# Highly Confident Local Structure Based Consensus Graph Learning for Incomplete Multi-view Clustering

Jie Wen[1†], Chengliang Liu[1†], Gehui Xu[1], Zhihao Wu[1], Chao Huang[2], Lunke Fei[3], Yong Xu[1,4*]

[1]Shenzhen Key Laboratory of Visual Object Detection and Recognition, Harbin Institute of Technology, Shenzhen, China
[2]School of Cyber Science and Technology, Shenzhen Campus of Sun Yat-sen University, Shenzhen, China
[3]School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China
[4]Pengcheng Laboratory, Shenzhen, China

## Abstract

*Graph-based multi-view clustering has attracted extensive attention because of the powerful clustering-structure representation ability and noise robustness. Considering the reality of a large amount of incomplete data, in this paper, we propose a simple but effective method for incomplete multi-view clustering based on consensus graph learning, termed as HCLS_CGL. Unlike existing methods that utilize graph constructed from raw data to aid in the learning of consistent representation, our method directly learns a consensus graph across views for clustering. Specifically, we design a novel confidence graph and embed it to form a confidence structure driven consensus graph learning model. Our confidence graph is based on an intuitive similar-nearest-neighbor hypothesis, which does not require any additional information and can help the model to obtain a high-quality consensus graph for better clustering. Numerous experiments are performed to confirm the effectiveness of our method.*

## 1. Introduction

In today's world, graphs are common tools to mine the intrinsic structure of data. Graph learning, as a powerful data analysis approach, has attracted increasing attention over the past few years [38, 42]. More and more machine learning and data mining tasks attempt to adopt graph learning to enhance the ability of structural representation and performance [35, 41, 43]. As an emerging data analysis and representation task, multi-view clustering also needs to deeply mine the geometric structure information of data samples [15, 36, 44]. That is also the core focus of this paper.

As we all know, multi-view data is collected from different sources or perspectives, which is a diverse description of the target, and this diversity endows the data with stronger

---
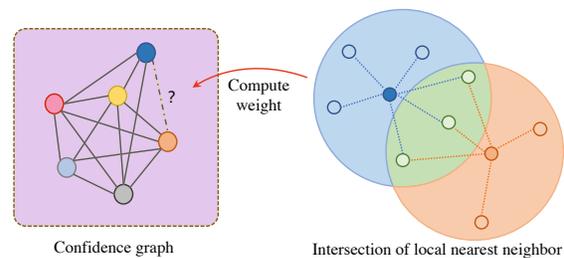*Corresponding author. †Co-first authors.

Figure 1. Our key motivation: A novel confidence graph whose edges are computed by the nodes' shared nearest neighbors.

discriminative and expressive power [19, 48]. For example, in the webpage recommendation tasks, we can assess the attributes of webpages from different elements such as audio, video, image, and text; in the autonomous driving scenario, ultrasonic radar, optical camera, and infrared radar together provide data basis for road traffic analysis. These different styles of multi-view data play an extremely important role in real-world applications and provide a decision-making basis for plentiful downstream tasks [4, 18]. To better mine and utilize multi-view data, researchers have proposed abundant multi-view learning methods. These methods can be broadly divided into the following categories according to different technical routes: multiple kernel kmeans based methods, canonical correlation analysis (CCA) based methods, subspace learning based methods, spectral clustering based methods, nonnegative matrix factorization (NMF) based methods, and deep learning based methods [2, 3, 5]. However, these conventional multi-view clustering methods are built on the premise that each sample has all complete views. Obviously, this idealized assumption is often violated in practice. Missing data is almost inevitable due to human oversight or the reasons from the application scenario itself [17]. Numerous solutions have been proposed for this incomplete multi-view clustering (IMC) problem. The classic PMVC (partial multi-view clustering) [17] seeks a common space to connect two in-

complete views. Liu *et al.* proposed a matrix factorization based model dubbed as LSIMVC [20], which directly learns a consistent low-dimensional representation of multi-view feature data and introduces prior indicator information to help avoid the negative effects of missing instances. Online multi-view clustering (OMVC) [27] and multiple incomplete Views Clustering (MIVC) [28] are designed to utilize the weight matrix to assign low weights to the missing instances in the joint matrix factorization based representation learning model. Wen *et al.* proposed a framework named UEAF (unified embedding aligned framework) [39] to recover missing views with forward and backward graph constraints and perform multi-view clustering simultaneously. In recent years, IMC methods based on graph learning are attracting increasing research enthusiasm. Compared with matrix factorization methods, graph based methods can more intuitively and effectively mine the intrinsic geometric structure of data. In this branch, Gao *et al.* [8] tried to fill the missing instances by the view average and learn a consistent cluster indicator matrix by kernel alignment method and the spectrogram theory. Wen *et al.* [37] proposed to perform spectral clustering on the subspace representation of each view and align all views in kernel space to obtain a consensus cluster indicator matrix.

The success of existing graph based IMC (GIMC) methods can be attributed to the fact that the discriminative grouping information of data naturally fits the clustering task. However, these GIMC methods still have some limitations. On the one hand, some methods are only applicable to two-view datasets or require at least one view to be complete. On the other hand, the way of constructing graphs directly from raw data is easily affected by noise, *e.g.*, graphs based on Gaussian kernel function are not necessarily suitable for all datasets while the self-representation of raw data is often susceptible to outliers. We believe that directly constructing or learning of individual graphs and consistent graph can reveal the internal structure of the data more flexibly and intuitively. In response to these issues, in this paper, we propose a highly confident local structure based consensus graph learning method (HCLS_CGL) for IMC tasks. Unlike existing GIMC methods that learn a clustering indicator matrix, our goal is to learn a cross-view consensus graph containing several connected subgraphs and revealing the intrinsic structure of the data intuitively. More importantly, in order to mine the pure data structure more deeply from the original graphs and to improve the noise resistance, we propose an innovative confidence graph constraint and fuse it with the consensus graph learning into one term. Overall, we summarize our contributions as follows:

- In this paper, we propose a model named HCLS_CGL for the IMC task, which is able to handle arbitrary view-missing situations. Extensive experiments are performed to confirm the effectiveness of our method.

- Different from the existing GIMC methods, we design a novel confidence nearest neighbor graph with the group-neighbors'similarity of the sample pair to guide the learning of consensus graph.

## 2. Preliminary

### 2.1. Notations

In our paper, we use uppercase $X$ to denote the matrix, and use $X_{:,i}$, $X_{i,:}$, and $X_{i,j}$ to represent its $i$-th column vector, $i$-th row vector, and $(i,j)$-th element, respectively. In addition, we define $\mathbf{1}$ and $I$ as the matrix with all 1 and vector with all 1, respectively. For a square matrix $Z \in \mathbb{R}^{n \times n}$, its Laplacian matrix is defined as $L_Z = D_Z - \frac{|Z|+|Z|^T}{2}$, where $D_Z$ is a diagonal matrix with the $i$-th diagonal element as $(D_Z)_{i,i} = \sum_{j=1}^{n} \frac{|Z_{i,j}|+|Z_{j,i}|}{2}$. $|Z|$ denotes the absolute operation on all elements of matrix $Z$. $Z \geq 0$ means all elements of matrix $Z$ are not less than 0. $rank(Z)$ denotes the rank of matrix $Z$.

