

# Incomplete and Unpaired Multi-View Graph Clustering with Cross-View Feature Fusion

Liang Zhao, Ziyue Wang, Xiao Wang, Zhikui Chen, Bo Xu

Dalian University of Technology, Dalian, China

{liangzhao, zkchen, BoXu}@dlut.edu.cn, wangziyue6306@mail.dlut.edu.cn, AlanWang\_1024@163.com

## Abstract

Due to its effectiveness and efficiency, graph-based multi-view clustering has recently attracted much attention. However, the multi-view data are often incomplete and unpaired in real-world applications as a consequence of data loss or corruption. Although efforts have been made through a series of methods to address the problems of incomplete or unpaired multi-view data, the following issues still persist: 1) Most existing methods only focus on the incomplete multi-view data or unpaired multi-view data, and exhibit weaknesses when addressing both incomplete and unpaired multi-view data simultaneously. 2) Some methods neglect the graph information of the data from different views during the learning process. To tackle these issues, we propose the Multi-view Graph Clustering framework with Cross-view Feature Fusion (MGCCFF), a novel approach for clustering incomplete and unpaired multi-view data. Specifically, MGCCFF learns soft clustering label information from complete data and utilizes this to capture category-level cross-view correspondences. It then learns latent representation enriched with cross-view information based on the established mappings. To obtain a multi-view graph structure under conditions of incomplete and unpaired data, MGCCFF innovatively integrates the concept of self-expression with the autoencoder architecture and exploits the latent relationships between labels and the graph structure, thereby enabling the generation of sparse and accurate graphical structure under multi-view conditions for the final clustering task. The experiments on incomplete and unpaired multi-view datasets demonstrate that MGCCFF outperforms state-of-the-art methods.

## Introduction

In recent real-world applications, the information of objectives is increasingly collected from different domains or multiple models. And multi-view learning has recently attracted much attention. Due to the effectiveness of graph information by signifying the relationships between different samples, multi-view graph learning has become a hot research topic (Gao et al. 2015; Huang et al. 2023; Zhu, Lu, and Zhou 2019; Nie et al. 2011; Li et al. 2022a; Nie, Li, and Li 2017; Zhan et al. 2018). Generally speaking, most multi-view graph learning methods aim to explore a consistent graph structure from multi-view data based on self-

representation learning (Wang, Yang, and Liu 2020; Kang et al. 2020; Zhang et al. 2020b, 2017). For example, CSMSC (Luo et al. 2018) derives the optimal consensus graph matrix from view-specific graph structure of each view through linear combination.

However, due to the sensor difference or corruption, multi-view data is often incomplete and unpaired in real applications. Most existing multi-view methods are impractical for problematic multi-view data because they assume that all samples are complete and paired. The visualization of complete, incomplete, unpaired, both incomplete and unpaired multi-view data is shown as Fig. 1, respectively.

Multi-view data with missing views are called incomplete multi-view data, which means the feature representation of samples on partial views is missing (Xu, Tao, and Xu 2015; Liu et al. 2019; Tang et al. 2024). Some approaches have been proposed to solve incomplete multi-view data. iCmSC (Wang et al. 2021) maximizes the intrinsic correlations among different views by deep Canonical Correlation Analysis (CCA) and explores the consistent graph structure for spectral clustering.

Multi-view data with incorrect alignment information between views is termed unpaired multi-view data, which implies, for a given unpaired multi-view dataset, the  $x_i^m$  from view  $m$  and  $x_i^n$  from view  $n$  may represent different objects (Zong, Zhang, and Liu 2018). In order to deal with this issue, some unpaired multi-view methods are proposed. PVC(unpaired) (Huang et al. 2020) delves the mapping relationship of the unpaired data in the latent subspace based on deep neural networks (DNNs) by improving the Hungarian algorithm. MvCLN (Yang et al. 2021) designs a noise-robust contrastive loss to explore the alignment information cross different views. MVC-UM (Yu et al. 2021) unifies Non-negative Matrix Factorization (NMF) and local structure constraints into a novel unpaired multi-view clustering framework.

Although remarkable success has been made, the above methods suffer from dealing with both incomplete and unpaired multi-view data. When directly facing and tackling incomplete and unpaired multi-view data, incomplete multi-view methods will misguidedly fuse multi-view features of different samples by obtaining incorrect cross-view mapping relationships, and unpaired multi-view methods will be unsuitable due to missing feature representations.

To tackle it, we propose a multi-view clustering framework for incomplete and unpaired multi-view data termed Multi-view Graph Clustering framework with Cross-view Feature Fusion (MGCCFF). Specifically, MGCCFF acquires the pseudo labels from the view-specific representations encoded on each view, securing the category-level cross-view correspondences and fusing feature representations across views, which is exploited to capture a simple and accurate graph structure for spectral clustering. Fig. 2 presents the overall framework of our proposed method. The main contributions are summarized as follows:

- We propose a novel multi-view clustering method for incomplete and unpaired multi-view data termed Multi-view Graph Clustering framework with Cross-view Feature Fusion (MGCCFF), which fuses the view-specific representations across different views and learns the graph information for spectral clustering.
- The pseudo labels are obtained by a discriminative clustering loss on each view, which is explored to acquire the category-level cross-view correspondences. Moreover, by integrating self-expression with autoencoders and exploiting the latent relationships among labels and spectral structures, MGCCFF achieves a sparse and accurate spectral structure under multi-view conditions for the final clustering task.
- Extensive experiments under various conditions verify the advantages of our proposed model MGCCFF. Compared with existing state-of-the-art methods, our model shows clearly superior performance on incomplete and unpaired multi-view data.

## Related Works

In this section, we briefly review the related incomplete multi-view methods and unpaired multi-view methods.

### Incomplete Multi-view Methods

In order to overcome the incomplete multi-view data learning scenario, some outstanding works in recent years are proposed, which can be generally grouped into two categories from the motivation (Wen et al. 2023a): 1) Ignoring the missing part of incomplete multi-view data and then performing feature fusion by focusing on the information of complete part among the available views. 2) Dedicating to recovering the missing part among incomplete multi-view data and integrating the fully multi-view information for a consensus representation or clustering.

iCmSC (Wang et al. 2021) implants the CCA into DNNs and explores the correlation information between different views through fully connected parts of incomplete multi-view data. IMSC\_AGL (Wen, Xu, and Liu 2020) adaptively constructs the graph of each view and designs a spectral constraint and a co-regularization term for low-dimensional common representation.

