Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
Volume 8, Number 4, Winter 2019, pp. 1-16
Received: 01/28/2018, Revised: 05/27/2019, Accepted:06/11/2019

# A Supervised Method for Building A Regularized Map for General Multi-View Multi-Manifold Learning

Faraein Aeini [1], Amir Masoud Eftekhari Moghadam [2*] and Fariborz Mahmoudi [3]
1- Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
2*- Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran.
3- Data scientist Advanced Analytics Department, General Motors, Warren, MI, USA.
[1] faraein.aeini@qiau.ac.ir, [2*]eftekhari@qiau.ac.ir, and [3] fariborz.mahmoudi@gm.com

Corresponding author's address: Amir Masoud Eftekhari Moghadam, Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran.

**Abstract-** In this paper, we consider the issue of automatic and unsupervised class-manifold selection in a multi-view multi-manifold space. General multi-manifold learning methods achieve multiple independent manifolds, so it is challenging for them to adjust the intra-class local manifold information and global inter-class discriminative structure. In this paper, we propose a multi-manifold embedding method, which can explicitly obtain multi-view multi-manifold structure while considering both intra-class compactness and inter-class separability without using the class label information. Furthermore, to the generalization of embedding to novel points, known as the out-of-sample extension problem in multi-view multi-manifold learning, we propose a supervised method for building a regularized map that provides an out-of-sample extension for general multi-view multi-manifold learning studied in the context of classification. Experimental results on face and object images demonstrate the potential of the proposed method for the classification of multi-view multi-manifold data sets and the proposed out-of-sample extension algorithm for the classification of manifold-modeled data sets.

**Keywords**- Multi-view multi-manifold learning, Dimensionality reduction, Out-of-sample extensions, Multi-view classification.

مجله علمی پژوهشی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
جلد ۸، شماره ۴، زمستان ۱۳۹۸ ، صفحه ۱۶-۱
دریافت: ۱۳۹۶/۱۱/۰۸، بازنگری: ۱۳۹۸/۰۳/۰۶، پذیرش: ۱۳۹۸/۰۳/۲۱

# ارائه نگاشت صریح و تنظیم‌شده‌ی باناظر برای یادگیری مالتی‌منیفولد داده‌های چندمنظری بدون برچسب

فرایین آیینی [۱] ، امیر مسعود افتخاری مقدم*[۲]، فریبرز محمودی
۱- دانشکده مهندسی کامپیوتر و فناوری اطلاعات، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.
*۲- دانشکده مهندسی کامپیوتر و فناوری اطلاعات، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.
۳- دپارتمان تجزیه و تحلیل پیشرفته داده‌های علمی، جنرال موتورز، وارن، ایالات متحده آمریکا.
[1] faraein.aeini@qiau.ac.ir, [2*]eftekhari@qiau.ac.ir, and [3] fariborz.mahmoudi@gm.com
* نشانی نویسنده مسئول: امیر مسعود افتخاری مقدم، قزوین، دانشگاه آزاد اسلامی واحد قزوین، دانشکده مهندسی کامپیوتر و فناوری اطلاعات.

**چکیده- در این مقاله، به مسئله‌ی انتخاب خودکار و بدون ناظر منیفولد طبقه در فضای مالتی‌منیفولد چندمنظری می‌پردازیم. مسئله‌ی طبقه‌بندی تصاویر چندمنظری برای یافتن منیفولد طبقه را می‌توان به‌عنوان مسئله‌ی یادگیری چندین منیفولد با تعدادی اشتراک بین منیفولدها در نظر گرفت. در حالت کلی مسئله‌ی یادگیری مالتی‌منیفولد با چندین زیر فضای مستقل کار می‌کند، بنابراین ایجاد تعادل میان اطلاعات درون منیفولد طبقه و ساختار متمایزکننده‌ی بین طبقه‌ها مشکل است. در این مقاله، روشی پیشنهاد می‌دهیم که بدون استفاده از اطلاعات برچسب نقاط داده با توجه به فشردگی درون طبقه‌ای و تفکیک‌پذیری برون طبقه‌ای، ساختار مالتی‌منیفولد چندمنظری را به‌دست می‌آورد. به‌علاوه، برای تعمیم تعبیه برای نقاط جدید که به‌عنوان مشکل «خارج از نمونه» شناخته می‌شود، نگاشت صریح و تنظیم‌شده‌ی باناظر برای کاهش بُعد غیرخطی ارائه کردیم که «توسعه خارج از نمونه» را برای یادگیری مالتی‌منیفولد چندمنظری در زمینه‌ی طبقه‌بندی انجام می‌دهد. نتایج آزمایش‌ها در دو دسته‌ی مدل‌سازی خودکار و بدون ناظر ساختار گراف مالتی‌منیفولد چندمنظری و نرخ بازشناسی بر روی چندین مجموعه داده‌ی چندمنظری برتری روش پیشنهادی را نسبت به دیگر روش‌های گزارش‌شده در مطالعات اخیر نشان می‌دهد.**

**واژه‌های کلیدی:** یادگیری مالتی‌منیفولد چندمنظری، کاهش بُعد غیر خطی، توسعه‌ی خارج از نمونه، طبقه‌بندی تصاویر چندمنظری.

## ۱- مقدمه

طی سال‌های اخیر، یادگیری منیفولد[۱] به‌صورت گسترده در حوزه‌ی یادگیری ماشین و بینایی کامپیوتر مورد مطالعه قرار گرفته است. بسیاری از داده‌ها در برنامه‌های کاربردی دنیای واقعی[۲] مانند چهره‌ها، ارقام دست‌نوشته و داده‌های مربوط به راه رفتن افراد می‌توانند بر روی منیفولد غیرخطی تعبیه‌شده[۳] درون فضای با ابعاد بالاتر قرار گیرند. تعبیه‌ی منیفولد با بُعد کم را می‌توان با کمک ساختار هندسی محلی منیفولد کشف کرد. اکثر الگوریتم‌های یادگیری منیفولد موجود، به‌عنوان پیش‌پردازش برای تعیین ساختار محلی، یک گراف همسایگی[۴] ایجاد می‌کنند. واضح است اگر گراف نتواند به‌طور صحیح ساختار را توصیف کند، نتیجه نهایی با نتیجه‌ی صحیح تفاوت قابل‌توجهی خواهد داشت. تعیین ساختار منیفولد بر مبنای روابط همسایگی اشیا، منجر به مسائل چالش‌برانگیز در یادگیری می‌شود. یکی از این مسائل مهم، یادگیری مالتی‌منیفولد است که به خاطر محدودیت کلی روش‌های یادگیری منیفولد برای تعبیه‌ی داده‌هایی که بر روی چندین منیفولد قرار دارند، ایجاد می‌شود. این مسئله از فرض اصلی روش‌های یادگیری منیفولد حاصل می‌شود که بیان می‌کند داده با چندین طبقه بر روی یک

منیفولد قرار می‌گیرد؛ اما در برنامه‌های کاربردی دنیای واقعی، داده‌های هر طبقه بر روی یک منیفولد مستقل قرار دارد. یادگیری مالتی‌منیفولد هنگامی‌که اشتراک‌هایی میان منیفولدها وجود دارد پیچیده‌تر می‌شود، مثل وظیفه‌ی بازشناسی تصاویر چندمنظری[۵] [۱]. به‌طورکلی، یادگیری چندمنظری یکی از مسائل یادگیری ماشین است که در آن داده توسط چندین مجموعه ویژگی مجزا نمایش داده می‌شود [۲, ۳]. در این مقاله، به نمایش‌های ناهمگون از یک شیء تمرکز نشده؛ بلکه، به چگونگی استفاده از اطلاعات متمایزکننده برای بهبود بازشناسی تصاویر چندمنظری می‌پردازیم که نمایش‌های همگن از یک شیء با زوایای دید مختلف وجود دارد. در یادگیری چندمنظری که مشکلات مربوط به وجود تصاویر متفاوت از یک شیء در زوایای دید متفاوت است، یک طبقه[۶] از تصاویر معمولاً شامل تصاویر با زوایای دید مختلف است که ممکن است شباهت کمی داشته باشند. از طرف دیگر، تصاویری در طبقه‌های دیگر با همان زاویه دید می‌توانند بسیار شبیه باشند؛ بنابراین، روش‌های یادگیری مبتنی بر گراف تحت تأثیر مشکل ناسازگاری[۷] طبقه‌ها قرار می‌گیرند.

یادگیری چندمنظری یکی از موضوعات مهم برای کاربردهایی مثل بازشناسی چهره[۸] و تشخیص عمل[۹] است. به‌عنوان مثال در بازشناسی چهره، همان‌طور که در شکل ۱ نشان داده شده است، فاصله بین دو چهره از روبه‌روی دو فرد مختلف کمتر از فاصله‌ی بین تصویر از روبه‌رو و تصویر از پهلوی یک فرد است. توانایی انطباق چندین زاویه دید از یک شیء[۱۰] برای یادگیری چندمنظری ضروری است. در یادگیری چندمنظری، انتخاب همسایه‌های صحیح به ایجاد یک گراف همسایگی قابل اعتماد منجر می‌شود.

تفاوت اصلی میان الگوریتم‌های یادگیری منیفولد معمول مانند Isomap [8]، LLE [۷] و LE [۵] و سایر روش‌های کاهش بُعد غیرخطی مانند Kernel PCA [۴] و SPCA [۵]، ویژگی حفظ‌کنندگی ساختار در نگاشت الگوریتم‌های یادگیری منیفولد است.

(الف) (ب) (ج)

**شکل ۱ : سه تصویر انتخاب شده از پایگاه داده چهره CAS-PEAL [۶]. الف) تصویر از رو به روی یک فرد که از او خواسته شده به پایین نگاه کند. ب) و ج) تصویر از روبه‌رو و از پهلوی فرد دیگر. اگر فاصله‌ی اقلیدسی بین تصاویر را اندازه‌گیری کنیم، تصویر (الف) و (ب) شبیه‌تر از (ب) و (ج) است.**

این هدف، با فرض اینکه نقاط داده نزدیک به‌هم در فضای با ابعاد بالا، در فضای کم بُعد نیز باید در کنار هم قرار بگیرند، انجام می‌شود. اینکه تنها معیار برای تعیین نگاشت حفظ‌کننده‌ی ساختار در روش‌های یادگیری منیفولد معمول، حفظ نقاط نزدیک به‌هم است، یک اشکال مهم ایجاد می‌کند. در این حالت هر دو نقطه که بردار ویژگی مشابه دارند، یعنی فاصله بین آن‌ها کوچک است، حتی اگر متعلق به دو طبقه مختلف باشند، باز هم به‌عنوان نقاط نزدیک در نظر گرفته می‌شوند؛ به‌عبارت‌دیگر این تفکر، معنای هندسی ساختار مالتی‌منیفولد را حذف می‌کند. برای مقابله با این مشکل، الگوریتم‌های انتخاب همسایگی باناظر مانند S-Isomap [۷]، S-LE [۸]، SNG [۹] و S-LLE [۱۰] از اطلاعات برچسب طبقه برای هدایت نگاشت داده‌ی اصلی در فضای تعبیه‌شده استفاده کردند. بیشتر روش‌های یادگیری مالتی‌منیفولد مانند M-Isomap [۱۱]، MM-LLE [۱۲]، MSGE [۱]، SH-NGC [۱۳] و فرآیند پیشنهادشده در [۱۴] از فرم با ناظر انتخاب همسایگی استفاده می‌کنند. سؤال مهم این است: «در فقدان دانش پیشین، چگونه باید همسایگان هم‌طبقه را به‌دست آوریم؟» در این مقاله، رابطه‌ی میان اشیا را بدون استفاده از اطلاعات پیشین برچسب‌های طبقه کشف می‌کنیم. هدف در این مقاله حفظ فاصله پیمایشی[۱۱] میان اشیا در فرآیند انتخاب همسایگان است. بدون ناظر بودن روش پیشنهادی اهمیت زیادی در وظیفه‌ی انتخاب همسایگان دارد.

در این مقاله، روشی بدون ناظر برای کشف خودکار منیفولد هر شیء در بازشناسی مالتی‌منیفولد چندمنظری پیشنهاد می‌دهیم. در روش پیشنهادی تصاویر مربوط به هر شیء را به‌عنوان یک منیفولد-طبقه نسبتاً مستقل تعریف می‌کنیم. راهکار پیشنهادی، برخلاف روش‌های یادگیری مالتی‌منیفولد معمول، مالتی‌منیفولد را با وجود اشتراکاتی میان منیفولدها در یک چارچوب کاری تعبیه گراف بدون ناظر یاد می‌گیرد. برای تعبیه گراف داده چندمنظری، روشی برای ساخت گراف مطلوب است که هم ساختار مالتی‌منیفولد چندمنظری و هم اطلاعات متمایزکننده را منعکس کند؛ بنابراین، ما روشی جدید و بدون ناظر طراحی کردیم که نه‌تنها به‌طور صریح ساختار منیفولد طبقه چندمنظری را حفظ کند، بلکه می‌تواند بین منیفولد محلی درون طبقه‌ای و اطلاعات متمایزکننده برون طبقه‌ای، بدون استفاده از اطلاعات برچسب طبقه، تعادل برقرار کند.

