

حداقل سازی خطای بازسازی تصاویر از طریق تطبیق توزیع و محدودیت رتبه-پایین

سمانه رضائی^۱، دانشجوی کارشناسی ارشد؛ جعفر طهمورث نژاد^۲، استادیار؛ وحید سلوک^۳، استادیار

۱- دانشکده مهندسی فناوری اطلاعات و کامپیوتر - دانشگاه صنعتی ارومیه - ارومیه - ایران - samanehrezaei@it.uut.ac.ir

۲- دانشکده مهندسی فناوری اطلاعات و کامپیوتر - دانشگاه صنعتی ارومیه - ارومیه - ایران - j.tahmores@it.uut.ac.ir

۳- دانشکده مهندسی فناوری اطلاعات و کامپیوتر - دانشگاه صنعتی ارومیه - ارومیه - ایران - v.solouk@it.uut.ac.ir

چکیده: هدف از تطبیق دامنه تصویری، یادگیری مدل‌های مقاوم برای داده‌های آزمایشی، با استفاده از انتقال دانش از داده‌های آموزشی است. در حالی که مجموعه‌های آموزشی و آزمایشی دارای توزیع‌های متفاوتی هستند. روش‌های موجود تلاش می‌کنند تا مسئله تغییر دامنه‌ها را با استفاده از تطبیق دامنه‌ها یا اعمال محدودیت‌های رتبه-پایین حل نمایند. در این مقاله، ما یک روش دو مرحله‌ای غیرنظارت شده با عنوان حداقل سازی خطای بازسازی تصاویر از طریق تطبیق توزیع و محدودیت رتبه-پایین پیشنهاد می‌دهیم که از هر دو روش تطبیق توزیع‌ها و محدودیت‌های رتبه-پایین برای فائق آمدن به اختلاف توزیع دامنه‌ها استفاده می‌کند. در مرحله اول، روش پیشنهادی ما داده‌های آموزشی و آزمایشی را به یک زیرفضای مشترک نگاشت می‌کند تا اختلاف توزیع حاشیه‌ای و شرطی دامنه‌ها حداقل شود. علاوه بر آن، EDA از خوشه‌بندی مستقل از دامنه برای تفکیک بین کلاس‌های مختلف بهره می‌برد. در مرحله دوم، برای حفظ ساختار داده در زیرفضای مشترک، EDA خطای بازسازی داده‌ها را با استفاده از محدودیت‌های رتبه-پایین و تنگ حداقل می‌کند. به طور کلی، EDA مسئله اختلاف دامنه‌ها را با پیچیدگی زمانی درجه سه حل می‌کند. روش پیشنهاد شده بر روی تنوعی از پایگاه داده‌های شناخته شده بصری ارزیابی می‌شود و کارایی آن با دیگر روش‌های به روز تطبیق دامنه‌ها مقایسه می‌شود. میانگین دقت EDA بر روی ۳۲ آزمایش ۶۸/۳۳٪ به دست آمده که نسبت به دیگر روش‌های به روز تطبیق دامنه، با بهبود ۴/۲۸٪ عملکرد بهتری دارد.

واژه‌های کلیدی: تطبیق دامنه تصویری، مسئله شیفت دامنه، تطبیق توزیع، محدودیت رتبه-پایین، تطبیق دامنه.

Image Reconstruction Error Minimization via Distribution Adaptation and Low-Rank Constraint

S. Rezaei¹, MSc Student; J. Tahmoresnezhad², Assistant professor; V. Solouk³, Assistant professor

1- Faculty of IT & Computer Engineering, Urmia University of Technology, Urmia, Iran, Email: samanehrezaei@it.uut.ac.ir

2- Faculty of IT & Computer Engineering, Urmia University of Technology, Urmia, Iran, Email: j.tahmores@it.uut.ac.ir

3- Faculty of IT & Computer Engineering, Urmia University of Technology, Urmia, Iran, Email: v.solouk@it.uut.ac.ir

Abstract: Visual domain adaptation aims to learn robust models for the test data by knowledge transferring from a training data where the training and test sets are from different distributions. Existing approaches attempt to solve domain shift problem with either adaptation across domains or performing low-rank constraints. In this paper, we propose a two-phases unsupervised approach referred as image reconstruction Error minimization via Distribution Adaptation and low-rank constraint (EDA), which benefits from both the distribution adaptation and the low-rank constraints to tackle distribution mismatch across domains. In the first phase, our proposed approach projects the training and test data onto a common subspace in which the marginal and conditional distribution differences of domains are minimized. Moreover, EDA benefits from domain invariant clustering to discriminate between various classes of data. In the second phase, for preserving data structure in the shared subspace, EDA minimizes the data reconstruction error using low-rank and sparse constraints. Overall, EDA solves the domain mismatch problem in cubic time complexity. The proposed approach is evaluated on variety of visual benchmark datasets and its performance is compared with the other state-of-the-art domain adaptation methods. The average accuracy of EDA on 32 experiments is determined 68.33% where outperforms other state-of-the-art domain adaptation methods with 4.28% improvement.

Keywords: Visual domain adaptation, Domain shift problem, Distribution adaptation, Low-rank constraint, Domain adaptation.

تاریخ ارسال مقاله: ۱۳۹۶/۱۲/۲۵

تاریخ اصلاح مقاله: ۱۳۹۷/۰۷/۱۵

تاریخ پذیرش مقاله: ۱۳۹۷/۱۲/۱۳

نام نویسنده مسئول: وحید سلوک

نشانی نویسنده مسئول: ایران - ارومیه - دانشگاه صنعتی ارومیه - دانشکده مهندسی فناوری اطلاعات و کامپیوتر

۱- مقدمه

میزان اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی می‌تواند متفاوت بوده و منشأهای مختلفی داشته باشد. در برخی مسائل، میزان اختلاف توزیع بین داده‌های آموزشی و آزمایشی قابل توجه نبوده و تأثیر چندانی بر کارایی مدل ندارد. به همین دلیل اعمال پردازش مازاد به سیستم جهت به حداقل رساندن اختلاف توزیع، مقرون به صرفه نمی‌باشد. اما در برخی مدل‌ها میزان اختلاف توزیع به اندازه‌ای است که میانگین صحت به دست آمده از مدل یادگیری ماشین، کمتر از میانگین صحت به دست آمده از مدل‌های تصادفی می‌باشد. برای اندازه‌گیری اختلاف توزیع بین دامنه‌ها، معیارهای مختلفی از جمله MMD ، KLD و BD وجود دارد. روش‌های KLD [۳] و BD [۴] از جمله روش‌های پارامتری و دو مرحله‌ای برای محاسبه میزان اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی هستند. اما روش MMD یک روش غیرپارامتری و تک مرحله‌ای است که بدون نیاز به محاسبه تابع چگالی احتمال داده‌های آموزشی و آزمایشی، اختلاف توزیع آن‌ها را در فضای $RKHS$ محاسبه می‌نماید [۵].

منشأ وجودی اختلاف توزیع بین داده‌های آموزشی و آزمایشی می‌تواند از اختلاف توزیع حاشیه‌ای یا اختلاف توزیع شرطی بین دامنه‌ها نشأت گیرد. در صورتی که دامنه‌های آموزشی و آزمایشی، فضای ویژگی یکسانی داشته باشند ولی احتمال رخداد ویژگی‌های یکسان در دامنه‌های مختلف، متفاوت باشد، اختلاف توزیع حاشیه‌ای رخ می‌دهد. اما در صورتی که توزیع برچسب‌های کلاس‌های مختلف در دو دامنه، متفاوت باشد اختلاف توزیع شرطی بین دامنه‌ای وجود دارد. برخی روش‌های یادگیری انتقالی [۳، ۶، ۷] با نگاشت داده‌های آموزشی و آزمایشی به یک زیرفضای جدید، اختلاف توزیع حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی را کاهش می‌دهند. مجموعه دیگری از روش‌ها [۸، ۹، ۱۰، ۱۱]، با کاهش اختلاف توزیع شرطی بین دامنه‌ها، آن‌ها را در مقابل تغییرات داده‌ای مقاوم می‌سازند. اما در صورتی که اختلاف توزیع بین دامنه‌ها زیاد باشد (برای نمونه، کلاسه‌بندی تصاویر) کاهش اختلاف توزیع حاشیه‌ای یا شرطی کافی نبوده و نیاز است که هر دو اختلاف توزیع هم‌زمان کاسته شوند. همچنین، از آنجایی که مدل ساخته شده بایستی بتواند داده‌های آزمایشی را با دقت بالایی کلاسه‌بندی نماید، علاوه بر کاهش اختلاف توزیع حاشیه‌ای و شرطی بین دامنه‌ها، مدل باید بتواند حاشیه کلاسه‌بندی مناسب و کافی بین کلاس‌های مختلف ایجاد کند. خوشه‌بندی مستقل از دامنه^{۱۲} می‌تواند داده‌ها را به گونه‌ای به زیرفضای جدید نگاشت کند که نمونه‌های هم‌کلاس از دامنه‌های مختلف در خوشه‌های یکسان قرار گیرند. بدین ترتیب، با استفاده از یک مدل کلاسه‌بند کلاسیک می‌توان برچسب نمونه‌های آزمایشی را با دقت بالایی پیش‌بینی نمود [۹].

نگاشت داده‌های آموزشی و آزمایشی به زیرفضای جدید، باید به گونه‌ای باشد که ساختار داده‌های آموزشی و آزمایشی از جمله واریانس و ساختار هندسی داده‌ها، در این زیرفضا حفظ شود [۱۲]. بدین منظور، برای هر نمونه دامنه آزمایشی در زیرفضای جدید، یک همسایگی

مسائل یادگیری در سیستم‌های هوشمند در دو حوزه هوش ماشینی و هوش مصنوعی مورد بررسی قرار می‌گیرند. در حوزه هوش ماشینی که جزو زیرشاخه‌های الکترونیک می‌باشد، یادگیری ماشین و پردازش تصویر به دنبال یافتن الگوریتم‌هایی برای کلاسه‌بندی داده‌ها و تصاویر با هدف بهبود یادگیری سیستم‌های هوشمند بوده، به طوری که جنبه سخت افزاری سیستم‌ها نیز مدنظر قرار می‌گیرد. در حوزه هوش مصنوعی، به‌عنوان زیرشاخه‌ای از رشته کامپیوتر، با بهره‌گیری از الگوریتم‌های یادگیری ماشین و پردازش تصویر، مدل‌هایی با کارایی بالا بدون در نظر گرفتن سخت افزار مورد استفاده و با استفاده از داده‌های آموزشی، آموزش داده می‌شود. این حوزه، پیوسته به دنبال بهبود کارایی مدل به‌منظور پیش‌بینی برچسب داده‌های آزمایشی می‌باشد. با این حال، ضروری است که در زمان ایجاد مدل، تعداد قابل قبولی از داده‌های آموزشی در دسترس بوده و داده‌های آموزشی و آزمایشی مشابهی برخوردار باشند، زیرا در صورتی که داده‌های آموزشی و آزمایشی از دو توزیع احتمالی متفاوت به دست آمده باشند، کارایی مدل به شدت تحت تأثیر قرار می‌گیرد [۱]. برای نمونه، فرض شود یک مدل بر روی تصاویر به دست آمده از دوربین‌های پلیس در روز ایجاد شده است. حال این مدل جهت کلاسه‌بندی تصاویر ثبت شده از همان دوربین‌ها در شب استفاده شود. بدیهی است که به دلیل وجود تفاوت‌های اساسی در دو مجموعه تصویر از جمله میزان نور، شفافیت، کیفیت و جهت تابش نور خورشید، مدل ایجاد شده در روز پاسخگوی تصاویر گرفته شده در شب نخواهد بود. به همین دلیل، مدل ایجاد شده خطای بالایی داشته و برای کاربردهای دنیای واقعی مناسب نخواهد بود. به چنین مشکلی در مسائل یادگیری ماشین و پردازش تصویر، شیفت دامنه‌ها^۱ یا مسائل بین دامنه‌ای^۲ گفته می‌شود.

