مقاله پژوهشی

HMCNet: یک چارچوب نوین برای خوشه‌بندی چندنما در شرایط ناقص بودن نماها

سندس بهادری

چکیده: خوشه‌بندی داده‌های چندنما در حضور نماهای ناقص یکی از چالش‌های مهم در یادگیری ماشین و داده‌کاوی است. روش‌های موجود غالباً در مواجهه با نرخ‌های بالای ناقص بودن نماها دچار افت عملکرد می‌شوند و قادر به استخراج ویژگی‌های مشترک و سازگار میان نماها نیستند. با توجه به آمار و نمونه‌های عینی بیان‌شده که نشان‌دهنده شیوع بالای ناقص‌بودن نماها در کاربردهای واقعی (با نرخ‌های بین ۱۰٪ تا ۵۰٪) و افت شدید عملکرد روش‌های سنتی در این شرایط است، در این پژوهش مدل HMCNet معرفی شده است. این مدل با بهره‌گیری از ترکیب ماژول توجه، کاهش ابعاد غیرخطی UMAP و تابع زیان ترکیبی مبتنی بر احتمال، قادر است ساختارهای محلی داده‌ها را حفظ کرده و سازگاری میان نماها را حتی در حضور نرخ‌های بالای ناقص‌بودن (تا ۵۰٪) تقویت نماید. برای ارزیابی، آزمایش‌های گسترده‌ای بر روی دادگان‌ها انجام شد. نتایج نشان داد که مدل پیشنهادی در معیارهای ACC، NMI و Purity به‌طور پایدار عملکردی برتر نسبت به روش‌های پیشرفته دارد. همچنین تحلیل‌های آزمایش‌های حذف اهمیت هر مؤلفه را در بهبود عملکرد مدل تأیید کردند. به‌طور کلی، یافته‌های این پژوهش نشان می‌دهند که HMCNet می‌تواند به‌عنوان یک چارچوب عمومی و قدرتمند برای خوشه‌بندی داده‌های چندنما در شرایط ناقص بودن نماها مورد استفاده قرار گیرد و مسیرهای جدیدی برای کاربردهای واقعی در حوزه‌های مختلف علمی و صنعتی فراهم سازد.

***کلیدواژه:* تابع زیان احتمالی، خوشه‌بندی چندنما، ماژول توجه، نماهای ناقص، نهفته‌سازی مبتنی بر کاهش ابعاد غیرخطی.**

۱- مقدمه

در عصر حاضر، داده‌های چندنمایی به‌عنوان یکی از مهم‌ترین منابع اطلاعاتی در حوزه‌های مختلف علمی و صنعتی شناخته می‌شوند [۱] و [۲]. این داده‌ها از یک نمونه منفرد، اما از زوایای گوناگون و نماهای متفاوت جمع‌آوری می‌شوند؛ برای مثال، یک تصویر می‌تواند با ویژگی‌های مختلفی مانند هیستوگرام گرادیان‌ها، توصیف‌گرهای موجک، یا ویژگی‌های مبتنی بر رنگ نمایش داده شود. همچنین در حوزه زیست‌پزشکی، یک نمونه ژنتیکی ممکن است هم داده‌های تصویری و هم داده‌های متنی داشته باشد. وجود نماهای مختلف باعث می‌شود که اطلاعات مکمل و متنوعی در اختیار پژوهشگران قرار گیرد و در نتیجه، تحلیل و خوشه‌بندی داده‌ها دقیق‌تر و جامع‌تر انجام شود.

با وجود این مزایا، یکی از چالش‌های اساسی در تحلیل داده‌های چندنمایی، ناقص بودن نماها است. در بسیاری از کاربردهای واقعی، همه نماها برای همه نمونه‌ها در دسترس نیستند؛ برای درک بهتر ابعاد این چالش، بررسی دادگان‌های استاندارد در حوزه خوشه‌بندی چندنما نشان می‌دهد که نرخ ناقص‌بودن نماها می‌تواند به‌طور قابل‌توجهی متغیر باشد. برای مثال، در دادگان Caltech-۱۰۱ که شامل تصاویر از شش نمای مختلف (مانند ویژگی‌های گابور، ویولت و توصیفگرهای محلی) است، بسته به روش نمونه‌برداری، بین ۱۰٪ تا ۳۰٪ از نماها ممکن است مفقود باشند. در دادگان YouTube Faces که برای تشخیص چهره از چندین فریم و زاویه مختلف استفاده می‌کند، نرخ فقدان داده در برخی نماها به بیش از ۴۰٪ می‌رسد . در حوزه داده‌های زیستی، دادگان TCGA که شامل داده‌های ژنومی، پروتئینی و بالینی از بیماران سرطانی است، به‌دلیل محدودیت‌های آزمایشگاهی و هزینه‌های بالا، به‌طور میانگین حدود ۲۰٪ تا ۳۵٪ از داده‌ها در یک یا چند نما مفقود هستند . حتی در دادگان‌های کنترل‌شده مانند MNIST که در آزمایش‌های خوشه‌بندی چندنما با اعمال مصنوعی ناقص‌بودن با نرخ‌های ۱۰٪، ۳۰٪ و ۵۰٪ مورد ارزیابی قرار می‌گیرند، عملکرد روش‌های سنتی در نرخ ۵۰٪ به‌طور میانگین حدود ۲۵٪ کاهش می‌یابد [۳].

این آمار نشان می‌دهد که ناقص‌بودن نماها نه‌تنها یک استثنا، بلکه یک قاعده در بسیاری از کاربردهای واقعی است. به‌عنوان مثال، ممکن است در یک پایگاه داده پزشکی، برخی بیماران تنها تصاویر MRI داشته باشند و داده‌های سی‌تی اسکن یا گزارش‌های بالینی آن‌ها ناقص باشد [۴]. در سیستم‌های تشخیص چهره چندزاویه‌ای، ممکن است یک فرد از نمای روبرو قابل شناسایی باشد اما از نمای نیمرخ به‌دلیل پوشش یا نورپردازی نامناسب، داده معتبری وجود نداشته باشد [۵]. این وضعیت که به آن خوشه‌بندی چندنمای ناقص گفته می‌شود، باعث می‌گردد روش‌های سنتی خوشه‌بندی چندنما که نیازمند داده کامل هستند، کارایی خود را از دست بدهند. بنابراین، توسعه روش‌های مقاوم و هوشمند برای خوشه‌بندی داده‌های ناقص به یک ضرورت علمی و عملی تبدیل شده است [۶].

در سال‌های اخیر، پژوهشگران رویکردهای متنوعی برای حل مسئله خوشه‌بندی چندنمای ناقص ارائه کرده‌اند. این رویکردها را می‌توان به چهار دسته کلی تقسیم کرد. نخست، روش‌های ماتریس‌فاکتوریزیشن که با استخراج یک ماتریس اجماعی از نماهای ناقص، نمایش مشترکی از داده‌ها ایجاد می‌کنند. مزیت اصلی این روش‌ها سادگی و قابلیت تفسیر نتایج است، اما ضعف آن‌ها در ناتوانی در مدل‌سازی روابط غیرخطی و پیچیده بین نماها و بی‌توجهی به ساختار محلی داده‌هاست [۲] و [۳]. دسته دوم، روش‌های مبتنی بر کرنل هستند که با استفاده از ترکیب کرنل‌های مختلف، نماهای ناقص را بازسازی کرده و یک کرنل اجماعی

این مقاله در تاریخ ۱۲ بهمن ماه ۱۴۰۴ دریافت و در تاریخ ۳ تیر ماه ۱۴۰۵ بازنگری شد.

سندس بهادری (نویسنده مسئول)، دانشکده فنی مهندسی، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران، (email: sondos.bahadori@iau.ac.ir).

برای خوشه‌بندی می‌سازند. نقطه قوت این روش‌ها توانایی در مدل‌سازی داده‌های غیرخطی است، اما انتخاب و تنظیم کرنل مناسب برای هر نما دشوار بوده و می‌تواند منجر به نتایج ناپایدار شود. در دسته سوم، روش‌های گراف‌محور قرار دارند که با یادگیری گراف‌های اجماعی، روابط همسایگی بین نمونه‌ها را حفظ می‌کنند. مزیت اصلی این رویکردها توانایی در نگه‌داشتن ساختار محلی داده‌هاست، اما ضعف آن‌ها در مقیاس‌پذیری محدود و هزینه محاسباتی بالا در داده‌های بزرگ و نویزی است. در نهایت، روش‌های مبتنی بر یادگیری عمیق با بهره‌گیری از خودرمزگذارها، شبکه‌های مولد یا یادگیری متضاد، نماهای ناقص را بازسازی کرده و نمایش‌های مشترک و غنی‌تری از داده‌ها استخراج می‌کنند. نقطه قوت این روش‌ها توانایی بالا در استخراج ویژگی‌های غیرخطی و مقاوم است، اما در عین حال نیازمند داده‌های زیاد، تنظیمات پیچیده و منابع محاسباتی سنگین هستند و در برابر داده‌های پرت یا ناقص ممکن است ناپایدار باشند.

با وجود پیشرفت‌های قابل توجه در هر یک از این دسته‌ها، سه چالش اساسی همچنان باقی مانده است:

ویژگی‌های زائد: بسیاری از روش‌ها ویژگی‌های نهفته را استخراج می‌کنند، اما این ویژگی‌ها همچنان ابعاد اضافی دارند که باعث کاهش دقت خوشه‌بندی می‌شود.

نادیده گرفتن ساختار محلی: روابط همسایگی بین نمونه‌ها که برای خوشه‌بندی حیاتی هستند، اغلب در نظر گرفته نمی‌شوند.

عدم استفاده از اطلاعات توزیع داده‌ها: بسیاری از مدل‌ها از اطلاعات احتمالی و توزیعی داده‌ها برای هدایت آموزش استفاده نمی‌کنند، در نتیجه در برابر نویز و داده‌های پرت آسیب‌پذیرند.

برای رفع این مشکلات، پژوهش‌های اخیر به سمت ترکیب خودرمزگذارهای چندنما با روش‌های کاهش ابعاد غیرخطی مانند UMAP و همچنین استفاده از مدل‌های آماری مقاوم مانند مدل مخلوط گاوسی حرکت کرده‌اند. این ترکیب باعث می‌شود هم ویژگی‌های سازگار و کم‌بعد استخراج شوند، هم ساختار محلی داده‌ها حفظ شود، و هم خوشه‌بندی نهایی در برابر نویز مقاوم‌تر باشد. علاوه بر این، معرفی توابع خطای جدید مانند تابع زیان احتمالی امکان استفاده از برچسب‌های نرم را فراهم کرده و کیفیت خوشه‌بندی را ارتقا داده است.

در این مقاله، ما یک چارچوب جدید به نام Hybrid Multi-view Clustering Network (HMCNet) معرفی می‌کنیم که نوآوری‌های زیر را در بر دارد:

- ترکیب ماژول توجه با خودرمزگذارهای چندنما برای وزن‌دهی پویا به نماهای مختلف.
- استفاده همزمان از UMAP و خطای سازگاری برای کاهش ابعاد و حفظ ساختار محلی و هم‌راستایی نماها.
- طراحی یک تابع خطای دوگانه شامل تابع زیان احتمالی و تابع خطای سازگاری برای خوشه‌بندی مقاوم‌تر.
- قابلیت اجرا بر روی داده‌های ناقص و کامل، به‌گونه‌ای که مدل در هر دو حالت عملکرد پایدار و دقیق دارد.

این نوآوری‌ها باعث می‌شوند HMCNet نه‌تنها بر محدودیت‌های روش‌های قبلی غلبه کند، بلکه یک چارچوب جامع و انعطاف‌پذیر برای خوشه‌بندی داده‌های چندنمایی در شرایط واقعی ارائه دهد. در ادامه مقاله، ابتدا مروری بر کارهای مرتبط خواهیم داشت، سپس روش پیشنهادی را به‌تفصیل معرفی می‌کنیم، و در نهایت نتایج آزمایش‌های گسترده بر روی چندین دادگان مرجع ارائه خواهد شد.

۲- کارهای مرتبط

تحلیل و خوشه‌بندی داده‌های چندنمایی در سال‌های اخیر توجه گسترده‌ای را به خود جلب کرده است. پژوهشگران برای مقابله با چالش‌های ناشی از ناقص بودن نماها، رویکردهای متنوعی ارائه کرده‌اند که می‌توان آن‌ها را در چند دسته کلی طبقه‌بندی کرد. در ادامه، هر یک از این دسته‌ها به‌تفصیل مورد بررسی قرار می‌گیرند.

