Adaptive Feature Imputation with Latent Graph for Deep Incomplete Multi-View Clustering

Jingyu Pu[1], Chenhang Cui[1], Xinyue Chen[1], Yazhou Ren[1, 2*], Xiaorong Pu[1, 2], Zhifeng Hao[3], Philip S. Yu[4], Lifang He[5]

[1]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
[2]Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, China
[3]College of Science, Shantou University, Shantou, China
[4]Department of Computer Science, University of Illinois Chicago, Chicago, IL, USA
[5]Department Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA
pujingyu0105@163.com, {chenhangcui, martinachen2580}@gmail.com,
{yazhou.ren, puxiaor}@uestc.edu.cn, haozhifeng@stu.edu.cn, psyu@uic.edu, lih319@lehigh.edu

Abstract

In recent years, incomplete multi-view clustering (IMVC), which studies the challenging multi-view clustering problem of missing views, has received growing research interests. Previous IMVC methods suffer from the following issues: (1) the inaccurate imputation for missing data, which leads to suboptimal clustering performance, and (2) most existing IMVC models merely consider the explicit presence of graph structure in data, ignoring the fact that latent graphs of different views also provide valuable information for the clustering task. To overcome these challenges, we present a novel method, termed adaptive feature imputation with latent graph for incomplete multi-view clustering (AGDIMC). Specifically, it captures the embedded features of each view by incorporating the view-specific deep encoders. Then, we construct partial latent graphs on complete data, which can consolidate the intrinsic relationships within each view while preserving the topological information. With the aim of estimating the missing sample based on the available information, we utilize an adaptive imputation layer to impute the embedded feature of missing data by using cross-view soft cluster assignments and global cluster centroids. As the imputation progresses, the portion of complete data increases, contributing to enhancing the discriminative information contained in global pseudo-labels. Meanwhile, to alleviate the negative impact caused by inferior impute samples and the discrepancy of cluster structures, we further design an adaptive imputation strategy based on the global pseudo-label and the local cluster assignment. Experimental results on multiple real-world datasets demonstrate the effectiveness of our method over existing approaches.

Introduction

The proliferation of multi-view data has created a compelling demand for researchers to analyze this intricate information. As an important unsupervised learning approach, multi-view clustering has been widely applied in real-world applications, such as recommendation systems, multimedia analysis, and bioinformatics (Nie, Cai, and Li 2017; Zhan et al. 2018; Chen et al. 2020; Huang, Kang, and Xu 2020; Xu et al. 2021). Most existing multi-view clustering methods heavily rely on the assumption of fully available data. However, it is a commonly encountered scenario that only partial data can be collected and transmitted owing to factors such as unstable sensors and damaged storage media. Therefore, an increasing attention is paid to partial multi-view clustering or incomplete multi-view clustering (IMVC) problems (Peng et al. 2019; Lin et al. 2021, 2022).

To handle the incompleteness issue, traditional incomplete multi-view clustering methods with satisfactory performance have been proposed. Typical strategies are mainly based on matrix decomposition (Li, Jiang, and Zhou 2014; Zhao, Liu, and Fu 2016; Hu and Chen 2019b), incomplete multiple kernel learning (Liu et al. 2020; Guo and Ye 2019) and graph-based methods (Xu et al. 2018; Wen et al. 2019, 2021). Improving the performance of incomplete multi-view clustering relies on the ability to learn more discriminative consensus representations despite the presence of incomplete view information. Most methods naturally focus on addressing the challenge of missing data by imputing, recovering, or inferring values for the missing samples.

However, the above traditional IMVC methods all exploit the shallow learning based methods, which are not effective enough to excavate the in-depth knowledge hidden in the data (Zhao et al. 2018). In recent years, deep incomplete multi-view clustering (DIMVC) have demonstrated remarkable progress by integrating clustering with the representation learning capabilities of deep models (Wen et al. 2019; Wang et al. 2020; Lin et al. 2022). Most existing DIMVC methods utilize the generalization capability of deep models to achieve the imputation for missing data (Yang et al. 2022).

Although existing DIMVC methods have achieved significant progress through imputation, they have at least two issues. On the one hand, the effectiveness of these methods is limited as they cannot exploit the topological information of samples. On the other hand, imputing missing data accurately based on observed data poses a significant challenge,

*Corresponding author

especially when a large number of samples are missing. Additionally, assessing the quality of imputation becomes complex due to the unavailability of ground truth for the missing data.

In this paper, we propose Adaptive Feature Imputation with Latent Graph for Deep Incomplete Multi-View Clustering (AGDIMC) to address the aforementioned issues. Specifically, we employ latent graphs to enable embedded features to capture topological information via GCN module, which mines the potential structure information. Concretely, to reduce the negative effects from low-quality imputed features, we design an adaptive imputation module to assess the quality of imputed features by measuring the difference between two soft pseudo assignments derived from original embedded features and imputed features, respectively. With this assessment, imputation can be circulatively adapted. After obtaining a more consistent representation from imputed features refined by weight graph fusion, the complementary cluster information contained in this representation transforms to supervised information with high confidence. This transformation aims to achieve consistent cluster assignments for all views. Moreover, to pay more attention to the reliable neighbors in latent graphs, we introduce the graph embedding constraint to preserve this relation. We illustrate the framework of our proposed AGDIMC in Fig. 1. In summary, our key contributions are as follows.

- We propose AGDIMC, a novel deep model for incomplete multi-view clustering which can capture topological information by leveraging both weight graph fusion (that containing complementary information) and latent graphs (capturing topological details) to perform deep incomplete multi-view clustering.
- We propose a novel adaptive imputation module that can evaluate the quality of imputation effectively in an unsupervised manner which reduces the negative effects from low-quality imputed features. Moreover, it boosts the confidence of assignments by improving the consistency of incomplete data.
- We propose a graph embedding constraint to reinforce the neighbor relationships, which helps prevent distortion of discriminative information during fusion. Extensive experiments demonstrate that our method achieves superior clustering performance compared to existing state-of-the-art methods.