یادگیری انتقالی^۲ و تطبیق دامنه‌ها^۳ از جمله روش‌های شناخته شده برای مقابله با مشکل شیفت دامنه‌ها در کلاسه‌بندی تصاویر می‌باشند. در واقع، یادگیری انتقالی با کاهش اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی، کارایی مدل‌های کلاسه‌بند را افزایش می‌دهد [۲]. برای نمونه، یادگیری انتقالی، تصاویر گرفته شده توسط دوربین‌های پلیس در ساعات روز و شب را به یک زیرفضای مشترک منتقل کرده تا اختلافات آماری^۴ ایجاد شده توسط محیط به حداقل برسد. بدین ترتیب، توزیع نمونه‌های آموزشی و آزمایشی در زیرفضای جدید به یکدیگر نزدیک شده و کارایی مدل ایجاد شده بهبود قابل ملاحظه‌ای پیدا می‌کند. همچنین یادگیری انتقالی با تطبیق پارامترهای مدل‌های آموزشی (برای مثال مدل SVM^۵ یا NN^۶) آن‌ها را در مقابل تغییرات دامنه‌ای مقاوم‌تر می‌کند. بدین ترتیب، در صورتی که مدل با داده‌های آزمایشی مواجه شود که نسبت به داده‌های آموزشی از اختلاف توزیع برخوردار هستند می‌تواند مقاومت بیشتری در مقابل افت بازدهی از خود نشان دهد.

شده و یا وزن‌دهی مجدد می‌شوند. روش انتخاب لندمارک^{۱۷}، یکی از روش‌های مبتنی بر نمونه می‌باشد [۱۴]. مجموعه لندمارک، شامل نمونه‌هایی از دامنه آموزشی بوده که دارای کم‌ترین اختلاف توزیع با نمونه‌های دامنه آزمایشی می‌باشند. (۲) روش‌های مبتنی بر مدل که در آن ابتدا یک کلاسه‌بند توسط داده‌های دامنه آموزشی آموزش داده شده و سپس پارامترهای مدل، به‌عنوان پارامترهای اولیه کلاسه‌بند داده‌های آزمایشی، مورد استفاده قرار می‌گیرند. (۳) روش‌های مبتنی بر ویژگی که در آن به‌منظور کاهش اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی، داده‌ها به فضای ویژگی مشترک با اختلاف توزیع کم‌تر نگاشت می‌شوند. نگاشت داده‌ها به زیرفضای جدید باید به‌گونه‌ای باشد که ساختار داده‌ها در زیرفضای جدید، حفظ شود. در فضای ویژگی جدید، یک کلاسه‌بند توسط داده‌های آموزشی آموزش داده می‌شود و سپس این مدل برای کلاسه‌بندی داده‌های آزمایشی مورد استفاده قرار می‌گیرد. اخیراً روش‌های ارائه‌شده مبتنی بر ویژگی، به دو دسته طبقه‌بندی شده‌اند [۱۳]: (۱) روش‌های مبتنی بر کاهش اختلاف توزیع بین داده‌های دامنه آموزشی و آزمایشی، (۲) روش‌های مبتنی بر بازسازی داده‌های آزمایشی توسط داده‌های آموزشی. مجموعه روش‌های دسته اول، به دنبال یافتن یک زیرفضای مشترک می‌باشند به‌طوری‌که در زیرفضای جدید، اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی، حداقل باشد. روش TJM^{18} ، ابتدا داده‌های آموزشی و آزمایشی را به یک زیرفضای کم بعد نگاشت کرده به‌طوری‌که اختلاف توزیع حاشیه‌ای بین داده‌های آموزشی و آزمایشی کاهش پیدا کند و سپس به نمونه‌هایی از دامنه آموزشی که بیش‌ترین مشابهت را از نظر توزیع با دامنه آزمایشی داشته باشند، با استفاده از نرم $l_{2,1}^{19}$ وزن بیش‌تری اختصاص می‌دهد [۶]. روش دیگر با عنوان VDA^{20} ، با کاهش ابعاد داده‌های آموزشی و آزمایشی به دنبال کاهش اختلاف توزیع حاشیه‌ای و شرطی بوده و همچنین در فضای جدید با روش خوشه‌بندی مستقل از دامنه، شکل هندسی و آماری داده‌های آموزشی در هر کلاس هدف را حفظ کرده و باعث افزایش دقت مدل بر روی داده‌های آزمایشی می‌شود [۹].

یکی از روش‌های دسته دوم با عنوان $LRSR^{21}$ ، به دنبال یافتن زیرفضایی است که در این زیرفضا هر یک از داده‌های آزمایشی به‌صورت خطی توسط داده‌های آموزشی قابل بازسازی بوده و خطای بازسازی داده‌های آزمایشی توسط داده‌های آموزشی کاهش یابد. بازسازی داده‌های آزمایشی توسط داده‌های آموزشی، توسط ماتریس بازسازی انجام می‌شود و بدین ترتیب اطلاعات عمومی^{۲۲} و محلی^{۲۳} داده‌ها به‌طور هم‌زمان حفظ می‌شود [۱۳، ۱۵]. در این مجموعه از روش‌ها کاهش اختلاف توزیع بین داده‌های آموزشی و آزمایشی در طول کاهش بعد نادیده گرفته می‌شود.

روش پیشنهادی در این مقاله، یک روش ترکیبی از دو دسته روش مبتنی بر کاهش اختلاف توزیع و مبتنی بر بازسازی داده‌ها می‌باشد. در روش پیشنهادی، ابتدا داده‌های دامنه آموزشی و آزمایشی به یک زیرفضا با ابعاد کم‌تر نگاشت شده به‌طوری‌که اختلاف توزیع حاشیه‌ای و شرطی

مشخص در این دامنه در نظر گرفته شده و این نمونه توسط همان همسایگی در دامنه آموزشی، بازسازی می‌شود. برای بازسازی نمونه‌های دامنه آزمایشی، از ماتریس ضرایب بازسازی^{۲۴} استفاده می‌شود. هر سطر از این ماتریس، نماینده یک نمونه از دامنه آموزشی بوده و هر ستون از این ماتریس، نشان‌گر یک نمونه از دامنه آزمایشی می‌باشد. مقادیر ماتریس ضرایب بازسازی، نشان دهنده میزان ارتباط هر نمونه از دامنه آموزشی و آزمایشی، با یکدیگر است. به‌منظور مرتبط شدن نمونه‌های یک کلاس در دامنه آموزشی و آزمایشی با یکدیگر، بایستی در ماتریس ضرایب بازسازی، نمایش تنک^{۲۴} و رتبه-پایین^{۲۵} مورد استفاده قرار گیرند [۱۳].

در این مقاله برای حل مسائل یادگیری انتقالی، روشی نوین با عنوان حداقل‌سازی خطای بازسازی داده‌ها از طریق تطبیق توزیع^{۱۶} و محدودیت رتبه-پایین (EDA) پیشنهاد می‌شود. روش پیشنهاد شده، دو هدف زیر را دنبال می‌کند: ۱- کاهش اختلاف توزیع حاشیه‌ای و شرطی بین داده‌های دامنه آموزشی و آزمایشی، ۲- بازسازی هر نمونه از دامنه آزمایشی در یک همسایگی مفروض، توسط نمونه‌های دامنه آموزشی در زیرفضای جدید. در مرحله اول روش EDA، با استفاده از روش تطبیق ویژگی، داده‌های آموزشی و آزمایشی به یک زیرفضای کم بعد نگاشت می‌شوند. در زیرفضای جدید، اختلاف توزیع حاشیه‌ای و شرطی بین دامنه‌های آموزشی و آزمایشی حداقل شده و از خوشه‌بندی مستقل از دامنه برای افزایش تفکیک‌پذیری بین کلاس‌های مختلف از یکدیگر استفاده می‌شود. در مرحله دوم، برای حفظ ساختار هندسی داده‌ها در زیرفضای جدید، داده‌های آزمایشی توسط داده‌های آموزشی، با استفاده از ماتریس ضرایب، بازسازی می‌شوند. به‌منظور کاهش خطای بازسازی مجدد داده‌ها در این زیرفضا، ماتریس ضرایب بازسازی داده‌ها، با اعمال محدودیت رتبه-پایین و نمایش تنک به‌دست می‌آید.

در ادامه، مقاله به‌صورت زیر سازماندهی شده‌است. در بخش دوم، روش‌های پیشین مورد مطالعه قرار گرفته‌است. در بخش سوم، روش ارائه‌شده در این مقاله به‌تفصیل توضیح داده شده‌است. در بخش چهارم، نتایج حاصل از روش ارائه‌شده با سایر روش‌های موجود، مورد مقایسه قرار گرفته‌است. در بخش پنجم، نتیجه‌گیری و پیشنهادات برای کارهای آتی گنجانده شده‌است.

۲- کارهای پیشین

به‌منظور بهبود صحت کلاسه‌بندی داده‌های دامنه آزمایشی، مجموعه متنوعی از روش‌های یادگیری انتقالی از طریق تطبیق دامنه‌ها، مورد استفاده قرار می‌گیرند. این مجموعه از روش‌ها به دنبال کاهش اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی می‌باشند. به‌طور کلی، روش‌های ارائه‌شده در حوزه تطبیق دامنه به سه دسته کلی تقسیم می‌شوند: (۱) روش‌های مبتنی بر نمونه که در آن به‌منظور کاهش اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی، نمونه‌هایی از دامنه آموزشی که موجب کاهش اختلاف توزیع داده‌های آموزشی و آزمایشی می‌شوند، انتخاب

اصلی داده‌ها، مورد جستجو قرار می‌گیرد. در زیرفضای جدید، اختلاف توزیع حاشیه‌ای و شرطی بین داده‌های آموزشی و آزمایشی حداقل شده و تفکیک پذیری بین کلاس‌ها در این زیرفضا افزایش می‌یابد. در شکل ۲، مرحله دوم EDA به صورت شماتیک نشان داده شده است. در مرحله دوم، داده‌های آزمایشی نگاشت شده توسط مرحله اول، توسط ماتریس ضرایب بازسازی رتبه-پایین و تنک، به گونه‌ای بازسازی می‌شوند که خطای بازسازی داده‌های آزمایشی نگاشت یافته توسط داده‌های آموزشی نگاشت یافته، حداقل شوند. در این مرحله به علت حفظ ساختار عمومی و محلی داده‌ها، دقت کلاسه‌بند برای پیش‌بینی برچسب‌های داده‌های دامنه آزمایشی افزایش می‌یابد.

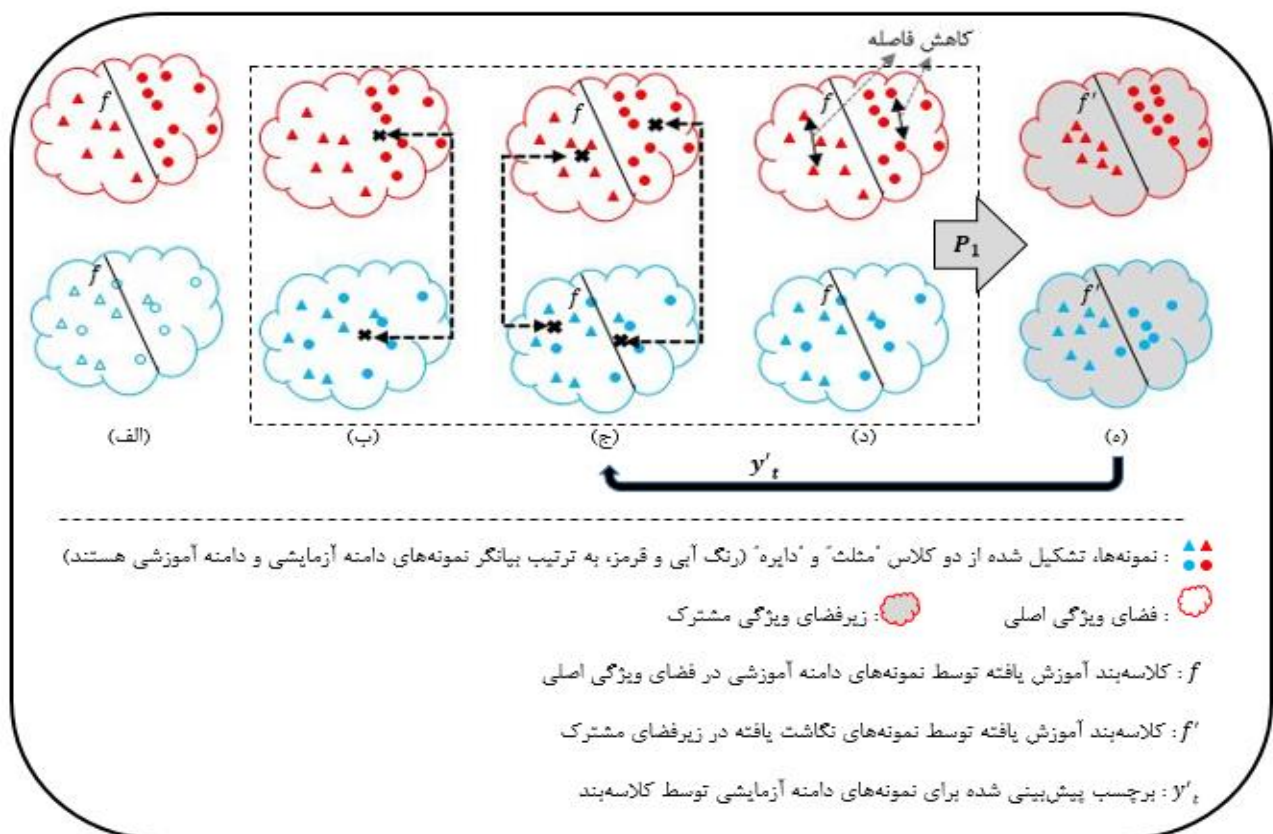
بین دامنه‌های آموزشی و آزمایشی، کاهش یابد. همچنین در زیرفضای جدید با بهره‌گیری از روش خوشه‌بندی مستقل از دامنه، فاصله بین کلاس‌های مختلف از یکدیگر افزایش می‌یابد. در مرحله دوم، به منظور کاهش خطای بازسازی داده‌ها در زیرفضای جدید، با استفاده از روش رتبه-پایین و نمایش تنک، خطای بازسازی داده‌ها حداقل می‌شود.