۲-۱ روش‌های ماتریس‌فاکتوریزیشن

روش‌های ماتریس‌فاکتوریزیشن در خوشه‌بندی چندنمای ناقص همچنان یکی از رویکردهای مهم هستند. مقالات اخیر نشان می‌دهند که این روش‌ها با افزودن منظم‌سازی‌های گرافی، وزن‌دهی مفهومی و تکنیک‌های هم‌تنظیمی تلاش کرده‌اند ضعف‌های سنتی مانند نادیده‌گرفتن ساختار محلی، بیش‌برازش و عدم هم‌ترازی نماها را برطرف کنند [۶]. از نقاط قوت اصلی این رویکرد می‌توان به سادگی و تفسیرپذیری بالا و امکان استخراج نمایش مشترک کم‌بعد اشاره کرد، اما ناتوانی در مدل‌سازی روابط غیرخطی و بی‌توجهی به ساختار محلی داده‌ها از مهم‌ترین محدودیت‌های آن‌هاست. مطالعه [۷] رویکرد وزن‌دهی مفهومی را معرفی می‌کند که با هم‌تنظیمی بین نماها، نمایش پایدارتر و مقاوم‌تری ایجاد می‌کند. مزیت اصلی آن کاهش بیش‌برازش و مدیریت داده‌های ناقص است. ضعف آن در پیچیدگی تابع هدف و نیاز به الگوریتم‌های تکراری پرهزینه برای همگرایی است.

در [۸] رویکردی مبتنی بر ماتریس‌فاکتوریزیشن را با منظم‌سازی گرافی ترکیب می‌کند تا ساختار همسایگی داده‌ها حفظ شود. هدف اصلی، افزایش مقاومت مدل در برابر داده‌های ناقص و نویزی است. نتایج نشان می‌دهد که این روش در داده‌های کوچک و متوسط عملکرد خوبی دارد، اما در داده‌های بزرگ هزینه محاسباتی بالایی دارد. در [۹] ترکیبی از ماتریس‌فاکتوریزیشن غیرمنفی و ساختار خودرمزگذار ارائه می‌دهد. با افزودن منظم‌سازی دوگانه (گراف داده و ویژگی‌ها)، تلاش می‌کند نمایش‌های کم‌بعدی دقیق‌تری یاد بگیرد. نقطه ضعف آن حساسیت به نویز و دشواری در تنظیم پارامترهاست.

در [۱۰] از NMF برای داده‌های چندمدلی و چندنما استفاده می‌کند تا همکاری بین نماها ایجاد شود. مزیت اصلی آن توانایی در مدیریت داده‌های غیرخطی و چندمدلی است. اما در داده‌های ناقص با نرخ بالای فقدان، عملکرد ضعیف‌تری نشان می‌دهد. در [۱۱]، ترکیبی از ماتریس اجماعی و ماتریس‌های اختصاصی نما را با منظم‌سازی گرافی ارائه می‌دهد. هدف آن حفظ ساختار هندسی داده‌ها و استخراج نمایش‌های سازگار و اختصاصی است. ضعف اصلی آن پیچیدگی زیاد در طراحی تابع هدف و نیاز به الگوریتم‌های چندمرحله‌ای برای بهینه‌سازی است.

۲-۲ روش‌های مبتنی بر کرنل

روش‌های مبتنی بر کرنل در خوشه‌بندی چندنمای ناقص تلاش می‌کنند با ترکیب چند کرنل از نماهای مختلف، یک کرنل اجماعی بسازند که بتواند ساختار داده‌ها را بهتر مدل کند [۲] و [۳]. مزیت اصلی این رویکردها توانایی در مدل‌سازی روابط غیرخطی و پیچیده بین نماهاست، اما ضعف آن‌ها معمولاً در انتخاب و تنظیم کرنل مناسب، هزینه محاسباتی بالا و حساسیت به داده‌های ناقص است.

در [۱۲] رویکردی ترکیبی ارائه شده است که داده‌های ناقص را ابتدا به فضای کم‌بعدی فرافکنی کرده و سپس با استفاده از کرنل‌ها و تجزیه تنسور، روابط بین نماها را مدل می‌کند. مزیت آن کاهش نویز و

ویژگی‌های زائد است، اما ضعفش پیچیدگی محاسباتی و نیاز به الگوریتم‌های بهینه‌سازی سنگین است. این روش [۱۳] از مکانیزم توجه بین‌نما و یادگیری خصمانه برای ساخت کرنل‌های مقاوم استفاده می‌کند. نقطه قوت آن توانایی در حفظ همزمان اطلاعات مشترک و اختصاصی نماهاست. ضعف آن نیاز به تنظیمات پیچیده و حساسیت به نرخ بالای داده‌های ناقص است.

در [۱۴] کرنل‌ها را در قالب یک تنسور کم‌رتبه مدل می‌کند تا همبستگی‌های مرتبه بالا بین نماها استخراج شود. مزیت آن توانایی در یادگیری ساختارهای پیچیده و مقاوم در برابر نویز است. ضعف آن هزینه محاسباتی بالا و دشواری در مقیاس‌پذیری است. در [۱۵] در حوزه داده‌های زیستی، MKL را برای ادغام داده‌های چنداُمیک به‌کار می‌گیرد. مزیت آن انعطاف‌پذیری در ترکیب داده‌های ناهمگون و عملکرد بهتر نسبت به مدل‌های پیچیده‌تر است. ضعف آن نیاز به انتخاب دقیق کرنل‌ها و وابستگی به کیفیت داده‌های ورودی است.

در [۱۶] یک برزیابی جامع برای ارزیابی تولید کرنل‌های یادگیری ماشین ارائه می‌دهد. مزیت آن فراهم‌کردن چارچوبی استاندارد برای مقایسه روش‌های MKL است. ضعف آن تمرکز بیشتر بر تولید کرنل و کمتر بر کاربرد مستقیم در خوشه‌بندی چندنماست. این مطالعه [۱۷] روشی مقیاس‌پذیر برای ترکیب چند کرنل آنلاین معرفی می‌کند تا سوگیری انتخاب کرنل کاهش یابد. مزیت آن کاهش خطاهای ناشی از انتخاب کرنل نامناسب است. ضعف آن محدودیت در داده‌های بسیار بزرگ و نیاز به تنظیم دقیق پارامترهاست.

۲-۳ روش‌های گراف‌محور

روش‌های گراف‌محور در IMC تلاش می‌کنند با ساخت یا یادگیری یک گراف اجماعی، روابط همسایگی بین نمونه‌ها را مدل کنند. توانایی در حفظ ساختار محلی داده‌ها و نمایش بهتر روابط بین نماها از مزایای اصلی این رویکردهاست. با این حال، هزینه محاسباتی بالا در داده‌های بزرگ، حساسیت به نویز، و دشواری در مقیاس‌پذیری از نقاط ضعف مشترک آن‌هاست [۱۸]. این روش‌ها معمولاً از تکنیک‌هایی مانند منظم‌سازی گراف لاپلاسی[1]، خوشه‌بندی طیفی یا جاسازی گراف[2] استفاده می‌کنند. با این حال، ضعف‌های مشترک آن‌ها شامل هزینه محاسباتی بالا در داده‌های بزرگ، حساسیت به نویز، و دشواری در مقیاس‌پذیری است. به همین دلیل، پژوهش‌های اخیر به سمت ترکیب گراف با یادگیری عمیق یا منظم‌سازی‌های چندگانه حرکت کرده‌اند.

در [۱۹] ماتریس‌فاکتوریزیشن را با منظم‌سازی گرافی ترکیب می‌کند تا ساختار محلی داده‌ها حفظ شود. مزیت آن توانایی در نگه‌داشتن روابط همسایگی است، اما ضعف آن هزینه محاسباتی بالا در داده‌های بزرگ است. در [۲۰] یک روش یادگیری تطبیقی گراف معرفی شده است که گراف اجماعی را به‌صورت پویا با داده‌های ناقص سازگار می‌سازد. مزیت آن انعطاف‌پذیری در مدیریت نماهای ناقص است، اما ضعف آن پیچیدگی الگوریتم و نیاز به تنظیم دقیق پارامترهاست .در [۲۱] از دو نوع منظم‌سازی گراف (گراف نمونه‌ها و گراف ویژگی‌ها) استفاده می‌کند تا نمایش‌های دقیق‌تری یاد بگیرد. مزیت آن حفظ همزمان ساختار داده و ویژگی‌هاست، اما ضعف آن افزایش پیچیدگی محاسباتی و دشواری در همگرایی است.

در [۲۲] یادگیری متضاد گرافی را برای داده‌های ناقص معرفی می‌کند. مزیت آن توانایی در استخراج نمایش‌های مقاوم و غنی است، اما ضعف آن نیاز به داده‌های زیاد و حساسیت به نرخ بالای فقدان نماهاست. در [۲۳] از یادگیری خودنظارتی برای ساخت گراف اجماعی استفاده می‌کند. مزیت آن کاهش وابستگی به برچسب‌ها و افزایش مقاومت در برابر نویز است، اما ضعف آن نیاز به طراحی دقیق وظایف خودنظارتی و هزینه محاسباتی بالاست. در [۲۴] روشی مقیاس‌پذیر برای داده‌های بزرگ معرفی می‌کند که از تقریب‌های گرافی استفاده می‌کند. مزیت آن توانایی در مدیریت داده‌های عظیم است، اما ضعف آن کاهش دقت نسبت به روش‌های دقیق‌تر است.

۲-۴ روش‌های مبتنی بر یادگیری عمیق

روش‌های مبتنی بر یادگیری عمیق در IMC تلاش می‌کنند با استفاده از خودرمزگذارها، شبکه‌های مولد یا یادگیری متضاد[3]، نماهای ناقص را بازسازی کرده و یک نمایش مشترک و غنی از داده‌ها استخراج کنند [۲۵]. مزیت اصلی این رویکردها توانایی در مدل‌سازی روابط غیرخطی پیچیده و استخراج ویژگی‌های مقاوم است. با این حال، ضعف‌های مشترک آن‌ها شامل نیاز به داده‌های زیاد، تنظیمات پیچیده، منابع محاسباتی سنگین، و حساسیت به داده‌های پرت یا نویزی است.

در [۲۶] از خودرمزگذارهای چندنما برای بازسازی داده‌های ناقص و هم‌ترازی نمایش‌های نهفته استفاده می‌کند. مزیت آن توانایی در یادگیری نمایش مشترک غیرخطی است، اما ضعف آن نیاز به تنظیم دقیق فراپارامترها و داده‌های آموزشی زیاد است. در [۲۷] از شبکه‌های مولد خصمانه برای بازسازی نماهای ناقص بهره می‌گیرد. مزیت آن توانایی در تولید داده‌های مصنوعی مقاوم و غنی است، اما ضعف آن ناپایداری آموزش GAN و حساسیت به نرخ بالای داده‌های ناقص است. در [۲۸] یادگیری متضاد را برای استخراج نمایش‌های مشترک مقاوم معرفی می‌کند. مزیت آن افزایش مقاومت در برابر نویز و داده‌های ناقص است، اما ضعف آن نیاز به طراحی دقیق وظایف متضاد و داده‌های متنوع برای آموزش است.

در [۲۹] از یک چارچوب دوگانه خودرمزگذار استفاده می‌کند که هم بازسازی نماها و هم یادگیری نمایش اجماعی را انجام می‌دهد. مزیت آن توانایی در ترکیب اطلاعات مشترک و اختصاصی نماهاست، اما ضعف آن پیچیدگی معماری و هزینه محاسباتی بالاست. در [۳۰] یادگیری خودنظارتی را با روش گرافی ترکیب می‌کند تا نمایش‌های مشترک بدون نیاز به برچسب یاد گرفته شوند. مزیت آن کاهش وابستگی به داده‌های برچسب‌دار و افزایش مقاومت در برابر نویز است، اما ضعف آن نیاز به طراحی دقیق وظایف خودنظارتی و هزینه محاسباتی بالا در داده‌های بزرگ است.

جدول ۱ این دسته‌بندی را به‌همراه نقاط قوت و ضعف هر دسته، و نیز نمونه روش‌های شاخص هر گروه به‌صورت خلاصه نشان می‌دهد.

HMCNet با ترکیب نقاط قوت رویکردهای پیشین و معرفی نوآوری‌های تازه، تلاش می‌کند بر محدودیت‌های موجود در خوشه‌بندی چندنمای ناقص غلبه کند. این مدل با بهره‌گیری از ماژول توجه برای وزن‌دهی پویا به نماها، استفاده از UMAP به‌منظور کاهش ابعاد و حفظ ساختار محلی داده‌ها، و طراحی یک تابع خطای دوگانه[4] شامل تابع زیان

1. Graph Laplacian Regularization
2. Graph Embedding
3. Contrastive Learning
4. Dual Loss

جدول ۱: مقایسه دسته‌بندی روش‌های خوشه‌بندی چندنمای ناقص.

