

روشی مبتنی بر ماشین یادگیری سریع با هسته غیرخطی برای انتخاب نمونه‌های اولیه در یادگیری چندبرچسبه مقیاس بزرگ

حمید کارگر شورکی^۱ و محمدعلی زارع چاهوکی^۲

چکیده

با وجود حجم عظیم محتوای چندرسانه‌ای در وب، ذخیره‌سازی و بازیابی آن‌ها با بکارگیری روش‌های یادگیری موجود با محدودیت‌هایی از جمله کمبود حافظه مواجه شده است. تاثیر گذاری محدودیت‌های مدنظر در روش‌های یادگیری دارای مرحله آموزش مانند ماشین بردار پشتیبان (SVM) و شبکه‌های عصبی تا جایی است که امکان بکارگیری این روش‌ها در کاربردهای مقیاس بزرگ تقریباً غیرممکن است. روش ماشین یادگیری سریع مبتنی بر هسته غیرخطی (KELM) یکی از روش‌های قدرتمند ارائه شده در حوزه یادگیری ماشین است. اساس مرحله یادگیری در این روش مبتنی بر ساخت ماتریس هسته‌ی نمونه‌های برچسب‌دار و محاسبه معکوس آن می‌باشد. از این رو، بکارگیری این روش در محیط‌های مقیاس بزرگ با وجود تعداد زیاد نمونه‌های برچسب دار امکان‌پذیر نیست. در این پژوهش به منظور حل مشکل مطرح شده در به‌کارگیری KELM در کاربردهای مقیاس بزرگ، روشی مبتنی بر انتخاب نمونه‌های اولیه با بهره‌گیری از KELM در مقیاس کوچک همسایگی هر نمونه آموزش ارائه شده است. با به‌کارگیری روش انتخاب نمونه‌های اولیه ارائه شده، حجم مجموعه آموزش کاهش می‌یابد. بنابراین امکان استفاده از روش یادگیری KELM در کاربردهای مقیاس بزرگ فراهم می‌شود. از آنجایی که کاربردهای حوزه چندرسانه‌ای وب به صورت چندبرچسبه می‌باشند، روش ارائه شده در انتخاب نمونه‌های اولیه، مبتنی بر کاربردهای چندبرچسبه مانند شرح‌گذاری خودکار تصاویر است. نتایج آزمایش‌های تجربی بر روی دادگان چندبرچسبه مقیاس بزرگ NUS-WIDE و نسخه‌های آن مانند Object، Lite و Scene بیانگر کارایی روش ارائه شده در حل محدودیت‌های بکارگیری KELM در کاربردهای چندبرچسبه مقیاس بزرگ با انتخاب نمونه‌های اولیه دارد.

کلیدواژه‌ها

ماشین یادگیری سریع، انتخاب نمونه‌های اولیه، یادگیری چندبرچسبه مقیاس بزرگ، شرح‌گذاری خودکار تصاویر

حال افزایش است. از طرفی به منظور ساماندهی، طبقه‌بندی و بازیابی رسانه‌های رقومی در کاربردهای واقعی مانند بازیابی چندرسانه‌ای، بیوانفورماتیک، پزشکی و صنعت نیاز به بکارگیری روش‌های یادگیری ماشین در پیش‌بینی برچسب و طبقه‌بندی موثر آن‌ها می‌باشد.

به منظور طبقه‌بندی و ساماندهی محتوای چندرسانه‌ای، روش‌های یادگیری ماشین متعددی ارائه شده است. روش‌های یادگیری قدرتمندی از جمله SVM، شبکه‌های عصبی و روش‌های مبتنی بر نمونه‌ها از جمله این روش‌ها می‌باشند. اساس کار این روش‌ها یافتن تابع نگاشتی از روی نمونه‌های موجود به

۱ مقدمه

امروزه با وجود رشد سریع استفاده از اینترنت و بکارگیری ابزارهای رقومی، حجم محتوای تولیدشده این حوزه روزبه‌روز در

این مقاله در مهرماه سال ۱۳۹۴ دریافت، در مردادماه سال ۹۵ بازنگری و در شهریورماه همان سال پذیرفته شد.

^۱ دانشجوی کارشناسی ارشد دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد.

رایانامه: hamidkargar@stu.yazd.ac.ir

^۲ دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد،

رایانامه: chahooki@yazd.ac.ir

نویسنده مسئول: محمدعلی زارع چاهوکی

مجموعه اصلی انتخاب شده و ماتریس هسته زیرمجموعه آموزش جدید ساخته می‌شود. عمده ایراد این روش، انتخاب تصادفی نمونه‌ها و در نظر گرفتن تاثیر برابر برای هر نمونه انتخاب شده در زیرمجموعه اصلی می‌باشد. انتخاب تصادفی هر نمونه منجر به حذف برخی نمونه‌های با ارزش و در نتیجه افت عملکرد روش طبقه‌بندی می‌شود.

در این مقاله، ما محدودیت حافظه در به‌کارگیری روش KELM در کاربردهای چندبرچسب مقیاس بزرگ را با استفاده از ایده انتخاب نمونه‌های اولیه^۲ (PS) بهبود داده‌ایم. فرض اصلی روش‌های انتخاب نمونه‌های اولیه، کاهش نمونه‌های مجموعه آموزش و حذف نمونه‌های بی‌تاثیر در فرآیند طبقه‌بندی با به‌کارگیری قاعده نزدیک‌ترین همسایه‌ها می‌باشد [۱۰]. از طرفی رویکرد حل محدودیت این پژوهش، به‌کارگیری یک روش انتخاب نمونه اولیه چندبرچسب با عملکرد خوب است. از اینرو ایده ما، ارائه روشی ترکیبی در انتخاب نمونه‌های اولیه در یادگیری چندبرچسب مبتنی بر KELM^۳ (MLPS-KELM) و استفاده از این روش در حل محدودیت حافظه KELM در کاربردهای مقیاس بزرگ است.

در بکارگیری روش‌های انتخاب نمونه‌های اولیه در کاربردهای چندبرچسب، مرحله تصمیم‌گیری حضور و یا عدم حضور نمونه‌های دارای برچسب در زیرمجموعه کاهش‌یافته از اهمیت بالایی برخوردار است. تصمیم‌گیری حضور و یا عدم حضور یک نمونه کاندید حذف در این روش‌ها بر اساس مقایسه بردار برچسب پیش‌بینی شده از روی K نزدیکترین همسایه نمونه کاندید حذف با بردار برچسب واقعی آن انجام می‌شود. بکارگیری روش سنتی در رای‌گیری حداکثری^۴ از K نمونه همسایه منجر به ساخت بردار برچسب غیر دقیق خواهد شد. از اینرو استفاده از یک روش یادگیری قدرتمند در این مرحله منجر به پیش‌بینی بردار برچسب دقیق از روی K نزدیکترین همسایه و بهبود عملکرد روش انتخاب نمونه اولیه خواهد شد. از آنجایی که روش یادگیری KELM دارای عملکرد قدرتمندی در پیش‌بینی برچسب‌ها می‌باشد [۵]، در مرحله تصمیم‌گیری روش ارائه‌شده برای انتخاب نمونه‌های اولیه از KELM استفاده شده است.

از طرفی اساس تصمیم‌گیری حضور و یا عدم حضور یک نمونه در زیرمجموعه کاهش‌یافته در روش‌های انتخاب نمونه اولیه در کاربرد چندبرچسب، مبتنی بر مقایسه سطح عدم شباهت بردار برچسب واقعی هر نمونه کاندید و بردار برچسب پیش‌بینی شده از همسایگی‌های آن نمونه می‌باشد. از آنجاکه بردار برچسب نمونه‌های چندبرچسب برداری صفر و یک است، این مقایسه بر اساس

برچسب‌های مفهومی آن‌ها و پیش‌بینی برچسب نمونه جدید با استفاده از تابع نگاشت یادگیری شده است [۱].

با وجود تعداد زیاد نمونه‌های موجود در کاربردهای واقعی، ابعاد بالای بردار ویژگی استخراج‌شده از نمونه‌ها، تعداد زیاد برچسب‌های مفهومی و خاصیت چند برچسب بودن برخی از کاربردها مانند طبقه‌بندی محتوای چندرسانه‌ای، روش‌های یادگیری با چالش‌ها و محدودیت‌های متعددی مواجه می‌شوند. از جمله محدودیت‌هایی که با وجود حجم و تعداد زیاد نمونه‌ها و ابعاد بالای بردار ویژگی به وجود می‌آید، محدودیت حافظه و زمان یادگیری است. همچنین عدم پیش‌بینی مناسب بردار برچسب‌های مفهومی چالش دیگر روش‌های یادگیری در مواجهه با کاربردهای چند برچسب و تعداد بالای برچسب‌ها می‌باشد [۲].

به‌منظور غلبه بر محدودیت زمان یادگیری در شبکه‌های عصبی روش ماشین یادگیری سریع^۱ (ELM) ارائه شده است [۳] و [۴]. در ELM محدودیت زمان یادگیری، با مقداردهی تصادفی به نرون‌های لایه پنهان حل شده است. نسخه هسته و چند کلاسه روش ELM در [۵] ارائه شده و صلاحیت و کارایی به‌کارگیری آن در کاربردهای مختلف نسبت به سایر روش‌های یادگیری از جمله نسخه عادی ELM، شبکه‌های عصبی و نسخه‌های متعدد SVM با انجام آزمایش‌های گسترده‌ای به اثبات رسیده است.

اساس مرحله یادگیری در KELM مبتنی بر محاسبه ماتریس هسته نمونه‌های برچسب‌دار آموزش و محاسبه معکوس این ماتریس برای بدست آوردن وزن لایه خروجی می‌باشد. از آنجایی که در کاربردهای مقیاس بزرگ تعداد نمونه‌های برچسب‌دار آموزش بسیار زیاد است، مرحله یادگیری روش KELM با چالش اساسی محدودیت حافظه و زمان یادگیری مواجه می‌شود. به عبارت دقیق‌تر با وجود تعداد زیاد نمونه‌های آموزش و همچنین ابعاد بالای نمونه‌ها، نگهداری ماتریس هسته ساخته شده در حافظه امکان‌پذیر نیست. از اینرو بکارگیری این روش قدرتمند در کاربردهای مقیاس بزرگ، سخت و در مواردی تقریباً غیرممکن است [۶].

چالش محدودیت حافظه در ساخت ماتریس هسته با وجود تعداد زیاد نمونه‌ها و ویژگی‌ها، در [۶]–[۹] مورد بررسی قرار گرفته است. رویکردهای مطرح شده در [۷]–[۹] مبتنی بر تخمین فضای هسته می‌باشد. در [۷] و [۸] نمونه‌های فضای هسته به یک فضای ویژگی تخمینی با ابعاد پایین‌تر نگاشت شده و مدل‌های خطی ساده در فضای نگاشت شده بکار گرفته می‌شوند. در [۹] نیز ایده مشابهی بکار گرفته شده و حالت شکل هسته با رویکردی فازی تخمین زده می‌شود.

رویکرد مطرح شده در [۶] مبتنی بر انتخاب تعدادی نمونه تصادفی از مجموعه آموزش اولیه می‌باشد. به عبارت دقیق‌تر در ابتدا زیرمجموعه‌ای شامل تعداد نمونه‌های کمتر از نمونه‌های

²Prototype Selection (PS)

³ Multi-Label Prototype Selection based on KELM (MLPS-KELM)

⁴Majority voting

¹Extreme Learning Machine (ELM)

روش‌های PS بر اساس رویکردهای مختلفی دسته‌بندی می‌شوند. این رویکردها شامل نحوه بدست آوردن زیرمجموعه نمونه‌های اولیه، نحوه جستجوی نمونه‌های اولیه و ارزیابی جستجو می‌باشد [۱۰]. روش‌های PS بر اساس رویکرد بدست آوردن زیرمجموعه نمونه‌های اولیه به سه دسته کلی تنک‌سازی^۲، اصلاح‌سازی^۳ و ترکیبی^۴ تقسیم‌بندی می‌شوند. هدف روش‌های تنک‌سازی حفظ نمونه‌های نزدیک به مرزهای طبقه‌بندی و حذف نمونه‌های دورتر از این مرزها است. از معروف‌ترین روش‌های تنک‌سازی می‌توان به CNN [۱۳]، MCNN [۱۴] و FCNN [۱۵] اشاره کرد.

تمرکز اصلی روش‌های اصلاح‌سازی، بهبود دقت طبقه‌بندی است. این روش‌ها با حذف نمونه‌های نویزی سعی دارند با کاهش حجم مجموعه آموزش دقت طبقه‌بندی و رگرسیون را افزایش دهند. از معروف‌ترین روش‌های اصلاح‌سازی می‌توان به ENN [۱۶]، AIKNN [۱۷] و RNGE [۱۸] اشاره کرد. روش انتخاب نمونه اولیه ENN بکارگرفته شده در این پژوهش از دسته روش‌های اصلاح‌سازی است.

در نهایت روش‌های ترکیبی با بهره‌گیری از ایده دودسته قبل سعی در یافتن زیرمجموعه کمینه نمونه‌های اولیه به همراه افزایش دقت طبقه‌بندی دارند. از معروف‌ترین روش‌های ترکیبی می‌توان به SSMA [۱۹] و RMHC [۲۰] اشاره کرد.

دومین رویکرد دسته‌بندی، نحوه جستجوی نمونه‌های اولیه است. روش‌های PS بر اساس این رویکرد به پنج گروه افزایشی^۵، کاهشی^۶، دسته‌ای^۷، تلفیقی^۸ و ثابت^۹ تقسیم‌بندی می‌شوند. در رویکرد افزایشی، فرآیند انتخاب نمونه‌ها با زیرمجموعه خالی شروع می‌شود. در هر مرحله نمونه‌ای که دارای شرایط خاص هر روش باشد به زیرمجموعه اضافه می‌شود. از آنجایی که انتخاب هر نمونه به صورت تصادفی است، روش‌های دارای این رویکرد ترتیب محور^{۱۰} هستند [۱۰]. از جمله معروف‌ترین روش‌های دارای این رویکرد می‌توان به CNN [۱۳]، FCNN [۱۵] و MCNN [۱۱] اشاره کرد.

در رویکرد کاهشی، فرآیند انتخاب نمونه‌ها با زیرمجموعه‌ای شامل تمام نمونه‌های آموزش آغاز می‌شود. در هر مرحله، نمونه‌ای که دارای شرایط خاص تعریف شده باشد، از زیرمجموعه حذف می‌شود. روش انتخاب نمونه اولیه ENN بکارگرفته شده در این پژوهش دارای این رویکرد می‌باشد. دیگر روش دارای این رویکرد RNGE [۱۵] است.

محاسبه عدم شباهت دودویی و در نظر گرفتن مقدار آستانه^۱ برای عدم شباهت صورت می‌گیرد [۱۱]. ایده مادر بخش مقایسه دو بردار برچسب در فرآیند انتخاب نمونه‌های اولیه، بررسی اثربخشی بکارگیری روابط عدم شباهت دودویی متعددی است.

بنابراین به طور کلی نوآوری‌های ارائه شده در این مقاله شامل: (۱) کاربردی‌سازی روش KELM در یادگیری چندبرچسب مقیاس بزرگ مبتنی بر انتخاب نمونه‌های اولیه، (۲) بهبود و ارائه یک روش انتخاب نمونه‌های اولیه چند برچسب جدید با بکارگیری روش پیش‌بینی KELM در مرحله تصمیم‌گیری حضور و یا عدم حضور نمونه‌ها و (۳) اثربخشی بکارگیری روابط عدم شباهت دودویی متعدد در مرحله تصمیم‌گیری حضور هر نمونه می‌باشد. عملکرد مطلوب رویکرد ارائه شده در حل محدودیت حافظه در بکارگیری روش KELM در طبقه‌بندی و همچنین انتخاب نمونه‌های اولیه مناسب، با انجام آزمایش‌های متعدد بر روی دادگان مقیاس بزرگ NUS-WIDE و سه دادگان NUS-WIDE-Object، NUS-WIDE-Lite و SNUS-WIDE-Scene در شرح‌گذاری خودکار مجموعه تصاویر در مقیاس بزرگ نشان داده شده است. نتایج آزمایش‌های صورت گرفته نشان از صلاحیت بکارگیری رویکرد ارائه شده در کاربردهای طبقه‌بندی و انتخاب نمونه‌های اولیه چندبرچسب مقیاس بزرگ دارد.