Related Work

Incomplete Multi-View Clustering

Multi-view clustering is a data analysis technique that clusters data points using multiple sets of distinctive features or viewpoints. However, dealing with incomplete multi-view data, where some views are missing or partially available, presents a challenge in transitioning from traditional multiple-view clustering to incomplete multi-view clustering. Unlike traditional multi-view clustering methods that assume complete and consistent information across all views, incomplete multi-view clustering (IMVC) (Wen et al. 2022; Trivedi et al. 2010) deal with scenarios where some views may be missing or partially observed for certain data points. Many methods aim to obtain consistent information from multiple views to handle the incomplete problem. (Lin et al. 2021) maximized mutual information between different views through contrastive learning, enabling the learning of informative and consistent representations. (Xu et al. 2019) simultaneously learned a common latent space and infers missing data using element-wise reconstruction and generative adversarial network (GAN). (Hu and Chen 2019a) leveraged instance alignment information to learn a common latent feature matrix for all views. (Xu et al. 2023) proposed an imputation-free deep IMVC method that incorporates distribution alignment in feature learning. The method employs adaptive feature projection to avoid the need for imputation of missing data.

Graph-based Incomplete Multi-View Clustering

The graph structure represents dependencies among data points, enables information propagation, and allows for diverse clustering techniques. In the context of handling multi-view data with missing information, graph-based incomplete multi-view clustering (Wen et al. 2018; Wu et al. 2018) emerges as a suitable method. It leverages graph-based techniques to capture relationships, handle missing data, and integrate multiple views. In recent years, many graph-related methods have been developed with the aim of completing incomplete graphs to obtain more accurate representations. (Zhou, Wang, and Yang 2019) tackled the problem by constructing complete graphs for each view using information from other views. This graph construction process takes into account the missing instances and effectively captures the relationships between the data points in each view. (Li, Wan, and He 2021) employed consensus graph learning to uncover data structures and adaptively weights the stretched base partition to retain useful information while minimizing noise. (Wen et al. 2020a) addressed the problem of incomplete multi-view data clustering by developing a joint framework for graph completion and consensus representation learning.

Methodology

Notation Given a multi-view dataset $X = \{X^v\}_{v=1}^V$ with N samples in V views, where $X^v = \{x_1^v, x_2^v, \ldots, x_N^v\} \in R^{N \times D_v}$ represents the instance set of the v-th view, d_v is the dimensionality of samples and N denotes the number of samples, and elements of the missing instances are denoted as $'NaN'$. K is the number of categories to be clustered. We denote n samples with complete data as a set X_C.

View-specific Feature Learning with Local Graph

Deep autoencoders have been widely used to extract high-level representations of raw features (Peng et al. 2019; Xu et al. 2021, 2022). Considering that different views contains valuable information pertaining to specific features and structures, we design several view-specific encoders $f_{\theta^v}^v$ and corresponding decoders $g_{\phi^v}^v$, where θ^v and ϕ^v are learnable parameters. For sample x_i^v from the v-th view, the latent rep-

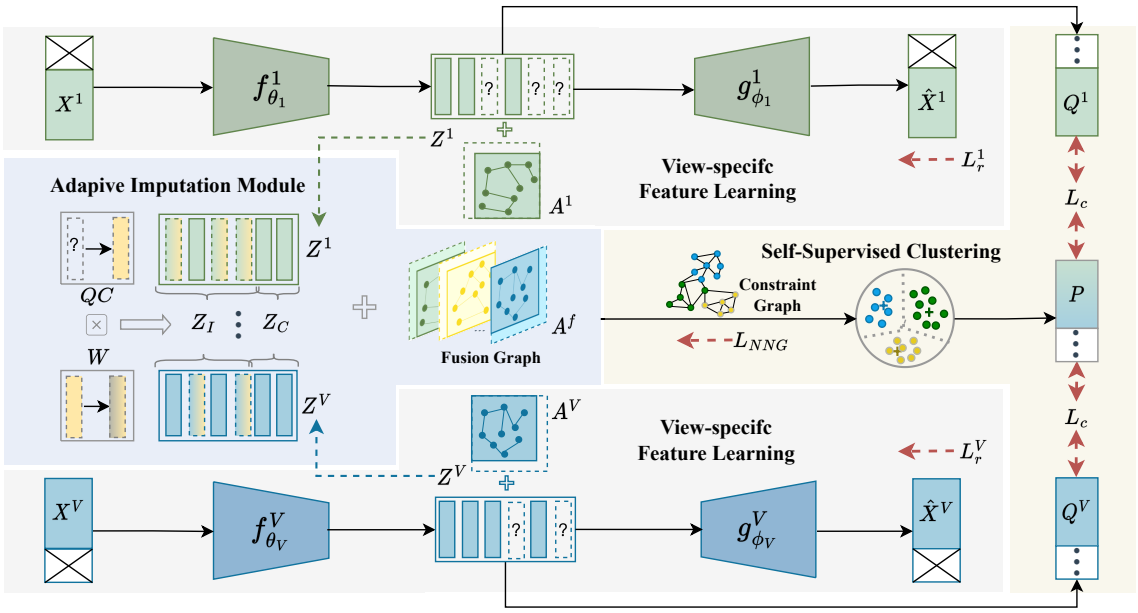


Figure 1: The framework of AGDIMC. For the v-th view, X^v denotes the input data, Z^v denotes the embedded features, A^v denotes the latent adjacency graph generated from Z^v, and Q^v denotes the cluster assignment distribution. A^f and P denote the global graph and unified target distribution respectively. L_r, L_c and L_{NNG} denote the reconstruction loss, clustering loss, and nearest neighbor constraint loss.