Multi-view data with $l$ views and $n$ samples is denoted by $\{Y^{(v)}\}_{v=1}^{l}$, where $Y^{(v)} \in R^{m_v \times n}$ represents the $v$-th view data and each column can be regarded as a feature representation of the corresponding sample in the $v$-th view. $m_v$ is the feature dimension of instances in the $v$-th view.

### 2.2. Laplacian rank based graph learning

Graph based spectral clustering is one of the mainstream branches in data clustering methodologies [23, 33]. It achieves the clustering results of data via three steps: affinity graph construction, eigenvalue decomposition on Lapcian matrix of graph, and kmeans clustering. In spectral clustering, the affinity graph is generally an $n \times n$ matrix for the data with $n$ samples and its elements generally reveal the relationships among samples, such as similarity degree or distance. Among these steps, the affinity graph construction is very crucial to the final clustering results. In recent years, to obtain a high-quality affinity graph that can reveal the intrinsic relationships among samples, various affinity graph learning methods have been proposed [7, 16, 21, 25], where the constrained Laplacian rank (CLR) based method [25] is one the representative works. CLR seeks to learn a new similarity matrix with exactly $c$ blocks corresponding to the $c$ clusters from the preconstructed similarity graph $S \in \mathbb{R}^{n \times n}$ of the data with $n$ samples and $c$ clusters as follows:

$$\min_{S} \|S - Z\|_F^2$$
$$s.t. 0 \leq S \leq 1, S1 = 1, rank(L_S) = n - c \quad (1)$$

where constraint $S1 = 1$ is introduced to avoid the zero vector for some rows of matrix $S$. Based on the Theorem that the number of the connected components in graph $S$ is equal to the number of the eigenvalue zero in its Laplacian matrix, introducing the rank constraint $rank(L_S) = n - c$

on the Laplacian matrix thus has the potential to obtain an optimal graph $S$ with exactly $c$ connected components that intuitively reflects the cluster structure of data [25].

## 3. The proposed method

### 3.1. Learning model of HCLS_CGL

Various graph learning based clustering methods demonstrate that learning a high-quality graph that reveals the intrinsic relationships of data is beneficial to obtain a higher clustering performance. To this end, based on the assumption that different views should comply with a consistent clustering result, many methods seek to learn a consensus high-quality graph $S \in \mathbb{R}^{n \times n}$ by fully considering the information of all views for multi-view clustering, where an intuitive approach is formulated as follows [24]:

$$\min_{\phi(S,\alpha)} \sum_{v=1}^{l} \alpha_v^r \left\| S - Z^{(v)} \right\|_F^2 + \lambda \|S\|_F^2 \qquad (2)$$

where $\phi(S,\alpha) = \{0 \leq S \leq 1, S1 = 1, rank(L_S) = n - c, 0 \leq \alpha_v \leq 1, \sum_{v=1}^{l} \alpha_v = 1\}$ denotes the boundary constraint of variables $S$ and $\alpha$. $l$ denotes the view number of the data. $Z^{(v)} \in \mathbb{R}^{n \times n}$ is the similarity graph, such as k-nearest neighbor graph, pre-constructed from the $v$-th view. After obtaining the consensus graph $S$ by optimizing (2), spectral clustering [23] or connected component search method [31] can be performed on $S$ to obtain the clustering result of the given data. $r$ is a tunable parameter to control the distribution of the coefficient vector $\alpha$.

**Problem and Motivation 1**: The consensus graph learning model formulated by (2) is succinctly and effectively for multi-view clustering. However, it fails to the application of multi-view clustering case with missing views. From the aspect of model formulation, the problem to the GIMC includes the diversity dimensions and unaligned structure of incomplete graphs constructed from the un-missing instances of all views. To address these issues, some works focus on recovering the missing elements of the incomplete graphs associated with the missing views [39]. However, these works generally have a high computational cost. In addition, it is impossible to obtain the perfect recovered graphs from the incomplete multi-view data, which may in turn decrease the clustering performance. To this end, in this paper, we seek to learn a consensus graph from the certain similarity information from the available instances rather than the recovered uncertain graphs. Specifically, based on the prior location indexes $O \in \{0,1\}^{n \times l}$ of missing views, where $O_{i,j} = 1$ denotes that the $j$-th view of the $i$-th sample is not missing, otherwise $O_{i,j} = 0$, we design the following formula to learn such a consensus graph $S \in R^{n \times n}$ from the certain graph information correspond-

ing to the un-missing views:

$$\min_{S,\alpha} \sum_{v=1}^{l} \left( \left\| G^{(v)} S G^{(v)T} - \tilde{Z}^{(v)} \right\|_F^2 \right) + \lambda_1 \|S\|_F^2$$
$$s.t. 0 \leq \alpha_v \leq 1, \sum_{v=1}^{l} \alpha_v = 1, S^T 1 = 1, \qquad (3)$$
$$0 \leq S \leq 1, diag(S) = 0, rank(L_S) = n - c$$

where $\lambda_1$ is a penalty parameter. $\tilde{Z}^{(v)} \in \mathbb{R}^{n_v \times n_v}$ is a similarity graph pre-constructed from the $v$-th view data with $n_v$ available instances. Element $\tilde{Z}_{i,j}^{(v)}$ represents a kind of similarity relationships between the $i$-th and $j$-th available instances in the $v$-th view. In our work, $\tilde{Z}^{(v)}$ is constructed by the recognizable k-nearest neighbor algorithm as follows:

$$\tilde{Z}_{i,j}^{(v)} = \begin{cases} e^{-\frac{\left\| x_i^{(v)} - x_j^{(v)} \right\|_2^2}{2}}, & if \ x_j^{(v)} \in \psi(x_i^{(v)}) \ or \\ & \qquad x_i^{(v)} \in \psi(x_j^{(v)}) \\ 0, & otherwise \end{cases} \qquad (4)$$

where $\psi\left(x_i^{(v)}\right)$ denotes the instance set composed of the k nearest neighbors of instance $x_i^{(v)}$.

In model (3), $G^{(v)} \in \{0,1\}^{n_v \times n}$ is a transformation matrix to align the pairwise similarity relationships of samples in graph $S$ and $\tilde{Z}^{(v)}$, which is constructed as follows according to the prior location indexes $O \in \{0,1\}^{n \times l}$:

$$G_{i,j}^{(v)} = \begin{cases} 1, & if \ the \ i\text{-}th \ available \ instance \ x_i^{(v)} \\ & of \ the \ v\text{-}th \ view \ belongs \ to \ the \ j\text{-}th \ sample \\ 0, & otherwise \end{cases}$$
$$(5)$$

With the definition of the transformation matrix $G^{(v)}$ as in Eq. (5), it is easy to prove that $G^{(v)} S G^{(v)T}$ can be regarded as a sub-graph of $S$ aligned with $\tilde{Z}^{(v)}$. In other words, the elements in $G^{(v)} S G^{(v)T}$ represent the similarity degrees of those available instances in the $v$-th view as in $\tilde{Z}^{(v)}$. In such a way, only the certain similarity information of the available instances in each view will be exploited to learn the consensus graph and the uncertain information associated with the missing instances in all views are excluded.

