نوع مقاله: پژوهشی

چارچوب نوین خودرمزگذار مبتنی بر ResNet برای خوشه‌بندی نیمه‌نظارتی با بهره‌گیری از محدودیت‌های زوجی

نور کاظم خودهیر [۱]، دانشجو، محمدعلی بالافر[۱]، استاد تمام[۲]، جعفر تنها[۳]، دانشیار، امین گلزاری اسکوئی[۴]، استادیار
۱- دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران- noorkadem669@gmail.com
۲- دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران- balafarila@tabrizu.ac.ir
۳- دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران- tanha@tabrizu.ac.ir
۴-دانشکده فناوری اطلاعات و مهندسی کامپیوتر دانشگاه صنعتی ارومیه، ارومیه - a.golzari@uut.ac.ir

چکیده: اخیراً استفاده از یادگیری عمیق برای خوشه‌بندی داده‌ها به دلیل توانایی آن در کشف ساختارهای پیچیده داده‌ها، مورد توجه بسیاری از پژوهشگران قرار گرفته است. در این مقاله، یک چارچوب نوین خودرمزگذار مبتنی بر ResNet برای خوشه‌بندی نیمه‌نظارتی پیشنهاد شده است که شامل معماری خودرمزگذار با اتصالات باقیمانده است. این چارچوب دو بخش اصلی دارد: رمزگذار مبتنی بر ResNet که بازنمایی‌های نهفته و معنادار از داده‌ها استخراج می‌کند و رمزگشا که وظیفه بازسازی دقیق داده‌های ورودی را بر عهده دارد. مدل پیشنهادی از یک تابع زیان ترکیبی بهره می‌برد که شامل میانگین مربعات خطا (MSE)، واگرایی کولبک-لایبلر (KLD)، محدودیت‌های زوجی نیمه‌نظارتی و برچسب‌های نیمه‌نظارتی است. این ترکیب نوآورانه، فرایند خوشه‌بندی را با استفاده از توزیع هدف به‌عنوان برچسب‌های نرم هدایت کرده و موجب بهبود پایداری مدل می‌شود. نتایج آزمایش‌ها بر روی مجموعه داده‌های عمومی نشان می‌دهد که مدل پیشنهادی به‌طور میانگین دقت ۹۶٫۸٪ در معیار دقت خوشه‌بندی و ۹۲٫۵٪ در معیار اطلاعات متقابل نرمال‌شده (NMI) را به دست آورده است. این نتایج حاکی از بهبود عملکرد مدل نسبت به روش‌های پیشین در خوشه‌بندی نیمه‌نظارتی است.

واژه‌های کلیدی: خوشه‌بندی نیمه‌نظارتی، خودرمزگذار مبتنی بر ResNet، محدودیت‌های زوجی، یادگیری عمیق برای خوشه‌بندی

A Novel ResNet-Based Autoencoder Framework for Semi-Supervised Clustering Using Pairwise Constraints

Noor Kadhim Khudhair, PhD student [1], Mohammadali Balafar, Full professor [2], Jafar Tanha, Associate professor [3], Amin Golzari Oskouei, Assistant professor [4]
1- Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, Email: noorkadem669@gmail.com
2- Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, Email: balafarila@tabrizu.ac.ir
3- Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, Email: tanha@tabrizu.ac.ir
4- Faculty of IT and Computer Engineering, Urmia University of Technology, Urmia, Iran, Email: a.golzari@uut.ac.ir

Abstract: Recently, the use of deep learning for data clustering has gained significant attention due to its ability to uncover complex structures within data. In this paper, a novel ResNet-based autoencoder framework for semi-supervised clustering is proposed. This framework utilizes an autoencoder architecture with residual connections, consisting of two main components: a ResNet-based encoder that extracts meaningful latent representations from the data, and a decoder that is responsible for accurately reconstructing the input data. The proposed model employs a composite loss function that integrates Mean Squared Error (MSE), Kullback-Leibler Divergence (KLD), semi-supervised pairwise constraints, and label-based loss. This innovative combination guides the clustering process using a target distribution as soft labels, ensuring the stability of the model. Experimental results on benchmark datasets demonstrate that the proposed model achieves an average clustering accuracy of 96.8% and 92.5% in Normalized Mutual Information (NMI). These results indicate a significant improvement in performance compared to existing methods in semi-supervised clustering.

Keywords: Semi-Supervised Clustering, ResNet-Based Autoencoder, Pairwise Constraints, Deep Learning for Clustering

تاریخ ارسال مقاله: ۱۴۰۳/۰۹/۱۳
تاریخ اصلاح مقاله: ۱۴۰۴/۰۶/۲۵
تاریخ پذیرش مقاله: ۱۴۰۴/۰۷/۱۲
نام نویسنده مسئول: محمدعلی بالافر
نشانی نویسنده مسئول: گروه کامپیوتر - دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران.

۱- مقدمه

تجزیه و تحلیل خوشه‌ای (Clustering) یکی از روش‌های مهم برای تحلیل مجموعه داده‌ها است که به درک ساختار ذاتی داده کمک می کند [۵-۱]. فرایند گروه‌بندی مجموعه‌ای از نقاط داده به گروه‌های متمایز که به آن‌ها خوشه (Cluster) گفته می‌شود، با عنوان خوشه‌بندی شناخته می‌شود [۱۱-۶ ,۳]. اهداف اصلی خوشه‌بندی شامل افزایش تنوع درون‌خوشه‌ای و کاهش هم‌زمان تنوع بین‌خوشه‌ای است [۱۴-۱۲ ,۱]. الگوریتم‌های شناخته‌شده خوشه‌بندی شامل تکنیک‌های خوشه‌بندی نرم، همان‌طور که در [۱۵] نشان داده شده و روش‌های خوشه‌بندی سخت، همان‌طور که در [۱۶ ,۱] توضیح داده شده، در این حوزه به‌خوبی شناخته‌شده هستند. علاوه بر این، نسخه‌های جایگزین این الگوریتم‌ها در ادامه مورد بحث قرار گرفته‌اند [۱۷, ۱۸].

با این حال، محدودیت‌های الگوریتم‌های سنتی خوشه‌بندی در زمان کاربرد در مجموعه داده‌های گسترده آشکار می‌شود [۲۰ ,۱۹] که این موضوع ناشی از ضعف در معیارهای شباهت پایه‌ای آن‌ها است [۲۱, ۲۲]. در نتیجه، این الگوریتم‌ها نمی‌توانند خوشه‌های بهینه را برای مجموعه داده‌های بزرگ تولید کنند، عمدتاً به دلیل پیچیدگی محاسباتی بالای آن‌ها. برای کاهش این چالش، تکنیک‌های کاهش ابعاد به‌منظور تبدیل داده‌های ورودی به فضای با ابعاد پایین‌تر اعمال می‌شوند، همان‌طور که تحلیل مؤلفه اصلی (PCA) نشان می‌دهد [۲۳]. با این حال، ویژگی‌های پیچیده ساختار داده‌ها همچنان چالشی برای اثربخشی روش‌های خوشه‌بندی موجود ایجاد می‌کنند [۲۴].

پیشرفت‌های اخیر در یادگیری بازنمایی عمیق امکان تحولات قابل‌توجهی را از طریق توانایی انجام تبدیلات غیرخطی فراهم کرده‌اند [۳۲-۲۵]. چارچوب‌های خوشه‌بندی مبتنی بر شبکه‌های عصبی عمیق که به آن‌ها خوشه‌بندی عمیق (DC[۱]) گفته می‌شود [۳۵-۳۳]، اخیراً توجه فزاینده‌ای را در محافل علمی به خود جلب کرده‌اند. بسیاری از روش‌های DC شامل یک فرایند دو مرحله‌ای هستند. مرحله اول بر دستیابی به بازنمایی‌های معنادار از داده‌های خام از طریق معماری‌های یادگیری عمیق متمرکز است، در حالی که مرحله دوم به دسته‌بندی این داده‌ها در فضای تعبیه‌شده به دست آمده مربوط می‌شود. در این مرحله دوم، می‌توان از روش‌های خوشه‌بندی سنتی در کنار معماری های یادگیری عمیق برای شناسایی خوشه‌ها استفاده کرد.

علاقه فزاینده محققان به روش‌های خوشه‌بندی عمیق به مزایای ذاتی آن‌ها نسبت داده می‌شود [۳۶]. نخست، ابعاد داده به‌طور قابل‌توجهی کاهش می‌یابد و ویژگی‌هایی ایجاد می‌شود که برای خوشه‌بندی بهینه مناسب هستند. در نتیجه، فرایندهای خوشه‌بندی در این فضای ویژگی تعبیه‌شده جدید با سرعت و کارایی بیشتری انجام می‌شوند. دوم، استخراج بازنمایی‌های جدید، عملکرد خوشه‌بندی را در مواجهه با مجموعه داده‌های پیچیده به‌طور قابل‌توجهی بهبود می‌بخشد.

با وجود برتری الگوریتم‌های خوشه‌بندی عمیق نسبت به روش‌های سنتی، این الگوریتم‌ها به صورت بدون نظارت عمل می‌کنند. ازآنجاکه در کاربردهای عملی غالباً برخی اطلاعات اولیه درباره برچسب‌های داده در دسترس است، مدیریت تأثیر نظارت جزئی بر نتایج مدل‌سازی در خوشه‌بندی اهمیت حیاتی پیدا می‌کند. خوشه‌بندی نیمه‌نظارت‌شده [۳۷] با اعمال دانش خارجی درباره برچسب‌های نقاط داده [۳۸]، مانند محدودیت‌های «باید-پیوند» و «نباید-پیوند»، عملکرد خوشه‌بندی بدون نظارت را بهبود می‌بخشد. تکنیک‌های متعددی از خوشه‌بندی نیمه‌نظارت‌شده که از محدودیت‌های زوجی استفاده می کنند، توسعه داده شده‌اند [۴۱-۳۹ ,۳۷]. با این حال، روش مؤثری که بتوان آن را بر اساس معماری ساده نظیر IDEC فرموله کرد و با استفاده از محدودیت‌های زوجی عملکرد خوشه‌بندی را به‌طور قابل‌توجهی بهبود بخشد، هنوز پیشنهاد نشده است.

با توجه به محدودیت‌های فوق‌الذکر، این مقاله یک چارچوب جدید خودرمزگذار مبتنی بر ResNet برای خوشه‌بندی نیمه‌نظارتی پیشنهاد شده است. این روش از معماری خودرمزگذار با اتصالات باقیمانده بهره می‌برد که شامل دو ماژول رمزگذار و رمزگشا است. رمزگذار مبتنی بر ResNet مسئول استخراج بازنمایی‌های نهفته و رمزگشا مسئول بازسازی داده‌های ورودی است. مدل پیشنهادی از یک تابع خطا ترکیبی شامل میانگین مربعات خطا، محدودیت‌های زوجی نیمه‌نظارتی، واگرایی کولبک-لایبلر و خطا مبتنی بر برچسب‌های نیمه‌نظارتی استفاده می‌کند. این ترکیب باعث بهبود خوشه‌بندی با در نظر گرفتن محدودیت‌های زوجی و بهره‌گیری از داده‌های برچسب‌دار می‌شود. همچنین، استفاده از توزیع هدف به عنوان برچسب‌های نرم، روند خوشه‌بندی را هدایت کرده و پایداری آن را تضمین می‌کند. نتایج نشان می‌دهد که این روش قادر است با ترکیب بازنمایی‌های نهفته و برچسب‌های نرم، دقت خوشه‌بندی را بهبود بخشد.

