طبقه‌بندی تصاویر با استفاده از نمایش تُنُک و تطبیق زیرفضا

فریماه شرافتی و جعفر طهمورث‌نژاد*

دانشکده مهندسی فناوری اطلاعات و کامپیوتر، دانشگاه صنعتی ارومیه، ارومیه، ایران

چکیده

در بیش‌تر الگوریتم‌های یادگیری ماشین و پردازش تصویر، فرض اولیه بر این است که توزیع احتمال داده‌های آموزشی (دامنه منبع) و آزمایش (دامنه هدف) یکسان است؛ اما در کاربردهای دنیای واقعی، برخی معیارها نظیر حالت تصویر، روشنایی یا کیفیت تصویر، موجب ایجاد اختلاف قابل‌توجهی بین دو مجموعهٔ آموزشی و آزمایش می‌شود. به همین دلیل، اغلب مدل‌های ایجادشده بر روی داده‌های آموزشی عملکرد ضعیفی بر روی داده‌های آزمایش خواهند داشت؛ بااین‌حال، روش‌های تطبیق دامنه، راه‌حل بسیار مؤثری برای کاهش اختلاف توزیع بین دامنه‌های آموزشی و آزمایش هستند. در این مقاله یک روش تطبیق دامنه با عنوان نمایش تُنُک و تطبیق زیرفضا (SRSA) پیشنهاد شده است، که با وزن‌دهی مجدد نمونه‌های آزمایش و نگاشت داده‌ها به یک زیرفضای جدید مشکل اختلاف توزیع داده‌ها را به‌خوبی مرتفع می‌سازد. SRSA با استفاده از یک نمایش تُنُک، بخشی از مجموعه داده‌های هدف را که ارتباط قوی‌تری با داده‌های منبع دارند، انتخاب می‌کند؛ علاوه‌بر آن، SRSA با نگاشت داده‌های تُنُک هدف و داده‌های منبع به زیرفضاهای مستقل، اختلاف توزیع آنها را درفضای به‌دست‌آمده کاهش می‌دهد؛ درنهایت با برروی‌هم‌گذاری زیرفضاهای نگاشت‌شده، SRSA اختلاف توزیع بین داده‌های آموزشی و آزمایش را به کمینه می‌رساند. ما روش پیشنهادی خود را با ترتیب‌دادن چهارده آزمایش بر روی پایگاه داده‌های بصری مختلف مورد ارزیابی قرارداده و با مقایسه نتایج به‌دست‌آمده، نشان داده‌ایم که SRSA عملکرد بهتری در مقایسه با جدیدترین روش‌های یادگیری ماشین و تطبیق دامنه دارد.

واژگان کلیدی: طبقه‌بندی تصویر، تطبیق دامنه‌های بصری، نمایش تُنُک، تطبیق زیرفضا.

Image Classification via Sparse Representation and Subspace Alignment

Farimah Sherafati & Jafar Tahmoresnezhad*

Faculty of IT & Computer Engineering, Urmia University of Technology, Urmia, Iran

Abstract

Image representation is a crucial problem in image processing where there exist many low-level representations of image, i.e., SIFT, HOG and so on. But there is a missing link across low-level and high-level semantic representations. In fact, traditional machine learning approaches, e.g., non-negative matrix factorization, sparse representation and principle component analysis are employed to describe the hidden semantic information in images, where they assume that the training and test sets are from same distribution. However, due to the considerable difference across the source and target domains result in environmental or device parameters, the traditional machine learning algorithms may fail.

Transfer learning is a promising solution to deal with above problem, where the source and target data obey from different distributions. For enhancing the performance of model, transfer learning sends the

* Corresponding author

*نویسندهٔ عهده‌دار مکاتبات

• تاریخ ارسال مقاله: ۱۳۹۷/۰۵/۲۷ • تاریخ پذیرش: ۱۳۹۸/۰۸/۱۲ • تاریخ انتشار: ۱۳۹۹/۰۶/۲۴ • نوع مطالعه: پژوهشی

[DOI: 10.29252/jsdp.17.2.58]

knowledge from the source to target domain. Transfer learning benefits from sample reweighting of source data or feature projection of domains to reduce the divergence across domains.

Sparse coding joint with transfer learning has received more attention in many research fields, such as signal processing and machine learning where it makes the representation more concise and easier to manipulate. Moreover, sparse coding facilitates an efficient content-based image indexing and retrieval.

In this paper, we propose image classification via Sparse Representation and Subspace Alignment (SRSA) to deal with distribution mismatch across domains in low-level image representation. Our approach is a novel image optimization algorithm based on the combination of instance-based and feature-based techniques. Under this framework, we reweight the source samples that are relevant to target samples using sparse representation. Then, we map the source and target data into their respective and independent subspaces. Moreover, we align the mapped subspaces to reduce the distribution mismatch across domains. The proposed approach is evaluated on various visual benchmark datasets with 14 experiments. Comprehensive experiments demonstrate that SRSA outperforms other latest machine learning and domain adaptation methods with significant difference.

Keywords: Image classification, Visual domains adaptation, Sparse representation, Subspace alignment.

۱- مقدمه

امروزه با گسترش روزافزون روش‌های مختلف جمع‌آوری اطلاعات دیجیتال مانند پویش‌گرها[1] و دوربین‌های دیجیتالی، پردازش تصویر کاربرد فراوانی یافته است. پردازش تصویر بیشتر به موضوع پردازش تصاویر دیجیتال گفته می‌شود که شاخه‌ای از علم رایانه است که با پردازش سیگنال‌های دیجیتال که نماینده تصاویر گرفته‌شده با دوربین دیجیتال یا پویش‌شده توسط پویش‌گر هستند، سر و کار دارد. با افزایش تصاویر دیجیتال مسأله طبقه‌بندی تصاویر به یکی از نیازهای اساسی تبدیل شده است. حوزه‌های کار با تصاویر در چهار زمینهٔ بهبود کیفیت ظاهری[2]، بازسازی تصاویر مختل‌شده[3]، فشردگی و رمزگذاری تصویر[4] و درک و طبقه‌بندی تصویر توسط ماشین متمرکز می‌شود.

یادگیری ماشین[5] مجموعه‌ای از الگوریتم‌های رایانه‌ای است که با استفاده از تصاویر آموزشی برچسب‌دار موجود، یک مدل یادگیری ساخته و از آن برای برچسب‌گذاری تصاویر آزمایش استفاده می‌کند. فرض اصلی در اغلب الگوریتم‌های یادگیری ماشین این است که داده‌های آموزشی و آزمایش دارای توزیع یکسانی هستند، اما در کاربردهای دنیای واقعی این فرض همیشه برقرار نبوده و ممکن است داده‌های آموزشی و آزمایش توزیع به‌طورکامل متفاوتی داشته باشند. در این حالت، مدل ساخته‌شده روی داده‌های آموزشی، عملکرد خوبی روی داده‌های آزمایش ندارد [1, 2].

یک راه حل ساده برای حل این مشکل جمع‌آوری مجدد داده‌های برچسب‌دار است؛ اما جمع‌آوری داده‌های برچسب‌دار نیازمند صرف هزینه و وقت بسیار بوده و بهترین راه حل ممکن نیست. راه حل دیگر استفاده از داده‌های برچسب‌دار از منابع دیگر است، که این راه حل هم منجربه مشکل دیگری به‌نام انتقال دامنه‌ها[6] می‌شود. انتقال دامنه‌ها درواقع به معنای اختلاف توزیع در داده‌های آموزشی و آزمایش است که باعث کاهش کارایی مدل می‌شود. برای کاهش اثرات انتقال دامنه‌ها بر روی مدل از یک راه حل شناخته‌شده به‌نام یادگیری انتقالی[7] یا تطبیق دامنه‌ها[8] استفاده می‌شود. یادگیری انتقالی با انتقال دانش از یک مجموعه برچسب‌دار به یک مجموعه بدون برچسب باعث افزایش کارایی مدل‌های طبقه‌بندی تصاویر می‌شود.

