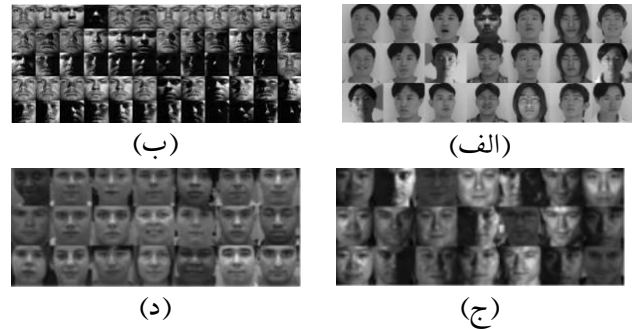
ما از چهار مجموعه داده تصاویر خاکستری چهره به شرح ذیل استفاده کرده ایم:

۱- Extended Yale (Yale Ext): این مجموعه شامل ۱۶۱۲۸ تصویر خاکستری از ۲۸ نفر در ۹ زاویه و ۶۴ وضعیت نور مختلف می باشد. در آزمایش‌های ما یک زیر مجموعه شامل ۱۷۷۴ تصویر مورد استفاده قرار گرفته است.

۲- PF01: این مجموعه شامل تصاویر خاکستری چهره ۱۰۳ نفر شامل ۵۳ مرد و ۵۰ زن در ۱۷ تصویر مختلف (۱ چهره عادی، ۴ حالت نوری مختلف، ۸ زاویه مختلف و ۴ حالت چهره) برای هر فرد می باشد.

۳- PIE: ما از مجموعه کاهش یافته شامل ۱۹۲۶ تصویر چهره‌ی ۶۸ نفر استفاده کرده ایم. تصاویر شامل تغییرات زاویه، نور و حالت چهره هستند.

۴- Feret: ما از زیرمجموعه‌ای از Feret استفاده کرده ایم. این زیرمجموعه شامل ۱۴۰۰ تصویر از ۲۰۰ نفر (۷ تصویر برای هر نفر) می باشد [21].



شکل ۴۷: نمونه‌هایی از مجموعه تصاویر چهره

(الف) PF01، (ب) Yale Ext، (ج) Pie، (د) Feret

۳-۲ - استخراج ویژگی، ساخت گراف مجاورت و انتشار برچسب

برای هر تصویر در هر مجموعه داده، علاوه بر تصویر اصلی سه نوع ویژگی LBP^v، دو نوع ویژگی کواریانس^h و یک ویژگی گابور^g استخراج شده است که مجموعاً ۷ ویژگی مختلف میشوند. سپس با روش L1-Robust [20] با ۷ مقدار آستانه (۰,۵، ۱، ۰,۹، ۰,۵، ۱، ۰,۹، ۱,۵ و ۲)، جمعاً تعداد ۴۹ گراف ساخته شده و در فرآیند انتخاب و ترکیب استفاده می‌شود.

در روش‌های انتشار برچسب هدف این است که برچسب‌داده‌های برچسب به داده‌های بدون برچسب به صورتی منتشر شوند که نمونه‌های نزدیک به هم برچسب‌های مشابه بگیرند، که بصورت ریاضیاتی می‌توان آن را به شکل یک مسئله مینیم کردن بصورت فرمول شماره (۲) بیان کرد. شیوه کار روش انتشار برچسب GFHF¹ [18] به همین صورت است.

$$\min \sum_{i,j=1}^n W_{ij} \|f_i - f_j\|_2^2 \quad (2)$$

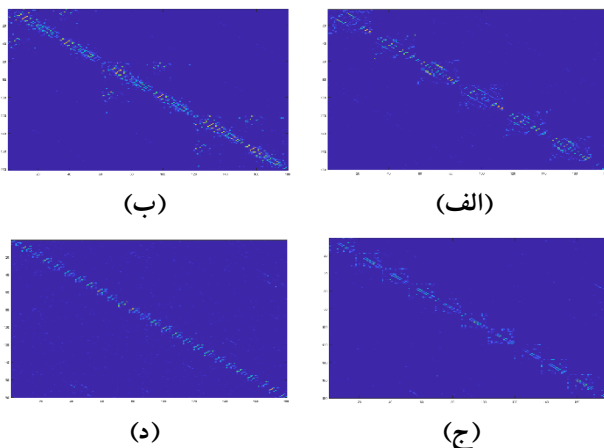
در فرمول (۲) شباهت W_{ij} و f_i و f_j نیز برچسب نمونه‌های x_i و x_j را نشان می‌دهد. در روش انتشار برچسب LGC¹¹ [13] یک بخش دیگر نیز به فرمول (۲) اضافه شده که در آن هدف این است که برچسب‌ها طوری بدست آیند که وقتی برچسب داده‌های برچسب دار را نیز مجدداً تخمین می‌زنیم به برچسب واقعی آنها نزدیک باشد. لذا روش انتشار برچسب LGC به صورت فرمول (۳) نوشته می‌شود.

$$\min \sum_{i,j=1}^n W_{ij} \|f_i - f_j\|_2^2 + \sum_{i=1}^l \|f_i - y_i\|_2^2 \quad (3)$$

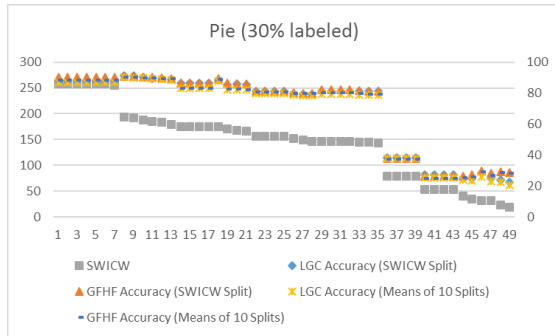
در فرمول (۳) y_i برچسب واقعی و f_i برچسب تخمین زده شده نمونه x_i می باشد. بعنوان معیاری برای مقایسه نتایج، با روش‌های انتشار برچسب GFHF و LGC براساس ده ترکیب مشخص از داده‌های برچسب دار و بدون برچسب، انتشار برچسب انجام شده و میانگین و انحراف معیار دقت گزارش شده است.

۳-۳ - رتبه‌بندی و انتخاب گراف‌ها

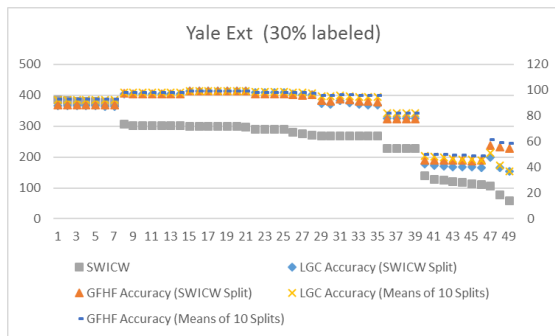
در این بخش نتایج رتبه‌بندی گراف‌ها با استفاده از معیار پیشنهادی SWICW، برای چهار مجموعه داده Feret، PF01، Pie، و Yale Ext مورد بررسی قرار گرفته است. برای بررسی شهودی عملکرد معیار پیشنهادی، برای تمام گراف‌ها داده‌های هر کلاس در کنار یکدیگر قرار داده شده‌اند، لذا با توجه به اینکه انتظار داریم در یک گراف خوب، وزن بین نمونه‌های یک کلاس با یکدیگر زیاد و با نمونه‌های کلاس‌های دیگر کم باشد، پس در صورت ترسیم بصری گراف باید در قطر تصویر مربع‌هایی منظم مشاهده کنیم. این موضوع در شکل (۳) که بزرگنمایی بخشی از گراف‌های چهار مجموعه داده مورد بررسی که بیشترین مقدار SWICW را دارند، به وضوح قابل مشاهده است. همچنین در شکل (۴) بزرگنمایی بخشی از گراف‌های چهار مجموعه داده مورد بررسی که کمترین مقدار SWICW را دارند، مشاهده شده و بصورت شهودی عملکرد معیار پیشنهادی را تایید می‌کند.



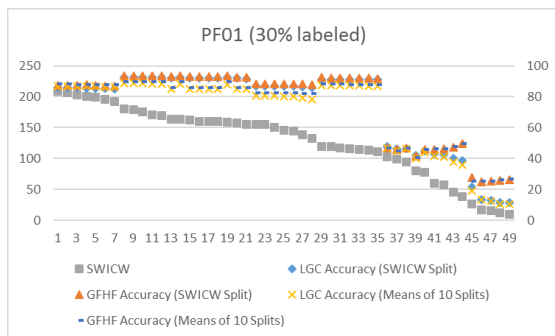
شکل ۴۸: بزرگنمایی بخشی از گرافی که بیشترین مقدار SWICW را دارد (الف) Pie، (ب) Yale EXT، (ج) PF01، (د) Feret



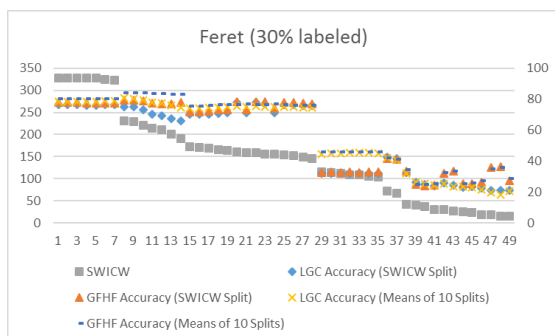
(الف)



(ب)



(ج)

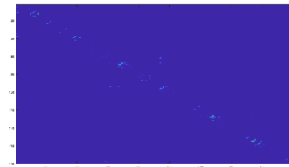


(د)

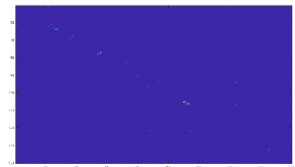
شکل ۵۰: مقایسه مقدار SWICW و دقت روش‌های انتشار برچسب GFHF و LGC

(برای یک ترکیب و میانگین ده ترکیب برچسب دار مختلف)

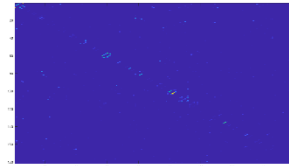
(الف) Pie، (ب) Yale Ext، (ج) PF01، (د) Feret



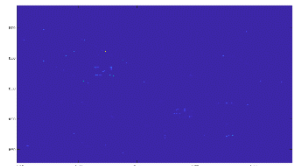
(ب)



(الف)



(د)



(ج)

شکل ۴۹: بزرگنمایی بخشی از گرافی که کمترین مقدار SWICW را دارد (الف) Pie، (ب) Yale EXT، (ج) PF01، (د) Feret

برای بررسی دقیق‌تر معیار پیشنهادی، همانطور که در بخش ۲-۳ شرح داده شده، برای ۴۹ گراف ساخته شده براساس ویژگی و پارامترهای مختلف، تطابق مقدار SWICW با دقت روش‌های انتشار برچسب LGC و GFHF برای یک ترکیب^{۱۱} داده‌های برچسب‌دار (۳۰ درصد) و بدون برچسب (۷۰ درصد) و میانگین ده ترکیب مختلف، بررسی شده است. ابتدا براساس یک ترکیب مشخص داده‌های برچسب دار و بدون برچسب مقادیر SWICW و دقت روش‌های انتشار برچسب LGC و GFHF محاسبه شده (SWICW Split) و سپس با توجه به اینکه ممکن است استفاده از ترکیب دیگری از داده‌های برچسب‌دار و بدون برچسب نتایج متفاوتی داشته باشد، برای بررسی اینکه بخش برچسب‌دار داده‌های مورد آزمایش تا چه حد می‌تواند تعمیمی از کل داده‌ها باشد، میانگین دقت روش‌های انتشار برچسب LGC و GFHF برای ده ترکیب مختلف از داده‌های برچسب‌دار و بدون برچسب، برای همه ۴۹ گراف محاسبه (Means of 10 Split) و سپس نتایج بدست آمده براساس مقدار SWICW بصورت نزولی مرتب و گزارش شده است. چنانکه بیان شد مقدار SWICW نشان دهنده قدرت اتصال نمونه‌های متعلق به یک کلاس است و از آنجاکه کلیه وزن‌ها نرمال شده‌اند با افزایش مقدار وزن نمونه‌های متعلق به کلاس‌های متفاوت، مقدار SWICW کاهش خواهد یافت. لذا انتظار داریم با کاهش مقدار SWICW، دقت انتشار برچسب نیز کاهش یابد. تناظر بین مقدار SWICW و دقت انتشار برچسب تا حد زیادی در شکل (۵) مشاهده می‌شود. همچنین همپوشانی زیاد دقت انتشار برچسب یک و ده ترکیب مختلف داده‌های برچسب‌دار و بدون برچسب، نشان می‌دهد که با استفاده از بخش برچسب‌دار گراف L1-R تا حد زیادی می‌توان در مورد کلیت آن قضاوت کرد.

چنانکه از نتایج مشخص است، با وجود اینکه در برخی قسمت‌های نمودار، معیار پیشنهادی با دقت انتشار برچسب در تناظر کامل نیست ولی در اکثر موارد بهترین گراف‌ها در ابتدای فهرست موارد انتخاب شده قرار می‌گیرند.

۳-۴- نتایج ترکیب گراف‌ها

مجموعه داده‌های Feret, PF01, Yale Ext, Pie

دقت انتشار برچسب				گراف	مجموعه
گراف‌های انتخاب شده		کل گراف‌ها			
LGC	GFHF	LGC	GFHF		
80.6±10	84.3±8.3	80.6±10	84.3±8.3	بالاترین	Feret7
95.1±7.1	96.4±5.1	86.3±8.7	89.9±7.1	UWAF	
94.4±7.7	95.3±6.5	85.2±9.9	88.4±8.8	OWAF	
88.7±4.7	89.8±4.3	88.7±4.7	89.8±4.3	بالاترین	PF01
98.2±1.5	98.7±1.1	94.7±2.9	96.3±1.8	UWAF	
98.4±1.4	98.9±1	95.8±2.4	97±1.6	OWAF	
90±1.9	90.3±1.6	90±1.9	90.3±1.6	بالاترین	Pie
96±1.3	96.2±1.1	91.5±2.1	92.6±1.2	UWAF	
95.9±1.5	96±1.1	92.7±1.8	93.6±1.2	OWAF	
99.6±0.2	99.6±0.2	99.6±0.2	99.6±0.1	بالاترین	Yale Ext
99.7±0.2	99.7±0.3	99.2±0.4	99.3±0.5	UWAF	
99.6±0.3	99.6±0.3	99.3±0.4	99.4±0.4	OWAF	

برای ترکیب گراف‌ها دو روش میانگین‌گیری ساده ($UWAF^{13}$) و میانگین‌گیری مرتب وزن دار (OWAF) که وزن‌های آن براساس مقادیر نرمال شده (بصورتی که مجموع یک بشود) SWICW تعیین شده، مورد استفاده قرار گرفته است. قبل از ترکیب گراف‌های انتخاب شده، چنانکه در بخش ۲ (روش پیشنهادی) شرح داده شده، برای بخش داده‌های برچسب دار گراف (W_{II}) لبه بین داده‌های متعلق به کلاس‌های متفاوت حذف شده، که این کار در بهبود نتایج موثر بوده است.

در جدول (۱) بالاترین دقت و نتایج ترکیب با دو روش برای کلیه گراف‌ها و گراف‌های انتخاب شده با معیار پیشنهادی SWICW نمایش داده شده است.

چنانکه از نتایج بخش قبل نیز مشخص بود، گراف با بهترین دقت در بین موارد انتخاب شده و معمولاً در ابتدا فهرست قرار دارد. لذا ۷۰٪ (۳۴) گراف از ابتدای لیست برای ترکیب انتخاب شده است. تعداد یا درصد گراف‌هایی که از ابتدا لیست انتخاب می‌شوند، پارامتر ورودی الگوریتم پیشنهادی ما می‌باشد.

نتایج جدول (۱) نشان می‌دهد در اکثر موارد، ترکیب گراف‌های انتخاب شده توسط روش پیشنهادی منجر به بهبود دقت، نسبت به ترکیب همه گراف‌ها شده است، به این معنا که بهتر از ترکیب همه گراف و یا نزدیک به آن است. لذا بکارگیری دو مرحله انتخاب بهترین گراف‌ها و ترکیب آنها در کنار یکدیگر می‌تواند منجر به کسب بهترین دقت در بین گراف‌های موجود و یا بالاتر از آن شود.

به دلیل اینکه روش ترکیب OWAF بر مبنای معیار پیشنهادی SWICW به گراف‌های بهتر وزن بیشتری می‌دهد، لذا در حالت ترکیب همه گراف‌ها، در اکثر موارد عملکرد روش ترکیب OWAF بهتر از UWAF است و این موید معیار پیشنهادی SWICW است. البته به دلیل اینکه تناظر SWICW و دقت انتشار در برخی موارد کاملاً برقرار نیست، استفاده از این نوع ترکیب برای موارد انتخاب شده (شامل بهترین گرافها) برتری نداشته است. لذا برای داده‌های تست شده انتخاب با معیار پیشنهادی و سپس ترکیب UWAF کارایی بهتری داشته است. این نتایج نشان می‌دهد که ترکیب گراف‌های انتخاب شده براساس معیاری پیشنهادی SWICW، حتی با ساده ترین روش ترکیب نیز می‌تواند منجر به بهبود نتایج شود.

البته با توجه به اینکه اکثر گراف‌های ترکیب شده دقت بالایی دارند و همچنین اکثر گراف‌هایی که در آنها وزن نمونه‌های یک کلاس با یکدیگر ضعیف است (مقدار SWICW کم است) با نمونه‌های سایر کلاس‌ها نیز ارتباط (مقدار) ضعیفی دارند. در مقایسه بین نتیجه ترکیب همه گراف‌ها و ترکیب گراف‌های انتخاب شده براساس SWICW در برخی موارد نتایج بهبود نداشته و یا بهبود قابل توجه نیست. چنانکه تعداد گراف‌های ضعیف که وزن نمونه‌های کلاس‌های مختلف با یکدیگر بیشتر از وزن نمونه‌های متعلق به کلاس‌های یکسان است، افزایش یابد، بهبود در نتایج انتخاب و ترکیب پیشنهادی مشهودتر خواهد بود.

جدول ۱۶: نتایج انتخاب و ترکیب گراف‌های ساخته شده براساس ویژگی‌ها و یا پارامترهای مختلف

۴- نتیجه

در این مقاله، ما یک روش رتبه‌بندی و تعیین وزن جهت انتخاب و ترکیب گراف‌ها پیشنهاد دادیم. نتایج آزمایش‌ها نشان می‌دهند که روش پیشنهادی برای انتخاب، بهترین گراف‌ها را در ابتدای فهرست و یا جزو اولین موارد قرار می‌دهد. همچنین ترکیب نیز در اکثر موارد نتیجه‌ای نزدیک به بهترین دقت و یا بالاتر از آن دارد.

در نهایت از نتایج مشخص است بکارگیری دو روش انتخاب و ترکیب پیشنهاد شده در کنار یکدیگر می‌تواند کسب نتیجه‌ای نزدیک به بالاترین دقت موجود در گراف‌ها و یا بهتر از آن منجر شود.

از آنجا که تخمین تاثیر وزن بین نمونه‌های کلاس‌های مختلف (ارتباطات اشتباه بین داده‌های کلاسهای مختلف، به این معنی که دو گره که متعلق به دو کلاس مختلف هستند و نباید بینشان لبه‌ای وجود داشته باشد)، در فرایند انتشار برچسب، روی دقت انتشار برچسب بررسی نشده (برچسب‌ها بخاطر وجود لبه‌های غلط، ممکن است به اشتباه منتشر شوند)، می‌تواند جزو کارهای آتی ما باشد. همچنین در مواردی که تعداد داده‌های برچسب دار خیلی کم باشد، احتمال اینکه روش پیشنهادی بتواند تخمین مناسبی از دقت انتشار برچسب گراف داشته باشد کاهش می‌یابد. لذا پیشنهاد روشی سریع برای انتخاب گراف‌ها که به داده‌های برچسب دار وابسته نباشد نیز از کارهای آتی ما خواهد بود.