اکثر روش‌های یادگیری منیفولد تنها براساس داده‌های آموزشی موجود، تعبیه منیفولدها را محاسبه می‌کنند. در روش‌های یادگیری منیفولد تابع نگاشت ویژگی غیرخطی صریحی[۱۲] برای انتقال ویژگی‌ها از فضای اصلی به فضای با ابعاد کم‌تر تعیین نمی‌شود. این

انتقال باید با استفاده از روش‌های درون‌یابی غیرخطی مثل شبکه رگرسیون تعمیم یافته[۱۳] یاد گرفته شود. درحالی‌که تعمیم تعبیه برای داده‌ی جدید خارج از داده آموزشی که به آن مشکل توسعه«خارج از نمونه[۱۴]» گفته می‌شود، در کاربردهای بازشناسی مهم است.

هدف در روش پیشنهادی، طبقه‌بندی بدون ناظر تصاویر چندمنظری در منیفولدهای مستقل و سپس حل مسئله‌ی بازشناسی آن‌ها است. در روش پیشنهادی برای ایجاد گراف همسایگی هر طبقه، تصاویر از منظرهای مختلف مربوط یک شیء را به مکان‌های نزدیک هم و تصاویر اشیا دیگر را به مکان‌های دور از آن نگاشت می‌دهیم. سپس به هر منیفولد- طبقه‌ی شناسایی شده، شماره برچسب منحصربه‌فردی اختصاص می‌دهیم. بنابراین در مرحله‌ی بازشناسی می‌توان به صورت باناظر عمل کرد. برای بازشناسی تصویر آزمون، یک نگاشت کاهش بُعد غیرخطی باناظر و تنظیم‌شده[۱۵] صریح از داده‌های در فضای با ابعاد بالا به داده‌های در فضای کم بُعد ارائه می‌دهیم. در اینجا، نوآوری‌های مقاله را فهرست‌وار بیان می‌کنیم:

- ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری
- ارائه نگاشت صریح و تنظیم‌شده‌ی باناظر برای کاهش بُعد غیرخطی

ادامه‌ی مقاله به‌صورت زیر سازمان‌دهی شده است. در بخش ۲ کارهای پیشین موردبررسی قرار خواهد گرفت. در بخش سوم روش پیشنهادی با جزئیات بیان می‌شود. در بخش چهارم نتایج عملی در رابطه با روش پیشنهادی و نیز مقایسه آن‌ها با روش‌های پیشین مورد بحث قرار می‌گیرد و در بخش نهایی به نتیجه‌گیری پرداخته شده است.

## ۲- کارهای پیشین

تحت فرض مالتی‌منیفولد چندمنظری، ضروری است که ساختار مالتی‌منیفولد را با اشتراکات ممکن آن در یک گراف مناسب ترکیب کنیم، به‌طوری‌که نقاط داده‌ی غیر هم‌طبقه با منظرهای مشابه تا حد امکان از هم دور شوند و نقاط داده‌ی هم‌طبقه با منظرهای متفاوت تا حد امکان به‌هم نزدیک شوند.

برخی روش‌های تعبیه گراف مثل LDE [۱۵]، LFDA [۱۶]، LSDA [۱۷]، CGE [۱۸] و MSGE [۱] برای یادگیری زیر فضاهای متمایزکننده پیشنهادشده‌اند. در M-Isomap و MM-LLE از اطلاعان متمایزکننده‌ی برچسب طبقه برای یادگیری مالتی‌منیفولد استفاده شد. در روش‌ها بررسی شده به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف اشاره‌ای نشد. در جدول۱ به بررسی رویکرد مسئله و ضعف هر روش می‌پردازیم.

MSGE قطعه‌های منیفولد-مودال (MMF) را برای هر طبقه تعریف می‌کند. با در نظر گرفتن رابطه‌ی میان منیفولد-مودال و منیفولد-طبقه، ساختار گراف را برای مسئله‌ی طبقه‌بندی تصویر چندمنظری تعیین می‌کند. در ساخت گراف همسایگیMSGE سه مرحله‌ی اصلی وجود دارد. ۱. اتصال رئوس درون MMFها با استفاده از روش $k$-NN. ۲. اتصال زوج MMFهای درون یک طبقه برای ایجاد گراف متصل با محدودیت نزدیک‌ترین رئوس همسایه در زوج MMFهای مختلف. ۳. اتصال نزدیک‌ترین رئوس همسایه از طبقه‌های مختلف و مودالیتی مشابه تا زمانی که اتصالی بین هر زوج MMF مشابه بین طبقه‌های مختلف ایجاد شود. معیار (۱) به فشردگی درون ساختار MMF کمک می‌کند و ویژگی‌های محلی منیفولد را حفظ می‌کند. معیار (۲) به حفظ همنوع بودن بین MMFها کمک می‌کند و هر طبقه را فشرده کند. معیار (۳) به تفکیک ساختار هندسی منیفولدهای مودال و به‌دست آوردن زیر فضای افتراقی‌تر کمک می‌کند. .

هدف اصلی در روش‌های یادگیری منیفولد کشف ساختار تعبیه منیفولدها با ابعاد کمتر از نقاط داده‌ی با ابعاد بالا است، به‌طوری‌که روابط میان نقاط داده حفظ شود. معمولاً چنین تعبیه‌هایی برای مصورسازی استفاده می‌شود. این روش‌ها، تابع نگاشت صریحی بین فضای با ابعاد بالا و فضای با ابعاد پایین، که به آن توسعه «خارج از نمونه» گفته می‌شود، فراهم نمی‌کنند. بنابراین تحقیقاتی برای یافتن روش‌هایی به‌منظور توسعه یادگیری منیفولد برای مدیریت نمونه جدید آغاز شده است. برخی از محققان مشکل توسعه «خارج از نمونه» را به‌عنوان یک مشکل رگرسیون غیرخطی خالص در نظر گرفتند که مورد توجه محققان قرار نگرفت. بنجیو و همکاران [۱۹] از روشی مبتنی بر نمونه‌برداری Nystrom برای نقطه دلخواه استفاده کرد. ونگا و ژائو [۲۰] از شبکه رگرسیون تعمیم‌یافته برای یادگیری نگاشت استفاده کرد. یانگ و همکاران [۲۱] از شبکه‌های عصبی پس انتشار (BP) استفاده شد. این روش‌ها با پیدا کردن همسایه‌های نزدیک در مجموعه داده اصلی و درون‌یابی مختصات تعبیه‌شده، به محاسبه‌ی مختصات تعبیه‌ی نقاط جدید می‌پردازند. DSPP [۲۲] نگاشتی غیرخطی و صریح برای برون‌یابی نمونه‌های خارج از داده فراهم می‌کند. DSPP با استفاده از رابطه‌ی کلی رگرسیون تنظیم‌شده تلاش می‌کند تا نمایش کاهش بُعد ساخته‌شده با نقاط و برچسب‌های شناخته‌شده داده را حفظ کند و درعین‌حال پیچیدگی$f$ را در فضای محیطی و نیز همواری$f$ را در حفظ نمایش

داده تُنُک در نظر بگیرد. Reg-S3DR [۲۳] روش کاهش بُعد خطی و نیمه نظارتی را با استفاده از اطلاعات طبقه و برای ساخت نگاشت صریح کاهش بعد پیشنهاد می‌دهد. Reg-S3DR مشابه روش DSPP ، از مدل رگرسیون تنظیم‌شده استفاده می‌کند. برای اینکه جملات تنظیم‌کننده متمایزکننده‌تر باشند، تنظیم‌کننده با استفاده از ساختار هندسی داده منیفولد و نیز اطلاعات برچسب تعریف شد.

## ۳- روش پیشنهادی

در این بخش، ما به رویکرد انتخاب کارای منیفولد متناظر با هر طبقه می‌پردازیم. برای کشف منیفولد هر طبقه ما به انتخاب یکنواخت همسایگان هم‌طبقه از تمام زوایای دید موجود می‌پردازیم. ما ابتدا به بیان مسئله‌ی طبقه‌بندی بدون ناظر مالتی‌منیفولد چندمنظری می‌پردازیم. پس از کشف هر منیفولد طبقه، برچسب منحصربه‌فردی به آن اختصاص می‌دهیم. سپس روش تنظیم‌شده و با ناظر پیشنهادی را برای نگاشت صریح تصاویر در فضای با ابعاد بالا به نمایش کم بُعد ارائه می‌دهیم.

### ۳-۱- بیان مسئله و ایده‌ی ما

فرض کنید که n شیء داریم و هر شیء با m تصویر از زوایای دید مختلف نمایش داده شده است. این داده‌ها دو هدف متفاوت را نشان می‌دهند: طبقه‌ی متغیر و زاویه‌های دید متغیر. تمام تصاویر در هر طبقه به‌عنوان یک منیفولد-طبقه در نظر گرفته می‌شود. در برنامه‌های کاربردی مانند تشخیص چهره، هدف این است که تعبیه منیفولد فرد با تغییرات در زاویه دید یادگرفته شود. اگر همه‌ی تصاویر را در یک مجموعه قرار داده و سعی کنیم آن‌ها را با استفاده از روش یادگیری مالتی‌منیفولد تعبیه کنیم، قادر نخواهیم بود تعبیه معناداری را به‌دست آوریم؛ زیرا ممکن است فاصله بین دو فرد مجزا در برخی از زوایای دید کوچک‌تر از فاصله بین یک فرد در زوایای دید مختلف باشد؛ بنابراین، برخی از منیفولدهای افراد ممکن است ادغام شوند یا نزدیک به یکدیگر باشند که منجر به مخلوط شدن همسایگان از منیفولدهای مختلف می‌شود.

ما چندین مجموعه از تصاویر اشیا مختلف داریم که هر مجموعه از تصاویر روی یک منیفولد غیرخطی قرار می‌گیرد. ما قصد داریم یک نمایش تعبیه‌شده‌ی مفید از این داده‌ها را بیابیم. در یادگیری مالتی‌منیفولد چندمنظری، ما انتظار داریم یک توصیف کلی از منیفولد هر طبقه بیابیم. براساس اصل کلیدی در یادگیری منیفولد، هر نقطه روی منیفولد می‌تواند توسط ترکیب خطی همسایگان هم‌طبقه‌اش تقریب زده شود. متأسفانه، مجموعه

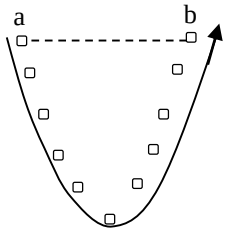
همسایگان ایده‌آل ناشناخته است و اگر ما چندین منیفولد که دارای اشتراک هستند در اختیار داشته باشیم، انتخاب یک همسایگی مناسب کار آسانی نیست. سه موضوع مهم در انتخاب همسایگان هم‌طبقه برای یادگیری مالتی‌منیفولد چندمنظری وجود دارد.

اول، انتخاب همسایگان بدون ناظر که در اکثر روش‌های یادگیری منیفولد بررسی شده است مثل LLE و Isomap مشکلاتی را در برخورد با داده‌های با چندین طبقه نشان می‌دهند. روش‌های یادگیری منیفولد باناظر پیشنهادشده، بر مشکلات روش‌های یادگیری منیفولد بدون ناظر موجود غلبه می‌کنند. بااین‌حال، برچسب‌گذاری وظیفه‌ای مشکل و هزینه‌بر است. ما داده‌های بدون برچسب زیادی از طبقه‌های مختلف و زوایای دید متفاوت داریم. روش‌های مالتی‌منیفولد که هسته‌ی آن‌ها معمولاً یک روش یادگیری منیفولد باناظر است، قادر به کار با این داده‌ها نیستند. در رویکرد ما برای حل این محدودیت، یافتن همسایگان معنایی پیشنهاد شد. برای این منظور به ارتباط میان منظرها در تصاویر هر فرد در طول چرخه‌ی دید تأکید می‌کنیم.