**Motivation 2**: Generally speaking, in most cases, any sample and its neighbor samples have a high probability of belonging to the same category. However, for the incomplete multi-view data, owing to the uncontrollable of missing views and noises, the preconstructed graph $\tilde{Z}^{(v)}$ maybe cannot reflect the real nearest neighbor relationships among samples. In this case, it is impossible to learn a high-quality graph from the noisy graphs $\left\{ \tilde{Z}^{(v)} \right\}_{v=1}^{l}$. To solve this problem, many methods seek to integrate the adaptive nearest neighbor graph learning model into the consensus graph learning framework. However, this approach will greatly increase the computational cost, complexity of model optimization, and model convergence epochs. **In our work, we**
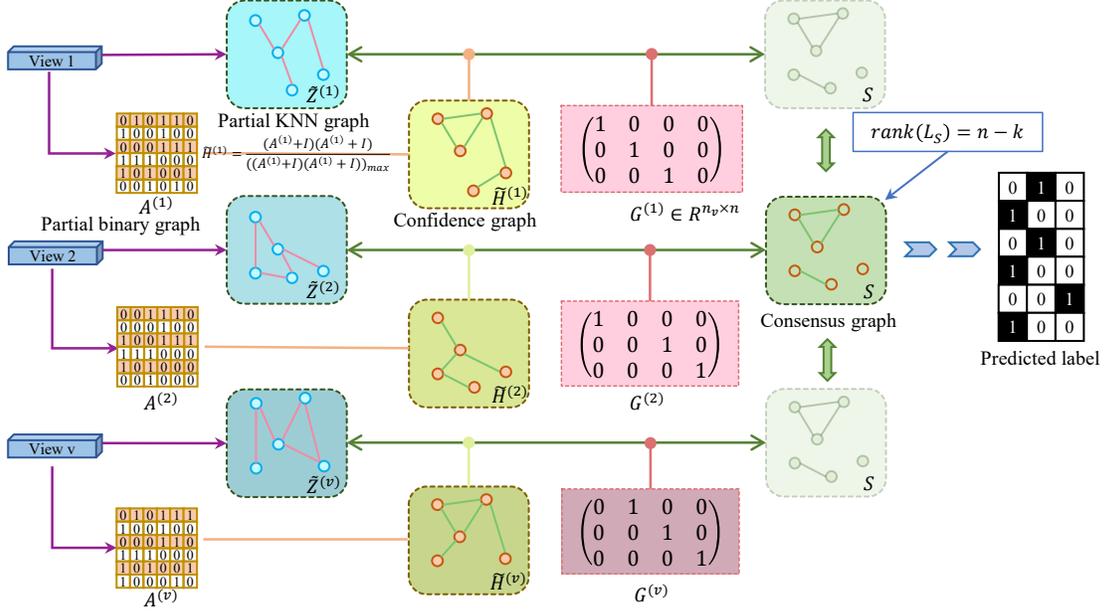
Figure 2. The schematic structure of the proposed HCLS_CGL.

**propose a simple but novel approach to solve this problem based on an assumption**: If samples $x_i$ and $x_j$ are intrinsic nearest neighbor to each other, then their nearest neighbor sets should have some common samples. In other words, as shown in Figure 1, if the more common nearest neighbors of samples $x_i$ and $x_j$ have, the higher probability of the two samples belong to the same class. Moreover, it is interesting to observe that such common nearest neighbors can be counted by $H_{i,j} = A_{i,:}A_{:,j}$ for samples $x_i$ and $x_j$, if $A$ is a special k-nearest neighbor graph with binary values constructed by the similar approach as in (4) with the diagonal elements as 1. By dividing the maximum element $H_{\max}$ of $H$, *i.e.*, $\tilde{H} = H/H_{\max}$, we obtain the confidence-nearest-neighbor graph $\tilde{H}$ whose element represents the confidence/probability of the corresponding two samples to be nearest neighbors to each other. Then, by constructing these confidence-nearest-neighbor graphs $\tilde{H}^{(v)} \in \mathbb{R}^{n_v \times n_v}$ from all incomplete views, we introduce them to guide the consensus graph learning model as follows:

$$\min_{S,\alpha} \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( G^{(v)} S G^{(v)T} - \tilde{Z}^{(v)} \right) \odot \tilde{H}^{(v)} \right\|_F^2 \right) + \lambda_1 \|S\|_F^2$$

$$s.t. 0 \le \alpha_v \le 1, \sum_{v=1}^{l} \alpha_v = 1, S^T I = I,$$

$$0 \le S \le 1, diag\,(S) = 0, rank\,(L_S) = n - c \tag{6}$$

where $\odot$ is the element-wise multiplication operation.

Indeed, since the number information of common nearest neighbors among samples can be regarded as a confidence degree to measure the nearest neighbor relation-

ships among samples, which provides another kind of structure information. Compared with the conventional nearest neighbor graphs $\tilde{Z}^{(v)}$, introducing these confidence structure information to the consensus graph learning model has the potential to obtain a better intrinsic consensus graph for clustering. The schematic structure and main ideas of our proposed HCLS_CGL is visually displayed in Figure 2.

### 3.2. Optimization

For (6), it is easy to prove that this problem is equivalent to the following optimization problem:

$$\min_{S,\alpha} \lambda_1 \|S\|_F^2$$
$$+ \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( S - G^{(v)T} \tilde{Z}^{(v)} G^{(v)} \right) \odot (G^{(v)T} \tilde{H}^{(v)} G^{(v)}) \right\|_F^2 \right)$$
$$s.t. 0 \le \alpha_v \le 1, \sum_{v=1}^{l} \alpha_v = 1, S^T I = I,$$
$$0 \le S \le 1, diag\,(S) = 0, rank\,(L_S) = n - c \tag{7}$$

In addition, in (7), the rank constraint $rank\,(L_S) = n - c$ can be transformed into the minimum optimization problem of $\min_{F^T F = I} \mathrm{Tr}\left( F^T L_S F \right)$ [25], then we can obtain the following equivalent optimization problem to (7) with the definition of $Z^{(v)} = G^{(v)T} \tilde{Z}^{(v)} G^{(v)}$ and $H^{(v)} = G^{(v)T} \tilde{H}^{(v)} G^{(v)}$:

$$\min_{S,\alpha,F} \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2 \right)$$
$$+ \lambda_1 \|S\|_F^2 + \lambda_2 Tr\left( F^T L_S F \right)$$
$$s.t. S^T I = I, 0 \le S \le 1, diag\,(S) = 0, \tag{8}$$
$$0 \le \alpha_v \le 1, \sum_{v=1}^{l} \alpha_v = 1, F^T F = I$$

Problem (8) has three variables to calculate. As it is difficult to calculate the analytical solutions to these variables, we adopt the alternating iterative optimization algorithm [46] to obtain the optimal approximate solutions by regarding problem (8) as three simple sub-problems with respect to these three variables as follows:

**Step 1**: Calculate variable $S$. According to the alternating iterative optimization algorithm, we can first fix variables $F$ and $\alpha$, and then obtain the following simplified optimization problem with respect to variable $S$:

$$\min_{S} \left( \sum_{v=1}^{l} \alpha_v^r \left( \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2 \right) + \lambda_1 \|S\|_F^2 + \lambda_2 Tr \left( F^T L_S F \right) \right) \quad (9)$$
$$s.t. S^T 1 = 1, 0 \le S \le 1, diag(S) = 0$$