PMVC\_CGAN (Wang et al. 2018) exploits AE and GAN to recover the missing views and fuse all information into the learned latent representation for clustering. In HCP-IMSC (Li et al. 2022b), a hypergraph is constructed from

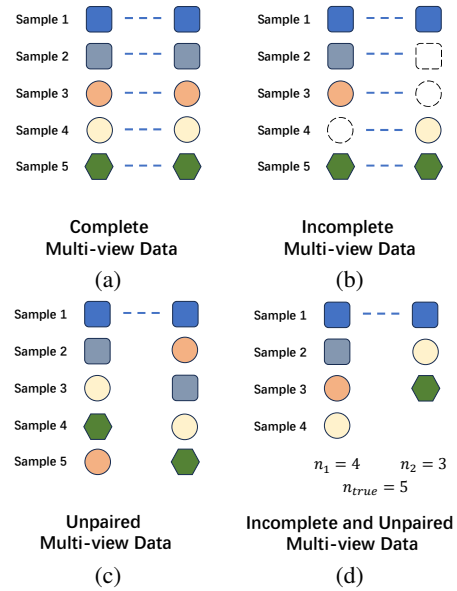


Figure 1: The visualizations of (a) complete, (b) incomplete, (c) unpaired and (d) incomplete and unpaired multi-view data. In the figure, different colors and different shapes denote different instances and different categories, respectively. The dotted line denotes the desired correspondences. And the dotted instances without color denote the missing instances.

the learned view-specific graph structure for reconstructing the samples with missing views by their neighbor samples.

### Unpaired Multi-view Methods

To tackle unpaired multi-view data, some researchers explore the correlations between cross-view data, including sample feature correspondences and graph structure correspondences across views. CMVNMF (Zhang et al. 2015) combines NMF and sample-level correspondences to learn the representations for clustering. MvCLN (Yang et al. 2021) establishes cross-view correspondences by designing noise-robust contrastive loss based on contrastive learning. MVC-UM (Yu et al. 2021) attempts to learn view-specific representations by integrating the graph structure correspondences and feature correspondences across views. UPMGC-SM (Wen et al. 2023b) utilizes the structural information from each view to refine cross-view graph structure correspondences.

### Incomplete and Unpaired Multi-view Methods

Although the aforementioned studies have yielded some encouraging results, almost all of them are only able to address either the issue of incompleteness or being unpaired. Addressing both incomplete and unpaired data simultaneously poses a significant challenge. SURE (Yang et al. 2023) integrates the issues of incompleteness and unpaired data into a unified category-level recognition framework. For each sample, SURE identifies its cross-view counterparts within the same category by employing a novel noise-robust contrastive learning paradigm, thereby establishing cross-view

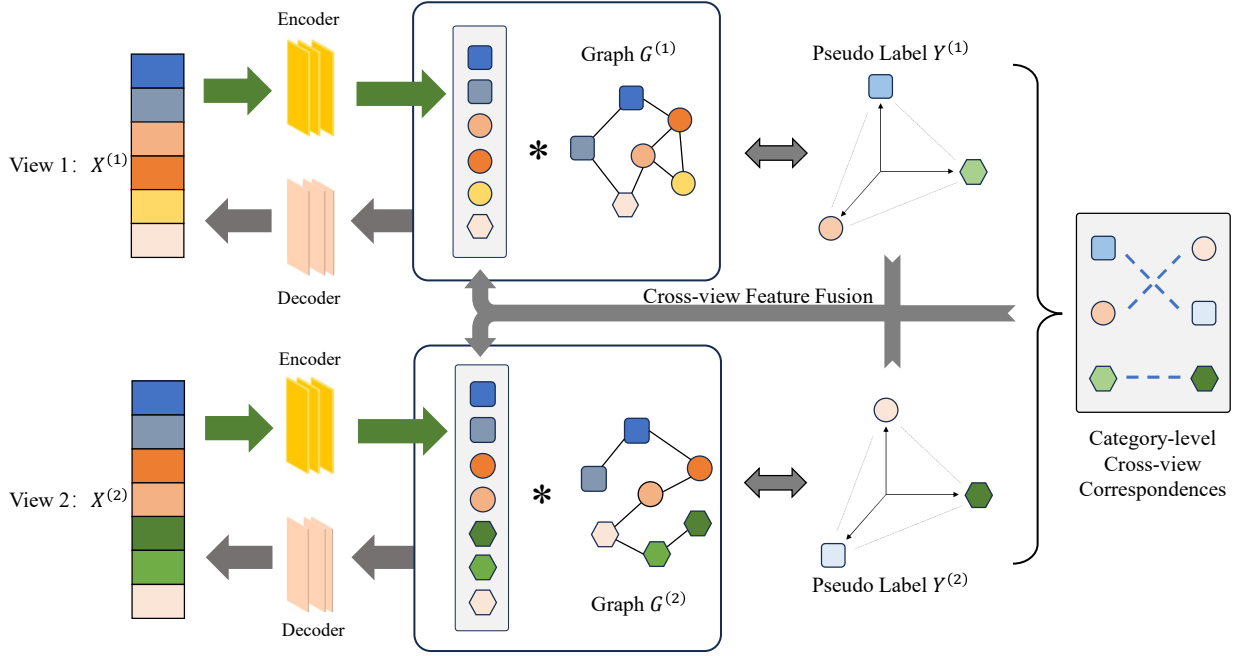


Figure 2: The overall framework of the proposed MGCCFF method.

correspondences at the category level and concurrently addressing both challenges. However, SURE does not effectively explore the spectral structure information inherent in the data itself.

### The Proposed Method

In this section, we introduce the proposed multi-view clustering framework for incomplete and unpaired multi-view data termed Multi-view Graph Clustering framework with Cross-view Feature Fusion (MGCCFF) as presented in Fig. 2.

