Received: 22 June 2022 | Revised: 13 September 2022 | Accepted: 11 October 2022

IET Computer Vision

DOI: 10.1049/cvi2.12152

ORIGINAL RESEARCH

# Self-inferring incomplete multi-view clustering

**Junjun Fan**[1] | **ZeQi Ma**[1] | **Jiajun Wen**[1] | **Zhihui Lai**[1] | **Weicheng Xie**[1] | **Wai Keung Wong**[2]

[1]Shenzhen University, Shenzhen, Guangdong, China

[2]The Hong Kong Polytechnic University, Hong Kong, Hong Kong

**Correspondence**
ZeQi Ma, Shenzhen University, Shenzhen, Guangdong, China.
Email: mzeqi@qq.com

**Abstract**

With the advantage of exploiting complementary and consensus information across multiple views, techniques for Multi-view Clustering have attracted increasing attention in recent years. However, it is common that data on some views is not completed in real-world applications, which brings the challenge of partial mapping between the views. To explore the information hidden in the local geometric structure and recover missing instances through mining the information hidden in existing instances, a self-inferring incomplete multi-view clustering algorithm is proposed. Firstly, the incomplete multi-view data is replenished directly and exploited as variables for inferring the missing instances. And then, a feature graph constraint is united in consensus learning. Besides, a similarity graph learning method is imposed to preserve the local manifold structure. At last, the inferred instances are filled in the missing instances for learning better consensus representation in the iterative process. Extensive experiment results show that this method can improve the clustering performance compared with the state-of-the-art methods.

## 1 | INTRODUCTION

In the big data era, the data generated from different sources or observed from different views are regarded as multi-view data [1, 2]. For example, images shared on the social platform are often published with relevant text description; surveillance videos are usually derived from cameras located in different locations to achieve more information from one place; human action information can be captured by both camera and wearable sensor. Existing works [3–5] concluded that these multi-view data share complementary and consensus information, which is beneficial to clustering and classification.

For better mining of the information hidden in multi-view data, Multi-view Clustering (MvC) has been proposed and is attracting more and more attention. The purpose of MvC is to integrate the information shared among different views and then cluster the data into several clusters [6]. In recent few years, there have been many works proposed to address problems of MvC, such as multi-kernel clustering [7], MvC based on graph [8], and MvC based on subspace learning [9] etc. For MvC, extracting the consensus information across all views is crucial for achieving better results. Thus, in ref. [10], the robust multi-view graph recovery was learned to aggregate a consensus graph to perform better clustering or classification. Besides, based on spectral clustering, De Sa [11] proposed a two independent view-based algorithm. Moreover, in ref. [12], Guo proposed to learn a convex subspace representation across all views and could be used as the input of standard clustering algorithms.

**Funding information:** Guangdong Science and Technology Plan, Grant/Award Numbers: 2020A1515110602, 2020A1515010707; National Natural Science Foundation of China, Grant/Award Number: 61703283; Guangdong Basic and Applied Basic Research Foundation, Grant/Award Numbers: 2021A1515011318, 2017A030310067; Shenzhen Municipal Science and Technology Innovation Council, Grant/Award Number: JCYJ20190808113411274; Shenzhen Visual Object Detection and Recognition Key Laboratory Open Project, Grant/Award Number: HITSZ20220287; Overseas High-Calibre Professional in Shenzhen, Grant/Award Number: 20190629729C; High-Level Professional in Shenzhen, Grant/Award Number: 20190716892H; Research Foundation for Postdoctor Worked in Shenzhen, Grant/Award Numbers: 707-0001300148, 707-0001310414; National Engineering Laboratory for Big Data System Computing Technology; Guangdong Laboratory of Artificial-Intelligence and Cyber-Economics (SZ); Shenzhen Institute of Artificial Intelligence and Robotics for Society; Scientific Research Foundation of Shenzhen University, Grant/Award Numbers: 2019049, 827-000526, 860-000002110328; Laboratory for Artificial Intelligence in Design (Project Code: RP3-3) under the InnoHK Research Clusters, Hong Kong SAR Government

The mentioned MvC algorithms require that every view should be complete in multi-view data. However, in real-world cases, it is common that data on some views is unavailable or is only partially available, which leads to incomplete multi-view data. For instance, it is hard to capture webpages containing texts or images in webpage clustering. It is noteworthy that incomplete multi-view data degrade performance of MvC algorithms. To address this issue, Rai et al. [13] proposed a general approach that allows the MvC to be applicable in incomplete multi-view data, in which only one view was complete and the auxiliary views were incomplete. Specifically, the kernel canonical correlation analysis-based MvC was taken as an example research to illustrate their approach. To handle more independent incomplete views, more and more novel methods have been proposed in recent years, such as partial MvC [14], multi-incomplete-view clustering (MivC) [15], online multi-view clustering algorithm (OMvC) [16], and doubly aligned incomplete multi-view clustering (DAIMC) [17]. However, these methods either failed to sufficiently explore the information hidden in the local geometric structure or were unable to recover missing instances through mining the information hidden in existing instances. Specially, Wen et al. proposed to infer missing instances by using an error matrix, combined with the incomplete view in a unified framework to learn a common latent representation [18]. Nevertheless, it also ignores the correlation between existing instances and missing instances.