در ادامه این مقاله بخش ۲ به مروری بر کارهای مرتبط صورت گرفته اختصاص دارد. در بخش ۳ الگوریتم‌های پایه‌ای ELM و KELM ارائه می‌شود. بخش ۴ به ارائه رویکرد ارائه شده در این پژوهش می‌پردازد. در بخش ۵ نتایج آزمایش‌های صورت گرفته در این پژوهش را مورد بررسی قرار می‌دهیم. در نهایت در بخش ۶ به بیان جمع‌بندی و نتیجه‌گیری از پژوهش صورت گرفته خواهیم پرداخت.

۲ مروری بر پژوهش‌های پیشین

در این بخش مروری بر پژوهش‌های مرتبط پیشین خواهیم داشت. بدین ترتیب در بخش ۲-۱ پژوهش‌های مرتبط با حوزه انتخاب نمونه‌های اولیه ارائه شده است. همچنین، بخش ۲-۲ به پژوهش‌های صورت گرفته در حوزه کاربرد ELM و KELM در یادگیری چندبرچسب مقیاس بزرگ اختصاص یافته است.

۲-۱ انتخاب نمونه‌های اولیه

انتخاب نمونه‌های اولیه (PS) فرآیند یافتن زیرمجموعه‌ای بهینه و کمینه از داده‌های اصلی با استفاده از قاعده نزدیک‌ترین همسایه‌ها است. زیرمجموعه کاهش یافته به‌عنوان مجموعه آموزش جدید در روش‌های یادگیری مورد استفاده قرار می‌گیرد. دسته‌بندی روش‌های انتخاب نمونه‌های اولیه در [۷] و [۹] ارائه شده است.

²Condensation

³Edition

⁴Hybrid

⁵Incremental

⁶Decremental

⁷Batch

⁸Mixed

⁹Fixed

¹⁰Order dependent

¹Threshold

نمونه‌های باارزش و در نظر گرفتن ارزش یکسان برای هر نمونه می‌باشد. در انتخاب زیرمجموعه نمونه‌ها، یافتن زیرمجموعه بهینه از اهمیت بالایی برخوردار بوده و منجر به بهبود عملکرد روش یادگیری می‌شود.

از روش ELM و نسخه هسته آن در کاربردهای چند برچسب مختلفی استفاده شده است [۲۴] تا [۳۰]. پژوهش‌های این حوزه با رویکردهای مختلف با مسائل چند برچسب برخورد کرده‌اند. تغییر روند یادگیری و تبدیل الگوریتم تک کلاس به چند کلاس [۵]، [۲۵]، [۲۶]، [۲۹]، [۳۰] و همچنین تبدیل مسئله چند برچسب به چند کلاس و به‌کارگیری الگوریتم پایه [۲۷] و [۲۸] رویکردهای اصلی در به‌کارگیری روش ELM در کاربرد چند برچسب می‌باشد. در [۳۰] روش ELM با رویکرد یکی در برابر یکی^۳ (OAO) و یکی در برابر همه^۴ (OAA) در تشخیص چهره بکار برده شده است. در این دو رویکرد هر مسئله به تعدادی مسئله دو کلاس تبدیل شده و سپس روش ELM عادی بر روی آن اعمال می‌شود. در رویکرد OAA به تعداد کلاس‌ها، ELM یادگیری شده و سپس به ازای هر نمونه آزمایش بالاترین خروجی به‌عنوان کلاس مدنظر برگردانده می‌شود. از طرفی در OAO به ازای هر دو کلاس با یکدیگر یک ELM یادگیری شده و خروجی بالاتر نشان‌دهنده کلاس مدنظر یک نمونه آزمایش می‌باشد. در [۳۱] نیز رویکرد OAA در برخورد با مسئله چند برچسب برای تشخیص محتوای مفهومی ویدیو بکار گرفته شده است.

دسته دیگری از روش‌ها با تغییر در روند یادگیری در ELM به ازای هر نمونه آزمایش رتبه^۵ مربوط به هر کلاس را به دست آورده و با تعریف یک مقدار آستانه، مقادیر رتبه بالاتر از مقدار آستانه را به‌عنوان برچسب نمونه آزمایش در نظر می‌گیرند [۲۵]، [۲۶]، [۲۹]، [۳۰]. در [۲۹] با تبدیل مسئله به یک محیط جدید و تغییر فرآیند یادگیری به نرم-۱ حداقل مربعات^۶، رتبه مربوط به هر نمونه آزمایش را به ازای هر کلاس بدست آورده و سپس با در نظر گرفتن مقدار آستانه تعریف‌شده، برچسب هر نمونه را مشخص می‌کند.

رویکرد مواجهه با مسائل چند برچسب در [۳۳] و [۳۲] بر اساس به‌کارگیری نسخه چند کلاس ارائه‌شده در [۵] و بدست آوردن رتبه نمونه‌های آزمایش و در نظر گرفتن مقدار آستانه در برچسب‌گذاری می‌باشد. مقدار آستانه در نظر گرفته‌شده در [۳۲] بر اساس یک مقدار ثابت از پیش تعریف‌شده و در [۳۳] بر اساس محاسبه مقادیر ویژه ماتریس رتبه خروجی و در نظر گرفتن مقادیر بزرگ‌تر از میانگین به‌عنوان برچسب نمونه آزمایش می‌باشد. در نظر گرفتن مقدار آستانه ثابت منجر به افت کارایی در برچسب‌گذاری نمونه‌های چند برچسب می‌شود. از این‌رو در این

در روش‌های دسته‌ای حذف نمونه‌ها به‌صورت دسته‌ای بوده و فرآیند حذف چندین نمونه در یک زمان معین صورت می‌گیرد. POP [۲۱] و AIKNN [۱۷] از معروف‌ترین روش‌های دارای این رویکرد می‌باشند.

تلفیقی و ثابت دو رویکرد دیگر نحوه جستجوی نمونه‌های اولیه است. روش‌های دارای رویکرد تلفیقی با یک زیرمجموعه تصادفی شامل چندین نمونه شروع به کار می‌کنند و در هر مرحله نمونه‌های دارای شرایط خاص می‌توانند به زیرمجموعه کاهش یافته اضافه شوند. روش‌های SSMA [۱۹] و CoCoIS [۲۲] از این رویکرد استفاده می‌کنند. در نهایت در رویکرد ثابت که زیرمجموعه تلفیقی می‌باشد، تعداد حذف و اضافه‌ها مشابه و در ابتدای مرحله اجرایی روش مشخص می‌شود. معروف‌ترین روش دارای این رویکرد، RMHC [۲۰] است.

در نهایت روش‌های PS از لحاظ ارزیابی جستجوی نمونه‌ها با استفاده از قاعده نزدیک‌ترین همسایه‌ها به دو دسته فیلتر^۱ و بسته‌بندی^۲ تقسیم‌بندی می‌شوند. در فیلتر قاعده نزدیک‌ترین همسایه‌ها روی زیرمجموعه جزئی از داده‌ها اعمال می‌شود. در صورتی که در بسته‌بندی عملیات یافتن نمونه‌های اولیه روی کل مجموعه داده‌هاست. به عبارت دقیق‌تر در فیلتر ابتدا مجموعه آموزش بر اساس رویکردی خاص شکسته شده، انتخاب نمونه‌ها صورت گرفته و سپس نمونه‌های انتخاب‌شده با هم ادغام می‌شوند. در صورتی که در بسته‌بندی فرآیند تقسیم و شکستن مجموعه آموزش وجود ندارد.

۲-۲ ELM و KELM در یادگیری چندبرچسب مقیاس بزرگ

همان‌طور که بیان شد، لازمه‌ی به‌کارگیری نسخه هسته روش‌های یادگیری ساخت ماتریس هسته نمونه‌های برچسب‌دار آموزش است. از آنجاکه در کاربردهای مقیاس بزرگ تعداد نمونه‌های آموزش بسیار زیاد است، نگهداری آن‌ها و ساخت ماتریس هسته با محدودیت‌های حافظه و پیچیدگی زمانی عملیات ماتریسی از قبیل محاسبه معکوس، ضرب و محاسبه مقادیر ویژه مواجه می‌شود. این محدودیت‌ها در روش‌های K SVM و KELM با ارائه راهکارهای مختلفی از جمله الگوریتم‌های تکرارشونده در حل مسئله بهینه‌سازی [۲۳]، تجزیه ماتریس هسته به ماتریس‌هایی با ابعاد کمتر [۲۴] و انتخاب نمونه‌های تصادفی و ساخت زیرمجموعه‌ای از مجموعه اصلی [۶]، [۲۲] و [۲۳] مورد بررسی قرار گرفته است. راهکارهای انتخاب نمونه و ساخت زیرمجموعه نمونه‌های آموزش، بدون هیچ‌گونه تغییری در روند یادگیری سعی در کاهش حجم مجموعه آموزش و کاهش پیچیدگی محاسباتی دارند. عمده ایراد انتخاب تصادفی نمونه‌ها، عدم توجه به

³One-against-one (OAO)

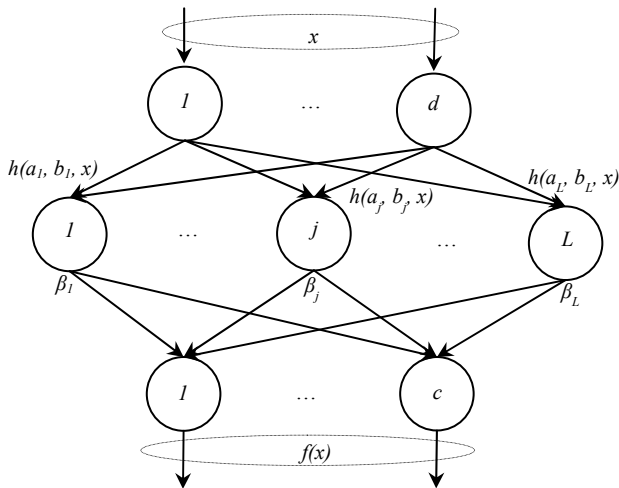
⁴One-against-all (OAA)

⁵Score

⁶l-norm least square

¹Filter

²Wrapper



شکل ۱ شمای کلی یک ELM با L نرون لایه پنهان.

در رابطه ۲، $H \in R^{N \times L}$ و $Y \in R^{N \times 1}$ به ترتیب ماتریس خروجی لایه پنهان و ماتریس برجسبها می باشد.

در فرآیند یادگیری در شبکه های تک لایه، پارامترهای شبکه به صورت تکرارشونده و با روش گرادینان نزولی بدست آمده و به روزرسانی می شوند. در صورتی که در ELM تعداد نرون های لایه پنهان، L و مقادیر پارامترها به صورت تصادفی مقداردهی می شود. به عبارت دقیق تر ELM یک رویکرد یادگیری بدون تنظیم^۳ در شبکه های پس انتشار را به اجرا می گذارد [۲۷]. از این رو فرآیند یادگیری به حل و بهینه سازی یک تابع خطی تبدیل شده و پیچیدگی محاسباتی در آن به شدت کاهش می یابد.

رویکرد یادگیری ELM در کمینه سازی خطای یادگیری و نرم ماتریسی وزن های خروجی در [۵] ارائه شده است. در این رویکرد مسئله بهینه سازی به صورت رابطه ۳ تعریف می شود.

$$\min_{\beta} \frac{1}{2} \|\beta\|_2^2 + \frac{C}{2} \sum_{i=1}^N \|\xi_i\|^2 \quad (3)$$

$$s.t. h(X_i)\beta = y_i^T - \xi_i^T \quad i = 1 \dots N$$

در رابطه ۳، ξ_i و y_i به ترتیب خطای تخمین و بردار برجسب نمونه آموزش X_i می باشد. همچنین C ضریب کنترل^۴ است. بر اساس نظریه KKT^۵ وزن های خروجی (β) با حل بهینه رابطه ۴ بدست می آید.

$$L_{ELM} = \frac{1}{2} \|\beta\|_2^2 + \frac{C}{2} \sum_{i=1}^N \|\xi_i\|^2 - \sum_{i=1}^N (h(X_i)\beta - y_i^T + \xi_i^T)\alpha_i \quad (4)$$

در رابطه ۴، $\alpha_i = [\alpha_{i1}, \dots, \alpha_{iL}] \in R^{1 \times L}$ پارامترهای نرون های لایه پنهان می باشد. بردار وزن های خروجی β با مشتق گیری از این رابطه نسبت به α_i ، ξ_i و β مساوی صفر قرار دادن بدست می آید. نسخه

پژوهش و در گام طبقه بندی نمونه های چند برجسب از یک رویکرد آستانه گذاری^۱ تطبیقی در برجسب گذاری نمونه ها استفاده شده است. رویکرد به کاررفته بر اساس ویژگی های دادگان از قبیل میانگین تعداد برجسبها عملیات برجسب گذاری را انجام می دهد.

۳ الگوریتم های پایه ای ELM و KELM

روش ELM [۳] و [۴] اولین بار به منظور غلبه بر مشکلات و محدودیت های به کارگیری روش شبکه های پس انتشار^۲ (BP) ارائه شد. دو ویژگی متمایز این روش نسبت به روش یادگیری BP، تعیین تعداد و مقداردهی تصادفی به نرون های لایه پنهان می باشد. به عبارت دقیق تر پارامترهای نرون های لایه پنهان در ابتدا و به صورت مستقل از نمونه های آموزش مقداردهی شده و فرآیند یادگیری تکرارشونده در شبکه های BP به یک فرآیند یادگیری تک مرحله ای تبدیل می شود. با به کارگیری این رویکرد، محدودیت های محاسباتی ELM در مواجهه با داده های مقیاس بزرگ کاهش قابل توجهی دارد. این روش در مهر و موم های اخیر مورد توجه بسیاری از پژوهشگران حوزه یادگیری قرار گرفته و نسخه های متعددی از آن در کاربردهای مختلف از جمله پردازش تصویر و بینایی ماشین [۳۰] تجزیه و تحلیل محتوای ویدئو [۳۱] و پزشکی [۳۴] ارائه و بکار گرفته شده است. نسخه های ارائه شده ELM از جمله I-ELM [۳۵]، EM-ELM [۳۶] و OS-ELM [۳۷] سعی در رفع محدودیت ها و به کارگیری این روش در کاربردهای مختلف یادگیری داشته اند.

با مجموعه آموزش داده شده $R^{N \times d}$ شامل N نمونه آموزش $Trn = \{(X_i, Y_i)\}_{i=1}^N$ ، تابع خروجی یک شبکه عصبی با L نرون لایه پنهان به صورت رابطه ۱ بیان می شود.

$$f(X_i) = \sum_{j=1}^L \beta_j h_j(\alpha_j, \beta_j, X_i) = h(X_i)\beta, i = 1 \dots N \quad (1)$$

که $\beta = [\beta_1, \beta_2, \dots, \beta_L]^T \in R^{1 \times L}$ بردار وزن های لایه خروجی، $h(X_i) = [h_1(\alpha_1, \beta_1, X_i), \dots, h_L(\alpha_L, \beta_L, X_i)] \in R^{1 \times L}$ اساس هر نمونه X_i و $h_j(\cdot)$ تابع تبدیل غیرخطی لایه پنهان مانند RBF می باشد. همچنین a_j و b_j پارامترهای نرون زام لایه پنهان است. شکل ۱ شمای کلی یک ELM را با L نرون لایه پنهان نشان می دهد.