resentation can obtain by:

$$z_i^v = f_{\theta^v}^v(x_i^v). \tag{1}$$

After that, z_i^v is decoded as $\hat{x}_i^v$ through the decoder $g_{\phi^v}^v$:

$$\hat{x}_i^v = g_{\phi^v}^v(z_i^v), \tag{2}$$

where $\hat{x}_i^v$ represents the reconstructed data. Thus, the feature learning encoder in each view is trained by optimizing the reconstruction loss for each node in the latent graph, and it can be formulated as:

$$L_r^v = \sum_{v=1}^{V} L_r^v = \sum_{v=1}^{V} \sum_{i=1}^{N} ||x_i^v - g_{\phi^v}^v(f_{\theta^v}^v(x_i^v))||^2. \tag{3}$$

To achieve a comprehensive representation of graph structures, our approach involves the dynamic generation of graphs within the latent space of each view. The rationale behind this lies in the limitations associated with composing fixed graphs solely based on raw data. When using fixed graphs, it is not possible to make adjustments based on the clustering results, which limits the accuracy and adaptability of the clustering process. Conversely, by leveraging the deep autoencoders ability to extract highly representative features, we construct graphs in the latent space.

To preserve the local structure information, we construct the k-nearest neighbor graph from complete data as follows:

$$A_{i,j}^v = \begin{cases} 1, & \text{if } \left(z_i^v, z_j^v \neq NaN\right) \& \\ & \left(z_i^v \in \psi\left(z_j^v\right) \text{ or } z_j^v \in \psi\left(z_i^v\right)\right) \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

where $\psi(z_i^v)$ denotes the nearest sample set to z_i^v.

After obtaining latent graphs from each view, a two-layer GCN module is utilized to explore the information contained in latent graphs that provide available information for the clustering task. The embedded feature of each view can be refined by its unique latent graph which contains discriminative information.

$$\tilde{z}_i^v = \sum_{m=0}^{2} \left(\tilde{D}^{-\frac{1}{2}} \tilde{A}^v \tilde{D}^{-\frac{1}{2}}\right)^m z_i^v, \tag{5}$$

where $\tilde{A} = I_n + A, \tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. Compared with other existing GCN-based methods, our adopted simple yet effective GCN module has better generalization capability as it does not require additional parameters.

To obtain clustering predictions, similar to the majority of existing deep embedded clustering models, we utilize the Student's t-distribution. This distribution helps determine the likelihood of the i-th example being assigned to the j-cluster in the v-th view:

$$q_{ij}^v = c_{\mu^v}^v(\tilde{z}_i^v) = \frac{(1 + \left\|\tilde{z}_i^v - \mu_j^v\right\|^2)^{-1}}{\sum_j (1 + \left\|\tilde{z}_i^v - \mu_j^v\right\|^2)^{-1}} \in Q^v. \tag{6}$$

In the v-th view, μ_j^v denotes the learnable cluster centroids and q_{ij}^v is considered as the probability that the embedded feature.

Adaptive Imputation Module

To reduce the risk of clustering performance degradation caused by incomplete data, we further propose a novel adap-

tive imputation module based on cluster-oriented information to recover the missing samples.

The view available and missing information is recorded in an indicator matrix $\mathbf{I} \in \{0,1\}^{N\times V}$, where $\mathbf{I}_{iv} = 1$ if the i-th sample is available in the v-th view, otherwise $\mathbf{I}_{iv} = 0$. After extracting embedded features from the view-specific autoencoders, let $\tilde{Z}^v = \{\tilde{z}_1^v, \tilde{z}_2^v, \ldots, \tilde{z}_n^v\} \in R^{n\times d_v}$. Then refined embedded features from all views are concatenated to form a unified representation, denoted as $\tilde{Z}$:

$$\tilde{Z} = [\tilde{Z}^1, \tilde{Z}^2, \ldots, \tilde{Z}^V] \in R^{N\times\sum_{v=1}^{V} d_v}. \tag{7}$$

Furthermore, Eq. (6) allows us to easily obtain the soft pseudo assignment for each view. Then, we acquire the global cluster assignments Q to assist us in adaptive imputation by following the formula:

$$q_i = \frac{\sum_{v=1}^{V} \mathbf{I}_{iv} q_i^v}{\sum_{v=1}^{V} \mathbf{I}_{iv}} \in Q. \tag{8}$$

Besides, we denote $\tilde{Z} = [\tilde{Z}_C; \tilde{Z}_I]$, we leverage only the observed complete data to learn a consensus representation that guides the imputation of missing samples from each view, thus avoiding the impact of inconsistent with missing views. The view-specific patterns W can be optimized by:

$$\begin{aligned} &\min_{W} \|\tilde{Z}_C - WQC\|_F^2 \\ &= \min_{\{W^v\}_{v=1}^{V}} \sum_{v=1}^{V} \sum_{\tilde{z}_i^v \in \tilde{Z}_C} \|\tilde{z}_i^v - W^v q_i C^v\|_2^2. \end{aligned} \tag{9}$$

We can leverage information from global cluster assignments Q, global centroid C, and view-specific patterns W to impute the unavailable embedded features $\tilde{z}_i^v$. In particular, it focuses on imputing the missing embedded features by considering the sample commonality and cross-view correspondences. In this case, when $\mathbf{I}_{iv} = 0$, the embedded feature of missing sample z_i^v can be imputed as follows:

$$\tilde{z}_i^v = W^v q_i C^v \in \tilde{Z}_I. \tag{10}$$

With the process of imputation, the consistency of embedded features improves. By learning from these multiple views simultaneously, we can uncover more comprehensive and accurate representations from incomplete data. To leverage the discriminative information across all views, we concatenate the embedded features obtained by expanding the data in Eq. (10) to update global features: $Z = [\tilde{Z}_C; \tilde{Z}_I]$.