۳- روش پیشنهادی

در این بخش، روش EDA برای حل مسئله یادگیری انتقالی، با جزئیات بیش‌تر توضیح داده شده است.

۳-۱- هدف تحقیق

در شکل ۱، مرحله اول از روش دو مرحله‌ای EDA، به صورت شماتیک نشان داده شده است. در مرحله اول، زیرفضایی با ابعاد کم‌تر از فضای



شکل ۱: نمایش مرحله اول در روش EDA به صورت شماتیک (نمایش بهتر به صورت رنگی). الف: EDA، به دنبال یافتن یک ماتریس نگاشت P_1 به منظور نگاشت دامنه‌های آموزشی و آزمایشی به یک زیرفضای مشترک است. نمونه‌های توخالی بیانگر نمونه‌های دامنه آزمایشی هستند که در ابتدا بدون برچسب هستند. ب: اختلاف توزیع حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی، با کاهش اختلاف بین میانگین تمام نمونه‌های دامنه‌های آموزشی و آزمایشی، محاسبه می‌شود. ج: اختلاف توزیع شرطی بین دامنه‌های آموزشی و آزمایشی به صورت اختلاف میانگین کلاس‌های مختلف محاسبه شده و کاهش داده می‌شود. د: با استفاده از روش خوشه‌بندی مستقل از دامنه، فاصله نمونه‌های دامنه آموزشی در هر کلاس از یکدیگر کاهش داده می‌شود. ه: در زیرفضای مشترک، اختلاف توزیع حاشیه‌ای و شرطی بین دامنه‌های آموزشی و آزمایشی، به حداقل رسیده و کلاس‌های مختلف در خوشه‌های مختلفی قرار گرفته‌اند. به عبارت دیگر، توزیع دامنه‌های آموزشی و آزمایشی به صورت آماری منطبق شده‌اند. در این مرحله، کلاسه‌بند NN بر روی نمونه‌های آموزشی، آموزش داده شده و سپس از این کلاسه‌بند به منظور برچسب‌گذاری دقیق‌تر برای نمونه‌های دامنه آزمایشی، استفاده می‌شود. بدین ترتیب، مرحله اول به صورت تکراری انجام شده و نمونه‌های دامنه آزمایشی، با دقت بیشتری برچسب‌گذاری می‌شوند.

۳-۳-۳- خوشه‌بندی مستقل از دامنه

با هدف افزایش دقت کلاسه‌بند آموزش داده شده توسط داده‌های دامنه آموزشی، نمونه‌های هر کلاس در دامنه آموزشی در یک خوشه^{۲۸} قرار داده شده و واریانس بین کلاس‌ها افزایش می‌یابد. در این صورت فاصله هر نمونه از میانگین نمونه‌های هر کلاس، کاهش یافته و تفکیک‌پذیری بین کلاس‌ها در دامنه آموزشی افزایش می‌یابد. فاصله نمونه‌های دامنه آموزشی از میانگین نمونه‌ها در هر کلاس در فضای

اصلی به صورت $S_W = \sum_{\forall c \in C} \sum_{x_i \in X_s^c} (x_i - \mu^c)^T (x_i - \mu^c)$ تعریف

می‌شود. در این رابطه، میانگین نمونه‌های کلاس c به صورت μ^c نشان داده شده و فاصله درون کلاسی در فضای جدید، به صورت رابطه ۳ تعریف می‌شود. به منظور افزایش تفکیک‌پذیری بین کلاس‌ها،

ماتریس نگاشت P_1 ، با هدف حداقل‌سازی مقدار D_3 در رابطه ۳ جستجو می‌شود.

$$D_3(X_s, \mu^c) = \arg \min_{P_1} (P_1^T S_W P_1) \quad (۳)$$

۳-۴- تطبیق توزیع هندسی

در روش‌های مبتنی بر بازسازی داده‌های آزمایشی توسط داده‌های آموزشی، فرض شده که هر داده در دامنه آزمایشی دارای یک همسایگی مشخص بوده و این داده توسط همان همسایگی در دامنه آموزشی بازسازی می‌شود. در این صورت، تابع توزیع احتمالی داده‌های آموزشی و آزمایشی با یکدیگر مشابه می‌باشند. به منظور بازسازی داده‌های آزمایشی، از ماتریس ضرایب بازسازی استفاده می‌شود. هر سطر از این ماتریس نشان‌گر یک داده از دامنه آموزشی بوده و هر ستون، نماینده یک داده از دامنه آزمایشی می‌باشد. مقادیر این ماتریس، فاصله هر نمونه آموزشی و آزمایشی از یکدیگر، در یک همسایگی مشخص می‌باشد. این مفهوم به صورت رابطه ۴ نوشته می‌شود.

$$P_2^T X_t = P_2^T X_s Z \quad (۴)$$

در رابطه ۴، Z ، ماتریس ضرایب بازسازی بوده و X_t و X_s به ترتیب، داده‌های آموزشی و آزمایشی نگاشت یافته در مرحله اول و با استفاده از ماتریس P_1 می‌باشند. $X_t \in R^{d \times n_t}$ و $X_s \in R^{d \times n_s}$ به ترتیب، داده‌های دامنه آموزشی و آزمایشی در زیرفضای جدید بوده که d ابعاد فضای جدید می‌باشد. به منظور بازسازی هر داده در دامنه آموزشی با یک همسایگی مشخص توسط همسایگی متناظر از دامنه آموزشی، بایستی ساختار ماتریس ضرایب بازسازی به صورت بلوک-شده^{۲۹} باشد [۱۳، ۱۹]. در صورتی که ماتریس ضرایب بازسازی، رتبه-پایین باشد، این شرط برقرار می‌شود. این مفهوم به صورت رابطه ۵، نوشته می‌شود:

باشد، $x_i, x_j \in X_t$ ، $(W_0)_{ij} = \frac{1}{n_t n_t}$ و در غیر این صورت، $(W_0)_{ij} = \frac{-1}{n_s n_t}$ می‌باشد.

۳-۳-۲- کاهش اختلاف توزیع شرطی

در صورتی که داده‌های دامنه آموزشی و آزمایشی با یکدیگر اختلاف توزیع زیادی داشته باشند، به منظور افزایش دقت کلاسه‌بند، بایستی علاوه بر کاهش اختلاف توزیع حاشیه‌ای، اختلاف توزیع شرطی نیز کاهش یابد [۹]. در روش MMD، به منظور محاسبه اختلاف توزیع شرطی، مجموع اختلاف میانگین نمونه‌های هر کلاس در دامنه آموزشی و آزمایشی، به صورت رابطه زیر محاسبه می‌شود:

$$D_2(X_s, X_t) = \left\| \frac{1}{n_s^c} \sum_{x_i \in X_s^c} P_1^T x_i - \frac{1}{n_t^c} \sum_{x_j \in X_t^c} P_1^T x_j \right\|^2 \quad (۲)$$

در رابطه ۲، P_1 ماتریس نگاشت به فضای هیلبرت بوده و D_2 ، فاصله بین میانگین دامنه‌های آموزشی و آزمایشی در هر کلاس، در فضای هیلبرت، می‌باشد. X_s^c ، داده‌های دامنه آموزشی در کلاس c بوده و X_t^c ، داده‌های دامنه آموزشی در کلاس c می‌باشد. n_s^c و n_t^c به ترتیب، تعداد نمونه‌های دامنه‌های آموزشی و آزمایشی در کلاس c می‌باشند. رابطه ۲، به صورت بسته، به شکل $D_2(X_s, X_t) = \arg \min_p \text{tr}(P_1^T X W_c X^T P_1)$ بازنویسی می‌شود.

در صورتی که در رابطه ۲، $x_i, x_j \in X_s^c$ باشد، $(W_c)_{ij} = \frac{1}{n_s^c n_s^c}$ و

در صورتی که $x_i, x_j \in X_t^c$ باشد، $(W_c)_{ij} = \frac{1}{n_t^c n_t^c}$ و در غیر این

صورت، $(W_c)_{ij} = \frac{-1}{n_s^c n_t^c}$ می‌باشد.

به دلیل عدم وجود داده‌های برچسب‌دار در دامنه آموزشی، محاسبه اختلاف توزیع شرطی به صورت مستقیم امکان‌پذیر نخواهد بود. به منظور حل این مشکل، یک کلاسه‌بند کلاسیک توسط داده‌های دامنه آموزشی، آموزش داده شده و سپس برای پیش‌بینی برچسب داده‌های دامنه آموزشی استفاده می‌شود. به دلیل وجود اختلاف توزیع بین داده‌های دامنه آموزشی و آزمایشی، در ابتدا این کلاسه‌بند دارای دقت پایینی خواهد بود. به منظور افزایش دقت کلاسه‌بند در پیش‌بینی برچسب دامنه آموزشی و افزایش تطبیق توزیع شرطی دامنه‌های آموزشی و آزمایشی، الگوریتم پیشنهادی در این مرحله، به صورت تکرار شونده خواهد بود [۱۸].

۳-۴-۱- حل مسئله بهینه‌سازی

به‌منظور پیاده‌سازی روش EDA، الگوریتم ۱، ارائه شده است. ورودی این الگوریتم، دامنه‌های آموزشی و آزمایشی در فضای اصلی و برجسب‌های دامنه آموزشی و پارامترهای λ و α و β هستند. در مرحله اول، به‌منظور نگاشت داده‌های آموزشی و آزمایشی به زیرفضایی با حداقل اختلاف توزیع حاشیه‌ای و شرطی و حداکثر تفکیک‌پذیری بین کلاس‌های هر دو دامنه، ماتریس نگاشت P_2 ، به‌صورت تکراری جستجو می‌شود. در مرحله دوم، با استفاده از الگوریتم ضرایب لاگرانژ اضافی غیردقیق^{۲۵}، متغیرهای Z ، P_2 ، E در رابطه ۷، محاسبه می‌شود [۲۲]. در این الگوریتم، در هر تکرار، با ثابت در نظر گرفتن تمام متغیرها به‌جز متغیر موردنظر، مقدار جدید متغیر مفروض محاسبه می‌شود [۱۳].

۳-۵- تحلیل زمانی اجرا

در این بخش، پیچیدگی زمانی روش EDA که یک الگوریتم دو مرحله ای می‌باشد، مورد بررسی قرار می‌گیرد. اگر فرض کنیم m ، ابعاد اصلی داده‌ها در دامنه‌های آموزشی و آزمایشی بوده و C ، تعداد کل کلاس‌ها باشد و n ، تعداد کل داده‌های آموزشی و آزمایشی به‌صورت $n = n_s + n_t$ باشد که n_s تعداد داده‌های دامنه آموزشی و n_t ، تعداد داده‌های دامنه آزمایشی است، پیچیدگی زمانی الگوریتم در مرحله اول، $O(mn + m^2 + Cn^2)$ می‌باشد. پیچیدگی زمانی الگوریتم در مرحله دوم (رابطه ۷)، به شرح زیر می‌باشد: در گام ۱۶ از الگوریتم ۱، مقدار پیچیدگی زمانی به‌صورت $O(d^2(n_t + n_s) + d^3 + dn_t c + dn_s c)$ به دست می‌آید. در گام ۱۷، مقدار پیچیدگی زمانی به‌صورت $O(dn_t c + n_s c n_t + n_s d c + n_s^3 + d^2 n_s^2)$ به دست می‌آید. در گام ۱۸، مقدار پیچیدگی زمانی، به‌صورت $O(n_t^3)$ است. بنابراین پیچیدگی زمانی در مرحله دوم، به‌صورت $O(d^2(n_s + n_t) + d^3 + n_s^3 + n_t^3)$ به‌طور کلی، پیچیدگی زمانی الگوریتم EDA، به‌صورت $O(n_s^3 + n_t^3)$ می‌باشد.