دسته روش	روش‌های	نقاط قوت اصلی	نقاط ضعف اصلی
ماتریس‌فاکتورینگ	PVC [۸], MIC [۹], DAIMC [۱۰], Graph-regularized NMF [۱۱]	سادگی و تفسیرپذیری بالا؛ امکان استخراج نمایش مشترک کم‌بعد از نماهای ناقص	ناتوانی در مدل‌سازی روابط غیرخطی پیچیده؛ بی‌توجهی به ساختار محلی داده‌ها؛ هزینه محاسباتی بالا در داده‌های بزرگ
روش‌های مبتنی بر کرنل	MKKM-IK [۱۲], LMKML [۱۴], LRKT-IMVC [۱۵]	توانایی مدل‌سازی روابط غیرخطی؛ قابلیت ترکیب چند کرنل از نماهای مختلف؛ انعطاف‌پذیری در داده‌های ناهمگون	انتخاب و تنظیم کرنل مناسب دشوار است؛ حساسیت به داده‌های ناقص و نویز؛ پیچیدگی محاسباتی بالا ((O(N3) ؛ نیاز به حداقل یک نمای کامل در برخی روش‌ها
روش‌های گراف‌محور	APMC [۱۹], IMG [۲۰], NACG_IMVC [۲۱], OIMVC-BGL [۲۲]	حفظ ساختار محلی داده‌ها؛ نمایش دقیق روابط همسایگی؛ قابلیت استفاده از اطالعات گراف دوبخشی برای کاهش محاسبات	هزینه محاسباتی بالا در داده‌های بزرگ (ساخت گراف کامل (O(N²)؛ حساسیت به نویز و پارامترها؛ نیاز به پس‌پردازش برای استخراج خوشه‌ها در برخی روش‌ها
روش‌های مبتنی بر یادگیری عمیق	DCP [۲۶], SMILE [۲۸], CDIMC-net [۳۰]	توانایی استخراج ویژگی‌های سطح‌بالا و مقاوم؛ مدل‌سازی روابط غیرخطی پیچیده؛ یادگیری نمایش‌های مشترک غنی	نیازمند داده‌های زیاد برای آموزش؛ تنظیم پیچیده فراپارامترها؛ منابع محاسباتی سنگین؛ حساسیت به داده‌های پرت و نویزی؛ ناپایداری در نرخ‌های بالای ناقص‌بودن

احتمال و یادگیری متضاد، قادر است نمایش‌های مشترک دقیق‌تر و خوشه‌بندی مقاوم‌تری ارائه دهد. بدین ترتیب، HMCNet به‌عنوان یک چارچوب جامع و انعطاف‌پذیر عمل می‌کند که هم در شرایط داده‌های ناقص و هم در شرایط داده‌های کامل، عملکردی پایدار و دقیق دارد.

۳- راهکار پیشنهادی

مدل پیشنهادی با نام Hybrid Multi-view Clustering Network (HMCNet) یک چارچوب جامع و مقاوم برای خوشه‌بندی داده‌های چندنمایی ناقص و کامل است. این مدل با ترکیب ماژول توجه، کاهش ابعاد غیرخطیUMAP، مدل مخلوط گاوسی و یک تابع خطای دوگانه طراحی شده است، معماری مدل پیشنهادی در شکل ۱ آمده است. در ادامه اجزای اصلی مدل توضیح داده می‌شوند.

۳-۱ تعریف مسئله

به منظور ساده‌سازی پردازش داده‌ها، هر نمونه از نماها با یک ماتریس مدل شده است. مساله خوشه‌بندی داده‌های چندنما به‌صورت زیر تعریف می‌شود:

$$X = \left\{x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(v)}\right\},\ x^{(v)} \hat{I} R^{n \times d_v} \qquad (1)$$

که $x^{(v)}$ نمای vام از n نمونه است. برخی نماها ممکن است ناقص باشند. هدف، خوشه‌بندی این داده‌ها به K گروه است.

۳-۲ ماژول استخراج ویژگی

۳-۲-۱ انکودر اختصاصی برای هر نما

هر نما ویژگی‌های خاص خودش را دارد (مثلاً تصویر با پیکسل‌ها، متن با کلمات، صدا با موج‌ها). اگر همه نماها یک انکودر مشترک داشته باشند، مدل نمی‌تواند تفاوت‌های ذاتی بین نماها را درک کند. با طراحی انکودر اختصاصی، هر نما ابتدا به‌طور مستقل فشرده و بازنمایی می‌شود، سپس این نمایش‌ها در مرحله بعدی (ماژول توجه و تجمیع) ترکیب می‌شوند. در معماری HMCNet، برای هر نما یک خودرمزگذار اختصاصی طراحی می‌شود:

$$H^{(v)} = f_{\theta^{(v)}}\left(X^{(v)}\right) \qquad (2)$$

که در آن $X^{(v)}$ داده‌های ورودی مربوط به نما v، $f_{\theta^{(v)}}$ تابع انکودر نما v که با پارامترهای $\theta^{(v)}$ آموزش داده می‌شود و $H^{(v)} \in R^{n \times d_v}$ ویژگی نهفته نما v است.

۳-۲-۲ ماژول توجه

همه نماها کیفیت یا اطلاعات یکسانی ندارند. با استفاده از ماژول توجه، مدل می‌تواند به‌صورت پویا یاد بگیرد که کدام نماها باید بیشتر در نمایش سازگار نهایی مشارکت کنند. این مکانیزم همان چیزی است که به آن ماژول توجه می‌گوییم.

برای وزن‌دهی پویا به نماها از معادله زیراستفاده می‌شود:

$$\alpha^{(v)} = \frac{\exp(w^T H^{(v)})}{\sum_{u=1}^{v} w^T H^{(u)}} \qquad (3)$$

که در آن w^T در واقع یک ضرب داخلی بین بردار وزن‌های قابل یادگیری و ویژگی‌های نهفته نما است که در طول آموزش بهینه می‌شود تا بتواند اهمیت هر نما را به‌صورت پویا تعیین کند، $\alpha^{(v)}$ وزن نما v است.

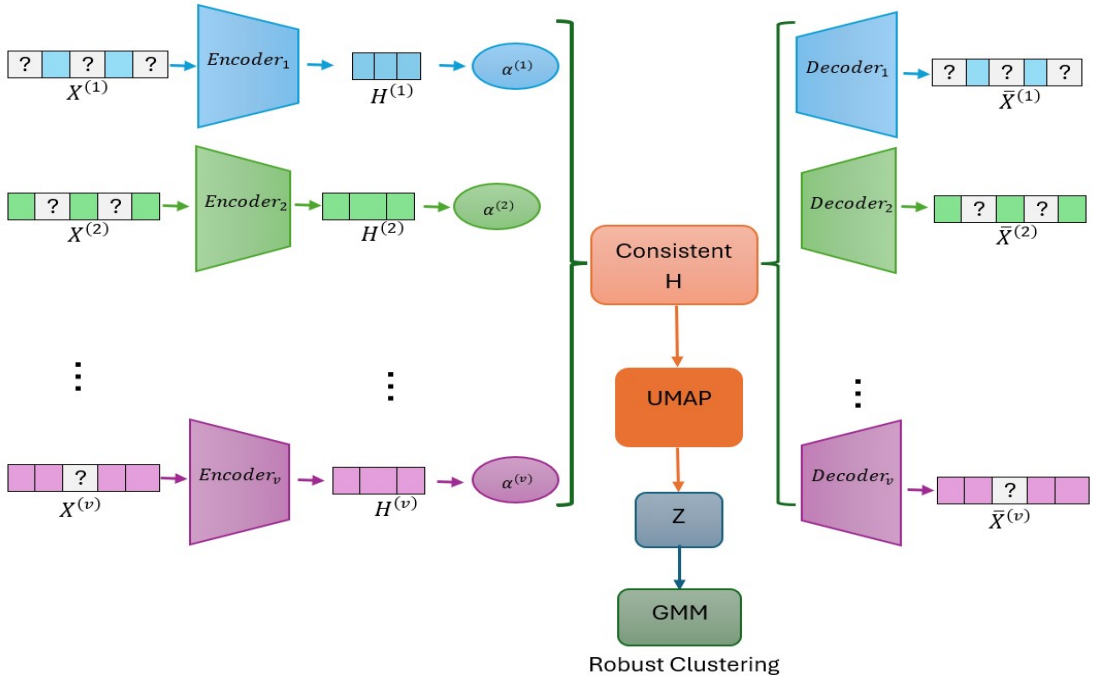
۳-۳ کاهش ابعاد و حفظ ساختار محلیUMAP

به منظور فیلتر نمودن ویژگی های اضافه و حفظ نمون وابستگی محلی نماها از مکانیزم UMAP استفاده شده است. در روش UMAP، هدف این است که ساختار محلی داده‌های پربعد در فضای کم‌بعد حفظ شود. برای این کار، UMAP ابتدا یک گراف وزن‌دار از داده‌های اصلی می‌سازد و سپس تلاش می‌کند گراف مشابهی در فضای کم‌بعد بازسازی کند.

تابع خطای UMAP به‌صورت زیر تعریف می‌شود:

$$L_{UMAP} = \sum_{i<j}\left[P_{ij}\log q_{ij} + (1-P_{ij})\log(1-q_{ij})\right] \qquad (4)$$

که P_{ij} شباهت بین نمونه‌های i و j در فضای پر بعد (بر اساس گراف K-NN و توزیع احتمال.)، q_{ij} شباهت بین همان نمونه‌ها در فضای کم‌بعد بر اساس فاصله بین بردارهای z_i و z_j است. این تابع خطا تلاش



شکل ۱: معماری مدل پیشنهادی.

می‌کند توزیع شباهت‌ها در فضای کم‌بعد q_{ij} تا حد ممکن به توزیع اصلی P_{ij} نزدیک باشد.

ویژگی‌های H به فضای کم‌بعدی Z نگاشت می‌شوند:

$$Z = UMAP(H) \quad (۵)$$

۳-٤ خوشه‌بندی مقاوم

۳-٤-۱ مدل مخلوط گاوسی

با توجه به اینکه مدل‌های مخلوط گاوسی می‌توانند توزیع‌های نمونه پیچیده را برازش کنند از توزیع GMM استفاده شده است. توزیع داده‌ها با ترکیب چند گاوسی به صورت زیر مدل می‌شود:

$$p(z) = \sum_{k=1}^{K} \pi_k N(z \mid \mu_k, \sigma_k) \quad (۶)$$

که در آن π_k وزن خوشه، μ_k میانگین و σ_k کوواریانس هستند.

۳-٤-۲ احتمال تعلق

برای هر نمونه z_i احتمال تعلق به خوشه k به صورت زیر محاسبه می‌شود:

$$P_{ik} = \frac{\pi_k N(z_i \mid \mu_k, \sigma_k)}{\sum_{j=1}^{K} \pi_j N(z_i \mid \mu_j, \sigma_j)} \quad (۷)$$

۳-۵ توابع خطا

برای حفظ ساختار کلی داده ها خطای بازسازی به صورت زیر محاسبه می‌شود:

$$L_R = \sum_{v=1}^{V} D^{(v)} \left\| X^{(v)} - \hat{X}^{(V)} \right\|^2 \quad (۸)$$

به منظور حفظ هم‌راستایی نماها، خطای سازگاری به صورت زیر تعریف می‌گردد:

$$L_C = \sum_{i=1}^{n} \sum_{v=1}^{V} \alpha^{(v)} \left\| h_i - H_i^{(v)} \right\|^2 \quad (۹)$$

خطای القاشده از احتمال بر اساس برچسب‌های نرم GMM به صورت زیر تعریف می‌شود:

$$L_P = \sum_{i=1}^{n} \sum_{j \in C_a, k \notin C_a} (1 - \left| P_{ia} - P_{ja} \right|) . \left\| h_i - h_j \right\|^2 \quad (۱۰)$$

مدل با ترکیب همه خطاها آموزش داده می‌شود و تابع هدف نهایی به صورت زیر تعریف می‌شود.

$$L = L_R + \alpha L_C + \beta L_P + \gamma L_{UMAP} \quad (۱۱)$$

که (α, β, γ) فراپارامترهای تنظیم وزن هستند.