#### Pseudo Label Learning

Considering the number of samples in different views is different for incomplete and unpaired multi-view data, the sample-level correspondences are unsuitable. Therefore, in our work, the category-level cross-view correspondences are adopted to align the unpaired multi-view instances, and in which obtaining pseudo label  $Y^v$  is essential.

To achieve this, feature extraction is first performed to obtain the view-specific representation  $S^v$  from the original data  $X^v$ , based on the encoder network  $f_{en}^v$  applied on each view according to

$$S^v = f_{en}^v(X^v) \quad (1)$$

Focusing on each view, for enforcing the accuracy of pseudo label  $Y$ , we adopt the Deep Divergence-based Clustering (DDC) loss (Kampffmeyer et al. 2019) for unsupervised clustering, which has shown state-of-the-art performance in single-view clustering. The DDC loss is composed of 3 terms for ensuring the inter-view separability and intra-view compactness, the closeness of pseudo labels to simplex corners, and the orthogonal requirements.

Configured with the number of clusters  $k$  and a matrix  $K$  storing Gaussian kernel elements of  $S$ , the first term of DDC loss encourages the inter-view separability and intra-view compactness:

$$L_1 = \frac{1}{k} \sum_{i=1}^{k-1} \sum_{j>i} \frac{Y_i^T K Y_j}{\sqrt{Y_i^T K Y_i Y_j^T K Y_j}} \quad (2)$$

in which  $Y_i$  means the pseudo label of  $i$ -th samples.

The second term aspires the effectiveness of the pseudo labels, i.e., the closeness of pseudo labels to standard simplex in  $\mathbb{R}^k$ :

$$L_2 = \frac{1}{k} \sum_{i=1}^{k-1} \sum_{j>i} \frac{M_i^T K M_j}{\sqrt{M_i^T K M_i M_j^T K M_j}} \quad (3)$$

where the  $M_i, M_j$  means the  $i$ -th,  $j$ -th columns of the matrix  $M = [M_{ia}]$  with  $M_{ia} = \exp(-\|Y_i - E_a\|)$ , and  $E_a$  represents the  $a$ -th corner of the simplex in  $\mathbb{R}^k$ .

The third term requires the orthogonal constraint, which can be formulated as:

$$L_3 = \text{triu}(Y^T Y) \quad (4)$$

where  $\text{triu}(\cdot)$  denotes the sum of the strictly upper triangular elements.

Then, the objective function of learning pseudo labels can be written as

$$L_{pl} = L_1 + L_2 + L_3 \quad (5)$$

#### Cross-view Feature Fusion

To establish the cross-view correspondences, we focus on the fully complete part of incomplete and unpaired multi-view data. Let the  $C_i^m$  denote the set of fully complete part

instances with  $i$ -th pseudo label in  $m$ -th view, the similarity of  $C_i^m$  and  $C_j^n$  can be calculated by  $n_{i,j}^{m,n}$ , the number of same instances from  $C_i^m$  and  $C_j^n$ . The km algorithm is utilized to explore the category-level cross-view correspondences  $P$  according to the similarity.  $P^{(m,n)}(i) = j \Leftrightarrow P^{(n,m)}(j) = i$  means the  $i$ -th cluster of  $m$  view corresponds to the  $j$ -th cluster of  $n$ -th view, and vice versa.

Considering about the requirement of category-level cross-view correspondences for dealing with category-level information, the centroid of  $S^m$  in instance set  $C_i^m$  is

$$\bar{S}_i^m = \frac{1}{|C_i^m|} \sum_{X_j} S_j^m \quad (6)$$

Then, for  $i$ -th instance in  $v$ -th view which has  $j$ -th pseudo label, the representation with cross-view information  $S_{fu_i}^v$  can be calculated as

$$S_{fu_i}^v = \frac{1}{V} \left( S_i^v + \sum_{m \neq v} \bar{S}_{P^{(v,m)}(j)}^m \right) \quad (7)$$

## Graph Structure Learning

To explore accurate graph structure combined cross-view information, the self-representation learning is integrated into AE structure. Assuming  $G$  as the affinity matrix, the requirement of self-representation learning idea on  $S_{fu}$  is  $S_{fu} = GS_{fu}$ . Considering the symmetry of the AE structure,  $GS_{fu}$  is utilized as the input of decoder  $f_{de}$ . The objective function of graph learning can be formulated as

$$L_{gl} = \|X - f_{de}(GS_{fu})\|_F^2 + \|S_{fu} - GS_{fu}\|_F^2 \quad (8)$$

However, the affinity matrix  $G$  is dense and inevitably includes noisy during graph learning. Since there is a natural correlation between groundtruth label and exact graph structure, the pseudo label  $Y$  and the affinity matrix  $G$  should interact on each other. The pseudo 0/1 graph structure  $A$  based on pseudo label  $Y$  can be calculated by

$$A = YY^T \quad (9)$$

After normalization, the pseudo affinity matrix  $\tilde{G}$  can be easily obtained that

$$\tilde{G}_{ij} = \frac{A_{ij}}{\sum_i^n 1\{A_{ij} = 1\}} \quad (10)$$

where  $1\{a = b\}$  means a discriminant function as  $1\{\text{true}\} = 1$  and  $1\{\text{false}\} = 0$ .

To ensure the sparseness and accuracy of  $G$ , the objective function of graph constraint can be written as

$$L_{gc} = \left\| G - \tilde{G} \right\|_F^2 \quad (11)$$

The whole process of our proposed method is shown in Algorithm 1.

## Experiments

In this section, the proposed MGCCFF is compared with the state-of-the-art incomplete multi-view methods and unpaired multi-view methods on real-world datasets with different completion rates and pairing rates, and the results are evaluated using ACC, NMI, and ARI.