To address the above-mentioned problems of existing methods, a novel algorithm is proposed, named self-inferring multi-view clustering (SIIMvC), to learn a common representation for all views while inferring the missing instances at the same time. Firstly, the missing instances of incomplete multi-view data are replenished with zeros. And then, the replenished multi-view data are exploited as variables and a feature graph is constructed for each view. Based on feature graphs, the missing instances can be inferred, so the consensus representation can be generated simultaneously. An improved consensus learning model and an improved similarity graph learning model are presented to better infer the missing instances and generate the consensus representation in the learning process.

The contributions of this paper are summarised as followings.

(1) A novel inferring method for missing instances in incomplete MvC is presented that the incomplete multi-view data are replenished with zeros and exploited as variables for inferring. Furthermore, a feature graph constraint is united in the consensus learning to ensure consistency between the inferred instances and the existing instances. With this method, the inferred instances can be used directly with the existing instance for consensus representation learning with alignment.
(2) An improved similarity graph learning model is introduced to preserve the local manifold structure, which benefits the consistence of the common representation. Furthermore, it is noted that the proposed similarity graph learning model has fewer parameters to be calculated compared to the most relevant state-of-the-art method of Unified Embedding Alignment Framework (UEAF).
(3) All terms of the proposed model are unified in an Incomplete Multi-View Clustering (IMC) framework, which can provide better performance for clustering. Extensive experiments show that the proposed SIIMvC outperforms the existing incomplete MvC methods.

The remaining content of this article is organised as follows. Section 2 introduces the related work briefly. Section 3 details the proposed method and the updating steps. In Section 4, extensive experiments and results are presented to verify that our method is effective. Section 5 provides conclusions of our article.

## 2 | RELATED WORK

In this section, some representative incomplete MvC methods are introduced and a unified framework proposed for combining incomplete view data and inferring view data is emphasised.

### 2.1 | Incomplete multi-view clusting methods

To solve the incomplete MvC problem, Rai et al. [13] first proposed an approach based on the kernel canonical correlation analysis to exploit incomplete multi-view data for clustering, in which only one view was complete, and the auxiliary views were incomplete. To deal with more independent incomplete views, the matrix factorisation technique was popular in incomplete MvC. Partial MvC [14] focuses on learning a common latent subspace for all views, in which instances belonging to the same example have the same representation. However, Partial MvC can only handle two independent views. MivC [15] extended matrix factorisation to more than two views through weighted non-negative matrix factorisation with $L_{2,1}$ regularisation. It is noted that MivC fills missing instances with average values of existing instances. In order to deal with large-scale incomplete views, Shao et al. [16] proposed OMvC seeks to reduce the influence of incomplete instances by using a dynamic weight setting. In DAIMC [17], the weighted matrix factorisation and a regression constraint were introduced to capture more information.

In the past few years, many novel works have been presented to better address the incomplete MvC problem by various methods. The Deep incomplete multi-view clustering network (DIMC-net) method integrates the deep multi-view encoding, view-missing information based weighted fusion, graph embedding, and Kullback-Leibler divergence based clustering into a joint deep network [19]. Different from the conventional deep learning based IMC methods, it only exploits the instances with complete views for network pre-training.

DIMC-net designs a more flexible graph embedded incomplete multi-view Autoencoder that can effectively exploit information of all incomplete views for pre-training. The inCOMPlete muLti-view clustEring via conTrastivE pRediction method learn the informative and consistent representation by maximising the mutual information across different views through contrastive learning, and recover the missing views by minimising the conditional entropy of different views through dual prediction [20]. The adaptive graph completion-based incomplete multi-view clustering (AGC_IMC) method develops a joint framework for graph completion and consensus representation learning, which mainly contains three components, that is, within-view preservation, between-view inferring, and consensus representation learning [21]. Furthermore, to reduce the negative influence of information imbalance, AGC_IMC introduces some adaptive weights to balance the importance of different views during the consensus representation learning. Aiming to deal with various incomplete clustering situations and be applied in large-scale datasets as well, Yu et al. [22] use an auto-weighted Sample-level Fusion with Anchors for Incomplete Multi-view Clustering, which can not only handle incomplete samples but also effectively explore the relationship between each instance and anchors. What's more, the robuSt mUlti-view clusteRing with incomplEte information method uses the available pairs as positives and randomly chooses some cross-view samples as negatives, providing unified solution to simultaneously handles Partially View-unaligned Problem and Partially Sample-missing Problem [23].

Unlike the previous works, this paper focuses on combining the common representation learning and missing view inferring into a unified framework for resolving incomplete MvC problem, which can improve the classification performance and rely on fewer parameters.