**جدول ۱: بررسی برخی از روش‌های پیشین یادگیری مالتی‌منیفولد**

| | باناظر/بدون ناظر | رویکرد حل مسئله | ضعف |
|---|---|---|---|
| LDE | بدون ناظر | ساخت گراف‌های درون طبقه‌ای و برون طبقه‌ای با استفاده از k-NN | ۱. عدم توجه به اطلاعات متمایز کننده‌ی سراسری<br>۲. عدم توجه به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف |
| CGE | بدون ناظر | ترکیب خطی ماتریس وزن محلی و ساخت ماتریس وزن برای به‌دست آوردن مدل زیرفضا | ۱. ساختار گراف ایجاد شدهاطلاعات متمایز کننده‌ی بین طبقه‌ای را نادیده می‌گیرد.<br>۲. عدم توجه به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف |
| LFDA | بدون ناظر | ایجاد گراف بین طبقه‌ای با اتصال کامل | ۱. تعداد اتصالات بین طبقه‌ای افزایش می‌یابد، نمی‌تواند به‌طور کارا ساختار هندسی محلی را حفظ کند.<br>۲. عدم توجه به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف |
| M-Isomap | باناظر | مبتنی بر ساخت گراف اصلی و گراف جریمه | ۱. تمرکز به افزایش فواصل محلی میان نقاط درون طبقه‌ای و عدم توجه به جدایی بین طبقه‌ای<br>۲. عدم توجه به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف |
| MM-LLE | باناظر | استفاده از SLLE در مرحله‌ی ساخت گراف همسایگی در LLE | ۱. عدم توجه به رابطه‌ی میان منظرهای مشابه در طبقه‌های مختلف |
| MSGE | باناظر | مبتنی بر ساخت قطعه‌های منیفولد-مودال و ایجاد ارتباط بین قطعه‌ها | ۱. استفاده از k-NN در ساخت قطعه منیفولد مودال معیار جامعی برای انتخاب تصاویر هم منظر نیست. |

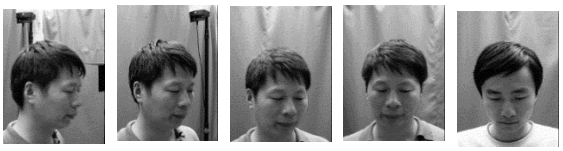
دوم، در حالت کلی گراف همسایگی براساس روابط محلی اشیا ساخته می‌شود؛ اما دو شیء که فاصله‌ی اقلیدسی کوچکی دارند ممکن است فاصله‌ی پیمایشی بزرگی داشته باشند. در این حالت اگر دو شیء را بر اساس فاصله‌ی اقلیدسی کوتاه میان آن‌ها به هم متصل کنیم، ممکن است توپولوژی سراسری میان نقاط را از دست بدهیم. این وضعیت معمولاً هنگامی رخ می‌دهد که نقاط روی منیفولد پیچ‌دار قرار دارند. این وضعیت در شکل ۲ نشان داده شد. شکل ۲ نشان می‌دهد که تفاوت زیادی بین فاصله‌ی پیمایشی و فاصله اقلیدسی برای ساخت گراف همسایگی برای داده‌ی روی منیفولد وجود دارد.




**شکل ۲ : بخشی از منیفولد یک بعدی. خط تیره‌ها فاصله‌ی اقلیدسی بین a و b را نشان می‌دهد. خط پررنگ فاصله‌ی پیمایشی بین a و b را نشان می‌دهد.**

سوم، در مجموعه داده‌های مالتی‌منیفولد چندمنظری ممکن است اشیا در برخی منظرها، همسایگان محلی از طبقه‌های مختلف داشته باشند. حال سؤالی که مطرح می‌شود این است: چگونه همسایگان غیر هم‌طبقه که از منظر مشابه‌ی تصویربرداری شدند را از میان همسایگان انتخاب شده برای یک شیء حذف کنیم؟ در راهکار ما برای حل این محدودیت، رابطه‌ی پنهان میان اشیا روی گراف‌های همسایگی در نظر گرفته شد. ایده‌ی پیشنهادی ما از این فرض ناشی می‌شود که شباهت میان اشیا خاصیت تعدی دارد؛ یعنی، هنگامی‌که دو شی توسط دنباله‌ای از اشیا واسط با فواصل کوچک به‌هم متصل شده‌اند ممکن است به‌عنوان اشیا مشابه در نظر گرفته شوند[۱۳, ۲۴]. برای توضیح بیشتر از شکل۳ استفاده می‌کنیم.

همان‌طور که در شکل۳ می‌توان دید، شکل۳ (الف) و (ب) تصویر چهره‌ی دو فرد مختلف را از یک منظر نشان می‌دهد و شکل ۳ (ب) تا (ه ) چهار تصویر چهره از منظر مختلف یک فرد را نشان می‌دهد. اگر فاصله‌ی بین تصویر (الف) و (ب)، و فاصله‌ی بین (ب) و (ه) را محاسبه کنیم، می‌توان دید که فاصله‌ی بین (الف) و (ب) کمتر از (ب) و(ه) است. اما از طرف دیگر، فاصله‌ی بین (ب) و (ج)، و نیز (ج) و (د)، و همچنین (د) و (ه) کمتر از فاصله‌ی (الف) و (ب) است. بر اساس این واقعیت که شباهت میان اشیا خاصیت تعدی دارد، می‌توان نتیجه گرفت که شکل۳ (ه) نسبت به تصویر (الف) کاندید بهتری برای همسایگی تصویر (ب) است.



(الف) (ب) (ج) (د) (ه)

**شکل ۳ : چهار نمونه از مجموعه داده‌ی CAS-PEAL [۶]**

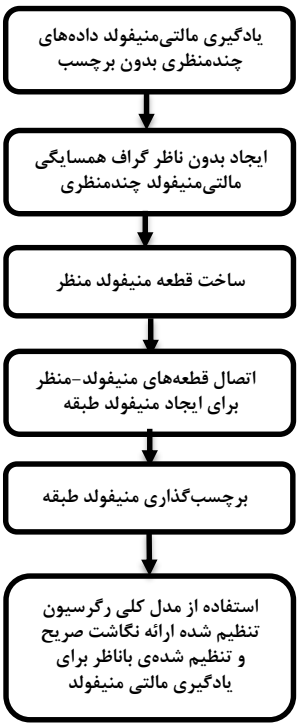
در اینجا ما «قطعه منیفولد-منظر» را به‌عنوان بخشی از منیفولد هر طبقه که دارای منظر مشابه هستند، تعریف می‌کنیم. با در نظر گرفتن رابطه‌ی بین قطعه‌های منیفولد-منظر، ساختار گراف مناسب را برای بازشناسی تصاویر چندمنظری ارائه می‌دهیم. با بیان رسمی‌تر، ما دو مرحله زیر را برای ساخت گراف متمایزکننده‌ی مالتی‌منیفولد داریم:

أ. اتصال $k$ شیء با منظر مشابه درون قطعه منیفولد-منظر با توجه به خاصیت تعدی در شباهت اشیا.

ب. اتصال جفت قطعه‌های منیفولد-منظر درون هر طبقه به‌عنوان گراف متصل با این محدودیت که نزدیک‌ترین جفت رئوس در قطعه‌های منیفولد-منظری که میانگین فاصله میان عناصر آن‌ها حداقل است، به‌هم متصل شوند.

معیار (أ) کمک می‌کند فشردگی ساختار هر قطعه منیفولد-منظر و در نتیجه ویژگی‌های ذاتی منیفولد محلی حفظ شود. معیار (ب) به حفظ هم‌طبقه بودن قطعه‌های منیفولد-منظر متصل شده و ایجاد یک طبقه فشرده کمک می‌کند.

پس‌ازاینکه توانستیم در رفتاری بدون ناظر مجموعه نقاط داده‌ی هر طبقه را در فضای مالتی‌منیفولد چندمنظری بیابیم، نیاز است نمونه داده‌هایی که در فضای با ابعاد بالا قرار دارند را به فضای با ابعاد کمتر نگاشت دهیم. این نگاشت باید ساختار میان نمونه داده‌ها را حفظ کند و نیز امکان تفکیک‌پذیری را در تعبیه یادگرفته شده افزایش دهد. اکثر روش‌های یادگیری منیفولد، تعبیه منیفولد را با توجه به نقاط داده آموزشی اولیه محاسبه می‌کنند؛ بنابراین تعمیم تعبیه برای نقاط جدید که به‌عنوان مشکل توسعه خارج از نمونه شناخته می‌شوند، در کاربردهای طبقه‌بندی و بازشناسی مورد توجه قرار گرفته است. در این مقاله از مدل رگرسیون تنظیم‌شده‌ی باناظر برای حل مسئله توسعه خارج از نمونه بهره برده‌ایم. شکل ۴ مراحل کلی حل مسئله را نشان می‌دهد.




**شکل ۴ : مراحل کلی الگوریتم پیشنهادی برای ارائه نگاشت صریح و تنظیم‌شده‌ی باناظر برای یادگیری مالتی‌منیفولد داده‌های چندمنظری بدون برچسب**

### ۳-۲- ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری

در این بخش از روش پیشنهادی، نه‌تنها قصد داریم نمونه‌های هم‌طبقه را بیابیم، بلکه قصد داریم بر مشکل نمونه‌های غیر هم‌طبقه با منظر مشابه نیز غلبه کنیم. جزئیات ساخت منیفولد هر طبقه به‌صورت زیر است.

برای ساخت iامین قطعه منیفولد-منظر با رفتاری بدون ناظر، از k نزدیک‌ترین همسایه‌ی شیء استفاده می‌کنیم. در روش پیشنهادی به‌جای انتخاب همسایگان اقلیدسی شیء از همسایگان معنایی آن استفاده می‌کنیم. استفاده از همسایگان اقلیدسی برای ساخت قطعه منیفولد-منظر دو مشکل را به همراه دارد.

اول، هنگامی‌که فاصله‌ی اقلیدسی بین دو شیء کم است، همیشه نمی‌توان نتیجه گرفت که فاصله‌ی پیمایشی بین آن‌ها روی منیفولد نیز کم است [۱۳].

دوم، انتخاب همسایگان تنها از ناحیه‌ی محلی اطراف یک شیء می‌تواند معنای هندسی توزیع کلی نقاط داده را نادیده بگیرد. چنین راهکاری تعبیه داده‌هایی مثل چهره که روی چندین منیفولد قرار می‌گیرند را نادیده می‌گیرد. هنگامی‌که دو منیفولد نزدیک به‌هم داریم، دو شیء غیر هم‌طبقه با منظر مشابه که روی دو منیفولد قرار دارند، ممکن است فاصله‌ی اقلیدسی کمی داشته باشند، اما فاصله‌ی پیمایشی آن‌ها کم نیست؛ بنابراین ما نیاز به تعریف جدیدی برای محاسبه‌ی فاصله‌ی میان نقاط داده روی چندین منیفولد داریم. در تعریف فاصله‌ی جدید، بدون استفاده از اطلاعات برچسب، باید بتوانیم نه‌تنها همسایگان محلی و هم‌طبقه، بلکه همسایگان هم‌طبقه‌ای که در ناحیه‌ی اطراف شیء قرار ندارند را نیز بیابیم. برای حفظ معنای هندسی میان کل نمونه داده‌ها توجه به همبستگی میان تمام نقاط داده پیشنهاد شد. هدف از روش پیشنهادی تمرکز روی معنای هندسی ساختار منیفولد در ساخت قطعه منیفولد-منظر در فضای مالتی‌منیفولد است. ایده‌ی اصلی در روش پیشنهادی از این واقعیت ناشی می‌شود که دو شی هنگامی‌که همسایگان مشترکی دارند، ممکن است به‌عنوان همسایه شناخته شوند.

برای یافتن $k$ نزدیک‌ترین همسایه‌ی معنایی شیء، ابتدا $k$ نزدیک‌ترین همسایه‌ی اقلیدسی شیء $x_i$ را با استفاده از روش $k$-NN انتخاب می‌کنیم. فهرست همسایگان انتخاب شده با $N(x_i)$ و فاصله‌ی اقلیدسی تا نمونه‌های کاندید با $D(x_i)$ نشان داده شد و به‌صورت نزولی مرتب شد. برای بهینه‌سازی همسایگان انتخاب شده، jامین عضو از $N(x_i)$ که با $x_{ij}$ نشان داده را انتخاب می‌کنیم، که $1 \le j \le k$. سپس فاصله‌ی بین $x_i$ تا همسایگان $x_{ij}$ که با $x_{ijp}$ نشان داده شده را محاسبه می‌کنیم، یعنی $D(x_i: x_{ijp}) = D(x_i: x_{ij}) + D(x_{ij}: x_{ijp})$. درصورتی‌که یکی از دو شرط زیر صحیح باشد فهرست همسایگی و فهرست فاصله را بروز رسانی و به‌صورت نزولی مرتب می‌کنیم.

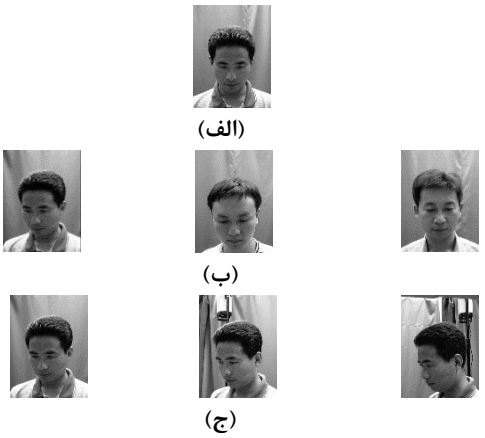
۱) آیا $D(x_{ij}: x_{ijp})$ کوچک‌تر از $\beta^{th}$-نزدیک‌ترین همسایه به $x_i$ در $N(x_i)$ است؟

۲) آیا $D(x_i: x_{ijp})$ کوچک‌تر از $D(x_i: x_{ijk})$ است که $k$ اندیس دورترین همسایه است؟

اگر حداقل یکی از شروط بالا برقرار باشد، دورترین همسایه در $N(x_i)$ حذف و $x_{ijp}$ در فهرست قرار می‌گیرد و فهرست به‌صورت نزولی مرتب خواهد شد. تصاویری که در یک قطعه منیفولد-منظر قرار گرفتند از ادامه‌ی روند انتخاب قطعه‌های منیفولد-منظر خارج می‌شوند. این کار به‌صورت سلسله مراتبی فضای جستجو را کاهش می‌دهد و منجر به کاهش زمان جستجو می‌شود. مقدار $\beta$ وابسته به مجموعه داده است. هنگامی‌که همسایگان هم‌طبقه به‌طور محلی در اطراف تصویر جستجو قرار ندارد، روش پیشنهادی مزیت خود را بیشتر نشان می‌دهد. با تنظیم مقدار $\beta$ می‌توان بین فاصله‌ی اقلیدسی مستقیم میان نمونه‌ها و فاصله‌ی پیمایشی با واسطه تعادل برقرار کنیم. با استفاده از روش پیشنهادی می‌توان علاوه بر

جستجوی محلی اطراف تصویر پرس‌وجو، میان نمونه داده‌هایی که در گراف به‌طور مستقیم به تصویر پرس‌وجو متصل نشده‌اند نیز کاوش کرد.