مراجع

- [1] B. Wang et al., "Similarity network fusion for aggregating data types on a genomic scale," Nat. Methods, vol. 11, no. 3, pp. 333–337, 2014.
- [2] A. Blum and S. Chawla, "Learning from labeled and unlabeled data using graph mincuts," 2001.
- [3] M. Szummer and T. Jaakkola, "Partially labeled classification with Markov random walks," in Advances in neural information processing systems, 2002, pp. 945–952.
- [4] T. Joachims, "Transductive learning via spectral graph partitioning," in Proceedings of the 20th International Conference on Machine Learning (ICML-03), 2003, pp. 290–297.
- [5] X. Zhu, "Semi-Supervised Learning: From Gaussian Fields to Gaussian Processes", 2003.
- [6] S. Thrun, L. K. Saul, and B. Schölkopf, Advances in Neural Information Processing Systems 16: Proceedings of the 2003

¹¹ Local and Global Consistency

¹² Split

¹³ UnWeighted Averaging Fusion

- [7] Conference, vol. 16. MIT press, 2004.
- [7] M. Herbster, M. Pontil, and L. Wainer, "Online learning over graphs," in Proceedings of the 22nd international conference on Machine learning, 2005, pp. 305–312.
- [8] V. Sindhwani, P. Niyogi, and M. Belkin, "Beyond the point cloud: from transductive to semi-supervised learning," in Proceedings of the 22nd international conference on Machine learning, 2005, pp. 824–831.
- [9] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," J. Mach. Learn. Res., vol. 7, no. Nov, pp. 2399–2434, 2006.
- [10] Y. Bengio, O. Delalleau, and N. Le Roux, "l1 label propagation and quadratic criterion," 2006.
- [11] M. Karasuyama and H. Mamitsuka, "Multiple graph label propagation by sparse integration," IEEE Trans. Neural Networks Learn. Syst., vol. 24, no. 12, pp. 1999–2012, 2013.
- [12] B. Wang and J. Tsotsos, "Dynamic label propagation for semi-supervised multi-class multi-label classification," Pattern Recognit., vol. 52, pp. 75–84, 2016.
- [13] J. Gui, R. Hu, Z. Zhao, and W. Jia, "Semi-supervised learning with local and global consistency," Int. J. Comput. Math., vol. 91, no. 11, pp. 2389–2402, 2014.
- [14] R. G. W. Verhaak et al., "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1," Cancer Cell, vol. 17, no. 1, pp. 98–110, 2010.
- [15] C. G. A. Network, "Comprehensive molecular portraits of human breast tumours," Nature, vol. 490, no. 7418, p. 61, 2012.
- [16] C. G. A. R. Network, "Comprehensive genomic characterization of squamous cell lung cancers," Nature, vol. 489, no. 7417, p. 519, 2012.
- [17] C. G. Willett, D. T. Chang, B. G. Czito, J. Meyer, and J. Wo, "Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012.(5)," Int. J. Radiat. Oncol. Biol. Phys., vol. 86, no. 1, 2013.
- [18] X. Zhu, J. Lafferty, L. Cs, and C. M. U. Edu, "Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions," in Proceedings International Conference on Machine Learning, 2003, pp. 912–919.
- [19] F. Nie, L. Tian, R. Wang, and X. Li, "Multiview Semi-Supervised Learning Model for Image Classification," IEEE Trans. Knowl. Data Eng., vol. PP, no. c, pp. 1–1, 2019.
- [20] Y. Fujiwara, Y. Ida, J. Arai, M. Nishimura, and S. Iwamura, "Fast algorithm for the lasso based l1-graph construction," in Proceedings of the VLDB Endowment, 2016, vol. 10, no. 3, pp. 229–240.
- [21] F. Dornaika, M. T. Kejani, and A. Bosaghzadeh, "Graph construction using adaptive Local Hybrid Coding scheme," Neural Networks, vol. 95, pp. 91–101, 2017.

زیرنویس‌ها

¹ Fusion

² manifold assumption

³ Similarity Network Fusion

⁴ Sparse Multiple Graph Integration

⁵ Sum of Within Class Weights

⁶ Ordered Weighted Averaging Fusion

⁷ Local Binary Patterns

⁸ Covariance

⁹ Gabor

¹⁰ Gaussian Field and Harmonic Functions



بازشناسی دیداری اشیاء از سیگنال تک ثبت EEG

مجتبی یاوندحسینی، فواد قادری*

آزمایشگاه تعامل انسان و کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران

fghaderi@modares.ac.ir

چکیده

ارتباط بین فعالیت‌های مغزی که با روش‌های مختلف مانند الکتروانسفالوگرافی (EEG) ثبت می‌شوند با مسائل دنیای واقعی اهمیت بررسی آنها را دوچندان کرده است. با این وجود سطح بالای نوفه (noise) در سیگنال‌های ثبت شده EEG و اختلافات موجود بین داده‌های ثبت شده کاربران مختلف کار طبقه بندی سیگنال‌های دریافتی را با چالش جدی مواجه می‌کند. در این مطالعه ظرفیت برخی از الگوریتم‌های معروف یادگیری ماشین برای یادگیری ویژگی‌های مهم و بارز از روی سیگنال‌های EEG که با روش تک ثبت از پاسخ‌های مغز به تصاویر محرک ضبط شده اند را ارزیابی کرده و ارایه می‌دهیم. علاوه بر بررسی و مطالعه‌ی روش‌های مختلف، نشان داده شده است که چگونه با نگاشت صحیح کل فضای داده‌ای به فضای ویژگی‌های بامعنی (IFS) کارایی تقریباً هم‌همی روش‌های دسته‌بندی سیگنال‌های EEG که در این تحقیق استفاده شده‌اند را بهبود دهیم. نتایج بررسی‌های ما از تمامی نتایج مشابه روی مجموعه داده‌ی بررسی شده بهتر است و توانایی مناسب و موفق برخی از الگوریتم‌های یادگیری ماشین در به دست آوردن ساختار مجزا و تفکیک شده‌ی اشیاء در دسته‌بندی سیگنال‌های ذخیره شده‌ی تک ثبت را نشان می‌دهد.

کلمات کلیدی

سیگنال‌های EEG تک ثبت؛ دسته‌بندی؛ بازشناسی دیداری اشیاء؛ محرک.

۱- مقدمه

کننده در یک دوره آزمایش نشان داده شده بدست آمده‌اند. شرکت کنندگان مجموعه‌ای از عکس‌های اشیاء با ترتیب و ساختار برنامه ریزی شده را مشاهده می‌کردند. پس از پیش پردازش و تمیز کردن داده‌ها از نوفه و آرتیفکت^۱ بردارهای ویژگی ساخته می‌شوند. این بردارهای ویژگی همچنان به خاطر ماهیت سیگنال‌های EEG به ویژه در روش تک ثبت، هم حجم بزرگی دارند هم حاوی اطلاعات غیر مفید راجع به مساله‌ی مربوطه هستند. این موارد باعث بروز مشکلاتی مثل بیش‌برازش و بار محاسباتی بالا در اکثر الگوریتم‌های دسته‌بندی شده و کارایی آنها را به شدت کم می‌کند. بنابراین ما در روش پیشنهادی که در بخش‌های بعدی توضیح داده خواهد شد به دنبال تشکیل مجموعه‌ی مؤثری از مولفه‌های مهم هستیم که علاوه بر حجم کم‌تر، بیشترین جداکنندگی را در طبقه‌بندی داشته باشند. از نکات مثبت و وجه تمایز این کار با کارهای قبلی دستیابی سریع به این مجموعه در مرحله آموزش مدل‌ها می‌باشد و در مرحله‌ی تست نیازی به این کار نیست. در ادامه‌ی بررسی‌های این مطالعه تعدادی از الگوریتم‌های معروف یادگیری ماشین برای دسته بندی سیگنال‌های EEG را هم با روش پیشنهادی هم بدون آن بررسی کرده و نتایج را با هم مقایسه می‌کنیم.

الکتروانسفالوگرافی (EEG) اندازه گیری غیر تهاجمی میدان‌های الکتریکی مغز است. الکترودها روی پوست سر قرار داده می‌شوند تا پتانسیل‌های ناشی از جریان الکتریکی در نرون‌های مجاور را ضبط کنند [۱]. با پردازش و طبقه‌بندی سیگنال‌های EEG در کاربردهای مختلف از بازی‌های ویدئویی گرفته تا کنترل اندام‌های مصنوعی، بازوهای رباتیک و کارهای تقریباً جدیدی مانند تشخیص گروه یا دسته‌ی اشیاء، مکان‌یابی اشیاء در تصاویر یا حاشیه‌نویسی و توصیف محتوای تصاویر، نتایج قابل قبولی حاصل شده است. در سیستم بینایی انسان تشخیص اشیاء از دسته‌های مختلف، در کسری از ثانیه با دقت قابل توجهی انجام می‌شود. سیستم بینایی انسان با مکانیسم‌های سریع و کارآمدی برای حل این موضوع تکامل یافته است [۲-۴].

در این تحقیق به بررسی دقت و سرعت برخی الگوریتم‌های یادگیری ماشین در دسته بندی سیگنال‌های EEG تک ثبت^۲ می‌پردازیم. نمونه‌های نوفه‌ای سیگنال از محرک‌های تصاویر که برای هر شرکت

¹ Neuron

² Single-trial

³ Noise

⁴ Artifact

۲- پژوهش‌های مرتبط

دسته‌بندی‌های مختلف وظایف تصور حرکتی^۱ در واسط‌های کامپیوتری مغز (BCI) و تحلیل‌های آماری آنها در [۷-۵] مقایسه شده است. بازشناسی اشیاء^۲ به طریق بخش بندی^۳ و تفکیک اشیاء از پیش زمینه‌ی تصویر، با سرعت و دقت تمایز بالا از گروه و دسته‌ی اشیاء، مستقل از تغییرات اندازه و موقعیت و محتوای پیش زمینه و همچنین فعالیت‌های ماشینی مثل ردیابی اشیاء^۴ در [۸] توسط دیکارلو^۵ و همکاران بررسی شده است. فیلترهای مکانی مشترک^۶ (CSP) و بررسی مولفه‌های مستقل^۷ (ICA) در [۹] به منظور جداسازی چهار کلاسه‌ی وظایف تصور حرکتی ارائه شده است. این وظایف حرکتی شامل حرکت‌های دست راست و دست چپ و دهان و پا می‌باشد. در این بررسی الگوریتم‌های مختلف فیلترهای مکانی مشترک و مولفه‌های مستقل مقایسه شده‌اند. مطالعه‌ی [۳] درک فضایی-زمانی در پردازش گروه و دسته‌ی اشیاء و تغییرات آنها در کل مغز را ارائه می‌دهد. در این مطالعه ده نفر در یک الگوی ضبط EEG پویا شرکت کرده و در آن گزارش می‌دهند که نقاط ثابت و شاخص دو شی که به طور پیوسته در هر آزمایش ارائه شده‌اند دارای رنگ یکسان یا متفاوت هستند. نویسندگان [۱۰] یک روش خودکار برای طبقه‌بندی حرکت‌های دست چپ یا راست ارائه دادند و بعد از پیش پردازش، ویژگی‌های استخراج شده با الگوریتم‌های ماشین بردار پشتیبان (SVM) و شبکه عصبی مصنوعی (ANN) دسته‌بندی کردند.

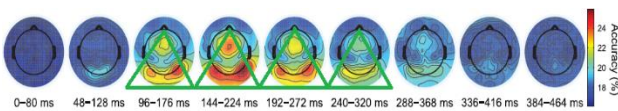
در بیشتر تحقیق‌ها و مطالعه‌های قبلی بر روی داده‌های EEG تعداد کلاس‌ها یا دسته‌هایی که مورد بررسی قرار گرفته‌اند کم بوده‌اند، اما در مطالعه [۱۱] کانشیرو^۸ و همکاران در طبقه‌بندی‌های مختلف و در مقیاس بزرگ کار کرده‌اند. آنها تجزیه و تحلیل بازنمایی شباهت در تغییرهای پردازش اشیاء با استفاده از دسته‌بندی‌های سیگنال‌های EEG تک‌ثبت را بررسی کرده‌اند.

۳- روش پیشنهادی

بسیاری از ویژگی‌های غیرمفید داده‌های EEG در مرحله تمیز کردن و پیش‌پردازش آماده‌سازی داده‌های EEG برای کارهای دسته‌بندی یادگیری ماشینی حذف می‌شوند. پس از ایجاد بردارهای ویژگی از داده‌های تمیز، کاهش ابعاد و استخراج ویژگی‌های دارای بیشترین جداکنندگی، نقش اصلی را در بهبود عملکرد مدل‌های طبقه‌بندی بازی می‌کند. برای دستیابی به این هدف، روش‌ها و الگوریتم‌هایی مانند استفاده از تحلیل مؤلفه‌های اصلی^۹ (PCA) و تحلیل افتراقی خطی^{۱۰} (LDA) و غیره وجود دارد [۱۲-۱۴].

در برخی از مساله‌های طبقه‌بندی EEG تعیین الکترودهای مفید کار مهمی است. نادیده گرفتن کانال‌های غیر ضروری مسیر توسعه سیستم‌های BCI سیار را هموار می‌کند [۱۶-۱۵]، اما در سیستم‌های جدید و پیچیده، مناطق مربوطه روی کاسه‌ی سر ممکن است مشخص نباشد و محدوده‌ی الکترودهای موثر به درستی قابل تعیین نباشد. از طرفی اگر هم با روش‌های خاصی این کار صورت پذیرد به صورت یک روش عمومی بوده و برای هر مجموعه داده‌ای ممکن است کارایی لازم را نداشته باشد.

روش موثر این است که داده‌های غیرمفید با نگاشت صحیح کل فضای داده‌ای به فضای ویژگی‌های بامعنی^{۱۱} (IFS) حذف گردند. این امر با انتخاب ناحیه‌هایی از الکترودهای سر که به درستی مرتبط با نرون‌های فعال و درگیر در قشر بینایی مغز^{۱۲} در کار مربوطه هستند، حاصل می‌شود. برای دستیابی به این هدف پس از آموزش اولیه مدل با کل فضای داده‌ای، دقت همه‌ی تک الکترودها در روی همین داده‌های آموزشی تعیین می‌گردد. سپس با انتخاب یک آستانه‌ی دقت، مجموعه‌ای از الکترودهای موثر و دقیقاً مرتبط با وظیفه‌ی در حال انجام و آموزش تشکیل شده و از روی آنها بردارهای ویژگی ساخته می‌شود. پس از این مرحله، مجدداً با استفاده از الگوریتم‌های کاهش ابعاد تعداد ویژگی‌ها کاسته شده و به مجموعه‌ای به نام فضای ویژگی‌های بامعنی (IFS) دست خواهیم یافت. در حقیقت این روش به بهبود عملکرد الگوریتم‌های کاهش ابعاد در فضای ویژگی‌های حاوی اطلاعات مفید کمک می‌کند. مدل نهایی مجدداً روی این ویژگی‌های بامعنی آموزش داده می‌شود. برای فاز تست فقط نیاز به نگهداری شماره الکترودهایی که در مرحله‌ی قبل پیدا شدند داریم. برای عمومی‌سازی^{۱۳} و انتخاب حاشیه اطمینان بهتر می‌توانیم مقدار آستانه را با بخشی از داده‌ها برای اعتبارسنجی مشخص کنیم. پس از ترسیم دقت تک الکترودها می‌توان دید که برای وظیفه‌ی مساله‌ی مورد بررسی، کدام ناحیه‌ی قشر بینایی انسان فعال شده است. در شکل ۱ این نواحی به وضوح نشان داده شده است.



شکل ۱. مناطق فعال قشر مغز در پنجره‌های زمانی مختلف برای نمایش دقت دسته‌بندی تک‌تک الکترودها که فضای ویژگی‌های بامعنی (IFS) را نشان می‌دهد. تصویر اصلی از [۱۱] تهیه شده است.

۳-۱- کاهش ویژگی‌ها

هدف اصلی کاهش ابعاد و ویژگی‌ها، کم کردن تعداد ویژگی‌ها و کوچک‌تر کردن ابعاد مساله از طریق حذف ویژگی‌های تکراری و وابسته به هم می‌باشد. ما برای رسیدن به این هدف ویژگی‌ها را از فضای با ابعاد بالا به فضای پایین‌تر نگاشت کرده و انتقال می‌دهیم.

¹¹ Informative Feature Space (IFS)

¹² Brain visual cortex

¹³ Generalization

¹ Motor imagery tasks

² Object recognition

³ Segmentation

⁴ Object tracking

⁵ DiCarlo

⁶ Common Spatial Filtering

⁷ Independent component analysis

⁸ Kaneshiro

⁹ Principal Component Analysis

¹⁰ Linear Discriminant Analysis

تعریف می‌شود. به عبارت دیگر، با توجه به داده‌های آموزش با برچسب داده شده، یک ابر صفحه بهینه را به نمایش می‌گذارد که نمونه‌های جدید را طبقه‌بندی می‌کند و می‌تواند در طبقه‌بندی یا مباحث رگرسیون مورد استفاده قرار گیرد، هر چند که این الگوریتم بیشتر در مسائل طبقه‌بندی استفاده می‌شود. در فضای دو بعدی این ابر صفحه یک خط است که یک صفحه را به دو قسمت تقسیم می‌کند به نحوی که هر کلاس در یک طرف قرار بگیرد. یادگیری ابر صفحه‌ها در SVM خطی با تبدیل مساله به کمک روش‌های جبر خطی انجام می‌شود. SVM روشی به نام ترفند هسته^۷ دارد و از توابعی تشکیل شده که فضای ورودی با ابعاد بالا را به یک فضای کم بعدتر تبدیل می‌کند. درحقیقت به کمک توابعی که هسته نامیده می‌شود یک مساله‌ی غیرقابل جدا شدن را به مساله‌ی قابل جدا شدن تبدیل می‌کنند. هسته‌های چندجمله‌ای یا نمای خط جدا ساز را در ابعاد بالاتر محاسبه می‌کنند و بیشتر در مسائل غیر خطی مفید هستند.

- **LDA:** این روش تعمیم یافته‌ی جداکننده‌ی خطی فیشر^۸ است [۱۴، ۱۳]. این روش روی داده‌های دارای برچسب به دنبال یافتن ترکیبی خطی از ویژگی‌هایی است که دو یا چند کلاس از اشیاء یا رویدادها را توصیف یا جدا می‌کنند و در آمار، تشخیص الگو، یادگیری ماشین و غیره مورد استفاده قرار می‌گیرد. روش دسته‌بندی که با لحاظ برچسب‌ها برای نگاشت فضای ویژگی‌های با ابعاد بالا به یک فضای با ابعاد پایین‌تر به منظور مدل‌سازی تفاوت‌های دسته‌ها مورد استفاده قرار می‌گیرد. LDA یک روش بسیار رایج، ساده و مفید برای مسائل دسته‌بندی نظارت شده است.
- **OVA(LDA):** یکی در برابر همه (OVA)، یکی در برابر بقیه^۹ (OVR)، یا یکی در مقابل همه^{۱۰} (OAA) [۲۰] یک مدل ترکیبی بررسی بر مبنای دسته‌بندی دوگانه با نام‌های مختلف است. استراتژی شامل آموزش یک دسته‌بند جدا در هر کلاس است که نمونه‌های آن کلاس به عنوان نمونه‌های مثبت و تمام نمونه‌های دیگر به عنوان منفی تلقی می‌شوند. این روش در برخی از مسائل طبقه‌بندی در مقیاس بزرگ استفاده می‌شود. جایی که چندین کلاس برای یک نمونه قابل پیش بینی باشد برچسب‌های تک کلاسه به تنهایی می‌تواند منجر به ابهاماتی شود. در این گونه مسائل الگوریتم به جای دادن فقط یک برچسب به هر کلاس به دسته‌بندی کننده‌های مینا نیاز دارد تا بر اساس رای یک تصمیم مطمئن واقعی بگیرد. دسته‌بندی کننده‌های مینا می‌تواند یکی از دسته‌بندهای دودودی مناسب و کارآمد مانند LDA باشد.
- **RF:** جنگل تصادفی [۲۲، ۲۱] شامل تعداد زیادی از درختان تصمیم‌گیری^{۱۱} تکی است که به عنوان یک گروه فعالیت می‌

کاهش ابعاد و ویژگی‌ها راهی است که با کاهش پیچیدگی مدل از مشکل بیش برازش پیشگیری کرده و سرعت آموزش و اجرایی مدل را بیشتر می‌کند. برای این منظور ما از دو روش PCA و LDA که از الگوریتم‌های معروف کاهش ابعاد و ویژگی هستند استفاده می‌کنیم. PCA: تحلیل مؤلفه‌های اصلی یکی از روش‌های موثر کاهش ابعاد بدون نظارت است که برچسب‌ها را در نظر نمی‌گیرد. مانند [۱۱]، ما از کاهش ابعاد در ماتریس داده X با استفاده از PCA و تجزیه مقدار منفرد^۲ (SVD) استفاده کردیم، بدین ترتیب ماتریس $X = USV^T$ به ماتریس $XV = US$ در فضای مولفه‌ها برده می‌شود [۱۷، ۱۲]. LDA: تحلیل افتراقی خطی (LDA)، یک روش کاهش ابعاد نظارت شده است که برچسب‌ها را در نظر می‌گیرد. نسخه اصلی جداکننده‌های خطی توسط رونالد فیشر^۳ برای تجزیه و تحلیل جداکننده‌های دوگانه در سال ۱۹۳۶ تهیه شده است [۱۳]. کاهش ابعاد یا پیش بینی دسته‌ی اعضا، از کاربردهای جداکننده‌های خطی در مسائل دسته‌بندی می‌باشد. این الگوریتم در کاربردهای کاهش ابعاد، به دنبال تشکیل مجموعه‌ی مؤثری از مولفه‌های مهم است که بیشترین جداکنندگی را داشته باشند. [۱۴].

۳-۲- دسته بندی تک ثبت

به طور معمول تهیه‌ی بیشتر مجموعه‌داده‌های EEG با مجموعه‌ای از مشاهده‌های برچسب زده شده برای هر آزمایش انجام می‌شوند. همچنین به منظور بهبود عملکرد در برابر نوفه‌ی بیشتر مدل‌های موجود تجزیه و تحلیل سیگنال‌های EEG، براساس میانگین‌گیری بیش از چندین تکرار آزمایش، انجام می‌شوند [۱۸]. اما در این مطالعه، مدل‌های طبقه بندی بر روی بردارهای ویژگی تک ثبت اعمال می‌شوند. با این روش سیگنال‌های داده EEG از کیفیت مناسبی مانند مجموعه داده‌های EEG معمولی برخوردار نیستند. نکته مثبت این است که مدل آموخته شده با این نمونه سیگنال‌های نوفه‌ای با ویژگی‌های ضعیف، به صورت واقعی قابل استفاده خواهد بود.