In (9), we can simply infer that $Tr\left(F^T L_S F\right) = \frac{1}{2} \sum_{i,j=1}^{n} \|F_{i,:} - F_{j,:}\|_2^2 S_{i,j}$ when $S$ is a positive matrix. Defining $E_{i,j} = \|F_{i,:} - F_{j,:}\|_2^2$, then we can transform (9) as follows:

$$\min_{S} \left( \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2 \right) + \lambda_1 \|S\|_F^2 + \lambda_2 Tr \left( F^T L_S F \right) \right)$$

$$\Leftrightarrow \min_{S} \left( \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2 \right) + \lambda_1 \|S\|_F^2 + \frac{\lambda_2}{2} \sum_{i,j=1}^{n} \|F_{i,:} - F_{j,:}\|_2^2 S_{i,j} \right) \quad (10)$$

$$\Leftrightarrow \min_{S} \sum_{i,j=1}^{n} \left( S_{i,j} - \frac{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} Z_{i,j} - \frac{\lambda_2}{4} E_{i,j}}{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} + \lambda_1} \right)^2$$

It is obvious that (10) is equivalent to the following $n$ independent sub-optimization problems with respect to each column of matrix $S$:

$$\min_{S_{:,j}} \sum_{i=1}^{n} \left( S_{i,j} - \frac{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} Z_{i,j} - \frac{\lambda_2}{4} E_{i,j}}{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} + \lambda_1} \right)^2 \quad (11)$$
$$s.t. \sum_{i=1}^{n} S_{i,j} = 1, 0 \le S_{i,j} \le 1, S_{j,j} = 0$$

According to the Lagrangian algorithm, we can obtain the following closed form solution for (11):

$$S_{i,j} = \begin{cases} (T_{i,j} + \eta_j)_+, & i \ne j \\ 0, & i = j \end{cases} \quad (12)$$

where $T_{i,j} = \frac{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} Z_{i,j} - \frac{\lambda_2}{4} E_{i,j}}{\sum_{v=1}^{l} \alpha_v^r H_{i,j}^{(v)2} + \lambda_1}$, function $(a)_+$ sets negative $a$ to zero and preserves the non-negative $a$. According to the boundary constraint $\sum_{i=1}^{n} S_{i,j} = 1, 0 \le S_{i,j} \le 1, S_{j,j} = 0$, we can obtain $\eta_j = \frac{1 - \sum_{i=1,i\ne j}^{n} T_{i,j}}{n-1}$.

**Step 2**: Calculate variable $F$. Fixing variables $S$ and $\alpha$ in problem (8), we can obtain the following optimization problem with respect to variable $F$:

$$\min_{F^T F = I} Tr\left(F^T L_S F\right) \quad (13)$$

Problem (13) can be optimized via Eigenvalue decomposition. If $\delta_1 \le \delta_2 \le \delta_3 \ldots \le \delta_n$ are the $n$ Eigenvalues of the Laplacian matrix $L_S$ and the corresponding Eigenvectors are represented as $u_1 \le u_2 \le u_3 \ldots \le u_n$, then the optimal solution $F$ to problem (13) can be constructed as $F = [u_1, u_2, \ldots, u_c] \in \mathbb{R}^{n \times c}$ with $c$ Eigenvectors corresponding to the $c$ minimum Eigenvalues.

**Step 3**: Calculate $\alpha$. Similar to the previous two steps, we can fix variables $S$ and $F$, then obtain the following optimization problem with respect to variable $\alpha$:

$$\min_{\alpha} \sum_{v=1}^{l} \left( \alpha_v^r \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2 \right) \quad (14)$$
$$s.t. 0 \le \alpha_v \le 1, \sum_{v=1}^{l} \alpha_v = 1$$

Defining $p_v = \left\| \left( S - Z^{(v)} \right) \odot H^{(v)} \right\|_F^2$, the Lagrange function of (14) can be formulated as follows [46]:

$$\min_{\alpha,\mu} \sum_{v=1}^{l} \alpha_v^r p_v - \mu \left( \sum_{v=1}^{l} \alpha_v - 1 \right) \quad (15)$$

Setting the derivative of problem (15) with respect to variable $\alpha_v$ to zero, we can obtain:

$$\alpha_v = \frac{\mu^{\frac{1}{r-1}}}{(r p_v)^{\frac{1}{r-1}}} \quad (16)$$

Then according to constraint $\sum_{v=1}^{l} \alpha_v = 1$, we can further obtain the optimal solution to variable $\alpha_v$ as follows:

$$\alpha_v = \frac{p_v^{\frac{1}{1-r}}}{\sum_{v=1}^{l} p_v^{\frac{1}{1-r}}} \quad (17)$$

The complete optimization algorithm is summarized in Algorithm 1. In addition, for our method, inspired by the popular spectral clustering algorithm Ratio cut [11], we can directly perform kmeans on the consensus representation $F$ derived by the consensus graph $S$ in (13) to obtain the final clustering result.

**Algorithm 1** HCLS_CGL (solving (6))

---

**Input**: Incomplete multi-view data $\left\{Y^{(v)} \in R^{m_v \times n}\right\}_{v=1}^{l}$ whose missing views are set as 'NaN'; Index matrix $O \in \{0,1\}^{n \times l}$, which indicates the location of missing views; Parameters $\lambda_1$ and $\lambda_2$.

**Initialization**: Construct the k-nearest neighbor graph $\tilde{Z}^{(v)}$ from the observed instance set of each view via Eq. (4) and k-nearest neighbor binary graph $A^{(v)}$, respectively. Then calculate the neighbor confidence weight matrix $\tilde{H}^{(v)}$ with probability property via $\tilde{H}^{(v)} = \frac{(A^{(v)}+I)(A^{(v)}+I)}{((A^{(v)}+I)(A^{(v)}+I))_{\max}}$

    **while** not converged **do**

        Update the consensus graph $S$ according to (12);

        Update matrix $F$ by solving (13);

        Update variable $\alpha$ via (17);

    **end while**

**Output**: $S$ and $F$.

---

### 3.3. Computational complexity analysis

From the previous section, the objective optimization problem (8) is decomposed into three simple sub-problems. For Step 1, we can find that the updating formula for the corresponding variable only takes some element-wise multiplication and division operations on matrices and vectors. In Step 2, the corresponding optimal solution is calculated by Eigenvalue decomposition, which generally takes $O(n^3)$ for an $n \times n$ matrix. Fortunately, we do not need to calculate all eigenvectors and eigenvalues of the Laplacian matrix $L_S$ in (13) but just to calculate the first $c$ minimum eigenvalues and corresponding eigenvectors to construct the optimal solution $F$ [40]. For this goal, an efficient eigenvalue decomposition algorithm with package 'eigs' proposed in [40] is a good choice, which only takes $O(cn^2)$ to obtain the optimal $F$ for problem (13). For Step 3, its updating formula also only includes some element-wise division operations for vectors. By ignoring the efficient element-wise based operations for matrix and vector, we can obtain that the total computational complexity of the above alternating iterative optimization approach is approximately $O(tcn^2)$, where $t$ denotes the iteration number.

## 4. Experiments

In this section, we first introduce our experimental settings in detail, including datasets and competitors, and then give extensive experimental results and analysis. Please refer to the Supplementary materials for further experimental results, including **parameter** and **convergence** analysis.