روش پیشنهادی این مقاله شامل چندین نوآوری مهم در حوزه خوشه‌بندی نیمه‌نظارتی مبتنی بر یادگیری عمیق است که به شرح زیر می‌باشند:

۱) استفاده از اتصالات باقیمانده (Residual Connections) در معماری خودرمزگذار به‌عنوان بخشی از فرآیند استخراج بازنمایی داده‌ها یک نوآوری کلیدی است. این ویژگی امکان یادگیری بازنمایی‌های نهفته‌ی عمیق‌تر و معنادارتر را فراهم کرده و باعث بهبود عملکرد خوشه‌بندی به‌ویژه در مواجهه با داده‌های پیچیده می‌شود.

۲) این روش از محدودیت‌های «باید-پیوند» و «نباید-پیوند» برای هدایت فرآیند خوشه‌بندی استفاده می‌کند. ترکیب این محدودیت‌ها در تابع زیان، امکان بهبود دقت خوشه‌بندی و

[۱] Deep Clustering

حفظ ساختار معنایی داده‌ها در فضای ویژگی را فراهم می کند.

۳) تابع زیان پیشنهادی شامل چهار جزء است: میانگین مربعات خطا (MSE) برای بازسازی داده‌ها، واگرایی کولبک-لایبلر (KLD) برای تنظیم خوشه‌ها، محدودیت‌های زوجی (SSC[2]) برای مدیریت روابط محلی بین نقاط داده و زیان مبتنی بر برچسب‌های نیمه‌نظارتی (SSL[3]). این ترکیب، امکان یادگیری هم‌زمان بازنمایی‌های مؤثر و بهینه‌سازی فرآیند خوشه‌بندی را فراهم می‌کند.

۴) روش پیشنهادی با استفاده از ماتریس محدودیت زوجی (A)، روابط محلی و سراسری نقاط داده را در فضای تعبیه‌شده حفظ می‌کند. این ویژگی، اطمینان حاصل می کند که نقاط داده با برچسب مشابه به یکدیگر نزدیک باشند و نقاط داده با برچسب متفاوت فاصله بیشتری داشته باشند.

این نوآوری‌ها روش پیشنهادی را از سایر روش‌های خوشه‌بندی عمیق و نیمه‌نظارتی متمایز کرده و آن را به یک راه‌حل مؤثر و انعطاف‌پذیر برای مسائل خوشه‌بندی در کاربردهای واقعی تبدیل می‌کند.

مقاله به این ترتیب سازمان‌دهی شده است که در بخش ۲ کارهای مرتبط بیان می‌شود، در بخش ۳روش پیشنهادی معرفی می‌شود، در بخش ۴ آزمایش‌ها بیان می‌شود و در نهایت بخش ۵ به نتیجه‌گیری و کارهای آینده می‌پردازد.

۲- کارهای مرتبط

الگوریتم‌های خوشه‌بندی عمیق اغلب از خود رمزگذارها برای یادگیری بازنمایی‌های جدید از داده‌های خام استفاده می‌کنند. معمولاً در این الگوریتم‌ها، تابع خطا خود رمزگذار به تابع خطا خوشه‌بندی اضافه می‌شود تا به‌صورت مشترک بازنمایی داده‌ها و نمایندگان خوشه یاد گرفته شود. در این حالت، برای ایجاد ثبات در این الگوریتم‌ها، مرحله پیش آموزش شبکه ضروری است [۴۲]. پیش آموزش شامل ۲ مرحله است. مرحله اول آموزش خود رمزگذار است و مرحله دوم استفاده از یک روش خوشه‌بندی (نظیر k-means) بر روی بازنمایی‌های یاد گرفته شده از مرحله اول است. سپس، از پارامترهای خود رمزگذار یاد گرفته شده (وزن‌ها و بایاس) و نمایندگان خوشه (که از طریق اجرا k-means بر روی بازنمایی‌های یاد گرفته شده تولید می‌شود) برای تنظیم اولیه پارامترهای خود رمزگذار و نمایندگان خوشه، قبل از آموزش چارچوب خوشه‌بندی عمیق استفاده می‌شود.

یکی از معروف‌ترین الگوریتم‌های این دسته DCN[4] است [۴۳]. این الگوریتم مبتنی بر k-means، یکی از اولین الگوریتم‌های DC مبتنی بر خود رمزگذار است. DCN ابتدا یک خود رمزگذار را آموزش می‌دهد سپس از پارامترهای یاد گرفته شده برای مقداردهی اولیه خود رمزگذار و نمایندگان خوشه استفاده می‌کند. در این روش خطا خود رمزگذار و خطا خوشه‌بندی به روش تناوبی بهینه می‌شود. این الگوریتم ادعا می‌کند k-means دوست است به این معنی که بازنمایی‌های نهایی به یک فضای مناسب برای الگوریتم k-means نزدیک می‌شوند.

برخی روش‌ها سعی دارند، در تولید بازنمایی‌های جدید، ساختار محلی داده‌ها را حفظ کنند که کار بسیار چالش‌برانگیزی است. DEN[5] [۴۴] یکی از این الگوریتم‌ها است که ابتدا سعی می‌کند خود رمزگذار را از قبل آموزش دهد و سپس محدودیت محلی را اعمال می‌کند. IDEC[6][۴۵] روش دیگری است که سعی در حفظ ساختار محلی داده‌ها دارد. این الگوریتم خوشه‌بندی نرم، خود رمزگذار را برای یادگیری بازنمایی‌های خوشه نیز به کار می‌گیرد. علاوه بر این، تابع خطا استفاده شده برای خوشه‌بندی KLD[7] است که بین برچسب‌های نرم و برچسب‌های پیش‌بینی‌شده محاسبه می‌شود. در این الگوریتم تابع خطا نهایی، ترکیبی از خطا خوشه‌بندی و خطا خود رمزگذار است. IDEC نتایج بسیار خوبی را برای انواع مختلف مجموعه‌داده (به‌عنوان مثال تصویر، متن و غیره) به دست آورده است.

DMC[8] [۴۶] یک رویکرد خوشه‌بندی چندمنظوره مبتنی بر یادگیری عمیق است که یک تابع خطا را پیشنهاد می‌کند که شامل سه عنصر است. اولین عنصر تابع خطا خوشه‌بندی است که باعث می‌شود بازنمایی‌های تولید شده مناسب خوشه‌بندی باشد. دوم عنصر تابع خطا خود رمزگذار است که به دستیابی به بازنمایی‌های جدید از داده‌ها کمک می‌کند و سومین عنصر خطای محلی است که بازنمایی‌های تولید شده از داده‌ها را مفید و معنادار می‌کند. روش دیگری نیز در [۴۷] معرفی شده است که با استفاده از نمونه‌های مثبت و منفی برای هر تصویر خوشه‌بندی داده‌ها را انجام می‌دهد. در نظر گرفتن یک نمونه مثبت و یک نمونه منفی بازنمایی‌های ایجادشده را به‌صورت قابل توجهی بهبود داده است[۴۸]. رویکردهای فعلی مذکور به سختی روابط معنایی بین نمونه‌ها (روابط محلی و سراسری) را در نظر می‌گیرند. علاوه بر این، از آنجایی که ویژگی‌های عمیق در لحظه به‌روزرسانی می شوند، تکیه‌بر این روابط بین نمونه‌ها ممکن است نتیجه عکس داشته باشد. برای مقابله با این مسئله، در [۴۹] روشی به نام تطبیق نزدیک‌ترین همسایه برای تطبیق نمونه‌ها با نزدیک‌ترین همسایگانشان از هر دو سطح محلی (دسته‌ای) و سراسری (کلی) ارائه شده است. همسایگان محلی در هر دسته از طریق فضای تعبیه‌شده محاسبه می‌شود و همسایگان سراسری بر اساس کل داده‌ها و فضای تعبیه‌شده محاسبه می‌شوند.

[2] Semi-supervised Similarity Constraint

[3] Semi-supervised Label

[4] Deep Clustering Network

[5] Deep Embedding Network

[6] Improved Deep Embedded Clustering

[7] Kullback–Leibler Divergence

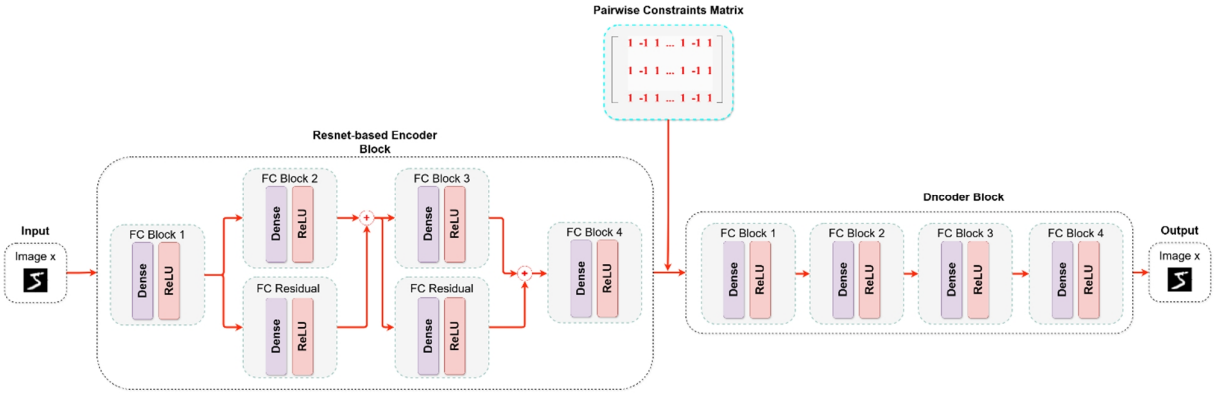
[8] Deep Multi-objective Clustering

۳- روش پیشنهادی

چارچوب خوشه‌بندی نیمه‌نظارت‌شده عمیق پیشنهادی، مشابه با الگوهای خوشه‌بندی مبتنی بر خودرمزگذار، شامل دو مرحله‌ی مجزا است. این بخش تحلیل جامعی از هر یک از این مراحل ارائه می‌دهد. در مرحله‌ی اول، شبکه‌ی خودرمزگذار آموزش داده می‌شود که این گام، بخش اساسی در مدل‌های خوشه‌بندی عمیق است. هدف این مرحله کاهش ابعاد داده و استخراج ویژگی‌های مهم و مرتبط با خوشه‌بندی است. در مرحله‌ی دوم (مرحله‌ی خوشه‌بندی)، از مدل آموزش‌دیده در مرحله‌ی قبل استفاده می‌شود. هدف این گام تعیین برچسب‌های مرتبط با خوشه‌ها است. در طول مرحله‌ی دوم روش پیشنهادی، یک تابع هزینه‌ی نوآورانه به کار گرفته می‌شود.