۳-۳- مدل‌ها

هر الگوریتم یادگیری ماشین دارای مزایا و معایبی در یادگیری و برخورد با مسائل مربوطه است. ما از بعضی از آنها که برای تجزیه و تحلیل داده‌های EEG معروف و قدرتمند هستند استفاده می‌کنیم: ماشین‌های بردار پشتیبان (SVM)، تحلیل افتراقی خطی (LDA)، روش چندکلاسه‌ی یکی در برابر همه^۴ (OVA)، جنگل تصادفی^۵ (RF)، تقویت شیب حداکثری^۶ (XGB). توضیح هر یک از روش‌ها در ادامه ارائه شده است:

- **SVM:** ماشین بردار پشتیبان [۱۹] دسته‌بندی کننده جدا ساز است که با داده‌های دارای برچسب آموزش می‌بیند (یادگیری نظارت شده) و بطور رسمی توسط یک ابرصفحه‌ی جداکننده

¹ Label

² Singular Value Decomposition

³ Ronald Fisher

⁴ One Versus All

⁵ Random forest

⁶ Extreme Gradient Boosting

⁷ Kernel

⁸ Fisher

⁹ One Versus Rest

¹⁰ One Against All

¹¹ Decision Tree

شده [۲۹] به خاطر اینکه امکان تحلیل داده به روش تک ثبت وجود دارد، مناسب‌تر می‌باشد. در حالیکه شرکت‌کننده‌ها در حال مشاهده هفتاد و دو عکس از اشیاء واقعی به عنوان محرک^۳ بودند پاسخ‌های مغزی آنها ضبط شده است. ده شرکت‌کننده از سنین ۲۱ تا ۵۷ سال (میانگین سنی ۳۰،۵ سال؛ ۳ نفر زن؛ ۱ نفر چپ‌دست) در این آزمایش شرکت کردند. دسته‌بندی تصاویر در شش رده گروهی عمده انجام شده است: بدن انسان (HB)، صورت انسان (HF)، بدن حیوانات (AB)، صورت حیوانات (AF)، سبزیجات میوه ای (FV) و اشیاء بی جان (IO). مجموعه داده با روش تک ثبت و با دستگاه EGI's HydroCel Geodesic Sensor Net، ۱۲۸ کاناله ثبت شده است [۳۰]. فرکانس نمونه‌گیری داده‌ها ۶۲،۵ هرتز و تعداد نمونه‌ها در هر آزمایش ۳۲ نقطه زمانی می‌باشد. داده‌ها از همه الکترودها و تمام ۳۲ نمونه‌ی زمانی از ۰-۴۹۶ میلی‌ثانیه در بردارهای ویژگی قرار می‌گیرند. بردارهای ویژگی از نمونه سیگنال‌های زمانی در هر الکتروده تشکیل می‌شوند. بزرگی آنها خیلی به تعداد آزمایش‌ها وابسته است. در این مطالعه، مجموعه کامل ۱۲۴ الکتروده و ۳۲ نمونه‌ی زمانی در هر الکتروده، یک بردار ویژگی به طول $3,968 = 32 \times 124$ را تولید می‌کند، و به روش IFS، زیرمجموعه‌ای از ۱۲۴ الکتروده و ۳۲ نمونه‌ی زمانی در هر الکتروده یک بردار ویژگی از طول $536,1 = 32 \times 48$ تولید می‌کند. بزرگی این تعداد باید با ۵۱۸۴ آزمایش برای هر شرکت‌کننده مقایسه شود. بردار ویژگی‌های با ابعاد بالا علاوه بر فرایند آموزشی وقت گیر، ممکن است منجر به مشکلات بیش برآزش مدل روی داده‌های غیر مفید یا نوفه‌ها شود. از طرف دیگر، همبستگی زیاد بین الکترودهای مجاور ممکن است باعث افزونگی اطلاعات نیز شود. این موارد از چالش‌های ماهیتی مجموعه داده‌های سیگنال‌های EEG است که در روش تک ثبت دوچندان نیز می‌شود.

۴-۲- نحوه‌ی دسته‌بندی

در مرحله دسته‌بندی، مدل‌ها در حالت‌های مختلف زیر آموزش داده شده‌اند:

- 6C: دسته‌بندی رده گروهی در شش حالت برای کل داده‌ها
 - 72X: دسته‌بندی بر مبنای همه‌ی محرک‌های تصویری استفاده شده در آزمایش‌ها در مقابل یکدیگر
 - 12F: دسته‌بندی دوازده سطحی، داخل گروهی و برای همه‌ی کلاس‌های موجود در گروه صورت انسان می‌باشد.
 - 12O: دسته‌بندی دوازده سطحی، داخل گروهی و برای همه‌ی کلاس‌های موجود در گروه اشیاء می‌باشد.
 - 2FO: دسته‌بندی دو سطحی برای گروه صورت انسان در مقابل گروه اشیاء انجام می‌پذیرد.
 - 2AI: دسته‌بندی دو سطحی برای دو دسته‌ی جانداران در مقابل دسته‌ی بی‌جان‌ها انجام می‌پذیرد.
- کلیه‌ی نتایج روش‌های بررسی شده در جدول ۱ و در قالب مقایسه‌ای به صورت دو مرحله‌ای آورده شده است:

کنند [۲۳]. درختان تصمیم‌گیری بخش‌های ساختاری مدل جنگل تصادفی هستند. در جنگل تصادفی هر درخت جداگانه پیش‌بینی کلاس را انجام می‌دهد و کلاس با بیشترین آرا پیش‌بینی نهایی مدل می‌باشد. در حقیقت تعداد زیادی از مدل‌ها (درختان) نسبتاً ناهمپوشان که به عنوان گروهی فعالیت می‌کنند پیش‌بینی‌هایی گروهی را تولید می‌کنند که در شرایط مشابه دقیق‌تر از هر پیش‌بینی فردی است. دلیل این نتیجه‌ی شگفت‌انگیز این است که در حالی که ممکن است برخی از درختان اشتباه باشند، بسیاری از درختان دیگر درست خواهند بود، بنابراین به عنوان گروهی درختان قادر به حرکت در جهت صحیح هستند.

- **XGB**: تقویت شیب حداکثری [۲۴] یکی دیگر از روش‌های انجام وظایف یادگیری ماشینی نظارت شده مانند طبقه‌بندی و رگرسیون است. در ماشین‌های تقویت‌کننده شیب، شبیه جنگل‌های تصادفی، رایج‌ترین نوع مدل مبنایی مورد استفاده درخت تصمیم‌گیری است. تقویت‌کننده شیب یک یادگیرنده ترکیبی مبتنی بر درخت تصمیم‌گیری است. این بدان معناست که مدل نهایی را بر اساس مجموعه‌ای از مدل‌های تکی ایجاد می‌کند. قدرت پیش‌بینی این مدل‌های تکی ضعیف و مستعد ابتلا به بیش‌برآزش است، اما ترکیب بسیاری از این مدل‌های ضعیف در یک گروه منجر به یک نتیجه کلی بسیار بهتر خواهد شد. یک حالت خاصی از تقویت که در آن خطاها با الگوریتم کاهش شیب به حداقل می‌رسند. پیاده‌سازی این روش نام‌های مختلفی دارد. به طور معمول با ماشین‌های تقویت‌کننده شیب (GBM) و XGBoost روبرو می‌شویم. XGBoost با اجرای خاصی از روش تقویت‌کننده شیب، از محبوبیت خاصی برخوردار است و از تقریب‌های دقیق‌تری برای یافتن بهترین مدل درخت استفاده می‌کند. با تعدادی از ترفندهای عالی که به کار می‌برد به ویژه با داده‌های ساخت‌یافته، فوق‌العاده موفق عمل می‌کند. بسیاری از مسائل دنیای واقعی چالش داده‌های گم شده دارند که حاوی اطلاعات ارزشمندی در مورد هدف بوده‌اند. در این حالت XGBoost روش خاصی برای مدیریت و مقابله با مقادیر گم‌شده دارد؛ در حین آموزش، با یادگیری وظیفه‌های فرعی، انتساب و انتقال اطلاعات برای هر ویژگی را انجام می‌دهد.

۴-۱- آزمایش‌ها و نتایج

۴-۱-۱- مجموعه‌داده

در سال‌های اخیر چندین مجموعه داده با تعداد بی‌سابقه‌ای از تصاویر و حاشیه‌نوشتار^۴ها برای درک صحنه‌های بصری و به ویژه وظیفه‌های هدف-گرا مانند بخش‌بندی و یا دسته‌بندی اشیاء تولید شده است [۲۵-۲۸]، ولی برای موضوع این تحقیق مجموعه‌داده‌ی اخیراً منتشر

¹ Missing Value

² Annotation

³ Stimuli

تصویری در مقابل یکدیگر (72X) $1,39 = 1 \div 72$ درصد است. برای دسته‌بندی دوازده سطحی داخل گروهی (12F یا 12O) $8,33 = 1 \div 12$ درصد و برای دسته‌بندی دو سطحی دو کلاسه (2FO یا 2AI) $50,0 = 1 \div 2$ درصد است.

تمام نتایج مورد بررسی در این مطالعه در جدول ۱ با جزئیات دقت و زمان صرف شده برای هر حالت دسته‌بندی، در دو مرحله یکبار با رویکرد پیشنهادی و یکبار بدون آن نشان داده شده است. برای مقایسه مقدار بهبود، بهترین نتایج بررسی‌های انجام شده در کلاس‌های مختلف، در مقایسه با نتایج [۱۱] آورده شده است. هر دو بررسی در جدول ۲ دقت بالاتر از سطح شانس را نشان می‌دهد و نتایج بررسی‌های ما برهمه‌ی نتایج کار [۱۱] برتری دارد. تفاوت خاصی در محیط پیاده‌سازی وجود ندارد اما الگوریتم‌های [۱۱] در Matlab^۲ و کلیه‌ی الگوریتم‌های ما در پایتون^۳ پیاده‌سازی شده است. به جز مدل XGB که از کتابخانه XGBoost استفاده شده است [۳۱] همه مدل‌های دیگر از کتابخانه Scikit [۳۳] در پایتون به کارگیری شده‌اند.

- بدون IFS: بردارهای ویژگی با الگوریتم‌های کاهش ابعاد PCA و LDA کوچک‌تر شده و سپس طبقه‌بندی‌کننده‌ها بردارهای کاهش یافته را آموزش می‌بینند.
- با IFS: قبل از کاهش بردارهای ویژگی با دو الگوریتم ذکر شده در حالت قبل، از کاهش ابعاد اولیه با روش IFS که در بخش روش پیشنهادی توضیح داده شده است، استفاده می‌کنیم.

۴-۳- بهبود کارایی

با انتخاب مناسب الگوریتم بهینه‌سازی خودکار، می‌توان عملکرد مدل را بهتر کرد. اما این فرایند وقت گیر است و بستگی به حجم فضای ویژگی‌های یادگیری، زیرساخت محاسباتی یا بودجه محاسباتی و غیره دارد. اگر مدل بسیار ساده باشد و بتواند سریع اجرا شود، یک جستجوی تصادفی ممکن است کافی باشد، اما برای مدل‌های در مقیاس بزرگ به اندازه کافی کارآمد نخواهد بود. در سیگنال‌های با ابعاد ویژگی بالا و کلاس‌های زیاد، بهینه‌سازی خودکار پارامترها بسیار پیچیده و وقت گیر هستند. در این تحقیق نیز تعدادی از الگوریتم‌های یادگیری در حالت‌های گوناگون مورد بررسی قرار گرفتند و بهینه‌سازی خودکار پارامترها زمان زیادی را می‌طلبید. از این‌رو ما با دانش پیرامونی فضای مساله، بهینه‌سازی پارامترها را انجام دادیم. برای نمونه در طبقه‌بندی دوگروهه، جنگل تصادفی (RF) با بردارهای ویژگی کاهش یافته از طریق IFS و PCA، در کلیه مجموعه داده‌های شرکت کنندگان به روش اعتبارسنجی متقابل ۱۰-بخشی^۱، با ۱۰۰۰ برآوردکننده (تعداد درختان تصمیم‌گیری در جنگل [۳۳]) اعمال شد و دقت در ۴۱۶۷ ثانیه ۶۲،۹۸ درصد بود و با ۱۰ برآوردکننده دقت ۶۰،۷۱ درصد در ۲۸۴ ثانیه بدست می‌آمد. نتیجه‌ی بهتر ۶۴،۶۷ درصد با ۱۰۰ برآوردگر در ۶۳۲ ثانیه حاصل گردید. بر خلاف انتظار نتیجه‌ی حاصل شده با ۱۰۰۰ برآوردکننده ضعیف و زمان بر بود! به عنوان یک روش ساده، در الگوریتم‌های مبتنی بر درخت تصمیم می‌توانیم تعداد برآوردکننده‌ها را تا جاییکه مشکل بیش‌برازش پیش نیاید به اندازه کافی بزرگتر از تعداد کلاس‌ها در نظر بگیریم. در [۱۱] تعداد دقیق مؤلفه‌های اصلی PCA در طی اعتبارسنجی متقاطع تودرتو مشخص شده ولی ما آن را مقدار ثابت و بهینه ۲۰۰ در نظر گرفتیم. همچنین در کلیه‌ی نتایج پارامتر مؤلفه کاهش ابعاد LDA برابر تعداد کلاس‌های منهای یک بهترین نتیجه را داشته است.

۴-۴- نتایج

برای ارزیابی عملکرد مدل‌های یادگیری ماشین روی سیگنال‌های EEG تک ثبت، مدل‌های خود را در برابر نتایج [۱۱] بررسی و محک می‌زنیم. به همان شیوه‌ی [۱۱]، تمام مدل‌های دسته‌بندی برای مجموعه داده‌های هر شرکت‌کننده جداگانه به صورت اعتبارسنجی متقاطع ۱۰-بخشی انجام می‌شود [۳۱]. دقت نهایی، میانگین دقت نتایج دسته‌بندی داده‌های همه‌ی شرکت‌کننده‌ها می‌باشد. سطح شانس در تجزیه و تحلیل آماری، برای دسته‌بندی گروهی شش سطحی (6C) $16,67 = 1 \div 6$ درصد است. برای دسته‌بندی همه‌ی محرک‌های

² <https://www.mathworks.com>

³ <https://www.python.org>

¹ 10-Fold cross-validation

بردار ویژگی استفاده می‌کند. تمام نتایج دسته‌بندی شده در دو حالت ارائه شده است: با رویکرد پیشنهادی IFS و بدون آن. برای کلیه ی نتایج مقدار اول دقت و مقدار بعد از علامت (/) زمان محاسبات بر حسب ثانیه می باشد.

جدول ۱: به صورت خلاصه نتایج دقت و زمان محاسبه‌ی میانگین دسته‌بندی‌های داده‌های ده شرکت کننده با اعتبارسنجی متقاطع ۱۰-بخشی برای همه‌ی حالت‌ها نشان داده شده است. دسته‌بندی داده‌ها از همه الکترودها و با استفاده از تمام نقاط زمانی از بازه‌ی ۰-۴۹۶ میلی‌ثانیه در کنار هم در ایجاد

Classification Type	Methods	Method (Accuracy / Time)				
		SVM	LDA	OVA(LDA)	RF	XGB
	LDA, PCA	LDA, PCA	LDA, PCA	LDA, PCA	LDA, PCA	
6 C - IFS 6 C + IFS		25.50 / 5978, 40.35 / 1488 42.29 / 697, 47.12 / 1094	27.23 / 7599, 40.09 / 706 42.38 / 715, 46.63 / 272	27.80 / 7296, 39.46 / 836 42.04 / 690, 46.33 / 400	26.70 / 7264, 33.34 / 6080 41.85 / 1508, 38.59 / 5631	41.97 / 1385, 41.88 / 1173 41.64 / 768, 31.32 / 470
72 X - IFS 72 X + IFS		02.82 / 6719, 11.94 / 2092 17.94 / 1132, 20.63 / 1499	12.72 / 641, 13.10 / 656 22.27 / 260, 22.33 / 261	03.90 / 8137, 11.51 / 2556 12.96 / 1426, 20.06 / 2287	03.11 / 7386, 03.76 / 5283 16.86 / 2002, 16.12 / 2100	03.20 / 3645, 03.11 / 3856 09.42 / 1271, 06.09 / 1409
12 F - IFS 12 F + IFS		20.90 / 216, 19.68 / 134 21.85 / 106, 24.46 / 85	20.95 / 214, 20.95 / 105 22.28 / 106, 25.85 / 60	20.91 / 216, 20.80 / 180 20.95 / 109, 26.25 / 135	20.11 / 453, 16.87 / 1049 20.63 / 307, 20.10 / 980	16.63 / 315, 12.02 / 273 16.89 / 185, 14.90 / 224
12 O - IFS 12 O + IFS		34.65 / 215, 32.23 / 132 34.67 / 106, 43.05 / 84	34.60 / 215, 34.56 / 104 35.32 / 105, 45.62 / 59	33.26 / 216, 32.85 / 180 33.21 / 108, 43.84 / 135	33.01 / 460, 26.08 / 848 31.84 / 336, 34.27 / 782	28.42 / 299, 16.06 / 270 25.09 / 183, 20.38 / 223
2 FO - IFS 2 FO + IFS		79.53 / 682, 81.96 / 224 51.26 / 293, 84.93 / 149	79.58 / 668, 81.99 / 171 55.17 / 300, 84.87 / 94	79.58 / 652, 82.01 / 161 55.17 / 259, 84.93 / 92	79.21 / 844, 77.05 / 1159 55.18 / 444, 79.64 / 1063	79.36 / 691, 70.13 / 249 55.22 / 319, 72.76 / 168
2 AI - IFS 2 AI + IFS		62.59 / 6514, 70.44 / 1298 71.46 / 769, 73.74 / 835	63.46 / 6197, 69.49 / 714 71.00 / 658, 72.76 / 266	63.30 / 721, 69.54 / 695 71.00 / 674, 72.75 / 265	63.34 / 6908, 60.89 / 1049 69.75 / 693, 64.67 / 632	63.35 / 6187, 62.40 / 723 71.10 / 738, 65.11 / 320

موثر در دسته‌بندی‌ها برخوردار است. همپوشانی خطی^۱ بین کلاس‌ها سرعت و دقت LDA را در بیشتر بررسی‌های ما کم کرده است. به عنوان مطالعات آتی می‌توان رویکرد پیشنهادی را روی مجموعه داده‌های دیگر و حتی به صورت ترکیبی با سایر روش‌های موجود بررسی کرد.

سپاسگزاری

از زحمات‌های دکتر بلر کانیشیرو^۲ از دانشگاه استنفورد برای راهنمایی در ایجاد شرایط آزمون یکسان برای مطابقت نتیجه الگوریتم پایه‌ی ما با نتیجه‌ی آثار مرتبط در [۱۱] سپاسگزاری می‌کنیم؛ همچنین از او و همکارانش برای انتشار مجموعه‌داده‌ی مفید EEG [۲۹] در دانشگاه استنفورد تشکر می‌کنیم.

مراجع

- [1] Biasucci A., Franceschiello B., Murray MM., Electroencephalography. Curr Biol. (2019) 29:R80-R85.
- [2] Thorpe S., Fize D., & Marlot C., Speed of processing in the human visual system. Nature 381, 520-522 (1996).
- [3] Sofer, I., Crouzet, S. M. & Serre, T. Explaining the timing of natural scene understanding with a computational model of perceptual categorization. PLoS Comput. Biol. 11 (2015).
- [4] Francois C., Deep Learning with Python, Visualizing what convnets learn. 1st. Greenwich, CT, USA: Manning Publications Co., 2017. isbn: 1617294438, 9781617294433.
- [5] Oana D.E., Anca M.L., Comparison of classifiers and statistical analysis for EEG signals used in brain computer interface motor task paradigm, Int. J. Adv. Res. Artif. Intell. (IJARAI) 4 (2015) 8-12.
- [6] Jochumsen M., Niazi I.K., Dremstrup K., Nlandu K.E., Detecting and classifying three different hand movement types through electroencephalography recordings for neurorehabilitation, Med. Biol. Eng. Comput. 53 (2015) 1-11.

¹ Column linearity

² Dr. Blair Kaneshiro

جدول ۲: بهترین نتایج بررسی‌های انجام شده در حالت‌های مختلف در مقایسه با نتایج [۱۱] آورده شده است.

Best Results	Category / Class					
	6 C	72 X	12 F	12 O	2 FO	2 AI
[11]	40.68	14.46	18.30	28.87 [⊗]	81.06	-----
our	47.12	22.33	26.25	45.62	84.93	73.74

علامت [⊗] حذف داده‌های یکی از شرکت‌کننده‌ها در کلاس 12O در نتایج [۱۱] را نشان می‌دهد.

۵- جمع بندی و نتیجه گیری

ناحیه بارگیری و فعال مغز انسان برای گروه‌هایی از کارها در قسمت‌های خاصی از مغز می‌باشد و استفاده‌ی هوشمندانه از داده‌ی الکترودهای موجود در این نواحی می‌تواند منجر به افزایش دقت طبقه‌بندی کننده‌ها شود. بررسی‌های ما نشان می‌دهد که چگونه می‌توانیم با انتخاب الگوریتم‌های مناسب و بهینه سازی عملکرد آنها و ایجاد فضای ویژگی‌های بامعنی (IFS) به کمک دانش فضای مساله، توانایی مدل‌های خود را افزایش دهیم.

خلاصه‌ی بهترین نتایج بررسی‌های ما که در جدول ۲ آمده، بر همه نتایج کار [۱۱] برتری دارد و دلیلی بر کارایی مناسب رویکرد پیشنهادی می‌باشد. با توجه به حالت‌های متنوع بررسی مجموعه‌ی داده‌ی مورد مطالعه و اکتساب نتایج برتر در همه‌ی آنها، شاید بتوان نتیجه گرفت که ما با این رویکرد پیشنهادی می‌توانیم کارایی تقریباً بیشتر طبقه‌بندی کننده‌های سیگنال‌های EEG را در کارهای مربوطه و خاص افزایش دهیم.