ترکیب چهار تابع خطای L_R، L_C ، L_P و L_{UMAP} در (۱۱) بر اساس سه اصل اساسی در خوشه‌بندی چندنمای ناقص طراحی شده است [۳۵]. نخست، اصل حفظ اطلاعات که توسط تابع بازسازی L_R تأمین می‌شود و تضمین می‌کند نمایش نهفته هر نما، اطلاعات کافی برای بازسازی داده اصلی را حفظ کند؛ در غیر این صورت، مدل ممکن است ویژگی‌های مهم هر نما را از دست بدهد. دوم، اصل هم‌راستایی که توسط تابع سازگاری L_C تأمین می‌شود و نمایش‌های نماهای مختلف را به‌یکدیگر نزدیک می‌کند تا به یک مفهوم مشترک اشاره کنند؛ این اصل در داده‌های چندنما که در آن نماها مکمل یکدیگرند، ضروری است. سوم، اصل تفکیک‌پذیری و حفظ ساختار محلی که به‌ترتیب توسط توابع L_P و L_{UMAP} تأمین می‌شوند؛ تابع احتمالی با استفاده از برچسب‌های نرم GMM، نمونه‌های هم‌خوشه را به‌هم نزدیک و نمونه‌های خوشه‌های مختلف را از هم دور می‌کند، در حالی که UMAP ساختار همسایگی داده‌ها را در فضای کم‌بعد حفظ می‌نماید. این چهار تابع به‌صورت مکملی ترکیب شده‌اند؛ برای مثال، L_R به‌تنهایی نمی‌تواند هم‌راستایی نماها را تضمین کند، اما L_C این کار را انجام می‌دهد. همچنین، L_C به‌تنهایی ممکن است باعث از دست رفتن اطلاعات اختصاصی هر نما شود که L_R آن را جبران می‌کند. وزن‌های α، β و γ در (۱۱) نقش تنظیم اهمیت نسبی هر یک از این اهداف را بر عهده دارند و انتخاب آن‌ها در زیربخش ۴-۴ به‌تفصیل توضیح داده شده است. وزن تابع UMAP به‌صورت ثابت و بدون نیاز به تنظیم مجدد در نظر گرفته شده است، زیرا پارامترهای

Algorithm HMCNet:

Input :
Multi-view dataset $X = \left\{ X^{(1)}, X^{(2)}, X^{(3)}, \ldots, X^{(v)} \right\}$
Number of clusters K
Hyperparameters α, β, γ
Output:
Cluster assignments C
Network parameters $\{ \theta^{(v)} \}$

Phase 1: Pre-training (Feature Extraction)

1: For each view $v = 1$ to V :
Initialize view-specific encoder $f_\theta(v)$ and decoder $g_\theta(v)$
2: For each sample i :
Compute latent features $H_i^{(v)} = f_\theta(v)(X^{(v)})$
3: Compute attention weights $\alpha^{(v)}$ for each view
4: Fuse latent features into consistent representation:

$$h_i = \sum_{v=1}^{v} \alpha^{(v)} D_i^{(v)} H_i^{(v)}$$

5: Reconstruct input views:

$$\hat{X}_l^{(v)} = g_{\theta^{(v)}}(h_i)$$

6: Compute losses :

$$L_R = \sum_{v=1}^{V} D^{(v)} \left\| X^{(v)} - \hat{X} \right\|^2$$

$$L_C = \sum_{i=1}^{n} \sum_{v=1}^{v} \alpha^{(v)} \left\| h_i - H_i^{(v)} \right\|^2$$

7:Update network parameters by minimizing:

$$L_{Pre} = L_R + \alpha LC$$

Phase 2: Retraining (Clustering & Embedding)

1: Initialize parameters with pre-trained values
2: For each iteration until convergence:
a. Compute consistent features h_i
b. Apply UMAP embedding:

$$Z = UMAP(H)$$

c. Cluster Z using GMM:
Obtain probabilities p_{ik} for each sample i
d. Construct triplets (anchor, positive, negative)
e. Compute losses:

$$L_P = \sum_{i=1}^{n} \sum_{j \in C_a, k \notin C_a} (1 - \left| P_{ia} - P_{ja} \right|) . \left\| h_i - h_j \right\|^2$$

$$L = L_R + \alpha L_C + \beta L_P + \gamma L_{UMAP}$$

f. Update network parameters by minimizing

Phase 3: Output

Return final cluster assignments C from GMM
Return optimized network parameters $\{ \theta^{(v)} \}$

شکل ۲: شبه کد HMCNet.

UMAP (مانند تعداد همسایگی و ابعاد خروجی) به‌صورت جداگانه تنظیم می‌شوند و تأثیر آن در تابع هدف با ضریب ثابت ۱ لحاظ شده است.

۳-۶ مراحل آموزش

فرآیند آموزش مدل پیشنهادی شامل دو مرحله اساسی است (شکل ۲). تنظیم فراپارامترها در این دو مرحله بر اساس استراتژی‌های متفاوتی انجام می‌شود:

مرحله ۱-پیش‌آموزش[1]: در این مرحله، خودرمزگذارها با استفاده از تابع زیان $L_{Pre} = L_R + \alpha LC$ به مدت ۱۰۰ دوره با نرخ یادگیری ۰/۰۰۱ آموزش داده می‌شوند. انتخاب ۱۰۰ دوره بر اساس معیار توقف زودهنگام[2] با پشتیبانی ۲۰ دوره انجام شده است؛ به این معنا که اگر به مدت ۲۰ دوره متوالی، مقدار خطای اعتبارسنجی کاهش نیابد، آموزش متوقف می‌شود. فراپارامتر α در این مرحله با جستجوی شبکه‌ای در محدوده [۰/۱-۰/۹] و با هدف بیشینه‌سازی شباهت بین نمایش‌های نماهای مختلف با استفاده از معیار تحلیل همبستگی کانونی[3] (CCA) تنظیم شده است.

مرحله ۲-بازآموزی[4]: در این مرحله، تمام مؤلفه‌های مدل شامل UMAP، GMM و توابع زیان ترکیبی، به‌صورت همزمان به مدت ۲۰۰ دوره با نرخ یادگیری ۰/۰۰۰۵ آموزش داده می‌شوند. انتخاب ۲۰۰ دوره بر اساس مشاهده همگرایی کامل تابع زیان کل، L، در دادگان‌های مختلف صورت گرفته است. فراپارامترهای β و γ در این مرحله به‌صورت همزمان تنظیم می‌شوند. وزن تابع UMAP در تابع هدف نهایی (۱۱) به‌صورت ضریب ثابت ۱ در نظر گرفته شده است و نیازی به تنظیم به‌عنوان فراپارامتر جداگانه ندارد، زیرا تأثیر آن از طریق پارامترهای مستقل UMAPتعداد همسایگی k و ابعاد خروجی که پیش‌تر با استفاده از معیارهای اعتمادپذیری[5] وتداوم[6] تنظیم شده‌اند، اعمال می‌شود.

استراتژی کاهش نرخ یادگیری: در هر دو مرحله، از یک برنامه کاهش نرخ یادگیری مبتنی بر دوره[7] با کاهش ۰/۹ به ازای هر ۲۰ دوره استفاده شده است تا از گیر افتادن در بهینه‌های محلی جلوگیری شود.

بهینه‌ساز: از بهینه‌ساز Adam با پارامترهای پیش‌فرض ($\varepsilon = 1e^{-8}, \beta_2 = 0/999, \beta_1 = 0/9$) استفاده شده است.

۳-۷ انتخاب فراپارامترهای UMAP

فراپارامترهای UMAP شامل تعداد همسایگی[8] و ابعاد فضای

1. Pre-training
2. Early Stopping
3. Cannonical Correlation Analysis
4. Retraining
5. Trustworthiness
6. Continuity
7. StepLR
8. N_Neighbors

جدول ۲: تحلیل پیچیدگی زمانی مؤلفه‌های مدلHMCNet.

مؤلفه مدل	پیچیدگی زمانی	پارامترهای تأثیرگذار	روش بهینه‌سازی
خودرمزگذارهای اختصاصی	$O(N \times h \times d_v)$	N : تعداد نمونه‌ها، d_v : ابعاد نما، h : اندازه لایه پنهان	پردازش دسته‌ای، موازی‌سازی
UMAP [۵]	$O(N \log N)$ تا $O(N^2)$	N: تعداد نمونه‌ها	جستجوی همسایگی تقریبی
GMM [۳۱]	$O(N \times K \times iter \times d_z)$	N : تعداد نمونه‌ها، K : تعداد خوشه‌ها، $iter$: تعداد تکرارها، d_z : ابعاد نهفته	استفاده از کوواریانس قطری به‌جای کامل
تابع زیان ترکیبی	$O(N \times K \times d_z)$	N : تعداد نمونه‌ها، V : تعداد نماها، d_z : ابعاد نهفته	محاسبات برداری
کل مدلHMCNet	$O(N \log N)$	-	ترکیب بهینه‌سازی‌های فوق

جدول۳: مقایسه پیچیدگی زمانی روش‌های مختلف خوشه‌بندی چندنمای ناقص.

روش	پیچیدگی زمانی	زمان آموزش (MNIST Subset)
PVC (ماتریس‌فاکتورینگ)	$O(N^3)$	~۴۵۰ ثانیه
UEF (گراف‌محور)	$O(N^2 \times d)$	~۳۲۰ ثانیه
CDIMC-net (یادگیری عمیق)	$O(N \times d \times h \times epochs)$	~۴۰۰ ثانیه
DCP (یادگیری عمیق)	$O(N \times d \times h \times epochs)$	~۳۸۰ ثانیه
IMCNET-MP	$O(N^2 \times \log N)$	~۲۵۰ ثانیه
LMKML (کرنل)	$O(N^3)$	~۵۲۰ ثانیه
HMCNet (پیشنهادی)	$O(N \log N)$	**~۱۸۰ ثانیه**

کم‌بعد[1] می‌شوند. این پارامترها به‌صورت جداگانه و با استفاده از معیار اعتمادپذیری و تداوم تنظیم شده‌اند:

تعداد همسایگی، k، در بازه [۵, ۵۰] با گام ۵ جستجو شد. مقدار $k = ۱۵$ به دلیل ایجاد تعادل بین حفظ ساختار محلی (با k کوچک) و پایداری توپولوژیکی (با k بزرگ) انتخاب شد.

ابعاد خروجی، d_z، در بازه [۱۶, ۱۲۸] با گام ۱۶ جستجو شد. مقدار $d_z = ۶۴$ انتخاب گردید زیرا افزایش بیشتر ابعاد، بهبود ناچیزی در کیفیت خوشه‌بندی (NMI کمتر از ۱٪ افزایش) داشت در حالی که هزینه محاسباتی را به‌طور قابل توجهی افزایش می‌داد. همچنین، پارامترهای تعداد مؤلفه‌های گاوسی[2] برابر با تعداد کلاس‌های واقعی هر دادگان در نظر گرفته شده است. این انتخاب بر اساس دانش قبلی[3] از دادگان‌های استاندارد صورت گرفته است. در کاربردهای واقعی که تعداد کلاس‌ها نامعلوم است، می‌توان از معیار اطلاعات بیزی[4] یا معیار AIC برای انتخاب تعداد بهینه مؤلفه‌ها استفاده کرد.

۳-۸ تحلیل پیچیدگی زمانی

برای ارزیابی کارایی محاسباتی مدل HMCNet، پیچیدگی زمانی مؤلفه‌های مختلف آن و همچنین کل مدل، با روش‌های رقیب مقایسه شده است .جدول ۲ پیچیدگی هر مؤلفه از مدل پیشنهادی را به‌صورت مجزا نشان می‌دهد

همان‌طور که در جدول ۲ مشاهده می‌شود، مؤلفه UMAP با پیچیدگی $O(N \log N)$ و GMM با کوواریانس قطری با پیچیدگی $O(N \times K \times d_z)$، بیشترین تأثیر را در کاهش پیچیدگی کل مدل دارند. خودرمزگذارهای اختصاصی نیز با پیچیدگی خطی نسبت به تعداد نمونه‌ها، با استفاده از پردازش دسته‌ای (Batch Processing) و موازی‌سازی قابل بهینه‌سازی هستند. تابع زیان ترکیبی نیز با پیچیدگی $O(N \times K \times d_z)$ که با v تعداد نماهاست، به‌دلیل ماهیت برداری محاسبات، زمان ناچیزی را به کل فرآیند اضافه می‌کند. در مجموع، پیچیدگی کل مدل HMCNet از مرتبه $O(N \log N)$است که در عمل با استفاده از بهینه‌سازی‌های پیاده‌سازی (مانند جستجوی همسایگی تقریبی در UMAP و پردازش دسته‌ای در خودرمزگذارها) به‌طور قابل‌توجهی کاهش می‌یابد.

در جدول۳، پیچیدگی زمانی مدل پیشنهادی با روش‌های پیشرفته موجود مقایسه شده است. نتایج نشان می‌دهد که HMCNet با پیچیدگی $O(N \log N)$عملکردی قابل‌رقابت با کارآمدترین روش‌های موجود دارد و نسبت به روش‌های مبتنی بر گراف کامل و طیفی ($O(N^2)$ تا $O(N^3)$) بهبود چشمگیری ارائه می‌دهد. این مزیت عمدتاً ناشی از جایگزینی گراف‌سازی کامل با UMAP و استفاده از GMM با ساختار کوواریانس قطری است که نیاز به محاسبات ماتریسی سنگین را کاهش می‌دهد.