۴- نتایج

در این بخش، آزمایش‌های انجام شده به‌منظور ارزیابی روش EDA به‌تفصیل توضیح داده می‌شود.

۴-۱- معرفی پایگاه داده

به‌منظور ارزیابی کارایی روش پیشنهادی، از دو پایگاه داده زیر استفاده شده است: ۱- آفیس [۲۳] و کالتک [۲۴]، ۲- چهره (پای^{۲۶}) [۲۵].

$$\min_{P_2, Z} \text{rank}(Z) \quad \text{s.t.} \quad P_2^T X_t = P_2^T X_s Z \quad (5)$$

مقدار رتبه در یک ماتریس، توسط تعداد مقادیر ویژه^{۳۰} غیر صفر ماتریس مفروض، محاسبه می‌شود. نرم هسته‌ای^{۳۱} یک ماتریس به‌صورت $\|Z\|_* = \|\sigma(Z)\|_1$ محاسبه می‌شود. با توجه به این‌که مسئله در رابطه ۵، یکی از مسائل سخت^{۳۲} بوده و توسط متدهای متداول قابل حل نمی‌باشد، برای محاسبه کم‌ترین رتبه در یک ماتریس، می‌توان از نرم هسته‌ای ماتریس استفاده کرد که عبارت جایگزین برای رابطه ۵ می‌باشد [۱۳، ۲۰]. بدین ترتیب، رابطه ۵ به‌صورت زیر بازنویسی می‌شود:

$$\min_{P_2, Z} \|Z\|_* \quad \text{s.t.} \quad P_2^T X_t = P_2^T X_s Z \quad (6)$$

دومین شرط به‌منظور افزایش تطابق داده‌های آموزشی و آزمایشی با یکدیگر، نمایش تنک می‌باشد. بدین ترتیب در زیرفضای جدید، به‌منظور بازسازی هر داده آزمایشی، آن دسته از داده‌هایی از دامنه آموزشی که بیش‌ترین ارتباط را با داده مفروض دارند، وزن بیش‌تری دریافت می‌کنند [۲۱]. به‌منظور کاهش تأثیر داده‌های پرت^{۳۳} در دامنه آزمایشی، یک ماتریس خطا در نظر گرفته شده که مقادیر این ماتریس، میزان فاصله یک داده پرت در دامنه آزمایشی از نزدیک‌ترین داده در دامنه آموزشی می‌باشد [۱۹]. به‌عبارت‌دیگر، هدف از استفاده ماتریس خطا، کاهش تأثیر انتقال منفی^{۳۴} در یادگیری انتقالی می‌باشد. رابطه نهایی در مرحله دوم با استفاده از روابط ۴، ۵ و ۶، به‌صورت رابطه زیر بازنویسی می‌شود.

$$\min_{P_2, Z, E} \|Z\|_* + \alpha \|Z\|_1 + \beta \|E\|_1 \quad (7)$$

$$\text{s.t.} \quad P_2^T X_t = P_2^T X_s Z + E$$

در رابطه ۷، $\|Z\|_1$ نرم l_1 برای ماتریس ضرایب بازسازی بوده که به‌صورت $\|Z\|_1 = \sum_{i,j} |Z_{i,j}|$ محاسبه می‌شود [۱۳]. با اعمال این شرط، داده‌های آزمایشی توسط آن دسته از داده‌های آموزشی بازسازی می‌شوند که بیش‌ترین ارتباط را با داده آزمایشی موردنظر داشته و اطلاعات ساختار محلی داده‌های آزمایشی، حفظ می‌شود. به‌منظور اعمال شرط تنک بودن بر روی ماتریس E ، از نرم l_1 استفاده شده که در رابطه ۷، به‌منظور کنترل میزان تأثیر این عبارت از دو پارامتر α و β استفاده شده است.

رابطه ۷، به‌عنوان تابع هدف، در جستجوی ماتریس نگاشت P_2 به یک زیرفضای مشترک و ماتریس ضرایب بازسازی Z و همچنین ماتریس خطا E بوده به‌طوری‌که در این زیرفضا، بازسازی داده‌های آزمایشی به نحو مناسبی انجام شود.

برای مثال در یک آزمایش، دامنه پای ۱ به عنوان دامنه آموزشی در نظر گرفته شده و پای ۳ به عنوان دامنه آزمایشی در نظر گرفته شده است. بدین ترتیب، در این پایگاه داده، ۲۰ آزمایش طراحی شده است.

الگوریتم ۱: الگوریتم روش پیشنهادی، EDA

۱	(مرحله اول)
۲	ورودی: داده‌های آموزشی و آزمایشی X ، برچسب‌های دامنه آموزشی γ_s ، پارامترهای تنظیم پذیری $\alpha, \beta, \gamma, \lambda$.
۳	خروجی: داده‌های نگاشت یافته دامنه‌های آموزشی و آزمایشی محاسبه ماتریس ضرایب MMD. به منظور کاهش اختلاف توزیع حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی، با استفاده از فرمول زیر:
۴	محاسبه فاصله بین نمونه‌های دامنه آموزشی از یکدیگر در هر کلاس، با استفاده از فرمول زیر:
۵	ایجاد برداری با اعضای $n_s \times 1$ با ابعاد $n \times 1$ که در آن، n مجموع تعداد نمونه‌های دامنه آموزشی و آزمایشی می‌باشد.
۶	محاسبه ماتریس مرکزیت H توسط فرمول زیر:
۷	تکرار تا نیل به همگرایی:
۸	• حل معادله $(X \sum_{c=0}^C W_c X^T + S_W + \lambda I) P_1 = (X H X^T) P_1$
۹	انتخاب K تا از بردارهای ویژه به عنوان ماتریس نگاشت. آموزش کلاسه‌بند استاندارد f توسط نمونه‌های نگاشت شده دامنه آموزشی $\{P_1^T X_s, \gamma_s\}$
۱۰	• به منظور به‌روزرسانی مقادیر، محاسبه مجدد ماتریس ضرایب MMD با استفاده از فرمول زیر:
۱۱	توقف حلقه
۱۲	خارج کردن نمونه‌های دامنه آموزشی و آزمایشی نگاشت شده، به عنوان خروجی مرحله اول
۱۳	(مرحله دوم)
۱۴	مقداردهی اولیه پارامترها
۱۵	تکرار تا برقراری شرط همگرایی:
۱۶	• ثابت نگهداشتن مقدار تمامی متغیرها و محاسبه مجدد مقدار متغیر P_2 با در نظر گرفتن شرایط فعلی.

پایگاه داده آفیس، مجموعه‌ای از تصاویر اشیاء مختلف از سه دامنه آمازون (A)، وبکم (W) و DSLR (D) بوده که این تصاویر از نظر کیفیت، رنگ، روشنایی، نوع زمینه با یکدیگر متفاوت می‌باشند. تصاویر دامنه آمازون، متشکل از تصاویر اشیای دانلود شده از وبسایت می‌باشند (برای مثال وبسایت amazon.com). در این تصاویر، اشیاء در مرکز تصویر با پس‌زمینه سفید قرار دارند. دامنه DSLR، شامل تصاویر اشیاء با کیفیت وضوح بالا بوده که این تصاویر توسط دوربین‌های دیجیتال حرفه‌ای ثبت شده‌اند. دامنه وبکم، از تصاویر با کیفیت وضوح پایین تشکیل شده که این تصاویر توسط دوربین وبکم ثبت شده‌اند. پایگاه داده کالتک ۲۵۶ (C)، یکی از پایگاه داده‌های استاندارد در مسائل تشخیص اشیاء می‌باشد. این پایگاه داده، شامل ۳۰۶۰۷ تصویر بوده که این تصاویر در ۲۵۶ کلاس قرار دارند. در هر کدام از آزمایش‌ها، یکی از دامنه‌ها به عنوان دامنه آموزشی و دیگری به عنوان دامنه آزمایشی در نظر گرفته شده است. بدین ترتیب، در این پایگاه داده، ۱۲ آزمایش طراحی شده است که شامل ۱۰ کلاس می‌باشد. به عنوان آزمایش اول، دامنه کالتک به عنوان دامنه آموزشی و دامنه آمازون به عنوان دامنه آزمایشی، استفاده شده است $(C \rightarrow A)$. در آزمایش دوم و سوم، دامنه کالتک به عنوان دامنه آموزشی و دامنه‌های وبکم و DSLR، به ترتیب به عنوان دامنه‌های آزمایشی انتخاب شده‌اند. در آزمایش‌های چهارم، پنجم و ششم، دامنه آمازون به عنوان دامنه منبع و سه دامنه دیگر، به ترتیب به عنوان دامنه‌های آزمایشی در نظر گرفته شده‌اند. در آزمایش‌های هفتم، هشتم و نهم، دامنه وبکم به عنوان دامنه آموزشی برای آموزش مدل استفاده شده و دامنه‌های کالتک، آمازون و DSLR، برای ارزیابی کارایی مدل استفاده شده‌اند. در آزمایش‌های دهم، یازدهم و دوازدهم، با استفاده از دامنه DSLR، مدل آموزش داده شده و سپس با استفاده از سه دامنه دیگر، میزان کارایی مدل بررسی می‌شود.

پایگاه داده پای، یکی از پایگاه داده‌های مورد استفاده در مسائل تشخیص چهره می‌باشد. این پایگاه داده، از ۴۱۳۶۸ تصویر چهره تشکیل شده که ساین تصاویر 32×32 پیکسل بوده و تعداد کلاس‌ها در این پایگاه داده ۶۸ می‌باشد. هر کلاس مربوط به چهره یک شخص خاص بوده که در شرایط نورپردازی مختلف ثبت شده است. این پایگاه داده از پنج دامنه تشکیل شده که هر کدام از این دامنه‌ها از نظر زاویه تصویربرداری متفاوت می‌باشند. ۱- پای ۱ (P1)، مربوط به تصویربرداری از حالت چپ است. ۲- پای ۲ (P2)، مربوط به تصویربرداری از حالت بالا است. ۳- پای ۳ (P3)، مربوط به تصویربرداری از حالت پایین است. ۴- پای ۴ (P4)، مربوط به تصویربرداری از حالت روبرو است. ۵- پای ۵ (P5)، مربوط به تصویربرداری از حالت راست است. تمامی این پنج دامنه به دلیل متفاوت بودن زاویه تصویربرداری دارای توزیع متفاوتی بوده که برای هر آزمایش، یکی از این پنج دامنه به عنوان دامنه آموزشی و هر کدام از چهار دامنه دیگر به عنوان دامنه آزمایشی در نظر گرفته می‌شود.

وجود دارند که پارامترهای α و β ، پارامترهای تنظیم‌پذیری نرم l_1 بوده و پارامتر γ ، پارامتر تنظیم‌پذیری مدل می‌باشد. مقدار بهینه پارامتر k ، در بخش ارزیابی پارامترها بررسی شده است. برای مقدار بهینه λ ، در تمام آزمایش‌ها، مقدار 0.05 به‌دست آمده است. مقدار بهینه پارامترهای α ، β و γ ، در بخش ارزیابی پارامترها بررسی شده‌است. به‌منظور پیاده‌سازی روش پیشنهادی، از نرم‌افزار متلب با نسخه ۲۰۱۶a استفاده‌شده و این الگوریتم، بر روی سیستمی با پردازنده Core i3 و حافظه چهار گیگا بایت، مورد پردازش قرار گرفته است.

۴-۴- ارزیابی نتایج

جداول ۱ و ۲ به ترتیب، نشان‌دهنده نتایج روش EDA و الگوریتم‌های مورد مقایسه بر روی پایگاه داده‌های آفیس و کالتک و پای می‌باشند. ستون اول در جداول ۱ و ۲، بیانگر آزمایش‌های طراحی شده به‌منظور ارزیابی عملکرد روش‌های پیشنهادی در حوزه یادگیری انتقالی می‌باشد. بدین ترتیب، دامنه آموزشی در سمت چپ فلش و دامنه آزمایشی در سمت راست فلش قرار گرفته‌اند. برای مثال، در آزمایش $A \rightarrow C$ ، دامنه کالتک به‌عنوان دامنه آموزشی و دامنه آمازون به‌عنوان دامنه آزمایشی در نظر گرفته شده‌اند.