درکل مسائلی که در حوزه یادگیری انتقالی مطرح می‌شوند با توجه به نوع داده مجموعه آزمایش به دو دسته کلی تقسیم می‌شوند: (۱) یادگیری بدون‌نظارت[9]، (۲) یادگیری نیمه‌نظارت‌شده[10]. در یادگیری بدون‌نظارت اگرچه مجموعه داده آموزشی به‌طورکامل برچسب‌دار است ولی مجموعه داده آزمایش به‌طورکامل بدون برچسب بوده و مدل ایجادشده بایستی بتواند این مجموعه را برچسب‌گذاری کند. در یادگیری نیمه‌نظارت‌شده مجموعه آموزشی به‌طورکامل دارای برچسب بوده و علاوه‌بر آن بخش کوچکی از مجموعه داده آزمایش نیز دارای برچسب است. در این مجموعه، تعداد داده‌های برچسب‌دار مجموعه آزمایش، به‌قدری نیست که بتوان با استفاده از آن یک مدل دقیق و درست ایجاد کرد.

[1] Scanners
[2] Enhancement
[3] Restoration
[4] Compression and Coding
[5] Machine Learning
[6] Domains Shift
[7] Transfer Learning
[8] Domains Adaptation
[9] Unsupervised Learning
[10] Semi-supervised Learning

[DOI: 10.29252/jsdp.17.2.58]

در چنین شرایطی از یادگیری انتقالی جهت انتقال دانش بین دو مجموعه داده استفاده می‌شود. در این مقاله، بر روی مسأله یادگیری بدون‌نظارت تمرکز شده است.

در سال‌های اخیر شاهد توسعه روزافزون استفاده از مفهوم نمایش تُنُک[1] در کاربردهای مختلف پردازش تصویر هستیم. استفاده از نمایش تُنُک در مسائل طبقه‌بندی تصاویر نتایج موفقیت‌آمیزی به همراه داشته است. ایده اصلی در روش‌های طبقه‌بندی تصاویر مبتنی‌بر نمایش تُنُک، نمایش هر داده به‌صورت ترکیب خطی تُنُک از سایر داده‌ها است؛ به‌گونه‌ای که داده‌های مشابه با داده مورد نظر در این ترکیب خطی بیشترین وزن را به خود اختصاص دهند [3].

در این مقاله، یک روش برای تطبیق دامنه بدون‌نظارت بین داده‌های آموزشی و آزمایش به‌نام طبقه‌بندی تصاویر با استفاده از نمایش تُنُک و تطبیق زیرفضا (SRSA[2]) پیشنهاد شده است. روش پیشنهادی با استفاده از نمایش تُنُک، نمونه‌هایی از دامنه منبع که شباهت بیشتری را با نمونه‌های دامنه هدف دارند انتخاب می‌کند، و با استفاده از PCA[3] [4] که یک روش شناخته‌شده برای کاهش ابعاد فضای ویژگی داده‌ها می‌باشد، داده‌های انتخاب‌شده از دامنه منبع را به زیرفضای منبع و داده‌های دامنه هدف را به زیرفضای هدف نگاشت می‌کند، به گونه‌ای که خطای بازسازی[4] داده‌ها در زیرفضاهای جدید کمینه باشد. همچنین روش پیشنهادی برای کاهش تغییرات بین دامنه‌ای، زیرفضاهای دامنه‌های منبع و هدف را با یکدیگر تطبیق می‌دهد. این هدف با یافتن یک ماتریس نگاشت امکان‌پذیر می‌شود، به‌طوری که با اعمال این ماتریس نگاشت بر روی داده‌های انتخاب‌شده از دامنه منبع در زیرفضای جدید، اختلاف میان داده‌های آموزشی و آزمایش در زیرفضاهای مجزا کاهش می‌یابد.

به‌طور کلی هر کدام از مراحل روش پیشنهادی شامل فرایندهای زیر جهت ایجاد تطبیق بین دامنه‌های منبع و هدف هستند: ۱) برخلاف روش‌های دیگر که از همه داده‌های دامنه منبع برای ساخت طبقه‌بند استفاده می‌کنند، روش ما با استفاده از نمایش تُنُک ابتدا داده‌های دامنه منبع را که شباهت کمی به داده‌های دامنه هدف دارند، حذف می‌کند؛ سپس، داده‌های هدف را بر اساس ترکیب خطی تُنُک از داده‌های منبع نمایش می‌دهد، ۲) نگاشت دامنه‌های منبع و هدف به زیرفضای کم‌بُعد مبتنی‌بر نظریه کاهش بُعد و تطبیق زیرفضاهای جدید انجام می‌شود که توزیع دامنه‌ها را به یکدیگر نزدیک می‌کند.

عملکرد روش پیشنهادی بر روی مجموعه داده‌های تشخیص اشیا (آفیس و کالتک) و اعداد دست‌نویس (MNIST،USPS) مورد بررسی قرار گرفته و با روش‌های مشابه در این حوزه مقایسه شده است. نتایج به‌دست‌آمده از آزمایش‌های جامع انجام‌شده، نشان‌دهنده برتری قابل‌ملاحظه روش پیشنهادی است.

ادامه این مقاله به‌صورت زیر سازماندهی شده است. مروری بر کارهای گذشته در بخش دوم مقاله و تعاریف در بخش سوم آورده شده است. بخش چهارم روش پیشنهادی، آزمایش‌های جامع انجام‌گرفته بر روی پایگاه‌داده‌های مختلف در بخش پنجم و در انتها، نتیجه‌گیری و کارهای پیشنهادی آینده آورده شده است.

۲- مروری بر کارهای گذشته

تطبیق و انتقال دامنه‌ها دو مسأله اساسی در یادگیری ماشین هستند و در سال‌های اخیر توجه بسیاری از پژوهش‌گران را در حوزه‌هایی همچون پردازش زبان‌های طبیعی و بینایی ماشین به‌خود جلب کرده است. روش‌های زیادی برای کاهش اختلاف توزیع بین دامنه‌ها پیشنهاد شده است، ولی با این‌حال سه رویکرد کلی در این حوزه وجود دارد: رویکردهای مبتنی‌بر نمونه، رویکردهای مبتنی‌بر خصوصیت و رویکردهای مبتنی‌بر مدل. در این مقاله، از ترکیب رویکردهای مبتنی‌بر نمونه و مبتنی‌بر خصوصیت استفاده شده که در زیر به‌تفصیل توضیح داده شده است.

۲-۱- رویکردهای مبتنی‌بر نمونه

در رویکردهای مبتنی‌بر نمونه، هدف، انتخاب نمونه‌هایی از دامنه منبع است که توزیع نزدیک‌تری با نمونه‌های دامنه هدف داشته باشند تا اختلاف توزیع بین دو دامنه کاهش یابد. لندمارک [5] یکی از روش‌های مبتنی‌بر نمونه است که برای محاسبه میزان اختلاف توزیع بین دامنه‌ها از MMD[5] استفاده می‌کند و به نمونه‌های منبع که اختلاف توزیع کمتری با نمونه‌های هدف دارند، وزن بیشتری اختصاص می‌دهد.

[1] Sparse Representation
[2] image classification via Sparse Representation and Subspace Alignment (SRSA)
[3] Principal Component Analysis
[4] Reconstruction error
[5] Maximum Mean Discrepancy

[DOI: 10.29252/jsdp.17.2.58]

طبقه‌بندی‌کننده مبتنی‌بر نمایش تُنُک (SRC[1]) [6, 7] یکی دیگر از روش‌های مبتنی‌بر نمونه است که در سامانه‌های تشخیص چهره نتایج قابل‌توجهی را به‌دست‌آورده است. ایده اصلی SRC به این صورت است که هر تصویر چهره به‌صورت ترکیب خطی تُنُک از سایر تصاویر چهره نمایش داده می‌شود. انتخاب تصاویر با استفاده از ضرایب آن‌ها صورت می‌گیرد؛ به‌گونه‌ای که تصاویر با ضرایب صفر حذف شده و تصاویری که ضرایب غیرصفر و وزن بیشتری به آن‌ها تعلق گرفته است، انتخاب می‌شوند.

روش SSSC[2] [8]، برای کاهش اختلاف توزیع بین دامنه‌های منبع و هدف، با یافتن یک ماتریس انتخاب، نمونه‌های منبع مشابه با نمونه‌های هدف را انتخاب می‌کند و با استفاده از نمایش تُنُک هر نمونه هدف را بر اساس یک ترکیب خطی تُنُک از نمونه‌های منبع نمایش می‌دهد.