It's worth noting that higher missing rates can drift the view information leading to inaccurate imputation. To address this limitation, we assess the imputed global features. If the clustering result in q_{ij}^* by Eq. (6) from imputed global features is inferior to the result q_{ij} without imputation, we abandon the imputation process and use the unimputed global features for subsequent training.

After that, each view latent graph A_{imp}^v is learned from imputed features by Eq. (4) as $A_{imp} = \left\{A_{imp}^1, A_{imp}^2, \ldots, A_{imp}^V\right\}$. By considering the discriminative information of each view, we can effectively capture the unique characteristics of different perspectives. The global graph, being more informative and extracting more complete relationships between different nodes, makes it natural to fuse these multi-view graphs into a more robust global graph. The global graph A^f plays a crucial role in capturing the underlying sample similarity present within the multi-view data. To mitigate the impact of low-quality and noisy views, we adopt a strategy of assigning distinct weights to different graphs.

Based on the above motivations, we develop a graph fusion method which can be formulated as follows:

$$A^f = \min_{A^f} \sum_{v=1}^{V} w_v ||A^f - A_{imp}^v||_F^2. \tag{11}$$

We use the inverse distance of fusion graph A and each latent graph A^v to obtain w_v. In our model, we assign the value of the exponential parameter of the inverse distance weighting method to 2, and then w_v can be adaptively computed as:

$$w_v = \frac{1}{||A^f - A_{imp}^v||_F}. \tag{12}$$

By doing so, our method progressively obtains clearer clustering structure from the latent features as training process forwards. Meanwhile, by assigning smaller weights to the less reliable views, the negative affect of noise graphs is reduced effectively. To utilize the A^f and imputed global feature Z as the input of the GCN module we mentioned, the refined global representation H with more comprehensive information can be obtained.

Through this cluster-oriented method, we can extract consistent imputations and mitigate the risk of clustering performance degradation caused by inconsistent imputations. The adaptive imputation module is as effective as learning solely from complete data in terms of generalization risk.

Self-Supervised Clustering Layer

After obtaining the refined global features $H = \{h_1, h_2, \ldots, h_N\} \in R^{N\times\sum_{v=1}^{V} d_v}$, through global graph convolution module, we impose the nearest neighbor graph constraints on the common latent representation of all views. This serves to reinforce local neighbor relationships from complete data. The following graph constraint is considered:

$$L_{NNG}^v = \frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \|h_i - h_j\|_2^2 A_{i,j}^v. \tag{13}$$

We apply K-means(MacQueen 1967) to calculate the cluster centroids c_j:

$$\min_{c_1,c_2,\ldots,c_K} \sum_{i=1}^{N} \sum_{j=1}^{K} \|h_i - c_j\|^2. \tag{14}$$

Then, the soft pseudo assignment t_{ij} between each global embedding and each cluster centroid with Student's t-distribution is defined as

$$t_{ij} = \frac{(1 + \|h_i - c_j\|^2)^{-1}}{\sum_j (1 + \|h_i - c_j\|^2)^{-1}}. \tag{15}$$

To increase the discriminability of the soft pseudo assignments, the global target distribution P is computed by

$$p_{ij} = \frac{(t_{ij}^2 / \sum_i t_{ij})}{\sum_j (t_{ij}^2 / \sum_i t_{ij})}. \tag{16}$$

The clustering loss of each view is the KL divergence between the unified target distribution P and its own cluster assignment distribution Q^v

$$L_c^v = D_{KL}\left(P \| Q^v\right) = \sum_{i=1}^{N} \sum_{j=1}^{K} p_{ij} \log \frac{p_{ij}}{q_{ij}^v}. \tag{17}$$

To learn an accurate assignment for clustering, it is necessary to introduce an integrated objective to guide the learning process. To this end, we jointly optimize the deep autoencoder embedding and clustering learning, and the total objective function is defined as:

$$L^v = L_r^v + \alpha L_c^v + L_{NNG}^v, \tag{18}$$

where $0 < \alpha < 1$ is a trade-off coefficient that controls the degree of distorting embedded space. We set $\alpha = 0.5$ for all experiments. Minimizing KL divergence between Q and P will make the distribution of Q sharper and mine the information in different views. After obtaining soft cluster assignments from multiple views, the highly confident predictions will guide the training process, which can also avoid the interference of a few wrong predictions. When the whole training process is completed, we compute the pseudo-label P once again and the final clustering assignment y_i for the i-th sample is:

$$y_i = \arg\max_j (p_{ij}). \tag{19}$$

Optimization

Algorithm 1 summarizes the optimization procedure of AGDIMC. It is composed of two procedures, $i.e.$, initialization, and fine-tuning. In the initialization stage, for each view, we firstly pretrain autoencoders $f_{\theta^v}^v$ and $g_{\theta^v}^v$ by optimizing the reconstruction loss in Eq. (3). During the finetuning stage, the embedded features of missing samples can be imputed by Eq. (8) and Eq. (10). Based on the unified representation H, the soft pseudo assignment can be obtained by Eq. (6). Intuitively, the quality of imputed features can be assessed by measuring the difference between two soft pseudo assignments derived from original embedded features and imputed features, respectively. For every T iteration, if imputed features with high consistency, the whole network of AGDIMC is trained by optimizing the objective function Eq. (18). Otherwise, we focus on training on complete data to get P that helps to generate informative embeded features. When the whole training process is completed, we compute the pseudo-label P once again to get the result by Eq.(19). The workflow of AGDIMC is summarized in Algorithm 1.