---

### Algorithm 1: The proposed MGCCFF

---

**Input:** Multi-view dataset  $X$ , cluster number  $K$ , epoch  $E$ ;  
**Output:** Cluster assignments  $L$ , latent representation  $S$ , affinity matrix  $G$ ;

- 1: Pretrain the autoencoder separately in each view by optimizing Eq. (5).
- 2: **while** not reaching  $E$  epochs **do**
- 3: Calculate the pseudo labels  $Y$  by Eq. (5).
- 4: Calculate the representation with cross-view information  $S$  by Eq. (7).
- 5: Calculate the pseudo affinity matrix  $\tilde{G}$  by Eq. (10).
- 6: Calculate the affinity matrix  $G$  by Eq. (11) and Eq. (8).
- 7: **end while**
- 8: Apply the outputs to subsequent clustering tasks.

---

## Datasets

We evaluate our model MGCCFF on datasets with 50%, 70% complete rates and 10%, 30%, 50%, 70% paired rates including

- **Handwritten** (Zhang et al. 2020a) consists 2000 samples of 10 classes. Two different types of features pix and profile correlations are used as two views.
- **BDGP** (Cai et al. 2012) contains image and text modalities of 2500 samples of 5 categories.
- **Caltech101-20** (Li et al. 2015) consists of 2386 images with 20 classes and 6 views, we employ 254 dimension CENTRIST feature and 512 dimension GIST feature as two views for experiments.
- **BBCSport** (Luo et al. 2018) is composed of news articles in 5 topical areas from BBC website, which is associated with 2 views.
- **Yale** (Luo et al. 2018) consists 165 grayscale images of 15 individuals, from which 3 types of features are extracted. we employ 4096 dimension feature and 3304 dimension feature as two views for experiments.

## Compared Methods

In order to show the excellent performance of the proposed MGCCFF, this paper introduces some single-view methods, several state-of-the-art incomplete multi-view algorithms, and unpaired multi-view algorithms for experiments. Single-view methods include KMeans, Spectral Clustering (SC), Non-negative Matrix Factorization (NMF), and AE.

Incomplete multi-view algorithms include:

**iCmSC** (Wang et al. 2021) combines the AE and CCA for a shared graph structure guided by self-representation learning idea. **CDIMC-net** (Wen et al. 2020) embeds the graph information into the original data representation and learns clustering results by designed self-paced means. **COMPLETER** (Lin et al. 2021) introduces the mutual information and utilizes the autoencoder structure to obtain consistent feature representations. **CRTC** (Wang et al. 2022) constructs an end-to-end multi-view clustering framework by integrating graph learning embedding idea and inferring

Paired Rates		10%		30%		50%		70%	
Methods		NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
Incomplete Methods	iCmSC	40.27±1.01	35.43±1.37	45.15±0.42	39.70±0.76	47.97±0.97	43.09±1.35	54.68±0.44	47.00±0.76
	CDIMC-net	46.43±0.37	41.15±0.61	48.18±0.18	42.19±0.14	53.60±0.36	48.29±0.63	58.39±0.23	49.69±1.10
	COMPLETER	44.62±0.49	38.19±0.34	50.46±0.16	41.44±0.94	52.13±0.47	45.45±0.36	60.20±0.11	55.60±0.95
	CRTC	51.45±1.45	45.01±2.12	53.46±0.64	46.61±0.92	59.37±2.52	52.13±3.11	60.50±0.67	54.44±0.87
	UAGCN	50.58±0.22	41.46±0.21	52.42±0.32	44.12±0.23	57.42±0.13	49.35±0.21	61.39±0.21	55.96±0.23
	IMRL-AGI	51.26±0.13	46.14±0.19	53.57±0.23	48.68±0.15	58.92±0.21	51.17±0.12	62.59±0.17	56.20±0.22
Unpaired Methods	MvCLN	40.15±1.54	33.48±1.74	43.25±1.07	35.16±2.04	47.11±1.55	40.27±1.76	49.47±1.11	40.28±2.05
	MVC-UM	17.01±1.51	3.71±2.61	22.41±1.34	9.63±0.74	23.02±1.47	10.68±2.59	28.93±1.37	15.03±0.78
	UPMGC-SM	47.31±0.64	37.11±0.70	50.94±1.24	44.84±2.96	54.09±0.66	43.82±0.73	57.81±1.27	50.66±2.94
	DGPPVC	50.42±0.15	39.05±0.23	53.40±0.13	44.43±0.22	57.94±0.15	50.05±0.21	61.08±0.17	56.49±0.21
<b>MGCCFF</b>		<b>67.61±0.16</b>	<b>58.84±0.63</b>	<b>71.16±0.06</b>	<b>63.41±0.10</b>	<b>75.57±0.18</b>	<b>66.14±0.65</b>	<b>79.29±0.10</b>	<b>71.17±0.21</b>

Table 1: Clustering average results and standard deviations of incomplete methods and unpaired methods on the Handwritten dataset with **50% complete rate** and different paired rates.

Paired Rates		10%		30%		50%		70%	
Methods		NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
Incomplete Methods	iCmSC	46.85±0.97	41.55±1.41	50.84±0.38	45.76±0.73	50.75±1.04	46.30±1.41	58.01±0.40	50.67±0.78
	CDIMC-net	53.17±0.35	48.01±0.66	55.06±0.14	48.50±1.17	57.31±0.33	51.22±0.61	61.48±0.16	52.85±0.91
	COMPLETER	51.08±0.47	44.61±0.35	53.88±0.21	47.42±0.92	54.75±0.49	49.10±0.36	63.13±0.16	58.56±0.61
	CRTC	58.18±1.04	52.01±1.91	60.42±0.68	53.12±0.96	61.51±1.01	55.10±2.28	64.28±0.65	57.26±0.93
	UAGCN	56.43±0.21	45.32±0.22	59.22±0.12	50.32±0.21	62.31±0.11	55.24±0.20	65.37±0.20	60.71±0.23
	IMRL-AGI	57.27±0.12	53.21±0.15	60.77±0.22	55.62±0.15	63.92±0.19	58.17±0.12	66.31±0.16	61.59±0.17
Unpaired Methods	MvCLN	48.82±1.54	42.16±1.78	51.46±1.03	43.91±2.04	52.87±1.54	46.24±1.75	56.03±1.09	47.55±2.00
	MVC-UM	25.58±1.49	12.42±2.60	28.16±1.30	18.45±0.69	29.58±1.48	16.43±2.60	35.02±1.31	21.95±0.75
	UPMGC-SM	55.57±0.61	45.73±0.70	59.72±1.25	45.95±2.92	60.29±0.64	49.95±0.74	63.16±1.27	57.80±2.92
	DGPPVC	57.31±0.14	48.22±0.21	60.57±0.11	52.13±0.20	63.89±0.15	57.13±0.22	66.17±0.16	62.47±0.20
<b>MGCCFF</b>		<b>78.23±0.15</b>	<b>69.41±0.62</b>	<b>81.97±0.08</b>	<b>74.24±1.51</b>	<b>84.83±0.24</b>	<b>78.61±0.71</b>	<b>88.84±0.13</b>	<b>86.11±0.41</b>