برای تأیید عملی این تحلیل، زمان اجرای مدل‌ها بر روی دادگان MNIST Subset با ۴۰۰۰ نمونه اندازه‌گیری شد. زمان آموزش کل مدل HMCNet حدود ۱۸۰ ثانیه و زمان استنتاج به‌ازای هر نمونه حدود ۲۵ میلی‌ثانیه است که برای کاربردهای عملی مناسب می‌باشد. قابل ذکر است که زمان آموزش روش LMKML با پیچیدگی $O(N^3)$ حدود ۵۲۰ ثانیه گزارش شده است که بیش از ۲/۵ برابر زمان آموزش مدل پیشنهادی است.

٤- نتایج آزمایش‌ها و تحلیل

برای ارزیابی جامع عملکرد مدل پیشنهادی HMCNet، مجموعه‌ای از آزمایش‌ها بر روی دادگان‌های استاندارد و پرکاربرد انجام شده است. این دادگان‌ها به دلیل تنوع در تعداد نماها، میزان ناقص بودن داده‌ها و

1. N_ Components
2. GMM
3. Prior Knowledge
4. BIC

جدول ۴: دادگان‌های مورد استفاده.

نام دادگان	تعداد نمونه	تعداد نما	تعداد کلاس‌ها
Handwritten	۲۰۰۰	۵	۱۰
BDGP	۲۵۰۰	۲	۵
MNIST Subset	۴۰۰۰	۲	۱۰
Caltech-۲۰	۲۳۸۶	۶	۲۰
Fashion	۱۰۰۰۰	۳	۱۰

پیچیدگی ساختار ویژگی‌ها انتخاب شده‌اند تا توانایی مدل در شرایط مختلف به‌طور کامل بررسی شود. نتایج با چندین روش پیشرفته موجود شامل PVC، UEAF، CDIMC-net، DCP، IMCNet-MP و LMKML مقایسه شده است. این مقایسه‌ها امکان تحلیل دقیق نقاط قوت و ضعف روش‌های مختلف را فراهم کرده و جایگاه مدل پیشنهادی را در میان رویکردهای پیشرفته روشن می‌سازد.

مدل پیشنهادی در دو مرحله آموزش داده شد. در مرحله پیش‌تمرین، شبکه به مدت ۱۰۰ دوره[1] با نرخ یادگیری ۰/۰۰۱ آموزش دید تا نمایش‌های اولیه پایدار استخراج شوند. سپس در مرحله بازآموزی، فرآیند آموزش به مدت ۲۰۰ دوره با نرخ یادگیری کاهش‌یافته ۰/۰۰۰۵ ادامه یافت تا همگرایی دقیق‌تر و بهینه‌سازی نهایی حاصل شود. برای کاهش ابعاد و حفظ ساختار محلی داده‌ها از الگوریتم UMAP با تعداد همسایه‌های $k=۱۵$ و ابعاد خروجی $d_z=۶۴$ استفاده شد. در مرحله خوشه‌بندی، الگوریتم GMM به کار گرفته شد که تعداد خوشه‌ها برابر با تعداد کلاس‌های واقعی هر دادگان در نظر گرفته شد.

فراپارامترهای مدل به صورت $\alpha=۰/۵$ ، $\beta=۱/۰$ و $\gamma=۰/۵$ تنظیم شدند. α وزن مربوط به تابع بازسازی نماها است که کیفیت نمایش نهفته هر نما را تضمین می‌کند؛ β وزن اصلی تابع خطای سازگاری است که هم‌ترازی بین نماها را تقویت می‌کند؛ و γ وزن تابع مبتنی بر احتمال است که توزیع خوشه‌ها را پایدارتر کرده و از بیش‌برازش جلوگیری می‌کند. این تنظیمات به گونه‌ای انتخاب شدند که تعادلی میان بازسازی دقیق، حفظ سازگاری بین نماها و مقاومت در برابر داده‌های ناقص برقرار گردد.

٤-١ مجموعه داده‌ها

جدول ۴ ویژگی‌های اصلی دادگان‌های مورد استفاده در آزمایش‌ها را نشان می‌دهد. این دادگان‌ها شامل Handwritten، BDGP، MNIST Subset، Caltech-۲۰ و Fashion هستند که به‌طور گسترده در مطالعات خوشه‌بندی چندنما به کار گرفته می‌شوند. هر دادگان از نظر تعداد نمونه، تعداد نماها، ابعاد ویژگی‌ها و تعداد خوشه‌ها متفاوت است و این تنوع امکان ارزیابی جامع عملکرد مدل پیشنهادی را فراهم می‌سازد.

برای شبیه‌سازی شرایط دنیای واقعی و ارزیابی عملکرد مدل در سناریوهای مختلف، ناقص‌بودن به‌صورت مصنوعی بر روی دادگان‌های کامل اعمال شده است. فرآیند به این صورت است که برای هر نما، درصد مشخصی از نمونه‌ها به‌صورت تصادفی انتخاب شده و کل ویژگی‌های مربوط به آن نما حذف می‌شوند. به عبارت دیگر، اگر یک نمونه برای یک نما به‌عنوان "ناقص" انتخاب شود، تمام ابعاد ویژگی‌های آن نما برای آن

1. Epoch

جدول ۵: مشخصات ناقص‌بودن مصنوعی اعمال‌شده بر روی هر دادگان.

نام دادگان	نرخ‌های ناقص‌بودن اعمال‌شده	نحوه اعمال	تعداد تکرار
Handwritten	٪۱۰ ،٪۳۰ ،٪۵۰	تصادفی	۳
BDGP	٪۱۰ ،٪۳۰ ،٪۵۰	تصادفی	۳
MNIST Subset	٪۱۰ ،٪۳۰ ،٪۵۰	تصادفی	۳
Caltech-۲۰	٪۱۰ ،٪۳۰ ،٪۵۰	تصادفی+ ساختاریافته	۳
Fashion	٪۱۰ ،٪۳۰ ،٪۵۰	تصادفی	۳

نمونه با مقدار صفر یا (NaN) جایگزین می‌شوند. این روش که در ادبیات موضوع با عنوان نمای کامل مفقود[2] شناخته می‌شود، سناریوی رایج‌تری نسبت به حذف تصادفی تک‌ویژگی‌ها در کاربردهای واقعی است . برای هر دادگان، سه نرخ ناقص‌بودن مختلف در نظر گرفته شده است: ٪۱۰ نرخ پایین، ٪۳۰ نرخ متوسط و ٪۵۰ نرخ بالا. این نرخ‌ها بر اساس مطالعات پیشین در حوزه خوشه‌بندی چندنمای ناقص انتخاب شده‌اند تا توانایی مدل را در سطوح مختلف چالش‌برانگیزی ارزیابی کنند . برای اطمینان از عدم سوگیری، فرآیند انتخاب نمونه‌های ناقص به‌صورت تصادفی و با بذرهای مختلف (۳ بار تکرار) انجام شده و نتایج نهایی به‌صورت میانگین با انحراف معیار گزارش شده‌اند. همچنین، برای هر دادگان، یک سناریوی فقدان ساختاریافته[3] نیز در نظر گرفته شده است که در آن نماهای خاصی برای کلاس‌های خاصی حذف می‌شوند تا سناریوی واقعی‌تر (مانند فقدان یک نوع تصویربرداری خاص برای یک نوع بیماری خاص) شبیه‌سازی شود. مشخصات ناقص‌بودن مصنوعی اعمال‌شده بر روی هر دادگان در جدول ۵ آمده است.

کلیه نتایج گزارش‌شده ادامه مقاله، مربوط به نرخ ناقص‌بودن (٪۳۰) نرخ متوسط است که به‌عنوان نرخ مرجع در اکثر مطالعات مشابه انتخاب می‌شود.

٤-٢ معیارهای ارزیابی

برای ارزیابی نتایج مدل پیشنهادی با سایر روش ها از سه معیار ارزیابی ACC ، NMI و Purity استفاده شده است. این سه معیار در کنار یکدیگر تصویری جامع از کیفیت خوشه‌بندی ارائه می‌دهند و امکان مقایسه دقیق روش‌های مختلف را فراهم می‌سازند.

معیار ACC میزان تطابق بین برچسب‌های پیش‌بینی‌شده و برچسب‌های واقعی را نشان می‌دهد.

$$ACC=\frac{۱}{n}\sum_{i=۱}^{n}۱\{y_i=\max(c_i)\} \qquad (۱۲)$$

NMI با محاسبه اطلاعات متقابل نرمال‌شده، میزان هم‌پوشانی و سازگاری بین خوشه‌ها و کلاس‌های واقعی را ارزیابی می‌کند.

$$NMI=\frac{۲.I(Y;C)}{H(Y)+H(C)} \qquad (۱۳)$$

Purity نیز با بررسی غالب‌ترین کلاس در هر خوشه، میزان خلوص خوشه‌ها را مشخص می‌سازد.

2. Complete View Missing
3. Structured Missingness

جدول ۶: هایپرپارامترهای مدل HMCNet.

هایپرپارامتر	نماد	مقدار نهایی	محدوده جستجو	گام جستجو
وزن تابع بازسازی	α	۰/۵	[۰/۱–۰/۹]	۰/۱
وزن تابع سازگاری	β	۱/۰	[۰/۵–۲/۰]	۰/۲۵
وزن تابع احتمالی	γ	۰/۵	[۰/۱–۱/۰]	۰/۱
تعداد همسایگی UMAP	k	۱۵	[۵–۵۰]	۵
ابعاد خروجیUMAP	d_z	۶۴	[۱۶–۱۲۸]	۱۶

جدول ۷: تأثیر هایپرپارامترها بر عملکرد با تحلیل حساسیت(مقادیر ACC بر روی دادگان MNIST Subset).

مقدار α (با $\beta = ۱/۰ \ \gamma = ۰/۵$)	ACC	مقدار β (با $\alpha = ۰/۵ \ \gamma = ۰/۵$)	ACC	مقدار γ (با $\alpha = ۰/۵ \ \beta = ۱/۰$)	ACC
۰/۱	۷۴/۲	۰/۵۰	۷۲/۸	۰/۱	۷۵/۳
۰/۲	۷۶/۸	۰/۷۵	۷۶/۱	۰/۲	۷۷/۹
۰/۳	۷۸/۵	۱/۰۰	۸۰/۱	۰/۳	۷۹/۴
۰/۴	۷۹/۲	۱/۲۵	۷۹/۸	۰/۴	۸۰/۲
۰/۵	۸۰/۱	۱/۵۰	۸۰/۱	۰/۵	۸۰/۱
۰/۶	۷۹/۷	۱/۷۵	۷۹/۲	۰/۶	۷۹/۸
۰/۷	۷۸/۹	۲/۰۰	۷۷/۵	۰/۷	۷۹/۰
۰/۸	۷۷/۳	-	-	۰/۸	۷۸/۲
۰/۹	۷۵/۶	-	-	۰/۹	۷۶/۸

$$Purity = \frac{1}{n}\sum_{j=1}^{K}\max\left|C_j \cap L_l\right| \quad (۱۴)$$

۴-۳ روش‌های مقایسه

مدل HMCNet با روش های پیشرفته زیر مقایسه شده است:

- PVC: این روش مبتنی بر NMF است و تلاش می‌کند داده‌های ناقص چندنما را با تجزیه ماتریس و یادگیری فضای اجماعی خوشه‌بندی کند. مزیت آن سادگی و تفسیرپذیری است، اما در داده‌های غیرخطی و بزرگ محدودیت دارد [۳۲].
- UEAF: این مدل نماهای ناقص را بازسازی کرده و با یادگیری گراف اجماعی، هم‌ترازی نهفته بین نماها را حفظ می‌کند. مزیت آن مدیریت داده‌های ناقص و حفظ ساختار محلی است [۳۳].
- CDIMC-net: یک شبکه عمیق با رویکرد شناختی است که نمونه‌های مطمئن‌تر را ابتدا آموزش می‌دهد .مزیت آن استخراج ویژگی‌های سطح بالا و کاهش اثر نویز است [۳۳].
- DCP: این روش سازگاری پویا بین نماها را حفظ می‌کند تا داده‌های ناقص بهتر مدل شوند. مزیت آن توانایی در مدیریت الگوهای فقدان پیچیده است، اما نیازمند منابع محاسباتی بالاست [۲۶].
- IMCNet-MP: این مدل با استفاده از الگوهای چندسطحی و تنسورهای دووجهی، روابط مرتبه بالا بین نماها را استخراج می‌کند. مزیت آن مدیریت نرخ بالای داده‌های ناقص و یادگیری ساختارهای پیچیده است [۳۴].
- LMKML: این روش با یادگیری کرنل‌های چندگانه و منظم‌سازی رتبه پایین، کرنل اجماعی مقاوم در برابر داده‌های ناقص می‌سازد. مزیت آن توانایی در مدل‌سازی روابط غیرخطی است، اما هزینه محاسباتی بالایی دارد [۳۴].