## 2.2 | The UEAF method

In this subsection, we emphatically introduce the UEAF, which provides a framework for robust incomplete MvC [18]. UEAF creates an error matrix for the missing views recovering, which enables the alignment of the incomplete views for the consensus representation learning and to exploit the hidden information of the missing views. Moreover, a reverse graph regularisation term is developed to learn the common latent representation. The overall model can be written as follows:

$$\min_{\Upsilon} \sum_{v=1}^{v_n} \left(\alpha^{(v)}\right)^r \begin{pmatrix} \|X^{(v)} + E^{(v)} W^{(v)} - U^{(v)} P\|_F^2 \\ + \lambda_1 Tr\left(E^{(v)T} L^{(v)} E^{(v)}\right) \\ + \lambda_2 Tr\left(P L_{S^2} P^T\right) \end{pmatrix}$$
$$\text{s.t. } U^{(v)T} U^{(v)} = I, \sum_{v=1}^{v_n} \alpha^{(v)} = 1, \alpha^{(v)} \geq 0, \forall i, S_{i,:}\mathbf{1} = 1,$$
$$0 \leq S_{i,j} \leq 1, S_{i,i} = 0, rank(L_S) = n - c. \tag{1}$$

where $\Upsilon = \{E^{(v)}, U^{(v)}, P, S, \alpha^{(v)}\}$, $\lambda_1$, $\lambda_2$ are penalty parameters. $X^{(v)} \in R^{m_v \times n}$ includes the existing instances and the missing instances, where the missing instances are filled with zeros. $m_v$ denotes the feature dimensionality of $v$th view, $v_n$ denotes the number of views, and $n$ is the total number of instances. $P \in R^{c \times n}$ is the common representation of all the views and $U^{(v)} \in R^{m_v \times c}$ is the basis matrix for each view. $c$ is equal to the number of objective cluster numbers. $E^{(v)} \in R^{m_v \times n_v^m}$ denotes the error matrix, which is used to infer the missing instances of the $v$th view, $n_v^m$ denotes the number of missing instances of the $v$th view. $L^{(v)}$ is the Laplacian matrix of graph $G^{(v)}$. $G^{(v)}$ is defined as follows:

$$G_{i,j}^{(v)} = \begin{cases} 1, & \overline{X}_{i,:}^{(v)} \in \Psi\left(\overline{X}_{j,:}^{(v)}\right) \; or \; \overline{X}_{j,:}^{(v)} \in \Psi\left(\overline{X}_{i,:}^{(v)}\right) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $\Psi\left(\overline{X}_{j,:}^{(v)}\right)$ denotes the set of $k$ nearest neighbours of the $j$th feature. $\overline{X}$ is the set of the available instances of the $v$th view. $W^{(v)} \in R^{n_v^m \times n}$ is defined as follows:

$$W_{i,j}^{(v)} = \begin{cases} 1, & \text{if jth instance is the ith missing instance in the vth view,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

$S \in R^{n \times n}$ denotes the nearest neigbour graph with each element representing the similarity probability corresponding to another instance. $L_S$ represents the Laplacian matrix of similarity graph $S$. $L_{S^2}$ represents the Laplacian matrix of graph $(S \odot S)$ and $\odot$ is the Hadamard product. $\alpha^{(v)}$ is a positive weight to balance the significance of the $v$th view.

The UEAF use an error matrix to infer the missing instances and combine them with the incomplete view in a unified framework to learn a common latent representation. However, view alignment is necessary, and the correlation between existing instances and missing instances is not fully explored. Inspired by the UEAF, we proposed a SIIMvC method to learn the common representation for all views while inferring the missing instances at the same time, in which there is no need for view alignment, and the correlation between existing instances and missing instances can be learned in the missing instances inferring process.

# 3 | PROPOSED METHOD

To better mine the latent information hidden in the existing instances for incomplete MvC, we propose SIIMvC, a novel method to replenish the incomplete multi-view data with zeros and employ them as variables for inferring the missing instances through the existing instances. Furthermore, an improved consensus learning model and a similarly graph learning model are proposed in the inferring framework to learn the consensus representation of the incomplete multi-view data. A computational complexity analysis of the proposed SIIMvC is also presented at the end of this section.

## 3.1 | Consensus learning

Clustering results are usually driven directly from the common consensus representation. Therefore, consensus representation learning is a significant method in incomplete MvC. However, most of the existing works [14–17] use a simple strategy of filling the missing instances with the average of existing instances, and thus ignore the relation between existing instances and the missing instances. Although the UEAF [18] algorithm uses an error matrix to infer missing instances, the relation between existing instances and missing instances is still outside their consideration. In this work, we propose an improved model that can simultaneously infer the missing instances and learn the consensus representation as following.