۳-۱- پیش‌آموزش شبکه

شکل ۱ روش پیشنهادی را نشان می‌دهد. در روش پیشنهادی از معماری مبتنی بر ResNets استفاده می‌شود. این معماری اغلب برای مسائل با ناظر استفاده می‌شوند. روش‌های خوشه‌بندی عمیق بسیار محدودی همراه با اتصالات باقی‌مانده (ResNet) ارائه شده‌اند. در این مقاله، ما یک چارچوب جدید خود رمزگذار مبتنی بر ResNet پیشنهاد می‌کنیم.



شکل ۱- چارچوب خود رمزگذار پیشنهادی. تعداد نودها در لایه‌های رمزگذار FC Block 1، FC Block 2، FC Block 3 و FC Block 4 به ترتیب ۵۰۰، ۵۰۰، ۲۰۰۰ و ۱۰ است. در لایه‌های رمزگشا این ترتیب برعکس است به طوری که در لایه‌های رمزگشا FC Block 1، FC Block 2، FC Block 3 و FC Block 4 به ترتیب ۱۰، ۲۰۰۰، ۵۰۰ و ۵۰۰ است. از اتصالات باقیمانده بین لایه‌های با تعداد نود ۵۰۰ و ۲۰۰۰ استفاده شده است.

روش پیشنهادی شامل یک معماری خود رمزگذار با اتصالات باقیمانده است. مشابه مدل‌های مبتنی بر خود رمزگذار، از دو ماژول مختلف به نام‌های رمزگذار و رمزگشا استفاده می‌شود. ماژول رمزگذار مسئول استخراج بازنمایی‌های تعبیه شده است و ماژول رمزگشا از این بازنمایی های استخراج‌شده برای بازسازی تصویر ورودی x استفاده می‌کند. برخلاف مدل‌های مبتنی بر خود رمزگذار موجود، در این روش، ما از یک بلوک رمزگذار مبتنی بر ResNet برای استخراج بازنمایی‌های تعبیه‌شده استفاده می‌کنیم.

در معماری پیشنهادی، نمونه iام (x_i) از لایه‌های تمام متصل عبور کرده و بازنمایی‌های جدید $(f_\theta(x_i))$ تولید می‌شود (رابطه ۱). در این رابطه، W_x و b_x وزن‌ها و بایاس مربوط به لایه‌هایی است که نمونه x از آن عبور می‌کند و σ نشان‌دهنده تابع فعال‌سازی RELU است.

$$f_\theta(x_i) = \sigma(W_x\, x_i + b_x) \quad (۱)$$

همان‌طور که در شکل ۱ نشان داده شده است، هنگام اضافه کردن یک اتصال باقیمانده، خروجی دو لایه میانی متوالی باید با هم جمع شوند. برای انجام این کار، خروجی هر دو لایه باید در یک بعد باشند؛ بنابراین، باید یک نگاشت غیرخطی از لایه قبلی به ابعادی که با خروجی‌های لایه‌های بعدی مطابقت دارد، انجام می‌شود.

جدول ۱- خلاصه‌ای از معماری پیشنهادی و نوع لایه‌ها

نام لایه	نوع لایه	فرمول ریاضی
Encoder 1	Dense + ReLU	$Eh_1 = \sigma(W_{e1}\, x_i + b_{e1})$
Residual 1	Dense + ReLU	$r_1 = \sigma(W_{r1} Eh_1 + b_{r1})$
Encoder 2	Dense + ReLU	$Eh_2 = \sigma(W_{e2} Eh_1 + b_{e2})$
Encoder 3	Dense + Residual + ReLU	$Eh_3 = \sigma(W_{e3}(Eh_2 + r_1) + b_{e3})$
Residual 2	Dense + ReLU	$r_2 = \sigma(W_{r2} Eh_2 + b_{r2})$
Encoder 4	Dense + Residual + ReLU	$f_\theta(x_i) = \sigma(W_{e4}(Eh_3 + r_2) + b_{e4})$
Clustering Layer	Soft Assignment	$q_{ij} = \frac{\left(1 + \|f_\theta(x_i) - r_j\|^2\right)^{-1}}{\sum_j \left(1 + \|f_\theta(x_i) - r_j\|^2\right)^{-1}}$
Decoder 1	Dense + ReLU	$Dh_1 = \sigma(W_{d1} f_\theta(x_i) + b_{d1})$
Decoder 2	Dense + ReLU	$Dh_2 = \sigma(W_{d2} Dh_1 + b_{d2})$
Decoder 3	Dense + ReLU	$Dh_3 = \sigma(W_{d3} Dh_2 + b_{d3})$
Decoder 4	Dense	$g_\eta \circ f_\theta(x_i) = \sigma(W_{d4} Dh_3 + b_{d4})$

همانطوریکه در شکل ۱، نشان داده شده است، خروجی لایه اول $(h1)$ وارد اولین بلوک ResNet شده و خروجی حاصل از این بلوک ResNet با خروجی لایه دوم $(h2)$ جمع می‌شود. نتیجه حاصل به عنوان ورودی لایه بعدی در نظر گرفته می‌شود. این مکانیسم برای لایه‌های بعدی نیز تکرار می شود. با در نظر گرفتن ۴ لایه در ماژول رمزگذار و ۲ لایه

ResNet اطلاعات و روابط حاصل از معماری پیشنهادی در جدول ۱ خلاصه شده است.

$f_\theta(x_i)$ از لایه‌ی تمام متصل و لایه خوشه‌بندی عبور داده می‌شود. لایه‌های تمام متصل رمزگشا دقیقاً شبیه به لایه‌های تمام متصل رمزگذار است با این تفاوت که ترتیب لایه‌ها برعکس شده است (رابطه ۲).

$$g_\eta \circ f_\theta(x_i) = \sigma(\acute{W}_x f_\theta(x_i) + \acute{b_x}) \quad ۲$$

در رابطه ۲، $g_\eta \circ f_\theta(x_i)$ نسخه بازسازی‌شده نمونه iام به دست آمده از لایه تمام متصل است.

در روش پیشنهادی از تابع خطا پیشنهادی (رابطه ۳) برای آموزش شبکه استفاده می‌کنیم. این تابع خطا ترکیبی از تابع خطا MSE [۹] و SSC است. منظور از MSE میانگین مربع خطا و منظور از SSC محدودیت‌های نیمه نظارت شده است. این تابع زیان مدل را ملزم می سازد که به روابط زوجی میان نقاط داده پایبند باشد. هنگامی که انتظار می‌رود دو نقطه داده به یک خوشه تعلق داشته باشند، این عبارت جریمه‌ای برای فاصله‌های زیاد میان بازنمایی‌های نهفته‌ی آن‌ها اعمال می‌کند و بدین ترتیب نزدیکی در فضای نهفته را تقویت می نماید. برعکس، اگر پیش‌بینی شود که دو نقطه داده در خوشه‌های متفاوت قرار گیرند، این عبارت جریمه‌ای برای فاصله‌های کم اعمال کرده و آن‌ها را به حفظ جدایی قابل‌توجه تشویق می‌کند. این مکانیسم تضمین می‌کند که فرآیند خوشه‌بندی با دانش موجود در مورد روابط میان نقاط داده همخوانی داشته باشد.

تابع خطا روش پیشنهادی از طریق رابطه ۳ محاسبه می‌شود.

$$\mathcal{L}_r = \mathcal{L}_{MSE} + \mathcal{L}_{SSC} \quad ۳$$

تابع خطا بازسازی ($\mathcal{L}_{MSE}$) در این الگوریتم از طریق رابطه ۴ محاسبه می‌شود.

$$\mathcal{L}_{MSE}(\chi, \eta, \theta) = \sum_{x \in \chi} \delta^I(x, g_\eta \circ f_\theta(x)) \quad ۴$$

در رابطه فوق $f_\theta(x)$ نشان‌دهنده داده‌ها در فضای تعبیه‌شده (خروجی رمزگذار) و $g_\eta \circ f_\theta(x)$ نسخه بازسازی‌شده داده‌های به دست آمده از لایه خروجی است (خروجی رمزگشا). δ^I نشان‌دهنده معیار فاصله است که می‌تواند فاصله اقلیدسی، کسینوسی و ... باشد.

تابع خطا $\mathcal{L}_{SSC}$ در این الگوریتم از طریق رابطه ۵ محاسبه می‌شود.

$$\mathcal{L}_{SSC} = \sum_{n=1}^{N} \sum_{\acute{n}=1}^{N} a_{n\acute{n}} \|z_n - z_{\acute{n}}\| \quad ۵$$

در این رابطه منظور از z_n و $z_{\acute{n}}$ به ترتیب نمونه nام و $\acute{n}$ام در فضای تعبیه‌شده (خروجی رمزگذار) است و منظور از $a_{n\acute{n}}$ محدودیت‌های زوجی بین دو نمونه است. برای فرموله‌سازی این محدودیت‌ها ابتدا ماتریس A را معرفی می‌کنیم.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1\acute{n}} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n\acute{n}} \end{bmatrix}_{n \times \acute{n}}$$

در این ماتریس، اگر دو جفت نمونه در یک کلاس باشند مقدار ۱، اگر در کلاس‌های مختلفی باشند مقدار منفی ۱ و در غیر این صورت (برچسب‌دار نباشند) مقدار ۰ در نظر گرفته می‌شود.

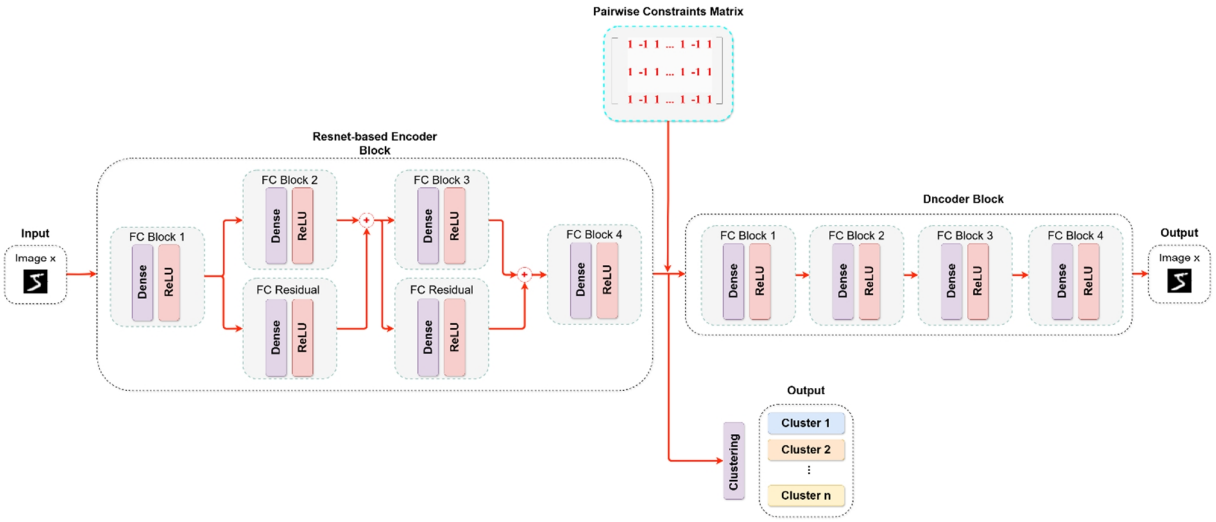
محدودیت‌های زوجی تعیین می‌کنند که آیا یک جفت نمونه داده در یک دسته مشابه یا در دسته‌های متفاوت طبقه‌بندی می‌شوند. انتظار می‌رود نقاط داده‌ای که برچسب یکسان دارند، در نزدیکی یکدیگر قرار بگیرند، در حالی که نقاط داده‌ای که از دسته‌های مختلف منشأ می‌گیرند، در فضای ویژگی نهفته فاصله زیادی از یکدیگر داشته باشند. لازم به ذکر است که ماتریس محدودیت زوجی A قبل از شروع آموزش و بر اساس برچسب‌های اولیه موجود در داده‌ها تشکیل می‌شود. مقدار هر درایه از این ماتریس به‌صورت ثابت بوده و در طول اجرای الگوریتم تغییر نمی‌کند. این امر از نوسانات ناشی از تغییر برچسب‌ها در طول آموزش جلوگیری کرده و ثبات مدل را تضمین می‌نماید.