فرآیند یادگیری در شبکه شامل کمینه سازی خطای نرم $\|H\beta - Y\|_2$ است که در آن:

$$H = \begin{bmatrix} h(X_1) \\ \dots \\ h(X_N) \end{bmatrix}, Y = \begin{bmatrix} y_1^T \\ \dots \\ y_N^T \end{bmatrix} \quad (2)$$

³Tuning free

⁴Regularization factor

⁵Karush-Kuhn-Tucker (KKT)

¹Thresholding

²Back Propagation (BP)

KELM در مجموعه‌های چندبرچسب مقیاس بزرگ با روش انتخاب نمونه‌های اولیه ارائه شده در این پژوهش مرتفع شده است. از این رو در بخش ۴-۱ روش‌های پایه‌ای مورد استفاده در روش انتخاب نمونه‌های اولیه ارائه شده در این پژوهش را شرح می‌دهیم. در بخش ۴-۲ به بیان گام‌های مختلف رویکرد MLPS-KELM ارائه شده در این پژوهش پرداخته می‌شود.

۴-۱ ENN و ENN چندبرچسب

روش اصلاح‌سازی نزدیک‌ترین همسایه^۲ (ENN) [۱۶] در سال ۱۹۷۲ توسط ویلسون^۳ ارائه شد. این روش از دسته روش‌های اصلاح‌سازی است و از رویکرد جستجوی کاهشی استفاده می‌کند. همانطور که در بخش ۲-۱ بیان شد، هدف عمده روش‌های اصلاح‌سازی حذف نمونه‌های نویز مجموعه و افزایش دقت و کاهش حجم هم‌زمان است.

فرض کنید $X = \{x_1, x_2, \dots, x_d\} \in R^{1 \times d}$ نمونه‌ای d بعدی است که در آن x_i مقدار ویژگی i ام نمونه X است. از طرفی نمونه X دارای بردار برچسب $Y = \{y_1, y_2, \dots, y_l\} \in R^{1 \times c}$ می‌باشد و $y_i = l$ نشان دهنده تعلق برچسب l ام به نمونه X است. مجموعه داده D شامل دو بخش آموزش Trn و آزمایش Tst است و S ، زیرمجموعه کاهش یافته Trn پس از اعمال روش PS است.

روش ENN فرآیند یافتن زیرمجموعه S را با $S = Trn$ آغاز می‌کند. به ازای هر نمونه X_i از S ، در صورتی که بردار برچسب Y_i با بردار برچسب اکثریت K نمونه همسایه‌اش برابر نباشد، حذف می‌شود. این روند یک بار و به ازای تمام نمونه‌های آموزش انجام می‌شود. در انتهای این فرآیند نمونه‌های باقیمانده در S نمونه‌های اولیه هستند و می‌توان از S به عنوان مجموعه آموزش کاهش یافته جدید در فرآیند یادگیری استفاده کرد.

به منظور بهبود عملکرد الگوریتم‌های طبقه‌بندی بر روی مجموعه‌های چند برچسب نامتعادل، رویکرد ENN چندبرچسب^۴ (MLENN) ارائه شده است [۱۱]. در این رویکرد به منظور تعادل‌سازی مجموعه‌ها از روش ENN استفاده می‌شود. فرض این رویکرد در تعادل‌سازی مجموعه داده‌های نامتعادل، شناسایی و حفظ نمونه‌های عدم تعادل در مجموعه است. بدین منظور در ابتدا تمام این نمونه‌ها شناسایی شده و سپس با استفاده از اعمال ENN روی مابقی نمونه‌ها، مجموعه‌ها متعادل می‌شوند. نمونه‌هایی عدم تعادل با دو معیار $LbIIIR$ و $DSIR$ شناسایی می‌شوند. معیار $LbIIIR$ نشان‌دهنده میزان عدم تعادل هر برچسب است. رابطه ۷ نحوه محاسبه این معیار را نشان می‌دهد. در رابطه ۷، γ برچسبی خاص از مجموعه برچسب‌های Y ، و Y_i بردار برچسب نمونه i ام است. مقدار $LbIIIR$ بالاتر نشان‌دهنده میزان عدم تعادل بالاتر است. از

عادی و هسته بدست آمده برای β به ترتیب در روابط ۵ و ۶ نشان داده شده است.^۱

$$\beta \quad (5)$$

$$\begin{aligned} \beta &= H^T \left(\frac{1}{C} I + H^T H \right)^{-1} Y \\ &= H^T \left(\frac{1}{C} I + K_M \right)^{-1} Y \end{aligned} \quad (6)$$

در روابط ۵ و ۶، I ماتریس همسانی و $K_M = \{K_M(X_i, X_j)\} = \{h(X_i) \cdot h(X_j)\}$ ، $i, j = 1 \dots N$ ماتریس هسته نمونه‌های آموزش می‌باشد.

از آنجایی که عبارت $H^T \left(\frac{1}{C} I + K_M \right)^{-1}$ یک ماتریس $N \times N$ می‌باشد، پیچیدگی محاسبه معکوس این عبارت از مرتبه $O(N^3)$ است. برای مثال در مجموعه با ۱۰۰۰۰۰ نمونه آموزش و با فرض اجرای هر دستور در زمان واحد به $(100000 \times 100000 \times 100000)$ زمان برای محاسبه معکوس ماتریس نیاز است. با این پیچیدگی بالا و تعداد زیاد نمونه‌ها محاسبه معکوس ماتریس عملی نیست. از این رو کاهش تعداد نمونه‌های آموزش، N ، محاسبه معکوس عبارت مدنظر را عملی می‌سازد.

پس از بدست آمدن ماتریس وزن‌های خروجی β در مرحله آموزش، رتبه هر کلاس به ازای هر نمونه آزمایش با استفاده از رابطه ۱ بدست می‌آید. سپس در مورد نحوه برچسب‌گذاری تصمیم‌گیری می‌شود. در مسائل تک کلاسه بزرگ‌ترین رتبه و در مسائل چند کلاسه و چند برچسب بر اساس یک آستانه خاص عملیات برچسب‌گذاری صورت می‌گیرد.

صلاحیت به‌کارگیری نسخه هسته روش ELM در مقایسه با سایر الگوریتم‌های یادگیری از جمله نسخه‌های متعدد SVM (مانند LS-SVM و P-SVM) و شبکه‌های عصبی تک لایه در [۵] نشان داده شده است. همچنین بر اساس نتایج آزمایش‌های ارائه شده، KELM دارای دقت طبقه‌بندی بهتری نسبت به ELM می‌باشد.

۴ رویکرد MLPS-KELM ارائه شده

در این بخش رویکرد MLPS-KELM ارائه شده در این پژوهش ارائه می‌شود. رویکرد ارائه شده، سعی در بهبود روش انتخاب نمونه‌های اولیه ENN چند برچسب با به‌کارگیری روش پیش‌بینی KELM در مرحله تصمیم‌گیری خودش دارد. از طرفی با به‌کارگیری روش انتخاب نمونه‌های اولیه و کاهش حجم نمونه‌های آموزش، به‌کارگیری روش KELM در محیط‌های چندبرچسب مقیاس بزرگ کاربرد سازی شده است. به عبارت دیگر، محدودیت حافظه روش

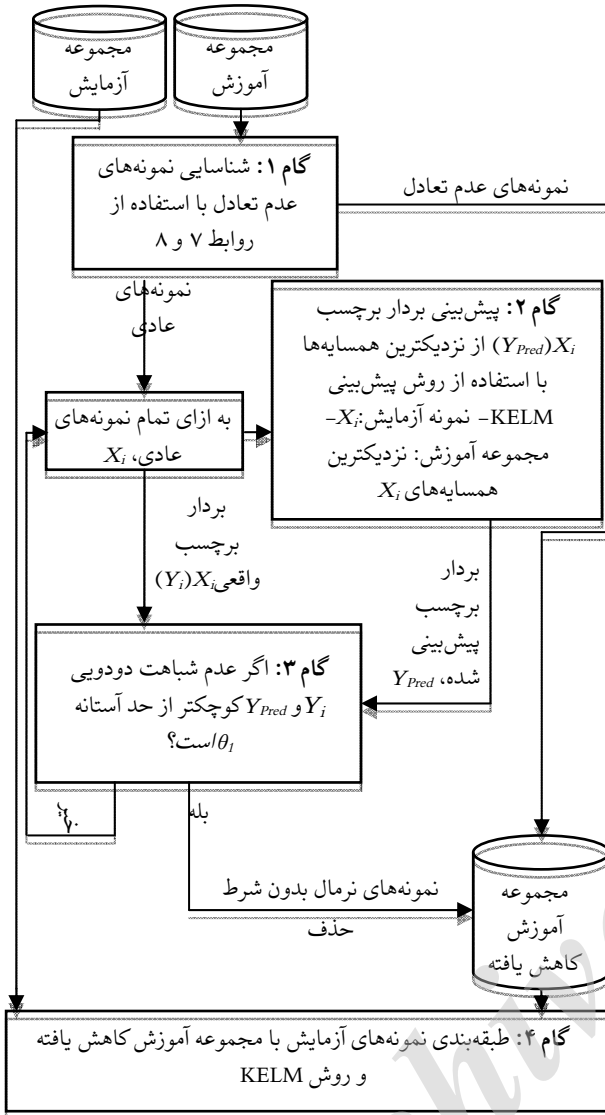
²Edited Nearest Neighbor (ENN)

³Wilson

⁴Multi-Label ENN

^۱جزئیات بیشتر در مورد نحوه محاسبات صورت گرفته و بدست آمدن هر رابطه در [۵].

می‌دهد. در ادامه این بخش هر یک از این گام‌ها مورد بررسی قرار خواهند گرفت.



شکل ۲ شمای کلی گام‌های اصلی روش MLPS-KELM ارائه شده.

گام ۱: شناسایی و نگهداری نمونه‌های منجر به عدم تعادل مجموعه آموزش: مجموعه آموزش و آزمایش مقیاس بزرگ شامل تعداد بالای نمونه‌های چند برچسب به الگوریتم MLPS-KELM داده می‌شود. در گام اول، نمونه‌های آموزشی عدم تعادل مجموعه شناسایی می‌شوند. در هر مجموعه چند برچسب با افزایش تعداد نمونه‌ها و افزایش تعداد برچسب‌ها سطح تعادل مجموعه کاهش می‌یابد [۱۱]. در این مجموعه‌ها تعدادی از نمونه‌ها منجر به عدم تعادل شده و نبود این نمونه‌ها سطح عدم تعادل مجموعه را بالا می‌برد. در نتیجه دقت به‌کارگیری هر روش طبقه‌بندی چند برچسب را کاهش می‌دهند (جزئیات در بخش ۴-۱ بیان شده است). از این‌رو بایستی این نمونه‌ها شناسایی شده و به مجموعه کاهش یافته اضافه شود. در این گام با استفاده از روابط ۷ و ۸ این نمونه‌ها شناسایی شده و به زیرمجموعه کاهش یافته اضافه می‌شوند. به عبارت دیگر زیرمجموعه کاهش یافته پیش از انجام هرگونه عملیاتی شامل این نمونه‌ها خواهد بود.

طرفی $DSIR$ نشان‌دهنده میزان عدم تعادل یک مجموعه داده است و با میانگین‌گیری از تمام مقادیر $LbIIR$ بدست می‌آید. رابطه ۸ نحوه محاسبه این معیار را نشان می‌دهد.

$$LbIIR(y) = \frac{y'_c = y_1 \cdot \argMax(\sum_{i=1}^{size(D)} k(y', Y_i))}{\sum_{i=1}^{size(D)} k(y, Y_i)} \quad (7)$$

$$k(y', Y_i) = \begin{cases} 1 & \text{if } y \in Y_i \\ 0 & \text{otherwise} \end{cases}$$

$$DSIR = \frac{\sum_{y=y_1}^{y_c} (LbIIR(y))}{Size(Y)} \quad (8)$$

ایده اصلی [۱۱] در تعادل‌سازی مجموعه‌ها، حفظ نمونه‌های دارای برچسبی با مقدار $LbIIR$ بزرگتر از مقدار $DSIR$ است. این نمونه‌ها منجر به عدم تعادل مجموعه شده و حذف آن‌ها دقت طبقه‌بندی را تحت تاثیر قرار می‌دهد. سپس به منظور تعادل‌سازی، الگوریتم ENN بر روی مابقی نمونه‌ها اعمال شده و نمونه‌هایی که شرط ENN را نداشته باشند حذف می‌شود. از آنجاکه کاربرد MLENN در مجموعه‌های چند برچسب می‌باشد، در مرحله تصمیم‌گیری در مورد حضور و یا عدم حضور یک نمونه از یک رابطه عدم شباهت دودویی و یک مقدار آستانه از پیش تعریف شده استفاده می‌شود. به این ترتیب روش ENN چند برچسب ارائه شده است.

در مجموعه‌های چند برچسب با افزایش تعداد نمونه‌ها و برچسب‌ها و مقیاس بزرگ شدن مجموعه، سطح عدم تعادل افزایش می‌یابد. از آنجایی که تمرکز اصلی رویکرد MLPS-KELM ارائه شده در این پژوهش مبتنی بر کاربردهای چند برچسب مقیاس بزرگ می‌باشد، از اینرو در گام اول رویکرد ارائه شده، با استفاده از دو معیار $DSIR$ و $LbIIR$ نمونه‌های عدم تعادل شناسایی و حفظ می‌شوند. سپس روش انتخاب نمونه‌های اولیه ارائه شده بر روی مابقی نمونه‌ها اعمال می‌شود. روش انتخاب نمونه‌های اولیه ارائه شده در این پژوهش مبتنی بر ENN چند برچسب با مرحله پیش‌بینی برچسب نمونه کاندید حذف با استفاده از KELM می‌باشد. همچنین در مرحله مقایسه بردار برچسب واقعی و پیش‌بینی شده، از روابط محاسبه عدم شباهت دودویی مختلف استفاده می‌شود.

۴-۲ MLPS-KELM

روش MLPS-KELM ارائه شده در این پژوهش شامل ۴ گام اصلی می‌باشد. سه گام اول شامل فرآیند کاهش حجم و انتخاب نمونه‌های اولیه با استفاده از روش ENN چند برچسب ارائه شده است. از طرفی گام آخر فرآیند طبقه‌بندی نمونه‌های آزمایش را با استفاده از مجموعه آموزش کاهش یافته صورت می‌دهد. شکل ۲ شمای کلی گام‌های روش MLPS-KELM ارائه شده را نشان

می‌توان در مورد شباهت و یا عدم شباهت دو بردار برچسب تصمیم‌گیری نمود. به منظور محاسبه عدم شباهت دودویی روابط متعددی ارائه شده است [۳۸]. در مرحله تصمیم‌گیری حضور و یا عدم حضور یک نمونه در روش MLPS-KELM ارائه شده، اثربخشی روابط متعدد محاسبه عدم شباهت دودویی ارائه شده در شکل ۲، تفاوت الگو^۴ و وتر^۵ مورد بررسی قرار گرفته است. این روابط بر اساس کیفیت محاسبه عدم شباهت مورد مقایسه قرار گرفته‌اند. برخی از این روابط مانند فاصله همینگ دارای منطق محاسبه ساده ولی دقت محاسبه پایین می‌باشند. برخی دیگر به مانند فاصله وتر^۵ دارای منطق پیچیده‌تر و دقت بالاتری هستند. فرض کنید Y_i و Y_j دو بردار برچسب دودویی دو نمونه i و j باشند. همچنین بر اساس تعریف قرار می‌دهیم: A تعداد دودویی‌های $Y_i=1$ و $Y_j=1$ ، B تعداد دودویی‌های $Y_i=1$ و $Y_j=0$ ، C تعداد دودویی‌های $Y_i=0$ و $Y_j=1$ و D تعداد دودویی‌های $Y_i=0$ و $Y_j=0$. به عبارت دقیق‌تر، $A=Y_i^T Y_j$ ، $B=Y_i^T (I-Y_j)$ ، $C=(I-Y_i)^T Y_j$ و همچنین $D=(I-Y_i)^T (I-Y_j)$. روابط ۹ تا ۱۴ به ترتیب نحوه محاسبه عدم شباهت‌های همینگ، تغییرات، تفاوت اندازه، تفاوت شکل، تفاوت الگو و وتر^۵ را نشان می‌دهد.

$$\frac{B+C}{A+B+C+D} \quad (9)$$

$$\frac{B+C}{4 \times (A+B+C+D)} \quad (10)$$

$$\frac{(B+C)^2}{(A+B+C+D)^2} \quad (11)$$

$$\frac{(A+B+C+D) \times (B+C) - (B-C)^2}{(A+B+C+D)^2} \quad (12)$$

$$\frac{4 \times B \times C}{(A+B+C+D)^2} \quad (13)$$

$$\sqrt{2 \times \left(1 - \frac{A}{\sqrt{(A+B)(A+C)}}\right)} \quad (14)$$

در میان روابط بررسی شده، بهترین رابطه از لحاظ دقت محاسبه انتخاب شده است. در مورد نحوه انتخاب بهترین رابطه در بخش نتایج تجربی بحث شده است.