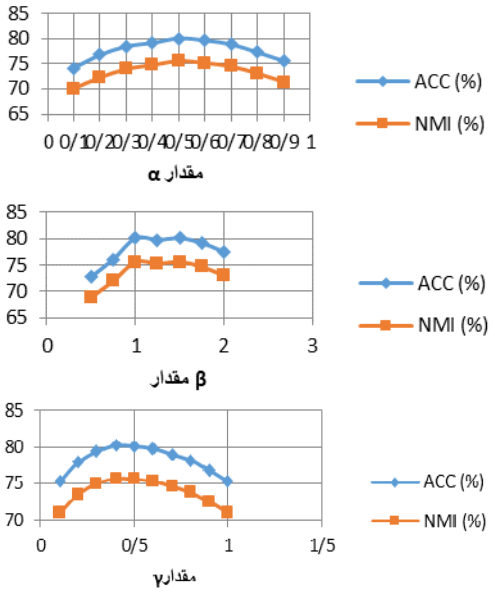
۴-۴ تنظیم فراپارامترها

فراپارامترهای مدل HMCNet در جدول ۶ آمده است. برای بررسی دقیق‌تر تأثیر فراپارامترهای مدل بر عملکرد نهایی، یک تحلیل حساسیت جامع انجام شد .جدول ۷، نتایج این تحلیل را بر روی دادگان MNIST Subset نشان می‌دهد.

با توجه به جدول ۷، بررسی ستون مربوط به α نشان می‌دهد که افزایش α از ۰/۱ به ۰/۵، معیار ACC را از ۷۴/۲٪ به ۸۰/۱٪ بهبود می‌بخشد که نقش مثبت تابع بازسازی را تأیید می‌کند. با این حال، افزایش بیشتر α به ۰/۹ باعث کاهش ACC به ۷۵/۶٪ می‌شود که بیانگر غلبه تابع بازسازی بر سایر اهداف و کاهش توانایی مدل در ایجاد نمایش مشترک میان نماهاست. در ستون مربوط به β، افزایش از ۰/۵ به ۱/۵، ACC را از ۷۸/۲٪ به ۸۰/۱٪ افزایش می‌دهد که نشان‌دهنده اهمیت بالای حفظ سازگاری بین نماهاست، اما افزایش به ۲/۰، ACC را به ۷۷/۵٪ کاهش می‌دهد که ناشی از غلبه بیش از حد تابع سازگاری است. در ستون مربوط به γ، افزایش از ۰/۱ به ۰/۵، ACC را از ۷۵/۳٪ به ۸۰/۱٪ افزایش می‌دهد که تأثیر مثبت تابع زیان احتمالی را نشان می‌دهد، اما افزایش بیشتر به ۱/۰، عملکرد را کاهش می‌دهد .نکته مهم :عملکرد مدل در بازه‌های α در [۰/۴–۰/۶]، β در [۱/۲۵–۱/۷۵] و γ در [۰/۴–۰/۶] دارای نوسان کمتر از ۱/۵٪ در ACC است که پایداری بالای مدل را نسبت به انتخاب دقیق فراپارامترها نشان می‌دهد.

روش تنظیم: فراپارامترها با استفاده از جستجوی شبکه‌ای جامع[1] بر روی مجموعه اعتبارسنجی که به‌صورت تصادفی ۲۰٪ از داده‌های آموزشی انتخاب شده بود، تعیین گردید. معیار انتخاب، بیشینه‌سازی معیار NMI بر روی داده‌های اعتبارسنجی بود. برای هر ترکیب از فراپارامترها، مدل بر روی ۸۰٪ داده‌های آموزشی (با اعمال نرخ ناقص‌بودن ۳۰٪ برای شبیه‌سازی شرایط واقعی) آموزش داده شده و بر روی ۲۰٪ اعتبارسنجی

1. Grid Search



شکل ۳: منحنی تغییرات عملکرد مدل نسبت به تغییر هر یک از هایپرپارامترهای اصلی بر روی دادگان MNIST Subset.

ارزیابی می‌شد. این فرآیند با استفاده از اعتبارسنجی متقاطع ۵-تایی[1] انجام شد تا از انتخاب فراپارامترهای پایدار و مستقل از تقسیم‌بندی داده‌ها اطمینان حاصل شود. بنابراین، تأکید می‌شود که مقادیر فراپارامترها صرفاً به‌صورت تجربی انتخاب نشده‌اند، بلکه حاصل یک فرآیند جستجوی سیستماتیک با اعتبارسنجی متقابل بوده و تحلیل حساسیت نیز پایداری مدل را در بازه‌های انتخابی تأیید کرده است. همچنین، برای اطمینان از عدم وابستگی نتایج به تنظیمات خاص، فرآیند جستجوی شبکه‌ای با بذرهای تصادفی مختلف سه بار تکرار شد که نتایج بسیار مشابهی (انحراف معیار کمتر از ۰٫۵٪ در NMI) به‌دست آمد.

لازم به ذکر است که برای دادگان‌های مختلف، این مقادیر با استفاده از همان روش جستجوی شبکه‌ای بر روی مجموعه اعتبارسنجی هر دادگان، به‌طور جداگانه بازتنظیم شدند. با این حال، به دلیل پایداری بالای مدل، تفاوت مقادیر بهینه در دادگان‌های مختلف کمتر از ۰٫۲ برای هر فراپارامتر بود که نشان‌دهنده قابلیت تعمیم‌پذیری مدل است و نیازی به تنظیم مجدد فراپارامترها برای هر دادگان جدید وجود ندارد.

تحلیل حساسیت: برای بررسی پایداری مدل نسبت به تغییر فراپارامترها، یک تحلیل حساسیت جامع انجام شد. در این تحلیل، هر فراپارامتر به‌صورت جداگانه در محدوده تعیین‌شده تغییر داده شد و سایر فراپارامترها در مقادیر بهینه خود ثابت نگه‌داشته شدند. نتایج نشان داد که مدل در بازه‌های ±۲۰٪ از مقادیر بهینه، عملکردی با نوسان کمتر از ۳٪ در معیار ACC دارد که بیانگر مقاومت مناسب مدل نسبت به تنظیمات فراپارامتری است. همچنین، تحلیل حساسیت نشان داد که مدل نسبت به تغییر β حساسترین و نسبت به تغییر γ کم‌حساس‌ترین واکنش را دارد که این موضوع با اهمیت نقش تابع سازگاری در داده‌های ناقص همخوانی دارد.

چرایی انتخاب مقادیر نهایی $\alpha = ۰٫۵$ ، $\beta = ۱٫۰$ و $\gamma = ۰٫۵$: انتخاب مقدار $\alpha = ۰٫۵$ به‌دلیل ایجاد تعادل بهینه بین بازسازی داده‌ها و سایر اهداف صورت گرفته است. تحلیل حساسیت نشان داد که در بازه [۰٫۴, ۰٫۶]، عملکرد مدل نوسانی کمتر از ۱٪ در ACC دارد و مقدار ۰٫۵ به‌عنوان نقطه میانی این بازه انتخاب شده است. مقدار $\beta = ۱٫۰$ اگر چه تحلیل حساسیت نشان داد $\beta = ۱٫۰$ نیز عملکرد مشابهی دارد، اما مقدار ۱٫۰ به دلیل سادگی تفسیر (نسبت ۱:۱ بین بازسازی و سازگاری) و جلوگیری از غلبه بیش از حد تابع سازگاری بر سایر توابع، انتخاب گردید. همچنین در بازه [۰٫۷۵–۱٫۵]، نوسان عملکرد کمتر از ۲٪ در ACC بود که پایداری مناسبی را نشان می‌دهد. مقدار $\gamma = ۰٫۵$ نیز به‌دلیل تأثیر مثبت تابع احتمالی در پایدارسازی خوشه‌بندی و جلوگیری از غلبه اطلاعات توزیعی GMM بر سایر اطلاعات انتخاب شده است. بازه [۰٫۴, ۰٫۶]، برای γ نیز نوسانی کمتر از ۱٫۵٪ در ACC دارد که تأییدکننده مقاومت مدل در این بازه است. وزن تابع UMAP در (۱۱) به‌صورت ضریب ثابت ۱ در نظر گرفته شده و نیازی به تنظیم مجدد ندارد، زیرا تأثیر آن از طریق پارامترهای تعداد همسایگی (k)و ابعاد خروجی که به‌صورت جداگانه با استفاده از معیارهای اعتمادپذیری و تداوم تنظیم شده‌اند، اعمال می‌شود. شکل ۳ منحنی تغییرات عملکرد مدل را نسبت به تغییر هر یک از فراپارامترهای اصلی نشان می‌دهد.

٤-٥ مقایسه نتایج

برای ارزیابی عملکرد مدل پیشنهادی، نتایج حاصل از آزمایش‌ها بر روی پنج دادگان واقعی شامل Handwritten، BDGP، MNIST Subset، Caltech-۲۰ و Fashion گزارش شده‌اند. همان‌طور که در جدول‌های ۸ تا ۱۲ مشاهده می‌شود، روش‌های کلاسیک مانند PVC در تمامی معیارها عملکرد ضعیف‌تری دارند که نشان‌دهنده محدودیت آن‌ها در مدیریت داده‌های ناقص و روابط غیرخطی است. روش UEAF با بهره‌گیری از هم‌ترازی نهفته و یادگیری گراف اجماعی، بهبود قابل توجهی نسبت به PVC ارائه می‌دهد، اما همچنان در دادگان‌های پیچیده‌تر مانند Caltech-۲۰ و Fashion دچار افت عملکرد می‌شود.

در مقابل، روش‌های مبتنی بر یادگیری عمیق نظیر CDIMC-net و DCP توانسته‌اند با استخراج ویژگی‌های سطح بالا و حفظ سازگاری پویا بین نماها، نتایج بهتری در معیارهای ACC، NMI وPurity کسب کنند. IMCNet-MP نیز با طراحی چندسطحی و یادگیری روابط مرتبه بالا، در اکثر دادگان‌ها عملکردی نزدیک به بهترین دارد. روش کرنل LMKML در دادگان‌های ساده‌تر مانند Handwritten و BDGP نتایج مناسبی ارائه می‌دهد، اما در داده‌های پیچیده‌تر به دلیل هزینه محاسباتی بالا و محدودیت در انتخاب کرنل، نسبت به روش‌های عمیق عقب می‌ماند.

در نهایت، مدل پیشنهادی HMCNet در تمامی دادگان‌ها بالاترین یا نزدیک به بالاترین مقادیر را در چهار معیار اصلی به دست آورده است. این برتری ناشی از ترکیب ماژول توجه برای وزن‌دهی پویا به نماها، استفاده از UMAP برای کاهش ابعاد با حفظ ساختار محلی، و طراحی خطای دوگانه برای تقویت سازگاری و پایداری خوشه‌بندی است. بنابراین، نتایج نشان می‌دهند که HMCNet نه تنها در داده‌های ساده بلکه در شرایط پیچیده و ناقص نیز عملکردی پایدار و برتر دارد.

برای بررسی تأثیر هم‌کنشی هایپرپارامترها با نرخ ناقص‌بودن داده‌ها، آزمایشی طراحی شد که در آن مدل با هایپرپارامترهای بهینه تنظیم شده در نرخ‌های مختلف ناقص‌بودن (از ۱۰٪ تا ۷۰٪) بر روی دادگان MNIST Subset ارزیابی گردید. نتایج در جدول ۶ نشان می‌دهد که هایپرپارامترهای بهینه در نرخ‌های پایین ناقص‌بودن (≤۳۰٪) کاملاً پایدار هستند (تغییر کمتر از ۱٪ در ACC). در نرخ‌های بالاتر (< ۵۰٪)، مقدار بهینه β به‌سمت مقادیر بزرگ‌تر (تا ۲/۰) تغییر می‌کند که نشان‌دهنده اهمیت بیشتر حفظ سازگاری بین نماها در شرایط فقدان شدید اطلاعات

1. Fold Cross-Validation

جدول ۸: مقایسه نتایج روش های مختلف بر روی دادگان Handwritten.

نوع روش	نام روش	ACC	NMI	Purity
ماتریس‌فاکتوریزیشن	PVC	۹۰/۱	۸۲/۴	۹۰/۱
گراف‌محور	UEAF	۹۲/۷	۸۵/۹	۹۲/۷
	CDIMC-net	۹۴/۵	۸۸/۲	۹۴/۵
یادگیری عمیق	DCP	۹۱/۳	۸۳/۷	۹۱/۳
	IMCNet-MP	۹۷/۶	۹۴/۳	۹۷/۶
مبتنی بر کرنل	LMKML	۹۸/۱	۹۵/۰	۹۸/۱
روش پیشنهادی	HMCNet	۹۳/۵	۸۶/۷	۹۳/۲

جدول ۹: مقایسه نتایج روش های مختلف بر روی دادگان BDGP.