۳-۲- گام خوشه‌بندی

مهم‌ترین دستاورد روش IDEC در استفاده نوآورانه از توابع زیان KLD و MSE نهفته است. این روش از نقاط داده با اطمینان بالا به‌عنوان یک مکانیسم نظارتی بهره می‌برد و توزیع متمرکزتری از نقاط در هر خوشه ایجاد می‌کند. با این حال، IDEC قادر به استفاده از محدودیت‌های زوجی تعریف‌شده توسط کاربر برای هدایت فرآیند خوشه‌بندی نیست. برای رفع این محدودیت، ما پیشنهاد می‌کنیم که محدودیت‌های زوجی در تابع هدف IDEC ادغام شوند تا فرآیندهای خوشه‌بندی و نهفته‌سازی به‌طور مؤثر هدایت شوند.

در مرحله خوشه‌بندی، از مدل توسعه‌یافته در گام پیشین استفاده می شود، با این تفاوت که عناصر مفهومی جدیدی معرفی شده‌اند: محدودیت‌های زوجی در فرآیند آموزش مدل گنجانده شده‌اند و یک عبارت جریمه به تابع زیان خوشه‌بندی اضافه شده است که امکان خوشه‌بندی نمونه‌ها تحت تأثیر نمونه‌های برچسب‌دار فراهم می‌کند. شکل این معماری در شکل ۲ ارائه شده است.

[9] Mean Squared Error



شکل ۲- چارچوب خوشه‌بندی پیشنهادی

در معماری پیشنهادی خروجی رمزگذار ($f_\theta(x_i)$) به لایه خوشه‌بندی پاس داده می‌شود. وزن‌های روی این لایه مراکز خوشه‌ها را مشخص می‌کند و خروجی این لایه میزان تعلق نمونه iام به خوشه jام را مشخص می‌کند (q_{ij}). رابطه ۶ میزان تعلق نمونه iام به خوشه jام را نشان می‌دهد.

$$q_{ij} = \frac{\left(1 + \|f_\theta(x_i) - r_j\|^2\right)^{-1}}{\sum_j \left(1 + \|f_\theta(x_i) - r_j\|^2\right)^{-1}} \quad ۶$$

در رابطه ۶، $f_\theta(x_i)$ نشان‌دهنده نمونه iام حاصل از خروجی رمزگذار و r_j نشان‌دهنده مرکز خوشه jام است. این رابطه تضمین می‌کند که مجموع تعلقات برای هر نمونه برابر ۱ باشد، یعنی: $\sum_j q_{ij} = 1$.

با استفاده از رابطه ۷ توزیع هدف یا به عبارتی بهتر شبه برچسب‌های واقعی تخمین زده می‌شوند. در هر تکرار این برچسب‌ها محاسبه و بروز می‌شوند. از این شبه برچسب‌ها برای آموزش شبکه استفاده می‌شود.

$$p_{i,j} = \frac{{q_{i,j}}^2/f}{\sum_j ({q_{i,j}}^2/f)} \quad ۷$$

در رابطه ۷، $\sum_i q_{i,j} = f$ است. این تعریف نیز خاصیت نرمال‌سازی را حفظ کرده و تضمین می‌کند که برای هر نمونه i ، رابطه $\sum_j p_{ij} = 1$ برقرار باشد.

مقادیر q_{ij} و $p_{i,j}$ به‌گونه‌ای تعریف شده‌اند که مجموع آن‌ها برای هر نمونه برابر با ۱ باشد و در نتیجه، به‌عنوان توزیع‌های احتمال معتبر برای تعلق نمونه‌ها به خوشه‌ها در نظر گرفته می‌شوند. این نرمال‌سازی برای محاسبه واگرایی KLD بین توزیع نرم پیش‌بینی‌شده و توزیع هدف ضروری است.

در این فاز از روش پیشنهادی از تابع خطا پیشنهادی (رابطه ۸) برای آموزش شبکه استفاده می‌کنیم. این تابع خطا ترکیبی از تابع خطا MSE و SSC در کنار KLD و SSL است.

$$\mathcal{L} = \mathcal{L}_{mse} + \alpha_0 \mathcal{L}_{LD} + \alpha_1 \mathcal{L}_{SSC} + \alpha_2 \mathcal{L}_{SL} \quad ۸$$

در رابطه ۸، α_0 و α_1 و α_2 ضریب هستند. تابع خوشه‌بندی ($\mathcal{L}_{KLD}$) از طریق رابطه ۹ به دست می‌آید. در الگوریتم پیشنهادی از KLD برای تابع خطا خوشه‌بندی استفاده می‌کنیم که بین برچسب‌های محاسبه‌شده نرم (رابطه ۶) و توزیع هدف (رابطه ۷) محاسبه می‌شود.

$$\mathcal{L}_{KLD} = KLD(P||Q) = \sum_i \sum_j p_{ij} \log\left(\frac{q_{ij}}{p_{ij}}\right) \quad ۹$$

تابع زیان برچسب نیمه‌نظارتی ($\mathcal{L}_{SSL}$) این تابع زیان از داده‌های برچسب‌دار برای هدایت فرآیند خوشه‌بندی استفاده می‌کند و تلاش می‌نماید نمونه‌های برچسب‌دار را به‌درستی دسته‌بندی کند. هدف مدل، به حداقل رساندن اختلافات بین شباهت خوشه‌ی پیش‌بینی‌شده برای نقاط داده برچسب‌دار و برچسب‌های واقعی آن‌ها است. این هم‌راستایی تضمین می‌کند که خوشه‌بندی با برچسب‌های شناخته‌شده همخوانی داشته باشد و با فراهم کردن نظارت مبتنی بر داده‌های برچسب‌دار، کیفیت خوشه‌بندی را بهبود می‌بخشد. رابطه ۱۰ این تابع خطا را نشان می‌دهد.

$$\mathcal{L}_{SSL} = \sum_{i=1}^{N} \sum_{j=1}^{K} b_i (q_{ij} - b_i f_{ij})^2 \quad ۱۰$$

در این رابطه $B = [b_i]$ یک بردار N بعدی است که در آن اگر نمونه برچسب دار باشد مقدار یک و در غیر این صورت مقدار صفر خواهد بود. همچنین $F = [f_{ij}]$ یک ماتریس برچسب است که در آن اگر نمونه iام به کلاس jام تعلق داشته باشد مقدار ۱ و در غیر این صورت مقدار ۰ خواهد بود.

شبه کد الگوریتم پیشنهادی در شکل ۳ نشان داد شده است. همان‌طور که در این الگوریتم مشاهده می‌شود توزیع هدف P به‌عنوان «برچسب واقعی» و «برچسب نرم» عمل می‌کند اما این برچسب به خروجی پیش‌بینی‌شده (q_{ij}) بستگی دارد؛ بنابراین، برای جلوگیری از بی‌ثباتی P، نباید در هر تکرار تنها با استفاده از یک دسته داده کوچک به‌روزرسانی شود. برای حل این مشکل، توزیع هدف (p) را با استفاده از

تمام نقاط تعبیه‌شده در هر T تکرار به‌روز می‌کنیم. برای به‌روز رسانی از روابط ۶ و ۷ استفاده می‌کنیم. هنگام به روز رسانی توزیع هدف، برچسب اختصاص داده شده به x_i توسط رابطه ۱۱ به دست می‌آید.

$$s_i = arg_j \, max \; q_{i,j} \qquad ۱۱$$

در این رابطه $q_{i,j}$ از طریق رابطه ۶ محاسبه می‌شود. اگر درصد تغییر برچسب نمونه‌ها بین دو به‌روزرسانی متوالی برای توزیع هدف کمتر از آستانه ε باشد، آموزش را متوقف می‌کنیم.

در هر تکرار، دسته‌ای از داده‌ها با توجه به اندازه دسته انتخاب می‌شود (مرحله ۹). این دسته از داده‌ها برای به‌روزرسانی پارامترهای شبکه (مرحله ۱۰) استفاده می‌شود.

Input: input data $\boldsymbol{\chi}$, Number of clusters K, Target distribution target interval T, stopping threshold ε, Maximum iterations $MaxIter$
Output: Cluster representatives R, Labels S
1: **for** $iter \in 0, 1, \dots, Maxiter$
2: **if** $iter \,\%\, T == 0$
3: Compute the embeddings for all samples
4: Compute target distribution (P) by Eq. (7)
5: Save last label assignments: $s_{old} = s$
6: Compute new label assignments by Eq. (11)
7: **if** $(sum\ (s_{old} \neq s) \,/\, n < \varepsilon)$
8: Stop training
9: Choose a batch of samples $s \in X$
10: Update network parameters on s

شکل ۳- شبه کد الگوریتم پیشنهادی

۴- آزمایش‌ها

در این تحقیق، کارایی روش پیشنهادی با استفاده از چندین مجموعه داده عمومی ارزیابی شده است. نتایج به‌صورت مقایسه‌ای در برابر طیفی از تکنیک‌های خوشه‌بندی عمیق بدون‌نظارت و نیمه‌نظارتی شامل روش‌های زیر قرار گرفته است:

- RICCINET[1] [۵۰],
- ADVDEC[2] [۵۱],
- SSDEC[3] [۳۷],
- DSCPS[4] [۴۱],
- DKMS[5] [۵۲],
- DBC[6] [۵۳],
- IDEC [۴۵],
- SDDC[7] [۵۴],
- ADC[8] [۵۵],
- DERC[9] [۵۶],
- DEEP_CLUSTER [۵۷],
- AutoEmbedder[10] [۴۰],
- JUL[11] [۵۸],
- CRDSSCPC[12] [۵۹],
- RDEIC-LFW-DSS[13] [۱۳],
- EDCWRN[14] [۱۷],
- VADE[15] [۶۰],
- DECCAE[16] [۶۱],
- DEC[17] [۶۲],
- DCDYNAE[18] [۱۰, ۶۳],
- SDKC[19] [۶۴],
- LSMVSC[20] [۶۵],
- DAC[21] [۶۶],
- SC-DEC[22] [۳۹].