گام ۴: طبقه‌بندی نمونه‌های بدون برچسب با مجموعه آموزش کاهش یافته: همانطور که بیان شد سه گام اول روش MLPS-KELM ارائه شده، فرآیند کاهش حجم مجموعه آموزش را صورت می‌دهد. به عبارت دقیق‌تر فرآیند انتخاب نمونه‌های اولیه و

گام ۲: پیش‌بینی برچسب نمونه‌های عادی از K نزدیک‌ترین همسایه‌های آن‌ها: پس از شناسایی نمونه‌های عدم تعادل در گام قبل، تصمیم‌گیری در مورد حضور و یا عدم حضور نمونه‌های باقیمانده در زیرمجموعه کاهش یافته صورت می‌گیرد. بدین ترتیب در ابتدا به ازای هر نمونه آموزش باقیمانده، K نزدیک‌ترین همسایه‌های آن در بین همه نمونه‌های آموزش یافت می‌شود. سپس بردار برچسب نمونه کاندید حذف ساخته می‌شود. در سایر روش‌های سنتی مانند ENN و MLENN ساخت بردار برچسب نمونه کاندید حذف بر اساس رای‌گیری حداکثری صورت می‌گیرد. رای‌گیری حداکثری منجر به پیش‌بینی بردار برچسب بی‌کیفیت خواهد شد. از اینرو ایده اصلی این گام در روش MLPS-KELM پیشنهادی، بکارگیری روش KELM در پیش‌بینی برچسب نمونه کاندید حذف از K نزدیک‌ترین همسایه‌هایش می‌باشد. اساس کار بکارگیری روش KELM در این گام، فرض پیش‌بینی دقیق بردار برچسب نمونه‌ها می‌باشد. بدین ترتیب روش KELM با پارامترهای تنظیم شده یادگیری شده و بردار برچسب نمونه کاندید حذف پیش‌بینی می‌شود. پیش‌بینی چندبرچسب بکار گرفته شده در این بخش بر اساس رابطه آستانه‌گذاری مطرح شده در گام ۴ می‌باشد. عمده تفاوت بکارگیری روش KELM در این بخش و بخش طبقه‌بندی (گام ۴)، تنظیم پارامترهای مجدد این روش در این گام می‌باشد. نحوه تنظیم پارامترهای هر گام در بخش نتایج تجربی مورد بررسی قرار خواهد گرفت.

گام ۳: مقایسه بردار برچسب پیش‌بینی شده و بردار برچسب واقعی: همانطور که قبلاً بیان شد، در مسائل چندبرچسب به ازای هر نمونه، بردار برچسبی نشان‌دهنده برچسب‌های نمونه وجود دارد. در این بردار تعلق و یا عدم تعلق یک برچسب به یک نمونه با صفر و یک نشان داده می‌شود. اساس کار روش‌های پایه‌ای استفاده شده در MLPS-KELM در تصمیم‌گیری حضور و یا عدم حضور یک نمونه مبتنی بر قاعده نزدیک‌ترین همسایه‌ها می‌باشد. در این تصمیم‌گیری برچسب نمونه با برچسب پیش‌بینی شده از نزدیک‌ترین همسایه‌هایش مقایسه شده و در صورت عدم شباهت، نمونه از مجموعه حذف می‌شود [۸] و [۱۳]. از اینرو مسئله مهم در بکارگیری این روش‌ها در داده‌های چند برچسب، مقایسه بردار برچسب هر نمونه $X_i \in \mathcal{S}$ با بردار برچسب پیش‌بینی شده K نزدیک‌ترین همسایه‌های آن است.

در مقایسه برچسب نمونه‌های تک برچسب تنها دو حالت ۰ (کاملاً مخالف) و ۱ (کاملاً موافق) پیش می‌آید. ولی در نمونه‌های دارای چند برچسب در نظر گرفتن مقادیر ۰ و ۱ سخت‌گیرانه است. در مقایسه بردار برچسب این نمونه‌ها درجه‌ای از عدم شباهت وجود دارد و مقدار عدم شباهت در بازه ۰ و ۱ متغیر است. بنابراین به منظور مقایسه بردار برچسب نمونه‌های چند برچسب و محاسبه عدم شباهت بایستی از روابط محاسبه عدم شباهت دودویی استفاده شود. با بکارگیری این روابط درجه عدم شباهت دو بردار برچسب بدست آمده و با تعیین یک آستانه

¹Vari

²Size difference

³Shape difference

⁴Pattern difference

⁵Chord

همسایه‌های نمونه‌کاندید حذف است که به صورت تجربی تعیین می‌شود. از طرفی θ_1 مقدار آستانه تصمیم‌گیری سطح عدم شباهت دو بردار برچسب و θ_2 مقدار آستانه برچسب‌گذاری بر اساس رتبه‌ها در رابطه ۱۵ است. مقدار θ_1 و θ_2 به صورت تجربی و در بازه ۰ تا ۱ انتخاب می‌شود. در مورد انتخاب مقدار بهینه آن‌ها در بخش نتایج تجربی بحث خواهد شد. در نهایت سه پارامتر KT ، C و KP مربوط به روش طبقه‌بندی KELM بوده و به صورت تجربی تنظیم می‌شود. خروجی Tst الگوریتم نیز مجموعه نمونه‌های آزمایش پس از عملیات برچسب‌گذاری می‌باشد.

الگوریتم MLPS-KELM

ورودی

Trn : مجموعه آموزش برچسب‌دار مقیاس بزرگ

Tst : مجموعه آزمایش بدون برچسب مقیاس بزرگ

K : تعداد نزدیکترین همسایه‌ها

θ_1 : مقدار آستانه در تصمیم‌گیری عدم شباهت

C : ضریب کنترل

KT : تابع هسته

KP : پارامتر تابع هسته

θ_2 : مقدار آستانه برچسب‌گذاری بر اساس رتبه‌ها

خروجی

Tst : مجموعه آزمایش برچسب‌گذاری شده

روند اصلی

گام ۱:

۱- محاسبه $LbIR$ برچسب‌های مجموعه Trn با استفاده از رابطه ۱

۲- محاسبه $DSIR$ مجموعه Trn با استفاده از رابطه ۲

۳- به ازای همه نمونه‌های X_i در Trn

۴- اگر X_i دارای برچسبی با مقدار $LbIR > DSIR$ است

۵- X_i به S اضافه می‌شود (زیرمجموعه کاهش یافته)

گام ۲:

۶- در غیر اینصورت

۷- یافتن K نزدیکترین همسایه‌های نمونه X_i در Trn

۸- $Y_{Pred} = KELM(X_i, Trn, C, KT, KP)$

گام ۳:

۹- اگر عدم تشابه چندبرچسبی Y_i و Y_{Pred} کوچکتر از θ_1 است

۱۰- X_i به S اضافه می‌شود

گام ۴:

۱۱- به ازای همه نمونه‌های X_i در Tst

۱۲- $SC_i = KELM(X_i, S, C, KT, KP)$

۱۳- به ازای همه $Score_i$ ‌های نمونه Z در SC_i

۱۴- برچسب‌گذاری نمونه Z با استفاده از بردار SC_i و رابطه ۹

شکل ۳ شبه کد الگوریتم روش MLPS-KELM ارائه شده.

همان‌طور که پیش‌تر بیان شد، عدم تعادل مجموعه‌های چندبرچسبه مقیاس بالا زیاد است. در رویکرد دوفازی مطرح‌شده، در ابتدا نمونه‌های عدم تعادل به زیرمجموعه کاهش یافته اضافه می‌شود. سپس در فاز دوم نمونه‌هایی که طبقه‌بندی آن‌ها برای طبقه‌بند سخت است، به زیرمجموعه کاهش یافته اضافه می‌شود. از آنجایی که تصمیم‌گیری در مورد نمونه‌ها در فاز دوم بدون حضور نمونه‌های با فرکانس پایین صورت می‌گیرد، ممکن است

کاهش حجم مجموعه آموزش در سه گام اول صورت می‌گیرد. پس از فرآیند انتخاب نمونه‌های اولیه، عملیات طبقه‌بندی نمونه‌های بدون برچسب با استفاده از روش KELM صورت می‌گیرد. فرض اصلی بکارگیری روش KELM در این گام صلاحیت عملکرد آن نسبت به سایر نسخه‌های ارائه شده از ELM و الگوریتم‌های یادگیری می‌باشد. از طرفی یکی از رویکردهای اصلی در این مقاله حل مشکلات حافظه و پیچیدگی روش KELM در کاربردهای چند برچسبه مقیاس بزرگ می‌باشد. در روش MLPS-KELM ارائه شده در دو مرحله پیش‌بینی برچسب نمونه‌کاندید حذف و مرحله طبقه‌بندی نمونه‌های بدون برچسب از روش KELM استفاده شده است.

به منظور طبقه‌بندی نمونه‌های بدون برچسب، در ابتدا الگوریتم KELM (توصیف شده در بخش ۳) با استفاده از مجموعه آموزش کاهش یافته و پارامترهای تنظیم‌شده یادگیری شده و ماتریس وزن‌های خروجی بدست می‌آید. سپس مقدار رتبه نمونه‌های بدون برچسب، با ضرب ماتریس فاصله هسته آن‌ها در ماتریس ضرایب β بدست می‌آید. در نهایت عملیات برچسب‌گذاری نمونه‌های بدون برچسب بر اساس رتبه‌های بدست آمده صورت می‌گیرد. روش‌های متعددی به منظور تصمیم‌گیری چند برچسبه ارائه شده است [۱]. این روش‌ها بر اساس رویکردهای آستانه‌گذاری عملیات برچسب‌گذاری را انجام می‌دهند. در نظر گرفتن مقدار آستانه ثابت منجر به افت کارایی و عملکرد روش‌های طبقه‌بندی چندبرچسبه می‌شود. در این بخش از روش MLPS-KELM ارائه شده، از یک رابطه آستانه‌گذاری تطبیقی استفاده شده است. فرض کنید بردار رتبه بدست آمده به ازای یک نمونه بدون برچسب به صورت $SC_i = [score_1, score_2, score_3, \dots, score_n] \in R^{1 \times C}$ باشد. این بردار شامل c مقدار رتبه متناظر با c برچسب می‌باشد. رابطه آستانه‌گذاری بکار رفته به منظور ساخت بردار برچسب دودویی نمونه X_i ، رتبه‌هایی را که θ_2 درصد مجموع همه رتبه‌ها باشند برابر یک قرار داده و مابقی را صفر در نظر می‌گیرد. رابطه ۱۵ فرآیند آستانه‌گذاری تطبیقی مورد استفاده را نشان می‌دهد.

$$\begin{cases} Y_i(j) = 1 & \text{if } SC_i(j) \geq \theta_2 \times \text{sum}(SC_i) \\ Y_i(j) = 0 & \text{Otherwise} \end{cases} \quad (15)$$

با بکارگیری این رابطه و تنظیم θ_2 بر اساس مشخصه‌های هر مجموعه داده از قبیل تعداد برچسب‌ها، برچسب‌گذاری نمونه‌های بدون برچسب صورت می‌گیرد. خروجی نهایی الگوریتم MLPS-KELM مجموعه نمونه‌های برچسب‌گذاری شده می‌باشد. شکل ۳ شبه کد الگوریتم روش MLPS-KELM ارائه شده را نشان می‌دهد.

همان‌طور که بیان شد این الگوریتم فرآیند انتخاب نمونه‌های اولیه و طبقه‌بندی را به صورت هم‌زمان انجام می‌دهد. در صورتی که پس از گام ۳ می‌توان خروجی S را دریافت کرده و به عنوان ورودی هر روش یادگیری در نظر گرفت. این الگوریتم دارای ۸ ورودی و یک خروجی است. دو مجموعه Trn و Tst مجموعه‌های آموزش و آزمایش ورودی می‌باشند. ورودی K نشان دهنده تعداد نزدیکترین

مجموعه‌ها، ۵ دسته ویژگی سراسری محاسبه شده است. این ۵ دسته ویژگی عبارت‌اند از: (۱) هیستوگرام^۶ رنگ LAB ۶۴ بعدی، (۲) ممان رنگ^۷ مبتنی بر بلاک ۲۲۵ بعدی، (۳) کورولوگرام^۸ رنگ HSV ۱۴۴ بعدی، (۴) هیستوگرام جهت لبه ۷۳ بعدی و (۵) موجک بافت^۹ ۱۲۸ بعدی. همچنین به ازای هر تصویر دسته ویژگی سراسری کیسه کلمات دیداری^{۱۰} ۵۰۰ بعدی نیز محاسبه شده است [۳۹]. در این پژوهش به‌مانند [۳۷] و [۳۸] به‌منظور نمایش هر تصویر از سه دسته ویژگی کورولوگرام رنگ HSV، هیستوگرام جهت لبه و موجک بافت استفاده شده است. از این رو هر تصویر با یک بردار ۳۴۵ بعدی نمایش داده می‌شود.

۲-۵ معیارهای ارزیابی

در این بخش به معرفی معیارهای ارزیابی مورد استفاده در شرح‌گذاری تصاویر و انتخاب نمونه‌های اولیه می‌پردازیم. **ارزیابی نتایج شرح‌گذاری مجموعه تصاویر:** از جمله معیارهای معروف مورد استفاده در ارزیابی مسائل طبقه‌بندی، دقت^{۱۱}، فراخوان^{۱۲} و F1 می‌باشد. اگر TP تعداد نمونه‌های مثبت درست مثبت پیش‌بینی شده، TN تعداد نمونه‌های منفی درست منفی پیش‌بینی شده، FP تعداد نمونه‌های منفی به اشتباه مثبت پیش‌بینی شده و FN تعداد نمونه‌های مثبت به اشتباه منفی پیش‌بینی شده باشد، مقادیر دقت، فراخوان و F1 با استفاده از رابطه‌های ۱۶ تا ۱۸ بدست می‌آید.

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (18)$$

در مسائل چند برچسب به ازای همه برچسب‌ها این معیارها محاسبه شده و میانگین آن‌ها به‌عنوان معیار ارزیابی نهایی در نظر گرفته می‌شود. برای نمونه میانگین دقت تمام برچسب‌ها (MAP) محاسبه به‌عنوان معیار ارزیابی در نظر گرفته می‌شود.