همچنین به عنوان یک نتیجه‌ی جانبی، نتایج بررسی‌های ما نشان می‌دهد که در حالت‌های مختلف مساله مورد مطالعه‌ی این مقاله، PCA نسبت به LDA از سرعت و اطلاع‌رسانی بهتری در انتخاب ویژگی‌های

- [29] Kaneshiro B., Guimaraes M. P., Kim H.-S., Norcia A. M. and Suppes P., EEG data analyzed in "A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification". Stanford Digital Repository. 2015 Available at: <http://purl.stanford.edu/bq914sc3730> KDB
- [30] [Ftp://ftp.egi.com/pub/support/Documents/net_layouts/hcgsn_128.pdf](ftp://ftp.egi.com/pub/support/Documents/net_layouts/hcgsn_128.pdf).
- [31] Hastie T, Tibshirani R, Friedman J., The Elements of Statistical Learning. 2nd ed. Springer; 2009.
- [32] "Python Package Index Pypi: xgboost". Retrieved 2016-08-01.
- [33] Pedregosa et al., Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.
- [7] Naeem M., Brunner C., Leeb R., Graimann B., Pfurtscheller G., Separability of four-class motor imagery data using independent components analysis, *J. Neural Eng.* 13 (2006) 208–216.
- [8] DiCarlo J, Zoccolan D, Rust N. How does the brain solve visual object recognition? *Neuron.* 2012;73(3):415–434.
- [9] Ghaemi A., Rashedi E., Mohammad Pourrahimi A., Kamandar M. and Rahdari F., Automatic channel selection in EEG signals for classification of left or right hand movement in Brain Computer Interfaces using improved binary gravitation search algorithm. *Biomedical Signal Processing and Control* 33 (2017) 109–118
- [10] Alomari M., Samaha A., AlKamha K., Automated classification of L/R hand movement EEG signals using advanced feature extraction and machine learning, *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* 4 (2013) 207–212.
- [11] Kaneshiro B, Perreau Guimaraes M, Kim H-S, Norcia AM, Suppes P., A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. *PLoS ONE* (2015); 10(8). e0135697. doi:10.1371/journal.pone.0135697
- [12] Golub GH, Reinsch C. Singular value decomposition and least squares solutions. *Numerische Mathematik.* 1970; 14(5):403–420.
- [13] Cohen J., Cohen P., West S. G. and Aiken L. S., *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences* 3rd ed. (2003). Taylor & Francis Group.
- [14] Green, S.B. Salkind, N. J. & Akey, T. M.. *Using SPSS for Windows and Macintosh: Analyzing and understanding data.* New Jersey: Prentice Hall (2008).
- [15] Lan T., Erdogmus D., Adami A., Pavel M. and Mathan S., Salient EEG channel selection in brain computer interfaces by mutual information maximization, *IEEE Eng. Med. Biol. 27th Ann. Conf. IEEE* (2005).
- [16] Schröder M.S., Navin L.T., Hinterberger T., Bogdan M., Hill N.J., Birbaumer N. and et al., Robust EEG channel selection across subjects for brain-computer interfaces, *EURASIP J. Appl. Signal Process.* (2005) 3103–3112.
- [17] Strang G. *Introduction to Linear Algebra.* 3rd ed. Wellesley-Cambridge Press; (2003)
- [18] Zhou H., Xuean M., Chaogui M., Liu J. and Jones P., "Classification of single-trial EEG signals," *The 3rd IEEE International Seminar on Medical Applications of Signal Processing* (Ref. No. 2005-1119), London, UK, 2005, pp. 135-140.
- [19] Vapnik V., "The support vector method of function estimation," in J.A.K. Suykens and J. Vandewalle (Eds) *Nonlinear Modeling: Advanced Black-Box Techniques*, Kluwer Academic Publishers, Boston, 1998: 55–85.
- [20] Bishop, Christopher M., *Pattern Recognition and Machine Learning.* Springer (2006).
- [21] Ho T. K., *Random Decision Forests.* Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.
- [22] Ho T. K., "The Random Subspace Method for Constructing Decision Forests". *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 20 (8) 1998: 832–844.
- [23] Opitz D., Maclin R., "Popular ensemble methods: An empirical study". *Journal of Artificial Intelligence Research.* 11 1999: 169–198.
- [24] Tianqi C. and Carlos G., XGBoost: A Scalable Tree Boosting System. In *22nd SIGKDD Conference on Knowledge Discovery and Data Mining*, 2016.
- [25] Torralba A., Fergus R., and Freeman W. T., 80 million tiny images: A large data set for nonparametric object and scene recognition. *TPAMI*, 30(11):1958–1970, 2008.
- [26] Deng J., Dong W., Socher R., Li L.-J., Li K. and Fei-Fei L., ImageNet: A large-scale hierarchical image database. In *CVPR*, pages 248–255, 2009.
- [27] Xiao J., Hays J., Ehinger K. A., Oliva A., and Torralba A., Sun database: Large-scale scene recognition from abbey to zoo. In *CVPR*, pages 3485–3492. IEEE, 2010.
- [28] Lin T.-Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P. and Zitnick C. L., Microsoft COCO: Common objects in context. In *ECCV*, pages 740–755, 2014.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند



۲۷ الی ۲۸ آذرماه ۱۳۹۸

ادغام هرمی تصاویر با نورگیری چندگانه با ویژگی‌های مبتنی بر رنگ و توزیع شدت روشنایی در دنباله تصاویر ورودی

عیلرضا اسدی، مهدی ازوجی

^۱ دانشجوی کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل

alirezaasadi@nit.ac.ir

^۲ استادیار دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل

m.ezoji@nit.ac.ir

چکیده

مهم‌ترین بخش الگوریتم‌های ادغام تصاویر با نورگیری چندگانه، انتخاب معیارها و ویژگی‌های کارآمد جهت استخراج نقشه‌های مناسب برای وزن‌دهی تصاویر ورودی است. در این مقاله، به کارگیری معیار رنگ دنباله تصاویر ورودی با نورگیری چندگانه جهت استخراج چنین ویژگی‌هایی پیشنهاد می‌شود. ایده پایه کار این است که تغییرات اندازه نورگیری بر ویژگی‌هایی چون روشنی و رنگ یک پیکسل اثر می‌گذارد. به همین منظور نقشه ویژگی استخراج شده مبتنی بر رنگ در کنار نقشه‌های وزنی مبتنی بر شدت روشنایی پیکسل‌ها در ساختاری هرمی با هم ترکیب شده و برای ساخت تصویری با گستره دینامیکی بالا استفاده شده است. بررسی و ارزیابی کمی و دیداری به ازای تصاویر ورودی متنوع، نشان از برتری روش پیشنهادی نسبت به روش‌های به‌روز دیگر دارد.

کلمات کلیدی

ادغام تصاویر، تصاویر با نورگیری چندگانه، ویژگی رنگ، نقشه‌های وزنی، ساختار هرمی

۱- مقدمه

های مبتنی بر پیچ. در روش‌های مبتنی بر پیکسل، تصویر خروجی با وزن‌دهی پیکسل‌های خوب از تصاویر ورودی ایجاد می‌شود و در روش‌های مبتنی بر پیچ، هر پیچ از تصویر خروجی بر اساس ویژگی‌های استخراج شده از پیچ‌های تصاویر ورودی ساخته می‌شود. در روش‌های مبتنی بر پیکسل، نقشه وزنی با ابعادی برابر با تصویر ورودی بیانگر اهمیت پیکسل‌های تصویر ورودی در ایجاد تصویر خروجی است. از این رو، پیدا کردن نقشه‌های وزنی مناسب اهمیت بسیار زیادی خواهد داشت. در پژوهش‌های این حوزه، رویکردهای گوناگونی برای تعریف نقشه‌های وزنی کارآمد معرفی شدند. برای نمونه سه معیار کیفیت به نام‌های جذب نور خوب، کنتراست و اشباع زیاد در [1] به منظور تعیین وزن هر پیکسل در ساخت تصویر خروجی مبتنی بر هرم لاپلاسی استفاده شدند [3]. در [4] مقدار گرادیان تصویر، برای تنظیم وزن‌ها و جهت گرادیان به منظور شناسایی حرکات در صحنه استفاده شدند.

در [5] الگوریتم ادغام تصاویر با نورگیری چندگانه مبتنی بر فیلترهای هدایت شده در لایه‌های جزئیات و کلیات تصویر طی میانگین‌گیری وزنی انجام شد. در [6] و [7] ادغام تک‌مقیاسی

ابزارهای تصویربرداری دیجیتال نسبت به صحنه‌های طبیعی که جزئیات روشن و تاریک را به طور هم‌زمان دربردارند، محدوده دینامیکی کمتری را پوشش می‌دهند و به همین دلیل تصاویر ثبت شده از این صحنه‌ها فاقد همه جزئیات روشن و تاریک صحنه است [1]. برای رفع این مشکل دو روش کلی وجود دارد. روش نخست تکنیک تصویربرداری با محدوده دینامیکی بالا است که در آن جهت ایجاد تصویر نهایی باید منحنی پاسخ دوربین معکوس شود که به محاسبات پیچیده نیاز دارد و از سوی دیگر باید محدوده دینامیکی تصویر خروجی ایجاد شده به منظور اهداف نمایشی فشرده شود. روش دیگر، تکنیک ادغام تصاویر با نورگیری چندگانه است که در آن نیازی به معکوس کردن تابع پاسخ دوربین نبوده و تصویر خروجی حاصل از آن به صورت مستقیم بر روی ابزارهای نمایشی متداول قابل نمایش است [2].

الگوریتم‌های ادغام تصاویر با نورگیری چندگانه را می‌توان به دو دسته تقسیم نمود که عبارت‌اند از روش‌های مبتنی بر پیکسل و روش-

m_n میانگین شدت روشنایی n امین تصویر ورودی است و مقدار σ_n به صورت تطبیقی از روی تصاویر ورودی مطابق با رابطه (۳) محاسبه می‌شود:

$$\sigma_n = \begin{cases} 1.5(m_{n+1} - m_n) & n = 1 \\ 0.75(m_{n+1} - m_{n-1}) & 1 < n < N \\ 1.5(m_n - m_{n-1}) & n = N \end{cases} \quad (3)$$

با افزایش اختلاف میانگین شدت روشنایی تصاویر بعد و قبل از تصویر ورودی، مقدار σ_n افزایش یافته و موجب می‌شود وزن سراسری بیشتری به آن تصویر تخصیص یابد.

در ادامه، [15] برای وزندهی به نواحی کم‌جزئیاتی که نورگیری خوبی دارند ولی از دید گرادین پنهان می‌مانند، از عکس گرادین منحنی هیستوگرام تجمعی تصاویر ورودی بهره گرفته است:

$$W_{2,n}(x,y) = \frac{Grad_n(I_n(x,y))^{-1}}{\sum_{n=1}^N Grad_n(I_n(x,y))^{-1} + \epsilon} \quad (4)$$

در رابطه (۴)، ϵ یک مقدار بسیار کوچک به منظور جلوگیری از تقسیم بر صفر و $Grad_n(I_n(x,y))$ بیانگر گرادین منحنی هیستوگرام تجمعی در سطح خاکستری پیکسل $I_n(x,y)$ است.

۲-۲- تابع وزنی پیشنهادی مبتنی بر معیار رنگ تصویر

بررسی‌ها نشان می‌دهد که تغییرات سطح نورگیری بر روی ویژگی‌هایی چون جزئیات روشنی محلی، تضاد و رنگ یک پیکسل اثر می‌گذارد [23]. از این رو، در مقاله حاضر تلاش شده است تا با سنجش این ویژگی‌ها به ترتیب مبتنی معیارهایی مانند تعداد عبور از صفر (ZC) [20]، لاپلاسین جمع اصلاح شده (SML) [24] و رنگ $(\hat{M}^{(3)})$ [22] نقشه ویژگی مناسبی را ارائه دهیم. از این رو، به مرور کوتاه این معیارها می‌پردازیم.

فرض کنید X تصویری با ابعاد $M \times N$ باشد. اگر $X_r(m,n)$ ، $X_g(m,n)$ ، $X_b(m,n)$ و $X_y(m,n)$ به ترتیب بیانگر بخش‌های قرمز، سبز، آبی و روشنایی پیکسل $X(m,n)$ ، و $B_{m,n}$ یک بلوک با مرکز (m,n) و ابعاد 5×5 باشد:

- $SML_X(m,n)$ تعریف می‌شود:

$$SML_X(m,n) = \sum_{(k,l) \in B_{m,n}} \{ |2X_y(k,l) - X_y(k-1,l) - X_y(k+1,l)| + |2X_y(k,l) - X_y(k,l-1) - X_y(k,l+1)| \} \quad (5)$$

- ویژگی عبور از صفر با توجه به رابطه (۶) قابل محاسبه است:

$$ZC_X(x,y) = ZC_V(x,y) + ZC_H(x,y) \quad (6)$$

که $ZC_V(x,y)$ ، تعداد عبور از صفر عمودی $X_y(k,l)$ و $ZC_H(x,y)$ ، تعداد عبور از صفر افقی $X_y(k,l)$ در $B_{m,n}$ است.

- معیار رنگ به صورت زیر رابطه (۷) محاسبه می‌شود:

$$\hat{M}_x^3(m,n) = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} + 0.3 \sqrt{\mu_{rg}^2 + \mu_{yb}^2} \quad (7)$$

تصویرها با تعیین بهینه پارامترهای وزندهی طی حل مسأله‌های بهینه‌سازی متناظر پیشنهاد شد. در [8] نقشه‌های وزنی با استفاده از انرژی کارآمد محلی و همبستگی بین لایه‌های هرم‌های لاپلاسین محاسبه شد. برای دستیابی به نقشه‌های مناسب وزنی، روش‌های گوناگون دیگری مبتنی بر محاسبه مقدار گرادین [9]، تخمین نور [10]، ساختار تصویر در فضای گرادین [11] و یا معیارهایی همچون کنتراست محلی و ثبات رنگ [12]، فیلترهای حفظ لبه [13] مانند فیلتر بازگشتی [14] و شدت روشنایی پیکسل‌ها و گرادین منحنی هیستوگرام تجمعی تصاویر ورودی [15] هم پیشنهاد شده است. در [16] ادغام تصاویر با استفاده از تجزیه و تحلیل تنک و طراحی یک دیکشنری مبتنی بر آن انجام می‌شود. در [17] رویکرد ادغام بر اساس تصحیح رنگ محلی پیشنهاد شده و وزن نهایی هر تصویر بر اساس محاسبه اشباع حاصل می‌گردد.

با بررسی ارتباط بین ادراک بینایی انسان و کیفیت تصاویر ویژگی‌هایی چون روشنایی سراسری، تضاد، میزان وضوح تصویر، جزئیات، طبیعی بودن و رنگ جهت طراحی معیارهای ارزیابی کیفیت تصاویر معرفی می‌شوند [18]، [19]، [20]، [21] و [22]. از سوی دیگر تغییرات سطح نورگیری تصاویر بر روی این ویژگی‌هایی و به ویژه جزئیات نور محلی، تضاد و رنگ یک پیکسل اثر می‌گذارد [23].

در ادامه این مقاله، در بخش دوم، به معرفی الگوریتم پیشنهادی و گام‌های استخراج نقشه‌های وزنی و شیوه ادغام پرداخته می‌شود. در بخش سوم، ارزیابی کمی و کیفی الگوریتم پیشنهادی در مقایسه با روش‌های دیگر انجام می‌شود. جمع‌بندی کار نیز در بخش پایانی ارائه خواهد شد.

۲- الگوریتم پیشنهادی

در روش پیشنهادی تصویر خروجی I_F مطابق با رابطه (۱)، مجموع وزندهی شده تصاویر ورودی یعنی I_n ها است:

$$I_F(x,y) = \sum_{n=1}^N W_n(x,y) I_n(x,y) \quad (1)$$

که N تعداد تصاویر ورودی، $I_n(x,y)$ شدت روشنایی پیکسل (x,y) ، در n امین تصویر ورودی و $W_n(x,y)$ وزن این پیکسل است. مهم‌ترین بخش یک الگوریتم ادغام تصاویر با نورگیری چندگانه، چگونگی طراحی $W_n(x,y)$ است چرا که مقدار آن باید با میزان نورگیری و نیز جزئیات در تصویر متناسب باشد.

در این مقاله، تلاش می‌شود تا نقشه وزنی مبتنی بر معیارهای جزئیات، تضاد و رنگ استخراج شده و با ماتریس‌های وزندهی ارائه شده در [15] که مبتنی بر شدت روشنایی پیکسل‌ها و گرادین منحنی هیستوگرام تجمعی تصاویر ورودی است، ادغام گردد. از این رو، به مرور کوتاه این وزندهی‌های پایه می‌پردازیم.

۲-۱- وزندهی مبتنی بر شدت روشنایی [15]

در [15] سنجش میزان نورگیری یک پیکسل نسبت به پیکسل‌های دیگر تصویر و نیز تصاویر دیگر دنباله، چنین انجام شده است:

$$W_{1,n}(x,y) = \exp\left(-\frac{(I_n(x,y) - (1-m_n))^2}{2\sigma_n^2}\right) \quad (2)$$

در روش پیشنهادی، وزن نهایی برای هر تصویر ورودی از یک دنباله با ادغام سه نقشه وزنی و نرمالیزه کردن آن به صورت رابطه (۱۱) محاسبه خواهد شد:

$$W_n(x, y) = \frac{W_{1,n}(x, y) \times W_{2,n}(x, y) \times W_{3,n}(x, y)}{\sum_{n=1}^N W_{1,n}(x, y) \times W_{2,n}(x, y) \times W_{3,n}(x, y) + \epsilon} \quad (11)$$

۲-۳- چارچوب پیشنهادی ادغام تصاویر

پس از محاسبه وزن نهایی مطابق با رابطه (۱۱)، عمل ادغام تصاویر را می‌توان با توجه به رابطه (۱) انجام داد. این رویکرد نتیجه خوبی نداشته و در تصویر خروجی حفره، بخش‌های غیرطبیعی و هاله‌هایی دیده می‌شود. در مقاله حاضر، برای رفع این مشکل همانند [1] از ادغام نقشه‌های وزنی در ساختار هرمی استفاده می‌شود. یعنی، رابطه (۱) در روندی چندمقیاسی با استفاده از تجزیه و تحلیل هرم تصاویر اعمال می‌شود.

اگر سطح l ام هرم لاپلاسین تصویر I به صورت $\{I\}_l^L$ و سطح l ام هرم گاوسی ماتریس W به صورت $\{W\}_l^G$ بیان شود، ادغام ضرایب در سطوح مختلف هرم به صورت رابطه (۱۲) انجام می‌شود:

$$L\{R\}_{ij}^l = \sum_{k=1}^N G\{W_n\}_{ij,k}^l L\{I\}_{ij,k}^l \quad (12)$$

سرانجام با ادغام هرم $\{R\}_l^L$ که حاصل مجموع حاصل ضرب سطوح متناظر از هرم گاوسی نقشه وزن نهایی و هرم لاپلاسین تصاویر ورودی است، تصویر خروجی به دست می‌آید که کیفیت مطلوبی دارد.

۳- آزمایش‌ها و ارزیابی

برای بررسی و ارزیابی کارکرد روش پیشنهادی و مقایسه آن با روش‌های دیگر ادغام تصاویر با نورگیری چندگانه از دنباله تصاویر مرجع [15] و [25] استفاده می‌کنیم. پایگاه داده [15] شامل ده دنباله است که تعداد تصاویر هر دنباله متغیر است و در مجموع هفتاد تصویر ورودی را شامل می‌شود. مرجع [25] افزون بر دنباله‌های پایگاه داده مرجع [15]، هفت دنباله دیگر هم دارد که در مجموع شامل صد و بیست تصویر ورودی است و تنوع بیشتری نسبت به دنباله‌های مرجع [15] دارد و از آن به منظور ارزیابی جامع‌تر روش پیشنهادی استفاده می‌شود. به طور کلی دنباله‌های مورد استفاده حالت‌های متنوعی از صحنه‌های طبیعی با محدوده دینامیکی بالا را پوشش می‌دهد مانند صحنه‌های شامل فضای داخل اتاق، صحنه‌های طبیعی با حضور نور زیاد، صحنه‌های طبیعی با حضور ساخته‌های دست انسان. از طرف دیگر تعداد تصاویر هر دنباله متغیر است تا کارکرد الگوریتم به ازای دنباله‌های با تعداد تصاویر ورودی متفاوت بررسی شود.

۳-۱- معیار ارزیابی MEF-SSIM

معیار ویژه ارزیابی کارکرد الگوریتم‌های ادغام تصاویر با نورگیری چندگانه MEF-SSIM [25] نام دارد که سال ۲۰۱۵ معرفی شد. این

که μ_{rg} ، μ_{yb} ، σ_{rg}^2 و σ_{yb}^2 به ترتیب مقدار میانگین $X_{rg}(k, l)$ ، $X_{yb}(k, l)$ ، واریانس $X_{rg}(k, l)$ و واریانس $X_{yb}(k, l)$ برای $(k, l) \in B_{m,n}$ را نشان می‌دهند. $X_{rg}(k, l)$ و $X_{yb}(k, l)$ به ترتیب طبق روابط (۸) و (۹) تعریف می‌شوند:

$$X_{rg}(k, l) = X_r(k, l) - X_g(k, l) \quad (8)$$

$$X_{yb}(k, l) = 1/2[X_r(k, l) + X_g(k, l)] - X_b(k, l) \quad (9)$$

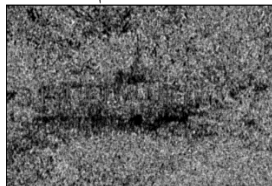
نمونه‌ای از معیارهای استخراجی فوق، برای تصویری آزمون در شکل ۱، نشان داده شده است. از شکل ۱ پیداست که در نقشه ویژگی حاصل از ZC تغییرات شدیدی دیده می‌شود و تفکیک پیکسل‌های خوب و بد به لحاظ جذب نور مناسب به خوبی مشخص نیست. این موضوع در دو نقشه ویژگی دیگر دیده نمی‌شود. نکته‌ای که قابل توجه است این است که در نقشه ویژگی حاصل از $\bar{M}^{(3)}$ نسبت به نقشه ویژگی حاصل از SML ، پیکسل‌های با جذب نور خوب و نواحی مختلف تصاویر به لحاظ داشتن رنگ مناسب به خوبی بازنمایی شده‌اند. همچنین نقشه ویژگی حاصل از $\bar{M}^{(3)}$ نسبت به دو معیار دیگر هموارتر بوده و تغییرات شدید در آن دیده نمی‌شود.