شایان ذکر است که نتایج روش‌های مقایسه‌شده در جداول ۷ تا ۱۰، از مقالات اصلی مربوط به هر روش استخراج شده‌اند و الگوریتم آن‌ها به‌صورت مجزا پیاده‌سازی نشده است. این اقتباس با هدف مقایسه عادلانه و استناد به بهترین نتایج گزارش‌شده برای هر روش صورت گرفته است.

برای مقایسه عملکرد روش پیشنهادی، مجموعه‌ای از روش‌های خوشه‌بندی عمیق و نیمه‌نظارتی انتخاب شده‌اند که عملکرد برجسته‌ای در مطالعات پیشین داشته‌اند. از جمله این روش‌ها می‌توان به IDEC (خوشه‌بندی عمیق با حفظ ساختار محلی)، ADC (خوشه‌بندی عمیق

[1] Riemannian Clustering Constrained Inference Network
[2] Adversarial Deep Embedded Clustering
[3] Semi-supervised Deep Embedded Clustering
[4] Deep Semi-supervised Clustering based on Pairwise Constraints and Sample similarity
[5] Deep K-Means with Sampling
[6] Discriminatively Boosted Clustering
[7] Semi-supervised Deep Density Clustering
[8] Active Deep Clustering
[9] Deep Embedded Clustering with Relative Entropy Minimization
[10] Automatic Deep Embedding System for Clustering
[11] Joint Unsupervised Learning
[12] Consistency Regularization for Deep Semi-supervised Clustering with Pairwise Constraints
[13] ResNet-based Deep Embedded Image Clustering with Local Feature Weighting and Dynamic Sample Selection
[14] Efficient Deep Clustering With Representation Weight and Neighbors
[15] Variational Deep Embedding
[16] Deep Embedded Clustering based on Contractive Autoencoder
[17] Deep Embedded Clustering
[18] Deep Clustering with a Dynamic Autoencoder
[19] Scalable Deep K-subspace Clustering
[20] Large-Scale Multi-View Subspace Clustering
[21] Deep Adversarial Clustering
[22] Soft-Constraint Deep Embedded Clustering

خصمانه)، RDEIC-LFW-DSS (خوشه‌بندی تعبیه‌شده مبتنی بر ResNet با وزن‌دهی ویژگی محلی)، AutoEmbedder (سیستم یادگیری تعبیه‌ای نیمه‌نظارتی)، SC-DEC (خوشه‌بندی نیمه‌نظارتی با توزیع هدف)، و سایر روش‌ها مانند VADE، JUL، EDCWRN و CRDSSCPC اشاره کرد. هر یک از این روش‌ها از ساختارهای مختلفی از جمله شبکه‌های خودرمزگذار، یادگیری خصمانه، محدودیت‌های زوجی و یا تنظیمات تعبیه‌ای بهره می‌برند. هدف از مقایسه با این روش‌ها، بررسی جامع کارایی مدل پیشنهادی در برابر تکنیک‌های متنوع و برجسته حوزه خوشه‌بندی عمیق است.

۴-۱- تنظیم پارامترهای روش پیشنهادی

مشابه تنظیمات استفاده‌شده در روش IDEC، ماژول رمزگذار مجموعه‌ای از لایه‌های متراکم (لایه‌های کاملاً متصل) با ابعاد $d-10-2000-500-500$ است که d بعد داده‌های ورودی است و ماژول رمزگشا همان لایه‌های کاملاً متصل ماژول رمزگذار است. با این تفاوت که ترتیب لایه‌ها برعکس شده است ($500-500-2000-10-d$). لازم به ذکر است که چون تصاویر ورودی به‌صورت ماتریس دوبعدی (۲۸× ۲۸ برای مجموعه‌داده‌های MNIST و FASHION و ۱۶× ۱۶ برای مجموعه‌داده USPS) هستند، قبل از ورود به لایه‌های Dense، این ماتریس‌ها به بردار یک‌بعدی (۷۸۴تایی برای مجموعه‌داده‌های MNIST و FASHION و ۲۵۶تایی برای مجموعه‌داده USPS) تبدیل شده و سپس وارد شبکه می‌شوند. همچنین از اتصالات باقیمانده بین لایه‌ها با تعداد نودهای ۵۰۰ و ۲۰۰۰ استفاده شده است. کل لایه‌های داخلی با استفاده از تابع غیرخطی ReLU فعال می‌شوند. ضریب تابع زیان خوشه‌بندی از مجموعه‌ای شامل مقادیر $\{0.1, 0.01, 0.001, 0.001, 0.0001\}$ انتخاب شده است. پس از انجام ارزیابی‌های تجربی، مقدار بهینه برای ضریب تابع زیان خوشه‌بندی برابر با ۰/۱ به‌عنوان مقدار نهایی انتخاب شد، چراکه در مقایسه با سایر مقادیر، پایداری بیشتری در فرآیند آموزش مدل ایجاد کرده و دقت خوشه‌بندی بالاتری به همراه داشت. انتخاب این مقدار باعث شد تا مدل بتواند به‌خوبی تعادل میان اجزای مختلف تابع زیان از جمله بازسازی داده، حفظ محدودیت‌های زوجی و اعمال نظارت جزئی را برقرار کند.

اندازه دسته، فواصل به‌روزرسانی (T) و آستانه همگرایی (ε) به ترتیب روی ۲۵۶، ۱۴۰ و ۰/۰۱ درصد تنظیم شده‌اند. بهینه‌ساز SGD [۱] با نرخ یادگیری [۲] ۰/۱ و تکانه [۳] ۰/۹ در چارچوب‌های پیشنهادی استفاده شده است.

در تمامی آزمایش‌های انجام‌شده، به‌منظور پیاده‌سازی چارچوب نیمه‌نظارتی، ۲۰ درصد از نمونه‌های هر کلاس به‌صورت تصادفی به‌عنوان داده‌های برچسب‌دار انتخاب شده‌اند.

۴-۲- مجموعه داده

برای ارزیابی عملکرد الگوریتم‌های پیشنهادی و انجام مقایسه‌ای جامع با سایر روش‌ها، از چهار مجموعه‌داده‌ی مرجع در حوزه‌ی بینایی ماشین استفاده شده است. این مجموعه‌داده‌ها شامل تصاویر دست‌نویس و پوشاک هستند و به دلیل تنوع و استفاده‌ی گسترده در پژوهش‌های یادگیری ماشین، گزینه‌های مناسبی برای ارزیابی روش‌های پیشنهادی محسوب می‌شوند. در ادامه، مشخصات هر یک از این مجموعه‌داده‌ها به‌طور خلاصه ارائه می‌شود.

- **MNIST:** شامل ۷۰۰۰۰ تصویر سیاه‌وسفید با ابعاد ۲۸×۲۸ پیکسل از ارقام دست‌نویس (۰ تا ۹) است. این مجموعه‌داده یکی از شناخته‌شده‌ترین منابع در حوزه‌ی یادگیری ماشین برای وظایف طبقه‌بندی و خوشه‌بندی محسوب می‌شود. این مجموعه‌داده از طریق لینک http://yann.lecun.com/exdb/mnist در دسترسی است.
- **FASHION-MNIST:** متشکل از ۷۰۰۰۰ تصویر سیاه‌وسفید با ابعاد ۲۸×۲۸ از ۱۰ دسته‌ی مختلف پوشاک است و به‌عنوان جایگزینی چالش‌برانگیز برای مجموعه‌ی MNIST معرفی شده است. این مجموعه‌داده از طریق لینک https://github.com/zalandoresearch/fashion-mnist در دسترسی است.
- **USPS:** این مجموعه شامل ۹،۲۹۸ تصویر سیاه‌وسفید با ابعاد ۱۶×۱۶ از ارقام دست‌نویس (۰ تا ۹) است که توسط اداره‌ی پست ایالات متحده گردآوری شده است. این مجموعه‌داده از طریق لینک https://www.kaggle.com/bistaumanga/usps-dataset در دسترسی است
- **MNIST-TEST:** شامل ۱۰،۰۰۰ تصویر سیاه‌وسفید با ابعاد ۲۸×۲۸ از ارقام دست‌نویس (۰ تا ۹) است و در واقع بخش آزمون مجموعه‌ی MNIST به شمار می آید. این مجموعه‌داده از طریق لینک http://yann.lecun.com/exdb/mnist در دسترسی است.

خلاصه‌ی آماری این مجموعه‌داده‌ها در جدول ۲ ارائه شده است.

جدول ۲- مجموعه‌داده

مجموعه‌داده	تعداد نمونه	تعداد ویژگی	تعداد کلاس
MNIST	۷۰۰۰۰	۷۸۴	۱۰

[۱] Stochastic Gradient Descent
[۲] Learning rate
[۳] Momentum

MNIST_TEST	۱۰۰۰۰	۷۸۴	۱۰
FASHION	۷۰۰۰۰	۷۸۴	۱۰
USPS	۹۲۹۸	۲۵۶	۱۰

۴-۳- معیار ارزیابی

در آزمایش‌ها انجام‌گرفته، از دو معیار دقت خوشه‌بندی و NMI[۱] برای اندازه‌گیری کارایی الگوریتم‌ها استفاده می‌شود. استفاده توأمان از این دو معیار امکان ارزیابی دقیق‌تر و عادلانه‌تری از عملکرد الگوریتم‌ها را فراهم می‌کند. معیارهای مذکور در بهترین حالت، یک و در بدترین حالت صفر هستند. در آزمایش‌ها این معیارها به‌صورت درصد بیان شده‌اند.

دقت خوشه‌بندی: معیار دقت خوشه‌بندی، درصد نمونه‌هایی را اندازه‌گیری می‌کند که به‌درستی به خوشه متناظر با برچسب واقعی خود تخصیص یافته‌اند. این معیار برای مقایسه مستقیم خوشه‌ها با برچسب‌های واقعی، از نگاشت بهینه بین خوشه‌ها و کلاس‌ها استفاده می‌کند [۶۷-۷۰]. این معیار در رابطه ۱۲ نشان داده می‌شود.

$$ACC = \frac{\sum_{k=1}^{K} d_k}{N} \times 100 \qquad ۱۲$$

در رابطه فوق d_k نشان‌دهنده تعداد نمونه‌هایی است که به‌درستی به خوشه K اختصاص داده شده است و N نشان دهنده تعداد کل نمونه‌ها در مجموعه‌داده است.