در میانگین‌گیری معیارهای ارزیابی مسائل چندبرچسب از دو رویکرد آماری میانگین‌گیری میکرو^{۱۳} و ماکرو^{۱۴} استفاده می‌شود. در میانگین‌گیری ماکرو، ابتدا بر اساس مقادیر TP، TN، FP و FN معیار ارزیابی مورد نظر هر برچسب به‌صورت جداگانه محاسبه

زیرمجموعه کاهش‌یافته نیز دچار عدم تعادل شود. یک راه اجتناب از این حالت حضور کل نمونه‌ها در فاز دوم الگوریتم است [۱۱]. به‌عبارت‌دیگر، در فاز اول ابتدا نمونه‌های عدم تعادل علامت‌گذاری می‌شود. این نمونه‌ها در فاز دوم و در تصمیم‌گیری حذف سایر نمونه‌ها شرکت داده شده و در انتهای فرآیند حذف همه نمونه‌ها به زیرمجموعه کاهش‌یافته اضافه می‌شود. با انجام این کار تعادل مابین کلاس‌ها در زیرمجموعه کاهش‌یافته تضمین می‌شود.

۵ نتایج تجربی

در این بخش نتایج ارزیابی اعمال روش MLPS-KELM ارائه شده بر روی چندین مجموعه داده چندبرچسب مقیاس بزرگ ارائه شده است. از آنجایی‌که مسئله شرح‌گذاری خودکار تصاویر، مسئله‌ای چند برچسب و چالش برانگیز است، به‌منظور نشان دادن صلاحیت روش ارائه شده از مجموعه تصاویر مسائل شرح‌گذاری خودکار تصاویر در مقیاس بزرگ استفاده شده است^۱. بدین منظور در بخش ۵-۱ مجموعه تصاویر مورد استفاده مورد بررسی قرار می‌گیرند. در بخش ۲-۵ معیارهای ارزیابی به‌کاررفته در این پژوهش ارائه می‌شوند. بخش ۳-۵ به تنظیم پارامترهای الگوریتم MLPS-KELM اختصاص دارد. در نهایت در بخش ۴-۵ نتایج ارزیابی روش MLPS-KELM ارائه و تحلیل خواهند شد.

۵-۱ دادگان مورد استفاده

به‌منظور ارزیابی عملکرد روش MLPS-KELM ارائه شده از مجموعه تصویر مقیاس بزرگ NUS-WIDE و نسخه‌های متعدد آن استفاده شده است [۳۹]. این مجموعه شامل ۴ نسخه NUS-WIDE-Object، NUS-WIDE-Scene، NUS-WIDE-Lite و NUS-WIDE می‌باشد. تصاویر و برچسب‌های این مجموعه‌ها از سایت فلیکر^۲ جمع‌آوری و به‌صورت دستی شرح‌گذاری مجدد شده است. جدول ۱ ویژگی‌های مجموعه تصاویر استفاده شده در این پژوهش را نشان می‌دهد.

جدول ۱: ویژگی‌های مجموعه تصاویر مورد استفاده در این پژوهش.

	NUS	Lite	Scene	Object
تعداد کل تصاویر	۲۶۹۶۴۸	۵۵۶۱۵	۳۴۹۲۶	۳۰۰۰۰
تعداد تصاویر آموزش	۱۶۱۷۸۹	۲۷۸۰۷	۱۷۴۶۳	۱۷۹۲۸
تعداد تصاویر آزمایش	۱۰۷۸۵۹	۲۷۸۰۸	۱۷۴۶۳	۱۲۰۷۲
تعداد برچسب‌ها	۸۱	۸۱	۳۳	۳۱

در این مجموعه‌ها به ازای هر تصویر دو دسته بردار ویژگی محلی^۳ و سراسری^۴ محاسبه شده است^۵. به ازای هر تصویر این

^۱ روش ارائه شده به صورت کلی دارای قابلیت اعمال در مسائل چندبرچسب مقیاس بزرگ می‌باشد.

^۲ www.flickr.com

^۳ Local

^۴ Global

^۵ <http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm>

^۶ Histogram

^۷ Color Moment

^۸ Correlogram

^۹ Wavelet Texture

^{۱۰} Bag of visual words

^{۱۱} Precision

^{۱۲} Recall

^{۱۳} Micro averaging

^{۱۴} Macro averaging

عدم شباهت دودویی ارائه و مورد بررسی قرار می‌گیرند. در آزمایش‌های صورت گرفته در این بخش، به منظور تعیین خط پایه تنظیم پارامترها از دو روش طبقه‌بندی استفاده شده است. در آزمایش‌های تنظیم پارامترهای مربوط به انتخاب نمونه‌های اولیه از قبیل تنظیم تعداد نزدیکترین همسایه‌ها (K) و θ_1 از روش KNN ($K=1$) به عنوان خط پایه استفاده شده است. استدلال اصلی این کار فرض بکارگیری روش KNN ($K=1$) پس از انتخاب نمونه‌های اولیه می‌باشد. از طرفی در تنظیم پارامترهای مربوط به روش KELM، از این روش به عنوان خط پایه تصمیم‌گیری استفاده شده است. همچنین در آزمایش‌های صورت گرفته در بخش بعد، نتایج اعمال این دو روش بر روی هر مجموعه تصویر ارائه شده است.

۵-۳-۱- بررسی اثربخشی بکارگیری روابط عدم شباهت دودویی و توابع هسته مختلف در روش MLPS-KELM

به منظور بررسی اثربخشی به کارگیری توابع هسته مختلف در KELM، آزمایشی بر روی مجموعه تصویر NUS-WIDE-Object صورت گرفته است. به منظور قابل اجرا بودن این آزمایش بر روی مجموعه Object در ابتدا فرآیند انتخاب نمونه‌های اولیه با پارامترهای ثابت پیش فرض توابع هسته مختلف صورت می‌گیرد. سپس روش KELM با همان تابع هسته بکاررفته در فرآیند انتخاب نمونه‌های اولیه اجرا می‌شود. استدلال اصلی در انجام این آزمایش انتخاب بهترین تابع هسته در روش KELM می‌باشد. بدین ترتیب ۴ تابع هسته خطی^۳، RBF، چندجمله‌ای^۴ و سیگموئید^۵ مورد بررسی بررسی قرار گرفته است. پارامترهای اولیه ثابت در نظر گرفته شده برای K ، θ_1 و C به ترتیب برابر ۵۰۰، ۰/۷۵، ۱۰۰ و ۰/۱۸ در نظر گرفته شده است. از طرفی پارامترهای هر تابع هسته نیز به ازای RBF، چندجمله‌ای و سیگموئید به ترتیب برابر ۵، (۱، ۰/۲) و (۱، ۰/۲) در نظر گرفته شده است. از آنجایی که اثربخشی توابع هسته در روش KELM مورد بررسی قرار گرفته است، معیار تصمیم‌گیری، نتایج شرح‌گذاری بکارگیری روش KELM می‌باشد. شکل ۴ نمودار مقایسه عملکرد شرح‌گذاری بکارگیری توابع هسته مختلف در MLPS-KELM را نشان می‌دهد.

همانطور که از نمودار شکل ۴ مشاهده می‌شود، بکارگیری تابع هسته RBF دارای بهترین عملکرد شرح‌گذاری می‌باشد. تابع هسته RBF تابع پیشنهادی [۶] در KELM می‌باشد. بنابراین در سایر آزمایش‌های صورت گرفته در این پژوهش از این تابع هسته استفاده شده است.

شده و سپس از مقادیر بدست آمده میانگین‌گیری می‌شود. در صورتی که در میانگین‌گیری میکرو در ابتدا از مقادیر TP ، TN ، FP و FN میانگین‌گیری شده و سپس معیار ارزیابی محاسبه می‌شود [۴۲]. روابط ۱۹ و ۲۰ نحوه محاسبه معیارهای ارزیابی میکرو و ماکرو را نشان می‌دهند. در این روابط c نشان دهنده تعداد برجسب‌هاست.

$$measure_{macro} = \frac{\sum_{i=1}^c measure(TP_i, TN_i, FP_i, FN_i)}{c} \quad (19)$$

$$measure_{micro} = measure\left(\sum_{i=1}^c TP_i, \sum_{i=1}^c TN_i, \sum_{i=1}^c FP_i, \sum_{i=1}^c FN_i\right) \quad (20)$$

در این پژوهش از میانگین‌گیری میکرو و ماکرو استفاده شده است. بدین ترتیب مقادیر دقت ماکرو، فراخوان ماکرو، F1 ماکرو، دقت میکرو، فراخوان میکرو و F1 میکرو محاسبه و به عنوان معیار ارزیابی نتایج شرح‌گذاری خودکار تصاویر استفاده شده است.

ارزیابی نتایج به کارگیری روش‌های PS:

در به کارگیری روش‌های انتخاب نمونه اولیه علاوه بر دقت طبقه‌بندی که با استفاده از معیارهای تعریف شده در بالا محاسبه می‌شود، میزان کاهش حجم نمونه‌ها و همچنین کاهش حجم به همراه عدم کاهش دقت طبقه‌بندی نیز از اهمیت بالایی برخوردار است. از این رو در ارزیابی این روش‌ها، از معیارهای ارزیابی دیگری با نام نرخ کاهش^۱ و نرخ تعادل^۲ استفاده می‌شود [۱۰].

نرخ کاهش به صورت تعداد نمونه‌های اولیه بدست آمده توسط هر الگوریتم PS بر تعداد نمونه‌های موجود در مجموعه آموزش اولیه تعریف می‌شود. نرخ تعادل نیز میزان تعادل مابین دقت طبقه‌بندی با به کارگیری مجموعه کاهش یافته به عنوان مجموعه آموزش و نرخ کاهش بدست آمده توسط روش PS را نشان می‌دهد. روابط ۲۱ و ۲۲ به ترتیب نحوه محاسبه این دو معیار را نشان می‌دهند.

$$Reduction Rate (RR) = \frac{\text{number of instance after apply PS algorithm}}{\text{number of training instance}} \quad (21)$$

$$Trade-off Rate = (Macro - Micro)F1 \times RR \quad (22)$$

ذکر این نکته ضروری است که پیاده‌سازی‌ها و ارزیابی‌های این مقاله با استفاده از نرم‌افزار MATLAB R2013b در سیستم Lenovo Ideapad با مشخصه‌های CPU CORE i5 و RAM 6GB صورت گرفته است.

۵-۳ تنظیم‌های اولیه آزمایش‌ها

در این بخش تنظیم‌های مربوط به آزمایش‌های صورت گرفته از قبیل تنظیم پارامترها و اثربخشی انواع توابع هسته و توابع محاسبه

³Linear

⁴Polynomial

⁵Sigmoid

¹Reduction rate

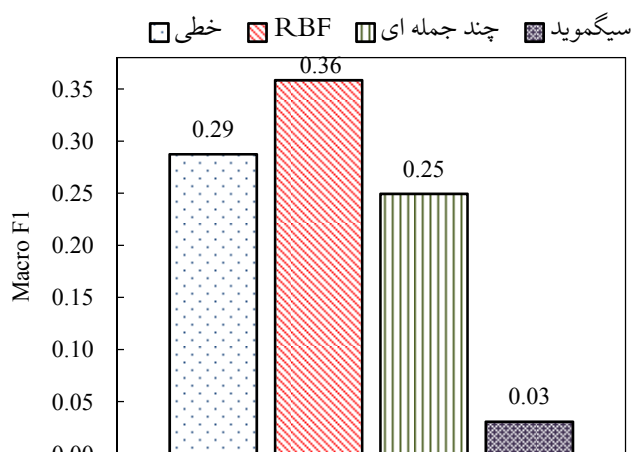
²Trade-off rate

بکارگیری مجموعه آموزش کاهش یافته با استفاده از این رابطه دارای بالاترین عملکرد شرح‌گذاری است. از این‌رو در سایر آزمایش‌های صورت گرفته از این رابطه استفاده شده است.

۵-۳-۲- تنظیم پارامترهای بهینه

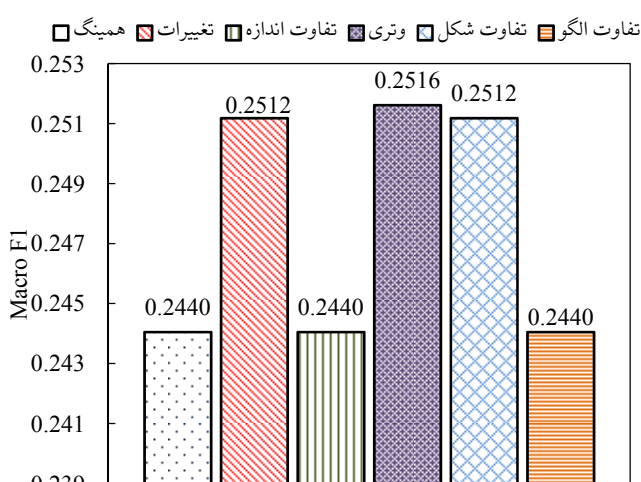
در الگوریتم MLPS-KELM ارائه شده دو دسته از پارامترها بایستی تنظیم شود. از آنجاکه تنظیم هر پارامتر بر روی کاهش حجم مجموعه آموزش و همچنین طبقه‌بندی مجموعه بدون برجسب آزمایش تاثیر زیادی دارد، پارامترهای مربوط به هر بخش به صورت جداگانه و بر اساس معیارهای تصمیم‌گیری جداگانه تنظیم شده است. به عبارت دقیق‌تر پارامترهای مربوط به کاهش حجم شامل θ_1, C, θ_2 و $KP(RBF)$ و پارامترهای طبقه‌بندی گام ۴ الگوریتم شامل C, θ_2 و $KP(RBF)$ به صورت جداگانه تنظیم می‌شوند. استدلال اصلی تنظیم جداگانه، بکارگیری روش KELM در دو گام از روش MLPS-KELM می‌باشد. تنظیم پارامترها در این بخش بر اساس روش اعتبارسنجی متقابل دو بخشی^۱ بر روی مجموعه آموزش صورت گرفته است. در این روش اعتبارسنجی، مجموعه آموزش به دو بخش تقسیم شده و هر بار یکی از بخش‌ها به عنوان مجموعه آموزش و دیگری به عنوان مجموعه اعتبارسنجی در نظر گرفته شده و الگوریتم اجرا می‌شود. سپس مجموعه‌های آموزش و اعتبارسنجی معکوس شده و الگوریتم مجدد اجرا می‌شود. در انتها میانگین نتایج بدست آمده به عنوان معیار تصمیم در نظر گرفته می‌شود. به مانند سایر آزمایش‌های صورت گرفته، خط پایه در تنظیم پارامترهای انتخاب نمونه‌های اولیه، روش $KNN(K=1)$ و در تنظیم پارامترهای طبقه‌بندی، روش KELM در نظر گرفته شده است.

در ابتدا پارامترهای مربوط به انتخاب نمونه‌های اولیه تنظیم شده است. در هر مرحله از تنظیم، سایر پارامترها را ثابت در نظر گرفته و پارامتر مورد نظر را تغییر داده و مقدار مربوط به بهترین نتیجه به عنوان پارامتر بهینه در نظر گرفته می‌شود. بدین منظور مقادیر θ_1 و θ_2 در بازه ۰ تا ۱، K در بازه ۵ تا ۵۲۵، C در بازه 2^{-5} تا 2^0 و $KP(RBF)$ در بازه ۱ تا ۱۰ در نظر گرفته شده است. در تمامی موارد تنظیم پارامتر در بخش انتخاب نمونه‌های اولیه، میانگین نرخ تعادل ماکرو و نرخ تعادل میکرو و در بخش طبقه‌بندی، میانگین $F1$ ماکرو و $F1$ میکرو به عنوان معیار تصمیم‌گیری در نظر گرفته شده است. از آنجاکه رفتار مجموعه تصاویر چند برجسبه استفاده شده در برخی از آزمایش‌ها مشابه می‌باشد، آزمایش مجددی صورت نگرفته و نتایج آزمایش‌ها در مجموعه NUS-WIDE-Object به عنوان پارامتر بهینه در این مجموعه تصاویر در نظر گرفته شده است. به عبارت دیگر تغییر در پارامترها تغییری در نتیجه اعمال الگوریتم ایجاد نخواهد کرد. جدول ۲ پارامترهای بهینه تنظیم شده هر بخش از الگوریتم MLPS-



شکل ۴ مقایسه عملکرد شرح‌گذاری بکارگیری توابع هسته مختلف در روش MLPS-KELM روی مجموعه Object.