$$\min_{Z^{(v)},U^{(v)},V} \sum_{v=1}^{v_n} \left( \begin{array}{l} \|Z^{(v)} - U^{(v)}V\|_F^2 \\ + \lambda_1 \sum_{i,j}^{m_v} \|Z_{i,:}^{(v)} - Z_{j,:}^{(v)}\|_2^2 Q_{i,j}^{(v)} \end{array} \right) \quad (4)$$
$$\text{s.t. } U^{(v)T}U^{(v)} = I$$

where $\lambda_1$ is a positive penalty parameter, $Z^{(v)} \in R^{m_v \times n}$ represents instances from $v$th view, including the existing instances and the missing instances. It is noted that we initiate the missing instances with zero. $m_v$ is the feature dimensionality of the $v$th view, $n$ represents the number of the instances, $v_n$ is the number of views. $Z_{i,:}^{(v)}$ and $Z_{j,:}^{(v)}$ represents the $i$th row and the $j$th row vector of matrix $Z^{(v)}$ respectively. $U^{(v)} \in R^{m_v \times c}$ denotes the basis matrix of the $v$th view. $V \in R^{c \times n}$ is the consensus representation of all views, in which c is the cluster number of the latent representation. $I$ represents the identity matrix. The orthogonal constraint $U^{(v)T}U^{(v)} = I$ is used to avoid the trivial solution.

In model (4), $Q^{(v)} \in R^{m_v \times m_v}$ is the neighbour graph of feature dimensionality of initial $Z^{(v)}$ from the $v$th view. The definition of $Q^{(v)}$ is proposed as follows:

$$Q_{i,j}^{(v)} = \begin{cases} 1, & \tilde{Z}_{i,:}^{(v)} \in \phi\left(\tilde{Z}_{j,:}^{(v)}\right) \; or \; \tilde{Z}_{j,:}^{(v)} \in \phi\left(\tilde{Z}_{i,:}^{(v)}\right) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where $\phi\left(\tilde{Z}_{i,:}^{(v)}\right)$ indicates the set of k nearest neighbours of the $i$th feature. $\tilde{Z}^{(v)}$ represents the set of existing instances of $Z^{(v)}$ from $v$th view.

The proposed consensus learning model can thus be reformulated as:

$$\min_{Z^{(v)},U^{(v)},V} \sum_{v=1}^{v_n} \left( \begin{array}{l} \|Z^{(v)} - U^{(v)}V\|_F^2 \\ + \lambda_1 \mathrm{Tr}\left(Z^{(v)T} L_{\tilde{Z}^{(v)}} Z^{(v)}\right) \end{array} \right) \quad (6)$$
$$\text{s.t. } U^{(v)T}U^{(v)} = I$$

where $L_{\tilde{Z}^{(v)}}$ is the Laplacian matrix of graph $Q^{(v)}$, and it is calculated as $L_{\tilde{Z}^{(v)}} = D^{(v)} - Q^{(v)}$. $D^{(v)}$ is a diagonal matrix and its $i$th diagonal element can be calculated through $D_{i,j}^{(v)} = \sum_{j=1}^{m_v} Q_{i,j}^{(v)}$. Tr denotes the trace function.

By replenishing the incomplete multi-view data with zeros, all views can be aligned naturally and the missing instances can be inferred from the existing instances. Furthermore, the consensus representation can be learned simultaneously in the inferring process.

## 3.2 | Similarity graph learning

As is well known, the graph-based methods successfully preserve the local manifold structure [24]. For each instance, it is connected by other instances with the probability $s_{ij}$ in the similarity matrix, which means that they belong to the same class. In other words, if two instances are closer, they should be assigned higher probability $s_{ij}$. Therefore, learning a similarity matrix is crucial to the performance of the clustering [18, 24, 25]. However, for incomplete MvC problem, it is unable to exploit the local manifold structure through incomplete data instances from each view. Thus, in order to overcome this difficulty, we propose the following similarity graph learning model:

$$\min_{S} \sum_{v=1}^{v_n} \sqrt{\sum_{i,j}^{n} \|U^{(v)}V_{:,i} - U^{(v)}V_{:,j}\|_2^2 S_{i,j}} + \alpha \sum_{i}^{n} \|S_i\|_2^2$$
$$\text{s.t. } \forall i, S_{i,:}\mathbf{1} = 1, 0 \le S_{i,j} \le 1, S_{i,i} = 0,$$
$$rank(L_S) = n - c. \quad (7)$$

where $S \in R^{n \times n}$ denotes the similarity graph, where each element represents the similarity degree corresponding to another instance. Constraint $rank\ (L_S)$ can ensure that graph $S$ has c components, which is equal to the number of the multiplicity of the eigenvalue 0 of $L_S$ [26, 27]. $L_S = D_S - (S^T + S)/2$ is the Laplacian matrix. $D_S$ is a diagonal matrix and its $i$th diagonal element can be calculated by $\sum_j (S_{i,j} + S_{j,i})/2$. $\alpha$ is an automatically adjusted parameter, which is determined by the number of nearest neighbours [24]. *1* is a vector that all elements are 1.

The proposed similarity graph learning model can preserve the local manifold structure by mining the relationship among multi-view data. As a result, a better consensus representation can be learned through the well-learning similarity graph, where each element has c neighbours. The advantages mentioned above can be further beneficial to recover the missing instances.