### 4.1. Experimental settings

**Dataset description.** The experiments are carried out on the following five datasets. **BBCSport** [10] is a docu-

ment database with 737 samples covering 5 sports, which is from the BBCSport website. Four parts of it are regarded as 4 views. **3 Sources** [10], a well-known multi-view text database, contains 169 stories simultaneously reported by BBC, Reuters and The Guardian. **Caltech7** [1] is a subset of the famous Caltech101 [6] dataset. It contains 1474 images with seven categories, and six kinds of feature extracted from each image are used in the experiments. **Animal** [6, 45] is a big dataset with 10158 images across 50 classes. Two types of deep features are selected to denote each sample. **Aloi_deep** is our new dataset extracted from Aloi database [9]. Three classical neural networks, *i.e.*, ResNet50 [12], Vgg16 [29], and Inception-v3 [30] with pretrained weights on the ImageNet [26] are used to extract the deep features as three views from total 10800 images.

**Incomplete multi-view data construction.** In the experiment, we adopt two approaches to construct incomplete data for evaluation. Following [20, 39], for Animal dataset with only two views, we select $p\%$ of all samples as the paired samples with complete views and randomly delete the first view for half of the remaining samples and the second view for the other half of the remainder. For the other datasets, we randomly remove $p\%$ of the instances of each view but keep at least one view available for each sample.

**Competitors.** In the experiments, our HCLS_CGL is compared with nine state-of-the-art methods: BSV [47], Concat [47], MIC [28], DAIMC [13], OMVC [27], OPIMC [14], MKKM-IK-MKC [22], PIC [34], and UEAF [39]. We briefly introduce the methods not mentioned in the previous sections: **BSV** is a baseline method, which fills missing instances with view average and conduct k-means on each view to report the best results. **Concat** is another baseline method that concatenates all views into one before clustering. **DAIMC** is a weighted NMF model by aligning basis matrices and missing views. **OPIMC** improves OMVC and gets clustering results directly without K-means. **MKKM-IK-MKC** recovers incomplete kernel and performs multiple kernel k-means algorithm. **PIC** learns a consistency Laplacian matrix and performs spectral clustering to obtain final results. All competitors will be used with their recommended parameters or by performing a parameter search for better performance.

**Evaluation metrics.** Following previous works [13, 20, 39], clustering accuracy (ACC), normalized mutual information (NMI), and purity are adopted as metrics to evaluate the performance of the aforementioned methods.

### 4.2. Experimental results and analysis

The results of the ten methods on five datasets with different view-missing rates are presented in Tab. 1 to Tab. 3, from which we can draw the following observations:

1) Graph-based methods have clear advantages compared to other methods, for example, MKKM-IK-MKC,

| Dataset | Method\Rate | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 10% | 30% | 50% | 10% | 30% | 50% | 10% | 30% | 50% |
| BBCSport | BSV | 58.62±3.94 | 51.31±5.33 | 44.03±3.78 | 43.73±7.43 | 31.03±2.08 | 21.40±2.61 | 65.79±5.52 | 55.07±1.51 | 47.59±2.28 |
| | Concat | 70.62±3.76 | 58.72±5.42 | 33.21±2.19 | 61.69±6.72 | 38.92±7.87 | 18.61±1.44 | 80.59±4.59 | 63.24±5.82 | 37.00±1.54 |
| | MIC | 51.21±4.21 | 46.21±4.71 | 46.03±5.19 | 29.90±6.25 | 25.84±3.24 | 24.01±5.39 | 55.00±4.15 | 51.72±4.27 | 52.41±6.23 |
| | DAIMC | 68.62±4.59 | 63.45±10.97 | 56.89±5.59 | 56.62±4.60 | 50.17±9.91 | 37.89±6.22 | 76.90±5.89 | 71.72±10.76 | 61.03±5.08 |
| | OMVC | 53.33±3.21 | 51.38±3.06 | 48.79±3.10 | 30.64±2.00 | 41.57±2.79 | 40.63±2.45 | 56.49±2.81 | 59.20±2.12 | 57.47±2.80 |
| | OPIMC | 54.14±4.78 | 52.93±4.93 | 45.69±6.00 | 35.66±4.71 | 31.56±6.10 | 21.75±6.44 | 58.28±4.82 | 56.72±5.76 | 50.86±6.87 |
| | MKKM-IK-MKC | 77.55±2.01 | 75.66±3.01 | 67.07±3.51 | 72.91±3.29 | 64.42±4.69 | 53.52±4.74 | 88.76±2.01 | 84.03±3.22 | 77.00±3.55 |
| | PIC | 75.52±1.57 | 74.48±3.32 | 69.48±6.02 | 70.94±2.22 | 64.18±2.74 | 53.91±6.22 | 87.41±1.44 | 82.41±2.48 | 77.14±6.04 |
| | UEAF | 78.22±0.94 | 77.24±3.08 | 69.31±3.43 | 70.71±2.59 | 68.25±5.13 | 55.13±7.09 | 87.41±2.24 | 87.07±2.86 | 77.07±3.98 |
| | Ours | 92.41±5.43 | 80.00±4.20 | 76.38±4.90 | 83.71±6.45 | 68.70±4.47 | 61.34±4.47 | 92.41±5.43 | 86.21±1.72 | 82.59±2.82 |
| 3Sources | BSV | 56.90±3.69 | 47.38±3.07 | 39.24±3.08 | 50.07±1.22 | 34.46±4.07 | 22.34±1.91 | 68.14±1.67 | 57.63±1.32 | 48.99±0.63 |
| | Concat | 53.54±3.00 | 46.79±3.99 | 37.68±2.91 | 51.98±1.37 | 37.87±3.66 | 18.32±3.25 | 69.78±1.09 | 58.51±3.18 | 46.48±2.82 |
| | MIC | 49.11±3.60 | 47.69±7.61 | 42.49±8.63 | 37.23±6.13 | 38.62±3.81 | 26.08±7.42 | 57.28±3.36 | 61.30±4.28 | 52.31±4.96 |
| | DAIMC | 56.33±4.23 | 52.43±6.63 | 50.73±3.87 | 52.98±3.65 | 49.07±5.78 | 41.64±2.43 | 68.99±4.26 | 67.21±4.89 | 63.56±3.38 |
| | OMVC | 43.95±7.35 | 41.11±4.31 | 39.53±3.63 | 36.48±10.77 | 28.42±3.41 | 24.34±1.50 | 59.37±8.26 | 48.76±5.44 | 45.44±3.10 |
| | OPIMC | 55.73±2.85 | 54.20±4.48 | 43.08±6.98 | 40.62±2.28 | 38.83±3.86 | 22.69±3.83 | 64.73±1.70 | 64.26±2.03 | 53.61±4.36 |
| | MKKM-IK-MKC | 54.44±0.78 | 49.44±5.98 | 50.39±4.14 | 55.46±2.55 | 48.76±3.48 | 48.17±2.76 | 76.97±1.06 | 70.39±4.38 | 69.59±2.17 |
| | PIC | 71.83±4.59 | 70.29±4.25 | 55.50±4.15 | 67.02±5.77 | 63.66±3.20 | 52.01±3.81 | 80.24±3.39 | 79.05±1.59 | 73.85±2.49 |
| | UEAF | 62.60±2.73 | 55.62±5.69 | 52.78±4.53 | 56.47±4.37 | 52.06±2.33 | 45.19±6.54 | 75.50±3.20 | 71.95±4.16 | 67.69±5.28 |
| | Ours | 78.22±1.06 | 79.88±5.95 | 70.89±4.00 | 72.76±2.04 | 67.78±5.78 | 59.11±6.36 | 84.62±1.78 | 83.67±3.96 | 76.92±3.79 |