شکل ۵ مراحل انتخاب قطعه منیفولد-منظر براساس روش پیشنهادی را برای یکی از تصاویر در پایگاه داده‌ی CAS-PEAL نشان می‌دهد. شکل ۵(الف) تصویری است که تصمیم داریم همسایگان هم‌طبقه و با منظر مشابه آن را بیابیم. نتیجه‌ی انتخاب ۳-نزدیک‌ترین همسایه یک فرد نمونه در شکل ۵(ب) نشان داده شده است. شکل ۵(ج) نتیجه‌ی بهینه‌سازی همسایگان با روش پیشنهادی را نشان می‌دهد.



**(الف)**

**(ب)**

**(ج)**

**شکل ۵ : انتخاب قطعه منیفولد-منظر براساس روش پیشنهادی. الف) یک نمونه تصویر از مجموعه داده‌ی CAS-PEAL. ب) انتخاب ۳ نزدیکترین همسایه با روش *k*-NN ج) بهینه‌سازی همسایگان انتخاب شده بااستفاده از روش پیشنهادی**

پس از یافتن مجموعه قطعه‌های منیفولد-منظر به‌عنوان واحدهای مستقل اطلاعاتی، قصد داریم با اتصال قطعات منیفولد-منظر هم‌طبقه، منیفولد طبقه‌ی هر شخص که شامل تصاویر یک فرد از منظرهای مختلف است را بیابیم. یادآوری می‌کنیم که در فاز اول روش پیشنهادی اطلاعات برچسب در اختیار نیست.

مشکل در اتصال بدون ناظر تصاویر هم‌منظر این است که تصاویر هم‌منظر از طبقه‌های مختلف شباهت بیشتری نسبت به تصاویر هم‌طبقه با منظرهای مختلف دارند[۱]. اگرچه فاصله‌ی میان تصاویر هم‌منظر از طبقه‌های مختلف کم است اما فاصله‌ی میان تصاویر از طبقات مختلف با منظرهای غیرمشابه، حتی با اندکی تفاوت، زیاد است. با توجه به اینکه در مجموعه داده آزمون از هر منظر تصویر تعداد کمی تصویر وجود دارد، مشخص است که در یک منیفولد-منظر کشف شده تصاویر از نظر زاویه دید منظر اندکی متفاوت هستند. در روش پیشنهادی از این تفاوت برای اتصال قطعه‌های منیفولد-منظر به‌صورت یک گراف متصل استفاده کردیم. برای کم کردن تأثیر هم‌منظر بودن بر روی شباهت میان تصاویر از طبقه‌های مختلف، از میانگین فاصله‌ی میان عناصر زوج قطعات منیفولد-منظر استفاده شد. جزئیات روش پیشنهادی در شکل ۶ بیان شده است.

**الگوریتم ۱ اتصال قطعه‌های منیفولد-منظر برای ایجاد منیفولد طبقه**

**ورودی: منیفولد-منظر**

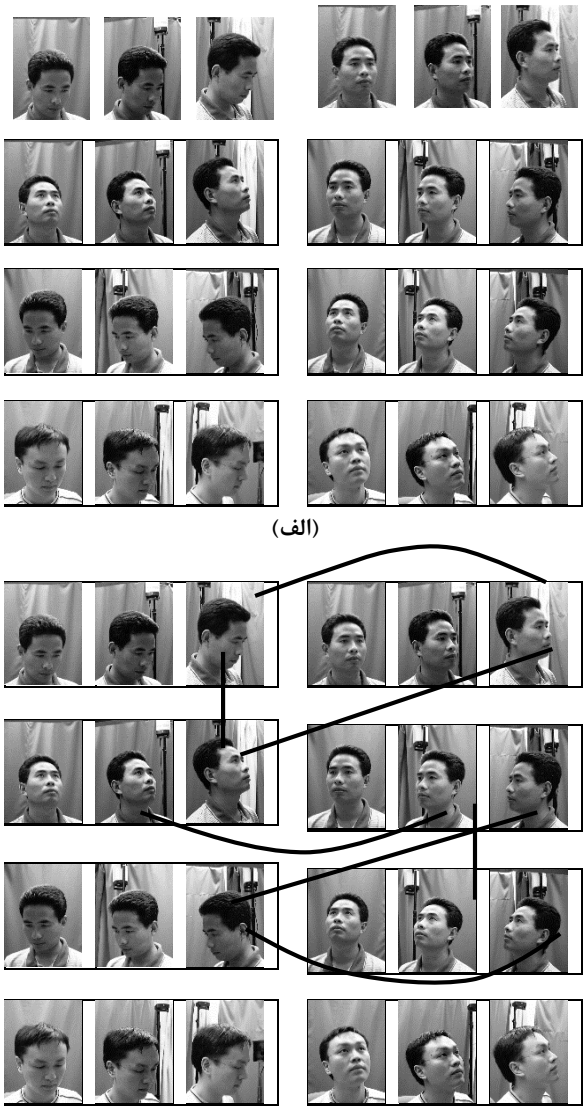
**خروجی : منیفولدهای طبقه**

۱. $k=1$
۲. **برای هر قطعه منیفولد منظر، میانگین فاصله میان عناصر زوج قطعات منیفولد منظر محاسبه می‌شود.**
۳. **$k$ اتصال بین قطعه منیفولد-منظر جاری و $k$ نزدیک‌ترین قطعات منیفولد-منظر همسایه‌ی متفاوت آن ایجاد می‌شود.**
۴. $k = k+1$
۵. **اگر گراف متصل حاصل نشد به شماره ۳ برو**
۶. **پایان**

**شکل ۶ : روش پیشنهادی برای ایجاد منیفولد-طبقه**

شکل ۷ نحوه‌ی اتصال قطعات منیفولد منظر را نشان می‌دهد. شکل ۷(الف) تعدادی از قطعه منیفولد-منظر متناظر با دو فرد را نمایش می‌دهد. شکل ۷(ب) اتصال قطعات منیفولد-منظر را با $k=2$ نشان می‌دهد. نکته مهم در روش پیشنهادی این است که حتماً بین هر دو قطعه منیفولد-منظر، به‌طور مستقیم یا غیرمستقیم، ارتباطی وجود دارد. به خاطر وجود خاصیت تعدی در شباهت میان قطعه‌های منیفولد-منظر، میانگین فاصله عناصر قطعه منیفولد-منظر هم‌طبقه مقدار کمتری نسبت به عناصر قطعه منیفولد-منظرهای غیر هم‌طبقه خواهند داشت؛ بنابراین می‌توان منیفولد طبقه را بدون داشتن اطلاعات برچسب طبقه به‌دست آورد. پس از یافتن دو قطعه منیفولد-منظر با کمترین میانگین فاصله میان عناصر آن‌ها، نزدیک‌ترین تصاویر آن‌ها را به‌هم متصل می‌کنیم. حال می‌توان به هر منیفولد طبقه یک شماره اختصاصی نسبت دهیم و مرحله‌ی ارائه‌ی نگاشت صریح و تنظیم‌شده برای کاهش بُعد غیرخطی را به‌صورت باناظر انجام دهیم.

پس از ایجاد گراف همسایگی برای تعیین منیفولدهای طبقه می‌توانیم، منیفولدها را برچسب‌گذاری کنیم. برای برچسب‌گذاری منیفولدهای هر طبقه از روش پیشنهادی در [۲۳] استفاده شد. در این شیوه از برچسب‌گذاری، برای هر منیفولد طبقه یک بردار تصادفی $d$-بُعدی که عناصر آن در بازه‌ی [۱و۰] است به‌عنوان برچسب در نظر می‌گیریم. در [۲۳] اثبات شد که احتمال اینکه بردارهای برچسب دور از هم قرار بگیرند بسیار بالا است. به تمام تصاویر انتخاب شده روی یک منیفولد، یک بردار برچسب مشترک اختصاص داده می‌شود.




**شکل ۷ : روش پیشنهادی برای ساخت منیفولد-طبقه. الف) تعدادی قطعه منیفولد-منظر استخراج شده با روش پیشنهادی از مجموعه داده CAS-PEAL. ب) نحوه‌ی اتصال قطعه‌های منیفولد-منظر.**

### ۳-۳- ارائه نگاشت صریح و تنظیم‌شده برای کاهش بُعد غیرخطی باناظر

در این بخش نگاشتی صریح و تنظیم‌شده برای کاهش بُعد غیرخطی باناظر پیشنهاد می‌دهیم. برای ارائه‌ی یک نگاشت کاهش بُعد با استفاده از مدل کلی رگرسیون تنظیم‌شده زیر را پیشنهاد می‌دهیم.

$$F^* = \arg min_F \{\frac{1}{l}\sum_{i=1}^{l} V(x_i, y_i, \mathrm{F}) + \gamma_k \|\mathrm{F}\|_k^2 + \gamma_I \|\mathrm{F}\|_I^2\} \qquad (۱)$$

که متغیر بهینه‌سازی نگاشت بردار $\mathrm{F}: R^N \to R^d$ است. $\gamma_k$ و $\gamma_I$ دو پارامتر تنظیم کننده است. پارامترهای تنظیم یکی از ورودی‌های مدل هستند که مقدار آن‌ها باید انتخاب شود. افزایش پارامتر تنظیم بیش برازش را کاهش می‌دهد. این کار با اضافه کردن بایاس به تخمین انجام می‌شود. روشی که در این مقاله برای انتخاب مقادیر پارامترهای تنظیم استفاده شد، انتخاب تصادفی زیرمجموعه‌ای از داده‌ها و بررسی واریانس تخمین است. سپس مقدار پارامترها را افزایش داده و تأثیر آن را بر تغییرات تخمین در نظر می‌گیریم. به یاد داشته باشید که مقدار پارامترها متناسب با زیرمجموعه‌ای است که انتخاب شده است. با احتمال بالا می‌توان از مقادیر کوچکتر برای رسیدن به مقادیر قابل مقایسه در مجموعه داده‌ی کامل استفاده کرد. در این مقاله برای پارامترهای تنظیم مقادیر مساوی ۰,۱ برای مجموعه داده‌ی COIL-20 و CAS-PEAL و مقدار ۰,۰۵ برای مجموعه داده‌ی FEI در نظر گرفته شد. $y_i \in R^d, i = 1, \dots, l$ نمایش کاهش بُعد یافته را نشان می‌دهد که در مرحله‌ی قبل اختصاص داده شد. $l$ تعداد تصاویر موجود در پایگاه داده است. ازآنجایی‌که به تمام نقاط داده‌ی روی یک منیفولد طبقه، یک بردار برچسب $d$ بُعدی نسبت داده شد، نمایش کاهش بُعد یافته‌ی نقاط هم‌طبقه تا حد ممکن به‌هم نزدیک می‌شود و ازآنجایی‌که بردار برچسب برای هر طبقه به‌صورت تصادفی انتخاب شد، نمایش کم بُعد از طبقه‌های مختلف تا حد ممکن از هم دور می‌شوند.

در تعمیم یک تعبیه، تمایل داریم مختصات داده‌های آموزشی یاد گرفته شده را حفظ کنیم، $\phi(x_i) = y_i$. بنابراین تابع زیان را به‌صورت $V(x_i, y_i, F) = \sum_{i=1}^{l} \|\phi(x_i) - y_i\|_2^2$ تعریف می‌کنیم. بنابراین نیاز به تابع هسته‌ای داریم که باعث شود بتوانیم داده اصلی با ابعاد D را با داده‌ی کاهش بُعد یافته‌ی $d$ قیاس کنیم. برای عرفی تابع هسته، از فرض اصلی در منیفولد که هر نقطه را می‌توان از تقریب خطی همسایگانش به‌دست آورد، استفاده کردیم. برای این هدف پیشنهاد می‌کنیم هر نقطه را با بردار ویژگی $d$ بُعدی که با استفاده از نقاط همسایه آن خواهیم ساخت، جایگزین کنیم. اگر $N(x_i)$ مجموعه‌ی $k$-نزدیک‌ترین همسایه‌ی هم‌طبقه‌ی $x_i$ را نشان دهد، بردار ویژگی یکه $x_i$، به‌صورت زیر پیشنهاد می‌شود:

$$\phi(x_i) = \left| \frac{x_i - x_j}{\|x_i - x_k\|} \right| \ , x_j \in N(x_i) \qquad (۲)$$

که $k$ دورترین همسایه‌ی $x_i$ است.