## 3.3 | Overall objective function

In this subsection, the above proposed models are incorporated into a unified incomplete multi-view learning framework named SIIMvC. And the overall objective function is written as follows:

$$\min_{Z^{(v)},U^{(v)},V,S} \left( \begin{array}{l} \sum_{v=1}^{v_n} \|Z^{(v)} - U^{(v)}V\|_F^2 \\ + \lambda_1 \mathrm{Tr}\left(Z^{(v)T} L_{\tilde{Z}^{(v)}} Z^{(v)}\right) \\ + \sum_{v=1}^{v_n} \sqrt{\sum_{i,j}^{n} \|U^{(v)}V_{:,i} - U^{(v)}V_{:,j}\|_2^2 S_{i,j}} \\ + \alpha \sum_{i}^{n} \|S_i\|_2^2 \end{array} \right)$$

$$\text{s.t. } U^{(v)T}U^{(v)} = I, \forall i, S_{i,:}\mathbf{1} = 1, 0 \le S_{i,j} \le 1, S_{i,i} = 0. \quad rank(L_S) = n - c. \tag{8}$$

Figure 1 shows the framework of the proposed SIIMvC. Firstly, the incomplete multi-view data are replenished with zeros and considered as variables for inferring missing instances. And then in the inferring process, the missing instances can learned under the constraint of the improved consensus learning model and similarity graph learning model. Based on the learning result, the consensus representation can be obtained for further clustering.

Through the constraint $U^{(v)T}U^{(v)} = I$ and the method in ref. [24], model (8) can be reformulated as:

$$\min_{Z^{(v)},U^{(v)},V,S} \left( \begin{array}{l} \sum_{v=1}^{v_n} \left( \|Z^{(v)} - U^{(v)}V\|_F^2 \right. \\ \left. + \lambda_1 \mathrm{Tr}\left(Z^{(v)T} L_{\tilde{Z}^{(v)}} Z^{(v)}\right)\right) \\ + \frac{1}{2}\gamma \sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j} + \alpha \sum_{i}^{n} \|S_i\|_2^2 \\ + 2\lambda_2 \mathrm{Tr}(F^T L_S F) \end{array} \right)$$

$$\text{s.t. } U^{(v)T}U^{(v)} = I, \forall i, S_{i,:}\mathbf{1} = 1, 0 \le S_{i,j} \le 1, S_{i,i} = 0, \quad F^T F = I. \tag{9}$$

where $\gamma = \frac{1}{\sqrt{\sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j}}}$ and $\lambda_2$ is a penalty parameter. $F$ denotes the cluster indicator matrix, and the last item with the constraint $F^T F = I$ is deduced through Ky Fan's theorem [28].

## 3.4 | Solution to the SIIMvC

In order to find the solution of the function (9), we exploit an alternating iteration procedure in which the objective function is optimised with only one variable while fixing the remaining