Table 2: Clustering average results and standard deviations of incomplete methods and unpaired methods on the Handwritten dataset with **70% complete rate** and different paired rates.

the missing part information. **UAGCN** (Zhao et al. 2023) combines unrestricted anchor graph construction with GCN to enhance incomplete multi-view clustering by effectively capturing graph structures even in high missing-rate scenarios. **IMRL-AGI** (Liu et al. 2024) combines an anchor graph-based GCN with an information bottleneck for robust incomplete multi-view representation learning.

Unpaired multi-view algorithms include:

**MvCLN** (Yang et al. 2021) establishes distance metrics between samples through contrastive learning and fuses cross-view feature representations for clustering. **MVC-UM** (Yu et al. 2021) considers the local distribution information of the data and utilizes sample-level correspondences to connect cross-view representations. **UPMGC-SM** (Wen et al. 2023b) extends self-representation learning into multi-view unpaired applications and learns cross-view graph mapping relationships for graph fusion. **DGPPVC** (Zhao et al. 2024) combines GCN and a dynamic graph structure for progressively learning partial view-aligned clustering.

Incomplete and unpaired multi-view algorithms include:

**SURE** (Yang et al. 2023) adopts a novel noise-robust contrastive learning paradigm for incomplete and unpaired data to capture categorical similarities and establish correspondences between views.

## Experimental Settings and Metrics

To measure experiments accurately, All algorithms are run 5 times and report the average results with the standard deviation. All the following experiments are implemented using sklearn library and PyTorch library on a windows 10 OS with an NVIDIA 3090 GPU.

To measure more accurately, three widely-used clustering metrics including Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI) are used to verify the clustering performance. The **best** values are highlighted in **bold**.

Since only SURE and MGCCFF can address both incomplete and unpaired problems simultaneously, for fair comparisons, the results of other tested methods that solve only one of these issues are reported under the following four settings:

- Setting 1 (Unpaired Methods): Since Unpaired Methods can only handle partially view-aligned data, for the experiments utilizing both incomplete and Unpaired data, the mean of the complete portions of the data will be used as imputation to address the incompleteness.
- Setting 2 (Incomplete Methods): For Incomplete Methods, we first apply the PCA method to reduce the dimen-

Paired Rates		10%		30%		50%		70%	
Methods		NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
Incomplete Methods	iCmSC	38.43±0.91	31.96±3.28	45.48±0.56	30.13±5.11	50.79±0.51	36.04±5.12	54.52±0.52	39.61±5.10
	CDIMC-net	37.94±0.66	26.16±2.28	42.46±0.61	38.43±2.18	47.96±0.63	44.39±2.21	51.09±0.60	47.59±2.22
	COMPLETER	41.16±1.51	27.24±4.47	53.13±1.47	41.95±2.86	58.42±1.52	47.46±2.84	62.26±1.49	51.33±2.83
	CRTC	40.15±4.78	32.55±5.53	54.72±4.99	42.16±6.44	60.18±4.95	47.92±6.24	64.05±4.93	51.57±6.48
	UAGCN	42.22±1.05	30.19±0.78	54.12±0.92	40.35±0.49	61.29±1.09	48.73±0.78	65.31±1.14	52.31±0.49
	IMRL-AGI	43.23±0.24	37.22±0.13	55.45±0.19	41.22±0.17	62.74±0.18	49.02±0.12	65.74±0.12	53.11±0.21
Unpaired Methods	MvCLN	44.16±0.84	38.13±2.16	54.18±0.51	45.91±0.31	61.45±0.51	53.45±0.35	65.53±0.50	57.87±0.30
	MVC-UM	28.43±0.64	19.16±1.46	48.31±0.61	35.84±1.34	55.63±0.61	43.32±1.38	59.80±0.69	47.92±1.40
	UPMGC-SM	48.16±1.01	36.69±3.16	57.16±2.14	42.71±4.81	64.56±2.19	49.97±4.81	69.37±2.17	54.20±4.81
	DGPPVC	51.27±0.14	40.62±0.19	58.38±0.16	46.24±0.17	63.47±0.18	53.13±0.22	68.67±0.18	59.21±0.19
<b>MGCCFF</b>		<b>62.48±0.34</b>	<b>48.97±1.03</b>	<b>71.91±0.08</b>	<b>65.41±0.13</b>	<b>76.61±0.06</b>	<b>66.46±0.12</b>	<b>80.54±0.03</b>	<b>69.44±0.06</b>

Table 3: Clustering average results and standard deviations of incomplete methods and unpaired methods on the BDGP dataset with **50% complete rate** and different paired rates.