در پایگاه داده آفیس و کالتک، متوسط دقت در روش EDA، $1/44\%$ نسبت به بهترین الگوریتم مورد مقایسه (DICD)، بهبود داشته است. در این پایگاه داده، متوسط دقت در روش EDA نسبت به روش پایه NN، $19/03\%$ بهبود یافته است. در پایگاه داده پای، متوسط دقت در روش EDA، $5/99\%$ نسبت به بهترین الگوریتم مورد مقایسه (DICD)، بهبود داشته و در ۱۹ حالت از ۲۰ حالت، دقت روش EDA بهتر بوده است. در ادامه به‌صورت مجزا کارایی روش پیشنهادی در آزمایش‌های مختلف و همچنین در مقایسه با روش‌های تطبیق دامنه ارزیابی می‌شود.

در آزمایش‌های طراحی شده بر روی پایگاه داده آفیس و کالتک، روش EDA، رفتارهای متفاوتی از خود نشان می‌دهد. این تغییر رفتار به دلیل اختلاف توزیع بین نمونه‌های دامنه‌های مختلف می‌باشد. برای مثال، اختلاف توزیع بین دامنه DSLR نسبت به دامنه‌های آمازون و وبکم متفاوت است. اختلاف توزیع DSLR نسبت به وبکم کم بوده و دقت بالایی به‌دست آمده است ($85/76\%$). همچنین، با در نظر گرفتن وبکم به‌عنوان دامنه آموزشی و DSLR به‌عنوان دامنه آزمایشی نیز، همچنان دقت بالایی به‌دست آمده است ($85/35\%$). این در حالی است که اختلاف توزیع DSLR نسبت به دامنه آمازون، زیاد بوده و دقت بالایی به‌دست نیامده است ($41/23\%$). همچنین، با در نظر گرفتن آمازون به‌عنوان دامنه آموزشی و DSLR به‌عنوان دامنه آزمایشی نیز، به دلیل اختلاف توزیع بالا، دقت بالایی به‌دست نیامده است ($47/13\%$). در پایگاه داده پای هم به دلیل وجود اختلاف توزیع متفاوت بین دامنه‌های

۱۷	• ثابت نگه‌داشتن مقدار تمامی متغیرها و محاسبه مجدد مقدار متغیر Z با در نظر گرفتن شرایط فعلی.
۱۸	• ثابت نگه‌داشتن مقدار تمامی متغیرها و محاسبه مجدد مقدار متغیر Z_1 با در نظر گرفتن شرایط فعلی.
۱۹	• ثابت نگه‌داشتن مقدار تمامی متغیرها و محاسبه مجدد مقدار متغیر Z_2 با در نظر گرفتن شرایط فعلی.
۲۰	• ثابت نگه‌داشتن مقدار تمامی متغیرها و محاسبه مجدد مقدار متغیر E با در نظر گرفتن شرایط فعلی.
۲۱	• به‌روزرسانی مقادیر سایر پارامترها.
۲۲	• چک کردن شرط همگرایی:
۲۳	$\ P_2^T X_t - P_2^T X_s Z - E\ _\infty < \epsilon, \ Z - Z_1\ _\infty < \epsilon, \ Z - Z_2\ _\infty < \epsilon$
۲۴	توقف حلقه
	خارج کردن نمونه‌های وارد شده به مرحله دوم پس از نگاهش با استفاده از ماتریس P_2 به‌عنوان خروجی الگوریتم.

۴-۲- الگوریتم‌های مورد مقایسه

روش EDA، یک روش ترکیبی از دو دسته روش می‌باشد. روش‌های دسته اول، به دنبال کاهش اختلاف توزیع بین داده‌های دامنه آموزشی و آزمایشی بوده و روش‌های دسته دوم، به دنبال کاهش خطای داده‌های آزمایشی توسط داده‌های آموزشی می‌باشند. کارایی روش EDA، با بهترین نتایج گزارش شده از کارایی روش‌های هر دو دسته مقایسه می‌شود. روش‌های مورد مقایسه در دسته اول، عبارت‌اند از: کلاسه‌بند نزدیک‌ترین همسایه (NN) [۳۰]، تحلیل اجزای اصلی (PCA) [۲۶]، TCA [۳]، GFK [۷]، TJM [۶]، JDA [۱۰]، CDDA [۲۷]، CLGA [۲۸]، DICD [۲۹]، روش‌های دسته دوم، عبارت‌اند از: LTSL [۱۹]، RDALR [۲۰]، LRSR [۱۳].

۴-۳- مفروضات پیاده‌سازی

برای مقایسه روش EDA با الگوریتم‌های شناخته‌شده در حوزه یادگیری انتقالی، دقت کلاسه‌بند بر روی داده‌های دامنه آزمایشی محاسبه می‌شود. این دقت به‌صورت رابطه ۸ تعریف می‌شود.

$$دقت = \frac{\sum_{i=1}^{n_t} f(x_i)}{n_t} = y(x_i) \quad (8)$$

در رابطه ۸، $f(x)$ ، برچسب پیش‌بینی شده برای نمونه x_i بوده و $y(x_i)$ ، برچسب واقعی نمونه می‌باشد. یک مدل NN توسط نمونه‌های نگاشت یافته به زیرفضای جدید، آموزش داده شده و سپس تمامی نمونه‌های دامنه آزمایشی، توسط این مدل برچسب‌گذاری می‌شوند. نسبت تعداد نمونه‌هایی از دامنه آزمایشی که برچسب پیش‌بینی شده توسط مدل و برچسب واقعی آن‌ها یکسان باشند، به‌عنوان دقت EDA در نظر گرفته می‌شود. در مرحله اول الگوریتم EDA، دو پارامتر وجود دارد: $k-1$: ابعاد فضای جدید و $2-\lambda$: پارامتر تنظیم‌پذیری نسبت می‌باشد. در مرحله دوم نیز، سه پارامتر α ، β و γ

ترتیب، متوسط دقت این روش در مقایسه با روش TCA، در پایگاه داده آفیس و کالتک و پای، به ترتیب، ۷/۳۷٪ و ۳۴/۳۳٪ بهبود یافته است.

روش GFK، به منظور کاهش اختلاف توزیع حاشیه‌ای بین داده‌های آموزشی و آزمایشی، ابعاد فضای اصلی داده‌ها را کاهش داده که به علت کاهش بیش از حد ابعاد، ساختار اصلی داده‌ها حفظ نمی‌شود. این درحالی است که در روش EDA، یک زیر فضای مشترک بین داده‌های دامنه آموزشی و آزمایشی ایجاد کرده که در این صورت ساختار داده‌ها حفظ می‌شود. میزان بهبود متوسط دقت در روش EDA نسبت به GFK بر روی پایگاه داده آفیس و کالتک و پای، به ترتیب ۷/۴۵٪ و ۴۳/۷۳٪ می‌باشد.

روش TJM، یک روش ترکیبی با بهره‌گیری از دو روش تطبیق ویژگی و انتخاب نمونه بوده که به‌طور هم‌زمان، دو هدف زیر را دنبال می‌کند: ۱- یک زیر فضای مشترک که دارای حداقل توزیع حاشیه‌ای بوده و واریانس به‌صورت حداکثر حفظ شده را ایجاد می‌کند، ۲- به‌منظور وزندهی مجدد داده‌های دامنه آموزشی از روش $l_{2,1}$ استفاده می‌کند. در این روش، به نمونه‌هایی از دامنه آموزشی که دارای بیش‌ترین شباهت از نظر توزیع به نمونه‌های دامنه آموزشی می‌باشد، وزن بیش‌تری اختصاص داده می‌شود. درحالی‌که EDA، اختلاف توزیع حاشیه‌ای و شرطی بین داده‌های دامنه آموزشی و آزمایشی را کاهش داده و ساختار محلی و عمومی داده‌ها را به‌طور هم‌زمان منطبق می‌کند.

مختلف، شرایط مشابه پایگاه داده آفیس وجود دارد. برای مثال، دامنه‌های P2 و P4، از نظر توزیع نمونه‌ها دارای شباهت زیادی می‌باشند، به همین دلیل، در آزمایش‌های $P2 \rightarrow P4$ و $P4 \rightarrow P2$ ، تنها برای ۷٪ از نمونه‌های دامنه آزمایشی، برچسب اشتباه پیش‌بینی شده‌است.

روش NN، یک روش پایه می‌باشد که به‌منظور کلاسه‌بندی داده‌ها در ابتدا فاصله اقلیدسی بین هر نمونه از دامنه آموزشی را نسبت به نمونه‌های دامنه آموزشی محاسبه می‌کند. سپس، با توجه به اینکه درجه همسایگی در این روش، یک در نظر گرفته شده‌است، برچسب نزدیک‌ترین نمونه از دامنه آموزشی، به‌عنوان برچسب هر نمونه از دامنه آموزشی اختصاص داده می‌شود [۳۰].

روش PCA، یک روش کاهش بعد بوده که داده‌های دامنه‌های آموزشی و آزمایشی را به زیرفضایی با ابعاد کم نگاشت می‌کند. این روش اختلاف توزیع بین دامنه‌های آموزشی و آزمایشی را کاهش نمی‌دهد و دقت این الگوریتم در مقایسه با الگوریتم‌های مطرح در حوزه یادگیری انتقالی غیرنظارتی بسیار پایین می‌باشد. متوسط بهبود دقت EDA نسبت به PCA در پایگاه داده آفیس و کالتک و پای، به ترتیب، ۱۰/۷۵٪، ۴۵/۲۳٪ می‌باشد.

روش TCA، یک روش برای کاهش اختلاف توزیع حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی است. در این روش، توجهی به کاهش اختلاف توزیع شرطی بین دامنه‌های آموزشی و آزمایشی نشده‌است. این درحالی است که در مرحله اول روش EDA، اختلاف توزیع شرطی و حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی، کاهش یافته و بدین

جدول ۱: دقت (%) کلاسه‌بند در پایگاه داده آفیس و کالتک (در ستون اول، دامنه آموزشی در سمت چپ فلش و دامنه آزمایشی در سمت راست

فلش قرار داده شده‌است.)

	NN [۳۰]	PCA [۲۶]	TCA (۲۰۱۱) [۲]	GFK (۲۰۱۲) [۷]	TJM (۲۰۱۴) [۶]	JDA (۲۰۱۳) [۱۰]	LTSL (۲۰۱۴) [۱۹]	RDALR (۲۰۰۹) [۲۰]	CDDA (۲۰۱۷) [۲۷]	LRSR (۲۰۱۶) [۱۳]	CLGA (۲۰۱۸) [۲۸]	DICD (۲۰۱۸) [۲۹]	EDA
$C \rightarrow A$	۲۳/۷۰	۳۶/۹۵	۴۵/۸۲	۴۱/۰۲	۴۶/۷۶	۴۴/۷۸	۲۵/۲۶	۳۸/۲۰	۴۸/۳۳	۵۱/۲۵	۴۸/۰۲	۴۷/۲۹	۵۴/۹۱
$C \rightarrow W$	۲۵/۷۶	۳۲/۵۴	۳۰/۵۱	۴۰/۶۸	۳۹/۹۸	۴۱/۶۹	۱۹/۳۲	۳۸/۶۴	۴۴/۷۵	۳۸/۶۴	۴۲/۳۷	۴۶/۴۴	۵۰/۵۱
$C \rightarrow D$	۲۵/۴۸	۳۸/۲۲	۳۵/۶۷	۳۸/۸۵	۴۴/۵۹	۴۵/۲۲	۲۱/۰۲	۴۱/۴۰	۴۸/۴۱	۴۷/۱۳	۴۹/۰۴	۴۹/۶۸	۵۲/۲۳
$A \rightarrow C$	۲۶/۰۰	۳۴/۷۳	۴۰/۰۷	۴۰/۲۵	۳۹/۴۵	۳۹/۳۶	۱۶/۹۲	۳۷/۷۶	۴۲/۱۲	۴۳/۳۷	۴۲/۳	۴۲/۳۹	۴۳/۲۸
$A \rightarrow W$	۲۹/۸۳	۳۵/۵۹	۳۵/۲۵	۳۸/۹۸	۴۲/۰۲	۳۷/۹۷	۱۴/۵۸	۳۷/۶۳	۴۱/۶۹	۳۶/۶۱	۴۱/۳۶	۴۵/۰۸	۴۶/۴۴
$A \rightarrow D$	۲۵/۴۸	۲۷/۳۹	۳۴/۳۹	۳۶/۳۱	۴۵/۲۲	۳۹/۴۹	۲۱/۰۲	۳۳/۱۲	۳۷/۵۸	۳۸/۸۵	۳۶/۳۱	۳۸/۸۵	۴۷/۱۳
$W \rightarrow C$	۱۹/۸۶	۲۶/۳۶	۲۹/۹۲	۳۰/۷۲	۳۰/۱۹	۳۱/۱۷	۳۴/۶۴	۲۹/۳۰	۳۱/۹۷	۲۹/۸۳	۳۲/۹۵	۳۳/۵۷	۳۱/۵۲
$W \rightarrow A$	۲۲/۹۶	۲۹/۳۵	۲۸/۸۱	۲۹/۷۵	۲۹/۹۶	۳۲/۷۸	۳۹/۵۶	۳۰/۰۶	۳۷/۲۷	۳۴/۱۳	۳۴/۵۷	۳۴/۱۳	۳۰/۶۹
$W \rightarrow D$	۵۹/۲۴	۷۷/۰۷	۸۵/۹۹	۸۰/۸۹	۸۹/۱۷	۸۹/۱۷	۷۲/۶۱	۸۷/۲۶	۸۷/۹۰	۸۲/۸۰	۹۲/۳۶	۸۹/۸۱	۸۵/۳۵
$D \rightarrow C$	۲۶/۲۷	۲۹/۶۵	۳۲/۰۶	۳۰/۲۸	۳۱/۴۳	۳۱/۵۲	۳۵/۰۸	۳۱/۷۰	۳۴/۶۴	۳۱/۶۱	۳۳/۶۶	۳۴/۶۴	۳۵/۸۰
$D \rightarrow A$	۲۸/۵۰	۳۲/۰۵	۳۱/۴۲	۳۲/۰۵	۳۲/۷۸	۳۳/۰۹	۳۹/۶۷	۳۲/۱۵	۳۳/۵۱	۳۳/۱۹	۸۹/۸۳	۳۴/۴۵	۴۱/۲۳
$D \rightarrow W$	۶۳/۳۹	۷۵/۹۳	۸۶/۴۴	۷۵/۵۹	۸۵/۴۲	۸۹/۴۹	۷۴/۹۲	۸۶/۱۰	۹۰/۵۱	۷۷/۲۹	۳۵/۹۹	۹۱/۱۹	۸۵/۷۶
میانگین	۳۱/۳۷	۳۹/۶۵	۴۳/۰۳	۴۲/۹۵	۴۶/۴۲	۴۶/۳۱	۳۴/۵۵	۴۳/۶۱	۴۸/۲۲	۴۵/۳۹	۴۸/۲۳	۴۸/۹۶	۵۰/۴۰