۲-۲- رویکردهای مبتنی‌بر خصوصیت

در رویکردهای مبتنی‌بر خصوصیت، فضای ویژگی[3] برای نمایش بهتر ویژگی‌های مشترک دو دامنه تغییر می‌کند. GFK[4] [9] و TCA[5] [10] از جمله این روش‌ها هستند. GFK داده‌های منبع و هدف را بر روی منیفلد گرسمن[6] نگاشت می‌کند و بین زیرفضاهای منبع و هدف یک مسیر کوتاه[7] بر روی منیفلد ایجاد می‌کند تا تغییرات دامنه از منبع تا هدف را مورد ارزیابی قرار دهد. از آن‌جایی که حرکت بر روی مسیر کوتاه بایستی به‌صورت روان و پیوسته باشد، GFK اندازه زیرفضا را کوچک درنظر می‌گیرد که موجب ازبین‌رفتن بخشی از داده‌های ورودی می‌شود. TCA اختلاف توزیع حاشیه‌ای بین دامنه‌ها را با استفاده از MMD کم می‌کند و با بیشینه‌کردن میزان واریانس داده‌ها، ویژگی‌های اصلی آن‌ها را حفظ می‌کند. SA[8] [11] داده‌های منبع را به زیرفضای منبع و داده‌های هدف را به زیرفضای هدف نگاشت می‌کند، و با استفاده از یک ماتریس نگاشت زیرفضاهای جدید را تطبیق می‌دهد و بدین صورت اختلاف توزیع بین دامنه‌ها را کاهش می‌دهد. SCA[9] [12] با بیشینه‌سازی اختلاف بین نمونه‌ها در رده‌های مختلف و کمینه‌سازی اختلاف بین نمونه‌های هر رده، دامنه‌ها را به زیرفضای جدید نگاشت می‌دهد. SDA[10] [13] هم‌زمان با ایجاد تطبیق بین زیرفضاهای دامنه منبع و هدف، اختلاف توزیع بین دامنه‌ها را از طریق تطبیق واریانس مؤلفه‌های اساسی دامنه‌ها به کمینه می‌رساند. VDA[11] [14] برای کاهش اختلاف توزیع حاشیه‌ای و شرطی بین دامنه‌ها، یک نمایش کم‌بُعد از دامنه‌های منبع و هدف ایجاد می‌کند. همچنین، VDA از خوشه‌بندی مستقل از دامنه[12] جهت حفظ ساختار هندسی و آماری بین دامنه‌ها در فضای جدید استفاده می‌کند.

روش پیشنهادی در این مقاله، یک راه حل بدون‌نظارت با بهره‌گیری از روش‌های مبتنی‌بر نمونه و مبتنی‌بر خصوصیت است. در این روش یک مدل طبقه‌بند توسط کاهش اختلاف توزیع بین نمونه‌های دامنه منبع و هدف با استفاده از نمایش تُنُک و همچنین با استفاده از تطبیق زیرفضاها ایجاد می‌شود که دارای کمینه خطای پیش‌بینی است.

۳- تعاریف و مفاهیم اولیه

۳-۱- هدف پژوهش

تطبیق دامنه یکی از مسائل مطرح در حوزه یادگیری ماشین است. این مقاله، به‌دنبال پیشنهاد روشی است که بتواند بر محدودیت‌های الگوریتم‌های یادگیری ماشین کلاسیک غلبه کند. در سال‌های اخیر استفاده از نمایش تُنُک در کاربردهای مختلف پردازش تصویر رواج زیادی پیدا کرده است؛ بنابراین، نمایش تُنُک به‌عنوان مدلی مناسب جهت مدل‌سازی داده‌ها در فضای جدید استفاده می‌شود.

۳-۲- تعریف دامنه[13] و وظیفه[14]

در یادگیری انتقالی هر دامنه از دو مفهوم زیر تشکیل شده است: فضای ویژگی X و توزیع احتمال حاشیه‌ای P(X). بنابراین دامنه D به‌صورت D={X,P(X)} نمایش داده می‌شود. بدین ترتیب، زمانی دو دامنه متفاوت هستند که یا فضای ویژگی و یا توزیع حاشیه‌ای متقاوتی داشته باشند.

برای هر دامنه D یک وظیفه T وجود دارد که شامل مجموعهٔ برچسب Y و تابع پیش‌بینی $f(x)$ است، یعنی

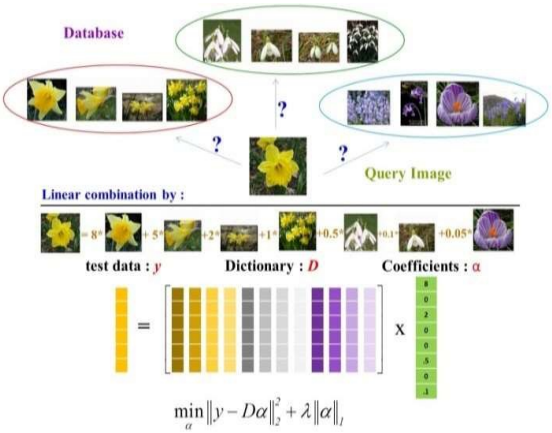
[1] Sparse Representation-based Classifier
[2] Sample Selection Sparse Coding
[3] Feature Space
[4] Geodesic Flow Kernel
[5] Transfer Component Analysis
[6] Grassman Manifold
[7] Geodesic
[8] Subspace Alignment
[9] Scatter Component Analysis
[10] Subspace Distribution Alignment
[11] Visual Domain Adaptation via transfer feature learning
[12] Domain invariant clustering
[13] Domain
[14] Task

[DOI: 10.29252/jsdp.17.2.58]

$T=\{Y,f(x)\}$. تابع $f(x)$ برای نمونه‌های ورودی x، مجموعه برچسب Y را پیش‌بینی می‌کند و می‌توان آن را به‌صورت توزیع احتمال شرطی $P(Y|x)$ بیان کرد. بدین ترتیب، زمانی دو وظیفه متفاوت هستند که یا مجموعه برچسب و یا تابع توزیع احتمال شرطی متفاوتی داشته باشند؛ بدین معنی که $Y_s \neq Y_t$ یا $P_s(Y_s|X_s) \neq P_T(Y_T|X_T)$.

۳-۳- نمایش تُنُک

امروزه نمایش تُنُک به‌عنوان ابزاری قوی و کارآمد برای نمایش سیگنال‌ها تبدیل شده و در کاربردهایی همچون نمونه‌برداری، بازیابی و طبقه‌بندی تصاویر مورد استفاده قرار گرفته است. هدف این نوع نمایش، تقریب داده‌های آزمایش (تصویر جدید) با استفاده از تعداد کمی (تُنُک) از داده‌های آموزشی است. طبقه‌بندی با استفاده از نمایش تُنُک به این صورت است که داده‌های آموزشی به‌عنوان دیکشنری[1] درنظر گرفته می‌شوند و داده‌های هدف به‌صورت ترکیب خطی از آن‌ها نمایش داده می‌شوند. همان‌طور که در شکل (۱) نشان داده شده، داده آزمایش y، بر اساس ترکیب خطی از دیکشنری (D) نمایش داده شده است.



(شکل-۱): نحوه عملکرد نمایش تُنُک. y داده آزمایش، D دیکشنری، α ماتریس ضرایب و λ پارامتر تنظیم می‌باشد. داده آزمایش y با استفاده از دیکشنری و ماتریس ضرایب، به‌صورت ترکیب خطی از دیکشنری نمایش داده می‌شود.

(Figure-1): Sparse representation methodology. y is test data, D shows dictionary, α is the coefficient matrix and λ illustrates the tuning parameter. Test data y is represented by a linear combination of the dictionary and coefficient matrix.

[1] Dictionary

دامنه منبع با $D_S=\{X_S,Y_S\}$ و دامنه هدف با $D_T=\{X_T,Y_T\}$ نمایش داده می‌شوند که در آن $X_s\in R^{d_s*n_s}$ و $X_T\in R^{d_T*n_T}$ به‌ترتیب داده‌های منبع و هدف هستند و $Y_s\in R^{n_s*1}$ و $Y_T\in R^{n_T*1}$ برچسب‌های منبع و هدف، d_s ابعاد نمونه‌های منبع، d_T ابعاد نمونه‌های هدف، n_s و n_T به‌ترتیب نشان‌دهنده تعداد نمونه‌های منبع و هدف است. فرض شده $d_s = d_T$ است و فضای ویژگی دامنه‌های منبع و هدف یکسان است.