Table 1. Results on **BBCSport** and **3Sources** databases with different incomplete rates. The 1st/2nd best results are marked in RED/BLUE.

| Dataset | Method\Rate | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 10% | 30% | 50% | 10% | 30% | 50% | 10% | 30% | 50% |
| Caltech7 | BSV | 43.89±1.37 | 39.06±1.26 | 38.31±1.68 | 39.66±2.23 | 31.63±1.51 | 26.81±1.38 | 84.08±1.23 | 75.25±0.71 | 68.97±0.49 |
| | Concat | 41.25±1.67 | 40.55±1.89 | 38.06±0.88 | 43.48±0.92 | 37.99±2.17 | 30.28±0.66 | 84.91±0.50 | 82.54±1.12 | 77.56±0.98 |
| | MIC | 44.07±4.97 | 38.01±2.12 | 35.80±2.34 | 33.71±2.66 | 27.35±1.69 | 20.44±0.98 | 78.12±1.76 | 73.31±0.72 | 68.26±1.40 |
| | DAIMC | 48.29±6.76 | 47.46±3.42 | 44.89±4.88 | 44.61±3.88 | 38.45±2.88 | 36.28±2.34 | 83.32±1.31 | 76.83±3.23 | 75.50±1.17 |
| | OMVC | 40.88±1.54 | 36.82±1.65 | 33.28±4.40 | 28.13±2.54 | 25.32±1.03 | 18.76±4.22 | 79.21±1.77 | 77.73±1.35 | 74.05±4.74 |
| | OPIMC | 49.24±2.89 | 48.34±4.36 | 44.12±5.85 | 42.98±1.02 | 41.54±2.38 | 35.98±2.77 | 84.89±0.69 | 83.70±1.80 | 80.64±2.06 |
| | MKKM-IK-MKC | 36.54±0.51 | 34.87±1.53 | 36.05±0.45 | 24.09±0.98 | 23.45±0.52 | 22.91±0.67 | 72.98±0.80 | 73.82±0.53 | 72.52±1.55 |
| | PIC | 58.82±2.95 | 58.24±1.20 | 56.50±2.93 | 41.73±3.93 | 44.44±3.12 | 43.51±1.50 | 83.99±0.54 | 83.89±0.53 | 83.64±0.55 |
| | UEAF | 50.82±4.05 | 42.71±0.84 | 36.32±4.22 | 39.44±2.07 | 31.07±1.99 | 24.02±1.37 | 81.49±1.78 | 78.26±2.12 | 76.29±1.93 |
| | Ours | 73.09±4.33 | 70.47±5.65 | 66.46±1.61 | 60.01±2.28 | 57.58±3.08 | 54.64±2.35 | 87.46±0.97 | 86.87±0.63 | 85.89±1.24 |
| Aloi_deep | BSV | 64.14±1.23 | 50.63±2.20 | 37.37±1.34 | 81.29±0.65 | 63.15±0.62 | 45.36±0.59 | 69.89±1.13 | 55.10±1.76 | 40.18±1.07 |
| | Concat | 71.07±2.90 | 59.60±1.26 | 41.39±1.30 | 89.75±1.24 | 77.47±0.46 | 68.63±0.97 | 76.52±2.39 | 64.44±0.88 | 44.99±1.54 |
| | MIC | 43.69±2.30 | 35.54±1.19 | 27.96±1.39 | 72.18±2.29 | 66.16±2.68 | 59.10±2.82 | 44.77±0.31 | 36.30±0.25 | 28.35±0.18 |
| | DAIMC | 84.07±1.27 | 81.99±1.32 | 69.00±2.75 | 95.66±0.38 | 94.78±0.23 | 87.70±1.63 | 87.65±0.85 | 85.64±0.82 | 72.61±0.63 |
| | OMVC | 63.13±1.43 | 51.02±1.45 | 35.18±0.62 | 80.99±1.37 | 69.54±0.87 | 57.91±0.80 | 67.58±1.30 | 55.59±1.36 | 39.37±0.72 |
| | OPIMC | 47.09±1.77 | 35.07±1.99 | 33.97±1.64 | 77.56±1.01 | 69.05±0.79 | 67.62±1.71 | 51.17±0.37 | 36.51±0.29 | 34.73±0.28 |
| | MKKM-IK-MKC | 83.23±1.16 | 83.80±1.86 | 83.56±1.54 | 95.52±0.26 | 95.44±0.43 | 95.03±0.46 | 86..90±1.06 | 87.06±1.56 | 86.58±1.42 |
| | PIC | 97.39±0.26 | 97.44±0.09 | 96.82±0.61 | 99.38±0.01 | 99.34±0.06 | 99.26±0.16 | 97.91±0.97 | 97.91±0.98 | 97.59±0.97 |
| | UEAF | 82.74±1.44 | 75.69±2.00 | 72.11±1.91 | 93.92±0.28 | 87.45±0.51 | 88.87±0.49 | 85.91±0.82 | 78.71±0.59 | 75.85±0.65 |
| | Ours | 97.75±0.08 | 97.60±0.08 | 97.38±0.25 | 99.38±0.01 | 99.35±0.04 | 99.28±0.12 | 97.99±0.01 | 97.97±0.03 | 97.82±0.24 |

Table 2. Results on **Caltech7** and **Aloi_deep** databases with different incomplete rates. The 1st/2nd best results are marked in RED/BLUE.

PIC, UEAF and our HCLS_CGL almost take the top two, which confirms that preserving local structure is beneficial for improving clustering performance.

2) Our HCLS achieves landslide victories on almost all datasets and all missing rates, *e.g.*, on the BBCSport dataset with a 10% missing rate, our method outperforms the second best UEAF by about 14% on the ACC metric.

3) For the vast majority of cases, an obvious rule is that the performance of the models gradually decreases with the increase of view-missing rates.

To compare these ten methods more intuitively, we plot their learned cluster indicator matrices in Fig. 3 using the t-SNE technique [32]. We ignore OPIMC because it learns cluster labels directly. As can be seen from Fig. 3, our HCLS_CGL is able to learn more discriminative representations compared to other competitors.

In addition, we plot the consensus graph learned from the incomplete BBCSport dataset with 10% missing views in Fig. 4. From that, we can observe four clear block structures in the figure, which correspond to the four clusters with many samples in the BBCSport dataset.