جدول ۲: دقت (%) کلاسه‌بند در پایگاه داده پای (در ستون اول، دامنه آموزشی در سمت چپ فلش و دامنه آزمایشی در سمت راست فلش قرار

داده شده‌است.)

	NN [۳۰]	PCA [۲۶]	TCA (۲۰۱۴) [۳]	GFK (۲۰۱۲) [۷]	TJM (۲۰۱۴) [۶]	JDA (۲۰۱۳) [۱۰]	LTSL (۲۰۱۴) [۱۹]	RDALR (۲۰۰۹) [۲۰]	CDDA (۲۰۱۷) [۲۷]	LRSR (۲۰۱۶) [۱۳]	CLGA (۲۰۱۸) [۲۸]	DICD (۲۰۱۸) [۲۹]	EDA
P1 → P2	۲۶/۰۹	۲۴/۸۰	۴۰/۷۶	۲۶/۱۵	۲۳/۸۷	۵۸/۸۱	۲۲/۹۶	۴۰/۷۶	۶۰/۲۲	۶۵/۸۷	۶۷/۸۳	۷۲/۹۹	۷۳/۱۷
P1 → P3	۲۶/۵۹	۲۵/۱۸	۴۱/۷۹	۲۷/۲۷	۲۸/۸۶	۵۴/۲۳	۲۰/۶۵	۴۱/۷۹	۵۸/۷	۶۴/۰۹	۶۳/۸۵	۷۲/۰۰	۷۷/۳۹
P1 → P4	۳۰/۶۷	۲۹/۲۶	۵۹/۶۳	۳۱/۱۵	۴۳/۳۷	۸۴/۵۰	۳۱/۸۱	۵۹/۶۳	۸۳/۴۸	۸۲/۰۳	۸۸/۹۵	۹۲/۲۲	۹۵/۵۵
P1 → P5	۱۶/۶۷	۱۶/۳۰	۲۹/۳۵	۱۷/۵۹	۱۹/۳۰	۴۹/۷۵	۱۲/۰۷	۲۹/۳۵	۵۴/۱۷	۵۴/۹	۶۱/۷۶	۶۶/۸۵	۷۲/۳۷
P2 → P1	۲۴/۴۹	۲۴/۲۲	۴۱/۸۱	۲۵/۲۴	۲۶/۱۴	۵۷/۶۲	۱۸/۲۵	۴۱/۸۱	۶۲/۳۳	۴۵/۵۴	۷۱/۴	۶۹/۹۳	۷۹/۵۹
P2 → P3	۴۶/۶۳	۴۵/۵۳	۵۱/۴۷	۴۷/۳۷	۳۷/۹۳	۶۲/۹۳	۱۶/۰۵	۵۱/۴۷	۶۴/۶۴	۵۳/۴۹	۷۲/۹۸	۶۵/۸۷	۷۴/۰۲
P2 → P4	۵۴/۰۷	۵۳/۳۵	۶۴/۷۳	۵۴/۲۵	۵۰/۵۳	۷۵/۸۲	۴۵/۱۵	۶۴/۷۳	۷۹/۹	۷۱/۴۳	۸۶/۲۴	۸۵/۲۵	۹۳/۹۹
P2 → P5	۲۶/۵۳	۲۵/۴۳	۳۳/۷۰	۲۷/۰۸	۲۱/۶۳	۳۹/۸۹	۱۷/۵۲	۳۳/۷۰	۴۴	۴۷/۹۷	۵۱/۲۳	۴۸/۷۱	۶۰/۱۷
P3 → P1	۲۱/۳۷	۲۰/۹۵	۳۴/۶۹	۲۱/۸۲	۲۸/۶۶	۵۰/۹۶	۲۲/۳۶	۳۴/۶۹	۵۸/۴۶	۵۲/۴۹	۷۰/۱۷	۶۹/۳۶	۶۸/۶۴
P3 → P2	۴۱/۰۱	۴۰/۴۵	۴۷/۷۰	۴۳/۱۶	۳۵/۹۷	۵۷/۹۵	۲۰/۲۶	۴۷/۷۰	۵۹/۷۳	۵۵/۵۶	۷۳/۴۸	۶۵/۴۴	۷۲/۰۷
P3 → P4	۴۶/۵۳	۴۶/۱۴	۵۶/۲۳	۴۶/۴۱	۵۱/۹۷	۶۸/۴۵	۵۷/۳۴	۵۶/۲۳	۷۷/۲	۷۷/۵	۸۹/۳۱	۸۳/۳۹	۹۲/۶۴
P3 → P5	۲۶/۲۳	۲۵/۳۱	۳۳/۱۵	۲۶/۷۸	۲۵/۳۱	۳۹/۹۵	۲۴/۵۷	۳۳/۱۵	۴۷/۲۴	۵۴/۱۱	۵۵/۵۱	۶۱/۴۰	۷۰/۲۲
P4 → P1	۳۲/۹۵	۳۱/۹۶	۵۵/۶۴	۳۴/۲۴	۴۵/۷۱	۸۰/۵۸	۵۱/۲۰	۵۵/۶۴	۸۳/۱	۸۱/۵۴	۸۹/۵۶	۹۳/۱۳	۹۴/۸۱
P4 → P2	۶۲/۶۸	۶۰/۹۶	۶۷/۸۳	۶۲/۹۲	۵۷/۵۸	۸۲/۶۳	۷۰/۱۰	۶۷/۸۳	۸۲/۲۶	۵۸/۳۹	۹۲/۹۴	۹۰/۱۲	۹۳/۶۲
P4 → P3	۷۳/۲۲	۷۲/۱۸	۷۵/۸۶	۷۳/۳۵	۷۱/۶۳	۸۷/۲۵	۷۲/۰۰	۷۵/۸۶	۸۶/۶۴	۸۲/۲۳	۹۳/۰۸	۸۸/۹۷	۹۳/۳۸
P4 → P5	۳۷/۱۹	۳۵/۱۱	۴۰/۲۶	۳۷/۳۸	۳۰/۹۴	۵۴/۶۶	۴۸/۲۸	۴۰/۲۶	۵۸/۳۳	۷۲/۶۱	۷۱/۶۳	۷۵/۶۱	۸۲/۹۷
P5 → P1	۱۸/۴۹	۱۸/۸۵	۲۶/۹۸	۲۰/۳۵	۲۷/۱۳	۴۶/۴۶	۱۳/۰۶	۲۶/۹۸	۴۸/۰۲	۵۲/۱۹	۵۷/۶۸	۶۲/۸۸	۷۰/۳۵
P5 → P2	۲۴/۱۹	۲۳/۳۹	۲۹/۹۰	۲۴/۶۲	۲۲/۶۵	۴۲/۰۵	۲۱/۶۱	۲۹/۹۰	۴۵/۶۱	۴۹/۴۱	۵۵/۴۳	۵۷/۰۳	۵۷/۴۶
P5 → P3	۲۸/۳۱	۲۷/۲۱	۲۹/۹۰	۲۸/۴۹	۲۸/۸۶	۵۳/۳۱	۱۷/۰۳	۲۹/۹۰	۵۲/۰۲	۵۸/۴۵	۵۸/۰۳	۶۵/۸۷	۷۴/۴۵
P5 → P4	۳۱/۲۴	۳۰/۳۴	۳۳/۶۴	۳۱/۳۳	۳۲/۵۹	۵۷/۰۱	۲۹/۵۹	۳۳/۶۴	۵۵/۹۹	۶۴/۳۱	۷۱/۸۵	۷۴/۷۷	۸۴/۸۳
میانگین	۳۴/۷۶	۳۳/۸۵	۴۴/۷۵	۳۵/۳۵	۳۵/۵۳	۶۰/۲۴	۳۱/۵۹	۴۴/۷۵	۶۳/۱	۶۳/۵۳	۷۲/۱۵	۷۳/۰۹	۷۹/۰۸

از این‌رو، میزان متوسط دقت در روش پیشنهادی نسبت به روش TJM، بر روی پایگاه داده آفیس و کالتک، ۳/۹۸٪ و بر روی پایگاه داده پای ۴۳/۵۵٪ بهبود یافته است.

روش JDA، یک زیرفضای مشترک ایجاد کرده که در این زیرفضا، به‌طور هم‌زمان اختلاف توزیع حاشیه‌ای و شرطی کاهش می‌یابد. این در حالی است که EDA، علاوه بر کاهش هم‌زمان توزیع حاشیه‌ای و شرطی، از خوشه‌بند مستقل از دامنه نیز استفاده کرده و توزیع هندسی داده‌های آموزشی و آزمایشی را نیز تطبیق می‌دهد. متوسط دقت روش EDA نسبت به JDA، بر روی پایگاه داده‌های آفیس و کالتک و پای، به ترتیب ۴/۰۹٪ و ۱۸/۸۴٪ بهبود یافته است.

الگوریتم LTSL، هیچ تضمینی برای حفظ ساختار محلی داده‌ها و بازسازی داده‌های آزمایشی به‌صورت خطی توسط داده‌های آموزشی ندارد این در حالی است که EDA، توسط اعمال شرط تنک بودن بر روی ماتریس ضرایب بازسازی، می‌تواند ساختار محلی داده‌ها را حفظ کند. میزان بهبود متوسط دقت در روش EDA نسبت به LTSL بر روی پایگاه داده‌های آفیس و کالتک و پای، در مقایسه با روش LTSL، ۱۵/۸۵٪ و ۴۷/۴۹٪ می‌باشد.

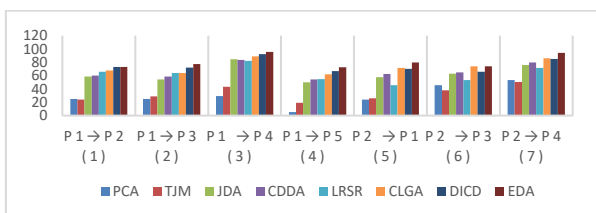
الگوریتم RDALR، به‌منظور کاهش اختلاف بین داده‌های آموزشی و آزمایشی، داده‌های آموزشی را به دامنه آزمایشی انتقال

داده که این کار ممکن است موجب هم‌پوشانی کلاس‌های مختلف در دامنه آزمایشی شود. این در حالی است که در EDA، نگاشت داده‌ها به یک زیرفضای مشترک باعث کاهش اختلاف توزیع بین داده‌های آموزشی و آزمایشی می‌شود. متوسط دقت در الگوریتم EDA، در مقایسه با روش RDALR، بر روی پایگاه داده‌های آفیس و کالتک و پای، ۶/۷۹٪ و ۳۴/۳۳٪ بهبود یافته است.