۴-۳- کاهش بُعد

برای به‌دست‌آوردن زیرفضاهای مشترک یا مستقل می‌توان از روش‌های کاهش بعد برای قراردادن داده‌ها در یک زیرفضای تعبیه‌شدهٔ جدید استفاده کرد. تجزیه مؤلفه‌های اصلی (PCA) یک روش کاربردی برای استخراج ویژگی از داده‌ها (سیگنال یا تصویر) است. به تعبیر دیگر می‌توان با PCA کاهش بعد انجام داده و ویژگی‌های جدید استخراج کرد. PCA داده‌ها را به‌گونه‌ای به یک زیرفضای جدید نگاشت می‌کند که ساختار اصلی داده‌ها حفظ شود. بدین ترتیب، داده‌ها بر روی اجزای اصلی خود نگاشت می‌شوند که شامل اجزایی است که دارای واریانس بالایی هستند. از بین اجزای به‌دست‌آمده، اجزایی که دارای حداکثر واریانس باشند به‌عنوان اجزای اصلی جهت نگاشت داده‌ها به فضای کم‌بعد استفاده می‌شوند.

تابع بهینه‌سازی PCA با استفاده از رابطه $H=I_{n_s+n_t}-\frac{1}{n_s+n_t}\vec{1}\vec{1}^T$ به‌دست می‌آید، که $\max_{A^TA=I} tr(A^TXHX^TA)$ تابع مرکزیت بوده، I ماتریس همانی و $\vec{1}$ بردار ستونی از یک‌ها است [4]. tr نشان‌دهنده حاصل جمع عناصر قطر اصلی ماتریس است.

۴- روش پیشنهادی

در این مقاله یک روش نوین با عنوان SRSA پیشنهاد شده است که با استفاده از نمایش تُنُک نمونه‌های منبع مرتبط با نمونه‌های هدف را انتخاب کند. SRSA یک نمایش کم‌بعد برای نمونه‌های منبع انتخاب‌شده و نمونه‌های هدف، ایجاد می‌کند و برای کاهش بعد داده‌ها از PCA استفاده می‌کند. سپس، SRSA زیرفضاهای ایجادشده منبع و هدف را با یکدیگر تطبیق داده تا توزیع آنها بر روی یکدیگر منطبق شود.

[DOI: 10.29252/jsdp.17.2.58]

۴-۱- انتخاب نمونه

SRSA برای انتخاب نمونه‌ها از یک نمایش تُنُک استفاده می‌کند و از نرم $l_{2,1}$ برای محدودکردن ضرایب ماتریس انتخاب A بهره می‌برد [15]. SRSA با استفاده از رابطه زیر، ماتریسی به‌دست می‌آورد که اختلاف توزیع بین داده‌های منبع و هدف را کمینه می‌سازد و می‌توان هر نمونه هدف را به‌صورت ترکیب خطی از کمینه نمونه‌های منبع نمایش داد.

$$\min_{\mathbf{A}} \| \mathbf{X}_T - \mathbf{X_S A} \|_F^2 + \rho \| \mathbf{A} \|_{2,1} \qquad (۱)$$

که در این رابطه، $\|.\|_F$ نرم فروبنیوس[1] و ماتریس $\mathbf{A} \in R^{n_s * n_T}$ یک ماتریس تُنُک است، که با استفاده از مسأله بهینه‌سازی بالا به‌دست می‌آید و ρ پارامتر تنظیم پراکندگی است. با استفاده از نرم $l_{2,1}$ هر سطر ماتریس همزمان صفر یا غیرصفر می‌شود که سطرهای غیرصفر به‌عنوان نمونه‌های منبع مرتبط با نمونه‌های هدف انتخاب و نمونه‌های منبع انتخاب‌شده با $\boldsymbol{X_{s1}}$ نمایش داده می‌شوند.

۴-۲- تولید زیرفضا

با وجود اینکه داده‌های منبع و هدف در فضای Dبُعدی یکسان قرار دارند، با این حال توزیع حاشیه‌ای متفاوتی دارند. بنابراین، داده‌های منبع و هدف با استفاده از PCA به زیرفضاهای مربوط به خود نگاشت داده می‌شوند تا اختلاف توزیع آن‌ها کاهش یابد. برای هر دامنه، d بردار ویژه مربوط به بزرگ‌ترین مقادیر ویژه PCA انتخاب می‌شوند. این بردارهای ویژه به‌عنوان زیرفضاهای منبع و هدف استفاده می‌شوند که به‌ترتیب با $\boldsymbol{P_S}$ و $\boldsymbol{P_T}$ نمایش داده می‌شوند ($\mathbf{P_s}, \mathbf{P_T} \in \mathbf{R^{D*d}}$).

۴-۳- تطبیق دامنه با استفاده از تطبیق زیرفضا

استراتژی اصلی در زیرفضای مبتنی‌بر تطبیق دامنه، این‌گونه است که داده‌های منبع و هدف به یک زیرفضای مشترک نگاشت می‌شوند، اما در این حالت فقط ویژگی‌های مشترک از دو دامنه استخراج می‌شوند و بعضی از اطلاعات که مختص هر کدام از دامنه‌ها هستند از بین می‌روند. با این‌حال، SRSA داده‌های منبع و هدف هر کدام از دامنه‌ها را به زیرفضای مربوط به خود نگاشت داده و سپس، زیرفضاهای به‌دست‌آمده از هر دامنه را با یکدیگر تطبیق می‌دهد که برای این هدف از استراتژی تطبیق زیرفضا استفاده می‌کند.

با استفاده از PCA، نمونه‌های منبع انتخاب‌شده ($\boldsymbol{X_{s1}}$) و نمونه‌های هدف ($\boldsymbol{X_t}$) به زیرفضاهای مجزا نگاشت داده می‌شوند؛ سپس، زیرفضاهای منبع ($\boldsymbol{P_S}$) و هدف ($\boldsymbol{P_T}$) با استفاده از ماتریس نگاشت Mبا یکدیگر تطبیق داده می‌شوند. ماتریس نگاشت M با حداقل‌سازی واگرایی ماتریس برگمن[2] زیر به‌دست می‌آید:

$$F(M) = \| P_S M - P_T \|_F^2 \qquad (۲)$$

فرم بسته رابطه (۲) را می‌توان به‌صورت زیر نوشت:

$$M^* = argmin_M (F(M)) \qquad (۳)$$

که M^* مقدار بهینه ماتریس تطبیق را نشان می‌دهد. به‌دلیل این‌که نرم فروبنیوس برای عملیات متعامدبودن[3] ثابت[4] است، فرمول (۲) را می‌توان به‌صورت زیر بازنویسی کرد:

$$F(M) = \| P_S' P_S M - P_S' P_T \|_F^2 = \| M - P_S' P_T \|_F^2 \qquad (۴)$$

با توجه به رابطه (۴)، می‌توان نتیجه گرفت که M^* بهینه با استفاده از رابطه $M^* = P_S' P_T$ به‌دست می‌آید و $X_a = P_S P_S' P_T$ است که X_a دامنه منبع تطبیق‌شده با دامنه هدف است.

ماتریس نگاشت M^* سامانه مختصات زیرفضای منبع را با سامانه مختصات زیرفضای هدف، با استفاده از بردارهای پایه منبع تطبیق می‌دهد. عملکرد روش تطبیق زیرفضا در شکل (۲) نشان داده شده است.

۴-۴- الگوریتم روش SRSA

با مشتق‌گیری از رابطه (۱) و مساوی قراردادن آن با صفر، مقدار A بهینه به‌صورت زیر به‌دست می‌آید:

$$\mathbf{A} = (\rho \mathbf{R} + \mathbf{X_S^T X_S})^{-1} \mathbf{X_S^T X_T} \qquad (۵)$$

که در رابطه بالا، $\mathbf{R} = \frac{1}{2 \| \mathbf{A}^i \|_2}$ یک ماتریس قطری و وابسته به A است و A با یک روش تکراری به‌دست می‌آید. الگوریتم روش SRSA با جزئیات کامل، در الگوریتم (۱) ارائه شده است.