$\|\mathrm{F}\|_k^2$ معیاری برای پیچیدگی نگاشت F است و $\|\mathrm{F}\|_I^2$ توانایی F را در حفظ ساختار داده اندازه‌گیری می‌کند. براساس تئوری یادگیری آماری [۲۵, ۲۶]، تنظیم‌کننده‌ی $\|\mathrm{F}\|_k^2$ معمولاً به‌عنوان نُرم تابع در فضای هیلبرت با هسته بازآفرین (RKHS) تعریف می‌شود. برای یک هسته‌ی مثبت نیمه معین k(u,v)، یک فضای هیلبرت با هسته بازآفرین $H_k$ وجود دارد. هر تابع $f_s \in H_k$ می‌تواند به‌عنوان یک

ترکیب خطی از توابع هسته $f_s(.) = \sum_i \eta_i^s k(u_i,.)$ بیان کرد. برای نگاشت بردار $F = [f_1, \dots, f_d]^T$، $\|F\|_k^2$ را می‌توان به‌صورت زیر تعریف کرد:

$$\|F\|_k^2 = \sum_{s=1}^{d} \|f_s\|_k^2 = \sum_{s=1}^{d} \sum_{i,j} \eta_i^s \eta_j^s K(u_i, u_j) \qquad (۳)$$

برای متمایزکننده‌تر کردن $F$، تنظیم کننده‌ی $\|F\|_I^2$ باید با استفاده از ساختار هندسی داده منیفولد و نیز اطلاعات برچسب تعریف شود. برای این منظور، وزن بین نقاط داده به‌صورت زیر تعریف می‌شود:

$$w_{i,j} = \begin{cases} \kappa & \text{اگر } x_i \text{ و } x_j \text{ هم طبقه باشند} \\ -\kappa & \text{در غیر اینصورت} \end{cases} \qquad (۴)$$

برای تعریف وزن $w_{i,j}$ از دانش پیشین برچسب طبقه استفاده می‌کنیم و نشان می‌دهیم که آیا $x_i$ و $x_j$ هم‌طبقه هستند یا خیر. مقدار $\kappa$ باید نسبتاً بزرگ انتخاب شود. در آزمایش‌های ما، مقدار $\kappa$ به‌صورت تجربی چهار انتخاب شده است. نگاشت متمایزکننده باید به نحوی طراحی شود که اگر دو نقطه‌ی هم‌طبقه به نقاط دور از هم نگاشت داده شوند و یا نقاط غیر هم‌طبقه به نقاط نزدیک به‌هم نگاشت داده شوند تابع هدف جریمه سنگینی را متحمل شود. یک راه حل منطقی برای یافتن نگاشت متمایزکننده، بهینه‌سازی تابع هدف زیر است:

$$\min \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \left\| F(\phi(x_i)) - F(\phi(x_j)) \right\|^2 \qquad (۵)$$

براین اساس می‌توان تنظیم کننده‌ی تابع $\|F\|_I^2$ را به‌صورت زیر تعریف نمود:

$$\|F\|_I^2 = \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \left\| F(\phi(x_i)) - F(\phi(x_j)) \right\|^2 \qquad (۶)$$

با ادغام روابط بالا، نگاشت صریح پیشنهادی را می‌توان به‌صورت زیر بیان کرد:

$$F^* = \arg min_F \left\{ \frac{1}{l} \sum_{i=1}^{l} \|\phi(x_i) - y_i\|_2^2 + \gamma_k \|F\|_k^2 + \frac{\gamma_I}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \left\| F(\phi(x_i)) - F\left(\phi(x_j)\right) \right\|^2 \right\} \qquad (۷)$$

با استفاده از تئوری زیر می‌توان راه‌حلی صریح برای رابطه‌ی (۷) بیابیم:

تئوری۳-۱: [۲۷] حداقل کننده‌ی مسئله‌ی بهینه‌سازی (۷) به‌صورت زیر قابل توسعه است:

$$F^*(\emptyset) = \sum_{i=1}^{N} \alpha_i \; k(\emptyset_i, \emptyset) \qquad (۸)$$

که $\alpha_i = [\alpha_{1i}, \dots, \alpha_{di}]^T \in R^d$، $\emptyset$ بردار ویژگی ساخته شده با استفاده از فاصله‌ی پیمایشی و $k(.,.)$ تابع هسته است. با جایگزینی روابط (۲) و (۷) در رابطه‌ی (۶) خواهیم داشت :

$$A^* = \arg min_{A \in R^{d \times N}} \{ \frac{1}{l} tr((AKJ - Y)(AKL - Y)^T + \gamma_k tr(AKA^T) + \frac{\gamma_I}{2N^2} tr(AKLKA^T) \} \qquad (۹)$$

که $A = [\alpha_{1i}, \dots, \alpha_{di}] \in R^{d \times N}$ ماتریس ضرایب برای توسعه‌ی F، $Y = [y_1, \dots, y_l] \in R^{d \times N}$ ماتریس هدف، $K = \left(k(\emptyset_i, \emptyset_j)\right) \in R^{N \times N}$ ماتریس هسته مثبت معین، $L = S - W \in R^{N \times N}$ ماتریس گراف لاپلاسین را نشان می‌دهد. $W = w_{ij} \in R^{N \times N}$، S ماتریس قطری است که عناصر قطر اصلی $S_{ii} = \sum_{j=1}^{N} w_{ij}$ است، I ماتریس همانی، و $J \in R^{N \times N}$ ماتریس انتخاب قطری است l عنصر اول قطر یک و بقیه عناصر قطر صفر است. از تابع هدف (۹) مشتق گرفته و برابر با مقدار صفر قرار می‌دهیم، به‌صورت زیر

$$\frac{1}{l}(Y - AKJ)(KJ)^T + \gamma_k AK + \frac{\gamma_I}{N^2} AKLK = 0 \qquad (۱۰)$$

رابطه‌ی (۱۰) را به‌صورت زیر می‌توان بازنویسی کرد:

$$A\left(KJ - \gamma_k lI - \frac{l}{N^2}\gamma_I KL\right) = Y \qquad (۱۱)$$

در نتیجه، کمترین مربع رابطه‌ی (۱۱) به‌صورت زیر به‌دست می‌آید:

$$A^* = Y(KJ - \gamma_k lI - \frac{\gamma_I l}{N^2} KL) \qquad (۱۲)$$

بنابراین نگاشت کاهش بُعد موردنظر به‌صورت زیر به‌دست خواهد آمد:

$$Z = F^*(\emptyset) = \sum_{i=1}^{N} \alpha_i^* K(\emptyset_i, \emptyset) \qquad (۱۳)$$

که $\alpha_i^*$، iامین ستون از $A^*$ است. در رابطه‌ی بالا از هسته‌ی خطی برای نگاشت کاهش بُعدF استفاده شد، $k(u_i, u_j) = u_i^T u_j$. هر نقطه‌ی $x_i$، با بردار کم بُعدش $z_i = A\Phi^T \phi_i$ می‌توان نشان داد که $\Phi = [\phi_1, \dots, \phi_N]$ است.

## ۴- نتایج عملی

در این مقاله، دو گروه مختلف از آزمایش‌ها انجام شد. در گروه اول، نتایج ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری

برای مقایسه‌ی روش پیشنهادی و MSGE [1] ارائه شد. درگروه دوم، نتایج بازشناسی اشیا و نیز بازشناسی چهره برای مقایسه‌ی روش پیشنهادی با چندین روش یادگیری منیفولد که در تحقیقات اخیر به آن‌ها اشاره شده، مورد بررسی قرار گرفت. برای پیاده سازی روش پیشنهادی از جعبه‌ابزار کاهش بعد متلب DR Toolbox [۲۸] استفاده شد. در آزمایش‌ها برای نشان دادن مفهوم چندمنظری، سه پایگاه داده انتخاب شد که تصاویر آن‌ها از منظرهای مختلف گرفته‌شده‌اند. برای بررسی دقت بازشناسی از هر فرد به‌طور تصادفی سه تصویر را از منظرهای مختلف انتخاب کردیم و به‌عنوان تصاویر آزمایشی در نظر گرفتیم. باقی تصاویر به‌عنوان نمونه‌های آموزشی در نظر گرفته می‌شوند. این آزمایش‌ها چهار بار برای تصاویر آزمایشی تصادفی مختلف انجام می‌شود و نتایج میانگین گرفته می‌شود.

### ۴-۱- توصیف مجموعه داده‌ها

در آزمایش‌ها سه مجموعه داده انتخاب شد که برای نشان دادن مفهوم مالتی‌منیفولد چندمنظری، تصاویر آن‌ها از نظر زاویه دید تغییر می‌کند. در ادامه سه مجموعه داده استفاده شده توصیف می‌شود:

۱. *Columbia Object Image Library* (COIL-20): براساس مطالعات اخیر در رابطه با مفهوم یادگیری مالتی‌منیفولد [۲۹-۳۱] ما COIL-20 را به‌عنوان یک نمونه از مجموعه داده‌ی مالتی‌منیفولد و چندمنظری انتخاب کردیم. این مجموعه داده حاوی ۱۴۴۰ تصویر سطح خاکستری از ۲۰ شیء است. هر شیء درحالی‌که دوربین ثابت است، از طریق چرخش ۳۶۰ درجه در ۷۲ موقعیت تصویربرداری شد. نمونه‌هایی از این مجموعه داده در شکل ۸ نشان داده شده است.

**شکل ۸ : نمونه‌ای از تصاویر مجموعه داده‌ی COIL-20.**

۲. پایگاه داده‌ی چهره CAS-PEAL: پایگاه داده‌ی چهره CAS-PEAL، یک پایگاه داده‌ی بزرگ از افراد چینی است که تصاویر از نظر زاویه‌ی سر، حالت چهره، متعلقات و نورپردازی متفاوت‌اند [۶]. پایگاه داده‌ی چهره‌ی CAS-PEAL شامل ۹۹۵۹۴ تصویر از ۱۰۴۰ نفر است (۵۹۵ مرد و ۴۴۵ زن). برای هر فرد نُه دوربین در بازه‌های برابر در یک راستای نیم‌دایره‌ی افقی قرار داده شده است تا به‌طور هم‌زمان تصاویری از منظرهای مختلف گرفته شود. از هر فرد خواسته شده به بالا و پایین نگاه کند تا تعدادی تصویر از دو زاویه دیگر گرفته شود. ما به‌طور تصادفی ۵۰ طبقه از تصاویر را انتخاب و آزمایش‌ها را روی آن‌ها ارزیابی کردیم. برخی تصاویر از سه فرد تحت زوایای دید مختلف در شکل ۹ نشان داده شده است.

**شکل ۹ : نمونه‌هایی از تصاویر چهره پایگاه داده CAS-PEAL تحت تغییرات منظر [۶]**

۳. پایگاه داده‌ی چهره FEI: پایگاه داده‌ی چهره FEI، یک پایگاه داده‌ی چهره‌ی برزیلی است که شامل مجموعه‌ای از تصاویر افراد است که از نظر ظاهر، مدل مو و زاویه‌ی دید متفاوت‌اند [۳۲]. پایگاه داده شامل تصاویر ۲۰۰ نفر است و از هر فرد ۱۴ تصویر وجود دارد. تمام تصاویر رنگی هستند و با پس‌زمینه سفید از موقعیت چرخش صورت به سمت راست تا ۱۸۰ درجه چرخش گرفته‌شده‌اند. تعداد تصاویر مرد و زن برابر است. شکل ۱۰ نمونه‌هایی از تغییرات تصاویر در پایگاه داده‌ی FEI را نشان می‌دهد.

**شکل ۱۰ : نمونه‌هایی از تغییرات تصاویر پایگاه داده‌ی چهره FEI [۳۳]**

### ۴-۲- ایجاد بدون ناظر ساختار گراف مالتی‌منیفولد چندمنظری

در این بخش، نتایج انتخاب قطعه منیفولد-منظر نشان داده می‌شود. روش MSGE که روشی شبیه به روش پیشنهادی است برای مقایسه انتخاب شد. MSGE برای حفظ ساختار محلی، قطعه‌های MMF را استخراج می‌کند. هر MMF از اتصال هر نمونه داده به $k$-نزدیک‌ترین همسایه‌ی اقلیدسی آن ایجاد می‌شود. باوجوداینکه هدف از انتخاب MMFها انتخاب تصاویر هم منظر بود اما این اتفاق به‌درستی نخواهد افتاد؛ زیرا MMF مبتنی بر روش $k$-NN است. واضح است که $k$-NN در انتخاب همسایگی، تصاویر هم‌منظر و غیر هم‌طبقه را به تصاویر هم‌طبقه با اندکی تفاوت در منظر، ترجیح می‌دهد. بنابراین MMFها

حاوی تصاویری از طبقه‌های مختلف خواهد بود. در روش پیشنهادی قطعه منیفولد-منظر با توجه به اینکه شباهت میان تصاویر خاصیت تعدی دارد انتخاب می‌شود و علاوه بر شباهت میان تصویر آزمون و تصاویر همسایه، به شباهت میان تصاویر همسایه کاندید نیز توجه می‌شود. مشکل بعدی MSGE در اتصالMMF ها است. اتصال MMFها نیز براساس نزدیک‌ترین تصاویر به عناصر MMFها انجام می‌شود. واضح است ممکن است اتصالاتی بین MMFهای هم‌منظر و غیر هم‌طبقه اتفاق بیافتد. در روش پیشنهادی به‌جای تأثیر یک تصویر، میانگین تصاویر موجود در هر قطعه بر اتصال تأثیر می‌گذارند تا بتوانیم تأثیر تصاویر غیر هم‌طبقه و هم‌منظر را کاهش دهیم.