(الف) تصویر پرنور



(ب) تصویر کم‌نور



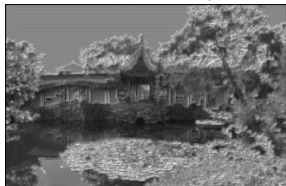
(ت) عبور از صفر ب



(پ) عبور از صفر الف



(ج) لاپلاسین جمع اصلاح شده ب



(ث) لاپلاسین جمع اصلاح شده الف



(ه) رنگ ب



(ج) رنگ الف

شکل ۱: تصاویر کم‌نور و پرنور به همراه نقشه‌های وزنی متناظر

این‌چنین از بین این سه معیار استفاده از معیار $\bar{M}^{(3)}$ را برای استخراج ویژگی و استفاده از نقشه ویژگی حاصل از آن در کنار توابع وزن‌دهی مرجع [15] پیشنهاد می‌کنیم، یعنی، نقشه وزن سوم به صورت رابطه (۱۰) خواهد بود:

$$W_{3,n}(x, y) = \bar{M}^{(3)}(x, y) \quad (10)$$

House	0.9259	0.9526	0.9226
Kluki	0.9600	0.9510	0.9474
Lamp	0.9484	0.9318	0.9249
Landscape	0.9840	0.9793	0.9633
LightHouse	0.9693	0.9761	0.9749
Madison	0.9809	0.9432	0.9473
Memorial	0.9661	0.9298	0.9084
Office	0.9892	0.9684	0.9422
Tower	0.9786	0.9764	0.9172
Venice	0.9661	0.9574	0.9414
Average MEF-SSIM	0.9678	0.9554	0.9378

آزمایش ۲: پس از ایجاد ماتریس وزن نهایی طبق رابطه (۱۱)، می‌توان تصویر نهایی را با تجزیه و تحلیل هرم تصویر مطابق با رابطه (۱۲) ایجاد نمود. در جدول ۲، مقادیر MEF-SSIM حاصل از روش پیشنهادی را در مقایسه با روش مینا یعنی روش مرجع [15] نشان داده شده است.

جدول ۲: مقادیر MEF-SSIM روش پیشنهادی و روش مرجع [15]

Input Sequence	[15]	Proposed Method
Balloons	0.9709	0.9708
BelgiumHouse	0.9661	0.9669
Cadik	0.9634	0.9647
Candle	0.9713	0.9714
Cave	0.9738	0.9779
ChineseGarden	0.9913	0.9914
Farmhouse	0.9791	0.9789
House	0.9621	0.9660
Kluki	0.9747	0.9771
Lamp	0.9518	0.9519
Landscape	0.9806	0.9752
LightHouse	0.9792	0.9790
Madison	0.9810	0.9800
Memorial	0.9667	0.9663
Office	0.9905	0.9909
Tower	0.9873	0.9882
Venice	0.9724	0.9717
Average MEF-SSIM	0.9743	0.9746

مطابق با جدول ۲، دیده می‌شود که روش پیشنهادی به ازای ۱۰ دنباله از ۱۷ دنباله نسبت به مرجع [15] کارایی بهتری دارد. از سوی دیگر، مقدار میانگین بهتری به ازای همه دنباله‌های ورودی نسبت به آن ارائه می‌دهد. می‌توان گفت که الگوریتم پیشنهادی در برابر تغییر محتوای دنباله ورودی مقاوم‌تر است.

آزمایش ۳: به منظور مقایسه روش پیشنهادی با سایر روش‌ها دنباله‌های مرجع [2] را در نظر گرفته و مقادیر MEF-SSIM خروجی روش پیشنهادی و سایر روش‌ها را در جدول ۳ آوردیم. با توجه به جدول ۳، دیده می‌شود که روش پیشنهادی نسبت به سایر روش‌ها مقدار میانگین MEF-SSIM بهتری ارائه داده و کارایی آن ثابت می‌شود.

جدول ۳: عملکرد روش پیشنهادی و روش‌های دیگر بر دنباله‌های [2]

مرجع	Average MEF-SSIM
[11]	0.910
[26]	0.944
[5]	0.965
[9]	0.852
[13]	0.960

مدل با دریافت یک پچ از تصویر خروجی الگوریتم پیشنهادی به نام y و به دنبال آن با انجام یک تجزیه و تحلیل مبتنی بر پچ بر روی دنباله تصاویر ورودی مورد آزمایش، تصویر \hat{x} را ایجاد کرده و با توجه به رابطه (۱۳)، مشابه با رویکرد SSIM به ارزیابی کیفیت محلی تصویر y می‌پردازد:

$$S(\{x_k\}, y) = \frac{2\sigma_{\hat{x}y} + C}{\sigma_{\hat{x}}^2 + \sigma_y^2 + C} \quad (13)$$

در رابطه (۱۳) $\sigma_{\hat{x}y}$ ، σ_y^2 ، $\sigma_{\hat{x}}^2$ به ترتیب بیانگر واریانس محلی \hat{x} ، واریانس محلی y و کوواریانس محلی بین \hat{x} و y است. C یک مقدار کوچک مثبت است که به منظور جلوگیری از اثر اشباع سیستم دیداری در نواحی با کنتراست پایین استفاده شده است. این روند برای کل تصویر چنین انجام می‌شود:

$$Q(Y) = \frac{1}{M} \sum_{j=1}^M S(\{x_k\}(j), y(j)) \quad (14)$$

در رابطه (۱۴)، z اندیس پچ محلی و M تعداد کل پچ‌ها است. از آنجا که ساختار روشنایی نسبی مقیاس بزرگ نمی‌تواند با رویکرد پنجره گذاری محلی به طور کامل دریافت شود، از رویکرد SSIM به صورت چندمقیاسه مطابق با رابطه (۱۵) استفاده گردید:

$$Q(Y) = \prod_{l=1}^L [Q_l(Y)]^{\beta_l} \quad (15)$$

در رابطه (۱۵)، l مجموع تعداد مقیاس‌ها و β_l وزن اختصاص داده شده به مقیاس l ام است.

مقدار Q یک عدد در بازه $[0,1]$ است که هرچه مقدار آن به ۱ نزدیک‌تر باشد به مفهوم حفظ بهتر جزئیات در تصویر خروجی و کیفیت بهتر آن است.

۳-۲- ارزیابی الگوریتم پیشنهادی

در این مقاله، برای ارزیابی کارایی الگوریتم پیشنهادی آزمایش‌های گوناگونی انجام شده است.

آزمایش ۱: ابتدا اثر هر یک از وزن‌های W_1 ، W_2 و W_3 را به صورت جداگانه در ادغام تصاویر بررسی شده و نتایج در جدول ۱ نشان داده شده است. دیده می‌شود که کمترین مقدار برابر با 0.8706 و بیشترین مقدار برابر با 0.9908 است. یعنی، ماتریس‌های وزنی به کار گرفته شده، اطلاعات مناسبی از نورگیری و میزان جزئیات تصویر را دربر دارد.

جدول ۱: بررسی اثر هر وزن به صورت جداگانه روی MEF-SSIM

Input Sequence	W_1	W_2	W_3
Balloons	0.9704	0.9389	0.9295
BelgiumHouse	0.9653	0.9450	0.9428
Cadik	0.9630	0.9316	0.8706
Candle	0.9719	0.9533	0.9467
Cave	0.9458	0.9619	0.9454
ChineseGarden	0.9908	0.9755	0.9578
Farmhouse	0.9771	0.9704	0.9604

شکل ۳، نشان داده شده است. همان‌گونه که دیده می‌شود، با در نظر گرفتن وزن $W3$ در کنار وزن‌های $W1$ و $W2$ کنتراست و جزئیات نواحی روشن تصویر بهبود یافته است.

تصویر خروجی حاصل از اعمال وزن $W1$ تصویر خروجی حاصل از اعمال وزن $W1W2$ تصویر خروجی حاصل از اعمال وزن $W1W2W3$ شکل ۳: اثر وزن $W3$ در کنار وزن $W1$ و $W2$ بر کیفیت تصویر خروجی

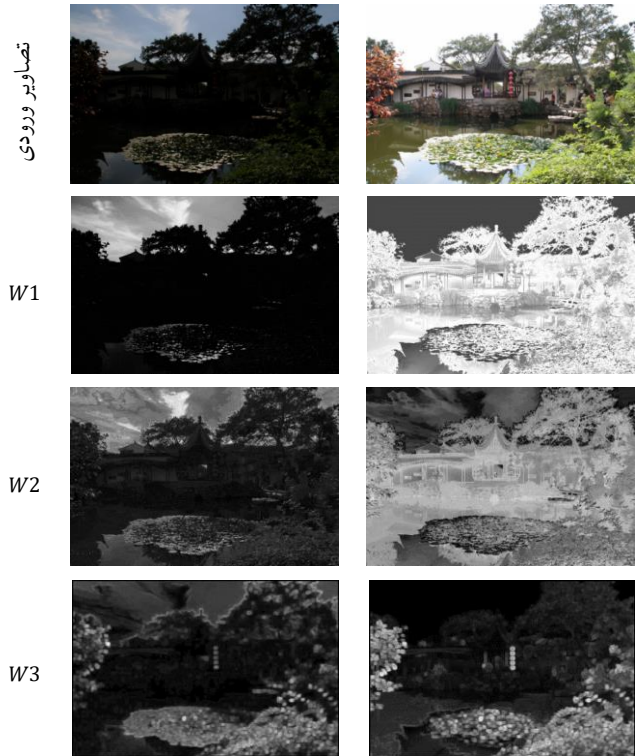
با توجه به شکل ۳ دیده می‌شود که به طور کلی کنتراست نواحی روشن در تصویر خروجی روش پیشنهادی نسبت به مرجع بهبود یافته و دیده می‌شود که جزئیات بیشتری از دنباله تصاویر ورودی به ویژه در نواحی خیلی روشن، در تصویر خروجی الگوریتم پیشنهادی حفظ شده است که نشان از برتری روش پیشنهادی نسبت به مرجع [15] دارد.

۴- جمع بندی

در این مقاله، روشی برای ترکیب تصاویر دنباله‌هایی با نورگیری چندگانه ارائه شده است که از اطلاعات رنگ، تضاد، شدت روشنایی و تابع توزیع تجمعی سطوح خاکستری تصویر بهره می‌گیرد. ارزیابی کارایی روش پیشنهادی طی آزمایش‌های مختلف، نشان‌گر این است که بهره‌گیری از نقشه‌های ویژگی پیشنهادی، هم از دید کیفی و هم از دید کمی به بهبود کارایی الگوریتم مبنای منجر شده است. بهبود وضعیت دیداری در تصاویر خروجی روش پیشنهادی، اغلب به دو صورت بهبود کنتراست نواحی روشن تصاویر خروجی و نیز بهبود پخش نور و حفظ جزئیات در تصویر دیده می‌شود.

[2]	0.969
Proposed Method	0.9746

آزمایش ۴: به منظور درک دیداری بهتر عملکرد هر کدام از توابع وزنی مورد استفاده، سه تصویر با سطوح نورگیری کم، متوسط و زیاد از یک دنباله، به همراه نقشه‌های وزنی مربوطه در شکل ۲ نشان داده شده است.



تصاویر ورودی



شکل ۲: تصاویر کم‌نور و پر نور به همراه نقشه‌های وزنی متناظر

همان‌گونه که انتظار می‌رفت، در نقشه‌های وزنی مبتنی بر شدت روشنایی پیکسل‌ها و گرادیان منحنی هیستوگرام تجمعی به رنگ پیکسل‌های تصویر توجه نمی‌شود. در شکل ۲، در نقشه‌های ویژگی حاصل از معیار رنگ به نقاطی از تصاویر ورودی که رنگ واضح‌تر و بیشتری دارند وزن بالایی اختصاص داده شده است.

آزمایش ۵: برای بررسی اثر وزن $W3$ ، تصویر خروجی الگوریتم به ازای حالات مختلف وزندهی مورد بررسی قرار گرفت. نتیجه آن در

- [23] Choi, S., O.-J. Kwon, and J. Lee, A Method for Fast Multi-Exposure Image Fusion. *Journal of IEEE Access*, 2017. **5**: p. 7371 - 7380.
- [24] Nayar, S.K. and Y. Nakagawa, Shape from focus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1994. **16**: p. 824–831.
- [25] Ma, K., K. Zeng, and Z. Wang, Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 2015. **24**(11): p. 3345-3356.
- [26] Li, Z.G., J.H. Zheng, and S. Rahardja, Detail-enhanced exposure fusion. *IEEE Transactions on Image Processing*, 2012. **21**(11): p. 4672-4676.
- [1] Mertens, T., J. Kautz, and F. Van Reeth. Exposure fusion: A simple and practical alternative to high dynamic range photography. in *Computer graphics forum*. 2009. Wiley Online Library.
- [2] Li, H. and L. Zhang. Multi-exposure fusion with CNN features. in *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018. IEEE.
- [3] Burt, P. and E. Adelson, The Laplacian pyramid as a compact image code. *IEEE Transactions on communications*, 1983. **31**(4): p. 532-540.
- [4] Zhang, W. and W.-K. Cham, Gradient-directed multiexposure composition. *IEEE Transactions on Image Processing*, 2011. **21**(4): p. 2318-2323.
- [5] Li, S., X. Kang, and J. Hu, Image fusion with guided filtering. *IEEE Transactions on Image processing*, 2013. **22**(7): p. 2864-2875.
- [6] Shen, R., et al., Generalized random walks for fusion of multi-exposure images. *IEEE Transactions on Image Processing*, 2011. **20**(12): p. 3634-3646.
- [7] Song, M., et al., Probabilistic exposure fusion. *IEEE Transactions on Image Processing*, 2011. **21**(1): p. 341-357.
- [8] Burt, P.J. and R.J. Kolczynski. Enhanced image capture through fusion. in *1993 (4th) International Conference on Computer Vision*. 1993. IEEE.
- [9] Raman, S. and S. Chaudhuri. Bilateral Filter Based Compositing for Variable Exposure Photography. in *Eurographics (short papers)*. 2009.
- [10] Vonikakis, V., O. Bouzos, and I. Andreadis. Multi-exposure image fusion based on illumination estimation. in *Proc. IASTED Int. Conf. on Signal and Image Processing and Applications*. 2011.
- [11] Gu, B., et al., Gradient field multi-exposure images fusion for high dynamic range image visualization. *Journal of Visual Communication and Image Representation*, 2012. **23**(4): p. 604-610.
- [12] Shen, J., et al., Exposure fusion using boosting Laplacian pyramid. *IEEE Trans. Cybernetics*, 2014. **44**(9): p. 1579-1590.
- [13] Li, S. and X. Kang, Fast multi-exposure image fusion with median filter and recursive filter. *IEEE Transactions on Consumer Electronics*, 2012. **58**(2): p. 626-632.
- [14] Gastal, E.S. and M.M. Oliveira. Domain transform for edge-aware image and video processing. in *ACM Transactions on Graphics (ToG)*. 2011. ACM.
- [15] Lee, S.-h., J.S. Park, and N.I. Cho. A Multi-Exposure Image Fusion Based on the Adaptive Weights Reflecting the Relative Pixel Intensity and Global Gradient. in *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018. IEEE.
- [16] Yang, Y., et al., Multi-Exposure Estimation and Fusion Based on A Sparsity Exposure Dictionary. *IEEE Transactions on Instrumentation and Measurement*, 2019.
- [17] Moriyama, D., et al. Saturation-Based Multi-Exposure Image Fusion Employing Local Color Correction. in *2019 IEEE International Conference on Image Processing (ICIP)*. 2019. IEEE.
- [18] Soundararajan, R. and A.C. Bovik, Survey of information theory in visual quality assessment. *Signal, Image and Video Processing*, 2013. **7**(3): p. 391-401.
- [19] Cohen, E. and Y. Yitzhaky, No-reference assessment of blur and noise impacts on image quality. *Signal, image and video processing*, 2010. **4**(3): p. 289-302.
- [20] Zhang, J. and T.M. Le, A new no-reference quality metric for JPEG2000 images. *IEEE Transactions on Consumer Electronics*, 2010. **56**(2): p. 743-750.
- [21] Panetta, K., C. Gao, and S. Agaian, No reference color image contrast and quality measures. *IEEE transactions on Consumer Electronics*, 2013. **59**(3): p. 643-651.
- [22] Hasler, D. and S.E. Suesstrunk. Measuring colorfulness in natural images. in *Human vision and electronic imaging VIII*. 2003. International Society for Optics and Photonics.



قطعه‌بندی تومورهای مغزی گلیوما مبتنی بر تصاویر MRI با استفاده از یادگیری عمیق

مطهره آقالری^۱، علی آقاگل زاده^۲، مهدی ازوجی^۳

^۱ دانشجوی کارشناسی ارشد مهندسی برق و کامپیوتر، گروه آموزشی مخابرات، دانشگاه صنعتی نوشیروانی، بابل،
motaahareh.aghafari@nit.ac.ir

^۲ استاد دانشکده مهندسی برق و کامپیوتر، گروه آموزشی مخابرات، دانشگاه صنعتی نوشیروانی، بابل
aghagol@nit.ac.ir

^۳ استادیار دانشکده مهندسی برق و کامپیوتر، گروه آموزشی الکترونیک، دانشگاه صنعتی نوشیروانی، بابل
m.ezaji@nit.ac.ir

چکیده

تومورهای مغزی گلیوما، اصلی‌ترین نوع تومورهای مغزی اند که از رشد غیرطبیعی سلول‌های گلیال ایجاد می‌شوند. این سلول‌ها در لایه‌های نورون‌ها قرار داشته و نقش محافظتی دارند. روش تصویربرداری MRI یکی از رایج‌ترین روش‌های مورد استفاده برای تشخیص و قطعه‌بندی تومورهای مغزی است. تشخیص مکان و ناحیه گسترش تومور، به طور معمول به صورت دستی انجام می‌شود که کاری بسیار وقت‌گیر و طاقت‌فرسا است. از این رو، وجود روش‌هایی برای قطعه‌بندی خودکار تومورهای مغزی ضروری است. در این مقاله ساختاری مبتنی بر شبکه عصبی عمیق، برای قطعه‌بندی خودکار تومورهای مغزی گلیوما پیشنهاد شده است. ارزیابی روش پیشنهادی بر روی پایگاه داده استاندارد BRATS2013 حاوی چهار نوع اسکن T1، T1c، T2 و تصاویر MRI، نشان می‌دهد که الگوریتم پیشنهادی می‌تواند با دقت % ۹۶/۳۵ مکان تومورهای مغزی را آشکار کند. همچنین اندازه سه معیار ضریب شباهت دایس، حساسیت و ویژگی، برای نتیجه‌ی قطعه‌بندی حاصل از روش پیشنهادی، به ترتیب % ۸۱/۲۳، % ۹۱/۱۱ و % ۹۶/۸۴ به دست آمد.

کلمات کلیدی

قطعه‌بندی، تومورهای مغزی، گلیوما، یادگیری عمیق، شبکه عصبی کانولوشنی

تومورهایی که نمی‌توان از نظر فیزیکی آن‌ها را از بین برد، می‌توان از اشعه درمانی و شیمی‌درمانی استفاده نمود [۲].