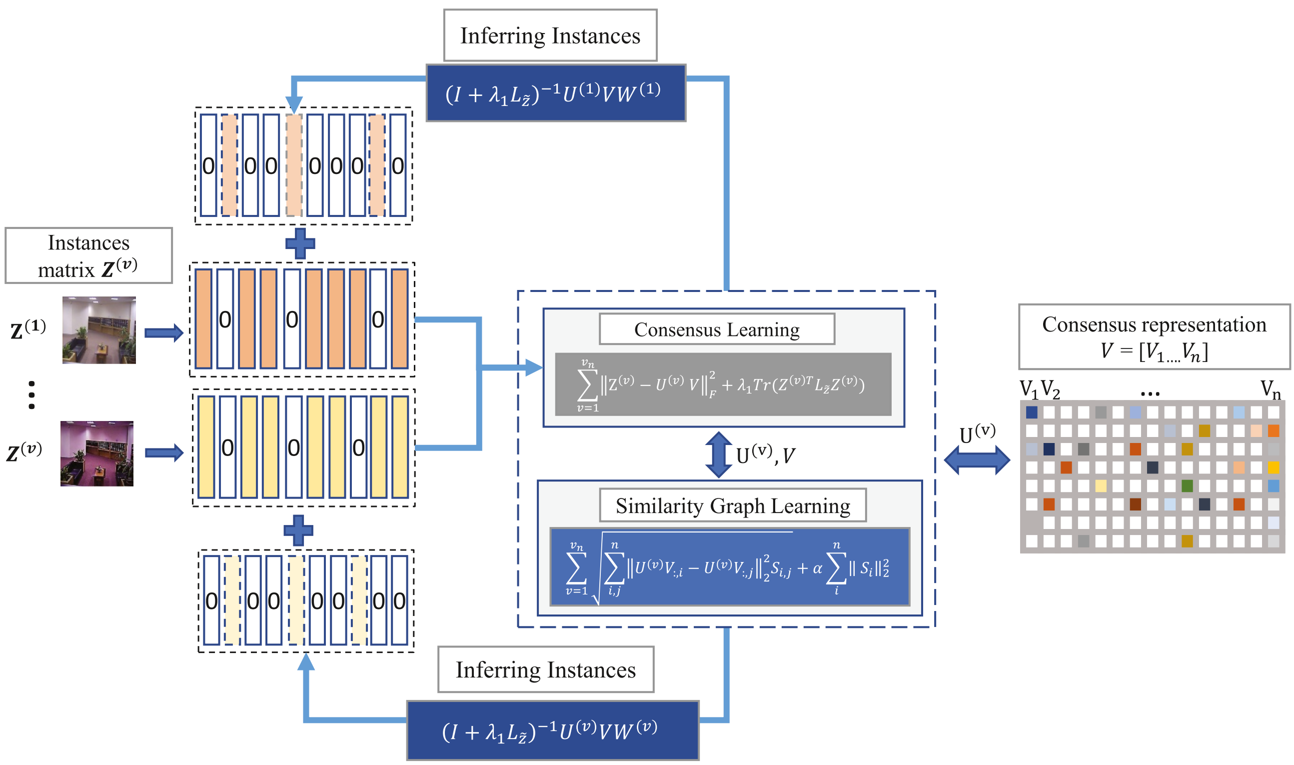



**FIGURE 1** The framework of the proposed SIIMvC. The incomplete multi-view data are replenished with zeros and considered as variables for inferring missing instances. In the inferring process, the improved consensus learning model and similarity graph learning model are incorporated into the proposed SIIMvC framework, and the consensus representation can be obtained for further clustering the learning result. SIIMvC, self-inferring multi-view clustering.

variables. The updated procedure of each variable is described in detail as following.

***Update V***: By fixing $U^{(v)}$, $Z^{(v)}$, $S$, $F$, the optimisation formula becomes:

$$\Gamma(v) = \sum_{v=1}^{v_n} \|Z^{(v)} - U^{(v)}V\|_F^2 + \frac{1}{2}\gamma \sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j}$$
$$= \sum_{v=1}^{v_n} \|Z^{(v)} - U^{(v)}V\|_F^2 + \gamma \mathrm{Tr}\left(VL_S V^T\right) \tag{10}$$

Let $\frac{\partial \Gamma(V)}{\partial V} = 0$, we can get the following equation:

$$V = \left(\sum_{v=1}^{v_n} U^{(v)T} Z^{(v)}\right)(I + \gamma L_S)^{-1} \tag{11}$$

***Update S***: By fixing $U^{(v)}$, $Z^{(v)}$, $V$, $F$, the optimisation formula becomes:

$$\min_S \frac{1}{2}\gamma \sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j} + \alpha \sum_{i}^{n} \|S_i\|_2^2 + 2\lambda_2 \mathrm{Tr}\left(F^T L_S F\right) \tag{12}$$
$$\text{s.t. } \forall i, S_{i,:}\mathbf{1} = 1, 0 \le S_{i,j} \le 1, S_{i,i} = 0, F^T F = I.$$

which can be further deduced as:

$$\min_S \frac{1}{2}\gamma \sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j} + \alpha \sum_{i}^{n} \|S_i\|_2^2 + \lambda_2 \sum_{i,j}^{n} \|F_{i,:} - F_{j,:}\|_2^2 S_{i,j} \tag{13}$$
$$\text{s.t. } \forall i, S_{i,:}\mathbf{1} = 1, 0 \le S_{i,j} \le 1, S_{i,i} = 0, F^T F = I.$$

Then, we denote $P_{i,j} = \frac{1}{2}\gamma \sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 + \lambda_2 \sum_{i,j}^{n} \|F_{i,:} - F_{j,:}\|_2^2$, and further denote $P_i$ as a vector with $j$th element equal to $P_{i,j}$. The problem (13) can be simplified in a vector form as:

$$\min_{\forall i, S_{i,:}\mathbf{1}=1, 0 \le S_{i,j} \le 1, S_{i,i}=0} \sum_{i}^{n} \|S_i + \frac{P_i}{2\alpha}\|_2^2 \tag{14}$$

This problem has a closed-form solution, which can be calculated through an efficient algorithm proposed in ref. [29].

***Update F***: By fixing $U^{(v)}$, $Z^{(v)}$, $V$, $S$, the optimisation formula becomes:

$$\min_F \mathrm{Tr}\left(F^T L_S F\right) \tag{15}$$

which can be calculated by the c eigenvectors corresponding to the first c smallest eigenvalues of $L_S$.

***Update γ***: By fixing $U^{(v)}$, $Z^{(v)}$, $V$, $S$, $F$, $\gamma$ is updated by $\gamma = \frac{1}{\sqrt{\sum_{i,j}^{n} \|V_{:,i} - V_{:,j}\|_2^2 S_{i,j}}}$.

***Update U^(v)***: By fixing $Z^{(v)}$, $S$, $V$, $F$, the optimisation formula becomes:

$$\min_{U^{(v)T}U^{(v)}=I} \|Z^{(v)} - U^{(v)}V\|_F^2 \tag{16}$$

which can be deduced as follows:

$$\min_{U^{(v)T}U^{(v)}=I} \|Z^{(v)} - U^{(v)}V\|_F^2$$
$$\Leftrightarrow$$
$$\min_{U^{(v)T}U^{(v)}=I} -2Tr\left(U^{(v)T} Z^{(v)} V^T\right) \tag{17}$$
$$\Leftrightarrow$$
$$\max_{U^{(v)T}U^{(v)}=I} \mathrm{Tr}\left(U^{(v)T} Z^{(v)} V^T\right)$$

According to ref. [30], we can obtain the optimal solution to Equation (17) as follows:

$$U^{(v)} = A^{(v)} R^{(v)T} \tag{18}$$

where $A^{(v)}$ and $R^{(v)}$ are the left and right singular matrices of the matrix $(Z^{(v)}V^T)$ respectively.