[1] Frobenius norm

[2] Bregman matrix divergence

[3] Orthonormal

[4] Invariant

[DOI: 10.29252/jsdp.17.2.58]

Algorithm 1. Algorithm of image classification via sparse representation and subspace alignment (SRSA)

Input: Source data X_S,Target data X_T, Source labels L_S, Subspace dimension d

Parameter: Regularization parameters ρ

Output: Predicted target labels L_T

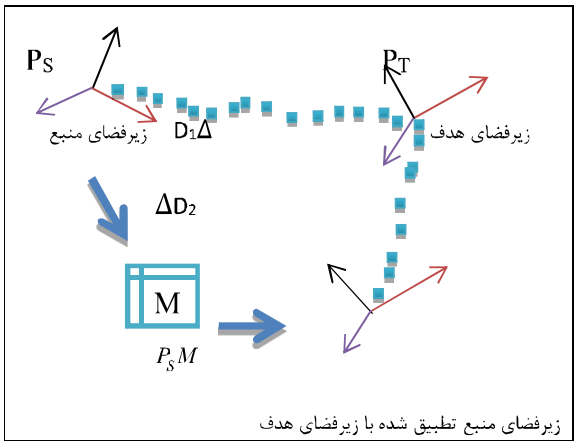
1: Repeat until the maximum number of iterations is reached

2: Compute selection matrix A via solving Eq. (5).

3: Compute diagonal matrix $N = \frac{1}{2 \| A^i \|_2}$ based on A.

4: END Repeat

5: Return $L_T = kNN(X_S A P_S' P_T, X_T P_T, L_S)$



(شکل-۲): عملکرد روش تطبیق زیرفضا. زیرفضای منبع با P_S و زیرفضای هدف با P_T نشان داده شده است. زیرفضای منبع با زیرفضای هدف تطبیق داده شده و $X_a = P_S M$ است.

(Figure-2): The performance of the subspace matching method. The source subspace is represented by P_S and the target subspace is showed by P_T. The source subspace is aligned with target subspace via $X_a = P_S M$.

در الگوریتم SRSA، هدف پیش‌بینی برچسب‌های نمونه‌های دامنه هدف با استفاده از نمایش تُنُک و تطبیق زیرفضا است. SRSA در ابتدا نمونه‌های منبع را با استفاده از ماتریس انتخاب A به‌دست‌آورده، سپس، زیرفضاهای هر کدام از دامنه‌های منبع و هدف را به‌طور مجزا محاسبه کرده و با یکدیگر تطبیق می‌دهد و درنهایت با استفاده از طبقه‌بند NN برچسب نمونه‌های هدف را پیش‌بینی می‌کند.

۵- تنظیمات اولیه محیط آزمایش

در این بخش مجموعه داده‌ها، نتایج آزمایش‌ها و عملکرد SRSA با دیگر الگوریتم‌ها مورد مقایسه، مورد ارزیابی قرار گرفته است.

۵-۱- مجموعه‌داده‌های دنیای واقعی

کارایی روش پیشنهادی در این مقاله، بر روی دو نوع پایگاه‌داده زیر مورد ارزیابی قرار گرفته است: ۱) آفیس و کالتک[1]، ۲) اعداد (USPS و MNIST).

مجموعه‌داده آفیس [16]، و کالتک [17]، از مجموعه‌داده‌های بسیار مشهور برای تطبیق دامنه هستند. مجموعه‌داده آفیس از سه دامنه آمازون[2]، وبکم[3] و DSLR[4] تشکیل شده است که شامل تصاویری با کیفیت، رنگ و روشنایی متفاوت هستند. تصاویر هر مجموعه‌داده تفاوت‌های زیادی با یکدیگر دارند. دامنه آمازون شامل تصاویری است که از سایت‌های تجاری بارگیری شده است. دامنه وبکم شامل تصاویری با وضوح پایین است که توسط دوربین‌های کم‌کیفیت وب گرفته شده‌اند و دامنه DSLR شامل تصاویری با وضوح بالا است که توسط دوربین‌های دیجیتالی حرفه‌ای گرفته شده‌اند. تصاویر موجود در دامنه کالتک، از وب‌سایت گوگل جمع‌آوری شده‌اند. با وجود این‌که دامنه‌های مختلف توزیع‌های متفاوتی دارند اما در آزمایش‌های ما از ده ردهٔ معنایی مشترک استفاده شده است. بر روی پایگاه‌داده آفیس و کالتک، دوازده آزمایش طراحی شده است که در هر یک از این آزمایش‌ها یکی از مجموعه داده‌ها (به‌عنوان نمونه وبکم)، به‌عنوان دامنه منبع و یکی از سه مجموعه داده دیگر به‌عنوان دامنه هدف انتخاب می‌شوند.

پایگاه داده اعداد شامل دو دامنه USPS [18] و MNIST [19] است. دامنه USPS، شامل اعداد دست‌نویس پویش‌شده از نامه‌های اداره پست آمریکا در اندازه‌های ۱۶*۱۶ پیکسلی است که ۷۲۹۱ داده آموزشی و ۲۰۰۷ داده آزمایش دارد. دامنه MNSIT، شامل اعداد دست‌نویس جمع‌آوری‌شده از دانش‌آموزان دبیرستانی آمریکا و کارمندان سازمان‌های مالیاتی آمریکا در اندازه ۲۸*۲۸ پیکسلی است که شامل شش‌هزار داده آموزشی و هزار داده آزمایش است. این پایگاه‌داده شامل ده رده مختلف می‌باشد و به‌منظور آزمایش هر دو دامنه در شرایط یکسان، دامنه USPS_vs_MNIST(U_M) ایجاد شده است که به‌طور تصادفی ۱۸۰۰ نمونه از داده‌های دامنه USPS به‌عنوان داده‌های آموزشی و دوهزار نمونه از داده‌های دامنه MNIST به‌عنوان داده‌های آزمایش به‌کار گرفته می‌شود. با جابه‌جایی نمونه‌های آموزشی و آزمایش دامنه

[1] Office and Caltech
[2] Amazon
[3] Webcam
[4] Digital single-lens reflex

[DOI: 10.29252/jsdp.17.2.58]

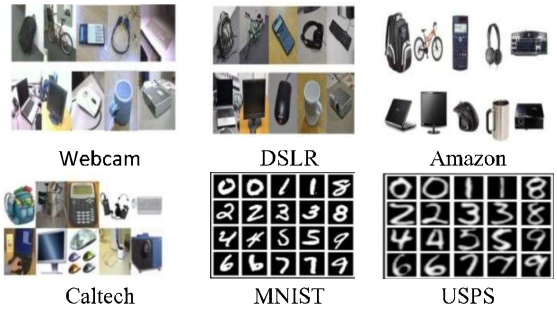
MNIST_vs_USPS(M_U) برای یک آزمایش دیگر ایجاد شده است.

به‌طور کلی کارایی الگوریتم پیشنهادی بر روی چهارده مجموعه تصاویر بین دامنه‌ای مختلف مورد ارزیابی قرار گرفته است. اطلاعات جزئی‌تر از پایگاه‌داده‌ها در جدول (۱) و تصاویر نمونه از هر دامنه در شکل (۳) نشان داده شده است.

(جدول-۱): معرفی پایگاه‌داده‌ها

(Table-1): The datasets

پایگاه داده	تعداد نمونه‌ها	تعداد ابعاد	تعداد کلاس‌ها	اختصار
USPS	1800	256	10	U
MNIST	2000	256	10	M
Amazon	958	800	10	A
Webcam	295	800	10	W
DSLR	157	800	10	D
Caltech256	1123	800	10	C



(شکل-۳): تصاویر نمونه از پایگاه‌داده‌های آفیس و کالتک و اعداد

(Figure-3): Sample images from Office & Caltech and digits datasets

۲-۵- مفروضات پیاده‌سازی

برای به‌دست‌آوردن تنظیمات مدل، روش SRSA با مقادیر مختلف پارامترها مورد ارزیابی قرار گرفته است. طبقه‌بند پایه در تمام آزمایش‌ها، طبقه‌بند NN است. به‌علاوه، پیاده‌سازی روش پیشنهادی SRSA، توسط نرم‌افزار Matlab انجام گرفته است. دقت طبقه‌بندی بر روی نمونه‌های دامنه هدف برای مقایسه الگوریتم‌ها توسط رابطه زیر محاسبه می‌شود:

$$\text{Accuracy} = \frac{|\{x: x \in D_t \wedge f(x) = y(x)\}|}{n_T}$$

که D_T، f(x) ،y(x) و n_t به‌ترتیب نشان‌دهنده دامنه هدف، تابع پیش‌بینی، برچسب واقعی داده و تعداد نمونه‌های دامنه هدف هستند.