به منظور بررسی اثربخشی و تصمیم‌گیری در مورد انتخاب مناسب‌ترین رابطه محاسبه عدم شباهت دودویی در مرحله مقایسه دو بردار برجسب پیش‌بینی شده و بردار برجسب واقعی (خط ۹ الگوریتم MLPS-KELM)، آزمایش دیگری بر روی مجموعه NUS-WIDE-Object صورت گرفته است. در این آزمایش رابطه‌های ۹ تا ۱۴ مورد مقایسه قرار گرفته‌اند. بدین ترتیب در هر مورد با بکارگیری هر رابطه، روش MLPS-KELM با پارامترهای ثابت پیش فرض اجرا و ارزیابی می‌شود. خط پایه شرح‌گذاری در این آزمایش روش $KNN(K=1)$ در نظر گرفته شده و معیار تصمیم‌گیری در مورد انتخاب مناسب‌ترین تابع محاسبه عدم شباهت دودویی، $F1$ ماکرو در نظر گرفته شده است. شکل ۵ نمودار مقایسه عملکرد شرح‌گذاری بکارگیری توابع محاسبه عدم شباهت دودویی مختلف در MLPS-KELM را نشان می‌دهد.



شکل ۵ مقایسه عملکرد شرح‌گذاری بکارگیری توابع محاسبه عدم شباهت دودویی مختلف در روش MLPS-KELM روی مجموعه Object.

همانطور که از نمودار شکل ۵ مشاهده می‌شود، رابطه محاسبه عدم شباهت دودویی وتری نسبت به سایر روابط از عملکرد شرح‌گذاری بالاتری برخوردار است. به عبارت دقیق‌تر

^۱2-fold cross validation

در رابطه عدم شباهت دودویی بررسی شده در این پژوهش (وتری) مورد مقایسه قرار گرفته است. از طرفی در بکارگیری این روش در فرآیند شرح‌گذاری مجموعه تصاویر در مقیاس بزرگ، این روش با روش‌های [۳۹] KNN (K=1)، [۴] ELM، [۱۱] MLENN، [۶] RKELM (وتری) و [۶] MLENN (وتری) مقایسه شده است. استدلالات مقایسه مشابه در دو روش MLENN و MLENN (وتری)، مقایسه عملکرد انتخاب نمونه‌های موثر در روش ارائه شده نسبت به دو رویکرد دیگر است.

۵-۴-۱- نتایج NUS-WIDE-Object

در این بخش نتایج ارزیابی روش ارائه شده بر روی مجموعه NUS-WIDE-Object مورد بررسی قرار می‌گیرد. همان‌طور که بیان شد مقایسه در دو حوزه عملکرد انتخاب نمونه‌های اولیه و عملکرد طبقه‌بندی صورت گرفته است. جدول ۳ مقایسه نتایج و عملکرد روش MLPS-KELM ارائه شده در سه بخش انتخاب نمونه‌های اولیه و کاهش حجم، شرح‌گذاری و نرخ تعادل مابین کاهش حجم و شرح‌گذاری را در مجموعه NUS-WIDE-Object نشان می‌دهد. همان‌طور که در جدول ۳ مشاهده می‌شود، رویکرد MLENN مقدار کاهش حجمی برابر ۶۱/۳۳٪ دارد. به عبارت دیگر با به‌کارگیری این ایده، تعداد نمونه‌های آموزش مجموعه تصویر NUS-WIDE-Object از ۱۷۹۲۸ به ۶۹۳۳ نمونه کاهش یافته است. با تغییر ایده ارائه شده در [۱۱] با بکارگیری رابطه عدم شباهت دودویی وتری، بهبود ۹/۲۷٪ در کاهش حجم حاصل شده است. (تغییر صورت گرفته در این روش بخشی از ایده اثربخشی بکارگیری روابط عدم شباهت دودویی مختلف بررسی شده در این پژوهش می‌باشد) همچنین با بکارگیری ایده پیش‌بینی برچسب با روش KELM و استفاده از الگوریتم MLPS-KELM برای کاهش حجم، بهبود ۱۲/۸۲٪ حاصل شده است.

جدول ۳ مقایسه عملکرد روش MLPS-KELM در سه بخش شرح‌گذاری، انتخاب نمونه‌های اولیه و نرخ تعادل مابین کاهش حجم و شرح‌گذاری در مجموعه NUS-WIDE-Object (تعداد نمونه‌های آموزش = ۱۷۹۲۸)

نرخ تعادل شرح‌گذاری / کاهش حجم		عملکرد کاهش حجم		عملکرد شرح‌گذاری						
نرخ تعادل میکرو (%)	نرخ تعادل ماکرو (%)	نرخ کاهش (%)	تعداد نمونه‌ها پس از کاهش	F1 میکرو (%)	فراخوان میکرو (%)	دقت میکرو (%)	F1 ماکرو (%)	فراخوان ماکرو (%)	دقت ماکرو (%)	روش
-	-	-	-	۳۱/۰۷	۳۰/۸۰	۳۱/۳۴	۲۴/۴۱	۲۴/۸۹	۲۳/۹۴	[۳۶] KNN
-	-	-	-	۳۹/۵۴	۵۰/۳۴	۳۲/۵۶	۲۰/۵۵	۳۲/۰۳	۱۵/۱۳	[۴] ELM
-	-	-	-	۳۴/۶۰	۳۹/۰۳	۳۱/۰۸	۲۰/۶۲	۲۶/۸۳	۱۶/۷۴	[۶] RKELM
۱۹/۱۹	۱۴/۹۵	۶۱/۳۳	۶۹۳۳	۳۱/۲۹	۳۰/۰۹	۳۲/۵۸	۲۴/۳۷	۲۳/۲۷	۲۵/۵۹	[۸] MLENN
۲۰/۹۹	۱۷/۴۵	۷۰/۶	۵۲۶۶	۲۹/۷۳	۳۰/۱۵	۲۹/۳۳	۲۴/۷۲	۲۴/۵۵	۲۴/۸۹	MLENN(Chord)
۲۲/۶۶	۱۸/۸۸	۷۴/۱۵	۴۶۳۴	۳۰/۵۶	۲۹/۷۱	۳۱/۴۵	۲۵/۵۰	۲۵/۲۱	۲۵/۷۰	MLPS-KNN
۳۰/۴۵	۲۶/۹۲	۷۴/۱۵	۴۶۳۴	۴۱/۰۶	۴۱/۰۵	۴۱/۰۷	۳۶/۳۱	۳۸/۹۰	۳۴/۰۵	MLPS-KELM

۱. بدون کاهش حجم و انتخاب نمونه‌های اولیه در مجموعه آموزش

۲. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KNN (K=1)

۳. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KELM

KELM ارائه شده در مجموعه‌های NUS-WIDE-Object، NUS-WIDE-Scene، NUS-WIDE-Lite و NUS-WIDE را نشان می‌دهد.

جدول ۲ پارامترهای بهینه تنظیم شده هر بخش از الگوریتم MLPS-KELM در مجموعه‌های NUS-WIDE-Object، NUS-WIDE-Scene، NUS-WIDE-Lite و NUS-WIDE.

پارامتر	Object	Scene	Lite	NUS	انتخاب نمونه‌های اولیه
K	۲۵	۲۵	۲۵	۲۵	
θ_1	۰/۷۵	۰/۷۵	۰/۷۵	۰/۷۵	
θ_2	۰/۴	۰/۵۵	۰/۷۵	۰/۷۵	
C	۲۱	۲۱	۲۱	۲۱	
KP(RBF)	۲/۵	۲/۵	۲/۵	۲/۵	
θ_2	۰/۱۸	۰/۰۸	۰/۰۵	۰/۰۷	شرح‌گذاری
C	۲۲	۲۲	۲۲	۲۲	
KP(RBF)	۴/۵	۴/۵	۴/۵	۴/۵	

۵-۴-۲- تحلیل و ارزیابی نتایج تجربی

در این بخش به بیان نتایج تجربی حاصل از اعمال الگوریتم MLPS-KELM ارائه شده بر روی هر مجموعه تصویر می‌پردازیم. همان‌طور که بیان شد، انتخاب نمونه‌های اولیه موثر و همچنین کاربردی سازی روش KELM در مجموعه‌های چندبرچسب مقیاس بزرگ و بهبود دقت طبقه‌بندی در این مجموعه‌ها دو ایده اصلی روش MLPS-KELM ارائه شده می‌باشند. از اینرو در آزمایش‌های صورت گرفته برای هر مجموعه نتایج کاهش حجم و انتخاب نمونه‌های اولیه و همچنین نتایج طبقه‌بندی در هر مجموعه مورد مقایسه قرار گرفته است.

در بکارگیری MLPS-KELM در فرآیند انتخاب نمونه‌های اولیه، این روش با روش‌های MLENN [۱۱] و MLENN با تغییر

مجموعه تصاویر در مقیاس بزرگ اهمیت کاربردی سازی روش KEM را در کاربردهای چندبرچسب مقیاس بزرگ به وضوح نشان می‌دهد. ذکر این نکته ضروری است که روش KEM بدون ایده کاهش حجم مجموعه آموزش مطرح شده در این پژوهش با محدودیت‌های حافظه مواجه شده و اعمال آن بر روی مجموعه آموزش مقیاس بزرگ کامل تا حدودی غیرممکن است.

در نهایت همان‌طور که جدول ۳ نشان می‌دهد، عملکرد کلی معیارهای تعادل مابین دقت و کاهش حجم در روش ارائه شده نسبت روش کاهش حجم MLENN بهبود قابل توجهی دارد. بهبود نرخ تعادل ماکرو و میکرو در روش MLPS-KNN نسبت به روش MLENN به ترتیب برابر ۳,۹۳٪ و ۳,۴۷٪ می‌باشد. همچنین بهبود این دو معیار در MLPS-KEM نسبت به MLENN به ترتیب برابر ۱۱,۹۷٪ و ۱۱,۲۶٪ می‌باشد.

۵-۴-۲- نتایج NUS-WIDE-Scene

در این بخش نتایج ارزیابی روش ارائه شده بر روی مجموعه NUS-WIDE-Scene مورد بررسی قرار می‌گیرد. نتایج این بخش نیز به‌مانند نتایج ارائه شده در بخش قبل می‌باشد. جدول ۴ مقایسه نتایج و عملکرد روش MLPS-KEM ارائه شده در سه بخش انتخاب نمونه‌های اولیه، شرح‌گذاری و نرخ تعادل مابین کاهش حجم و شرح‌گذاری را در مجموعه NUS-WIDE-Scene نشان می‌دهد.

همان‌طور که از جدول ۴ مشاهده می‌شود، رویکرد MLENN تعداد نمونه‌ها را از ۱۷۴۶۳ به ۱۳۷۳۷ کاهش داده است. در صورتی که این تعداد در روش MLPS-KEM به ۶۸۶۷ نمونه رسیده است. به عبارت دیگر بهبودی ۳۹,۳۴٪ نسبت به روش MLENN از خود نشان داده است.

بکارگیری ایده‌های مطرح شده در این پژوهش دو مزیت عمده در بهبود دقت شرح‌گذاری خودکار مجموعه تصاویر در مقیاس بزرگ به همراه داشته است. بهبود نتایج در MLPS-KNN نشان از انتخاب نمونه‌های اولیه موثر در روش ارائه شده نسبت به سایر روش‌های انتخاب نمونه‌های اولیه دارد. هدف اصلی ارائه نتایج روش ارائه شده با بکارگیری روش KNN، مقایسه تاثیر نحوه انتخاب نمونه‌های اولیه در روش‌های دارای رویکرد انتخاب نمونه‌های اولیه و طبقه‌بند مبتنی بر نمونه‌ها می‌باشد. دقت شرح‌گذاری بالای روش MLPS-KEM نسبت به سایر روش‌های شرح‌گذاری، اهمیت کاربردی سازی روش KEM را در مجموعه‌های چندبرچسب مقیاس بزرگ نشان می‌دهد. از طرفی بهبود دقت روش MLPS-KEM نسبت به سایر نسخه‌های ارائه شده از ELM نشان از صلاحیت بکارگیری روش KEM در پژوهش جاری دارد.

همان‌طور که مشاهده می‌شود، بهبود دقت صورت گرفته در معیار F1 ماکرو در MLPS-KNN نسبت به روش‌های KNN، MLENN و MLENN(Chord) به ترتیب برابر ۱,۰۹٪، ۱,۱۳٪ و ۰,۷۸٪ می‌باشد. عدم بهبود در معیار F1 میکرو در این روش نسبت به برخی از روش‌ها از جمله KNN به دلیل نسبت بالای کاهش حجم نمونه‌ها در روش ارائه شده می‌باشد. این افت عملکرد با تغییر روش KNN با KEM در گام ۴ الگوریتم جبران شده است.

بهبود دقت روش MLPS-KEM در معیار F1 ماکرو نسبت به روش‌های KNN، MLENN، ELM، MLENN(Chord) و RKEM به ترتیب برابر ۱۱,۹۹٪، ۱۵,۷۶٪، ۱۱,۹۴٪، ۱۱,۵۹٪ و ۱۵,۶۹٪ می‌باشد. از طرفی بهبود صورت گرفته در معیار F1 میکرو نسبت به این روش‌ها به ترتیب برابر ۹,۹۹٪، ۱,۵۲٪، ۹,۷۷٪، ۱۱,۳۳٪ و ۶,۴۶٪ می‌باشد. بهبود قابل توجه دقت روش MLPS-KEM نسبت به سایر روش‌های شرح‌گذاری

جدول ۴ مقایسه عملکرد روش MLPS-KEM در سه بخش شرح‌گذاری، انتخاب نمونه‌های اولیه و نرخ تعادل مابین کاهش حجم و شرح‌گذاری در مجموعه NUS-WIDE-Scene (تعداد نمونه‌های آموزش = ۱۷۴۶۳)

نرخ تعادل شرح‌گذاری / کاهش حجم		عملکرد کاهش حجم		عملکرد شرح‌گذاری						
نرخ تعادل میکرو (%)	نرخ تعادل ماکرو (%)	نرخ کاهش (%)	تعداد نمونه‌ها پس از کاهش	F1 میکرو (%)	فراخوان میکرو (%)	دقت میکرو (%)	F1 ماکرو (%)	فراخوان ماکرو (%)	دقت ماکرو (%)	روش
-	-	-	-	۴۹/۶۵	۴۷/۸۷	۵۱/۵۶	۲۴/۴۱	۲۳/۴۶	۲۵/۴۸	KNN ^[۳۶]
-	-	-	-	۵۹/۰۸	۵۴/۷۰	۶۴/۲۲	۲۴/۷۸	۲۸/۵۱	۲۱/۹۲	ELM ^[۴]
-	-	-	-	۵۵/۴۰	۴۹/۲۷	۶۳/۲۸	۲۳/۱۷	۲۲/۲۲	۲۴/۲۱	RKEM ^[۶]
۱۰/۶۵	۵/۲۴	۲۱/۳۴	۱۳۷۳۷	۴۹/۸۹	۴۶/۶۱	۵۳/۶۶	۲۴/۵۶	۲۲/۵۹	۲۶/۹۱	MLENN ^[۸]
۲۹/۹۲	۱۵/۰۵	۶۱/۷۰	۶۶۹۴	۴۸/۴۹	۴۵/۸۹	۵۱/۴۱	۲۴/۳۹	۲۱/۹۴	۲۷/۴۵	MLENN(Chord) ^۲
۳۰/۵۷	۱۵/۰۹	۶۰/۶۸	۶۸۶۷	۵۰/۳۷	۴۶/۱۸	۵۵/۴۱	۲۴/۸۶	۲۱/۹۵	۲۸/۶۶	MLPS-KNN ^۲
۳۶/۵	۲۰/۱۷	۶۰/۶۸	۶۸۶۷	۶۰/۱۵	۵۵/۱۱	۶۶/۲۱	۳۳/۲۴	۳۲/۹۰	۳۳/۵۹	MLPS-KEM ^۳

۱. بدون کاهش حجم و انتخاب نمونه‌های اولیه در مجموعه آموزش

۲. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KNN (K=1)

۳. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KEM

نسبت به روش MLENN به ترتیب برابر ۱۴,۹۳٪ و ۲۵,۸۵٪ و نسبت به روش MLENN(Chord) به ترتیب برابر ۵,۱۲٪ و ۶,۵۸٪ می‌باشد.