***Update Z^(v)***: By fixing $U^{(v)}$, $S$, $V$, $F$, the optimisation formula becomes:

$$\Gamma\left(Z^{(v)}\right) = \|Z^{(v)} - U^{(v)}V\|_F^2 + \lambda_1 Z^{(v)T} L_{\tilde{Z}^{(v)}} Z^{(v)}$$
$$\text{s.t. } U^{(v)T}U^{(v)} = I \tag{19}$$

which can be simplified as follows:

$$\Gamma\left(Z^{(v)}\right) = \mathrm{Tr}\left(Z^{(v)T} Z^{(v)} - 2Z^{(v)T} U^{(v)} V\right) + \lambda_1 Z^{(v)T} L_{\tilde{Z}^{(v)}} Z^{(v)} \tag{20}$$

Let $\frac{\partial \Gamma\left(Z^{(v)}\right)}{\partial Z^{(v)}} = 0$, we can obtain:

$$Z^{(v)} = \left(I + \lambda_1 L_{\tilde{Z}^{(v)}}\right)^{-1} U^{(v)} V \tag{21}$$

In order to preserve the existing data, we update the equation of $Z^{(v)}$ as follows:

$$Z^{(v)} = \widehat{Z}^{(v)} + Y^{(v)} W^{(v)} \tag{22}$$

where $\widehat{Z}^{(v)} \in R^{m_v \times n}$, which fills the missing instances with zero, is the initialisation of $Z^{(v)}$. $Y^v = \left(I + \lambda_1 L_{\tilde{Z}^{(v)}}\right)^{-1} U^{(v)} V$ is what we obtain in Equation (21). And $W^v \in R^{n \times n}$ is defined as follows:

$$W_{i,i}^{(v)} = \begin{cases} 1, & \text{if ith instance is not in the vth view,} \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

From Equation (22) we can obtain the inferred instances in every updating step. Algorithm 1 summarises the solution steps of the proposed SIIMvC.

**Algorithm 1 SIIMvC (solution to Equation [9])**

**Input:** Incomplete multi-view data matrix $\{X^{(v)} \in R^{m_v \times n}\}_{v=1}^{v_n}$, which fills missing instances with 0. Feature neighbour graph $Q = \{Q^{(v)}\}_{v=1}^{v_n}$ in consensus learning. Parameters $\lambda_1$, $\lambda_2$ and diagonal matrix $W = \{W^{(v)} \in R^{n \times n}\}_{v=1}^{v_n}$.
**Initialisation:** Instances matrix $\{Z^{(v)} \in R^{m_v \times n}\}_{v=1}^{v_n}$ whose values equal to data matrix $X$. Orthogonal matrix $U^{(v)}$ with random values and consensus matrix $V$ is initialised by using $U^{(v)}$. Graph $S$ is initialised through matrix $V$ and $F$ is initialised based on $S$.
**while** not converged **do**
  1. Update $V$ using (11);
  2. Update $S$ using (14);
  3. Update $F$ using (15);
  4. Update $\gamma$ and $\alpha$;
  **for** $v$ from 1 to $v_n$ **do**
    5. Update $U^{(v)}$ using (15);
    6. Update $Z^{(v)}$ using (22);
  **end for**
**end while**
**return** $V$;

# 4 | EXPERIMENT

## 4.1 | Dataset description

In this subsection, we introduce five widely used multi-view datasets, that is, the BUAA-visnir face dataset (BUAA) [31], the Digit dataset [32], the 3Sources dataset [33], the BBCSport dataset [34] and the Berkeley Drosophila Genome Project (BDGP) dataset [35]. They are chosen to evaluate the performance of the proposed algorithm SIIMvC, compared with several popular state-of-the-art methods.

(1) **BUAA Dataset**: Similar to ref. [36], a subset of BUAA dataset that is composed of 90 visual images and 90 near infrared images of the first 10 volunteers is used to evaluate our algorithm in our experiments. Two types of images are considered as two views in our experiments.
(2) **Digit Dataset**: It consists of 2000 hand-written digits (0–9). Six types of features grouped in this dataset, each of which can be considered as a view. The feature dimensions of these views are 240, 76, 216, 47, 64 and 6 respectively.
(3) **3Sources Dataset**: This dataset is derived from three sources in online news. In total, 948 news articles were collected, which cover 416 different news stories. For our experiments, a subset containing 169 articles was randomly selected to evaluate our algorithm.
(4) **BBCSport Dataset**: There are 737 news articles collected from the BBC Sport website in this dataset. We chose the four-view dataset for our experiments to evaluate the proposed algorithm and other compared incomplete MvC-based methods. The subset has 116 instances, and the feature dimensions for four different views are 1991, 2063, 2113 and 2158 respectively.
(5) **BDGP Dataset**: There are 2500 instances in BDGP, which can be divided into five classes. The BDGP dataset has four views, and consists of texture feature and the visual feature extracted from lateral, dorsal and ventral images.