۳-۵- تأثیر پارامترها

برای به‌دست‌آوردن تنظیمات مدل، SRSA را با مقادیر مختلف پارامترها بررسی کرده‌ایم. دو پارامتر در این روش وجود دارد که مقادیر بهینه آن‌ها در جدول (۲) آورده شده است. پارامتر ρ، پارامتر تنظیم تُنُکی می‌باشد که پارامتر تأثیرگذار در انتخاب نمونه‌های منبع است و تعداد نمونه‌های انتخاب‌شده را از دامنه منبع مشخص می‌کند. پارامتر K، ابعاد زیرفضای جدید را مشخص می‌کند. پارامتر ρ در محدوده [۰/۵، ۰/۰۰۱] و پارامتر K در محدوده [۲۲۰، ۱۰] درنظر گرفته شده‌اند. روش SRSA با پارامترهای بیان‌شده بر روی چهارده مجموعه‌داده مورد آزمایش قرار می‌گیرد.

دقت طبقه‌بندی و تعداد نمونه‌های انتخاب‌شده با توجه به پارامتر ρ در جدول (۳) نشان داده شده است. هنگامی که ρ =۰/۵ است ماتریس A تُنُکی کمتری دارد و در نتیجه نمونه‌های منبعی که توزیع نزدیک‌تری با نمونه‌های هدف دارند انتخاب می‌شوند.

نتایج به‌دست‌آمده بر روی پایگاه‌داده‌های آفیس و کالتک در شکل (۴) نشان‌دهنده این است که پایگاه‌داده آفیس و کالتک دارای رفتار متفاوتی نسبت به مقادیر مختلف پارامتر K هستند. برای به‌دست‌آوردن مقدار محدوده بهینه پارامتر، محدوده‌ای انتخاب می‌شود که دارای بهینه‌ترین دقت بر روی تعداد بیشتری از پایگاه داده‌ها باشد. برای بیشتر پایگاه‌داده‌ها محدوده بهینه پارامتر K در پایگاه‌داده آفیس و کالتک، [۴۰، ۶۰] می‌باشد. نتایج به‌دست‌آمده برای مقادیر مختلف K، در پایگاه داده اعداد نشان‌دهنده این است که پایگاه‌داده اعداد دارای حساسیت زیادی نسبت به مقادیر مختلف K و بهترین محدوده برای پارامتر K در پایگاه‌داده اعداد، نیز [۴۰، ۶۰] است.

(جدول-۲): مقادیر بهینه پارامترها برای دو پایگاه‌داده بصری (آفیس و کالتک، اعداد)، ρ پارامتر تنظیم تُنُکی، K ابعاد زیرفضای جدید

(Table-2): Optimal parameters for two visual datasets (office & Caltech, digits), ρ is the sparsity parameter and K is the new subspace dimension

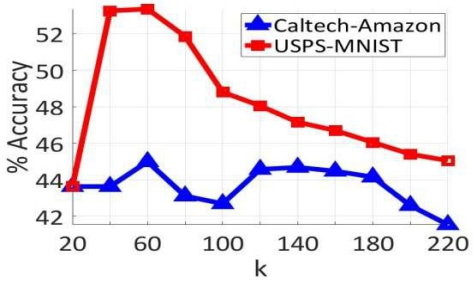
SRSA	l_1 norm	Orj	S-T
54.35	53.95	54.40	U-M
68.56	66.11	66.44	M-U
61.46	60.03	60.42	Avg

[DOI: 10.29252/jsdp.17.2.58]

(جدول-۳): تأثیر پارامتر تنظیم تُنُکی ρ بر تعداد نمونه‌های انتخاب‌شده از دامنه منبع در پایگاه‌داده اعداد

(Table-3): The effect of sparsity regularization on the number of selected samples from the source domain on the digit dataset

ρ	0.001	0.005	0.01	0.05	0.1	0.5
متوسط دقت U-M	54.40	54.40	54.40	54.40	54.40	54.35
U-M #	1800	1800	1800	1800	1800	1780
متوسط دقت M-U	66.44	66.44	66.44	66.44	66.44	68.56
M-U #	2000	2000	2000	2000	2000	1366



(شکل-۴): ارزیابی دقت پایگاه‌داده‌های USPS-MNIST، Caltech-Amazon با مقادیر مختلف پارامتر k

(Figure-4): USPS-MNIST, Caltech-Amazon datasets accuracy assessment with different values of k parameter

۵-۴- ضرورت هر پارامتر در الگوریتم

در این بخش، ضرورت دو پارامتر بر روی پایگاه‌داده اعداد بررسی شده و نتایج آن در جدول (۴) گزارش شده است.

(جدول-۴): ضرورت هر پارامتر در الگوریتم

S-T: دامنه منبع و هدف، Orj: همه نمونه‌های منبع

(Table-4): The necessity of each parameter in the algorithm

پارامترها	K	ρ
آفیس و کالتک	50	0.5
اعداد	50	0.5

۱. استفاده از همه نمونه‌های منبع (Orj): انتخاب همه نمونه‌های منبع برای آموزش مدل به‌جای انتخاب برخی از آن‌ها.

۲. استفاده از نرم l_1 به‌جای نرم $l_{2,1}$: با استفاده از نرم l_1 به‌جای نرم $l_{2,1}$ تعداد ضرایب غیرصفر در ماتریس $\mathbf{A}$ افزایش یافته و به‌طوردقیق نمی‌توان هر نمونه هدف را با تعداد کمی از نمونه‌های منبع نمایش داد.

نتایج آزمایش‌ها حاکی از این است که استفاده از نمایش تُنُک برای انتخاب نمونه‌ها تأثیر به‌سزایی بر روی دقت مدل دارد. استفاده از همه نمونه‌های منبع کارایی مدل را کاهش می‌دهد، به‌طوری‌که انتخاب برخی از نمونه‌ها با استفاده از نمایش تُنُک برای ساخت مدل پیش‌بینی، متوسط بهبود دقت را ۱/۰۴ افزایش می‌دهد.

همچنین، همان‌طور که در جدول نشان داده شده است، استفاده از نرم l_1 برای تعیین ضرایب ماتریس تُنُک کارایی را به‌طور ملموس پایین می‌آورد. زمانی که از نرم $l_{2,1}$ استفاده شود، به‌دلیل این‌که نمونه‌های منبع مرتبط با هدف دقیقاً مشخص می‌شود، متوسط دقت مدل ۱/۴۳ افزایش می‌یابد.

۵-۵- مقایسه با الگوریتم‌های موجود

الگوریتم SRSA با هشت الگوریتم زیر مورد مقایسه قرار گرفته است: طبقه‌بند نزدیک‌ترین همسایه (NN[1])، SA [9]، PCA [3]، DAM [20]، SDA [11]، GFK [7]، TCA [8] و SCA [10]. به‌دلیل اینکه تمامی روش‌های نام‌برده‌شده (به‌جز NN)، روش‌های کاهش بُعد هستند، از طبقه‌بند استاندارد نزدیک‌ترین همسایه برای آموزش یک طبقه‌بند بر روی داده‌های منبع جهت پیش‌بینی داده‌های بدون برچسب دامنه هدف استفاده می‌شود.

دقت طبقه‌بندی SRSA و هشت روش دیگر بر روی پایگاه‌داده‌های آفیس و کالتک و اعداد در جداول (۵ و ۶) نشان داده شده است. همان‌طورکه مشاهده می‌شود SRSA کارایی بهتری نسبت به هشت روش دیگر از خود نشان داده و در بیشتر حالات عملکرد مطلوبی ارائه می‌دهد. نتایج گزارش‌شده شامل میانگین برای همه روش‌ها است.

در پایگاه‌داده اعداد، SRSA دارای ۴/۹ متوسط بهبود دقت نسبت‌به بهترین الگوریتم مورد مقایسه SCA و ۶/۱۴ نسبت‌به الگوریتم استاندارد NN است. در پایگاه‌داده آفیس و کالتک، SRSA دارای ۱/۴۳ متوسط بهبود دقت نسبت به بهترین الگوریتم مورد مقایسه SCA و ۱۵/۲۲ نسبت به الگوریتم استاندارد NN است. در ادامه SRSA را با تمام روش‌ها به‌صورت جزئی مقایسه می‌کنیم.