تعداد همسایگی در هر قطعه منیفولد-منظر نقش مهمی بر اینکه آیا گراف تشخیص داده شده برای فرد می‌تواند به‌طور دقیق ساختار اصلی منیفولد فرد را نشان دهد، دارد. موفقیت تمام روش‌های یادگیری منیفولد مبتنی بر گراف به‌شدت وابسته به تعداد مناسب همسایگی و روش پیشنهادی ما نیز وابسته به تعداد همسایگی در هر قطعه منیفولد-منظر است. بااین‌حال، انتخاب کارای این تعداد برای هر مجموعه داده، مسئله‌ای باز است. هنوز روشی دقیق برای کشف تعداد صحیح همسایگی برای مجموعه داده‌های مختلف وجود ندارد. ما برای بررسی کارایی روش پیشنهادی به بررسی تأثیر تعداد قطعه منیفولد-منظر بر دقت گراف همسایگی پرداختیم. جدول۲ دقت در ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری را با مقادیر مختلف k روی پایگاه داده‌های COIL-20 نشان می‌دهد.

**جدول ۲: دقت در ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری COIL-20 با مقادیر مختلف *k***

| k | روش پیشنهادی | | MMF | |
|---|---|---|---|---|
| | میانگین | واریانس | میانگین | واریانس |
| ۳ | ۱ | ۰ | ۰/۹۹ | ۰/۰۰۰۱۸ |
| ۶ | ۱ | ۰ | ۰/۹۶ | ۰/۰۰۰۲ |
| ۹ | ۱ | ۰ | ۰/۹۲ | ۰/۰۰۰۵ |

دقت ایجاد گراف همسایگی روی CAS-PEAL در جدول ۳ و دقت ایجاد گراف همسایگی روی FEI در جدول ۴ نشان داده شد. نتایج عملی نیز مباحث تئوری ارائه‌شده در بالا را تأیید می‌کند.

هنگامی‌که تعدادی تصویر هم منظر برای یک شیء وجود دارد مثل مجموعه داده‌های COIL-20، MMF می‌تواند نتایج بهتری نسبت به زمانی که از هر منظر تنها یک تصویر برای هر شیء وجود دارد مثل مجموعه داده‌ی CAS-PEAL به‌دست آورد. مجموعه داده‌ی COIL-20 برخلاف مجموعه داده‌های چهره حاوی تصاویر اشیا با هندسه متفاوت هستند. همان‌طور که از نتایج قابل‌مشاهده است، MMF در این حالت نتایج بهتری نسبت به مجموعه داده چهره به‌دست آورده است.

**جدول ۳: دقت در ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری CAS-PEAL با مقادیر مختلف *k***

| k | روش پیشنهادی | | MMF | |
|---|---|---|---|---|
| | میانگین | واریانس | میانگین | واریانس |
| ۳ | ۰/۹۹ | ۰/۰۰۰۰۱۷ | ۰/۲۹۹۶۲۲ | ۰/۰۰۳۲ |
| ۶ | ۰/۹۹ | ۰/۰۰۰۰۴ | ۰/۲۳۳۰۵۸ | ۰/۰۰۶۵ |
| ۹ | ۰/۹۸ | ۰/۰۰۰۸۹ | ۰/۲۰۲۳۸۵ | ۰/۰۹۵ |

**جدول ۴: دقت در ایجاد بدون ناظر گراف همسایگی مالتی‌منیفولد چندمنظری FEI با مقادیر مختلف *k***

| k | روش پیشنهادی | | MMF | |
|---|---|---|---|---|
| | میانگین | واریانس | میانگین | واریانس |
| ۳ | ۰/۹۵ | ۰/۰۰۰۰۰۱۶ | ۰/۵۹ | ۰/۰۰۰۰۱۴ |
| ۶ | ۰/۹۱ | ۰/۰۰۰۰۰۴۲ | ۰/۴۲ | ۰/۰۰۰۰۳۶ |
| ۹ | ۰/۸۷ | ۰/۰۰۰۰۹۳ | ۰/۳۳ | ۰/۰۰۰۸۱ |

واضح است که k-NN در انتخاب همسایگی، تصاویر هم‌منظر و غیر هم‌طبقه را به تصاویر هم‌طبقه با منظر متفاوت، ترجیح می‌دهد. بنابراینMMFها حاوی تصاویری از طبقه‌های مختلف خواهند بود.

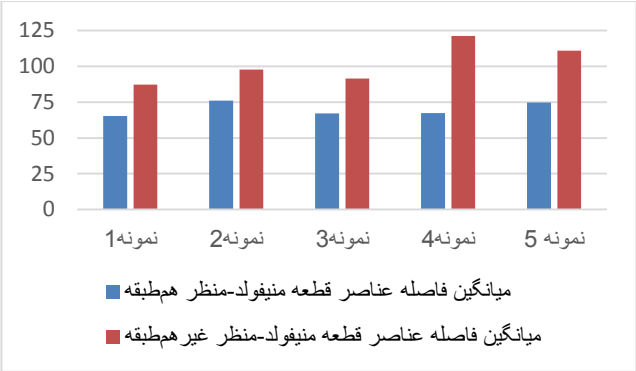
در روش پیشنهادی قطعه منیفولد-منظر با توجه به اینکه شباهت میان تصاویر خاصیت تعدی دارد، انتخاب می‌شود و علاوه بر شباهت میان تصویر آزمون و تصاویر همسایه، به شباهت میان تصاویر همسایه کاندید نیز توجه می‌شود. مشکل بعدی MSGE در اتصال MMFها است. اتصالMMF ها نیز بر اساس نزدیک‌ترین تصاویر به عناصرMMF ها انجام می‌شود. واضح است ممکن است اتصالاتی بین MMFهای هم‌منظر و غیر هم‌طبقه اتفاق بیافتد. در روش پیشنهادی به‌جای تأثیر یک تصویر، میانگین تصاویر موجود در هر قطعه بر اتصال تأثیر می‌گذارند تا بتوانیم تأثیر تصاویر غیر هم‌طبقه و هم‌منظر را کاهش دهیم.

ازآنجایی‌که در مجموعه داده‌ی CAS-PEAL تنها دو تصویر با منظر مشابه برای هر فرد وجود دارد و در MSGE انتخاب عناصر قطعه منیفولد-مودال تنها بر اساس روش $k$-NN است، انتخاب مقدار $k$ برابر با سه و بیشتر موجب انتخاب عناصر قطعه منیفولد-مودال از طبقه‌های دیگر می‌شود. همان‌طور که نتایج عملی نیز تأیید می‌کند کارایی روش پیشنهادی و MSGE به‌شدت وابسته به مقدار $k$ است و برای تمام مجموعه داده‌های چندمنظری مناسب نیست. روش پیشنهادی عناصر قطعه منیفولد-منظر را با توجه با خاصیت تعدی

در شباهت میان عناصر انتخاب می‌کند، بنابراین شانس بیشتری برای انتخاب عناصر هم‌طبقه دارد. هنگامی‌که تغییرات کوچکی در موقعیت اشیا در طول یک چرخه‌ی دید داریم، مثل مجموعه داده‌ی COIL-20 و CAS-PEAL روش پیشنهادی به‌خوبی می‌تواند کارایی خود را نشان دهد. مجموعه داده‌ی COIL-20 برخلاف مجموعه داده‌های چهره حاوی تصاویر اشیا با هندسه متفاوت هستند و برای هر شیء ۱۵ تصویر با منظر بسیار مشابه وجود دارد. همان‌طور که از نتایج قابل‌مشاهده است،MMF در این حالت نتایج بهتری نسبت به مجموعه داده چهره به‌دست آورده است. برتری روش پیشنهادی هنگامی‌که هندسه‌ی اشیا شبیه به هم است و نیز از هر شیء تنها یک تصویر هم منظر وجود دارد، یعنی در پایگاه داده CAS-PEAL، FEI چشم‌گیر است. در CAS-PEAL نُه دوربین در بازه‌های برابر در یک راستای نیم‌دایره‌ی افقی قرارگرفته شد و ۲۱ تصویر از هر فرد ذخیره شد. در FEI از هر فرد ۱۳ تصویر در راستای نیم‌دایره افقی داریم. در FEI نیز بازه‌ی تغییر زاویه‌ی سر برای تمام افراد برابر است اما بیشتر از بازه‌ی تغییر زاویه چرخش در CAS-PEAL است. به خاطر تغییرات کم در زاویه‌ی چرخش و خاصیت تعدی در شباهت میان تصاویر، روش پیشنهادی در تعداد همسایگی بیشتر نیز می‌تواند نتایج قابل قبولی را برای CAS-PEAL به‌دست آورد. اما در مجموعه تصاویر FEI، شباهت میان تصاویر هم‌منظر از طبقه‌های مختلف بر خاصیت تعدی در شباهت میان تصاویر هم‌طبقه غلبه می‌کند و در تعداد همسایگی بالا روی FEI نتایج مطلوبی را به‌دست نخواهد آورد.

برتری روش هنگامی‌که هندسه‌ی اشیا شبیه به‌هم است و نیز از هر شیء تنها یک تصویر هم منظر وجود دارد، یعنی در پایگاه داده CAS-PEAL و FEI چشم‌گیر است.

اتصال قطعه منیفولدهای منظر در روش MSGE براساس نزدیک‌ترین تصویر در دو قطعه منیفولد-منظر انجام شد. در روش پیشنهادی استفاده از میانگین فاصله میان عناصر منیفولدهای منظر برای اتصال منیفولدها پیشنهاد شد. برای نشان دادن برتری استفاده از میانگین فاصله میان عناصر قطعه منیفولدها، در شکل ۱۱، نمودار میانگین فاصله‌ی عناصر در قطعه منیفولد-منظر هم‌طبقه نسبت به نمودار میانگین فاصله‌ی عناصر در قطعه منیفولد-منظر غیر هم‌طبقه برای پنج تصویر که از منظرهای مختلف به‌صورت تصادفی انتخاب شده ارائه شد.




**شکل ۱۱ : نمودار میانگین فاصله‌ی عناصر در قطعه منیفولد-منظر هم‌طبقه نسبت به نمودار میانگین فاصله‌ی عناصر در قطعه منیفولد-منظر غیر هم‌طبقه**

### ۴-۳- بازشناسی چهره

برای مقایسه، چندین روش یادگیری منیفولد مثل LLE، Isomap، LE، S-LE، S-Isomap و SNG-LLE استفاده شد. اگرچه تمام روش‌های یادگیری منیفولد توانایی یافتن نمایش داده با بُعد کم را دارند، اما در بسیاری از آن‌ها با خطا در معنای فاصله‌ی تخمین زده شده روبه‌رو هستیم. این خطا موجب نادیده گرفتن توزیع داده روی چندین منیفولد می‌شود. ما روش پیشنهادی را با روش‌های یادگیری مالتی‌منیفولد نیز مقایسه کردیم. این روش‌ها بر پایه‌ی ساخت گراف همسایگی جدید و تلاش برای حفظ بهتر فاصله‌ی پیمایشی درون طبقه‌ای و برون طبقه‌ای عمل می‌کنند پنج روش یادگیری مالتی‌منیفولد زیر برای ارزیابی مورد استفاده قرار گرفت:

۱. الگوریتم M-Isomap، k جزء متصل در مجموعه داده را تعیین می‌کند (که k تعداد همسایگان است) و سپس Isomap استاندارد را روی هر یک از این اجزاء متصل، به‌صورت مستقل، اعمال می‌کند.

۲. SLLE یک روش انتخاب همسایگی باناظر برای داده‌ی مالتی‌منیفولد است که به‌طور مصنوعی فاصله‌ی بین اشیا متعلق به طبقه‌های مختلف را افزایش می‌دهد اما فاصله‌ی میان اشیا هم‌طبقه را بدون تغییر باقی می‌گذارد.

۳. M-LLE از روش SLLE برای انتخاب همسایگی در فاز اول LLE استفاده می‌کند.

۴. SH-NGC-LLE از روشSNG برای انتخاب همسایگی در فاز اول LLE استفاده می‌کند. SNG روش یادگیری منیفولد باناظر است که با نزدیک‌کردن تصاویر هم‌طبقه و دور کردن تصاویر غیرهم‌طبقه سعی در بهبود یادگیری منیفولد دارد.

۵. MSGE با استفاده از روش k-NN، قطعه منیفولد-مودال (MMF) را به‌عنوان بخشی از منیفولد مودال در یک طبقه تعریف

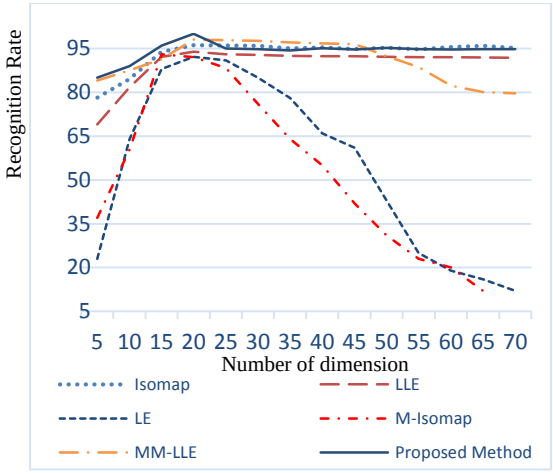
می‌کند. سپس، MSGE، قطعه MMFهای یک طبقه و همچنین نزدیک‌ترین رئوس همسایه در دو منیفولد-طبقه و یک منیفولد مودال را به‌هم متصل می‌کند.