نوع روش	نام روش	ACC	NMI	Purity
ماتریس‌فاکتوریزیشن	PVC	۸۲/۴	۷۵/۱	۸۲/۴
گراف‌محور	UEAF	۸۷/۲	۷۹/۳	۸۷/۲
	CDIMC-net	۹۰/۵	۸۲/۷	۹۰/۵
یادگیری عمیق	DCP	۹۲/۱	۸۴/۰	۹۲/۱
	IMCNet-MP	۸۸/۹	۸۰/۲	۸۸/۹
مبتنی بر کرنل	LMKML	۹۶/۶	۸۷/۸	۹۵/۴
روش پیشنهادی	HMCNet	۹۷/۰	۸۵/۵	۹۶/۰

جدول ۱۰: مقایسه نتایج روش های مختلف بر روی دادگان MNIST Subset.

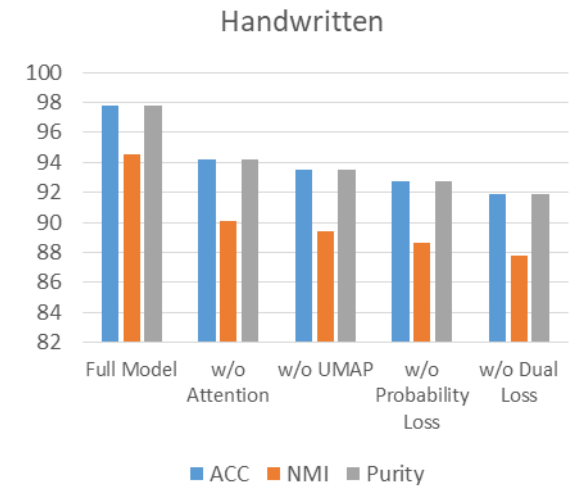
نوع روش	نام روش	ACC	NMI	Purity
ماتریس‌فاکتوریزیشن	PVC	۶۵/۷	۶۰/۲	۶۵/۷
گراف‌محور	UEAF	۷۰/۴	۶۴/۱	۷۰/۴
	CDIMC-net	۷۴/۸	۶۸/۵	۷۴/۸
یادگیری عمیق	DCP	۷۶/۹	۷۱/۲	۷۶/۹
	IMCNet-MP	۷۲/۳	۶۶/۰	۷۲/۳
مبتنی بر کرنل	LMKML	۷۸/۲	۷۵/۵	۷۸/۲
روش پیشنهادی	HMCNet	۸۰/۱	۸۶/۸	۸۰/۱

جدول ۱۱: مقایسه نتایج روش های مختلف بر روی دادگان Caltech.

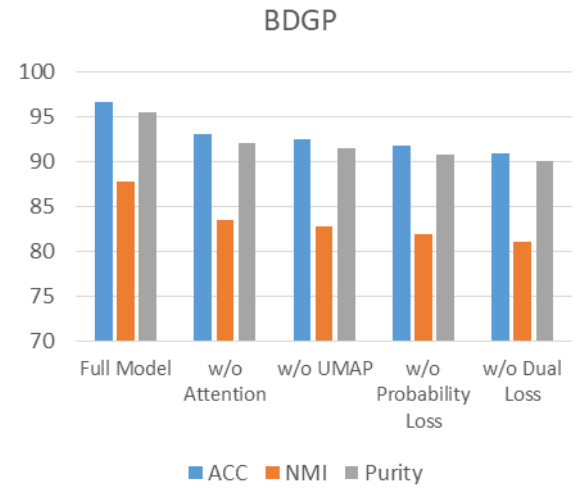
نوع روش	نام روش	ACC	NMI	Purity
ماتریس‌فاکتوریزیشن	PVC	۵۸/۹	۵۵/۳	۷۰/۲
گراف‌محور	UEAF	۶۲/۷	۵۹/۱	۷۳/۵
	CDIMC-net	۶۵/۸	۶۲/۴	۷۶/۰
یادگیری عمیق	DCP	۶۷/۹	۶۴/۲	۷۸/۱
	IMCNet-MP	۶۳/۴	۶۰/۰	۷۴/۰
مبتنی بر کرنل	LMKML	۶۸/۶	۹۶/۶	۸۱/۶
روش پیشنهادی	HMCNet	۷۰/۲	۷۰/۸	۸۲/۴

جدول ۱۲: مقایسه نتایج روش های مختلف بر روی دادگان Fashion.

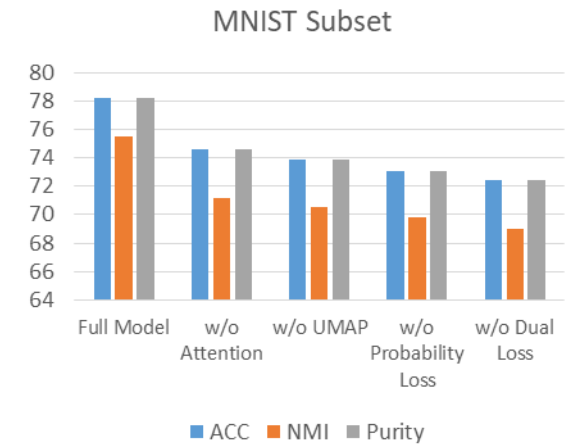
نوع روش	نام روش	ACC	NMI	Purity
ماتریس‌فاکتوریزیشن	PVC	۷۲/۱	۶۶/۵	۷۲/۱
گراف‌محور	UEAF	۷۵/۸	۶۹/۹	۷۵/۸
	CDIMC-net	۷۷/۹	۷۱/۸	۷۷/۹
یادگیری عمیق	DCP	۷۸/۶	۷۲/۵	۷۸/۶
	IMCNet-MP	۷۴/۳	۶۸/۵	۷۴/۳
مبتنی بر کرنل	LMKML	۸۵/۱	۷۸/۲	۸۵/۱
روش پیشنهادی	HMCNet	۷۲/۱	۶۶/۵	۷۲/۱



شکل ۴: نتایج مطالعه حذفی بر روی دادگان Handwritten.



شکل ۵: نتایج مطالعه حذفی بر روی دادگان BDGP.

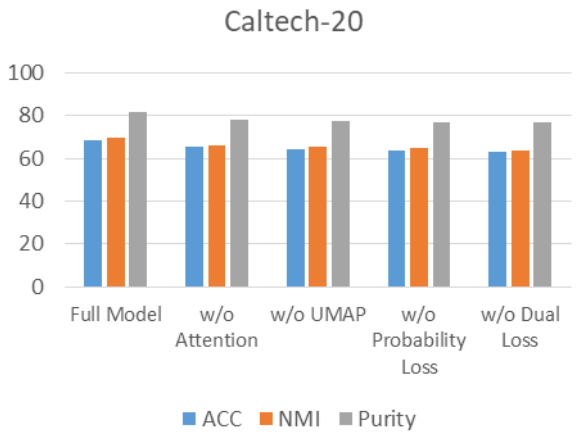


شکل ۶: نتایج مطالعه حذفی بر روی دادگان MNIST Subset.

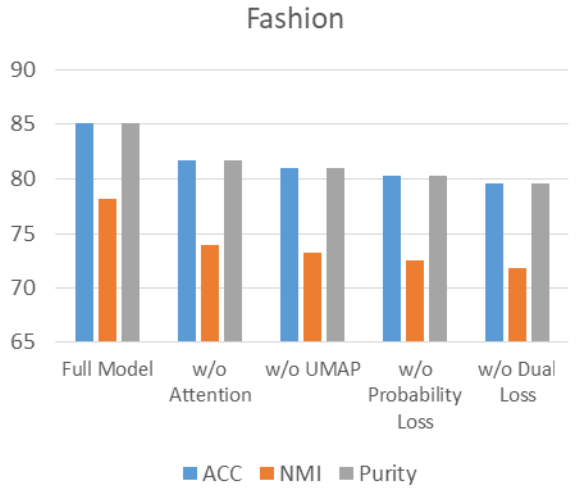
است. این مشاهده نشان می‌دهد که برای کاربردهایی با نرخ‌های بالای ناقص‌بودن، بازتنظیم هایپرپارامتر β می‌تواند عملکرد مدل را بهبود بخشد.

٤-٦ آزمایش‌های حذف

برای بررسی نقش هر مؤلفه در عملکرد مدل پیشنهادی HMCNet، مجموعه‌ای از آزمایش‌های حذف طراحی شد. در این آزمایش‌ها، ماژول‌های کلیدی مدل به‌صورت جداگانه حذف یا جایگزین شدند و نتایج



شکل ۷: نتایج مطالعه حذفی بر روی دادگان Caltech-۲۰.



شکل ۷: نتایج مطالعه حذفی بر روی دادگان Fashion.

بر روی دادگان‌های Handwritten، BDGP، MNIST Subset، Caltech-۲۰ و Fashion گزارش گردید.

بدون ماژول توجه: حذف مکانیزم توجه منجر به کاهش قابل توجه در معیارهای ACC و NMI شد، زیرا وزن‌دهی پویا به نماها نقش مهمی در استخراج ویژگی‌های سازگار دارد.

بدون UMAP: جایگزینی UMAP با یک روش خطی مانند PCA باعث افت عملکرد در معیار ARI شد، که نشان‌دهنده اهمیت حفظ ساختار محلی داده‌ها در کاهش ابعاد است.

بدون تابع زیان احتمالی: حذف این مؤلفه موجب کاهش پایداری خوشه‌بندی و حساسیت بیشتر به نویز شد، به‌ویژه در دادگان‌های پیچیده مانند Caltech-20.

بدون خطای دوگانه فقط بازسازی: استفاده صرف از بازسازی نماها باعث شد مدل نتواند سازگاری بین نماها را به‌خوبی حفظ کند و در نتیجه عملکرد کلی کاهش یافت.

نتایج مطالعه حذفی[1] ارائه‌شده در شکل‌های ۴ تا ۸ نشان می‌دهند که هر مؤلفه به‌طور مستقل سهم قابل توجهی در بهبود عملکرد دارد، اما ترکیب همزمان آن‌ها در چارچوب HMCNet بهترین نتایج را به دست می‌دهد. به‌ویژه، ماژول توجه و تابع زیان احتمالی بیشترین تأثیر را در افزایش دقت و پایداری خوشه‌بندی داشته‌اند. نتایج تحلیل حساسیت در

1. Ablation Study

جدول ۶ نشان داد که در نرخ‌های پایین ناقص‌بودن (≤ ۳۰٪)، هایپرپارامترهای بهینه کاملاً پایدار هستند و تغییرات آن‌ها تأثیر معناداری بر عملکرد مدل ندارد.

۷-٤ عملکرد در نرخ‌های مختلف ناقص‌بودن

برای ارزیابی جامع‌تر عملکرد مدل پیشنهادی در سناریوهای مختلف، آزمایش‌هایی با سه نرخ ناقص‌بودن ۱۰٪، ۳۰٪ و ۵۰٪ بر روی دادگان MNIST Subset (به‌عنوان دادگان نماینده) انجام شد. شکل ۸ عملکرد مدل HMCNet را در مقایسه با روش‌های رقیب در این نرخ‌ها نشان می‌دهد.

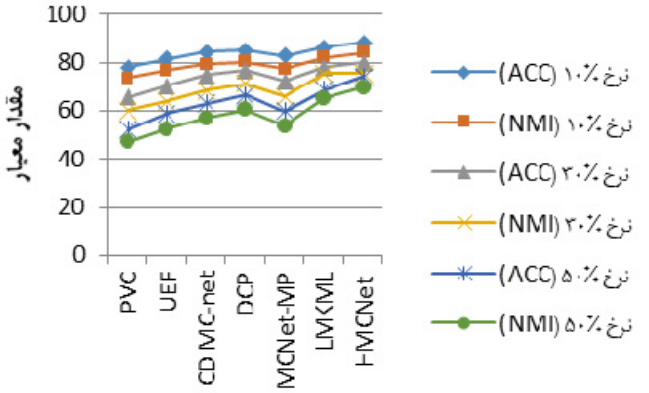
همان‌طور که در شکل ۹ مشاهده می‌شود، در نرخ ناقص‌بودن ۱۰٪ (نرخ پایین)، اکثر روش‌ها عملکرد نسبتاً مشابهی دارند و تفاوت بین بهترین و ضعیف‌ترین روش حدود ۱۰٪ است. با افزایش نرخ ناقص‌بودن به ۳۰٪ (نرخ متوسط)، شکاف عملکردی بین روش‌ها بیشتر شده و روش‌های سنتی مانند PVC و UEF کاهش عملکرد محسوسی (حدود ۱۵٪–۱۲٪) نشان می‌دهند. در نرخ ۵۰٪ (نرخ بالا)، عملکرد روش‌های سنتی به‌شدت کاهش می‌یابد (بیش از ۲۵٪ افت ACC نسبت به نرخ ۱۰٪)، در حالی که HMCNet با افت کمتر از ۱۵٪، پایداری بالایی از خود نشان می‌دهد. این مزیت عمدتاً ناشی از ترکیب ماژول توجه (که وزن‌دهی پویا به نماهای موجود را ممکن می‌سازد) و تابع زیان احتمالی (که از برچسب‌های نرم GMM برای پایدارسازی خوشه‌بندی در حضور اطلاعات محدود استفاده می‌کند) است. همچنین، روش LMKML به‌عنوان یک روش مبتنی بر کرنل، عملکرد نسبتاً خوبی در نرخ‌های پایین و متوسط دارد، اما در نرخ ۵۰٪ با افت حدود ۱۷٪ مواجه می‌شود که نشان‌دهنده وابستگی آن به کرنل‌های کامل است.