PCA یک روش کاهش بُعد است که یک فضای مشترک کم‌بعد بین دامنه‌ها ایجاد می‌کند. PCA اختلاف توزیع بین دامنه‌ها را چندان کاهش نمی‌دهد و عملکرد ضعیفی نسبت به دیگر الگوریتم‌های تطبیق دامنه نشان می‌دهد و دارای دقت پایینی است. در پایگاه‌داده اعداد، SRSA دارای ۵/۸۷ متوسط بهبود دقت نسبت به PCA و در

[1] Nearest neighbor

[DOI: 10.29252/jsdp.17.2.58]

پایگاه‌داده آفیس و کالتک، دارای ۶/۹۴ متوسط بهبود دقت و عمده‌ترین دلیل برتری SRSA نسبت به PCA استفاده از تطبیق زیرفضا برای کاهش اختلاف بین دامنه‌ها است.

TCA یک روش تطبیق خصوصیات است که با استفاده از اجزای انتقال، یک نمایش مشترک بین دامنه‌های منبع و هدف ایجاد می‌کند که در نمایش جدید، اختلاف توزیع حاشیه ای بین دامنه‌های منبع و هدف کاهش می‌یابد. در این روش به اختلاف توزیع شرطی بین دامنه های منبع و هدف توجهی نشده و همچنین از داده‌های برچسب‌دار دامنه منبع در ایجاد فضای جدید استفاده نمی‌شود؛ درحالی‌که در روش SRSA ، علاوه‌بر کاهش اختلاف توزیع حاشیه‌ای بین دامنه‌ها، از برچسب‌های نمونه‌های دامنه منبع برای ایجاد تفکیک‌پذیری بهتر بین رده‌ها استفاده شده است. به همین دلیل، عملکرد روش SRSA نسبت به روش TCA بهتر است. در پایگاه‌داده اعداد، SRSA دارای ۷/۷۹ متوسط بهبود دقت نسبت به TCA و در پایگاه‌داده آفیس و کالتک، دارای ۳/۵۶ متوسط بهبود دقت است.

DAM یک روش تطبیق دامنه چندمنبعی است که یک طبقه‌بند مقاوم برای پیش‌بینی برچسب نمونه‌های آزمایش با به‌کارگیری یک مجموعه از طبقه‌بندهای پایه از پیش یادگرفته‌شده ایجاد می‌کند. در پایگاه‌داده اعداد، SRSA دارای ۱۳/۷ متوسط بهبود دقت نسبت به DAM و در پایگاه‌داده آفیس و کالتک، دارای ۴/۰۴ متوسط بهبود دقت است.

در روش GFK، زیرفضاهایی از دامنه‌های منبع و هدف ایجاد می‌شود که در این زیرفضاها اختلاف توزیع حاشیه‌ای بین دامنه‌های منبع و هدف به کمینه می‌رسد، اما به‌دلیل اینکه زیرفضاهای نگاشت‌شده دارای ابعاد بسیار کمی است، باعث از بین‌رفتن بخشی از داده شده و نمایش خوبی از داده‌ها ایجاد نمی‌شود. در پایگاه‌داده اعداد، SRSA دارای ۴/۶۲ متوسط بهبود دقت نسبت به GFK و در پایگاه‌داده آفیس و کالتک، دارای ۳/۶۴ متوسط بهبود دقت است. از جمله دلایل بهبود این است که در روش SRSA یک زیرفضا برای هر کدام از دامنه‌ها ایجاد می‌شود که ساختار داده‌ها را نیز حفظ می‌کند و همچنین با استفاده از نمایش تُنُک فقط از برخی از نمونه‌های دامنه منبع برای ساخت طبقه‌بند استفاده می‌کند.

SA نیز یکی از روش‌های کاهش بعد در تطبیق بین دامنه‌ها است که زیرفضاهای به‌دست‌آمده را با یکدیگر تطبیق می‌دهد. در پایگاه‌داده اعداد، SRSA دارای ۸/۷۳ متوسط بهبود دقت نسبت به SA و در پایگاه‌داده آفیس و کالتک، دارای ۱/۸۷ متوسط بهبود دقت است. دلیل بهبود قابل‌ملاحظه SRSA نسبت به SA این است که SRSA از نمایش تُنُک برای انتخاب نمونه‌های منبع استفاده می‌کند.

SDA یک روش تطبیق دامنه بدون‌نظارت است که شامل تطبیق توزیع به تطبیق زیرفضا است که علاوه‌بر کاهش اختلاف بین دو زیرفضا، اختلاف داده را نیز کم می‌کند. در پایگاه‌داده اعداد، SRSA دارای ۱۱/۲ متوسط بهبود دقت نسبت به SDA و در پایگاه‌داده آفیس و کالتک، دارای ۲/۰۷ متوسط بهبود دقت و عمده‌ترین دلیل برتری SRSA نسبت به SDA، علاوه‌بر استفاده از تطبیق زیرفضا، استفاده از نمایش تُنُک برای انتخاب نمونه‌ها است.

(جدول-۵): دقت (%) طبقه‌بند در پایگاه‌داده آفیس و کالتک

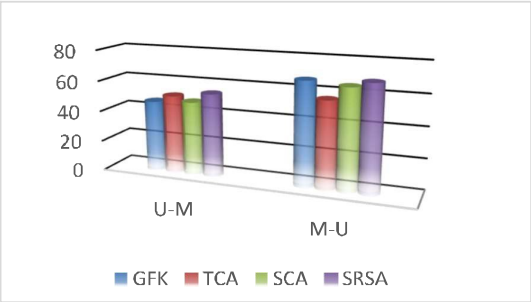
(Table-5): Classification accuracy on Office & Caltech dataset

SRSA	SCA (2016)	SDA (2015)	SA (2013)	GFK (2012)	DAM (2012)	TCA (2011)	PCA (2002)	NN	Dataset
44.57	45.62	**49.69**	49.27	41.02	42.69	45.82	36.95	23.7	C_A
39.32	**40**	38.98	**40**	40.68	34.58	30.51	32.54	25.76	C_W
52.87	47.13	40.13	39.49	38.85	33.12	35.67	38.22	25.48	C_D
38.91	39.72	39.54	39.98	**40.25**	35.35	40.07	34.73	26	A_C
40.68	34.92	30.85	32.22	38.98	31.86	35.25	35.59	29.83	A_W
37.58	**39.49**	33.76	33.76	36.31	36.31	34.39	27.39	25.48	A_D
30.37	31.08	34.73	**35.17**	30.72	33.84	29.92	26.36	19.86	W_C
32.15	29.96	**39.25**	**39.25**	29.75	37.58	28.81	29.35	22.96	W_A
90.45	87.26	75.80	75.16	80.89	80.89	85.99	77.07	59.24	W_D
32.15	30.72	**35.89**	34.55	30.28	32.41	32.06	29.65	26.27	D_C
31.21	31.63	38.73	**39.87**	32.05	34.34	31.42	32.05	28.5	D_A
88.81	84.41	76.95	76.95	75.59	77.63	86.44	75.93	63.39	D_W
46.59	45.16	44.52	44.72	42.95	42.55	43.03	39.65	31.37	متوسط دقت

[DOI: 10.29252/jsdp.17.2.58]

SCA یک روش تطبیق و تعمیم دامنه‌ها است که اختلاف بین دامنه‌ها را کمینه، اختلاف بین نمونه‌ها در رده‌های مختلف را بیشینه و اختلاف بین نمونه‌های هر رده را کمینه می‌سازد. در پایگاه‌داده اعداد، SRSA دارای ۴/۹ متوسط بهبود دقت نسبت به SCA و در پایگاه‌داده آفیس و کالتک، دارای ۱/۴۳ متوسط بهبود دقت و عمده‌ترین دلیل برتری SRSA نسبت به SCA استفاده از تطبیق زیرفضا برای کاهش اختلاف بین دامنه‌ها است.

نمودارهای مربوط به مقایسه دقت طبقه‌بند بر روی ۱۴ پایگاه‌داده با روش‌های GFK، TCA و SCA در شکل‌های (۵ و ۶) نشان داده شده است.