۵-۴-۳- نتایج NUS-WIDE-Lite

در این بخش نتایج ارزیابی روش ارائه شده بر روی مجموعه NUS-WIDE-Lite مورد بررسی قرار می‌گیرد. آزمایش‌های صورت گرفته در این بخش نیز به‌مانند دو بخش قبلی می‌باشد. جدول ۵ نتایج عملکرد و مقایسه ارائه شده در کاهش حجم و انتخاب نمونه‌های اولیه موثر روی مجموعه NUS-WIDE-Lite را نشان می‌دهد.

همانطور که از جدول ۵ مشاهده می‌شود، رویکرد MLENN تعداد نمونه‌ها را از ۲۷۸,۰۷ به ۲۱۰,۳۶ کاهش داده است. بهبود نرخ کاهش در روش MLPS-KELM ارائه شده نسبت به این روش ۵۴,۷۹٪ می‌باشد.

نرخ کاهش حجم بالای روش MLPS-KELM با بهبود در دقت طبقه‌بندی نمونه‌های بدون برچسب همراه است. بهبود دقت در معیار F1 ماکرو در MLPS-KNN نسبت به روش‌های KNN، MLENN و MLENN(Chord) به ترتیب برابر ۰,۰۵٪، ۰,۱۲٪ و ۱,۰۵٪ می‌باشد. همچنین بهبود در معیار F1 میکرو نسبت به این روش‌ها به ترتیب برابر ۲,۱۱٪، ۱,۳۱٪ و ۱,۴۳٪ می‌باشد.

بهبود دقت MLPS-KELM در معیار F1 ماکرو نسبت به روش‌های KNN، ELM، MLENN، MLENN(Chord) و RKELM به ترتیب برابر ۶,۷۲٪، ۸,۴۷٪، ۶,۷۹٪، ۷,۷۲٪ و ۹,۶۶٪ می‌باشد. از طرفی بهبود صورت گرفته در معیار Micro F1 نسبت به این روش‌ها به ترتیب برابر ۹,۶٪، ۰,۲۶٪، ۸,۸٪، ۸,۹۲٪ و ۴,۵۷٪ می‌باشد.

نرخ کاهش حجم بالای روش ارائه شده با بهبود در دقت طبقه‌بندی نمونه‌های بدون برچسب همراه است. بهبود دقت صورت گرفته در معیار F1 ماکرو در MLPS-KNN نسبت به روش‌های KNN، MLENN، MLENN(Chord) به ترتیب برابر ۰,۴۵٪، ۰,۳٪ و ۰,۴۷٪ می‌باشد. همچنین بهبود در معیار F1 میکرو در این روش به ترتیب برابر ۰,۷۲٪، ۰,۴۸٪ و ۱/۸۸٪ می‌باشد.

بهبود دقت در روش MLPS-KELM نیز در معیارهای ماکرو و میکرو نسبت به سایر روش‌های شرح‌گذاری مجموعه تصاویر در مقیاس بزرگ در جدول ۴ قابل مشاهده است. بهبود دقت این روش در معیار F1 ماکرو نسبت به روش‌های KNN، MLENN، ELM و MLENN(Chord) به ترتیب برابر ۸,۸۳٪، ۸/۴۶٪، ۸,۶۸٪، ۸,۸۵٪ و ۱۰,۰۷٪ می‌باشد. از طرفی بهبود صورت گرفته در معیار F1 میکرو نسبت به این روش‌ها به ترتیب برابر ۱۰,۵٪، ۱۰,۰۷٪، ۱۰,۲۶٪، ۱۱,۶۶٪ و ۴,۷۵٪ می‌باشد. همان‌طور که مشاهده می‌شود در مجموعه NUS-WIDE-Scene نیز به‌مانند NUS-WIDE-Object بهبود قابل توجهی در دقت شرح‌گذاری روش MLPS-KELM نسبت به سایر روش‌های شرح‌گذاری مجموعه تصاویر در مقیاس بزرگ مشاهده می‌شود. این بهبود نشان از اهمیت بکارگیری روش KELM در کاربردهای چندبرچسب مقیاس بزرگ از جمله مسئله شرح‌گذاری مجموعه تصاویر در مقیاس بزرگ می‌باشد.

در نهایت همان‌طور که مشاهده می‌شود بهبود نرخ تعادل ماکرو و میکرو در روش MLPS-KNN نسبت به روش MLENN به ترتیب برابر ۹,۸۵٪ و ۱۹,۹۲٪ و نسبت به روش MLENN(Chord) به ترتیب برابر ۰,۰۴٪ و ۰,۶۵٪ می‌باشد. همچنین این بهبود در روش MLPS-KELM که از روش KELM در طبقه‌بندی نمونه‌های آزمایش بدون برچسب استفاده می‌کند،

جدول ۵ مقایسه عملکرد روش MLPS-KELM در سه بخش شرح‌گذاری، انتخاب نمونه‌های اولیه و نرخ تعادل مابین کاهش حجم و شرح‌گذاری در مجموعه NUS-WIDE-Lite (تعداد نمونه‌های آموزش = ۲۷۸,۰۷)

نرخ تعادل شرح‌گذاری / کاهش حجم		عملکرد کاهش حجم		عملکرد شرح‌گذاری						
نرخ تعادل میکرو (%)	نرخ تعادل ماکرو (%)	نرخ کاهش (%)	تعداد نمونه‌ها پس از کاهش	F1 میکرو (%)	فراخوان میکرو (%)	دقت میکرو (%)	F1 ماکرو (%)	فراخوان ماکرو (%)	دقت ماکرو (%)	روش
-	-	-	-	۴۶/۲۰	۴۴/۷۵	۴۷/۷۵	۱۸/۹۱	۱۸/۶۴	۱۹/۱۹	KNN ^[۳۶]
-	-	-	-	۵۵/۵۴	۴۹/۱۳	۶۳/۸۷	۱۷/۱۶	۱۷/۹۲	۱۶/۴۶	ELM ^[۴]
-	-	-	-	۵۱/۲۳	۴۳/۶۸	۶۱/۹۳	۱۵/۹۷	۱۴/۵۱	۱۷/۷۶	RKELM ^[۶]
۱۱/۴۲	۴/۵۸	۲۴/۳۰	۲۱۰,۳۶	۴۷	۴۴/۴۱	۴۹/۹۲	۱۸/۸۴	۱۸/۱۵	۱۹/۶۰	MLENN ^[۸]
۳۴/۶۰	۱۳/۲۲	۷۳/۸۰	۷۲۷۱	۴۶/۸۸	۴۷/۲۳	۴۶/۵۳	۱۷/۹۱	۱۷/۹۴	۱۷/۸۹	MLENN(Chord) ^۱
۳۸/۲۱	۱۴/۹۹	۷۹/۰۹	۵۸۱۴	۴۸/۳۱	۴۳/۵۳	۵۴/۲۷	۱۸/۹۶	۱۷/۱۶	۲۱/۱۷	MLPS-KNN ^۲
۴۴/۲۷	۲۰/۲۷	۷۹/۰۹	۵۸۱۴	۵۵/۸۰	۴۶/۹۶	۶۹/۲۹	۲۵/۶۳	۲۳/۳۶	۲۸/۳۷	MLPS-KELM ^۳

۱. بدون کاهش حجم و انتخاب نمونه‌های اولیه در مجموعه آموزش

۲. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KNN (K=1)

۳. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KELM

جدول ۶ مقایسه عملکرد روش MLPS-KELM در سه بخش شرح‌گذاری، انتخاب نمونه‌های اولیه و نرخ تعادل مابین کاهش حجم و شرح‌گذاری در مجموعه NUS-WIDE (تعداد نمونه‌های آموزش = ۱۶۱۷۸۹)

نرخ تعادل شرح‌گذاری / کاهش حجم		عملکرد کاهش حجم		عملکرد شرح‌گذاری						روش
نرخ تعادل میکرو (%)	نرخ تعادل ماکرو (%)	نرخ کاهش (%)	تعداد نمونه‌ها پس از کاهش	F1 میکرو (%)	فراخوان میکرو (%)	دقت میکرو (%)	F1 ماکرو (%)	فراخوان ماکرو (%)	دقت ماکرو (%)	
-	-	-	-	۵۵/۱۹	۵۳/۷۰	۵۶/۷۷	۱۶/۰۵	۱۵/۶۹	۱۶/۴۳	KNN ^۱ [۳۶]
-	-	-	-	۲۶/۹۴	۲۴/۱۸	۳۰/۴۱	۷/۱۴	۸/۵۱	۶/۱۶	ELM ^۱ [۴]
-	-	-	-	۳۱/۶۴	۲۳/۲۵	۴۹/۴۹	۷/۶۱	۵/۵۰	۱۲/۳۶	RKELM ^۱ [۶]
۳۰/۷۲	۸/۷۸	۵۵/۷۲	۷۱۶۳۹	۵۵/۱۳	۵۰/۹۴	۶۰/۰۶	۱۵/۷۶	۱۴/۲۱	۱۷/۶۸	MLENN ^۱ [۸]
۲۱/۷۱	۹/۷۰	۶۷/۲۶	۵۲۹۷۳	۳۲/۲۸	۳۴/۹۰	۳۰/۰۲	۱۴/۴۳	۱۴/۶۷	۱۴/۱۹	MLENN(Chord) ^۱
۲۸/۷۴	۱۳/۱۱	۹۰/۴۳	۱۵۴۸۸	۳۱/۷۸	۲۴/۷۲	۴۴/۴۹	۱۴/۵۰	۱۰/۷۴	۲۲/۲۸	MLPS-KNN ^۱
۳۶/۹۳	۲۰/۵۵	۹۰/۴۳	۱۵۴۸۸	۴۰/۸۴	۳۰/۹۶	۶۰	۲۲/۷۲	۱۹/۵۹	۲۷/۰۵	MLPS-KELM ^۱

۱. بدون کاهش حجم و انتخاب نمونه‌های اولیه در مجموعه آموزش

۲. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KNN (K=1)

۳. انتخاب نمونه‌های اولیه و سپس شرح‌گذاری با روش KELM

ELM، MLENN، MLENN(Chord) و RKELM به ترتیب برابر ۶،۶۷٪، ۱۵،۵۸٪، ۶،۹۶٪، ۸،۲۹٪ و ۱۵،۱۱٪ می‌باشد. در نهایت بهبود نرخ تعادل ماکرو در روش MLPS-KNN نسبت به روش MLENN برابر ۴،۳۳٪ و نسبت به روش MLENN(Chord) برابر ۳،۴۱٪ می‌باشد. همچنین بهبود نرخ تعادل ماکرو و میکرو در روش MLPS-KELM نسبت به روش MLENN به ترتیب برابر ۱۱،۷۷٪ و ۶،۲۱٪ و نسبت به روش MLENN(Chord) به ترتیب برابر ۱۰،۸۵٪ و ۱۵،۲۲٪ می‌باشد.

۵-۴-۵- مقایسه عملکرد شرح‌گذاری با سایر روش‌ها

به منظور نشان دادن صلاحیت بکارگیری روش ارائه شده در شرح‌گذاری خودکار مجموعه تصاویر در مقیاس بزرگ نتایج شرح‌گذاری این روش در مجموعه‌های NUS-WIDE-Lite و NUS-WIDE با برخی از رویکردهای ارائه شده در این مجموعه تصاویر از جمله [۴۳]LNP، [۴۴]EGSSC، [۴۵]SGSSL، [۴۶]LSMP، [۴۷]BIA و [۴۸]HASH-SISO مقایسه شده است. در این مقایسه، در ابتدا مجموعه آموزش کاهش یافته و سپس مجموعه تصویر آزمایش شرح‌گذاری می‌شود. در تمامی روش‌های مورد مقایسه از تعداد تصاویر آموزش و آزمایش و ویژگی‌های مشابه با این پژوهش استفاده شده است. از آنجایی که در برخی از مقالات مورد نظر فقط دقت ماکرو در دسترس بوده است، نتایج این معیار با یکدیگر مقایسه شده است. شکل ۶ الف و ب به ترتیب نمودار مقایسه‌ای روش ارائه شده با سایر روش‌ها را در مجموعه تصاویر NUS-WIDE-Lite و NUS-WIDE نشان می‌دهد.

از طرفی همانطور که مشاهده می‌شود بهبود نرخ تعادل ماکرو و میکرو در روش MLPS-KNN نسبت به روش MLENN به ترتیب برابر ۱۰،۴۱٪ و ۲۶،۷۹٪ و نسبت به روش MLENN(Chord) به ترتیب برابر ۱،۷۷٪ و ۳،۶۱٪ می‌باشد. همچنین این بهبود در روش MLPS-KELM نسبت به روش MLENN به ترتیب برابر ۱۵،۶۹٪ و ۳۲،۸۵٪ و نسبت به روش MLENN(Chord) به ترتیب برابر ۷،۰۵٪ و ۹،۶۷٪ می‌باشد.

۵-۴-۴- نتایج NUS-WIDE

در نهایت در این قسمت نتایج ارزیابی روش ارائه شده بر روی مجموعه NUS-WIDE مورد بررسی قرار می‌گیرد. نحوه انجام آزمایش‌های این بخش نیز به مانند بخش‌های قبل می‌باشد. جدول ۶ نتایج عملکرد و مقایسه ارائه شده در کاهش حجم و انتخاب نمونه‌های اولیه موثر روی مجموعه NUS-WIDE را نشان می‌دهد. موثر بودن ایده‌های ارائه شده در کاهش حجم و انتخاب نمونه‌های اولیه و همچنین دقت طبقه‌بندی به وضوح قابل مشاهده است. همان‌طور که مشاهده می‌شود، با به کارگیری MLENN نرخ کاهش حجمی برابر ۵۵،۷۲٪ به دست می‌آید. در صورتی که در روش MLPS-KELM ارائه شده این نرخ برابر ۹۰،۴۳٪ می‌باشد. به عبارت دیگر بهبود در نرخ کاهش حجم این روش نسبت به MLENN(Chord) به ترتیب برابر ۳۴،۷۱٪ و ۲۳،۱۷٪ می‌باشد.