۵- نتیجه گیری و راهکارهای آتی

در این پژوهش، مدل HMCNet برای خوشه‌بندی داده‌های چندنما در شرایط ناقص بودن نماها معرفی و ارزیابی شد. نتایج آزمایش‌ها بر روی پنج دادگان استاندارد نشان داد که مدل پیشنهادی به‌طور پایدار عملکردی برتر نسبت به روش‌های موجود دارد و توانسته است در معیارهای ACC، NMI و Purity بهبود قابل توجهی ایجاد کند. این برتری ناشی از ترکیب ماژول توجه، UMAP Embedding و تابع زیان ترکیبی مبتنی بر احتمال است که به حفظ ساختار محلی داده‌ها و افزایش سازگاری میان نماها کمک می‌کند.

علاوه بر این، تحلیل‌های آزمایش‌های حذف اهمیت هر مؤلفه را در بهبود عملکرد مدل تأیید کردند. حذف هر بخش از معماری منجر به افت محسوس در نتایج شد، به‌ویژه ماژول توجه و تابع زیان ترکیبی مبتنی بر احتمال که بیشترین سهم را در افزایش دقت و پایداری خوشه‌بندی داشتند. بنابراین، می‌توان نتیجه گرفت که HMCNet نه تنها در شرایط داده‌های کامل بلکه در نرخ‌های بالای ناقص بودن نیز کارایی بالایی دارد و می‌تواند به‌عنوان یک چارچوب عمومی برای مسائل خوشه‌بندی چندنما در حوزه‌های علمی و صنعتی مورد استفاده قرار گیرد.

با وجود نتایج مطلوب، مدل HMCNet دارای محدودیت‌هایی است که ذکر آن‌ها برای ارزیابی منصفانه و هدایت پژوهش‌های آتی ضروری است. مدل دارای پنج هایپرپارامتر اصلی $(\alpha,\beta,\gamma,k,d_z)$ است که تنظیم آن‌ها نیازمند جستجوی شبکه‌ای زمان‌بر و مجموعه اعتبارسنجی جداگانه است. معماری فعلی برای داده‌های ایستا طراحی شده و برای داده‌های جریانی که نمونه‌ها به‌تدریج وارد می‌شوند، نیاز به بازطراحی دارد. مؤلفه



شکل۹: نتایج عملکرد در نرخ‌های مختلف ناقص‌بودن.

UMAP به‌عنوان یک مرحله جداگانه و غیرقابل‌تمایز طراحی شده است که مانع از آموزش سرتاسری مدل و بهینه‌سازی همزمان تمام مؤلفه‌ها می‌شود. بر اساس محدودیت‌های شناسایی‌شده، جهت زیر برای پژوهش آینده پیشنهاد می‌شود: جایگزینی UMAP با یک لایه قابل‌تمایز مانند خودرمزگذارهای عمیق با ساختار منظم‌سازی شده برای امکان آموزش سرتاسری.

مراجع

[1] K. Berahmand, S. Bahadori, M. N. Abadeh, Y. Li, and Y. Xu, "SDAC-DA: Semi-supervised deep attributed clustering using dual autoencoder," *IEEE Trans. on Knowledge and Data Engineering,* vol. 36, no. 11, pp. 6989-7002, 2024.

[2] X.-L. Li, M.-S. Chen, C.-D. Wang, and J.-H. Lai, "Refining graph structure for incomplete multi-view clustering *IEEE Trans. on Neural Networks and Learning Systems,* vol. 35, no. 2, pp. 2300-2313, 2022.

[3] Y. Chen, X. Zhao, Z. Zhang ,Y. Liu, J. Su, and Y. Zhou, "Tensor learning meets dynamic anchor learning: From complete to incomplete multiview clustering," *IEEE Trans. on Neural Networks and Learning Systems,* vol. 35, no. 11, pp. 15332-15345, 2023.

[4] J. Wen, *et al.*, "Structural deep incomplete multi-view clustering network," in *Proc, of the 30th ACM Int. Conf. on Information & Knowledge Management*, pp. 3538-3542, Virtual Event Queensland, Australia, 1-5 Nov. 2021.

[5] L. McInnes, J. Healy, and J. Melville, *Umap: Uniform Manifold Aapproximation and Projection for Dimension Reduction, arXiv preprint arXiv:1802.03426,* 2018.

[6] C. Leng, H. Zhang, G. Cai, I. Cheng, and A. Basu, "Graph regularized Lp smooth non-negative matrix factorization for data representation," *IEEE/CAA Journal of Automatica Sinica,* vol. 6, no. 2, pp. 584-595, 2019.

[7] G. A. Khan, "Complementary incomplete weighted concept factorization methods for multi-view clustering," *Knowledge and Information Systems,* vol. 66, no. 12, pp. 7469-7494, 2024.

[8] J. Wen, *et al.*, "Graph regularized and feature aware matrix factorization for robust incomplete multi-view clustering," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 34, no. 5, pp. 3728-3741, 2023.

[9] Y. Ban, Y. Cai, and Z. Huang, "Autoencoder-like non-negative matrix factorization with dual-graph constraints for multi-view clustering," *International Journal of Machine Learning and Cybernetics,* pp. 1-16, 2025.

[10] Y. Khalafaoui, N. Grozavu, B. Matei, and L. -W. Goix, "Multi-modal multi-view clustering based on non-negative matrix factorization," in *Proc. IEEE Symp. Series on Computational Intelligence*, pp. 1386-1391, Singapore, Singapore, 4-7 Dec. 2022.

[11] Z. Shu, Q. Long, L. Zhang, Z. Yu, and X.-J. Wu, "Robust graph regularized NMF with dissimilarity and similarity constraints for ScRNA-seq data clustering," *Journal of Chemical Information and Modeling,* vol. 62, no. 23, pp. 6271-6286, 2022.

[12] W .Lv, C. Zhang, H. Li, X. Jia, and C. Chen, "Joint projection learning and tensor decomposition-based incomplete multiview clustering," *IEEE Trans. on Neural Networks and Learning Systems,* vol. 35, no. 12, pp. 17559-17570, Dec. 2024.

[13] W. Li and C. Zhu, "Cross-view attention with adversarial learning for incomplete multi-view clustering," *Multimedia Systems,* vol. 31, no. 5, p. 347, 2025.

[14] T. Wu, S. Feng, and J. Yuan, "Low-rank kernel tensor learning for incomplete multi-view clustering," in *Proc. of the AAAI Conf. on Artificial Intelligence*, pp. 15952-15960, 20-27 Feb. 2024.

[15] M. Briscik, G. Tazza, L. Vidács, M. -A. Dillies, and S. Déjean, "Supervised multiple kernel learning approaches for multi-omics data integration," *BioData Mining*, vol. 17, Article ID: 53, 2024.

[16] Z. Wen, *et al.*, *MultiKernelBench: A Multi-Platform Benchmark for Kernel Generation*, arXiv eprints, pp. arXiv–2507, 2025.

[17] R. Zhang, Y. Li, Y. Gui, D. J. Armaghani, and M. Yari, "A stacked deep multi-kernel learning framework for blast induced flyrock prediction," *International Journal of Rock Mechanics and Mining Sciences,* vol. 177, Article ID: 105741, 2024.

[18] S. Wang, *et al.*, "Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph," in *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pp. 9776-9785, New Orleans, LA, USA, 18-24 Jun. 2022.

[19] J. Wen, Z. Zhang, Y. Xu, and Z. Zhong, "Incomplete multi-view clustering via graph regularized matrix factorization," in *Proc. of the European Conf. on Computer Vision Workshops*, pp. 593-608, Munich, Germany, 8-14 Sept. 2018.

[20] J. Cao, J. Hu, and R. Zhang, "Adaptive graph learning algorithm for incomplete multi-view clustered image segmentation," *Engineering Applications of Artificial Intelligence,* vol. 139, Article ID: 109264, 2025.

[21] Z. Shu, B. Li, C. Hu, Z. Yu, and X.-J. Wu, "Robust dual-graph regularized deep matrix factorization for multi-view clustering," *Neural Processing Letters,* vol. 55, no. 5, pp. 6067-6087, 2023.

[22] D. Zhuang, *et al.*, "Robust graph contrastive learning for incomplete multi-view clustering," in *Proc. of the Thirty-Fourth Int. Joint Conf. on Artificial Intelligence*, pp. 7282-7290, Montreal, Canada, 16-22 Aug. 2025.

[23] W. Xia, Q. Wang, Q. Gao ,X. Zhang, and X. Gao, "Self-supervised graph convolutional network for multi-view clustering," *IEEE Trans. on Multimedia,* vol. 24, pp. 3182-3192, 2021.

[24] Y. Wen *et al.*, "Scalable incomplete multi-view clustering with structure alignment," in *Proc. of the 31st ACM Int. Conf. on Multimedia*, pp. 3031-3040, Ottawa, Canada, 29 Oct.- 3 Nov 2023.

[25] J. Xu, *et al.*, "Adaptive feature projection with distribution alignment for deep incomplete multi-view clustering," *IEEE Trans. on Image Processing,* vol. 32 ,pp. 1354-1366, 2023.

[26] W. Yan, K. Liu, W. Zhou, and C. Tang, "Deep incomplete multi-view clustering via dynamic imputation and triple alignment with dual optimization," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 35, no. 4, 3250 – 3261, Apr. 2024.

[27] Y. Li and B. Wu, "Joint generative and alignment adversarial learning for robust incomplete multi-view clustering," *Neural Networks,* vol. 194, Article ID: 108141, Feb. 2026.

[28] J. Yin, P. Wang, S. Sun, and Z. Zheng, "Incomplete multi-view clustering via multi-level contrastive learning," *IEEE Trans. on Knowledge and Data Engineering*, vol. 37, no. 8, pp. 4716 - 4727, Aug. 2025.

[29] Q. Qi, Z. Qiu, Y. Yan, Y. Lu, and H. Wang, "IMC-Det: Intra-inter modality contrastive learning for video object detection," *International Journal of Computer Vision,* vol. 133, no. 2, pp. 890-909, 2025.

[30] J. Wen *et al.*, "Dimc-net: Deep incomplete multi-view clustering network," in *Proc. of the 28th ACM Int. Conf. on Multimedia*, pp. 3753-3761, Seattle, WA, USA, 12-16 Oct. 2020.

[31] C. M. Bishop, *Pattern Recognition and Machine Learning,* Springer Science+Business Media, 2006.

[32] S. -Y. Li, Y. Jiang, and Z. -H. Zhou, "Partial multi-view clustering," in *Proc. of 28th AAAI Conf. on Artificial Intelligence*, pp. 1968-1974, Québec, Canada, 27-31 Jul. 2014.

[33] W. Jie, *et al.,* "CDIMC-net: Cognitive deep incomplete multi-view clustering network," in *Proc. of the 29th Int. Joint Con. on Artificial Intelligence*, pp. 3230-3236, Yokohama, Japan, Jan. 2020.

[34] J. Luo and C. Zhu, "Incomplete multiview clustering with bipartite tensors," *Applied Intelligence,* vol. 55, no. 13, Article ID: 924, 2025.

سندس بهادری تحصیلات مقاطع کارشناسی و کارشناسی ارشد و دکتری خود در رشته سیستم‌های نرم افزاری در سال‌های ۱۳۸۴ و ۱۳۸۷ و ۱۳۹۷ از دانشگاه بوعلی همدان، و دانشگاه آزاد اسلامی به پایان رسانده است و هم‌اکنون استادیار دانشکده مهندسی کامپیوتر دانشگاه آزاد اسلامی واحد ایلام می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: تحلیل شبکه‌های پیچیده، محاسبات نرم و کاربردهای آن در توسعه نرم‌افزارهای بزرگ.