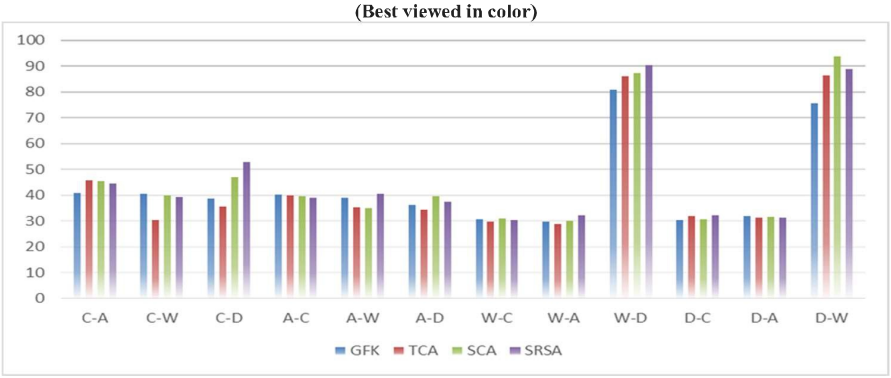


(شکل-۵): مقایسه دقت طبقه‌بندی بر روی پایگاه داده اعداد با استفاده از روش‌های GFK، TCA، SCA و SRSA (نمایش بهتر به‌صورت رنگی)

(Figure-5): The comparison of the classification accuracy on Digit dataset using GFK, TCA, SCA and SRSA methods (Best viewed in color)

(شکل-۶): مقایسه دقت طبقه‌بندی بر روی پایگاه داده آفیس و کالتک با استفاده از روش‌های GFK، TCA، SCA و SRSA (نمایش بهتر به‌صورت رنگی)

(Figure-6): The comparison of the classification accuracy on Office & Caltech dataset using GFK, TCA, SCA, and SRSA (Best viewed in color)



(جدول-۶): دقت (٪) طبقه‌بند در پایگاه‌داده اعداد

(Table-6): The classification accuracy on Digit dataset

SRSA	SCA (2016)	SDA (2015)	SA (2013)	GFK (2012)	DAM (2012)	TCA (2011)	PCA (2002)	NN	Dataset
54.35	48	35.70	41.50	46.45	42.69	51.05	44.95	44.7	U_M
68.56	65.11	65	63.95	67.22	52.83	56.28	66.22	65.94	M_U
61.46	56.56	50.35	52.73	56.84	47.76	53.67	55.59	55.32	متوسط دقت

۶- نتیجه‌گیری

در این مقاله یک روش بدون‌نظارت جهت حل مسأله شیفت در دامنه‌ها، با عنوان طبقه‌بندی تصاویر با استفاده از نمایش تُنُک و تطبیق زیرفضا برای تطبیق دامنه‌های بصری ارائه شد. SRSA یک روش نوین است که در دو مرحله تلاش می‌کند اختلاف توزیع بین داده‌های آموزشی و آزمایش را کاهش دهد. درواقع SRSA مدلی ایجاد می‌کند که با به‌کمینه‌رساندن ناسازگاری داده‌ها، مقاومت و تحمل‌پذیری بیشتری در برابر تغییرات توزیع داده‌ها بین دامنه‌های منبع و هدف داشته باشد. SRSA در گام نخست با استفاده از نرم $l_{2,1}$ نقش نمونه‌هایی از داده‌های منبع را که عامل ایجاد اختلاف بین دو دامنه هستند، کم‌رنگ‌تر می‌کند؛ علاوه‌بر این، SRSA با تطبیق زیرفضاهای ایجادشده اختلاف توزیع بین زیرفضاها را به کمینه می‌رساند. برای ارزیابی عملکرد SRSA، آزمایش‌های مختلفی بر روی انواع پایگاه‌داده‌های واقعی ترتیب داده شده و نتایج ارزیابی‌ها نشان از برتری قابل‌ملاحظه SRSA در مقایسه با روش‌های تطبیق دامنه بدون‌نظارت در این حوزه است. برای ادامه کار از نظریه منیفلدها[1] جهت برچسب‌گذاری بهتر نمونه‌های هدف بهره خواهیم برد.

[1] Manifolds

[DOI: 10.29252/jsdp.17.2.58]

7- References ۷- مراجع

[1] Y. Zhang, T. Liu, M. Long, and M. I. Jordan, "Bridging Theory and Algorithm for Domain Adaptation", arXiv preprint arXiv:1904.05801, 2019.

[2] S. Rezaei and J. Tahmoresnezhad, "Discriminative and domain invariant subspace alignment for visual tasks", *Iran Journal of Computer Science*, pp. 1-12, 2019.

[3] W. Kumagai and T. Kanamori, "Risk bound of transfer learning using parametric feature mapping and its application to sparse coding", *Machine Learning*, pp. 1-34, 2019.

[4] I. Jolliffe "Principal component analysis", Wiley, vol. 2, pp. 433-459, 2002.

[5] M. Jing, J. Li, J. Zhao, and K. Lu, "Learning distribution-matched landmarks for unsupervised domain adaptation", *In International conference on database systems for advanced applications*, pp. 491-508, 2018.

[6] A. Li, D. Chen, Z. Wu, G. Sun, and K. Lin, "Self-supervised sparse coding scheme for image classification based on low rank representation", *PloS one*, Vol. 13(6), e0199141, 2018.

[7] A. Mirjalili, V. Abootalebi, M. T. Sadeghi, "mproving the performance of sparse representation-based classifier for EEG classification," *JSDP* , pp. 43-55, 2015.

[8] X. Li, M. Fang, J. J. Zhang and J. Wu, "Sample selection for visual domain adaptation via sparse coding", *Signal Processing: Image Communication*, vol. 44, pp. 92-100, 2016.

[9] B. Gong, Y. Shi, F. Sha and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2066-2073, 2012.

[10] S. J. Pan, I. W. Tsang, J. T. Kwok and Q. Yang, "Domain adaptation via transfer component analysis", *IEEE Trans. Neural Netw*, vol. 22, no. 2, pp. 199–210, 2011.

[11] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, "Unsupervised visual domain adaptation using subspace alignment", in Proc. *IEEE International Conference on Computer vision*, pp. 2960-2967, 2013.

[12] M. Ghifary, D. Balduzzi, W. B. Kleijn, and M. Zhang, "Scatter component analysis: A unified framework for domain adaptation and domain generalizayion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-1, 2016.

[13] B. Sun and K. Saenko, "Subspace distribution alignment for unsupervised domain adaptation", in Proc. *British Machine Vision Conference*, 2015.

[14] J. Tahmoresnezhad and S. Hashemi, "Visual domain adaptation via transfer feature learning", *Knowledge and Information Systems*, vol. 50, no. 2, pp. 585-605, 2017.

[15] X. Li, M. Fang, J. J. Zhang and J. Wu, "Sample selection for visual domain adaptation via sparse coding", *Signal Processing: Image Communication*, vol. 44, pp. 92-100, 2016.

[16] K. Saenko, B. Kulis, M. Fritz and T. Darrell, "Adapting visual category models to new domains", *Proceedings of the European Conference on Computer Vision*, pp. 213-226, 2010.

[17] G.Griffin, A. Holub and P. Perona, "Caltech-256 object category dataset", Technical Report7694, 2007.

[18] J. J. Hull, "A", IEEE Trans. Pattern Anal. Mach. Intell, vol. 16, no. 5, pp. 550–554, 1994.

[19] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proc. *IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[20] L. Duan, D. Xu, I.W. Tsang, "Domain adaptation from multiple sources: a domain-dependent regularization approach", *IEEE* Trans. Neural Netw. Learn. Syst, vol. 23, no. 3, pp. 504–518, 2012.

فریماه شرافتی مدرک کارشناسی خود را در رشته مهندسی فناوری اطلاعات از دانشگاه ارسنجان دریافت کردند. ایشان با ادامه تحصیل موفق به اخذ مدرک کارشناسی ارشد در رشته مهندسی فناوری اطلاعات از دانشگاه صنعتی ارومیه شده‌اند. علایق پژوهشی ایشان شامل حوزه‌های یادگیری ماشین، یادگیری انتقالی است.

نشانی رایانامه ایشان عبارت است از:

farimah.sherafati@it.uut.ac.ir

جعفر طهمورث‌نژاد مدرک دکترای مهندسی کامپیوتر-هوش مصنوعی خود را در سال ۱۳۹۴ از دانشگاه شیراز دریافت کردند. ایشان در راستای فعالیت‌های علمی خود، در حال حاضر به‌عنوان استادیار دانشگاه صنعتی ارومیه درحال فعالیت هستند. علایق پژوهشی ایشان شامل حوزه‌های یادگیری ماشین، یادگیری انتقالی، داده‌کاوی و سیستم‌های امنیتی است.

نشانی رایانامه ایشان عبارت است از:

j.tahmores@it.uut.ac.ir

[DOI: 10.29252/jsdp.17.2.58]