بهبود دقت در معیار F1 ماکرو در MLPS-KNN نسبت به روش MLENN(Chord) برابر ۰،۰۷٪ می‌باشد. همان‌طور که بیان شد، عدم بهبود در معیار F1 میکرو در این روش نسبت به برخی روش‌ها از جمله KNN به دلیل نسبت بالای کاهش حجم نمونه‌ها در روش ارائه شده نسبت به سایر روش‌ها می‌باشد. بهبود دقت در MLPS-KELM در معیار F1 ماکرو نسبت به روش‌های KNN،

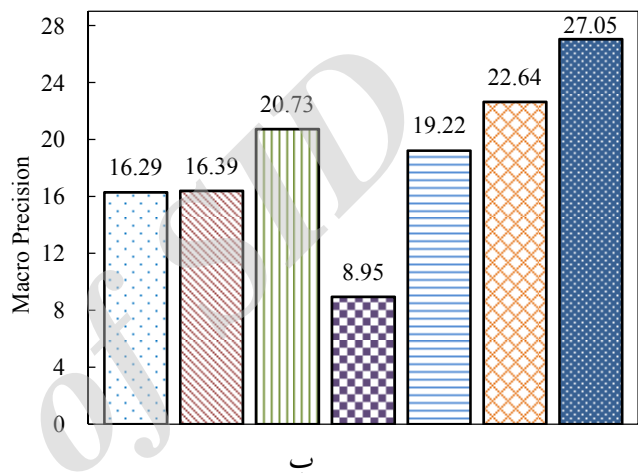
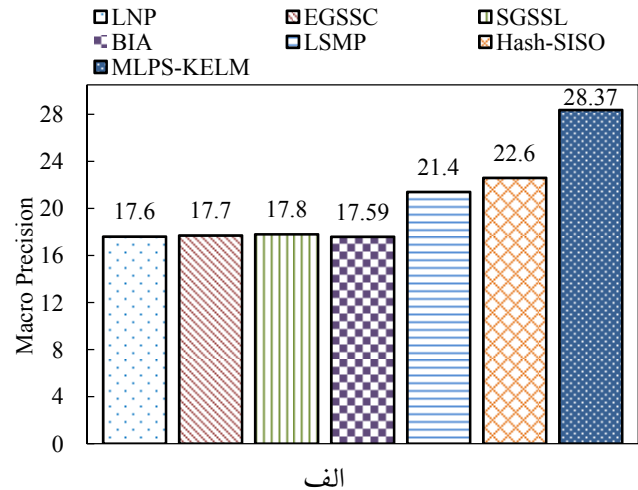
محدودیت کاهش حجم مجموعه آموزش با ارائه نسخه‌ای بهبودیافته از روش انتخاب نمونه‌های اولیه ENN می‌باشد. از آنجایی‌که روش یادگیری KELM دارای عملکرد قدرتمندی در پیش‌بینی برجسب‌ها می‌باشد، در مرحله تصمیم‌گیری روش ارائه‌شده برای انتخاب نمونه‌های اولیه از KELM نیز استفاده شده است.

به بیان دقیق‌تر، در این پژوهش دو موضوع مهم مورد بررسی قرار گرفت. (۱) حل مشکلات و محدودیت‌های حافظه در روش KELM و به عبارتی کاربردی سازی الگوریتم KELM مبتنی بر انتخاب نمونه‌های اولیه و (۲) ارائه الگوریتم انتخاب نمونه‌های اولیه بهبودیافته چند برجسب با به‌کارگیری روش پیش‌بینی برجسب در مرحله تصمیم‌گیری حذف نمونه‌ها، دو نوآوری اصلی ارائه‌شده در حل مشکلات مطرح‌شده در KELM و انتخاب نمونه‌های اولیه در مقایسه با سایر روش‌ها با انجام آزمایش‌های متعدد بر روی دادگان مقیاس بزرگ در شرح‌گذاری خودکار مجموعه تصاویر در مقیاس بزرگ نشان داده شد.

در پژوهش‌های آتی بر آنیم تا با ترکیب رویکرد انتخاب نمونه‌های اولیه، KELM و ویژگی‌ها قدم موثرتری در حل مشکلات و محدودیت‌های حافظه و دقت روش KELM برداریم. بدین ترتیب با به‌کارگیری هم‌زمان عملیات انتخاب نمونه‌های اولیه و انتخاب ویژگی، عملکرد انتخاب نمونه‌های اولیه و همچنین دقت طبقه‌بندی بهبود می‌یابد.

مراجع

- [1] M.-L. Zhang and Z.-H. Zhou, "A review on multi-label learning algorithms," *Knowl. Data Eng. IEEE Trans.*, vol. 26, no. 8, pp. 1819–1837, 2014.
- [2] R. Li, J. Lu, Y. Zhang, Z. Lu, and W. Xu, "A framework of large-scale and real-time image annotation system," in *Artificial Intelligence, 2009. ICAI'09. International Joint Conference on*, 2009, pp. 576–579.
- [3] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.
- [4] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, 2004, vol. 2, pp. 985–990.
- [5] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *Syst. Man, Cybern. Part B Cybern. IEEE Trans.*, vol. 42, no. 2, pp. 513–529, 2012.
- [6] W. Deng, Q. Zheng, and K. Zhang, "Reduced kernel extreme learning machine," in *Proceedings of the 8th*



شکل ۶ نمودار مقایسه عملکرد شرح‌گذاری روش MLPS-KELM ارائه‌شده با سایر روش‌های شرح‌گذاری در مجموعه تصاویر NUS-WIDE-Lite (الف) و NUS-WIDE (ب).

همانطور که نمودارهای شکل ۶ نشان می‌دهد، عملکرد شرح‌گذاری روش MLPS-KELM ارائه شده از سایر روش‌های ارائه شده در شرح‌گذاری خودکار مجموعه تصاویر در مقیاس بزرگ بهتر می‌باشد. این در حالی است که علاوه بر دقت بالای شرح‌گذاری، روش ارائه شده عملکرد خوبی در کاهش حجم مجموعه آموزش نیز دارا می‌باشد. از اینرو این روش صلاحیت خوبی در به‌کارگیری در محیط‌های مقیاس بزرگ و واقعی دارد. بهبود عملکرد این روش نسبت به روش‌های LNP، EGSSC، SGSSL، LSMP، BIA و HASH-SISO در مجموعه تصویر NUS-WIDE-Lite به ترتیب برابر ۱۰/۷۷٪، ۱۰/۶۷٪، ۱۰/۵۷٪، ۶/۹۷٪، ۱۰/۷۸٪، ۱۰/۵۷٪، ۱۰/۷۶٪، ۱۰/۶۶٪، ۶/۳۲٪، ۱۸/۱٪، ۷/۸۳٪ و ۴/۴۱٪ می‌باشد.

۶ نتیجه‌گیری و پیشنهاد پژوهش

در این پژوهش محدودیت حافظه در به‌کارگیری روش KELM در کاربردهای مقیاس بزرگ با استفاده از روش انتخاب نمونه‌های اولیه چندبرجسب مورد بررسی قرار گرفت. رویکرد اصلی حل این

- [22] N. García-Pedrajas, J. A. R. Del Castillo, and D. Ortiz-Boyer, "A cooperative coevolutionary algorithm for instance selection for instance-based learning," *Mach. Learn.*, vol. 78, no. 3, pp. 381–420, 2010.
- [23] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to Platt's SMO algorithm for SVM classifier design," *Neural Comput.*, vol. 13, no. 3, pp. 637–649, 2001.
- [24] G. Wu, Z. Zhang, and E. Y. Chang, "Kronecker Factorization for Speeding up Kernel Machines," in *SDM*, 2005, pp. 611–615.
- [25] K.-M. Lin and C.-J. Lin, "A study on reduced support vector machines," *Neural Networks, IEEE Trans.*, vol. 14, no. 6, pp. 1449–1459, 2003.
- [26] Q. Liu, Q. He, and Z. Shi, "Extreme support vector machine classifier," in *Advances in Knowledge Discovery and Data Mining*, Springer, 2008, pp. 222–233.
- [27] Z. Lin, J. Cao, T. Chen, Y. Jin, Z.-L. Sun, and A. Lendasse, "Extreme Learning Machine on High Dimensional and Large Data Applications," *Math. Probl. Eng.*, vol. 501, p. 624903, 2015.
- [28] F. Cao, B. Liu, and D. S. Park, "Image classification based on effective extreme learning machine," *Neurocomputing*, vol. 102, pp. 90–97, 2013.
- [29] Y. Kongsorot, P. Horata, and K. Sunat, "Applying Regularization Least Squares Canonical Correlation Analysis in Extreme Learning Machine for Multi-label Classification Problems," in *Proceedings of ELM-2014 Volume 1*, Springer, 2015, pp. 377–396.
- [30] W. Zong and G.-B. Huang, "Face recognition based on extreme learning machine," *Neurocomputing*, vol. 74, no. 16, pp. 2541–2551, 2011.
- [31] B. Lu, G. Wang, Y. Yuan, and D. Han, "Semantic concept detection for video based on extreme learning machine," *Neurocomputing*, vol. 102, pp. 176–183, 2013.
- [32] W. Zheng, Y. Qian, and H. Lu, "Text categorization based on regularization extreme learning machine," *Neural Comput. Appl.*, vol. 22, no. 3–4, pp. 447–456, 2013.
- [33] L. Feng, J. Wang, S. Liu, and Y. Xiao, "Multi-label dimensionality reduction and classification with extreme learning machines," *Syst. Eng. Electron. J.*, vol. 25, no. 3, pp. 502–513, 2014.
- [34] J. Kim, H. Shin, Y. Lee, and M. Lee, "Algorithm for classifying arrhythmia using Extreme Learning Machine and principal component analysis," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 3257–3260.
- [35] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *International Conference on Computer Recognition Systems CORES 2013*, 2013, pp. 63–69.
- [7] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Advances in neural information processing systems*, 2007, pp. 1177–1184.
- [8] A. Vedaldi and A. Zisserman, "Efficient additive kernels via explicit feature maps," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 34, no. 3, pp. 480–492, 2012.
- [9] J. Chanussot, I. Nyström, and N. Sladoje, "Shape signatures of fuzzy star-shaped sets based on distance from the centroid," *Pattern Recognit. Lett.*, vol. 26, no. 6, pp. 735–746, 2005.
- [10] S. Garcia, J. Derrac, J. R. Cano, and F. Herrera, "Prototype Selection for Nearest Neighbor Classification: Taxonomy and Empirical Study," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 34, no. 3, pp. 417–435, 2012.
- [11] F. Charte, A. J. Rivera, M. J. del Jesus, and F. Herrera, "MLeNN: A First Approach to Heuristic Multilabel Undersampling," in *Intelligent Data Engineering and Automated Learning—IDEAL 2014*, 2014, pp. 1–9.
- [12] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*. Springer, 2015.
- [13] P. E. Hart, "The Condensed Nearest Neighbor Rule," *IEEE Trans. Inf. Theory*, vol. 14, no. 3, pp. 515–516, 1968.
- [14] V. S. Devi and M. N. Murty, "An incremental prototype set building technique," *Pattern Recognit.*, vol. 35, no. 2, pp. 505–513, 2002.
- [15] F. Angiulli, "Fast Nearest Neighbor Condensation for Large Data Sets Classification," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 11, pp. 1450–1464, 2007.
- [16] D. L. Wilson, "Asymptotic properties of nearest neighbor rules using edited data," *Syst. Man Cybern. IEEE Trans.*, no. 3, pp. 408–421, 1972.
- [17] I. Tomek, "An experiment with the edited nearest-neighbor rule," *IEEE Trans. Syst. Man. Cybern.*, no. 6, pp. 448–452, 1976.
- [18] J. S. Sánchez, F. Pla, and F. J. Ferri, "Prototype selection for the nearest neighbour rule through proximity graphs," *Pattern Recognit. Lett.*, vol. 18, no. 6, pp. 507–513, 1997.
- [19] S. García, J. R. Cano, and F. Herrera, "A memetic algorithm for evolutionary prototype selection: A scaling up approach," *Pattern Recognit.*, vol. 41, no. 8, pp. 2693–2709, 2008.
- [20] D. B. Skalak, "Prototype and feature selection by sampling and random mutation hill climbing algorithms," in *Proceedings of the eleventh international conference on machine learning*, 1994, pp. 293–301.
- [21] J. C. Riquelme, J. S. Aguilar-Ruiz, and M. Toro, "Finding representative patterns with ordered projections," *Pattern Recognit.*, vol. 36, no. 4, pp. 1009–1018, 2003.



حمید کارگر شورکی کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۹۲ از دانشگاه بیرجند دریافت کردند. سپس کارشناسی ارشد خود را در همان رشته و گرایش در سال ۱۳۹۴ از دانشگاه یزد اخذ نمودند. حوزه‌های پژوهشی ایشان شامل یادگیری ماشین، بینایی ماشین (شرح‌گذاری خودکار تصاویر، بازیابی تصاویر و نظارت دیداری) است.



محمدعلی زارع چاهوکی کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۷۸ از دانشگاه شهید بهشتی دریافت کردند. سپس کارشناسی ارشد و دکتری خود را در همان رشته و گرایش در سال‌های ۱۳۸۳ و ۱۳۹۲ از دانشگاه تربیت مدرس اخذ نمودند. در حال حاضر ایشان استادیار گروه مهندسی کامپیوتر دانشگاه یزد هستند. حوزه‌های پژوهشی ایشان شامل یادگیری ماشین، مهندسی نرم‌افزار (متدولوژی و کاربردهای مرتبط با یادگیری)، بینایی ماشین (شرح‌گذاری خودکار تصاویر، بازیابی تصاویر و نظارت دیداری) و نهم‌نگاری است.

Neural Networks, IEEE Trans., vol. 17, no. 4, pp. 879–892, 2006.

- [36] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *Neural Networks, IEEE Trans.*, vol. 20, no. 8, pp. 1352–1357, 2009.
- [37] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *Neural Networks, IEEE Trans.*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [38] S.-S. Choi, S.-H. Cha, and C. C. Tappert, "A survey of binary similarity and distance measures," *J. Syst. Cybern. Informatics*, vol. 8, no. 1, pp. 43–48, 2010.
- [39] T. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng, "NUS-WIDE: A Real-World Web Image Database from National University of Singapore," in *Proceedings of the ACM International Conference on Image and Video Retrieval*, 2009, pp. 0–8.
- [40] Y. Yang, F. Wu, F. Nie, H. T. Shen, Y. Zhuang, and A. G. Hauptmann, "Web and Personal Image Annotation by Mining Label Correlation With Relaxed Visual Graph Embedding," *IEEE Trans. image Process.*, vol. 21, no. 3, pp. 1339–1351, 2012.
- [41] Z. Ma, F. Nie, Y. Yang, J. R. R. Uijlings, N. Sebe, and S. Member, "Web Image Annotation Via Subspace-Sparsity Collaborated Feature Selection," *IEEE Trans. Multimed.*, vol. 14, no. 4, pp. 1021–1030, 2012.
- [42] A. Özgür, L. Özgür, and T. Güngör, "Text categorization with class-based and corpus-based keyword selection," in *Computer and Information Sciences-ISCIS 2005*, Springer, 2005, pp. 606–615.
- [43] F. Wang and C. Zhang, "Label propagation through linear neighborhoods," *Knowl. Data Eng. IEEE Trans.*, vol. 20, no. 1, pp. 55–67, 2008.
- [44] A. Subramanya and J. A. Bilmes, "Entropic graph regularization in non-parametric semi-supervised classification," in *Advances in Neural Information Processing Systems*, 2009, pp. 1803–1811.
- [45] J. Tang, R. Hong, S. Yan, and T. Chua, "Image Annotation by k NN-Sparse Graph-based Label Propagation over Noisily-Tagged Web Images," vol. 1, no. 1, pp. 111–126, 2010.
- [46] X. Chen, Y. Mu, S. Yan, and T.-S. Chua, "Efficient large-scale image annotation by probabilistic collaborative multi-label propagation," in *Proceedings of the international conference on Multimedia*, 2010, pp. 35–44.
- [47] A. Makadia, V. Pavlovic, and S. Kumar, "Baselines for image annotation," *Int. J. Comput. Vis.*, vol. 90, no. 1, pp. 88–105, 2010.
- [48] J. Huang, H. Liu, J. Shen, and S. Yan, "Towards efficient sparse coding for scalable image annotation," in *Proceedings of the 21st ACM international conference on Multimedia*, 2013, pp. 947–956.