معیار NMI: NMI معیاری مبتنی بر تئوری اطلاعات است که میزان اطلاعات مشترک بین خوشه‌های تولیدشده و برچسب‌های واقعی را محاسبه می‌کند. این معیار مستقل از ترتیب یا نگاشت خوشه‌ها و کلاس‌ها است و مقدار آن بین ۰ (عدم وابستگی کامل) تا ۱ (هم‌پوشانی کامل) قرار می‌گیرد [۲۱, ۷۱, ۷۲]. این معیار در رابطه ۱۳ تعریف شده است.

$$NMI\ (R,Q) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} p(i,j)\ \log\frac{p(i,j)}{p(i)\,p(j)}}{\sqrt{H(R)H(Q)}} \times 100 \qquad ۱۳$$

در معادله فوق R و Q دو پارتیشن از مجموعه‌داده هستند. فرض می‌کنیم R و Q به ترتیب دارای خوشه‌های I و J باشند. $p(i)$ احتمال این است که یک نمونه تصادفی انتخاب‌شده از مجموعه‌داده‌ها، به خوشه‌ی R_i در پارتیشن R اختصاص داده شود. $p(i,j)$ احتمال اینکه یک نمونه هم به خوشه R_i در R و هم به خوشه Q_i در Q متعلق باشد را نشان می‌دهد. $H(R)$ آنتروپی مرتبط با تمام احتمالات $p(i)$ در پارتیشن R را نشان می‌دهد.

[۱] Normalized Mutual Information

۴-۴- آزمایش ۱: تحلیل تابع خطا محدودیت نیمه‌نظارتی ($\mathcal{L}_{ssc}$)

در این آزمایش، تأثیر عبارت محدودیت نیمه‌نظارتی که در تابع زیان خوشه‌بندی پیشنهادی ادغام شده است، برای ارزیابی تأثیر آن بر کارایی روش پیشنهادی بررسی می‌شود. نتایج در جدول‌های ۳ و ۴ ارائه شده است.

جدول ۳- تأثیر $\mathcal{L}_{ssc}$ معیار دقت

روش	MNIST	FASHION	MNIST_TEST	USPS
با استفاده از $\mathcal{L}_{ssc}$	۹۹،۱	۸۸،۱	۹۹	۹۹
بدون استفاده از $\mathcal{L}_{ssc}$	۹۲،۱	۶۰،۱	۹۴،۶	۹۴،۴

جدول ۴- تأثیر $\mathcal{L}_{ssc}$ معیار NMI

روش	MNIST	FASHION	MNIST_TEST	USPS
با استفاده از $\mathcal{L}_{ssc}$	۹۷،۲	۸۱،۱	۹۶،۵	۹۷،۱
بدون استفاده از $\mathcal{L}_{ssc}$	۹۲،۷	۶۱،۱	۹۲،۲	۹۲،۴

ادغام عبارت محدودیت نیمه‌نظارتی در تابع زیان پیشنهادی، اثربخشی روش پیشنهادی را به‌طور قابل‌توجهی افزایش می‌دهد. این محدودیت باعث بهبود میانگین ۱۱ درصد در دقت خوشه‌بندی و ۹ درصد در اطلاعات متقابل نرمال‌شده (NMI) می‌شود. نتایج چشمگیری در دقت و NMI برای مجموعه داده‌هایی مانند FASHION ثبت شده است. نتایج به‌دست‌آمده نشان می‌دهند که $\mathcal{L}_{ssc}$ به مدل پیشنهادی کمک می‌کند تا نمایش‌هایی را یاد بگیرد که با اهداف خوشه‌بندی همخوانی دارند، یعنی کاهش فاصله درون‌خوشه‌ای و افزایش فاصله بین‌خوشه‌ای.

۴-۵- آزمایش ۲: تحلیل تابع خطا برچسب نیمه‌نظارتی ($\mathcal{L}_{ssL}$)

در این آزمایش، تأثیر عبارت برچسب نیمه‌نظارتی که در تابع زیان خوشه‌بندی پیشنهادی ادغام شده است، برای ارزیابی تأثیر آن بر کارایی روش پیشنهادی بررسی می‌شود. نتایج در جدول‌های ۵ و ۶ ارائه شده است.

جدول ۵- تأثیر $\mathcal{L}_{ssc}$ معیار دقت

روش	MNIST	FASHION	MNIST_TEST	USPS
با استفاده از $\mathcal{L}_{ssc}$	۹۹،۱	۸۸،۱	۹۹	۹۹
بدون استفاده از $\mathcal{L}_{ssc}$	۹۶،۲	۸۲،۴	۹۵،۷	۹۵،۱

جدول ۶- تأثیر $\mathcal{L}_{ssc}$ معیار NMI

روش	MNIST	FASHION	MNIST_TEST	USPS
با استفاده از $\mathcal{L}_{ssc}$	۹۷،۲	۸۱،۱	۹۶،۵	۹۷،۱
بدون استفاده از $\mathcal{L}_{ssc}$	۹۳،۵	۷۶،۷	۹۳،۵	۹۳،۵

ادغام عبارت برچسب نیمه‌نظارتی در تابع خطا پیشنهادی، اثربخشی روش پیشنهادی را به‌طور قابل‌توجهی افزایش می‌دهد. این محدودیت باعث بهبود میانگین ۳ درصد در دقت خوشه‌بندی و ۴ درصد در اطلاعات متقابل نرمال‌شده (NMI) می‌شود. نتایج چشمگیری در دقت و NMI برای مجموعه داده‌هایی مانند FASHION ثبت شده است. نتایج به‌دست‌آمده نشان می‌دهند که پس از $\mathcal{L}_{SSL}$ (آزمایش ۱)، تابع خطا $\mathcal{L}_{SSL}$ تأثیر مثبت بر نتایج دارد.

۴-۶- آزمایش ۳: روش پیشنهادی در مقایسه با روش‌های خوشه‌بندی

در این تحقیق، کارایی روش پیشنهادی با روش‌های خوشه‌بندی عمیق بدون نظارت مقایسه شده است. نتایج کمی در جداول ۷ و ۸ ارائه شده‌اند. این جداول نشان می‌دهند که روش پیشنهادی به‌طور مداوم عملکرد بهتری نسبت به سایر روش‌ها در تمامی مجموعه‌داده‌ها دارد. به‌ویژه، معیارهای عملکرد روش پیشنهادی نسبت به روش‌های شناخته‌شده‌ای مانند RDEIC-LFW-DSS، EDCWRN، DCDYNAE و ADVDEC که به ترتیب پس از آن قرار می‌گیرند، برتری قابل‌توجهی دارند.

جدول ۷- دقت روش پیشنهادی در مقایسه با روش‌های خوشه‌بندی

روش	USPS	MNIST_TEST	*FASHION*	MNIST
VADE	۵۶،۶	۲۸،۷	۵۷،۸	۹۴،۵
SDKC	-	۸۳،۳	۶۰،۰	۸۷،۱
RDEIC-LFW-DSS	۹۸،۳	۹۸،۷	۶۴،۲	۹۸،۸
LSMVSC	۶۷،۵	۵۷،۲	۵۱،۲	۵۵،۷
JUL	۹۵،۰	۹۶،۱	۵۶،۳	۹۶،۴
IDEC	۷۶،۱	۸۴،۶	۵۲،۹	۸۸،۱
EDCWRN	۹۸،۱	۹۸،۳	۶۲،۲	۹۸،۵
DKMS	۷۷،۰	-	-	۸۶،۲
DERC	۹۶،۴	۹۶،۵	۳۹،۲	۹۶،۵
DEEP_CIUSTER	۵۶،۲	۸۵،۴	۵۴،۲	۷۹،۷
DECCAE	۷۷،۳۱	–	۶۰،۹	۹۶،۳
DEC	۷۶،۲	۸۵،۶	۵۱،۸	۸۶،۳
DCDYNAE	۹۸،۱	۹۸،۷	۵۹،۱	۹۸،۷
DBC	۷۴،۳	-	-	۹۶،۴
DAC	-	۸۰،۴	–	۸۰،۱
ADVDEC	۹۸،۱	۹۸،۵	۵۸،۶	۹۸،۶
RICCINET	۸۱،۱	–	–	۹۱،۵۶
روش پیشنهادی	۹۹	۹۹	۸۸،۱	۹۹،۱

جدول ۸- NMI روش پیشنهادی در مقایسه با روش‌های خوشه‌بندی

روش	USPS	MNIST_TEST	*FASHION*	MNIST
VADE	۵۱،۲	۲۸،۷	۶۳،۰	۸۷،۶
SDKC	-	۷۷،۴	۶۲،۳	۷۸،۱
RDEIC-LFW-DSS	۹۵،۲	۹۶،۳	۶۸،۳	۹۶،۶
LSMVSC	۶۲،۱	۵۱،۶	۵۱،۹	۵۰،۳
JUL	۹۱،۳	۹۱،۳	۶۰،۸	۹۱،۳
IDEC	۷۸،۵	۸۰،۲	۵۵،۷	۸۶،۷
EDCWRN	۹۵،۰	۹۶،۲	۶۸،۰	۹۶،۲
DKMS	۷۸،۷	-	-	۸۰،۵
DERC	۹۲،۷	۹۱،۵	۳۹،۲	۹۱،۷
DEEP_CIUSTER	۵۴،۰	۷۱،۳	۵۱،۰	۶۶،۱
DECCAE	۸۰،۵۳	–	۶۶،۹	۹۰،۷
DEC	۷۶،۷	۸۳،۰	۵۴،۶	۸۳،۴

DCDYNAE	۹۴،۸	۹۶،۳	۶۴،۲	۹۶،۴
DBC	۷۲،۴	-	-	۹۱،۷
DAC	-	۷۸،۰	–	۷۸،۴
ADVDEC	۹۴،۸	۹۵،۷	۶۶،۲	۹۶،۱
RICCINET	۸۲،۱۰	-	-	۹۱،۲۴
روش پیشنهادی	۹۷،۱	۹۶،۵	۸۱،۱	۹۷،۲

۴-۷- آزمایش ۴: روش پیشنهادی در مقایسه با روش‌های نیمه‌نظارت شده خوشه‌بندی

کارایی روش پیشنهادی با روش‌های نیمه‌نظارت‌شده خوشه‌بندی عمیق نیز مورد ارزیابی قرار گرفته است. نتایج ارائه‌شده در جداول ۹ و ۱۰ نشان می‌دهند که روش پیشنهادی معیارهای عملکردی بالاتر از تمامی روش‌های دیگر قرار دارد. در بین سایر روش‌ها، CRDSSCPC و ADC عملکرد قابل‌توجهی داشته‌اند، اما همچنان نسبت به روش پیشنهادی پایین‌تر هستند.