به منظور انتخاب هر یک از روش‌های درمانی نیاز است که تصویری از مغز در اختیار پزشک قرار داده شود تا با قطعه‌بندی تصویر، مکان تقریبی و ناحیه گسترش تومور تعیین گردد. هر تومور مغزی به چهار بخش، ادم^{۱۹}، نکروز^{۱۱}، ناحیه‌ی افزایشی^{۱۱} و ناحیه‌ی غیرافزایشی^{۱۲}، تقسیم می‌شود که با تعیین این چهار بخش، مکان کلی تومور و ناحیه‌ی گسترش آن در تصویر مشخص می‌گردد [۲]. روش تصویربرداری MRI به دلیل عدم استفاده از اشعه و داروهای رادیواکتیو در مقایسه با روش‌های دیگر مانند CT و PET، برای بدن انسان بی‌ضرر بوده [۳] و دارای وضوح بالا است؛ لذا برای تشخیص مکان و ناحیه گسترش تومور استفاده می‌شود [۴]. قطعه‌بندی تومورهای گلیوما براساس داده‌های MRI به دلایلی چالش برانگیز است:

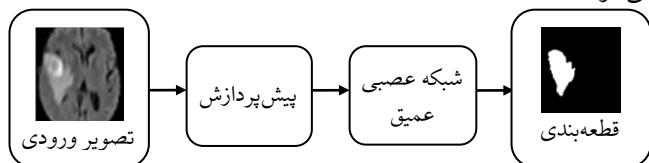
۱- مقدمه

در میان تومورهای مغزی، تومور مغزی گلیوما^{۱۶} به دلیل ساختار و اندازه‌ی متنوع، یکی از اصلی‌ترین نوع تومورهای مغزی به شمار می‌آید. سازمان بهداشت جهانی این نوع تومورها را به دو دسته تومورهای درجه پایین^{۱۷} و درجه بالا^{۱۸} طبقه‌بندی کرده است [۱]. تشخیص تومورهای درجه بالا به علت شباهت بافت‌های توموری و بافت‌های غیر تومور مغز، نسبت به تومورهای درجه پایین دشوارتر است. در گلیومای درجه پایین، بیماران می‌توانند امید به زندگی چندین ساله و در گلیومای درجه بالاتر، می‌توانند امید به زندگی تا یکسال را هم داشته باشند. روش‌های درمانی مختلفی برای تومورهای مغزی وجود دارد و جراحی رایج‌ترین روش درمانی محسوب می‌شود، اما برای کاهش رشد

شبکه تماماً متصل کانونلوشنی قطعه‌بندی شدند. سه نوع شبکه آبخاری کانونلوشنی جهت قطعه‌بندی خودکار تومورهای مغزی، توسط Wu و همکاران [۱۵] معرفی شد. معماری‌های آبخاری از دوشبکه‌ی کانونلوشن تشکیل شده که خروجی شبکه کانونلوشن اول به ورودی یا نقشه‌های ویژگی شبکه کانونلوشن دوم برای استخراج ویژگی‌های محتوایی^{۱۱۳} بیشتر از تصویر، ترکیب می‌شوند. Hussain و همکاران [۱۶] شبکه‌های عصبی کانونلوشنی Nexus را طی یک آموزش دو مرحله‌ای برای مقابله با مشکل عدم تعادل پیکسل‌های ناحیه‌ی توموری و سالم مغز به کار گرفتند. در مطالعه Chen و همکاران [۱۷]، دو معماری معروف شبکه‌های کانونلوشنی عمیق DeepMedic و Unet جهت قطعه‌بندی تومورهای مغزی ارائه گردید. در این مقاله، یک روش قطعه‌بندی خودکار مبتنی بر شبکه‌های عصبی عمیق معرفی شده در پژوهش Hussain و همکاران [۱۶]، جهت طبقه‌بندی پیکسل‌های تصویر به دو ناحیه‌ی توموری و غیرتوموری پیشنهاد شده است. معیارهای حساسیت و ویژگی به ترتیب ۹۱/۱۱٪ و ۹۶/۸۴٪ به دست آمده که نسبت به پژوهش Hussain و همکاران [۱۶] بر روی پایگاه داده‌ی BRATS2013 بهبود چشم‌گیری را نشان می‌دهد.

۲- روش پیشنهادی

در شکل (۱) قالب کلی روش پیشنهادی نشان داده شده است. مطابق این شکل تصویر ورودی، طی یک سری پردازش اولیه بهبود یافته و سپس از شبکه عصبی عمیق Linear-Middle-Linear، برای قطعه‌بندی تومورهای مغزی گلیوما مبتنی بر تصاویر MRI استفاده می‌شود.



شکل ۱: بلوک دیاگرام روش پیشنهادی

۲-۱- پیش‌پردازش

تصاویر MRI همراه با انحراف میدان بایاس تغییر می‌کنند که باعث تغییر شدت مقادیر پیکسل‌های بافت‌های یکسان در تصویر می‌شود. برای رفع این مشکل، از فیلتر N4ITK جهت اصلاح میدان بایاس و حذف مصنوعات در تصاویر MRI استفاده شده است [۱۸].

از آنجا که هر تصویر MRI، خود از برش^{۱۱۴} هایی تشکیل شده و مقادیر شدت در میان برش‌ها متفاوت است، علاوه بر اصلاح میدان بایاس، نرمال‌سازی برش‌ها نیز انجام می‌گیرد [۱۵].

اگر x برش اصلی تصویر MRI با میانگین μ و انحراف معیار σ باشد، x_n برش نرمالیزه آن طبق رابطه زیر، به دست می‌آید که میانگینی برابر صفر و انحراف معیاری برابر یک دارد:

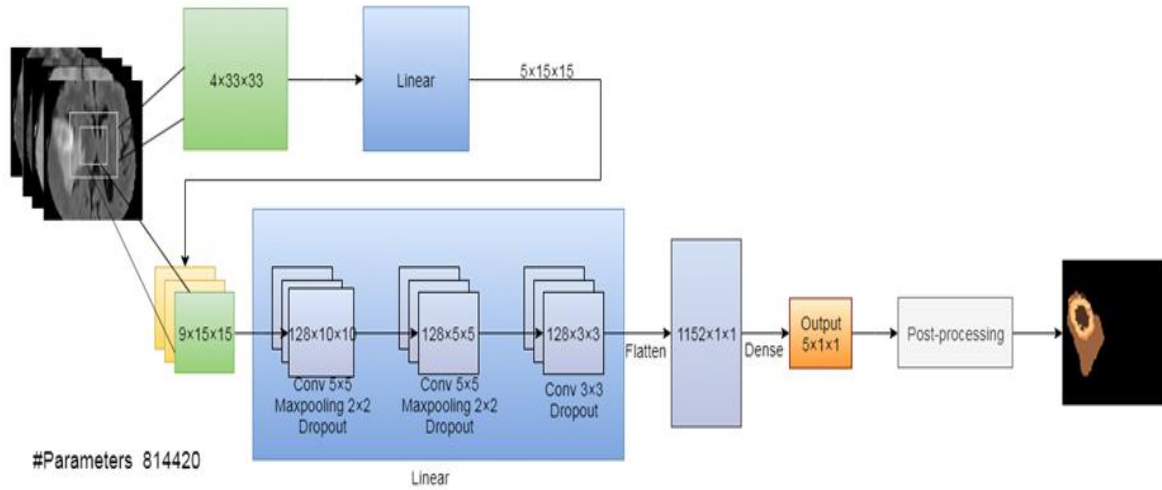
● گلیوما ظاهر یکسان با ضایعه‌های مغزی دیگر همچون سکتی مغزی را دارد [۵].

● تومورهای گلیوما با هجوم به بافت‌های اطراف مغز باعث ایجاد مرزهای فازی با بافت‌های سالم مغز می‌شوند [۵].

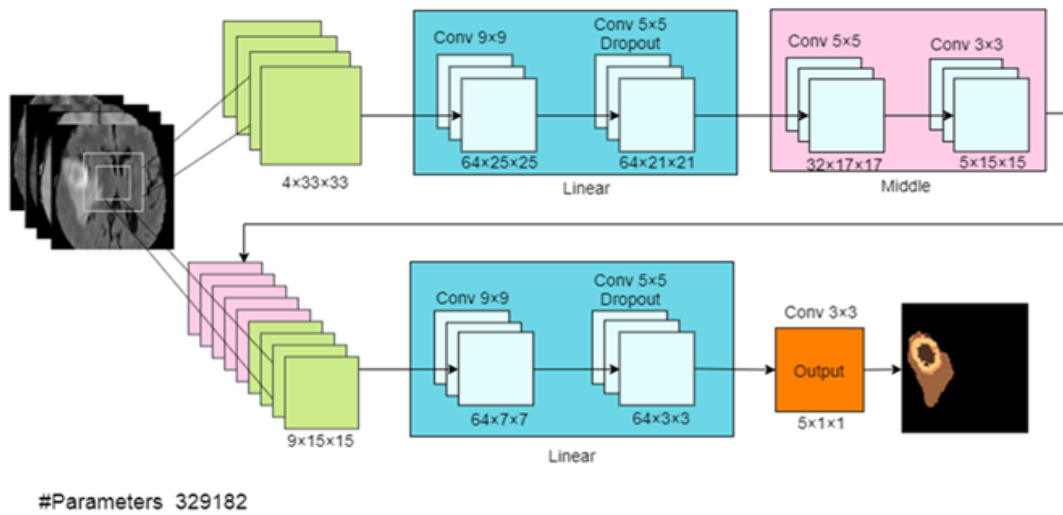
● گلیوما ممکن است در هر مکانی از مغز با شکل و اندازه متنوع ظاهر شود [۶].

با توجه به دلایل ذکر شده قطعه‌بندی دستی تومورهای مغزی بسیار دشوار است. لذا تلاش‌های زیادی برای توسعه روش‌های قطعه‌بندی تومورهای مغزی گلیوما به صورت نیمه‌خودکار و خودکار انجام گرفته است. در روش قطعه‌بندی نیمه‌خودکار با استفاده از الگوریتم‌های پردازش تصویر منطقه تقریبی تومور مشخص شده و پزشک نتایج را ارزیابی کرده و در صورت نارضایتی، فرایند را اصلاح یا تکرار می‌کند. در روش‌های خودکار، قطعه‌بندی تنها با استفاده از الگوریتم‌های پردازش تصویر و بدون دخالت پزشک انجام می‌شود [۷]. در ادامه به برخی از تلاش‌های انجام شده در این زمینه پرداخته می‌شود.

Gupta و همکاران [۱]، دو نوع قطعه‌بندی متفاوت را برای دو زیر مجموعه از تصاویر پیشنهاد نمودند که مجموعه اول ترکیبی از تصاویر T1c و T1 توسط آستانه‌گیری تطبیقی و عملیات ریخت‌شناسی و مجموعه دوم با ترکیبی از تصاویر FLAIR و T2 توسط فیلتر تشخیص لبه Canny قطعه‌بندی شدند. در پژوهش انجام شده توسط Devkota و همکاران [۳]، قطعه‌بندی مبتنی بر عمل‌گر ریخت‌شناسی فرسایش با عنصر ساختاری دایروی و عمل‌گر بستن با عنصر ساختاری مربعی پیشنهاد شده که ساده و سریع بوده اما دارای دقت قطعه‌بندی پایین در تومورهای درجه بالا است. Ilhan و همکاران [۸] از روش آستانه‌گذاری مبتنی بر میانگین‌گیری شدت مقادیر پیکسل‌های غیر صفر تصویر استفاده نمودند. نتیجه قطعه‌بندی، عمل‌کرد ضعیفی را در تومورهای درجه بالا به علت شباهت مقادیر شدت بافت‌های توموری و غیر تومور، نشان داد. در مطالعه Li و همکاران [۹]، قطعه‌بندی مبتنی بر رشد ناحیه معرفی گردید که شامل تخمین ناحیه‌ی توموری با استفاده از الگوریتم خوشه‌بندی فازی، انتخاب مراکز خوشه‌ها به عنوان ناحیه‌های اولیه و رشد ناحیه‌های اولیه براساس معیار شباهت پیکسل‌ها و موقعیت‌شان است. در مطالعه Rahima و همکاران [۱۰]، ابتدا اطلاعات فاز محلی با هدف تعیین مقدار تابع سطح، براساس ترکیبی از روش رشد ناحیه و عملیات ریخت‌شناسی استخراج شده، سپس تابع تعیین سطح بر روی مجموعه‌ای از تکرارها با استفاده از ناحیه‌های اولیه تکامل می‌یابد. Tong و همکاران [۱۱]، یک روش قطعه‌بندی براساس ویژگی‌های بافت و کدگذاری^{۱۱۵} تک مبتنی بر اسکن FLAIR تصاویر MRI را پیشنهاد نمودند که یادگیری بردارهای پایه توسط خوشه‌بندی فازی انجام گرفت. Nabizadeh و همکاران [۱۲] در پژوهش خود، روش ماشین بردار پشتیبان را برای استخراج ویژگی‌های تصویر و قطعه‌بندی آن به دو ناحیه‌ی توموری و غیر توموری، به کار بردند. Pereira و همکاران [۱۳] از یک شبکه عصبی کانونلوشنی با هسته‌هایی به ابعاد 3×3 برای قطعه‌بندی خودکار تومور مغزی استفاده نمودند. در پژوهش Shen و همکاران [۱۴] ترکیبی از چهار نوع تصاویر MRI توسط یک



(الف)



(ب)

شکل ۲: الف- معماری Linear-Linear [۱۶]، ب- معماری پیشنهادی Linear-Middle-Linear

همراه با آن یک تصویر قطعه‌بندی شده توسط پزشک، به پنج ناحیه‌ی ادم، نکروز، ناحیه‌ی افزایشی، غیرافزایشی و ناحیه‌ی غیر تومور وجود دارد.

مدل پیشنهادی در محیط Google colab و بستر پایتون با استفاده از کتابخانه‌ی keras پیاده‌سازی شده است.

۳-۱- معیارهای ارزیابی

برای ارزیابی کارایی روش پیشنهادی در قطعه‌بندی تومورهای مغزی گلیوما و تعیین مکان و ناحیه‌ی گسترش تومور از معیارهای دقت^۳، ضریب شباهت دایس، حساسیت و ویژگی بهره‌گرفته شده است. این معیارها به ترتیب با رابطه‌های (۲) تا (۵) اندازه‌گیری می‌شوند.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (2)$$

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Specifity} = \frac{TN}{TN + FP} \quad (5)$$

که در آن TP، FN، TN و FP به ترتیب نشان‌دهنده‌ی ناحیه‌ی توموری صحیح، ناحیه‌ی توموری ناصحیح، ناحیه‌ی غیر تومور صحیح و ناحیه‌ی غیر تومور ناصحیح است.

۳-۲- آزمایش‌ها

از آنجایی که تعداد پیکسل‌های بافت سالم در تصویر MRI، نسبت به بافت توموری بسیار بیشتر است، در این مقاله، آموزش شبکه‌ی عصبی عمیق پیشنهادی با انتخاب تصادفی تعداد یکسان از پچ‌های ناحیه‌های توموری و غیرتوموری مغز، توسط بهینه‌ساز Adam تنها در یک مرحله انجام گرفت که در مقایسه با آموزش دومرحله‌ای در [۱۶-۱۵]، سرعت آموزش شبکه افزایش یافت.

برای بررسی و ارزیابی روش پیشنهادی آزمایش‌های متعددی انجام شده است:

آزمایش ۱: ۷۰٪ داده‌ها به صورت تصادفی برای داده‌های آموزش و باقی داده‌ها برای آزمایش در نظر گرفته شدند. این کار ده بار تکرار شد و میانگین معیارهای ارزیابی به دست آمد. طبق نتایج گزارش شده در جدول (۱) روش پیشنهادی در مقایسه با مطالعات انجام شده [۳، ۸، ۲۰] به دقت بالاتری رسیده و نیز بر روی پایگاه داده‌ای مشابه با پژوهش‌های دیگر [۱۶-۱۳]، معیارهای حساسیت و ویژگی را بهبود داده است.

$$x_n = \frac{x - \mu}{\sigma} \quad (1)$$

۲-۲- معماری پیشنهادی Linear-Middle-Linear

معماری Linear-Linear معرفی شده در [۱۶] از دو شبکه کانولوشن مشابه به نام Linear تشکیل شده که در شکل (۲-الف) نشان داده شده است. در ساختار شبکه‌ی Linear، لایه‌های کانولوشن به همراه لایه‌ی بیشینه‌گیر^۱ به کار گرفته شده است. لایه‌ی بیشینه‌گیر، با حرکت دادن یک پنجره‌ی ۲×۲ بر روی نقشه‌های ویژگی خروجی هر لایه‌ی کانولوشن، ویژگی‌هایی با بیشترین مقدار را استخراج کرده و ابعاد نقشه‌های ویژگی را کاهش می‌دهد. در نتیجه برخی از ویژگی‌های استخراجی از دست می‌روند.

می‌دانیم که با افزایش تعداد لایه‌های کانولوشن، می‌توان ویژگی‌های بیشتری را استخراج کرده و دقت قطعه‌بندی را افزایش داد. از این رو برای بهبود عملکرد ساختار شبکه پیشنهادی در [۱۶]، در این مقاله تغییرات زیر انجام شده است:

- افزودن یک بلوک میانی شامل دو لایه کانولوشن با هسته‌هایی به ابعاد ۵×۵ و ۳×۳، همراه با حذف لایه‌های بیشینه‌گیر.
- جایگزینی لایه‌ی اتصال فراگیر با یک لایه کانولوشن.
- آموزش شبکه در یک مرحله با بهینه‌ساز Adam.
- کاهش تعداد مراحل پیش‌پردازش با حذف مرحله‌ی نرمال‌سازی پچ‌های استخراجی از تصویر.
- عدم استفاده از عملیات پس‌پردازش.

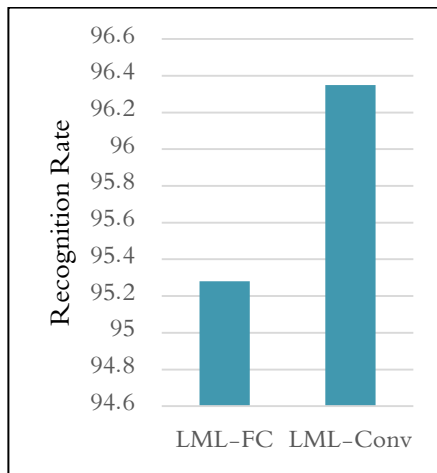
معماری پیشنهادی تحت عنوان Linear-Middle-Linear (LML) در شکل (۲-ب) نشان داده شده است.

در معماری LML، شبکه‌ی کانولوشن اول شامل دو لایه کانولوشن با هسته‌هایی به ابعاد ۹×۹ و ۵×۵ به همراه لایه‌های نرمال‌سازی دسته‌ای و یک لایه حذف تصادفی بوده که در ورودی پچ‌هایی از چهار نوع اسکن MRI به ابعاد ۳۳×۳۳ را می‌گیرد. خروجی شبکه اول به ابعاد ۲۱×۲۱ از یک بلوک میانی با دو لایه کانولوشن با هسته‌هایی به ابعاد ۵×۵ و ۳×۳ عبور کرده و خروجی به ابعاد ۱۵×۱۵×۱۵ را تولید می‌کند. این خروجی با پچ‌های استخراج شده از چهار نوع اسکن MRI به ابعاد ۱۵×۱۵ الحاق شده و در مجموع ۹ پچ وارد شبکه دوم می‌شود. در نهایت خروجی شبکه کانولوشن دوم از یک لایه‌ی اتصال کانولوشن عبور کرده که در مقایسه با لایه‌ی اتصال فراگیر دقت شبکه را افزایش می‌دهد [۱۵]. با به‌کارگیری تابع فعال‌ساز^۲ Softmax پس از لایه آخر، پیکسل میانی پچ ورودی به یکی از چهار ناحیه‌ی تومور و غیر تومور مغز طبقه‌بندی شده و مکان و ناحیه‌ی گسترش تومور تعیین می‌گردد [۱۶].

۳- آزمایش و ارزیابی

در ارزیابی روش پیشنهادی، پایگاه داده BRATS2013 [۱۹] استفاده شده است. این پایگاه داده شامل اسکن‌های T1، T1c، T2 و Flair از ۳۰ بیمار (۱۰ گلیوما درجه پایین و ۲۰ گلیوما درجه بالا) است که

اتصال کانولوشن دقت بالاتری را در قطعه بندی ناحیه کامل تومور، به علت استخراج ویژگی‌های بیشتر، دارد.



شکل ۴: مقایسه متوسط دقت شبکه LML در دو نوع اتصال فراگیر و کانولوشن در تشخیص ناحیه کامل تومور

در شکل (۵-الف) و (۵-ب) به ترتیب نمونه‌ای از تصویر قطعه‌بندی شده تومور مغزی گلیوما درجه پایین و بالا، توسط پزشک متخصص و روش پیشنهادی نشان داده شده است. همان‌گونه که در شکل (۵-ب) دیده می‌شود، قطعه‌بندی تومور مغزی گلیوما درجه بالا، به علت تشابه زیاد بین بافت‌های سالم و تومور در تصویر، نسبت به گلیوما درجه پایین دشوارتر است.

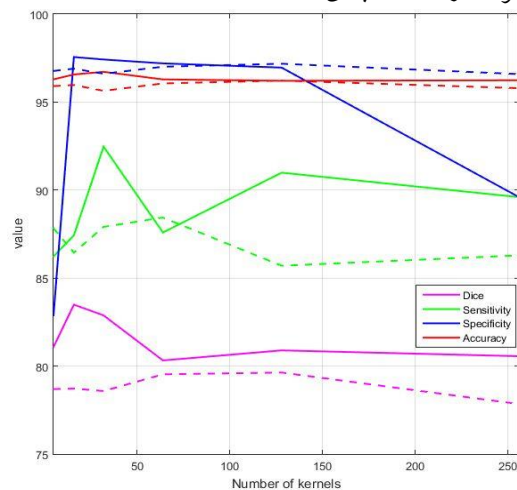
۴- نتیجه‌گیری

این مقاله به ارائه یک روش کارآمد جهت قطعه‌بندی خودکار تومورهای مغزی گلیوما و تعیین مکان و ناحیه‌ی گسترش تومور بر پایه اسکن‌های T1، T1c، T2 و تصاویر MRI پرداخته است. بدین منظور از فیلتر N4ITK جهت بهبود تصویر اولیه استفاده کرده و برای قطعه‌بندی از یادگیری عمیق مبتنی بر معماری LML بهره گرفته شده است. در معماری پیشنهادی، با افزودن یک بلوک میانی به ساختار [۱۶] و کاهش تعداد لایه‌های کانولوشن شبکه‌ی Linear، معیارهای دقت، حساسیت و ویژگی، همراه با کاهش تعداد پارامترهای شبکه، افزایش یافتند. همچنین آموزش یک مرحله‌ای شبکه توسط بهینه‌ساز Adam و جایگزینی لایه‌ی اتصال فراگیر با یک لایه‌ی کانولوشن، منجر به افزایش سرعت آموزش شبکه شده است. نیاز به یادآوری دارد که روش پیشنهادی نیاز به عملیات پس‌پردازش نداشته و با حذف یک مرحله پیش‌پردازش، متوسط زمان لازم جهت قطعه‌بندی تومور مغزی یک تصویر کاهش یافته است.