Algorithm 1: Optimization of the proposed AGDIMC

Input: Multi-view dataset X, cluster number K.
Output: Cluster assignments Y

```
Pretrain the autoencoder separately in each view by optimizing Eq. (3).
Initialize the pseudo-labels by Eq. (16)
Initialize the centriods c^v_{μ^v} by the decomposing the global cluster centroids.
while not reach the maximum iterations T_max do
    repeat
        Optimize Eq. (3) and Eq. (17), update θ^v, φ^v and μ^v_j of each view.
        Optimize Eq. (9), update W.
    until The iteration time is divisible by T
    Compute the q_ij by Eq. (8)
    Impute the missing sample by Eq. (10)
    Compute the q*_ij by Eq. (6) from imputed features.
    if arg max_j(q*_ij) > max_j(q_ij) then
        Update P by imputed features.
        Optimize Eq. (13), update the H.
    end if
end while
Compute the final pseudo-labels P by Eq. (16)
Compute for each sample by Eq. (19)
```

Experiments

Experimental Setup

Datasets Three widely used and publicly available multi-view datasets are used in our study:

BDGP (Cai et al. 2012) contains 2500 samples from 5 categories and each class has 500 samples. For each sample, the texture feature and three kinds of visual features extracted from the lateral, dorsal, and ventral images via the bag-of-words extractor are regarded as four views.

Handwritten Numerals (HW) represented by six kinds of features extracted from its binary image. Each class has 200 samples. Each instance has six visual views, including profile correlations, Fourier coefficients of the character shapes, Karhunen-Love, morphological features, pixel averages in 2 × 3 windows, and Zernike moments.

Reuters is comprised of 1200 articles in 6 categories (C15, CCAT, E21, ECAT, GCAT and M11), each providing 200 articles. For each article, it is written in five different languages (English, French, German, Italian, and Spanish).

Comparing Methods Comparison methods include 8 state-of-art methods, i.e., GIMC-FLSD (generalized incomplete multi-view clustering with flexible locality structure diffusion (Wen et al. 2020b)), CDIMC-net (cognitive deep incomplete multi-view clustering network (Wen et al. 2021)), DCP (completer: incomplete multi-view clustering via contrastive prediction (Lin et al. 2021)), HCP-IMSC (high-order correlation preserved incomplete multi-view subspace clustering (Li et al. 2022)), IMVC-CBG (highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph (Wang et al. 2022)), DSIMVC (deep safe incomplete multi-view clustering: Theorem and algorithm (Tang and Liu 2022)), LSIMVC (localized sparse incomplete multi-view clustering (Liu et al. 2022)), and PGP (self-guided partial graph propagation for incomplete multi-view clustering (Liu et al. 2023)).

	Missing rate	0.1			0.3			0.5			0.7		
	Evaluation metrics	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
BDGP	GIMC-FLSD(Wen et al. 2020b)	0.833	0.628	0.634	0.783	0.594	0.555	0.772	0.532	0.534	0.725	0.497	0.456
	CDIMC-net(Wen et al. 2021)	0.882	0.789	0.819	0.744	0.537	0.503	0.727	0.594	0.517	0.524	0.311	0.224
	DCP(Lin et al. 2021)	0.465	0.447	0.186	0.443	0.325	0.066	0.424	0.305	0.054	0.356	0.277	0.053
	HCP-IMSC(Li et al. 2022)	0.968	0.902	0.922	0.938	0.823	0.852	0.901	0.769	0.759	0.896	0.726	0.759
	IMVC-CBG(Wang et al. 2022)	0.392	0.242	0.154	0.374	0.221	0.106	0.363	0.176	0.056	0.342	0.018	0.068
	DSIMVC(Tang and Liu 2022)	0.963	0.905	0.912	0.956	0.886	0.910	0.941	0.829	0.859	0.917	0.791	0.807
	LSIMVC(Liu et al. 2022)	0.732	0.713	0.805	0.613	0.555	0.571	0.490	0.388	0.493	0.380	0.236	0.575
	PGP(Liu et al. 2023)	0.526	0.327	0.536	0.472	0.316	0.542	0.496	0.310	0.496	0.422	0.268	0.496
	AGDIMC(ours)	**0.979**	**0.937**	**0.949**	**0.962**	**0.908**	**0.912**	**0.961**	**0.883**	**0.906**	**0.922**	**0.829**	**0.816**
Reuters	GIMC-FLSD(Wen et al. 2020b)	0.478	0.292	0.208	0.469	0.283	0.202	0.473	0.275	0.202	0.494	**0.283**	0.216
	CDIMC-net(Wen et al. 2021)	0.310	0.130	0.075	0.303	0.080	0.051	0.263	0.053	0.027	0.253	0.050	0.027
	DCP(Lin et al. 2021)	0.323	0.141	0.06	0.315	0.158	0.042	0.232	0.137	0.013	0.224	0.095	0.014
	HCP-IMSC(Li et al. 2022)	0.342	0.175	0.085	0.357	0.209	0.111	0.407	0.219	0.136	0.386	0.218	0.133
	IMVC-CBG(Wang et al. 2022)	0.442	0.273	0.139	0.402	0.228	0.103	0.364	0.213	0.088	0.348	0.188	0.057
	DSIMVC(Tang and Liu 2022)	0.455	0.302	0.224	0.425	0.274	0.198	0.421	0.256	0.187	0.418	0.237	0.174
	LSIMVC(Liu et al. 2022)	0.382	0.209	0.131	0.398	0.221	0.145	0.303	0.152	0.062	0.243	0.107	0.029
	PGP(Liu et al. 2023)	0.575	0.348	**0.338**	0.516	0.306	**0.359**	0.245	0.088	**0.285**	0.261	0.108	0.215
	AGDIMC(ours)	**0.588**	**0.354**	0.294	**0.548**	**0.261**	0.245	**0.548**	**0.293**	0.236	**0.509**	0.256	**0.227**
HW	GIMC-FLSD(Wen et al. 2020b)	0.613	0.577	0.456	0.427	0.440	0.307	0.408	0.431	0.229	0.262	0.174	0.046
	DCP(Lin et al. 2021)	0.797	0.753	0.570	0.742	0.743	0.640	0.738	0.734	0.626	0.752	0.725	0.596
	CDIMC-net(Wen et al. 2021)	0.933	0.878	0.861	0.892	0.831	0.808	0.854	0.886	0.807	0.789	0.804	0.724
	HCP-IMSC(Li et al. 2022)	0.817	0.788	0.724	0.797	0.767	0.709	0.775	0.710	0.651	0.636	0.545	0.392
	IMVC-CBG(Wang et al. 2022)	0.572	0.598	0.444	0.512	0.509	0.320	0.471	0.473	0.237	0.426	0.406	0.144
	DSIMVC(Tang and Liu 2022)	0.810	0.798	0.725	0.778	0.772	0.686	0.762	0.736	0.650	0.747	0.732	0.648
	LSIMVC(Liu et al. 2022)	0.943	0.889	0.920	0.936	0.880	0.912	0.874	0.828	0.865	0.858	0.765	0.838
	PGP(Liu et al. 2023)	0.836	0.835	0.846	0.850	0.854	0.839	0.854	0.863	0.842	0.823	0.814	**0.845**
	AGDIMC(ours)	**0.973**	**0.943**	**0.941**	**0.956**	**0.912**	**0.906**	**0.946**	**0.882**	**0.881**	**0.926**	**0.855**	0.842