برای مقایسه‌ی دقت روش پیشنهادی و روش‌های ذکر شده در بالا، ما نقاط داده در فضای با ابعاد بالا را به فضای با بُعد کم تصویر کردیم. سپس، با استفاده از طبقه‌بندی کننده‌ی ۵-نزدیک‌ترین همسایه به پیش‌بینی برچسب طبقه تصویر پرس‌وجو پرداختیم.

لازم به یادآوری است که برخلاف روش‌های یادگیری منیفولد استاندارد، روش پیشنهادی یک تابع صریح برای نگاشت نقاط داده از فضای با ابعاد بالا به فضای با ابعاد کم، یا به‌عبارت‌دیگر راه‌حلی برای مشکل توسعه خارج از نمونه در یادگیری منیفولد، ارائه می‌دهد. برای دیگر روش‌های یادگیری منیفولد که دارای تابع صریح برای نگاشت داده نیستند، باید یک تابع نگاشت از نقاط داده در مجموعه آموزشی به نمونه‌های متناظر با آن‌ها در فضای با بُعد کم به‌دست آوریم. ما از شبکه رگرسیون تعمیم یافته برای تقریب تابع نگاشت استفاده کردیم. سپس، تصویر آزمون بدون برچسب جدید با استفاده از نگاشت کاهش بُعد یاد گرفته شده، نگاشت داده شد. برای هر مجموعه داده، به‌طور تصادفی سه تصویر از منظرهای مختلف را به‌عنوان تصویر آزمایشی انتخاب کردیم و باقی تصاویر به‌عنوان نمونه‌های آموزشی استفاده شد. آزمایشات چهار بار با تصاویر انتخاب شده به‌صورت تصادفی تکرار و از نتایج میانگین گرفته شد.

برای بررسی عملی محدوده‌ی مقدار d به عنوان بُعد کاهش یافته ما به مقایسه‌ی روش پیشنهادی و دو روش یادگیری مالتی منیفولد M-Isomap و MM-LLE و نیز تعدادی روش یادگیری منیفولد روی مجموعه داده‌ی COIL-20 پرداختیم. برای مقایسه تأثیر مقدارd بر الگوریتم‌های یادگیری، طیف وسیعی از مقادیر d پوشش داده شد، که در شکل ۱۲ نشان داده شده است. با توجه به مقادیر مختلف d بررسی شده، بهترین نتایج براساس مقدار بُعد ذاتی به‌دست خواهد آمد. از این رو، نرخ تشخیص برای مقایسه‌ی روش پیشنهادی و روش‌های پیشین فقط برای مقدار ذاتی d که با استفاده از تابع intrinsic_dim مبتنی بر روش بیشینه درست نمایی[۱۶] در جعبه‌ابزار کاهش بعد متلب DR Toolbox [۲۸] به دست آمد.

نتایج مقایسه‌ی روش پیشنهادی با دیگر روش‌های یادگیری منیفولد روی COIL-20 در شکل ۱۳ نشان داده شد. همان‌طور که در روش پیشنهادی و روش MSGE تعداد عناصر در قطعه منیفولد-منظر بر نرخ بازشناسی تأثیرگذار است، در باقی روش‌های مورد مقایسه نرخ بازشناسی به تعداد همسایه‌ی انتخاب شده وابسته است.

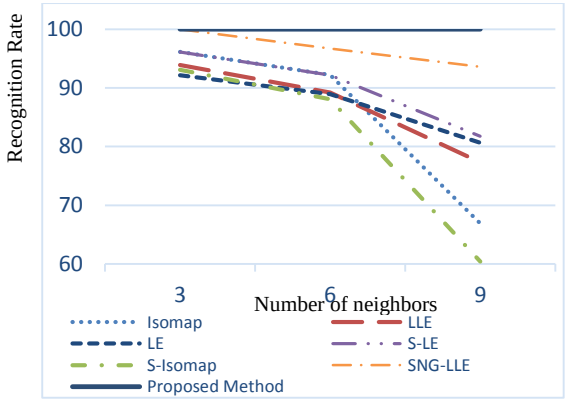



**شکل ۱۲ : بررسی تأثیر بُعد کاهش یافته بر نرخ بازشناسی**

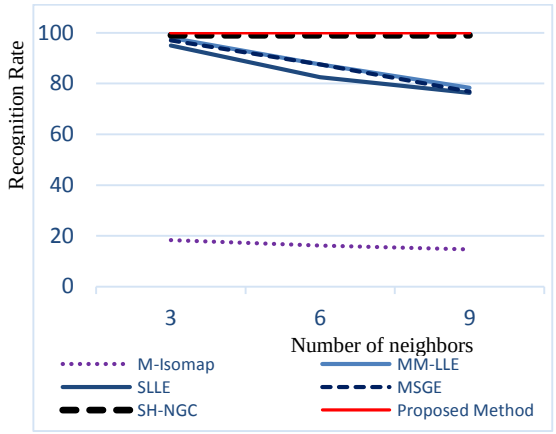
برای نشان دادن مزیت روش پیشنهادی در انتخاب صحیح همسایگی، نرخ بازشناسی با تعداد مختلف همسایگی مورد آزمایش قرار گرفت. لازم به ذکر است، نتایج آزمایشات ما با نتایج ارائه‌شده برای S-LE در [۸] مطابقت ندارد. دلیل این عدم انطباق را می‌توان به خاطر روش متفاوت مقایسه دانست. در S-LE، تمام مجموعه داده روی زیر فضای تعبیه تصویر می‌شود، سپس داده در فضای کم بُعد به دو بخش آموزشی و آزمایشی تقسیم می‌شود. نمونه‌های آزمایشی براساس نزدیک‌ترین همسایه در فضای تعبیه طبقه‌بندی می‌شوند. نتایج کاهش بُعد گزارش شده در S-LE، هم از داده‌ی آموزشی و هم از داده‌ی آزمایشی بهره می‌برد که این مطلب برای وظایف دنیای واقعی صحیح نیست؛ ما در اولین مرحله از بازشناسی، داده‌های آزمایشی را در اختیار نداریم، بلکه باید داده‌ی آموزشی را به فضای با بُعد کم تصویر کنیم، سپس تابع نگاشت از نقاط داده در فضای اصلی و نقاط متناظر آن‌ها در فضای هم بُعد را یاد بگیریم. سپس، داده‌ی آزمون جدید و بدون برچسب را با استفاده از تابع نگاشت یاد گرفته شده، نگاشت دهیم. بنابراین، نتایج ما نسبت به نتایج گزارش شده برای S-LE مقادیر کمتری دارد.

نتایج مقایسه‌ی روش پیشنهادی و روش‌های مالتی‌منیفولد روی مجموعه داده‌ی COIL-20 در شکل ۱۳(ب) نشان داده شده است. کارایی M-Isomap به‌شدت وابسته به انتخاب همسایگی k-NN است.

بنابراین در پایگاه داده‌هایی مثل COIL-20 (شکل۱۳(ب)) که شامل تصاویر مختلف با زاویه‌های دید مختلف است، همسایه‌ها ممکن است از طبقه‌های مختلف انتخاب شوند. از این‌رو، نرخ بازشناسی در M-Isomap براساس دقت انتخاب همسایگی می‌تواند تغییر کند.




(الف)




(ب)

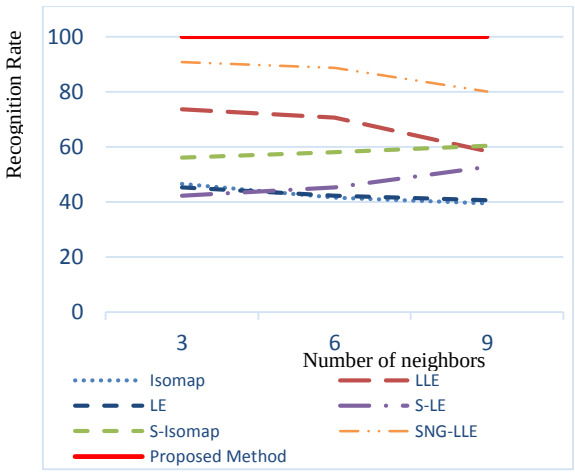
**شکل ۱۳ : مقایسه نرخ بازشناسی برای مجموعه داده‌ی COIL-20 با مقادیر مختلف تعداد همسایگان (k). الف) نرخ بازشناسی روش‌های یادگیری منیفولد و روش پیشنهادی. ب) نرخ بازشناسی روش‌های یادگیری مالتی‌منیفولد و روش پیشنهادی.**

نتایج آزمایش‌هایی که در شکل ۱۳(ب) خلاصه شده است نشان می‌دهد که SLLE و MM-LLE در وظایف دنیای واقعی کارایی بالایی ندارند. بنابراین استفاده از آن‌ها در انتخاب همسایگی برای داده‌ی مالتی‌منیفولد و چندمنظری مناسب نیست. SH-NGC-LLE با رفتاری باناظر تلاش می‌کند با تعریف معنایی جدید برای فاصله‌ی بین نقاط داده همسایگان هم‌طبقه از منظرهای مختلف را انتخاب کند. در SH-NGC-LLE تناسب میان نقاط داده و معنای هندسی توزیع داده حفظ می‌شود. کارایی MSGE وابسته به تعداد تصاویر آموزشی هم منظر برای هر شیء است. در MSGE برای انتخاب عناصر هر قطعه منیفولد-مودال از روش $k$-NN استفاده شد. در مجموعه داده‌هایی مثل COIL-20 که تعداد کمی از تصاویر هم‌طبقه، هم‌منظر هستند، نرخ بازشناسی برای MSGE با افزایش تعداد عناصر در قطعه منیفولد-مودال، کاهش می‌یابد. کارایی روش پیشنهادی نیز به تعداد تصاویر آموزشی هم منظر برای هر شیء وابسته است. در روش پیشنهادی برای انتخاب عناصر هم‌طبقه و هم‌منظر بر خاصیت تعدی در شباهت میان تصاویر مشابه تأکید دارد. با انتخاب دقیق عناصر منیفولد-منظر می‌توان ساختار محلی هر منیفولد و با اتصال عناصر منیفولد-منظر هم‌طبقه معنای هندسی را برای مدل‌سازی تصاویر مالتی‌منیفولد حفظ کرد. همان‌طور که در نتایج شکل ۱۳ (ب) قابل مشاهده است، نتایج روش پیشنهادی که یک روش بدون ناظر است با نتایج روش باناظر SH-NGC-[۱۳] LLE نزدیک است.

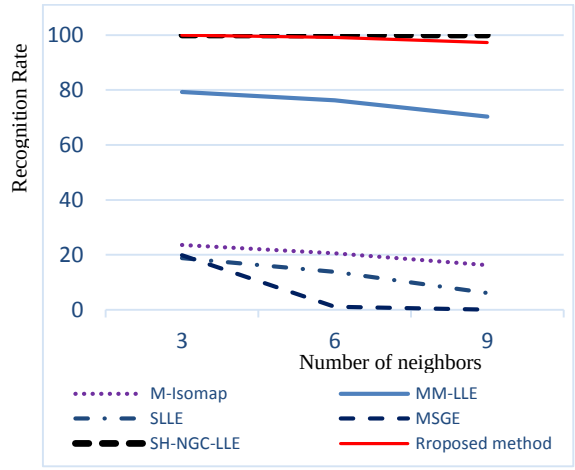
ازآنجایی‌که اشیا در طبقه‌های مختلف مجموعه داده‌های چهره، هندسه‌ی مشابهی دارند و تعداد افراد برای هر منظر در هر طبقه نسبتاً کم است، می خواهیم بررسی کنیم که چگونه عملکرد هر یک از روش‌های یادگیری منیفولد هنگامی‌که که تعداد همسایه‌ها افزایش می یابد، تغییر می‌کند. برای این هدف، از مجموعه داده‌های CAS-PEAL و FEI برای آزمایش روش‌های یادگیری مالتی‌منیفولد چندمنظری استفاده شد. آزمایشات روی مجموعه داده‌های چهره مشابه با COIL-20 است. نتایج آزمایشات روی CAS-PEAL در شکل ۱۴ نشان داده شد. نتایج مقایسه‌های روش پیشنهادی و روش‌های یادگیری منیفولد در شکل ۱۴(الف) و نتایج مقایسه‌ی روش پیشنهادی با روش‌های یادگیری مالتی‌منیفولد در شکل ۱۴(ب) نشان داده شد. نتایج آزمایش‌های انجام شده روی FEI در شکل ۱۵ ارائه شد. نتایج مقایسه‌های روش پیشنهادی و روش‌های یادگیری منیفولد در شکل ۱۴(الف) و نتایج مقایسه‌ی روش پیشنهادی با روش‌های یادگیری مالتی‌منیفولد در شکل ۱۴(ب) نشان داده شد. نتایج آزمایشات روی این مجموعه داده‌ها برتری روش پیشنهادی را نسبت به روش‌های مطرح یادگیری مالتی‌منیفولد نشان می‌دهد.