جدول ۱: مقایسه پژوهش‌های مرتبط با قطعه‌بندی تومورهای مغزی گلیوما

ویژگی	حساسیت	ضریب شباهت دایس	دقت	تعداد تصویر	پژوهش انجام شده
-	-	۸۹	۹۷/۱۳	۸۰	[۱]
-	-	-	۹۲	۱۹	[۳]
۸۳	۹۲	۸۷	-	۳۰	[۶]
۱۰۰	۹۴/۲۸	-	۹۶	۱۰۰	[۸]
۷۷	۸۳	۸۸	-	۳۰	[۱۳]
۸۴	۹۱	۸۸	-	۳۰	[۱۴]
۸۹	۸۵	۸۷	-	۳۰	[۱۵]
۹۳	۸۶	۸۵	-	۳۰	[۱۶]
-	-	-	۸۴/۷۲	۷۲	[۲۰]
۹۶,۸۴	۹۱,۱۱	۸۱,۲۳	۹۶,۳۵	۳۰	روش پیشنهادی LML

آزمایش ۲: تعداد هسته‌های 5×5 لایه کانولوشن میانی، طی پیش‌پردازش‌های اعمالی در [۱۶] (شامل استفاده از فیلتر N4ITK، نرمال‌سازی برش تصویر MRI و نرمال‌سازی پچ‌ها) و پیش‌پردازش پیشنهادی در این مقاله (حذف نرمال‌سازی پچ‌ها) مورد بررسی قرار گرفت. همان‌گونه که در شکل (۳) دیده می‌شود، با افزایش تعداد هسته‌ها، به دلیل بیش‌برازش شبکه، معیارهای ارزیابی کاهش می‌یابد. لذا بهترین نتیجه برای تعداد ۳۲ هسته‌ی لایه‌ی میانی به‌دست آمد. همچنین نرمال‌سازی پچ‌های استخراجی از تصویر، به دلیل کاهش بیش از حد مقادیر پیکسل‌های تصویر، منجر به کاهش معیارهای ارزیابی و افزایش متوسط زمان قطعه‌بندی شده است.



شکل ۳: نمودار تعداد هسته‌های 5×5 لایه کانولوشن میانی پس از پیش‌پردازش: [۱۶]--، پیشنهادی مقاله--

آزمایش ۳: نوع لایه‌ی آخر معماری LML، در دو حالت اتصال فراگیر و اتصال کانولوشنی، مورد بررسی قرار گرفت. در شکل (۴) متوسط دقت دو شبکه حاصل، گزارش شده که نشان می‌دهد، شبکه با

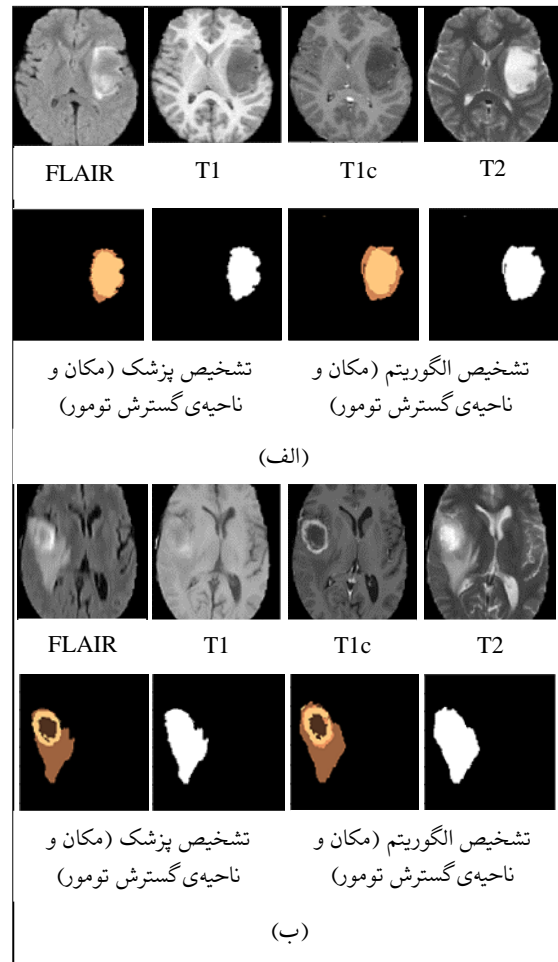
- [6] Zhao, X., et al., A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Medical image analysis, Vol. 43: pp. 98-111, 2018.
- [7] Bauer, S., Wiest, R., Nolte, L.-P., Reyes, M., A survey of MRI-based medical image analysis for brain tumor studies. Phys. Med. Biol., Vol.58, pp.97-129, 2013 .
- [8] Ilhan, U., Ilhan, A., Brain tumor segmentation based on a new threshold approach. Procedia computer science, Vol.120, pp. 580-587, 2017.
- [9] Li, Q., et al., Glioma segmentation using a novel unified algorithm in multimodal MRI images. IEEE Access: pp. 1-1, 2018.
- [10] Rahima, Z., et al. Segmentation of low-grade gliomas based on the growing region and level sets techniques. in 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE, 2018.
- [11] Tong, J., et al., MRI brain tumor segmentation based on texture features and kernel sparse coding. Biomedical Signal Processing and Control, 47: p. 387-392, 2019.
- [12] Nabizadeh, N., Kubat, M., "Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features," Computers & Electrical Engineering, Vol.45, pp. 286-301, 2015.
- [13] Pereira, S., et al., Brain tumor segmentation using convolutional neural networks in MRI images. IEEE transactions on medical imaging, Vol. 35, pp. 1240-1251, 2016.
- [14] Shen, H., Zhang, J., Zheng, W., Efficient symmetry-driven fully convolutional network for multimodal brain tumor segmentation. in 2017 IEEE International Conference on Image Processing (ICIP), IEEE, 2017.
- [15] Wu, G., et al., Sparse representation-based Radiomics for the diagnosis of brain tumors. IEEE transactions on medical imaging, Vol 37, pp. 893-905, 2018.
- [16] Hussain, S., Anwar, S.M., Majid, M., Segmentation of glioma tumors in brain using deep convolutional neural network. Neurocomputing, Vol. 282, pp. 248-261, 2018.
- [17] Chen, S., Ding, C., Liu, M., Dual-force convolutional neural networks for accurate brain tumor segmentation. Pattern Recognition, Vol.88, pp90-100, 2018.
- [18] Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., N4ITK: improved N3 bias correction, IEEE Trans. Med. Imaging, Vol. 29, pp.1310-1320, 2010 .
- [19] Menze B.H., Jakab A., Bauer S., et al., The multimodal brain tumor im- age segmentation benchmark (brats), IEEE Trans. Med. Imaging, Vol.34, No.10, pp.1993-2024, 2014 .
- [20] Sujan, M., Alam, N., Abdullah, S., Jahirul, M., A Segmentation based Automated System for Brain Tumor Detection. International Journal of Computer Applications, Vol.153, pp.41-49, 2016.
- [21] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, Vol.15, pp.1929-1958, 2014.

زیر نویس ها

¹ Max-pooling

² Activation function

³ Accuracy



شکل ۵: نمونه‌ای از کارایی روش پیشنهادی در قطعه‌بندی تصویر الف- گلیومای درجه پایین، ب- گلیومای درجه بالا
 ادم، نکروز، ناحیه‌ی افزایشی، غیرافزایشی

مراجع

- [1] Gupta, N., Bhatele, P., Khanna, P., Glioma detection on brain MRIs using texture and morphological features with ensemble learning. Biomedical Signal Processing and Control, Vol. 47, pp. 115-125, 2019.
- [2] Işın, A., Direkoğlu, C., Şah, M., Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Computer Science, 102: pp. 317-324, 2016.
- [3] Devkota, B., et al., Image segmentation for early stage brain tumor detection using mathematical morphological reconstruction. Procedia Computer Science, Vol.125: pp. 115-123, 2018.
- [4] Havaei, M., et al., Brain tumor segmentation with deep neural networks. Medical image analysis, Vol.35, pp. 18-31, 2017.
- [5] Goetz, M., Weber, C., Binczyk, F., Polanska, J., Tarnawski, R., Bobek-Billewicz, B., et al., DALSA: domain adaptation for supervised learning from sparsely annotated MR images. IEEE Trans. Med. Imaging, Vol.35, pp.184-196, 2016.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود

بهبود امنیت در سیستم مبتنی بر کدینگ CIOD بر مبنای انتخاب تصادفی آنتن

علیرضا بقائی پوری^۱، محمد ترابی^۲

^۱ دانشجوی دکتری، دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران، ایران
a_baghaeipouri@sbu.ac.ir

^۲ هیأت علمی، دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران، ایران
m_torabi@sbu.ac.ir

چکیده

در این مقاله به بررسی امنیت در سیستم CIOD (Coordinate Interleaved Orthogonal Designs) با چهار آنتن در فرستنده و یک آنتن در گیرنده پرداخته‌ایم. ماتریس کدینگ CIOD بر مبنای دو ماتریس یکسان الموتری انتخاب شده است. همچنین مدل ارسال و دریافت در سیستم یاد شده بیان گردیده و روابط بسته تئوری برای احتمال قطع ارتباط محرمانه بدست آورده شده است. علاوه بر این جهت بهبود امنیت در این سیستم، روش کم‌هزینه‌ای را به کار برده‌ایم که مبتنی بر انتخاب تصادفی آنتن فرستنده در هر اسلات زمانی است. انتخاب تصادفی آنتن با استفاده از کلیدی که وابسته به کانال بین فرستنده و گیرنده مجاز است، کنترل می‌شود و لذا شنودگر از آن اطلاعی ندارد. با استفاده از این روش ساده نیازی به افزایش توان سیگنال ارسالی یا اختصاص بخشی از توان به ارسال نویز نخواهیم داشت و با این حال می‌توان شنودگر را دچار مشکل کرد. در بخش شبیه‌سازی صحت روابط تئوری با استفاده از شبیه‌سازی بررسی شده و تطابق نتایج تئوری و شبیه‌سازی نمایش داده شده است. علاوه بر این اثر تغییر بهره کانال گیرنده مجاز و شنودگر بر روی معیارهای امنیت نشان داده شده است.

کلمات کلیدی

امنیت لایه فیزیکی، کدینگ فضا-زمان، CIOD، ظرفیت محرمانگی

۱- مقدمه

در [1] ارائه شده است از دو آنتن در فرستنده و یک آنتن در گیرنده استفاده می‌کند. برای این شبکه در [2] روشی مبتنی بر افزودن نویز مصنوعی به سیگنال جهت افزایش امنیت ارائه شده است. همچنین در [3] روشی بر اساس چرخش فاز تصادفی جهت افزایش امنیت استفاده شده است. ساختار سیستم^۲ CIOD نیز به عنوان همتایی برای سیستم OSTBC در مقالات مختلفی همچون [8]-[4] مورد بررسی قرار گرفته و نشان داده شده این سیستم می‌تواند عملکرد بهتری نسبت به سیستم OSTBC از نظر خطای بیت داشته باشد. این میان امنیت در سیستم CIOD کمتر مورد توجه واقع شده و نیاز به کار پژوهشی بیشتری در این حوزه احساس می‌شود.

با پیشرفت تکنولوژی و افزایش روزافزون کاربرد شبکه‌های بی‌سیم، مسئله امنیت در این شبکه‌ها نیز بیش از پیش مورد توجه قرار گرفته است. با شناخت ضعف‌های رمزنگاری، امنیت در لایه فیزیکی به عنوان مکملی جهت افزایش امنیت در سیستم‌های مخابراتی معرفی شد که هدف آن استفاده از ویژگی‌های کانال همچون نویز و محوشوندگی در راستای افزایش امنیت است.

ساختار و امنیت در سیستم^۱ OSTBC به عنوان شبکه‌ای پرکاربرد در مقالات زیادی از نظر امنیت بررسی شده است. به طور مثال شبکه الموتری نوعی شبکه مبتنی بر کدینگ فضا-زمان است و بر اساس آن چه

که در آن C_s ظرفیت محرمانگی، C_d و C_e به ترتیب ظرفیت کانال گیرنده مجاز و ظرفیت کانال شنودگر و $[x]^+ = \max(x, 0)$ است.

ظرفیت کانال گیرنده مجاز در سیستم CIOD با ۴ آنتن فرستنده و یک آنتن گیرنده به صورت زیر قابل بیان است [5].

$$C_d^{\text{CIOD}} = \frac{1}{2} \left[\log_2 \left\{ 1 + \frac{\rho}{2} \left(|h_0^d|^2 + |h_1^d|^2 \right) \right\} + \log_2 \left\{ 1 + \frac{\rho}{2} \left(|h_2^d|^2 + |h_3^d|^2 \right) \right\} \right]. \quad (2)$$

که در آن $h_i^d, (i = 0, \dots, 3)$ بهره کانال بین i -امین آنتن فرستنده و آنتن گیرنده مجاز است. علاوه بر این $\rho = P/\sigma_n^2$ ، در حالی که P

توان سیگنال ارسالی و σ_n^2 واریانس نویز در گیرنده است.

با استفاده از روابط (۱) و (۲) ظرفیت محرمانگی لحظه‌ای در سیستم CIOD را می‌توان به صورت زیر محاسبه نمود

$$C_s^{\text{CIOD}} = \left[\frac{1}{2} \log_2 \left\{ \frac{1 + \frac{\rho}{2} \left(|h_0^d|^2 + |h_1^d|^2 \right)}{1 + \frac{\rho}{2} \left(|h_0^e|^2 + |h_1^e|^2 \right)} \right\} + \frac{1}{2} \log_2 \left\{ \frac{1 + \frac{\rho}{2} \left(|h_2^d|^2 + |h_3^d|^2 \right)}{1 + \frac{\rho}{2} \left(|h_2^e|^2 + |h_3^e|^2 \right)} \right\} \right]^+. \quad (3)$$

که در آن $h_i^e, (i = 0, \dots, N-1)$ بهره کانال بین i -امین آنتن فرستنده و آنتن شنودگر است.

چون محاسبه روابط بسته برای معیارهای امنیت در سیستم CIOD مشکل است، ما برای محاسبه ظرفیت محرمانگی لحظه‌ای از تقریبی استفاده می‌کنیم که بر این اساس $\log(1 + \text{SNR}) \approx \log(\text{SNR})$. در بخش‌های بعد نشان خواهیم داد، نتایج بدست آمده با این تقریب با روابط بدست آمده از طریق شبیه‌سازی، بخصوص در SNRهای بالا، بسیار نزدیک هستند. علاوه بر این مشابه [10] شرط غیرمنفی بودن ظرفیت محرمانگی لحظه‌ای را حذف می‌کنیم. در [11] نشان داده شده است که این تقریب بر روی احتمال قطع محرمانه اثری نخواهد داشت. با استفاده از توضیحات ارائه شده رابطه (۳) را می‌توان به صورت زیر نوشت:

$$C_s^{\text{CIOD}} \approx \frac{1}{2} \left\{ \log_2 \left[\frac{|h_0^d|^2 + |h_1^d|^2}{|h_0^e|^2 + |h_1^e|^2} \right] + \log_2 \left[\frac{|h_2^d|^2 + |h_3^d|^2}{|h_2^e|^2 + |h_3^e|^2} \right] \right\}. \quad (4)$$

ما در این مقاله، سیستم CIOD را به دلیل افزایش کاربرد، همچنین توانایی در ارسال و دریافت داده با خطای پایین مدنظر قرار داده‌ایم و برای احتمال قطع ارتباط محرمانه در این شبکه روابط تحلیلی بسته‌ای بدست آورده‌ایم. در ادامه جهت افزایش امنیت در این شبکه از روش انتخاب تصادفی آنتن استفاده کرده و چگونگی بکارگیری آن را شرح داده‌ایم. در بخش شبیه‌سازی نیز صحت روابط ارائه شده و عملکرد روش پیشنهادی به کمک شبیه‌سازی بررسی و تصدیق شده است.

بخش‌های بعدی مقاله بدین صورت تنظیم شده است. در بخش ۲ مدل سیستم تحت بررسی به همراه پارامترها و متغیرهای لازم معرفی شده است. در بخش ۳ روابط مربوط به احتمال قطع ارتباط محرمانه بدست آمده است. در بخش ۴ نرخ خطای سمبل به همراه چگونگی انتخاب آنتن جهت ایجاد امنیت بررسی شده است. در بخش ۵ نتایج عددی و شبیه‌سازی‌های مختلفی جهت بررسی روابط و روش پیشنهادی ارائه گردیده و در نهایت نتیجه‌گیری انجام شده است.

۲- مدل سیستم

سیستمی که ما در این مقاله در نظر گرفته‌ایم شامل یک فرستنده (Alice)، با ۴ آنتن فرستنده است که قصد دارد با گیرنده مجاز (Bob) که دارای ۱ آنتن گیرنده است ارتباط برقرار کند. در طرف مقابل شنودگری قرار دارد که او نیز دارای ۱ آنتن گیرنده است و قصد دارد اطلاعات رد و بدل شده را شنود کند. هدف آن است که گیرنده مجاز قادر باشد اطلاعات ارسالی فرستنده را بدون خطا دریافت کند، در حالی که شنودگر در دریافت اطلاعات دچار مشکل است. فرض می‌کنیم بهره کانال بین i -امین آنتن فرستنده و آنتن گیرنده مجاز $h_i^d, (i = 0, \dots, 3)$ باشد، و بهره کانال i -امین آنتن فرستنده و آنتن شنودگر $h_i^e, (i = 0, \dots, 3)$ باشد. همچنین فرض می‌شود که کانال گیرنده مجاز و شنودگر شبه ایستا^۱ و دارای محوشوندگی از نوع تخت^۲ باشد و بهره کانال‌ها دارای توزیع گوسی مختلط با میانگین صفر و واریانس λ_d و λ_e باشد.

فرستنده از یک کدکننده فضا-زمان بر مبنای CIOD و ماتریس کدینگ الموتی استفاده می‌کند. داده کدشده را می‌توان به صورت ماتریس $X = [c_{li}]_{l=0, \dots, 3; i=0, \dots, 3}$ با ابعاد 4×4 نمایش داد که c_{li} سمبل ارسالی از i -امین آنتن در اسلات زمانی l -ام است.

۳- ارزیابی احتمال قطع ارتباط محرمانه

در این بخش ظرفیت محرمانگی و احتمال قطع ارتباط محرمانه در شبکه CIOD بررسی می‌گردد. لازم به ذکر است در سیستم CIOD در گروه اسلات زمانی اول نیمه اول آنتن‌ها ($L/2$ اول) فعال، و در گروه اسلات زمانی دوم نیمه دوم ($L/2$ دوم) فعال هستند. ظرفیت محرمانگی لحظه‌ای در یک شبکه به صورت زیر تعریف می‌گردد [9].

$$C_s = [C_d - C_e]^+, \quad (1)$$

حال تابع چگالی احتمال $Z_1 = \log_2(W)$ را می‌توان با استفاده از رابطه (۱۰) و رابطه مربوط به توزیع تابعی از متغیر تصادفی بدست آورد:

$$f_{Z_1}(z_1) = 2^{z_1} \ln(2) f_W(2^{z_1}) \\ = 2^{2z_1+1} \frac{3 \ln(2) \lambda_e^2 \lambda_d^2}{(\lambda_e 2^{z_1} + \lambda_d)^4} \quad (11)$$

مشابه گام‌هایی که از (۷) تا (۱۱) برداشته شد، می‌توان به سادگی تابع چگالی احتمال $X_2 = |h_2^d|^2 + |h_3^d|^2$ ، $Y_2 = |h_2^e|^2 + |h_3^e|^2$ و در نتیجه $Z_2 = \log_2(X_2/Y_2)$ را بدست آورد که نتیجه مطابق زیر خواهد بود:

$$f_{Z_2}(z_2) = 2^{2z_2+1} \frac{3 \ln(2) \lambda_e^2 \lambda_d^2}{(\lambda_e 2^{z_2} + \lambda_d)^4} \quad (12)$$

که این رابطه مشابه (۱۱) است.

با دانستن توابع چگالی احتمال Z_1 و Z_2 در روابط (۱۱) و (۱۲) در ادامه می‌توانیم رابطه مربوط به احتمال قطع محرمانه را بدست آوریم.