## 4.3. Ablation study

In this section, ablation experiments on the BBCSport and 3 Sources datasets are conducted to verify the effectiveness of the two main techniques, *i.e.*, adaptive view weighted learning and confidence neighbor graph regularization, where two degraded models, *i.e.*, our method without the corresponding two constraints, called 'no view weight' and 'no confidence neighbor graph', are compared. The experimental results are plotted in Fig. 5. From these two sub-figures, it is clearly to see that the proposed method performs better than its two degraded models.

| Dataset | Method\Rate | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 30% | 50% | 70% | 30% | 50% | 70% | 30% | 50% | 70% |
| Animal | BSV | 42.05±1.20 | 48.63±1.89 | 56.22±1.20 | 48.16±0.44 | 55.91±0.58 | 63.99±0.38 | 45.20±0.88 | 52.26±1.19 | 60.31±0.78 |
| | Concat | 42.79±0.67 | 49.34±1.39 | 53.99±0.99 | 55.46±0.16 | 59.31±0.38 | 63.88±0.35 | 48.12±0.45 | 53.24±0.88 | 59.26±0.81 |
| | MIC | 43.38±0.63 | 45.88±0.34 | 49.15±0.88 | 52.79±0.77 | 55.69±0.36 | 59.30±0.54 | 49.21±0.78 | 52.31±0.34 | 55.33±0.64 |
| | DAIMC | 50.18±2.18 | 53.87±1.36 | 56.42±1.37 | 55.03±1.03 | 59.36±1.16 | 62.76±0.46 | 54.82±1.57 | 59.51±1.65 | 62.12±1.04 |
| | OMVC | 42.51±0.89 | 43.98±0.77 | 46.39±1.02 | 50.77±0.63 | 53.11±0.83 | 55.38±0.46 | 47.33±0.66 | 50.42±0.91 | 52.97±0.76 |
| | OPIMC | 46.33±2.14 | 53.14±1.38 | 53.88±1.26 | 52.34±0.69 | 58.51±0.46 | 62.04±0.26 | 49.49±1.41 | 56.23±1.20 | 57.91±0.43 |
| | MKKM-IK-MKC | 51.77±0.48 | 57.75±0.38 | 61.18±0.59 | 56.54±0.33 | 61.66±0.22 | 66.28±0.27 | 56.14±0.48 | 62.14±0.41 | 66.40±0.53 |
| | PIC | 55.94±0.78 | 56.84±1.55 | 57.67±1.03 | 62.35±0.46 | 64.37±0.64 | 65.82±0.26 | 63.07±0.44 | 64.75±1.57 | 65.42±0.38 |
| | UEAF | 45.73±12.9 | 51.86±6.48 | 58.19±3.04 | 51.61±12.87 | 58.43±7.53 | 64.92±3.95 | 49.10±0.27 | 55.36±0.36 | 63.02±0.47 |
| | Ours | 58.72±1.87 | 61.74±0.37 | 61.29±1.10 | 63.12±0.59 | 67.88±0.47 | 67.25±0.48 | 64.21±1.33 | 68.52±0.56 | 67.98±0.54 |

Table 3. Results on **Animal** database with different paired sample rates. The 1st/2nd best results are marked in RED/BLUE.



(a) BSV  (b) Concat  (c) MIC  (d) DAIMC  (e) OMVC

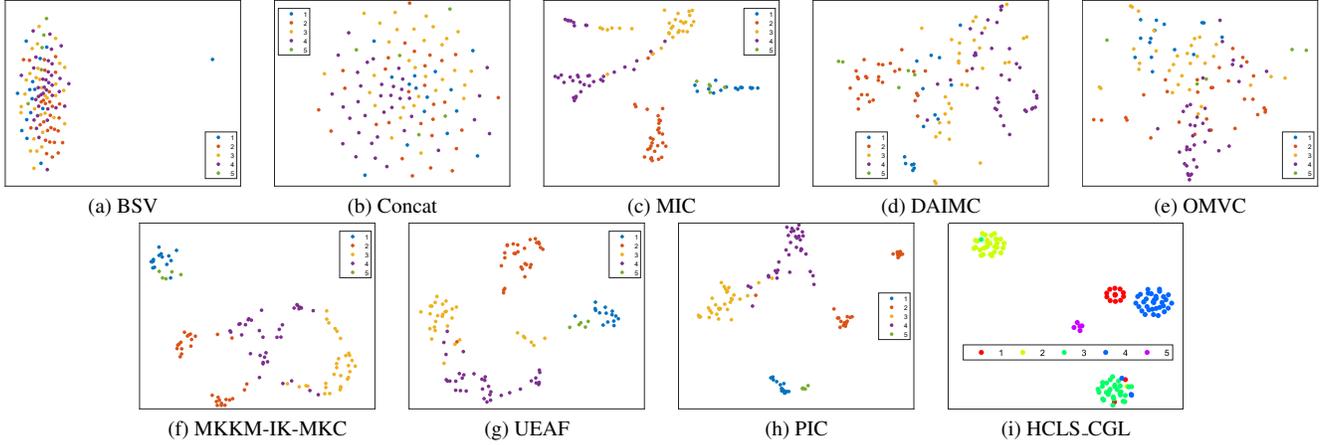(f) MKKM-IK-MKC  (g) UEAF  (h) PIC  (i) HCLS_CGL

Figure 3. T-SNE visualization of the consensus representation of samples obtained by different methods on the BBCSport dataset with 10% missing views, where the visualized consensus representation of our method is variable $F$ in our learning model.
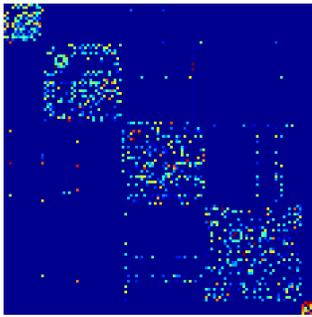


Figure 4. The consensus graph learned by our method from the BBCSport with 10% missing views.
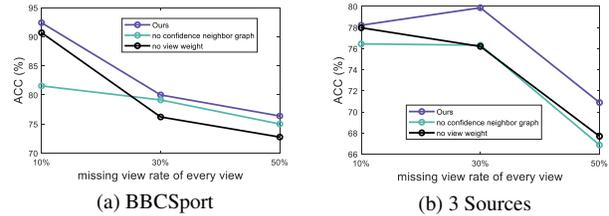


(a) BBCSport  (b) 3 Sources

Figure 5. The clustering ACC (%) obtained by the proposed method with its two degraded methods on the BBCSport and 3 Sources datasets with different missing view rates.

## 5. Conclusion

In this paper, we proposed a novel consensus graph learning based method, called HCLS_CGL, for the challenging IMC task. Different from the existing works which only consider the local pairwise similarity information of data, we provide a new way to explore the group-wise structure information among samples, which is called the confidence neighbor graph in our paper. The confidence neighbor graph constructed in our paper provides the nearest neighbor probability information of any two samples.

Extensive experimental results on five datasets demonstrate that the proposed highly confidence local structure based consensus graph learning method can obtain a better graph yet better performance than many IMC methods.

## Acknowledgments

# References

[1] Xiao Cai, Feiping Nie, and Heng Huang. Multi-view k-means clustering on big data. In *International Joint Conference on Artificial Intelligence*, pages 2598–2604, 2013. 6

[2] Guoqing Chao, Shiliang Sun, and Jinbo Bi. A survey on multiview clustering. *IEEE transactions on artificial intelligence*, 2(2):146–168, 2021. 1

[3] Man-Sheng Chen, Ling Huang, Chang-Dong Wang, Dong Huang, and S Yu Philip. Multiview subspace clustering with grouping effect. *IEEE Transactions on Cybernetics*, 52(8):7655–7668, 2020. 1

[4] Cheng Deng, Zongting Lv, Wei Liu, Junzhou Huang, Dacheng Tao, and Xinbo Gao. Multi-view matrix decomposition: a new scheme for exploring discriminative information. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015. 1

[5] Shijie Deng, Jie Wen, Chengliang Liu, Ke Yan, Gehui Xu, and Yong Xu. Projective incomplete multi-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2023. 1

[6] Li Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. In *Conference on Computer Vision and Pattern Recognition Workshop*, pages 178–178. IEEE, 2004. 6