Table 1: Clustering results of all methods on three datasets. The best and second-best results are highlighted with bold and underline, respectively.

	Components			BDGP			Reuters			HW		
	A	B	C	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
Item-1		✓	✓	0.932	0.824	0.807	0.527	0.304	0.211	0.907	0.883	0.869
Item-2	✓		✓	0.789	0.643	0.598	0.502	0.314	0.218	0.880	0.841	0.827
Item-3	✓	✓		0.913	0.783	0.797	0.462	0.268	0.204	0.938	0.902	0.895
Item-4	✓	✓	✓	0.979	0.937	0.949	0.588	0.354	0.294	0.973	0.943	0.941

Table 2: Ablation studies on three datasets.

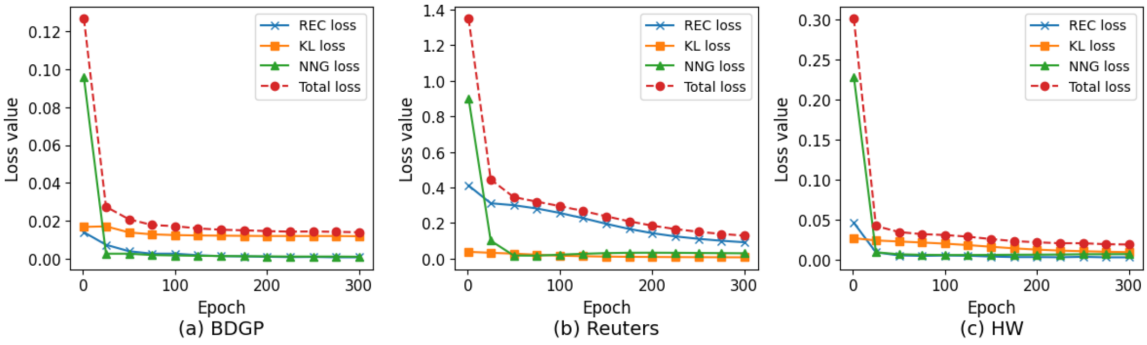
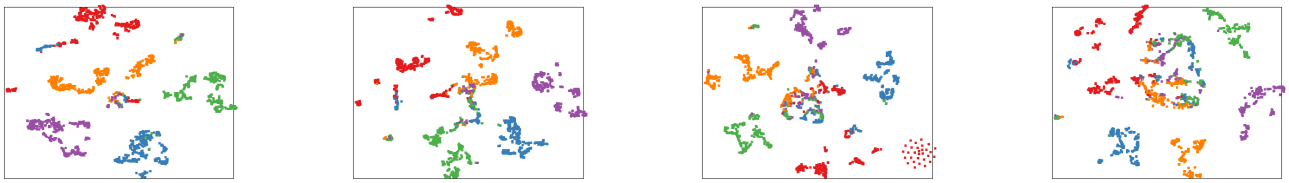


Figure 2: Convergence Analysis on BDGP, REU, and HW.

Implementation Details Following (Guo et al. 2017), we use the same fully connected (Fc) autoencoder structure on all three datasets. Specifically, for each view, the structure of the encoder is Input - Fc500 - Fc500 - Fc2000 - Fc10. Decoders are symmetric with the encoders of corresponding views. All the autoencoders are pre-trained for 2000 epochs. The trade-off coefficient α is set to 0.5 and the number of neighbors k applied in kNN graph algorithm is set to 10. The dimensionality of embeddings Z_v is reduced to 10. The activation function is ReLU (Glorot, Bordes, and Bengio 2011).



(a) Missing rate = 0.1 (b) Missing rate = 0.3 (c) Missing rate = 0.5 (d) Missing rate = 0.7

Figure 3: Visualization of the clustering results on BDGP with different missing rates.

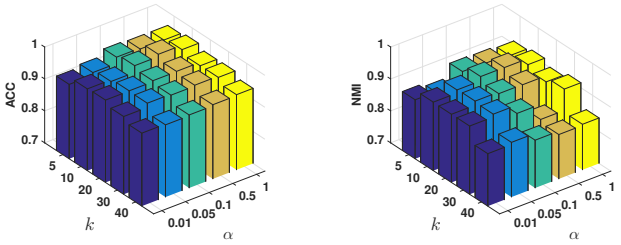


Figure 4: Clustering performance w.r.t. different parameter settings on BDGP dataset with missing rate = 0.1.