Paired Rates		10%		30%		50%		70%	
Methods		NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
Incomplete Methods	iCmSC	42.45±0.92	36.93±3.26	50.05±0.55	34.55±1.05	55.55±0.52	40.49±1.51	59.29±0.53	43.80±1.36
	CDIMC-net	42.28±0.66	31.11±2.34	47.21±0.62	42.74±2.18	52.99±0.60	47.75±2.22	56.30±0.60	51.52±2.18
	COMPLETER	45.74±1.49	32.15±4.48	57.33±1.46	46.24±2.90	62.81±1.50	52.24±2.91	66.24±1.54	51.64±2.92
	CRTC	45.12±4.81	36.73±5.55	59.12±5.01	47.08±6.44	64.79±2.02	52.43±4.61	68.41±4.11	55.48±6.84
	UAGCN	46.18±2.23	35.49±0.86	58.79±1.27	46.26±0.73	65.59±1.25	52.71±0.84	69.82±1.18	56.31±0.69
	IMRL-AGI	47.23±0.28	41.63±0.17	59.45±0.21	47.18±0.19	66.04±0.16	53.32±0.17	69.58±0.15	57.58±0.19
Unpaired Methods	MvCLN	49.90±0.80	43.54±2.19	59.33±0.54	51.61±0.33	66.43±0.52	58.71±0.34	71.24±0.72	63.34±1.41
	MVC-UM	33.78±0.67	24.92±1.49	54.20±0.65	41.19±1.38	61.56±0.33	48.39±1.34	66.49±0.54	53.98±1.36
	UPMGC-SM	53.19±1.01	41.97±3.18	62.57±2.15	47.84±4.82	70.21±1.49	54.69±4.83	72.19±2.01	60.49±4.86
	DGPPVC	56.73±0.18	46.59±0.24	63.52±0.23	51.34±0.21	69.53±0.27	58.31±0.22	73.62±0.35	65.19±0.35
<b>MGCCFF</b>		<b>66.92±0.37</b>	<b>53.17±1.04</b>	<b>76.71±0.09</b>	<b>69.84±0.12</b>	<b>79.98±0.03</b>	<b>72.98±0.09</b>	<b>81.06±0.08</b>	<b>75.94±0.12</b>

Table 4: Clustering average results and standard deviations of incomplete methods and unpaired methods on the BDGP dataset with **70% complete rate** and different paired rates.

sionality of the incomplete and unpaired data to a suitable latent space, allowing the Hungarian algorithm to be applied to establish correspondences of the partially view-aligned data. Following this, we conduct these Incomplete baselines on the realigned data.

- Setting 3 (Single view method): When comparing our proposed model to single-modal clustering algorithms, the input data for single-modal models are completely paired and 50% complete rate.
- Setting 4 (Incomplete and Unpaired Methods): For MGCCF and SURE, data with varying complete and pair rates are directly inputted.

### Clustering Results and Analysis

The clustering results of our proposed MGCCFF and several state-of-the-art incomplete multi-view methods and unpaired multi-view methods on Handwritten and BDGP datasets with different complete and paired rates are presented in Tables 1 to 4. With the increase of the complete rate and paired rate, the performance of all methods becomes better. Based on graph learning idea, the results of iCmSC, CRTC, UAGCN, IMRL-AGI, UPMGC-SM, and DGPPVC are better than those of other methods which neglect graph

information and indicate that graph information plays an extremely important role in incomplete and unpaired multi-view scenarios. Considering both incomplete and unpaired situations, our proposed method MGCCFF significantly outperforms other methods in different situations.

Besides, to further validate our model’s effectiveness, we conducted comparative experiments on three additional datasets with SURE, a method also capable of managing incomplete and unpaired data. From Table 5, our model achieved optimal experimental results on the Caltech101-20, BBCSport, and Yale datasets. In Table 5, the ‘-’ symbol indicates that experiments could not be conducted due to issues with the model setup.

The success of MGCCFF can be attributed to the model’s embedding, which learns the spectral structure of samples through self-expression ideas and an autoencoder structure. Furthermore, we use pseudo-labels to find class-level alignments, reducing the likelihood of introducing pairs of noisy samples. These measures help improve the overall performance.

### Ablation Study

To demonstrate the effectiveness of the ability to introduce cross-view information of our proposed methods, we carry

Complete Rate	Pair rate		10			30			50			70		
	Dataset	Method	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
50	Caltech101-20	SURE	22.82	24.76	9.81	29.30	28.97	14.57	28.69	29.40	15.45	28.21	31.37	15.78
		<b>MGCCFF</b>	<b>29.69</b>	<b>29.74</b>	<b>18.74</b>	<b>30.32</b>	<b>29.00</b>	<b>18.24</b>	<b>31.74</b>	<b>30.33</b>	<b>19.52</b>	<b>33.31</b>	<b>31.75</b>	<b>19.54</b>
	BBC	SURE	35.26	2.05	0.02	33.42	2.83	-0.32	32.87	3.76	-0.07	33.49	5.64	0.68
		<b>MGCCFF</b>	<b>36.49</b>	<b>11.64</b>	<b>7.25</b>	<b>37.61</b>	<b>13.73</b>	<b>6.66</b>	<b>38.29</b>	<b>15.53</b>	<b>9.62</b>	<b>42.63</b>	<b>17.19</b>	<b>12.02</b>
	Yale	SURE	-	-	-	30.30	34.77	6.27	32.85	37.44	9.37	-	-	-
		<b>MGCCFF</b>	<b>37.12</b>	<b>48.42</b>	<b>17.57</b>	<b>38.56</b>	<b>49.52</b>	<b>19.62</b>	<b>40.80</b>	<b>49.08</b>	<b>19.16</b>	<b>42.58</b>	<b>52.36</b>	<b>24.19</b>
70	Caltech101-20	SURE	25.46	27.75	14.39	29.96	29.88	16.25	29.29	31.05	16.78	29.85	31.16	17.41
		<b>MGCCFF</b>	<b>31.97</b>	<b>29.47</b>	<b>18.48</b>	<b>32.73</b>	<b>29.89</b>	<b>20.10</b>	<b>33.16</b>	<b>31.48</b>	<b>20.50</b>	<b>34.25</b>	<b>31.29</b>	<b>21.20</b>
	BBC	SURE	35.62	2.66	-0.01	33.68	3.01	-0.01	32.24	3.29	-0.03	32.68	3.40	0.91
		<b>MGCCFF</b>	<b>38.53</b>	<b>12.05</b>	<b>8.10</b>	<b>39.09</b>	<b>15.26</b>	<b>9.40</b>	<b>41.26</b>	<b>16.66</b>	<b>11.45</b>	<b>44.11</b>	<b>17.73</b>	<b>12.35</b>
	Yale	SURE	-	-	-	32.12	37.91	9.56	31.64	37.12	9.18	34.91	41.13	13.72
		<b>MGCCFF</b>	<b>38.13</b>	<b>48.48</b>	<b>19.36</b>	<b>39.13</b>	<b>48.80</b>	<b>19.64</b>	<b>41.33</b>	<b>51.16</b>	<b>22.03</b>	<b>43.87</b>	<b>53.30</b>	<b>24.25</b>

Table 5: Clustering average results of incomplete and unpaired methods on the different datasets with different complete rates and different paired rates.