[7] Jiashi Feng, Zhouchen Lin, Huan Xu, and Shuicheng Yan. Robust subspace segmentation with block-diagonal prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3818–3825, 2014. 2

[8] Hang Gao, Yuxing Peng, and Songlei Jian. Incomplete multi-view clustering. In *International Conference on Intelligent Information Processing*, pages 245–255. Springer, 2016. 2

[9] Jan-Mark Geusebroek, Gertjan J Burghouts, and Arnold WM Smeulders. The amsterdam library of object images. *International Journal of Computer Vision*, 61(1):103–112, 2005. 6

[10] Derek Greene and Pádraig Cunningham. Practical solutions to the problem of diagonal dominance in kernel document clustering. In *International Conference on Machine Learning*, pages 377–384, 2006. 6

[11] Lars Hagen and Andrew B Kahng. New spectral methods for ratio cut partitioning and clustering. *IEEE transactions on computer-aided design of integrated circuits and systems*, 11(9):1074–1085, 1992. 5

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. 6

[13] Menglei Hu and Songcan Chen. Doubly aligned incomplete multi-view clustering. In *International Joint Conference on Artificial Intelligence*, pages 2262–2268, 2018. 6

[14] Menglei Hu and Songcan Chen. One-pass incomplete multi-view clustering. In *AAAI Conference on Artificial Intelligence*, volume 33, pages 3838–3845, 2019. 6

[15] Zhao Kang, Zhiping Lin, Xiaofeng Zhu, and Wenbo Xu. Structured graph learning for scalable subspace clustering: From single view to multiview. *IEEE Transactions on Cybernetics*, 52(9):8976–8986, 2021. 1

[16] Zhao Kang, Chong Peng, and Qiang Cheng. Clustering with adaptive manifold structure learning. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pages 79–82. IEEE, 2017. 2

[17] Shao-Yuan Li, Yuan Jiang, and Zhi-Hua Zhou. Partial multi-view clustering. In *AAAI Conference on Artificial Intelligence*, volume 28, pages 1968–1974, 2014. 1

[18] Zhiping Lin, Zhao Kang, Lizong Zhang, and Ling Tian. Multi-view attributed graph clustering. *IEEE Transactions on knowledge and data engineering*, 2021. 1

[19] Chengliang Liu, Jie Wen, Xiaoling Luo, and Yong Xu. Incomplete multi-view multi-label learning via label-guided masked view-and category-aware transformers. *arXiv preprint arXiv:2303.07180*, 2023. 1

[20] Chengliang Liu, Zhihao Wu, Jie Wen, Yong Xu, and Chao Huang. Localized sparse incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 2022. 2, 6

[21] Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma. Robust recovery of subspace structures by low-rank representation. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):171–184, 2012. 2

[22] Xinwang Liu, Xinzhong Zhu, Miaomiao Li, Lei Wang, En Zhu, Tongliang Liu, Marius Kloft, Dinggang Shen, Jianping Yin, and Wen Gao. Multiple kernel $k$ k-means with incomplete kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(5):1191–1204, 2019. 6

[23] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14, 2001. 2, 3

[24] Feiping Nie, Jing Li, and Xuelong Li. Self-weighted multiview clustering with multiple graphs. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 2564–2570, 2017. 3

[25] Feiping Nie, Xiaoqian Wang, Michael Jordan, and Heng Huang. The constrained laplacian rank algorithm for graph-based clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30, 2016. 2, 3, 4

[26] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015. 6

[27] Weixiang Shao, Lifang He, Chun-ta Lu, and S Yu Philip. Online multi-view clustering with incomplete views. In *IEEE International Conference on Big Data (Big Data)*, pages 1012–1017. IEEE, 2016. 2, 6

[28] Weixiang Shao, Lifang He, and S Yu Philip. Multiple incomplete views clustering via weighted nonnegative matrix factorization with $l_{2,1}$ regularization. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 318–334. Springer, 2015. 2, 6

[29] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 6

[30] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016. 6

[31] Robert Tarjan. Depth-first search and linear graph algorithms. *SIAM journal on computing*, 1(2):146–160, 1972. 3

[32] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 7

[33] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007. 2

[34] Hao Wang, Linlin Zong, Bing Liu, Yan Yang, and Wei Zhou. Spectral perturbation meets incomplete multi-view data. *International Joint Conference on Artificial Intelligence*, pages 3677–3683, 2019. 6

[35] Yang Wang, Xuemin Lin, Lin Wu, Wenjie Zhang, Qing Zhang, and Xiaodi Huang. Robust subspace clustering for multi-view data by exploiting correlation consensus. *IEEE Transactions on Image Processing*, 24(11):3939–3949, 2015. 1

[36] Yang Wang and Lin Wu. Beyond low-rank representations: Orthogonal clustering basis reconstruction with optimized graph structure for multi-view spectral clustering. *Neural Networks*, 103:1–8, 2018. 1

[37] Jie Wen, Yong Xu, and Hong Liu. Incomplete multiview spectral clustering with adaptive graph learning. *IEEE transactions on cybernetics*, 50(4):1418–1429, 2018. 2

[38] Jie Wen, Zheng Zhang, Lunke Fei, Bob Zhang, Yong Xu, Zhao Zhang, and Jinxing Li. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022. 1

[39] Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Hong Liu. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *AAAI Conference on Artificial Intelligence*, volume 33, pages 5393–5400, 2019. 2, 3, 6

[40] Thomas G Wright and Lloyd N Trefethen. Large-scale computation of pseudospectra using arpack and eigs. *SIAM Journal on Scientific Computing*, 23(2):591–605, 2001. 6

[41] Wu Xi, Zhenwen Ren, and F Richard Yu. Parameter-free shifted laplacian reconstruction for multiple kernel clustering. *IEEE/CAA Journal of Automatica Sinica*, 2023. 1

[42] Feng Xia, Ke Sun, Shuo Yu, Abdul Aziz, Liangtian Wan, Shirui Pan, and Huan Liu. Graph learning: A survey. *IEEE Transactions on Artificial Intelligence*, 2(2):109–127, 2021. 1

[43] Muli Yang, Cheng Deng, and Feiping Nie. Adaptive-weighting discriminative regression for multi-view classification. *Pattern Recognition*, 88:236–245, 2019. 1

[44] Jiali You, Zhenwen Ren, Quansen Sun, Yuan Sun, and Xingfeng Li. Approximate shifted laplacian reconstruction for multiple kernel clustering. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 2862–2870, 2022. 1

[45] Changqing Zhang, Zongbo Han, Huazhu Fu, Joey Tianyi Zhou, Qinghua Hu, et al. Cpm-nets: Cross partial multi-view networks. In *Conference on Neural Information Processing Systems*, pages 559–569, 2019. 6

[46] Zheng Zhang, Li Liu, Fumin Shen, Heng Tao Shen, and Ling Shao. Binary multi-view clustering. *IEEE transactions on pattern analysis and machine intelligence*, 41(7):1774–1782, 2018. 5

[47] Handong Zhao, Hongfu Liu, and Yun Fu. Incomplete multi-modal visual data grouping. In *International Joint Conference on Artificial Intelligence*, pages 2392–2398, 2016. 6

[48] Xianbing Zhao, Yinxin Chen, Sicen Liu, and Buzhou Tang. Shared-private memory networks for multimodal sentiment analysis. *IEEE Transactions on Affective Computing*, 2022. 1