The batch size is set to the instance number like most GNN-based methods. We adopt Adam (Kingma and Ba 2014) to optimize the deep models with a learning rate of 0.001. For the comparing methods, we directly use the source codes provided by authors and the suggested parameter settings.

Experimental Results and Analysis

Comparison with Baselines In our study, we compared our model AGDIMC with several state-of-the-art algorithms that have shown superior performance in recent years. We evaluated the performance of AGDIMC on three benchmark datasets and four different missing rates: [0.1, 0.3, 0.5, 0.7]. The clustering performance of all methods on three datasets is presented in Table 1. Our results demonstrate that AGDIMC outperforms the baseline algorithms across these datasets. We also found that the clustering performance of all methods declines as the missing rate varies from 0.1 to 0.7. Nevertheless, our AGDIMC still achieves superior clustering performance in most cases.

The success of AGDIMC can be attributed to the incorporation of clustering techniques. Furthermore, we address the negative impact caused by inferior impute samples and the discrepancy of cluster structures. These measures help improve the overall performance.

Ablation Study To further verify the contributions of the proposed method, we conduct an ablation study that shows in Table 2. Components A, B, and C stand for local graph convolution and global graph fusion, adaptive imputation module, and nearest neighbor graph constraints, respectively. Item-1 does not apply the GCN module to refine embedded features that discard topological information, Item-2 is affected by low-quality imputed features without an adaptive imputation module. Besides, Item-3 is limited because the neighbor relationship is lost during the fusion process. The best performance is achieved by Item-4, which illustrates the significance of different parts in our framework.

Convergence and Training Process The convergence analysis was performed on the REC loss, KL loss, NNG loss, and total loss using the BDGP, REU, and HW datasets. Fig. 2 clearly demonstrates that as the number of epochs increases, all four loss functions exhibit a gradual convergence and reach a stable state. This provides strong evidence that our model is both stable and effective.

Visualization of Learning Process We conducted cluster result visualization analysis on BDGP. From Fig. 3, it can be observed that our model is capable of effectively separating embeddings from different categories. This visual demonstration highlights the effectiveness of our approach. Additionally, it can be noticed that as the missing rate increases, the clustering structures of the embedded features become less apparent.

Parameter Sensitivity In this analysis, we investigated the sensitivity of our clustering model's performance to two main hyperparameters, α and k. The BDGP dataset was used, with a missing rate of 0.1. We experimented with varying values of α and k and evaluated the clustering performance using appropriate metrics. From Fig. 4, we observed that within a certain range of α values, the clustering performance remained relatively stable. This suggests that our model is robust to changes in this hyperparameter and can produce consistent results.

Conclusion

In this paper, we propose an Adaptive Feature Imputation with Latent Graph for Deep Incomplete Multi-View Clustering (AGDIMC). We not merely consider the topological information that the latent graph can provide for the clustering task, but also utilize the fusion graph to capture the complementarity. Specifically, a novel adaptive imputation module is developed to recover the missing samples dynamically and improve the robustness of multi-view clustering on incomplete data. Besides this, it incorporates the graph embedding technique to preserve the local structure of data. Experimental results on multiple real-world multi-view datasets demonstrate the state-of-the-art performance of the proposed method.

Acknowledgments

This work is supported in part by National Key Research and Development Program of China (Nos. 2020YFC2004300 and 2020YFC2004302), National Natural Science Foundation of China (No. 61971052), and Shenzhen Science and Technology Program (Nos. JCYJ20230807120010021 and JCYJ20230807115959041). Philip S. Yu is supported in part by NSF under grant III-2106758. Lifang He is partially supported by the NSF grants (MRI-2215789, IIS-1909879, IIS-2319451), NIH grant under R21EY034179, and Lehigh's grants under Accelerator and CORE.

References

Cai, X.; Wang, H.; Huang, H.; and Ding, C. 2012. Joint stage recognition and anatomical annotation of drosophila gene expression patterns. *Bioinformatics*, 28(12): i16–i24.

Chen, M.-S.; Huang, L.; Wang, C.-D.; and Huang, D. 2020. Multi-view clustering in latent embedding space. In *AAAI*, volume 34, 3513–3520.

Glorot, X.; Bordes, A.; and Bengio, Y. 2011. Deep sparse rectifier neural networks. In *AISTATS*, 315–323. JMLR Workshop and Conference Proceedings.

Guo, J.; and Ye, J. 2019. Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering. In *AAAI*, volume 33, 118–125.

Hu, M.; and Chen, S. 2019a. Doubly aligned incomplete multi-view clustering. *arXiv*.

Hu, M.; and Chen, S. 2019b. One-pass incomplete multi-view clustering. In *AAAI*, volume 33, 3838–3845.

Huang, S.; Kang, Z.; and Xu, Z. 2020. Auto-weighted multi-view clustering via deep matrix decomposition. *Pattern Recognition*, 97: 107015.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv*.

Li, L.; Wan, Z.; and He, H. 2021. Incomplete multi-view clustering with joint partition and graph learning. *TKDE*, 35(1): 589–602.

Li, S.-Y.; Jiang, Y.; and Zhou, Z.-H. 2014. Partial multi-view clustering. In *AAAI*, volume 28.

Li, Z.; Tang, C.; Zheng, X.; Liu, X.; Zhang, W.; and Zhu, E. 2022. High-Order Correlation Preserved Incomplete Multi-View Subspace Clustering. *TIP*, 31: 2067–2080.

Lin, Y.; Gou, Y.; Liu, X.; Bai, J.; Lv, J.; and Peng, X. 2022. Dual contrastive prediction for incomplete multi-view representation learning. *TPAMI*, 45(4): 4447–4461.