احتمال قطع ارتباط محرمانه، بیانگر احتمال رخداد آن است که ظرفیت محرمانگی لحظه‌ای از نرخ مشخصی که ما آن را با R_t نمایش می‌دهیم، کمتر شود و بنابراین به صورت زیر بیان می‌شود [11]:

$$SOP = P(C_s < R_t) \quad (13)$$

احتمال قطع ارتباط محرمانه برای سیستم CIOD با ۴ آنتن فرستنده به صورت زیر قابل محاسبه است:

$$SOP^{CIOD} = P(C_s^{CIOD} < R_t) \\ = P\left(\frac{1}{2} \left\{ \log_2 \left[\frac{|h_0^d|^2 + |h_1^d|^2}{|h_0^e|^2 + |h_1^e|^2} \right] \right\} \right. \\ \left. + \frac{1}{2} \left\{ \log_2 \left[\frac{|h_2^d|^2 + |h_3^d|^2}{|h_2^e|^2 + |h_3^e|^2} \right] \right\} < R_t \right) \\ = P\left(\frac{1}{2} Z_1 + \frac{1}{2} Z_2 < R_t\right) \quad (14)$$

با جایگذاری توابع چگالی احتمال Z_1 و Z_2 عبارت فوق به صورت زیر می‌گردد:

فرض کنید بهره‌های کانال‌های گیرنده مجاز و شنودگر متغیرهای تصادفی گوسی مختلط، مستقل و با توزیع یکسان باشد. میانگین این متغیرها صفر و واریانس آن‌ها برای کانال گیرنده مجاز و شنودگر به ترتیب برابر λ_d و λ_e فرض می‌شود. در این صورت توزیع $|h_i^d|^2$ و $|h_i^e|^2$ نمایی و به ترتیب با میانگین λ_d و λ_e خواهد بود، که تابع چگالی احتمال آن‌ها را می‌توان به صورت زیر نمایش داد:

$$f_{|h_i^d|^2}(x) = \frac{1}{\lambda_d} e^{-x/\lambda_d}, \quad (5)$$

و

$$f_{|h_i^e|^2}(x) = \frac{1}{\lambda_e} e^{-x/\lambda_e}. \quad (6)$$

فرض کنید $Y_1 = |h_0^e|^2 + |h_1^e|^2$ ، $X_1 = |h_0^d|^2 + |h_1^d|^2$ و $Z_1 = \log_2(X_1/Y_1)$ در حالی که $Z_1 = C_s^{OSTBC}$. بنابراین تابع چگالی احتمال X_1 و Y_1 که جمع دو متغیر تصادفی i.i.d هستند را می‌توان به صورت $f_{X_1}(x_1) = f_{|h_0^d|^2}(x) * f_{|h_1^d|^2}(x)$ و $f_{Y_1}(y_1) = f_{|h_0^e|^2}(x) * f_{|h_1^e|^2}(x)$ به دست آورد در حالی که اپراتور * بیانگر کانولوشن است. نتیجه به صورت زیر خواهد شد:

$$f_{X_1}(x_1) = \int_0^{x_1} \frac{1}{\lambda_d^2} e^{-u/\lambda_d} e^{-(x_1-u)/\lambda_d} du = \frac{x_1 e^{-x_1/\lambda_d}}{\lambda_d^2}, \quad (7)$$

و

$$f_{Y_1}(y_1) = \int_0^{y_1} \frac{1}{\lambda_e^2} e^{-u/\lambda_e} e^{-(y_1-u)/\lambda_e} du = \frac{y_1 e^{-y_1/\lambda_e}}{\lambda_e^2}. \quad (8)$$

برای محاسبه تابع چگالی احتمال $Z_1 = \log_2(X_1/Y_1)$ ، در ابتدا بایستی مطابق زیر تابع توزیع احتمال $W = X_1/Y_1$ را بدست آوریم.

$$F_W(w) = P\left(\frac{X_1}{Y_1} < w\right) \\ = \int_0^\infty \int_0^{wy_1} f_{X_1}(x_1) f_{Y_1}(y_1) dx_1 dy_1 \\ = \int_0^\infty \int_0^{wy_1} \frac{x_1 y_1 e^{-x_1/\lambda_d} e^{-y_1/\lambda_e}}{\lambda_d^2 \lambda_e^2} dx_1 dy_1 \\ = 1 - \frac{3w \lambda_d^2 \lambda_e^2 + \lambda_d^3 \lambda_e}{\lambda_e (\lambda_e w + \lambda_d)^3} \quad (9)$$

سپس تابع چگالی احتمال W را می‌توان به سادگی با مشتق‌گیری از تابع توزیع احتمال آن بدست آورد که نتیجه به صورت زیر است

$$f_W(w) = \frac{d}{dw} F_W(w) = \frac{6w \ln(2) \lambda_e^2 \lambda_d^2}{(\lambda_e w + \lambda_d)^4}. \quad (10)$$

$P_2^i(d_j \rightarrow d_k)$ و $P_1^i(d_j \rightarrow d_k)$ به صورت زیر قابل محاسبه هستند:

$$P_1^i(d_j \rightarrow d_k) = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{\rho_i \kappa_1 \beta_1^2}{4 \sin^2 \theta}\right)^{-\frac{N_T N_R}{2}} \left(1 + \frac{\rho_i \kappa_2 \beta_2^2}{4 \sin^2 \theta}\right)^{-\frac{N_T N_R}{2}} d\theta, \quad (17)$$

و

$$P_2^i(d_m \rightarrow d_n) = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{\rho_i \kappa_1 \beta_1^2}{4 \sin^2 \theta}\right)^{-\frac{N_T N_R}{2}} \left(1 + \frac{\rho_i \kappa_2 \beta_2^2}{4 \sin^2 \theta}\right)^{-\frac{N_T N_R}{2}} d\theta, \quad (18)$$

در حالی که ρ_i بیانگر SNR-i-امین زیرکانال است و κ_t ($t=1,2$) مقیاس ماتریس‌های OSTBC جهت برقراری قید توان است. علاوه بر این در عبارات (17)، $[\beta_1, \beta_2]^T = [d_m - d_n]$ (18) و $[\beta_1, \beta_2]^T = [d_j - d_k]$ می‌باشد.

در CIOD در گروه اسلات زمانی اول، یعنی اولین $N_T/2$ اسلات زمانی، گروه اول آنتن‌ها فعال هستند و گروه اسلات زمانی دوم، یعنی $N_T/2$ اسلات زمانی بعد، گروه دوم آنتن‌ها فعال هستند. در این حالت، اولین روشی که جهت افزایش امنیت به ذهن می‌رسد ارسال نویز مصنوعی است اما این روش علیرغم مرسوم بودن، نیاز به توان مصرفی مضاعف دارد و PAPR^o را نیز تغییر می‌دهد. ما در ادامه روشی را معرفی می‌کنیم که نیازی به توان یا تغییر PAPR نداشته باشد. فرض کنید در ابتدای ارتباط، فرستنده و گیرنده سیگنال‌های راهنمایی را برای یکدیگر ارسال کنند. اگر زمان ارسال این سیگنال‌ها از زمان همدوس‌سازی کانال کمتر باشد، نشانگر کیفیت کانال (CQI¹) محاسبه شده در فرستنده و گیرنده مجاز یکسان خواهد بود. از طرفی چون شنودگر از نظر مکانی متفاوت از آن‌هاست CQI محاسبه شده متفاوتی خواهد داشت. مقدار CQI می‌تواند به عنوان مقدار اولیه یک مولد تصادفی از نوع LFSR^v مورد استفاده قرار گیرد. مقدار تصادفی تولید شده توسط مولد تصادفی سطری از یک جدول از پیش تعریف شده که حاوی تمام اولویت‌های ممکن برای آنتن‌های فرستنده است را انتخاب می‌کند. به عبارت دیگر سطر انتخاب شده مشخص می‌کند در گروه اسلات زمانی فعلی کدام آنتن‌ها بایستی فعال شوند و بالطبع در گروه اسلات زمانی بعدی، گروه بعدی آنتن‌ها فعال خواهند شد. با فعال کردن تصادفی آنتن‌ها، شنودگر حتی با دانستن ضرایب کانال قادر به تشخیص اطلاعات ارسالی نخواهد بود، چون در این حالت شنودگر نمی‌داند کدام یک از ضرایب را باید در تخمین سمبل ارسالی مورد استفاده قرار دهد. این درحالی است که گیرنده مجاز با دانستن مقدار تصادفی و در نتیجه آنتن‌های انتخابی قادر است مانند قبل سمبل‌های ارسالی را بدون مشکل تخمین بزند. با به کارگیری این روش ساده و کم-هزینه می‌توان امنیت سیستم را بالا برد.

فرض کنید Θ مانریس کد CIOD باشد. علاوه بر این W ماتریسی است که حاوی تمامی اولویت‌های ممکن برای اعداد بین ۱ تا N_T در سطرهایش باشد که N_T تعداد آنتن فرستنده است. مقدار

$$\begin{aligned} &= P\left(\frac{1}{2}Z_1 + \frac{1}{2}Z_2 < R_t\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{2R_t - z_2} f_{Z_1}(z_1) f_{Z_2}(z_2) dz_1 dz_2 \\ &= \int_{-\infty}^{\infty} F_{Z_1}(2R_t - z_2) f_{Z_2}(z_2) dz_2 \\ &= 1 - \int_{-\infty}^{\infty} \frac{3\lambda_e \lambda_d^2 2^{2R_t - z_2} + \lambda_d^3}{(\lambda_e 2^{2R_t - z_2} + \lambda_d)^3} \times \frac{6z_2 2^{2z_2} \ln(2) \lambda_e^2 \lambda_d^2}{(\lambda_e 2^{z_2} + \lambda_d)^4} dz_2 \end{aligned} \quad (15)$$

با کمی محاسبات ریاضی انتگرال فوق را می‌توان به صورت زیر محاسبه نمود:

$$\begin{aligned} \text{SOP}^{\text{CIOD}} &= 1 - \left[\frac{\lambda_d^4}{(\lambda_d^2 - \lambda_e^2 2^{2R_t})^6} \times \ln(\lambda_e / \lambda_d) \right. \\ &\quad \times 3\lambda_e^4 2^{4R_t+1} (3\lambda_e^2 2^{4R_t} + 7\lambda_e^2 \lambda_d^2 2^{2R_t+1} + 3\lambda_d^4) \\ &\quad \left. + \frac{\lambda_d^4}{(\lambda_d^2 - \lambda_e^2 2^{2R_t})^6} \left[+3\lambda_e^4 2^{4R_t+1} (\lambda_d^2 - \lambda_e^2 2^{2R_t})^2 \right. \right. \\ &\quad \left. \left. + \frac{2(\lambda_d^2 - 3\lambda_e^2 2^{2R_t})(\lambda_d^3 - \lambda_e \lambda_d^2 2^{2R_t})^3}{\lambda_d^3} \right. \right. \\ &\quad \left. \left. - 3\lambda_e^4 2^{4R_t+1} (\lambda_d^2 - \lambda_e^2 2^{2R_t})(3\lambda_e^2 2^{2R_t} + 5\lambda_d^2) \right. \right. \\ &\quad \left. \left. + 9\lambda_e^4 2^{4R_t+1} (\lambda_d^2 - \lambda_e^2 2^{2R_t})(3\lambda_e^2 2^{2R_t} + \lambda_d^2) \right. \right. \\ &\quad \left. \left. + 3\lambda_e^4 2^{4R_t+1} (3\lambda_e^4 2^{4R_t} + 7\lambda_e^2 \lambda_d^2 2^{2R_t+1} + 3\lambda_d^4) \ln\left(\frac{\lambda_d}{\lambda_e 2^{2R_t}}\right) \right. \right. \\ &\quad \left. \left. - \frac{3(\lambda_d^3 - \lambda_e^3 \lambda_d 2^{2R_t})^2 (9\lambda_e^4 2^{4R_t} - \lambda_d^2 \lambda_e^2 2^{2R_t+1} + \lambda_d^4)}{\lambda_d^2} \right] \right]. \end{aligned} \quad (16)$$

۴- نرخ خطای سمبل و روش پیشنهادی

با توجه به توضیحات ارائه شده، در CIOD بخش‌های حقیقی و موهومی هر سمبل در اسلات زمانی متفاوتی ارسال می‌شود. بنابراین محاسبه مقدار دقیق نرخ خطای سمبل دشوار است اما محاسبه SPEP نسبتاً ساده‌تر است [12]. مشابه آن‌چه در [7]، [12] انجام شده، ما SPEP را برای سیستم CIOD بدست می‌آوریم. فرض کنید d_j ، d_k ، d_m ، d_n ($j, k, m, n = 1, 2, \dots, L$) سیگنال‌های حاصل از منظومه \mathcal{S} با سایز L باشند. با استفاده از روش ارائه شده در [87]، می‌توان باند تجمیع برای نرخ خطای سمبل در سیستم CIOD با منظومه \mathcal{S} را با استفاده از $P_u^i(C) = P_{u,1}^i(C) + P_{u,2}^i(C)$ بدست آورد، در حالی که

$$P_{u,1}^i(C) = \frac{1}{L} \sum_{j=1}^{L-1} \sum_{k=j+1}^L P_1^i(d_j \rightarrow d_k)$$

و

$$P_{u,2}^i(C) = \frac{1}{L} \sum_{m=1}^{L-1} \sum_{n=m+1}^L P_2^i(d_m \rightarrow d_n)$$

پارامترهای

- [11] S. Yan, N. Yang, R. Malaney, and J. Yuan, "Transmit antenna selection with Alamouti coding and power allocation in MIMO wiretap channels," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 3, pp. 1656–1667, 2014.
- [12] D. N. Dào and C. Tellambura, "Decoding, performance analysis, and optimal signal designs for coordinate interleaved orthogonal designs," vol. 7, no. 1, pp. 48–53, 2008.

در شکل (۳)، نرخ خطای سمبل برای CIOD با ۴ آنتن بر حسب SNR آورده شده است. ماتریس کدینگ بر اساس مدل الموتی انتخاب شده است. نتایج تئوری و شبیه‌سازی ارائه شده. همان‌طور که مشاهده می‌شود روش پیشنهادی هیچ اثر مخربی بر عملکرد گیرنده مجاز ندارد اما می‌تواند عملکرد شنودگر را دچار مشکل کند که این کار منجر به بهبود امنیت در سیستم مورد بررسی می‌گردد.

زیرنویس‌ها

- ¹ Orthogonal Space-Time Block Coding
- ² Coordinate Interleaved Orthogonal Designs
- ³ Quasi-static
- ⁴ Flat
- ⁵ Peak-to-average power ratio
- ⁶ Channel quality indicator
- ⁷ Linear feedback shift register

۶- نتیجه‌گیری

در این مقاله، امنیت در سیستم CIOD را بررسی کردیم و با بیان مدل ارسال و دریافت سیگنال، برای احتمال قطع ارتباط محرمانه در این سیستم رابطه تئوری بسته ارائه دادیم. در ادامه از روش انتخاب تصادفی آنتن فرستنده جهت افزایش امنیت در این سیستم استفاده کرده و عملکرد این روش را در بخش شبیه‌سازی بررسی نمودیم. طبق انتظار مشاهده شد که روش پیشنهادی در عین سادگی نرخ خطای سمبل در شنودگر را به شدت افزایش می‌دهد. از طرفی رابطه احتمال قطع ارتباط محرمانه با شبیه‌سازی سیستم و کانال مورد بررسی قرار گرفت. تطابق نتایج تئوری و شبیه‌سازی حاکی از صحت روابط بیان‌شده است. علاوه بر این نشان داده شد که افزایش نسبت بهره کانال گیرنده مجاز به بهره کانال شنودگر، باعث کاهش احتمال قطع ارتباط محرمانه می‌گردد.

مراجع

- [1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [2] T. Allen and A. Tajar, "Reduced-Feedback AN-Aided Secure Alamouti MAC Transmissions," *Glob. Commun. Conf.*, 2016.
- [3] T. Allen, J. Cheng, and N. Al-Dhahir, "Secure space-time block coding without transmitter CSI," *IEEE Wirel. Commun. Lett.*, vol. 3, no. 6, pp. 573–576, 2014.
- [4] H. Lee, J. G. Andrews, R. W. Heath, and E. J. Powers, "The performance of space-time block codes from coordinate interleaved orthogonal designs over Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 653–664, 2009.
- [5] M. Z. A. Khan and B. S. Rajan, "Single-symbol maximum likelihood decodable linear STBCs," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2062–2091, 2006.
- [6] D. N. Dào and C. Tellambura, "Decoding, performance analysis, and optimal signal designs for coordinate interleaved orthogonal designs," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 1, pp. 48–53, 2008.
- [7] D. N. Dào and C. Tellambura, "On Space-Time Block Codes from Coordinate Interleaved Orthogonal Designs," *Proc. - IEEE Mil. Commun. Conf. MILCOM*, no. x, pp. 0–4, 2007.
- [8] Y.-H. K. Y.-H. Kim and M. Kaveh, "Coordinate-interleaved space-time coding with rotated constellation," *57th IEEE Semiannu. Veh. Technol. Conf. 2003. VTC 2003-Spring.*, vol. 1, pp. 732–735 vol.1, 2003.
- [9] J. Barros and M. D. Rodrigues, "Secrecy Capacity of Wireless Channels," *2006 IEEE Int. Symp. Inf. Theory*, vol. 1, pp. 356–360, 2006.
- [10] X. Chen, J. Chen, H. Zhang, Y. Zhang, and C. Yuen, "On Secrecy Performance of Multiantenna-Jammer-Aided Secure Communications with Imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8014–8024, 2016.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود

اعضای کمیته راهبردی

رئیس دانشگاه صنعتی شاهرود	پروفسور محمد مهدی فاتح
معاون پژوهشی دانشگاه صنعتی شاهرود	دکتر علی دستفان
رئیس قطب علمی سامانه‌های پردازش دیجیتال	پروفسور حسن غفوری فرد
دبیر کمیته دائمی کنفرانس	پروفسور کریم فائز
دبیر کنفرانس پنجم	پروفسور حمید حسن پور



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود

اعضای کمیته علمی

دانشگاه تربیت مدرس	پروفسور احسان الله کبیر
دانشگاه صنعتی امیرکبیر	پروفسور حسن غفوری فرد
دانشگاه صنعتی امیرکبیر	پروفسور کریم فائز
دانشگاه صنعتی شاهرود	پروفسور حمید حسن پور
دانشگاه شهید باهنر کرمان	پروفسور حسین نظام آبادی پور
پژوهشگاه ارتباطات و فناوری اطلاعات (مرکز تحقیقات مخابرات ایران)	پروفسور کامبیز بدیع
دانشگاه صنعتی نوشیروانی بابل	پروفسور علی آقاگل زاده
دانشگاه تربیت مدرس	پروفسور محمدحسن قاسمیان
دانشگاه صنعتی شاهرود	پروفسور محمد مهدی فاتح
دانشگاه صنعتی امیرکبیر	دکتر حسن طاهری
دانشگاه گیلان	دکتر اسدالله شاه بهرامی
دانشگاه یزد	دکتر محمد قاسم زاده
دانشگاه گنکور دیا	دکتر نعمت‌الله شیری
دانشگاه فردوسی مشهد	دکتر حیدر طوسی‌یان شان‌دیز
دانشگاه سمنان	دکتر فرزین یغمایی
دانشگاه آزاد اسلامی ساری	دکتر همایون مؤتمنی
دانشگاه مالزی	دکتر والاعلی روحانی
دانشگاه شاهد	دکتر نادر جعفری راد
دانشگاه یزد	دکتر ولی درهمی
دانشگاه صنعتی سجاد	دکتر جواد حمیدزاده
دانشگاه علم و صنعت ایران	دکتر بهروز مینائی
دانشگاه صنعتی شریف	دکتر حمید بیگی
دانشگاه صنعتی شاهرود	دکتر امین روشندل کاهو
دانشگاه صنعتی شاهرود	دکتر علیرضا ناظمی
دانشگاه صنعتی شاهرود	دکتر علی نجاتی کلاته
دانشگاه صنعتی شاهرود	دکتر داود شاهسونی
دانشگاه صنعتی شاهرود	دکتر حسین مروی

دانشگاه صنعتی شاهرود	دکتر علی دستفان
دانشگاه صنعتی شاهرود	دکتر علیرضا احمدی فرد
دانشگاه صنعتی شاهرود	دکتر علی سلیمانی
دانشگاه صنعتی شاهرود	دکتر امیدرضا معروضی
دانشگاه صنعتی شاهرود	دکتر وحید ابوالقاسمی
دانشگاه صنعتی شاهرود	دکتر هادی گرایلو
دانشگاه صنعتی شاهرود	دکتر سید مسعود میررضایی
دانشگاه صنعتی شاهرود	دکتر هدی مشایخی
دانشگاه صنعتی شاهرود	دکتر فاطمه جعفری نژاد
دانشگاه صنعتی شاهرود	دکتر مرضیه رحیمی
دانشگاه صنعتی شاهرود	دکتر محسن بیگلری

Prof. Firuz Zare
Prof. A. Beghdadi
Dr. M. Mahdavi
DR. Amin Zehabian
Dr. Amin Hajizadeh
Dr. Pooya Davari

University of Queensland
University Paris 13, France
Victoria University Sydney
Free University of Berlin
Aalborg University
Aalborg University



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۷ الی ۲۸ آذرماه ۱۳۹۸



دانشگاه صنعتی شاهرود

فهرست الفبایی نویسندگان

صفحه	نویسنده
	ا
۲۱۴.....	آقالری، مطهره
۲۰۸.....	اسدی، علیرضا
۱۳۸.....	الوندی، احد
	ب
۷۸.....	بابائی، مرضیه
۲۲۰.....	بقایی‌پور، علیرضا
	پ
۱۴۵.....	پرکوک، سحر
۱۸۸.....	پیوندی، کیمیا
	ت
۱۷۵.....	تاجیک، فاطمه
	ج
۸۵.....	جهانی سه قعله، علی
	ح
۱۱۷.....	حورعلی، سمیرا
	خ
۱۷۰.....	خدابخشی، زینب
	د
۷۲.....	دهقان سلماسی، نفیسه
	ر
۱۵۷.....	رستمی، زهرا
۱۲۹.....	روحانی، نفیسه
	ز

۱۶۴	زرگری، حامد
	س
۲۰	سرمدی، پدram
۴۵.۳۹	سلطانی، آزاده
۱۲۳	سوادى حسينى، مهدى
	ش
۲۷	شجاعى، بهنام
	ع
۱۰۳	عزيزى لارى، عبدالعزيز
۳۳	عطارى، سبحان
	ف
۱۸۱	فاتح، هاله
۱۱۰	فيروزيان، ايمان
	ق
۶۵	قاضى زاهدى، امين
	م
۱۵۲	مرتضايى، زهرا
۹۶	موسوى آذرنگ، سيد حسين
	ن
۹۱	نادرى اسرمى، زهره
۱۹۴	نامجوى، امير
	ی
۲۰۱	ياوند حسنى، مجتبى