جدول ۹- دقت روش پیشنهادی در مقایسه با روش‌های نیمه‌نظارت‌شده خوشه‌بندی عمیق

روش	USPS	MNIST_TEST	*FASHION*	MNIST
SSDEC	۷۶،۳۹	۸۶،۱	–	۸۶،۱۱
ADC	۹۸،۸۰	–	۸۸،۰	۹۹،۶
SC-DEC	۷۶،۶۲	–	–	۹۲،۱۱
AutoEmbedder	–	–	۹۱،۹	۹۸،۴
DSCPS	۷۷،۳	–	۶۲،۳	۹۷،۹
SDDC	۹۶،۷	۹۶،۱	۶۲،۶	
CRDSSCPC	۹۹،۱		۸۵،۷	۹۹،۳
روش پیشنهادی	۹۹	۹۹	۸۸،۱	۹۹،۱

جدول ۱۰- NMI روش پیشنهادی در مقایسه با روش‌های نیمه‌نظارت‌شده خوشه‌بندی عمیق

روش	USPS	MNIST_TEST	*FASHION*	MNIST
SSDEC	۷۷،۶۸	۸۲،۹	–	۸۲،۸۹
ADC	۹۶،۵۰	–	۸۲	۹۸،۷

SC-DEC	۷۵،۳۱	–	–	۸۴،۰۱
AutoEmbedder	–	–	۸۷،۰	۹۵،۰
DSCPS	۷۲،۱	–	۶۰،۸	۹۴،۲
SDDC	۹۱،۷	۹۰،۷	۶۳،۳	
CRDSSCPC	۹۷،۴	–	۷۹،۱	۹۷،۸
روش پیشنهادی	۹۷،۱	۹۶،۵	۸۱،۱	۹۷،۲

۵- نتیجه‌گیری

در این مقاله، یک روش نوآورانه برای خوشه‌بندی نیمه‌نظارتی مبتنی بر یادگیری عمیق ارائه شد که ترکیبی از معماری خودرمزگذار مبتنی بر ResNet، محدودیت‌های زوجی نیمه‌نظارتی و یک تابع زیان ترکیبی چندمنظوره است. روش پیشنهادی با هدف بهره‌برداری از اطلاعات برچسب‌های محدود و ساختار نهفته داده‌ها طراحی شده و توانسته است عملکرد قابل توجهی در مقایسه با روش‌های پیشین ارائه دهد. ارزیابی روش پیشنهادی بر روی مجموعه‌داده‌های متنوع، نشان‌دهنده کارایی و انعطاف‌پذیری بالای آن است. به‌ویژه، استفاده از توزیع هدف برای تنظیم شبه‌برچسب‌ها، نقش کلیدی در بهبود کیفیت خوشه‌بندی ایفا کرده است. همچنین، حفظ روابط محلی و سراسری در فضای تعبیه‌شده، منجر به تشکیل خوشه‌های معنادار و سازگار با ساختار داده‌ها شده است. روش پیشنهادی علاوه بر ارائه دقت بالا در مسائل خوشه‌بندی، پتانسیل زیادی برای کاربرد در حوزه‌های مختلف مانند تحلیل تصاویر، داده‌های پزشکی و پردازش زبان طبیعی دارد. با این حال، یکی از محدودیت‌های اصلی این روش، پیچیدگی محاسباتی آن در مجموعه‌داده‌های بسیار بزرگ است که می‌تواند در تحقیقات آینده مورد توجه قرار گیرد. در آینده، توسعه و بهینه‌سازی این مدل برای داده‌های حجیم و بررسی تأثیر استفاده از معماری‌های عمیق دیگر می‌تواند به بهبود عملکرد و گسترش دامنه کاربرد این روش منجر شود. همچنین، بررسی رویکردهای ترکیبی دیگر برای ادغام اطلاعات نظارتی و غیرنظارتی، افق‌های جدیدی را در زمینه خوشه‌بندی مبتنی بر یادگیری عمیق باز خواهد کرد. این مطالعه گامی مؤثر در جهت پیشرفت روش‌های خوشه‌بندی عمیق و نیمه‌نظارتی بوده و ابزار ارزشمندی برای تحلیل داده‌های پیچیده فراهم می‌کند.

مراجع

[1] A. Golzari Oskouei, N. Samadi, and J. Tanha, "Feature-weight and cluster-weight learning in fuzzy c-means method for semi-supervised clusteringImage 1," *Applied Soft Computing,* vol. 161, p. 111712, 2024/08/01/ 2024, doi: https://doi.org/10.1016/j.asoc.2024.111712.

[2] S. Hu, G. Zou, C. Zhang, Z. Lou, R. Geng, and Y. Ye, "Joint contrastive triple-learning for deep multi-view clustering," *Information Processing & Management,* vol. 60, no. 3, p. 103284, 2023/05/01/ 2023, doi: https://doi.org/10.1016/j.ipm.2023.103284.

[3] Y. Wang, J. Zou, K. Wang, C. Liu, and X. Yuan, "Semi-supervised deep embedded clustering with pairwise constraints and subset allocation," *Neural Networks,* vol. 164, pp. 310-322, 2023/07/01/ 2023, doi: https://doi.org/10.1016/j.neunet.2023.04.016.

[4] A. Golzari Oskouei, N. Samadi, J. Tanha, A. Bouyer, and B. Arasteh, "Viewpoint□Based Collaborative Feature□Weighted Multi□View Intuitionistic Fuzzy Clustering Using Neighborhood Information," *Neurocomputing,* vol. 617, p. 128884, 2025/02/07/ 2025, doi: https://doi.org/10.1016/j.neucom.2024.128884.

[5] A. G. Oskouei, N. Samadi, J. Tanha, and A. Bouyer, "SSFCM-FWCW: Semi-Supervised Fuzzy C-Means method based on Feature-Weight and Cluster-Weight learning," *Software Impacts,* vol. 21, p. 100678, 2024/09/01/ 2024, doi: https://doi.org/10.1016/j.simpa.2024.100678.

[6] M. A. Balafar, R. Hazratgholizadeh, and M. R. F. Derakhshi, "Active Learning for Constrained Document Clustering with Uncertainty Region," *Complexity,* vol. 2020, p. 3207306, 2020/05/20 2020, doi: 10.1155/2020/3207306.

[7] K. Berahmand, E. Nasiri, R. Pir mohammadiani, and Y. Li, "Spectral clustering on protein-protein interaction networks via constructing affinity matrix using attributed graph embedding," *Computers in Biology and Medicine,* vol. 138, p. 104933, 2021/11/01/ 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104933.

[8] C. Hu, T. Wu, S. Liu, C. Liu, T. Ma, and F. Yang, "Joint unsupervised contrastive learning and robust GMM for text clustering," *Information Processing & Management,* vol. 61, no. 1, p. 103529, 2024/01/01/ 2024, doi: https://doi.org/10.1016/j.ipm.2023.103529.

[9] J.-H. Chang and Y.-C. Leung, "Dynamic image clustering from projected coordinates of deep similarity learning," *Neural Networks,* vol. 152, pp. 1-16, 2022/08/01/ 2022, doi: https://doi.org/10.1016/j.neunet.2022.03.030.

[10] N. Mrabah, N. M. Khan, R. Ksantini, and Z. Lachiri, "Deep clustering with a dynamic autoencoder: From reconstruction towards centroids construction," *Neural Networks,* vol. 130, pp. 206-228, 2020.

[11] A. Bouyer, H. A. Beni, A. G. Oskouei, A. Rouhi, B. Arasteh, and X. Liu, "Maximizing Influence in Social Networks Using Combined Local Features and Deep Learning-Based Node Embedding," *Big Data,* vol. 0, no. 0, p. null, doi: 10.1089/big.2023.0117.

[12] K. Berahmand, A. Bouyer, and M. Vasighi, "Community Detection in Complex Networks by Detecting and Expanding Core Nodes Through Extended Local Similarity of Nodes," *IEEE Transactions on Computational Social Systems,* vol. 5, no. 4, pp. 1021-1033, 2018, doi: 10.1109/TCSS.2018.2879494.

[13] A. Golzari Oskouei, M. A. Balafar, and C. Motamed, "RDEIC-LFW-DSS: ResNet-based deep embedded image clustering using local feature weighting and dynamic sample selection mechanism," *Information Sciences,* vol. 646, p. 119374, 2023/10/01/ 2023, doi: https://doi.org/10.1016/j.ins.2023.119374.

[14] M. Hashemzadeh, A. Golzari Oskouei, and N. Farajzadeh, "New fuzzy C-means clustering method based on feature-weight and cluster-weight learning," *Applied Soft Computing,* vol. 78, pp. 324-345, 2019/05/01/ 2019, doi: https://doi.org/10.1016/j.asoc.2019.02.038.

[15] J. C. Bezdek, "Objective function clustering," in *Pattern recognition with fuzzy objective function algorithms*: Springer, 1981, pp. 43-93.

[16] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967, vol. 1, no. 14: Oakland, CA, USA, pp. 281-297.

[17] A. Golzari Oskouei, M. A. Balafar, and C. Motamed, "EDCWRN: efficient deep clustering with the weight of representations and the help of neighbors," *Applied Intelligence,* 2022/07/05 2022, doi: 10.1007/s10489-022-03895-5.

[18] A. Golzari Oskouei and M. Hashemzadeh, "CGFFCM: A color image segmentation method based on cluster-weight and feature-weight learning," *Software Impacts,* vol. 11, p. 100228, 2022/02/01/ 2022, doi: https://doi.org/10.1016/j.simpa.2022.100228.

[19] T. Qi, X. Feng, W. Wang, and X. Li, "Game theory based Bi-domanial deep subspace clustering," *Information Sciences,* vol. 617, pp. 150-164, 2022/12/01/ 2022, doi: https://doi.org/10.1016/j.ins.2022.10.067.

[20] C. Biswas, D. Ganguly, D. Roy, and U. Bhattacharya, "Weakly supervised deep metric learning on discrete metric spaces for privacy-preserved clustering," *Information Processing & Management,* vol. 60, no. 1, p. 103109, 2023/01/01/ 2023, doi: https://doi.org/10.1016/j.ipm.2022.103109.

[21] A. Golzari Oskouei, M. Hashemzadeh, B. Asheghi, and M. A. Balafar, "CGFFCM: Cluster-weight and Group-local Feature-weight learning in Fuzzy C-Means clustering algorithm for color image segmentation," *Applied Soft Computing,* vol. 113, p. 108005, 2021/12/01/ 2021, doi: https://doi.org/10.1016/j.asoc.2021.108005.

[22] A. Golzari Oskouei, M. A. Balafar, and C. Motamed, "FKMAWCW: Categorical fuzzy k-modes clustering with automated attribute-weight and cluster-weight learning," *Chaos, Solitons & Fractals,* vol. 153, p. 111494, 2021/12/01/ 2021, doi: https://doi.org/10.1016/j.chaos.2021.111494.

[23] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems,* vol. 2, no. 1, pp. 37-52, 1987/08/01/ 1987, doi: https://doi.org/10.1016/0169-7439(87)80084-9.

[24] C. Du, L. Tian, B. Chen, L. Zhang, W. Chen, and H. Liu, "Region-factorized recurrent attentional network with deep clustering for radar HRRP target recognition," *Signal Processing,* vol. 183, p. 108010, 2021/06/01/ 2021, doi: https://doi.org/10.1016/j.sigpro.2021.108010.

[25] S. Huang, Z. Kang, Z. Xu, and Q. Liu, "Robust deep k-means: An effective and simple method for data clustering," *Pattern Recognition,* vol. 117, p. 107996, 2021/09/01/ 2021, doi: https://doi.org/10.1016/j.patcog.2021.107996.

[26] Z. Khan and J. Yang, "Bottom-up unsupervised image segmentation using FC-Dense u-net based deep representation clustering and multidimensional feature fusion based region merging," *Image and Vision Computing,* vol. 94, p. 103871, 2020/02/01/ 2020, doi: https://doi.org/10.1016/j.imavis.2020.103871.