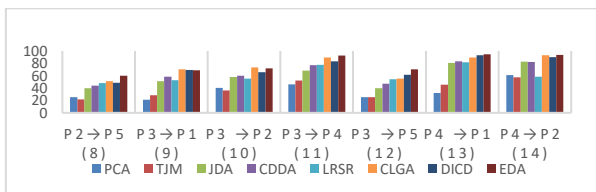
روش CDDA، به‌گونه‌ای نمونه‌های دامنه‌های آموزشی و آزمایشی را به‌صورت آماری تطبیق می‌دهد که کلاس‌های مختلف در هر دامنه بیش‌ترین فاصله را از یکدیگر داشته باشند. روش EDA، علاوه بر تطبیق آماری دامنه‌های آموزشی و آزمایشی، ساختار هندسی نمونه‌ها در هر دو دامنه را نیز تطبیق می‌دهد. روش EDA، در مقایسه با روش CDDA، به‌طور متوسط بر روی پایگاه داده‌های آفیس و کالتک و پای، ۲/۱۸٪ و ۱۵/۹۸٪ بهبود یافته است.

روش LRSR، با اعمال محدودیت‌های رتبه-پایین و تنک، به بازسازی نمونه‌های دامنه آزمایشی با استفاده از نمونه‌های دامنه آموزشی می‌پردازد. این در حالی است که روش EDA، علاوه بر کاهش خطای بازسازی نمونه‌ها در زیرفضای جدید، توزیع آماری دامنه‌های آموزشی و آزمایشی را نیز منطبق می‌کند. بدین ترتیب، متوسط دقت

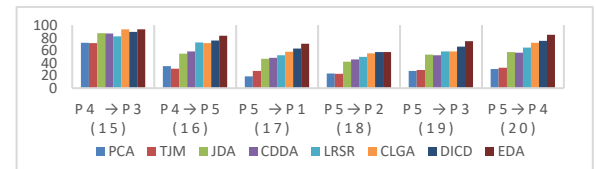
در مقایسه با روش‌های PCA و TJM، کارایی روش EDA، به ترتیب در ۱۲ و ۱۱ حالت بهبود داشته‌است. کارایی روش EDA در مقایسه با روش‌های JDA و CDDA، در ۹ حالت از ۱۲ حالت بهبود داشته‌است؛ در مقایسه با روش‌های LRSR، DICD و CLGA، کارایی روش EDA، به ترتیب در ۸، ۱۰ و ۷ حالت از ۱۲ حالت، بهتر شده‌است. در شکل ۴، دامنه‌های آموزشی و آزمایشی از پایگاه داده پای مشخص شده و کارایی روش EDA با سایر روش‌های مشخص شده، به‌طور مجزا مورد مقایسه قرار گرفته‌است. روش EDA، در مقایسه با روش DICD و CLGA، به ترتیب در ۱۹ و ۱۸ حالت از ۲۰ حالت، بهتر بوده و در مقایسه با سایر روش‌ها، در تمامی ۲۰ حالت، کارایی روش EDA، بهتر بوده‌است.



الف



ب



ج

شکل ۴: مقایسه میزان بهبود کارایی روش پیشنهادی توسط الگوریتم‌های مورد مقایسه در پایگاه داده پای (نمایش بهتر به صورت رنگی)

۴-۵- تأثیر پارامترها

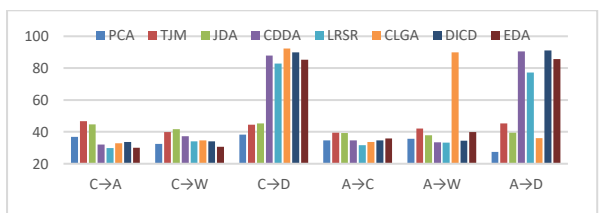
در مرحله اول الگوریتم EDA، در ۱۰ آزمایش تأثیر پارامتر λ مورد بررسی قرار گرفته و مقدار بهینه 0.05 به دست آمده‌است. در شکل ۵، تغییرات متوسط دقت روش EDA بر روی پایگاه داده آفیس و کالتک، به ازای مقادیر مختلف پارامتر λ نشان داده شده‌است. در مرحله دوم، کارایی روش EDA به ازای ۱۰ مقدار مختلف برای پارامترهای α ، β و γ بر روی پایگاه داده آفیس و کالتک، مورد ارزیابی قرار گرفته‌است. در شکل ۶، متوسط دقت روش EDA بر روی پایگاه داده آفیس و کالتک، به ازای ابعاد مختلف زیرفضای مشترک بررسی شده که با توجه به این نمودار، در صورتی که نمونه‌های آموزشی و آزمایشی به زیرفضایی با ابعاد ۲۰ نگاشت شوند، روش

روش EDA، بر روی پایگاه داده‌های آفیس و کالتک و پای، $0.5/0.1$ و $0.55/0.15$ بهبود یافته‌است.

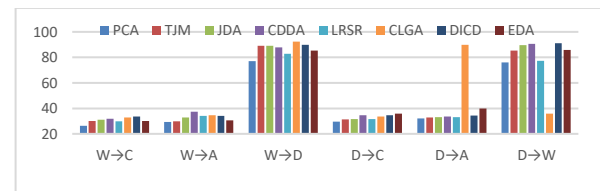
روش CLGA همانند روش EDA، دامنه‌های آموزشی و آزمایشی را در دو مرحله آماری و هندسی منطبق می‌کند. روش CLGA، در مرحله تطبیق توزیع آماری، توزیع‌های شرطی و حاشیه‌های دامنه‌های آموزشی و آزمایشی را منطبق می‌کند، در حالی که روش EDA در مرحله تطبیق توزیع آماری، علاوه بر تطبیق توزیع‌های شرطی و حاشیه‌ای، با استفاده از خوشه‌بندی مستقل از دامنه، کلاس‌های مختلف را از یکدیگر متمایز می‌کند. در مرحله دوم از روش CLGA، با استفاده از فرض منیفولد و گراف همسایگی، اطلاعات هندسی نمونه‌ها حفظ می‌شود. این در حالی است که روش EDA، به منظور حفظ اطلاعات هندسی نمونه‌ها، نمونه‌های دامنه آموزشی را با استفاده از نمونه‌های دامنه آزمایشی و محدودیت‌های رتبه-پایین و تنگ بازسازی می‌کند. بدین ترتیب، متوسط دقت روش EDA در مقایسه با روش CLGA، بر روی پایگاه داده‌های آفیس و کالتک و پای، $0.17/0.2$ و $0.93/0.6$ افزایش یافته‌است.

روش DICD با ایجاد یک زیرفضای مشترک، علاوه بر کاهش اختلاف توزیع شرطی و حاشیه‌ای، فاصله هر دو نمونه از یک کلاس یکسان را کاهش داده و در عین حال، فاصله هر دو نمونه از دو کلاس مختلف را افزایش می‌دهد. روش EDA، علاوه بر کاهش اختلاف توزیع شرطی و حاشیه‌ای و خوشه‌بندی مستقل از دامنه، از روش‌های متفاوتی برای کاهش خطای بازسازی نمونه‌ها استفاده کرده و بدین ترتیب، به‌طور متوسط بر روی پایگاه داده آفیس و کالتک و پای $0.44/0.1$ و $0.99/0.5$ بهبود یافته‌است.

به منظور مقایسه بهتر، در شکل ۳، دامنه‌های آموزشی و آزمایشی از پایگاه داده آفیس و کالتک مشخص شده و روش EDA، به‌طور مجزا با روش‌های PCA، TJM، JDA، CDDA، LRSR، CLGA و DICD مقایسه شده‌است.



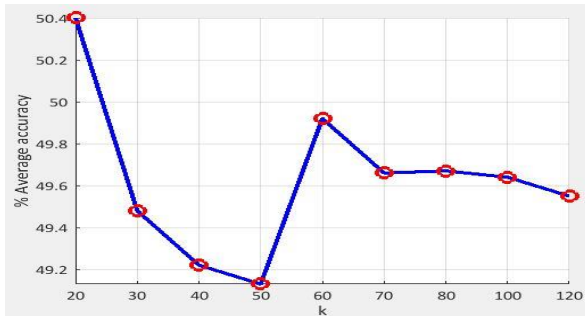
الف



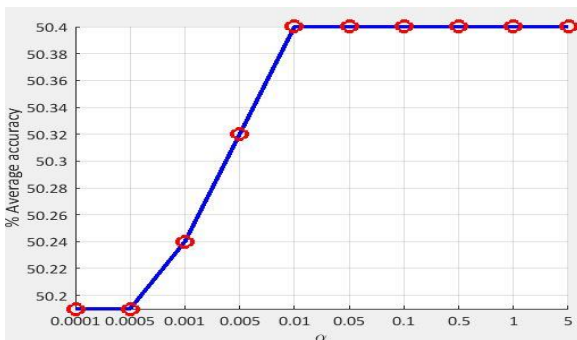
ب

شکل ۳: مقایسه میزان بهبود کارایی روش پیشنهادی توسط الگوریتم‌های مورد مقایسه در پایگاه داده آفیس و کالتک (نمایش بهتر به صورت رنگی)

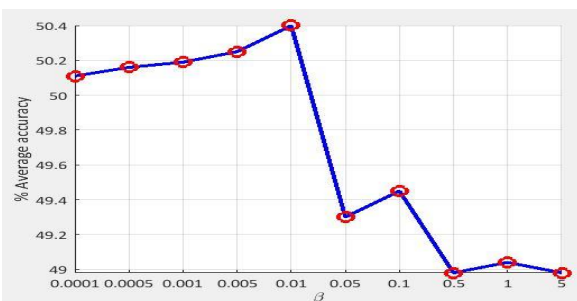
در صورتی که هر سه پارامتر به صورت هم‌زمان حذف شوند، کارایی روش پیشنهادی، در ۱۰ حالت از ۱۲ حالت در پایگاه داده آفیس و کالتک، پایین‌تر می‌باشد.



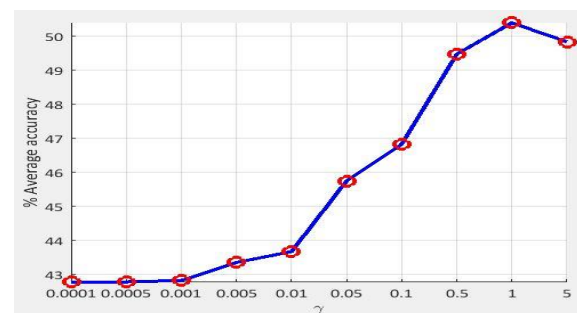
شکل ۶: ارزیابی دقت روش EDA برای مقادیر مختلف پارامتر k (نمایش بهتر به صورت رنگی)



شکل ۷: ارزیابی دقت روش EDA برای مقادیر مختلف پارامتر α (نمایش بهتر به صورت رنگی)



شکل ۸: ارزیابی دقت روش EDA برای مقادیر مختلف پارامتر β (نمایش بهتر به صورت رنگی)

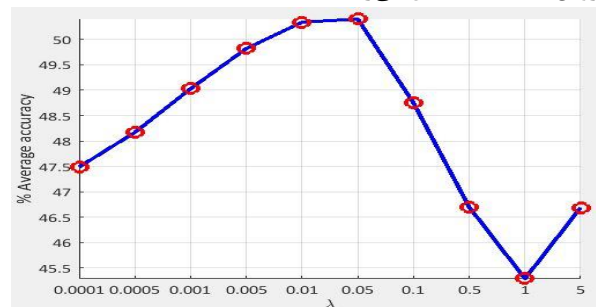


شکل ۹: ارزیابی دقت روش EDA برای مقادیر مختلف پارامتر γ (نمایش بهتر به صورت رنگی)

EDA به طور متوسط نیمی از نمونه‌های دامنه آزمایشی را به درستی برچسب‌گذاری می‌کند.

در شکل ۷، متوسط دقت روش EDA در صورت تغییر مقادیر پارامتر α نشان داده شده که با توجه به این نمودار، مقدار بهینه α برای پارامتر α به دست آمده است. پارامتر α ، میزان استفاده از اطلاعات محلی نمونه‌های موجود در همسایگی هر نمونه، به منظور بازسازی آن نمونه را نشان می‌دهد. مقدار α بیانگر این است که در صورتی که از اطلاعات محلی نمونه‌های کمتری در همسایگی هر نمونه برای بازسازی نمونه مورد نظر استفاده شود، کارایی روش پیشنهادی افزایش خواهد یافت.