ازآنجایی‌که در مجموعه داده‌ی FEI و CAS-PEAL تنها دو تصویر با منظر مشابه برای هر فرد وجود دارد، در MSGE که انتخاب عناصر قطعه منیفولد-مودال تنها براساس روش $k$-NN است، انتخاب مقدار $k$ برابر با سه موجب انتخاب عناصر قطعه منیفولد-مودال از طبقه‌های دیگر می‌شود. همان‌طور که نتایج عملی نیز تأیید می‌کند کارایی روش MSGE به شدت وابسته به مقدار $k$ است و برای تمام مجموعه داده‌های هم منظر مناسب نیست.

روش پیشنهادی عناصر قطعه منیفولد-منظر را با توجه با خاصیت تعدی در شباهت میان عناصر انتخاب می‌کند، بنابراین شانس بیشتری برای انتخاب عناصر هم‌طبقه دارد. روش‌های SLLE، MM-LLE، M-Isomap و SH-NGC-LLE برخلاف روش پیشنهادی و MSGE از برچسب نقاط داده برای افزایش فاصله‌ی میان نقاط هم‌طبقه استفاده می‌کنند و نتایج بهتری را به‌دست می‌آورند.

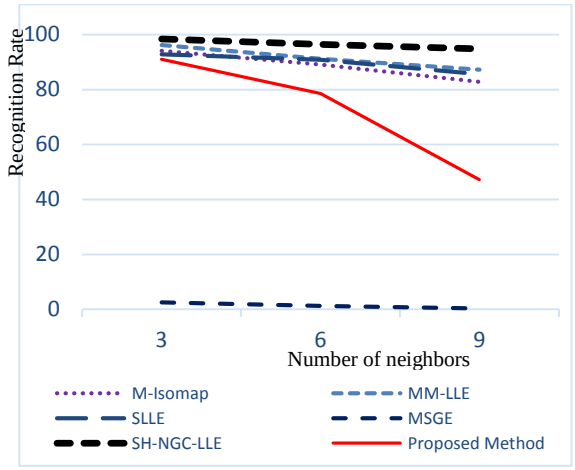



(الف)

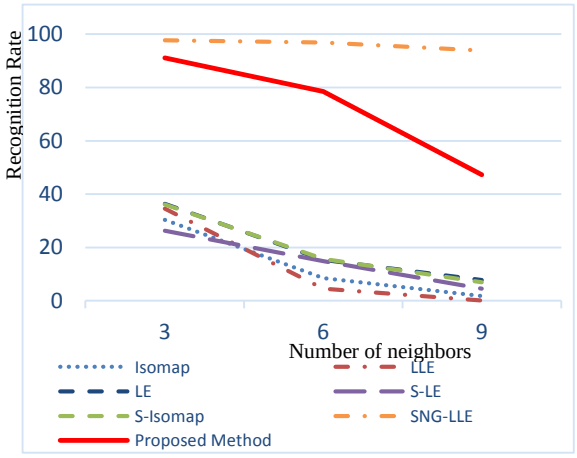



(ب)

**شکل ۱۴ : مقایسه نرخ بازشناسی برای مجموعه داده‌ی CAS-PEAL با مقادیر مختلف تعداد همسایگان (k). الف) نرخ بازشناسی روش‌های یادگیری منیفولد و روش پیشنهادی. ب) نرخ بازشناسی روش‌های یادگیری مالتی‌منیفولد و روش پیشنهادی.**




(الف)




(ب)

**شکل ۱۵ : مقایسه نرخ بازشناسی برای مجموعه داده‌ی FEI با مقادیر مختلف تعداد همسایگان (k). الف) نرخ بازشناسی روش‌های یادگیری منیفولد و روش پیشنهادی. ب) نرخ بازشناسی روش‌های یادگیری مالتی‌منیفولد و روش پیشنهادی**

## ۵- نتیجه‌گیری

در این مقاله، روشی بدون ناظر برای کشف خودکار منیفولد طبقه در بازشناسی مالتی‌منیفولد چندمنظری پیشنهاد شد. ایده‌ی روش پیشنهادی از نیاز روش‌های یادگیری مالتی‌منیفولد برای انتخاب منیفولدهای هم‌طبقه از میان تصاویر موجود از منظرهای مختلف مربوط به اشیا متفاوت ناشی شد. ما تصاویر مربوط به هر شیء که از زوایای دید مختلف تصویربرداری شده است را به‌عنوان یک منیفولد-طبقه نسبتاً مستقل طبقه‌بندی کردیم. برای تعبیه گراف داده چندمنظری، روشی برای ساخت گراف مطلوب است که هم ساختار مالتی‌منیفولد چندمنظری و هم اطلاعات متمایزکننده را منعکس کند. برای این منظور ما قطعه منیفولد-منظر را به‌عنوان یک واحد اطلاعاتی مستقل تعریف کردیم که حاوی تصاویر هم‌طبقه و هم‌منظر باشد. با توجه به اینکه در روش پیشنهادی اطلاعات برچسب در اختیار نیست، با توجه به خاصیت تعدی در شباهت میان تصاویر، قطعه‌های منیفولد-منظر انتخاب شد. برای حفظ رابطه‌ی بین منیفولد محلی درون طبقهی و اطلاعات متمایزکننده برون طبقه‌ای، بدون استفاده از اطلاعات برچسب طبقه، از میانگین فاصله‌ی بین قطعه‌های منیفولد-منظر استفاده کردیم. در انتها، برای ساخت یک نگاشت کاهش بُعد صریح، ما یک روش کاهش بُعد غیرخطی باناظر پیشنهاد دادیم. روش پیشنهادی ما تصاویر یک شیء را به مکان‌های نزدیک هم و تصاویر اشیا دیگر را به مکان‌های دور از آن نگاشت می‌دهد. هدف روش پیشنهادی ما، حل مسئله‌ی بازشناسی است، درحالی‌که تصاویر هر شیء از چندین زاویه دید تصویربرداری شده است. روش پیشنهادی با استفاده از سه مجموعه

تصاویر چندمنظری موردی ارزیابی و مقایسه با روش‌های پیشین قرار گرفت. نتایج به‌دست آمده بسیار امیدوار کننده است؛ نتایج، توانایی روش پیشنهادی را برای انتخاب تصاویر هم‌طبقه و هم‌منظر تأیید می‌کند. یافته‌های ما در نرخ بازشناسی توانایی روش در شرایط چندمنظری سخت مثل چهره را نشان می‌دهد.

## مراجع


[1] J. Li, Y. Wu, J. Zhao, and K. Lu, "Multi-manifold Sparse Graph Embedding for Multi-modal Image Classification," *Neurocomputing*, vol. 173, pp. 501–510, 2016.

[2] YangxiLi, XinShi, CuilanDu, YangLiu, and YonggangWen, "Manifold regularized multi-view feature selection for social image annotation," *Neurocomputing*, vol. 204, pp. 135-141, 2016.

[3] S. Sun, "A survey of multi-view machine learning," *Neural Comput & Applic*, vol. 23, pp. 2031–2038, 2013.

[4] B. Schölkopf, A.Smola, and K.-R.Müller, "Kernel principal component analysis," *Artificial NeuralNetworks-ICANN'97,Springer,Lausanne,Switzerland,*, pp. 583–588, 1997.

[5] W. Liu, H. Zhang, D. Tao, Y. Wang, and K. Lu, "Large-scale paralleled sparse principal component analysis," *Multimed.ToolsAppl*, pp. 1-13, 2013.

[6] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, *et al.*, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS*, vol. 38, pp. 149-161, 2008.

[7] X. Geng, D. C. Zhan, and Z. H. Zhou, "Supervised Nonlinear Dimensionality Reduction for Visualization and Classification," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS*, vol. 35, pp. 1098-1107, 2005.

[8] B. Raducanu and F.Dornaika, "A supervised non-linear dimensionality reduction approach for manifold learning," *Pattern Recognition*, vol. 45, pp. 2432–2444, 2012.

[9] F. Aeini, A. m. Eftekhari-Moghadam, and F. Mahmoudi, "Non Linear Dimensional Reduction Method based on Supervised Neighborhood Graph," presented at the 7th International Symposium on Telecommunications (IST'2014), Tehran, Iran, 2014.

[10] D. d. Ridder, O. Kouropteva, O. Okun, M. Pietikainen, and R. P. W. Duin, "Supervised locally linear embedding," presented at the Artificial Neural Networks and Neural Information Processing-ICANN/ICONIP 2003, 2003.

[11] Z. Zhang, T. W. S. Chow, and M. Zhao, "M-Isomap: Orthogonal Constrained Marginal Isomap for Nonlinear Dimensionality Reduction," *IEEE Transactions on Cybernetics* vol. 43, pp. 180 - 191, 2013.

[12] R. Hettiarachchi and J.F.Peters, "Multi-manifold LLE learning in pattern recognition," *Pattern Recognition*, vol. 48, pp. 2947–2960, 2015.

[13] F. Aeini, A. M. E. Moghadam, and F. MAhmoudi, "Supervised hierarchical neighborhood graph construction for manifold learning," *Signal, Image and Video Processing*, vol. 12, pp. 1-9, 2018.

[14] C.-S. Lee, A. Elgammal, and M. Torki, "Learning representations from multiple manifolds," *Pattern Recognition*, vol. 50, pp. 74-87, 2016.

[15] H.-T. Chen, H.-W. Chan, and T.-L. .Liu, "Local discriminant embedding and its variants," presented at the 2005 IEEE Computer Society Conferenceon Compute rVision and Pattern Recognition, CVPR, 2005.

[16] M. Sugiyama, "Local fisher discriminant analysis for supervised dimensionality reduction," presented at the Proceedings of the 23rd International Conferenceon Machine Learning, Pittsburgh, Pennsylvania, 2006.

[17] D. Cai, X. He, K. Zhou, J. Han, and H. Bao, "Locality sensitive discriminantan alysis," *IJCAI*, pp. 708–713, 2007.

[18] X. He, M. Ji, and H. Bao, "Graph embedding with constraints," *IJCAI*, vol. 9, pp. 1065–1070, 2009.

[19] Y. Bengio, J. F. Paiement, P. Vincent, O. Delalleau, N. L. Roux, and M. Ouimet, "Out-of-sample extensions for LLE, Isomap, MDS, eigenmaps, and spectral clustering," presented at the Proceedings of the Neural Information Processing Systems, MIT Press, Cambridge, MA, 2004.

[20] W. K. Wonga and H. T. Zhao, "Supervised optimal locality preserving projection," *Pattern Recognition* ,vol. 45 pp. 186–197, 2012.

[21] B. Yang, M. Xiang, and Y. Zhang, "Multi-manifold discriminant Isomap for visualization and classification," *Pattern Recognition*, vol. 55, pp. 215–230, 2016.

[22] N. Gu, M. Fan, H. Qiao, and B. Zhang, "Discriminative Sparsity Preserving Projections for Semi-Supervised Dimensionality Reduction," *IEEE SIGNAL PROCESSING LETTERS*, vol. 19, pp. 391-394, 2012.

[23] M. Fan, X. Zhang, Z. Lin, Z. Zhang, and H. Bao, "A Regularized Approach for Geodesic-Based Semisupervised Multimanifold Learning," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 23, pp. 2133-2147, 2014.

[24] B. Fischer and J. M. Buhmann, "Path-based clustering for grouping of smooth curves and texture segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol .25 ,pp. 513–518, 2003.

[25] T. Evgeniou, T. Poggio, M. Pontil, and A. Verric, "Regularization and statistical learning theory for data analysis," *Computational Statistics & Data Analysis*, vol. 38, pp. 421-432, 2002.

[26] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold Regularization: A Geometric Framework for Learning from Labeled and Unlabeled Examples," *J. Mach. Learn. Res.*, vol. 7, pp. 2399-2434, 2006.

[27] N. Gu, M. Fan, H. Qiao, and B. Zhang, "Discriminative Sparsity Preserving Projections for Semi-Supervised Dimensionality Reduction," *IEEE Signal Processing Letters*, vol. 19, pp. 391 - 394, 2012.

[28] L. v. d. Maaten, E. Postma, and J. v. d. Herik, "Dimensionality Reduction: A Comparative Review," Tilburg University Technical Report, TiCC-TR 2009-

[29] J. Lu, Y.-P. Tan, and G. Wang, "Discriminative Multimanifold Analysis for Face Recognition from a Single Training Sample per Person," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 39 - 51, 2013.

[30] M. Fornasier and F. Pitolli, "Adaptive iterative thresholding algorithms for magnetoence phalography (MEG)," *Comput.Appl.Math.*, vol. 211, pp. 386–395, 2008.

[31] F. Nie, H. Huang, X. Cai, and C. Ding, "Efficient and robust feature selection via joint ℓ2,1-norms minimization," *Advances in Neural Information Processing Systems*, 2010.

[32] C. EduardoThomaz and G. AntonioGiraldi, "A new ranking method for Principal Components Analysis and its application to face image analysis," *Image and Vision Computing*, vol. 28, pp. 902-913, 2010.

[33] T. C. E. and G. G. A., "A new ranking method for Principal Components Analysis and its application to face image analysis," *Image and Vision Computing*, vol. 28, pp. 902-913, 2010.


## پاورقی‌ها:

[1] Manifold learning
[2] Real-world
[3] Embedded
[4] Neighborhood graph
[5] Multi-view
[6] class
[7] inconsistency
[8] Face recognition
[9] Action recognition
[10] Object
[11] Geodesic
[12] explicit
[13] Generalized regression network
[14] Out-of-sample
[15] Regularized
[16] Maximum Likelihood Estimator