Datasets	Handwritten				BDGP			
	view 1		view 2		view 1		view 2	
Methods	NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
<i>k</i> -means	60.82±0.47	46.57±0.60	46.93±0.00	27.95±0.00	19.58±0.41	14.49±0.29	44.99±1.73	15.73±1.98
SC	57.51±0.02	36.03±0.03	48.93±0.00	34.64±0.00	22.22±0.00	17.99±0.00	53.14±0.00	34.60±0.00
NMF	65.32±0.93	53.32±1.11	49.03±0.04	21.93±0.04	16.05±0.12	9.22±0.02	46.26±1.68	17.15±1.68
AE	70.33±2.73	58.92±4.40	45.64±2.73	34.73±4.40	20.76±0.50	15.64±0.37	55.38±0.22	38.70±0.40
<b>MGCCFF(30%)</b>	81.39±0.05	73.39±0.02	70.33±0.11	58.92±0.14	30.15±0.00	22.19±0.00	76.03±0.00	66.33±0.00
<b>MGCCFF(50%)</b>	88.85±0.11	87.31±0.10	71.69±0.00	60.06±0.00	32.02±0.00	25.71±0.00	77.79±0.00	65.14±0.00
<b>MGCCFF(70%)</b>	<b>89.19±0.00</b>	<b>88.53±0.00</b>	<b>72.98±0.00</b>	<b>64.80±0.00</b>	<b>34.35±0.05</b>	<b>24.94±0.03</b>	<b>82.27±0.00</b>	<b>71.59±0.00</b>

Table 6: Clustering average results and standard deviations on single view of Handwritten and BDGP datasets with 50% complete rate.

$L_{pl}$	$L_{gl}$	$L_{gc}$	ACC	NMI	ARI
✓			27.28	28.76	14.71
✓	✓		28.23	29.02	14.97
✓		✓	29.61	28.35	16.49
	✓	✓	24.62	19.44	7.12
✓	✓	✓	<b>31.74</b>	<b>30.33</b>	<b>19.52</b>

Table 7: the ablation study on the Caltech101-20 dataset with 50% complete and 50% paired rate, the symbol "✓" indicates the MGCCFF with the inclusion of components.

out the ablation experiments on single view of multi-view dataset with different paired rates. Table 6 shows the clustering results of single-view methods on Handwritten and BDGP datasets with 50% complete rate, and the results of MGCCFF on single view with cross-view information when given different paired rates. Table 6 proves that the proposed MGCCFF can introduce valuable information by fusing incomplete and unpaired multi-view data and be applied on multi-view datasets with different complete and paired rates.

To further verify the contributions of the proposed method, we conduct an ablation study on the Caltech101-20 that shows in Table 7. In Table 7,  $L_{pl}$  stands for Pseudo Label Learning,  $L_{gl}$  denotes Cross-view Feature Fusion and

partial Graph Structure Learning, and  $L_{gc}$  represents affinity matrix-based graph structure learning.  $L_{pl}$  ensures inter-view separability and intra-view compactness through DDC loss to obtain better clustering results and learn pseudo labels. The absence of this component significantly impacts subsequent Graph Structure Learning via pseudo labels. Meanwhile,  $L_{gl}$  and  $L_{gc}$  ensure that the latent representations of each view are well integrated with cross-view sample information and graph structure information. Table 7 demonstrates the crucial importance of these modules, indicating that the removal of any module would result in a significant decrease in downstream clustering performance.

## Conclusion

In this paper, we propose a novel multi-view clustering framework for incomplete and unpaired multi-view data termed Multi-view Graph Clustering framework with Cross-view Feature Fusion (MGCCFF). MGCCFF captures the category-level cross-view correspondences based on pseudo labels for integrating cross-view information and learns the spectral structure for the final clustering task by exploring the latent relationships between pseudo-labels and the graph and imposing a sparse and residual constraint on the graph structure. Compared with existing state-of-the-art methods, extensive experiments under various conditions demonstrate the superiority on incomplete and unpaired multi-view data.

## Acknowledgments

This work is supported by the Science and Technology Project of Dalian City (2024JJ12GX025, 2023JJ12SN029, and 2023JJ11CG005).