[27] Q. Zhou, W. a. Zhou, and S. Wang, "Cluster adaptation networks for unsupervised domain adaptation," *Image and Vision Computing,* vol. 108, p. 104137, 2021/04/01/ 2021, doi: https://doi.org/10.1016/j.imavis.2021.104137.

[28] M. Asgari-Chenaghlu, M. R. Feizi-Derakhshi, L. Farzinvash, M. A. Balafar, and C. Motamed, "CWI: A multimodal deep learning approach for named entity recognition from social media using character, word and image features," *Neural Computing and Applications,* vol. 34, no. 3, pp. 1905-1922, 2022/02/01 2022, doi: 10.1007/s00521-021-06488-4.

[29] C. Buizza *et al.*, "Data Learning: Integrating Data Assimilation and Machine Learning," *Journal of Computational Science,* vol. 58, p. 101525, 2022/02/01/ 2022, doi: https://doi.org/10.1016/j.jocs.2021.101525.

[30] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks,* vol. 61, pp. 85-117,

[47] Z. Dang, C. Deng, X. Yang, and H. Huang, "Doubly Contrastive Deep Clustering," *arXiv preprint arXiv:2103.05484,* 2021.

[48] R. sarhan, M. A. Balafar, M. R. Feizi Derakhshi, and a. Golzari Oskouei, "Image classification based on unsupervised adversarial transfer learning and preserving theinter-class and intra-class distance," *Advanced Signal Processing,* pp. -, 2024, doi: 10.22034/jasp.2024.59731.1241.

[49] Z. Dang, C. Deng, X. Yang, K. Wei, and H. Huang, "Nearest neighbor matching for deep clustering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 13693-13702.

[50] L. Sun *et al.*, "Riccinet: Deep clustering via a riemannian generative model," in *Proceedings of the ACM on Web Conference 2024*, 2024, pp. 4071-4082.

[51] N. Mrabah, M. Bouguessa, and R. Ksantini, "Adversarial deep embedded clustering: on a better trade-off between feature randomness and feature drift," *IEEE Transactions on Knowledge and Data Engineering,* 2020.

[52] M. Moradi Fard, T. Thonet, and E. Gaussier, "Deep k-Means: Jointly clustering with k-Means and learning representations," *Pattern Recognition Letters,* vol. 138, pp. 185-192, 2020/10/01/ 2020, doi: https://doi.org/10.1016/j.patrec.2020.07.028.

[53] F. Li, H. Qiao, and B. Zhang, "Discriminatively boosted image clustering with fully convolutional auto-encoders," *Pattern Recognition,* vol. 83, pp. 161-173, 2018/11/01/ 2018, doi: https://doi.org/10.1016/j.patcog.2018.05.019.

[54] X. Xu, H. Hou, and S. Ding, "Semi-supervised deep density clustering," *Applied Soft Computing,* vol. 148, p. 110903, 2023/11/01/ 2023, doi: https://doi.org/10.1016/j.asoc.2023.110903.

[55] B. Sun, P. Zhou, L. Du, and X. Li, "Active deep image clustering," *Knowledge-Based Systems,* vol. 252, p. 109346, 2022/09/27/ 2022, doi: https://doi.org/10.1016/j.knosys.2022.109346.

[56] K. Ghasedi Dizaji, A. Herandi, C. Deng, W. Cai, and H. Huang, "Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5736-5745.

[57] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 132-149.

[58] J. Yang, D. Parikh, and D. Batra, "Joint unsupervised learning of deep representations and image clusters," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5147-5156.

[59] D. Huang, J. Hu, T. Li, S. Du, and H. Chen, "Consistency regularization for deep semi-supervised clustering with pairwise constraints," *International Journal of Machine Learning and Cybernetics,* vol. 13, no. 11, pp. 3359-3372, 2022/11/01 2022, doi: 10.1007/s13042-022-01599-3.

[60] Z. Jiang, Y. Zheng, H. Tan, B. Tang, and H. Zhou, "Variational deep embedding: An unsupervised and generative approach to clustering," *arXiv preprint arXiv:1611.05148,* 2016.

[61] B. Diallo, J. Hu, T. Li, G. A. Khan, X. Liang, and Y. Zhao, "Deep embedding clustering based on contractive autoencoder," *Neurocomputing,* vol. 433, pp. 96-107, 2021/04/14/ 2021, doi: https://doi.org/10.1016/j.neucom.2020.12.094.

[62] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *International conference on machine learning*, 2016: PMLR, pp. 478-487.

[63] N. Mrabah, N. M. Khan, R. Ksantini, and Z. Lachiri, "Deep clustering with a dynamic autoencoder," *arXiv preprint arXiv:1901.07752,* 2019.

[64] T. Zhang, P. Ji, M. Harandi, R. Hartley, and I. Reid, "Scalable deep k-subspace clustering," in *Asian Conference on Computer Vision*, 2018: Springer, pp. 466-481.

[65] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, and Z. Xu, "Large-scale multi-view subspace clustering in linear time," 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.neunet.2014.09.003.

[31] M. Aria, E. Nourani, and A. Golzari Oskouei, "ADA-COVID: Adversarial Deep Domain Adaptation-Based Diagnosis of COVID-19 from Lung CT Scans Using Triplet Embeddings," *Computational Intelligence and Neuroscience,* vol. 2022, p. 2564022, 2022/02/08 2022, doi: 10.1155/2022/2564022.

[32] W. Hu, C. Chen, F. Ye, Z. Zheng, and Y. Du, "Learning deep discriminative representations with pseudo supervision for image clustering," *Information Sciences,* vol. 568, pp. 199-215, 2021/08/01/ 2021, doi: https://doi.org/10.1016/j.ins.2021.03.066.

[33] J. R. Hershey, Z. Chen, J. L. Roux, and S. Watanabe, "Deep clustering: Discriminative embeddings for segmentation and separation," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 20-25 March 2016 2016, pp. 31-35, doi: 10.1109/ICASSP.2016.7471631.

[34] M. Xie, Z. Ye, G. Pan, and X. Liu, "Incomplete multi-view subspace clustering with adaptive instance-sample mapping and deep feature fusion," *Applied Intelligence,* vol. 51, no. 8, pp. 5584-5597, 2021.

[35] J. Xu, Y. Ren, G. Li, L. Pan, C. Zhu, and Z. Xu, "Deep embedded multi-view clustering with collaborative training," *Information Sciences,* vol. 573, pp. 279-290, 2021/09/01/ 2021, doi: https://doi.org/10.1016/j.ins.2020.12.073.

[36] D. Das, R. Ghosh, and B. Bhowmick, "Deep Representation Learning Characterized by Inter-Class Separation for Image Clustering," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 7-11 Jan. 2019 2019, pp. 628-637, doi: 10.1109/WACV.2019.00072.

[37] Y. Ren, K. Hu, X. Dai, L. Pan, S. C. H. Hoi, and Z. Xu, "Semi-supervised deep embedded clustering," *Neurocomputing,* vol. 325, pp. 121-130, 2019/01/24/ 2019, doi: https://doi.org/10.1016/j.neucom.2018.10.016.

[38] J. Tanha, N. Samadi, Y. Abdi, and N. Razzaghi-Asl, "CPSSDS: conformal prediction for semi-supervised classification on data streams," *Information Sciences,* vol. 584, pp. 212-234, 2022.

[39] M. S. AlZuhair, M. M. Ben Ismail, and O. Bchir, "Soft Semi-Supervised Deep Learning-Based Clustering," *Applied Sciences,* vol. 13, no. 17, p. 9673, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/17/9673.

[40] A. Q. Ohi, M. F. Mridha, F. B. Safir, M. A. Hamid, and M. M. Monowar, "AutoEmbedder: A semi-supervised DNN embedding system for clustering," *Knowledge-Based Systems,* vol. 204, p. 106190, 2020/09/27/ 2020, doi: https://doi.org/10.1016/j.knosys.2020.106190.

[41] X. Qin, C. Yuan, J. Jiang, and L. Chen, "Deep semi-supervised clustering based on pairwise constraints and sample similarity," *Pattern Recognition Letters,* vol. 178, pp. 1-6, 2024/02/01/ 2024, doi: https://doi.org/10.1016/j.patrec.2023.12.010.

[42] D. Erhan, A. Courville, Y. Bengio, and P. Vincent, "Why does unsupervised pre-training help deep learning?," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010: JMLR Workshop and Conference Proceedings, pp. 201-208.

[43] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, "Towards k-means-friendly spaces: Simultaneous deep learning and clustering," in *international conference on machine learning*, 2017: PMLR, pp. 3861-3870.

[44] P. Huang, Y. Huang, W. Wang, and L. Wang, "Deep embedding network for clustering," in *2014 22nd International conference on pattern recognition*, 2014: IEEE, pp. 1532-1537.

[45] X. Guo, L. Gao, X. Liu, and J. Yin, "Improved Deep Embedded Clustering with Local Structure Preservation," in *Ijcai*, 2017, pp. 1753-1759.

[46] D. Chen, J. Lv, and Y. Zhang, "Unsupervised multi-manifold clustering by learning deep representation," in *Workshops at the thirty-first AAAI conference on artificial intelligence*, 2017.

in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, vol. 34, no. 04, pp. 4412-4419.

[66] P. Zhou, Y. Hou, and J. Feng, "Deep adversarial subspace clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1596-1604.

[67] A. Golzari Oskouei, N. Abdolmaleki, A. Bouyer, B. Arasteh, and K. Shirini, "Efficient superpixel-based brain MRI segmentation using multi-scale morphological gradient reconstruction and quantum clustering," *Biomedical Signal Processing and Control,* vol. 100, p. 107063, 2025/02/01/ 2025, doi: https://doi.org/10.1016/j.bspc.2024.107063.

[68] A. G. Oskouei, M. A. Balafar, and T. Akan, "A Brain MRI Segmentation Method Using Feature Weighting and a Combination of Efficient Visual Features," in *Applied Computer Vision and Soft Computing with Interpretable AI*: Chapman and Hall/CRC, 2024, pp. 15-34.

[69] M. Aria, E. Nourani, and A. Golzari Oskouei, "ADA-COVID: Adversarial Deep Domain Adaptation-Based Diagnosis of COVID-19 from Lung CT Scans Using Triplet Embeddings," *Computational Intelligence and Neuroscience,* vol. 2022, no. 1, p. 2564022, 2022, doi: https://doi.org/10.1155/2022/2564022.

[70] K. Berahmand, F. Daneshfar, M. Dorosti, and M. J. Aghajani, "An Improved Deep Text Clustering via Local Manifold of an Autoencoder Embedding," 2022.

[71] A. Strehl and J. Ghosh, "Cluster ensembles---a knowledge reuse framework for combining multiple partitions," *Journal of machine learning research,* vol. 3, no. Dec, pp. 583-617, 2002.

[72] T. Akan, A. G. Oskouei, S. Alp, and M. A. N. Bhuiyan, "Brain magnetic resonance image (MRI) segmentation using multimodal optimization," *Multimedia Tools and Applications,* 2024/07/02 2024, doi: 10.1007/s11042-024-19725-4.