در شکل ۸، مقدار متوسط دقت روش EDA، در صورت تغییر مقادیر پارامتر β نشان داده شده که با توجه به این نمودار، روش EDA نسبت به پارامتر β ، حساس بوده و مقدار β موجب بهبود عملکرد روش پیشنهادی می‌شود. در شکل ۹، تأثیر پارامتر γ بر روی کارایی روش EDA نشان داده شده است. با توجه به این نمودار، به ازای مقدار ۱ برای پارامتر γ ، کارایی روش EDA به حداکثر خواهد رسید. بهینه‌ترین مقدار برای تکرار در مرحله اول، ۱۰ به دست آمده و تعداد بهینه تکرار در مرحله دوم، ۱۰۰ به دست آمده است. به منظور آگاهی از میزان اهمیت پارامترها در کارایی الگوریتم پیشنهادی، نتیجه حذف هر کدام از پارامترها بر روی پایگاه داده آفیس و کالتک، در جدول ۳ گنجانده شده است. در روش EDA، به منظور حفظ اطلاعات محلی داده‌ها در ماتریس ضرایب بازسازی، از پارامتر تنظیم پذیری α استفاده شده که با حذف این پارامتر، محدودیت تنک، از ماتریس ضرایب بازسازی، حذف می‌شود. با توجه به جدول ۳، در حالت $D \rightarrow A$ ، در صورتی که پارامتر α حذف شود، میزان کارایی EDA، ۱/۴۶٪ کاهش یافته است. پارامتر β ، پارامتر تنظیم‌پذیری محدودیت تنک بر روی ماتریس خطا می‌باشد. ماتریس خطا، باعث کاهش تأثیر انتقال منفی می‌شود. با حذف پارامتر β در حالت $D \rightarrow A$ ، میزان دقت EDA، ۰/۸۳٪ کاهش یافته است. پارامتر γ ، پارامتر تنظیم‌پذیری مدل بوده که با حذف این پارامتر، متوسط دقت روش پیشنهادی، ۷/۶۲٪ کاهش یافته که در چنین شرایطی، کارایی روش EDA به حداقل می‌رسد.



شکل ۵: ارزیابی دقت روش EDA برای مقادیر مختلف پارامتر λ (نمایش بهتر به صورت رنگی)

مجزا نگاشت داده شوند تا بدین ترتیب بتوانیم صحت پیش‌بینی برچسب دامنه آزمایشی را افزایش دهیم.

جدول ۳: تأثیر پارامترهای الگوریتم پیشنهادی بر کارایی الگوریتم در پایگاه داده آفیس و کالتک

	دقت EDA در صورت حذف پارامتر α	دقت EDA در صورت حذف پارامتر β	دقت EDA در صورت حذف پارامتر γ	دقت EDA در صورت حذف پارامتر β, α و γ	دقت EDA
$C \rightarrow A$	۵۴/۹۱	۵۴/۱۸	۵۵/۴۲	۵۲/۳۴	۵۴/۹۱
$C \rightarrow W$	۵۱/۱۹	۵۰/۸۵	۴۸/۸۱	۵۲/۵۴	۵۰/۵۱
$C \rightarrow D$	۵۱/۵۹	۴۷/۱۳	۴۹/۰۴	۴۲/۶۸	۵۲/۲۳
$A \rightarrow C$	۴۳/۳۷	۴۱/۲۳	۴۳/۴۶	۴۱/۱۴	۴۳/۲۸
$A \rightarrow W$	۴۶/۴۴	۴۷/۱۲	۴۱/۰۲	۴۳/۰۵	۴۶/۴۴
$A \rightarrow D$	۴۶/۵۰	۴۲/۶۸	۴۰/۷۶	۳۶/۳۱	۴۷/۱۳
$W \rightarrow C$	۳۱/۱۷	۳۰/۱۹	۲۹/۸۳	۲۹/۵۶	۳۱/۵۲
$W \rightarrow A$	۳۰/۶۹	۲۹/۷۵	۳۰/۷۹	۳۱/۵۲	۳۰/۶۹
$W \rightarrow D$	۸۵/۳۴	۸۲/۸۰	۶۲/۴۲	۶۸/۷۹	۸۵/۳۵
$D \rightarrow C$	۳۵/۲۶	۳۵/۶۲	۳۲/۲۴	۳۲/۰۶	۳۵/۸۰
$D \rightarrow A$	۳۹/۷۷	۴۰/۴۰	۲۹/۳۳	۳۴/۱۳	۴۱/۲۳
$D \rightarrow W$	۸۶/۱۰	۸۴/۰۷	۴۹/۴۹	۶۴/۰۷	۸۵/۷۶
میانگین	۵۰/۱۹	۴۸/۸۳	۴۲/۷۸	۴۴/۱۰	۵۰/۴۰

مراجع

- [1] Gong, B., Grauman, K. and Sha, F., Reshaping visual datasets for domain adaptation. In *Advances in Neural Information Processing Systems*, pp. 1286-1294, 2013.
- [2] PAN, S. J., AND YANG, Q., A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), pp. 1345-1359, 2010.
- [3] Pan, S.J., Tsang, I.W., Kwok, J.T. and Yang, Q., Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 22(2), pp.199-210, 2011.
- [4] Si, S., Tao, D. and Geng, B., Bregman divergence-based regularization for transfer subspace learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(7), pp.929-942, 2010.
- [5] Gretton, A., Borgwardt, K., Rasch, M.J., Scholkopf, B. and Smola, A.J., A kernel method for the two-sample problem. *arXiv preprint arXiv:0805.2368*, 2008.
- [6] Long, M., Wang, J., Ding, G., Sun, J. and Yu, P.S., Transfer joint matching for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1410-1417, 2014.
- [7] Gong, B., Shi, Y., Sha, F. and Grauman, K., Geodesic flow kernel for unsupervised domain adaptation. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 2066-2073, IEEE, 2012, June.
- [8] Satpal, S. and Sarawagi, S., Domain adaptation of conditional probability models via feature subseting. In *European Conference on Principles of Data Mining and Knowledge Discovery*, Springer, Berlin, Heidelberg, pp. 224-235, 2007, September.
- [9] Tahmoresnezhad, J. and Hashemi, S., Visual domain adaptation via transfer feature learning. *Knowledge and Information Systems*, 50(2), pp.585-605, 2017.
- [10] Long, M., Wang, J., Ding, G., Sun, J. and Philip, S.Y., Transfer feature learning with joint distribution adaptation. In *Computer Vision (ICCV), 2013 IEEE International Conference on* pp. 2200-2207, 2013, December.
- [11] Quanz, B., Huan, J. and Mishra, M., Knowledge transfer with low-quality data: A feature extraction issue. *IEEE Transactions on Knowledge and Data Engineering*, 24(10), pp.1789-1802, 2012.
- [12] Zhang, L., Zuo, W. and Zhang, D., LSDT: Latent sparse domain transfer learning for visual adaptation. *IEEE Transactions on Image Processing*, 25(3), pp.1177-1191, 2016.
- [13] Xu, Y., Fang, X., Wu, J., Li, X. and Zhang, D., Discriminative transfer subspace learning via low-rank and sparse representation. *IEEE Transactions on Image Processing*, 25(2), pp.850-863, 2016.
- [14] Gong, B., Grauman, K. and Sha, F., February. Connecting the dots with landmarks: Discriminatively learning domain-invariant features for unsupervised domain adaptation. In *International Conference on Machine Learning*, pp. 222-230, 2013.
- [15] Ma, Z., Yang, Y., Sebe, N. and Hauptmann, A.G., Knowledge adaptation with partially shared features for event detection using few exemplars. *IEEE transactions*

۵- نتیجه گیری

در این مقاله، یک روش یادگیری انتقالی بدون نظارت، با عنوان حداقل سازی خطای بازسازی داده‌ها از طریق تطبیق توزیع و محدودیت رتبه-پایین (EDA) پیشنهاد شده است. EDA، یک روش دو مرحله‌ای می‌باشد که در مرحله اول، با استفاده از روش تطبیق توزیع، داده‌های آموزشی و آزمایشی به یک زیرفضای جدید، نگاشت می‌شوند. در زیرفضای جدید، اختلاف توزیع شرطی و حاشیه‌ای بین دامنه‌های آموزشی و آزمایشی به صورت هم‌زمان، به حداقل رسانده می‌شود. همچنین، برای افزایش صحت کلاسه‌بند در نمایش جدید، خوشه‌بندی مستقل از دامنه، بر روی دامنه آموزشی اعمال می‌شود. در مرحله دوم، به منظور حفظ ساختار هندسی داده‌ها، از ماتریس ضرایب بازسازی با اعمال محدودیت رتبه-پایین و تنک استفاده می‌شود. کارایی روش EDA، توسط ۳۲ آزمایش بر روی پایگاه داده‌های بصری که شامل دامنه‌هایی با اختلاف توزیع قابل توجهی نسبت به یکدیگر هستند، سنجیده شده است. نتایج به دست آمده، نشان‌دهنده بهبود قابل ملاحظه‌ای در کارایی روش EDA نسبت به روش‌های مورد مقایسه در حوزه یادگیری ماشین و یادگیری انتقالی بر روی دامنه‌های مختلف می‌باشد.

برای ادامه راه، جهت توسعه روش EDA، در نظر داریم که در مرحله اول این روش، دامنه‌های آموزشی و آزمایشی، به دو زیرفضای

- [23] Saenko, K., Kulis, B., Fritz, M. and Darrell, T., Adapting visual category models to new domains. Springer, Berlin, Heidelberg. In *European conference on computer vision*, pp. 213-226, 2010, September.
- [24] Griffin, G., Holub, A. and Perona, P., Caltech-256 object category dataset, 2007.
- [25] Sim, T., Baker, S. and Bsat, M., The CMU pose, illumination, and expression (PIE) database. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on* pp. 53-58, 2002, May.
- [26] Jolliffe I, *Principal component analysis*, Wiley, vol. 2, pp. 433-459, 2002.
- [27] Luo, L., Wang, X., Hu, S., Wang, C., Tang, Y. and Chen, L., 2017. Close yet distinctive domain adaptation. *arXiv preprint arXiv:1704.04235*.
- [28] Liu, J., Li, J. and Lu, K., Coupled local-global adaptation for multi-source transfer learning. *Neurocomputing*, 275, pp.247-254, 2018.
- [29] Li, S., Song, S., Huang, G., Ding, Z. and Wu, C., Domain Invariant and Class Discriminative Feature Learning for Visual Domain Adaptation. *IEEE Transactions on Image Processing*, 27(9), pp. 4260-4273, 2018.
- [۳۰] مهرداد حیدری ارچلو، سید قدرت الله سیف السادات و مرتضی رزاز، «یک روش هوشمند تشخیص جزیره در شبکه توزیع دارای تولیدات پراکنده مبتنی بر تبدیل موجک و نزدیک‌ترین k -همسایگی (kNN)»، مجله مهندسی برق دانشگاه تبریز، جلد ۴۳، شماره ۱، صفحات ۱۵-۲۶، ۱۳۹۲.
- [۳۱] طاهره زارع بیدکی و محمد تقی صادقی، «بهینه‌سازی وزن‌ها در کرنل مرکب برای طبقه‌بند مبتنی بر نمایش تنک کرنلی»، مجله مهندسی برق دانشگاه تبریز، جلد ۴۷، شماره ۳، صفحات ۱۰۵۹-۱۰۷۲، ۱۳۹۶.
- [22] Lin, Z., Chen, M. and Ma, Y., The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv preprint arXiv:1009.5055*, 2010.

زیر نویس‌ها

¹⁹Norm

²⁰Visual Domain Adaptation

²¹Low-Rank and Sparse Representation

²²Global

²³Local

²⁴Task

²⁵Semi-Supervised

²⁶Unsupervised

²⁷Hilbert

²⁸Cluster

²⁹Block-wise

³⁰Singular value

³¹Nuclear norm

³²NP-Hard

³³Outlier

³⁴Negative transfer

³⁵Inexact Augmented Lagrange Multiplier (IALM)

³⁶Pie

¹Domain shift

²Cross Domain problems

³Transfer Learning

⁴Domain Adaptation

⁵Statistical differences

⁶Support vector machine (SVM)

⁷Nearest neighbor (NN)

⁸Maximum mean discrepancy (MMD)

⁹Kullback-Leibler divergence (KLD)

¹⁰Bregman divergence (BD)

¹¹Reproducing kernel hilbert space (RKHS)

¹²Domain invariant clustering

¹³Reconstruction coefficient matrix

¹⁴Sparse representation

¹⁵Low-rank

¹⁶Reconstruction Error minimization via Distribution Adaptation

¹⁷Landmark

¹⁸Transfer Joint Matching