## 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.
- Gao, H.; Nie, F.; Li, X.; and Huang, H. 2015. Multi-View Subspace Clustering. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 4238–4246.
- Huang, S.; Liu, Y.; Tsang, I. W.; Xu, Z.; and Lv, J. 2023. Multi-View Subspace Clustering by Joint Measuring of Consistency and Diversity. *IEEE Transactions on Knowledge and Data Engineering*, 35(8): 8270–8281.
- Huang, Z.; Hu, P.; Zhou, J. T.; Lv, J.; and Peng, X. 2020. Partially View-aligned Clustering. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems*, volume 33, 2892–2902. Curran Associates, Inc.
- Kampffmeyer, M.; Løkse, S.; Bianchi, F. M.; Livi, L.; Salberg, A.-B.; and Jenssen, R. 2019. Deep divergence-based approach to clustering. *Neural Networks*, 113: 91–101.
- Kang, Z.; Shi, G.; Huang, S.; Chen, W.; Pu, X.; Zhou, J. T.; and Xu, Z. 2020. Multi-graph fusion for multi-view spectral clustering. *Knowledge-Based Systems*, 189: 105102.
- Li, X.; Zhang, H.; Wang, R.; and Nie, F. 2022a. Multiview Clustering: A Scalable and Parameter-Free Bipartite Graph Fusion Method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1): 330–344.
- Li, Y.; Nie, F.; Huang, H.; and Huang, J. 2015. Large-Scale Multi-View Spectral Clustering via Bipartite Graph. In *AAAI Conference on Artificial Intelligence*.
- Li, Z.; Tang, C.; Zheng, X.; Liu, X.; Zhang, W.; and Zhu, E. 2022b. High-Order Correlation Preserved Incomplete Multi-View Subspace Clustering. *IEEE Transactions on Image Processing*, 31: 2067–2080.
- Lin, Y.; Gou, Y.; Liu, Z.; Li, B.; Lv, J.; and Peng, X. 2021. Completer: Incomplete multi-view clustering via contrastive prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11174–11183.
- Liu, X.; Zhu, X.; Li, M.; Wang, L.; Tang, C.; Yin, J.; Shen, D.; Wang, H.; and Gao, W. 2019. Late Fusion Incomplete Multi-View Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(10): 2410–2423.
- Liu, Z.; Wang, X.; Huang, X.; Li, G.; Sun, K.; and Chen, Z. 2024. Incomplete Multi-View Representation Learning Through Anchor Graph-Based GCN and Information Bottleneck. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.
- Luo, S.; Zhang, C.; Zhang, W.; and Cao, X. 2018. Consistent and specific multi-view subspace clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.
- Nie, F.; Li, J.; and Li, X. 2017. Self-weighted multiview clustering with multiple graphs. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI'17*, 2564–2570. AAAI Press. ISBN 9780999241103.
- Nie, F.; Zeng, Z.; Tsang, I. W.; Xu, D.; and Zhang, C. 2011. Spectral Embedded Clustering: A Framework for In-Sample and Out-of-Sample Spectral Clustering. *IEEE Transactions on Neural Networks*, 22(11): 1796–1808.
- Tang, J.; Yi, Q.; Fu, S.; and Tian, Y. 2024. Incomplete multi-view learning: Review, analysis, and prospects. *Applied Soft Computing*, 153: 111278.
- Wang, H.; Yang, Y.; and Liu, B. 2020. GMC: Graph-Based Multi-View Clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6): 1116–1129.
- Wang, Q.; Ding, Z.; Tao, Z.; Gao, Q.; and Fu, Y. 2018. Partial Multi-view Clustering via Consistent GAN. In *2018 IEEE International Conference on Data Mining (ICDM)*, 1290–1295.
- Wang, Q.; Lian, H.; Sun, G.; Gao, Q.; and Jiao, L. 2021. iCmSC: Incomplete Cross-Modal Subspace Clustering. *IEEE Transactions on Image Processing*, 30: 305–317.
- Wang, Y.; Chang, D.; Fu, Z.; Wen, J.; and Zhao, Y. 2022. Incomplete multiview clustering via cross-view relation transfer. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(1): 367–378.
- Wen, J.; Xu, Y.; and Liu, H. 2020. Incomplete Multiview Spectral Clustering With Adaptive Graph Learning. *IEEE Transactions on Cybernetics*, 50(4): 1418–1429.
- Wen, J.; Zhang, Z.; Fei, L.; Zhang, B.; Xu, Y.; Zhang, Z.; and Li, J. 2023a. 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 Xie, G.-S. 2020. CDIMC-net: Cognitive Deep Incomplete Multi-view Clustering Network. In Bessiere, C., ed., *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, 3230–3236. ijcai.org.
- Wen, Y.; Wang, S.; Liao, Q.; Liang, W.; Liang, K.; Wan, X.; and Liu, X. 2023b. Unpaired Multi-View Graph Clustering With Cross-View Structure Matching. *IEEE Transactions on Neural Networks and Learning Systems*, 1–15.
- Xu, C.; Tao, D.; and Xu, C. 2015. Multi-View Learning With Incomplete Views. *IEEE Transactions on Image Processing*, 24(12): 5812–5825.
- Yang, M.; Li, Y.; Hu, P.; Bai, J.; Lv, J.; and Peng, X. 2023. Robust Multi-View Clustering With Incomplete Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1): 1055–1069.
- Yang, M.; Li, Y.; Huang, Z.; Liu, Z.; Hu, P.; and Peng, X. 2021. Partially View-Aligned Representation Learning With Noise-Robust Contrastive Loss. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1134–1143.
- Yu, H.; Tang, J.; Wang, G.; and Gao, X. 2021. A Novel Multi-View Clustering Method for Unknown Mapping Relationships Between Cross-View Samples. In *Proceedings*

of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, KDD '21, 2075–2083. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383325.

Zhan, K.; Zhang, C.; Guan, J.; and Wang, J. 2018. Graph Learning for Multiview Clustering. *IEEE Transactions on Cybernetics*, 48(10): 2887–2895.

Zhang, B.; Qiang, Q.; Wang, F.; and Nie, F. 2020a. Fast multi-view semi-supervised learning with learned graph. *IEEE Transactions on Knowledge and Data Engineering*, 34(1): 286–299.

Zhang, C.; Fu, H.; Hu, Q.; Cao, X.; Xie, Y.; Tao, D.; and Xu, D. 2020b. Generalized Latent Multi-View Subspace Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(1): 86–99.

Zhang, C.; Hu, Q.; Fu, H.; Zhu, P.; and Cao, X. 2017. Latent Multi-View Subspace Clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4279–4287.

Zhang, X.; Zong, L.; Liu, X.; and Yu, H. 2015. Constrained NMF-Based Multi-View Clustering on Unmapped Data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Zhao, L.; Wang, Z.; Yuan, Y.; and Ding, F. 2023. Unrestricted Anchor Graph Based GCN for Incomplete Multi-View Clustering. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.

Zhao, L.; Xie, Q.; Li, Z.; Wu, S.; and Yang, Y. 2024. Dynamic Graph Guided Progressive Partial View-Aligned Clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 1–13.

Zhu, W.; Lu, J.; and Zhou, J. 2019. Structured general and specific multi-view subspace clustering. *Pattern Recognition*, 93: 392–403.

Zong, L.; Zhang, X.; and Liu, X. 2018. Multi-view clustering on unmapped data via constrained non-negative matrix factorization. *Neural Networks*, 108: 155–171.