Lin, Y.; Gou, Y.; Liu, Z.; Li, B.; Lv, J.; and Peng, X. 2021. Completer: Incomplete multi-view clustering via contrastive prediction. In *CVPR*, 11174–11183.

Liu, C.; Li, R.; Wu, S.; Che, H.; Jiang, D.; Yu, Z.; and Wong, H.-S. 2023. Self-Guided Partial Graph Propagation for Incomplete Multi-View Clustering. *TNNLS*.

Liu, C.; Wu, Z.; Wen, J.; Xu, Y.; and Huang, C. 2022. Localized Sparse Incomplete Multi-view Clustering. *TMM*.

Liu, X.; Li, M.; Tang, C.; Xia, J.; Xiong, J.; Liu, L.; Kloft, M.; and Zhu, E. 2020. Efficient and effective regularized incomplete multi-view clustering. *TPAMI*, 43(8): 2634–2646.

MacQueen, J. 1967. Classification and analysis of multivariate observations. In *BSMSP*, 281–297.

Nie, F.; Cai, G.; and Li, X. 2017. Multi-view clustering and semi-supervised classification with adaptive neighbours. In *AAAI*, volume 31.

Peng, X.; Huang, Z.; Lv, J.; Zhu, H.; and Zhou, J. T. 2019. COMIC: Multi-view clustering without parameter selection. In *ICML*, 5092–5101. PMLR.

Tang, H.; and Liu, Y. 2022. Deep safe incomplete multi-view clustering: Theorem and algorithm. In *ICML*, 21090–21110. PMLR.

Trivedi, A.; Rai, P.; Daumé III, H.; and DuVall, S. L. 2010. Multiview clustering with incomplete views. In *NeurIPS workshop*, volume 224, 1–8.

Wang, Q.; Lian, H.; Sun, G.; Gao, Q.; and Jiao, L. 2020. ICMSC: Incomplete cross-modal subspace clustering. *TIP*, 30: 305–317.

Wang, S.; Liu, X.; Liu, L.; Tu, W.; Zhu, X.; Liu, J.; Zhou, S.; and Zhu, E. 2022. Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph. In *CVPR*, 9776–9785.

Wen, J.; Yan, K.; Zhang, Z.; Xu, Y.; Wang, J.; Fei, L.; and Zhang, B. 2020a. Adaptive graph completion based incomplete multi-view clustering. *TMM*, 23: 2493–2504.

Wen, J.; Zhang, Z.; Fei, L.; Zhang, B.; Xu, Y.; Zhang, Z.; and Li, J. 2022. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2): 1136–1149.

Wen, J.; Zhang, Z.; Xu, Y.; Zhang, B.; Fei, L.; and Liu, H. 2019. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *AAAI*, volume 33, 5393–5400.

Wen, J.; Zhang, Z.; Xu, Y.; Zhang, B.; Fei, L.; and Xie, G.-S. 2021. Cdimc-net: Cognitive deep incomplete multi-view clustering network. In *IJCAI*, 3230–3236.

Wen, J.; Zhang, Z.; Xu, Y.; and Zhong, Z. 2018. Incomplete multi-view clustering via graph regularized matrix factorization. In *ECCV workshops*.

Wen, J.; Zhang, Z.; Zhang, Z.; Fei, L.; and Wang, M. 2020b. Generalized Incomplete Multi-view Clustering With Flexible Locality Structure Diffusion. *IEEE Transactions on Cybenetics*, doi: 10.1109/TCYB.2020.2987164.

Wu, J.; Zhuge, W.; Tao, H.; Hou, C.; and Zhang, Z. 2018. Incomplete multi-view clustering via structured graph learning. In *PRICAI*, 98–112. Springer.

Xu, C.; Guan, Z.; Zhao, W.; Wu, H.; Niu, Y.; and Ling, B. 2019. Adversarial incomplete multi-view clustering. In *IJCAI*, volume 7, 3933–3939.

Xu, J.; Li, C.; Peng, L.; Ren, Y.; Shi, X.; Shen, H. T.; and Zhu, X. 2023. Adaptive Feature Projection With Distribution Alignment for Deep Incomplete Multi-View Clustering. *TIP*, 32: 1354–1366.

Xu, J.; Ren, Y.; Li, G.; Pan, L.; Zhu, C.; and Xu, Z. 2021. Deep embedded multi-view clustering with collaborative training. *Information Sciences*, 573: 279–290.

Xu, J.; Ren, Y.; Tang, H.; Yang, Z.; Pan, L.; Yang, Y.; Pu, X.; Philip, S. Y.; and He, L. 2022. Self-supervised discriminative feature learning for deep multi-view clustering. *TKDE*, 1–12.

Xu, N.; Guo, Y.; Zheng, X.; Wang, Q.; and Luo, X. 2018. Partial multi-view subspace clustering. In *ACMM*, 1794–1801.

Yang, M.; Li, Y.; Hu, P.; Bai, J.; Lv, J.; and Peng, X. 2022. Robust multi-view clustering with incomplete information. *TPAMI*, 45(1): 1055–1069.

Zhan, K.; Niu, C.; Chen, C.; Nie, F.; Zhang, C.; and Yang, Y. 2018. Graph structure fusion for multiview clustering. *TKDE*, 31(10): 1984–1993.

Zhao, H.; Liu, H.; and Fu, Y. 2016. Incomplete multi-modal visual data grouping. In *IJCAI*, 2392–2398.

Zhao, L.; Chen, Z.; Yang, Y.; Wang, Z. J.; and Leung, V. C. 2018. Incomplete multi-view clustering via deep semantic mapping. *Neurocomputing*, 275: 1053–1062.

Zhou, W.; Wang, H.; and Yang, Y. 2019. Consensus graph learning for incomplete multi-view clustering. In *PAKDD*, 529–540. Springer.