## 4.2 | Compared methods

In this subsection, we briefly introduce 10 relevant methods that we used to compare with the proposed SIIMvC.

**BSV** [36]: The Best Single View (BSV) firstly fills missing instances with average of existing instances, and then performs k-means on each view.

**Concat** [36]: Like the BSV, the missing instances are filled with the average of existing instances in the Contat. After concatenating all views into one view, k-means is used to perform clustering on the concatenated single view.

**GPMvC** [37]: The GPMvC extends MvC into more than two views and imposes a graph laplacian regularisation term for consensus representation learning.

**OMvC** [16]: The OMvC is presented to deal with large-scale incomplete views. In order to reduce the negative effects of incompleteness, the OMvC introduces dynamic weight setting to give missing instances low weights.

**MivC** [38]: Through incorporating the weighted-NMF and a regularised method, the MivC learns a common subspace across all views.

**MVL_IV** [9]: The multi-view learning with incomplete views (MVL_IV) assumes that different views share a common subspace. To handle large-scale data and obtain fast convergence, MVL_IV integrates a new method to solve the objective function.

**MultiNMF** [39]: The key idea of the MultiNMF is formulating the joint matrix factorisation with a constraint to learn a common consensus.

**DAIMC** [17]: The DAIMC introduces weight matrices for each view respectively so that the case of more than two views can be adopted. Every basis matrix in the DAIMC is aligned through a regression term to reduce the impact of missing instances.

**UEAF** [18]: The UEAF constructs a unified framework, where an error matrix is used to infer missing instances and a reverse graph regularisation is introduced to preserve the local manifold structure.

**AGC_IMC** [21]: The AGC_IMC proposes a novel framework composing of with-in view preservation, between-view inferring and consensus representation learning.

In the experiment, we focus on better inferring the missing instances and generating the consensus representation in the same learning process. Firstly, each dataset is randomly removed 10%, 30% and 50% instances in each view to build the incomplete multi-view datasets. And then, each algorithm was conducted to perform the evaluation through 15 randomly generated incomplete groups, and then the average result would be recorded. Finally, to evaluate the performance of the consensus presentation learning by all the compared methods and the proposed SIIMvC, the K-means is employed as a classification method for obtaining the clustering result.

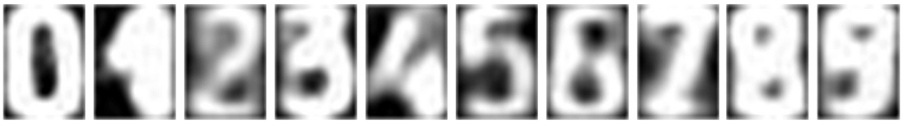
All experiments were implemented on the same software and hardware: Win 10 system, Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz 1.80 GHz, 8-GB RAM and Matlab 2020a.

## 4.3 | Experimental results and analysis

Figure 2 shows a demonstration of the missing instances and the reconstructed instances on the digit dataset under the missing rate of 0.3. It can be found that the details of the images are enhanced in the reconstructed instance compared to the missing instances.

To evaluate the performance of the proposed SIIMvC and other compared methods, the clustering accuracy (ACC), the normalised mutual information (NMI), and the purity are employed as the evaluation criterion. Tables 1–5 show the experimental results of the compared methods and the proposed SIIMvC on the BUAA dataset, the Digit dataset, the 3Sources dataset, the BBCSport dataset, and the BDGP dataset respectively. From these experimental results, interesting findings are listed as following.

**(a)** Missing instances



**(b)** Recovered instances

**FIGURE 2** The performance of inferring instances on digit dataset. In (a) are the missing instances of 'mfeat-pix' view in the digit dataset. And images in (b) are recovered instances obtained by our self-inferring multi-view clustering.

**TABLE 1** ACC, NMI and Purity of Several Algorithms on the BUAA Dataset under missing rate of 10%, 30% and 50% respectively

| | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** |
| BSV | 52.82 ± 2.47 | 46.53 ± 2.61 | 36.13 ± 1.78 | 57.69 ± 1.45 | 47.26 ± 3.6 | 32.94 ± 1.53 | 55.8 ± 2.17 | 49.87 ± 3.4 | 37.82 ± 1.69 |
| Concat | 52.02 ± 1.12 | 46.4 ± 2.49 | 36.02 ± 1.69 | 58.44 ± 1.24 | 48.81 ± 1.13 | 35.98 ± 2.54 | 54.67 ± 1.49 | 49.49 ± 2.81 | 38.24 ± 2.24 |
| GPMvC | 53.02 ± 1.91 | 38.02 ± 6.66 | 31.13 ± 2.49 | 55.19 ± 1.07 | 37.87 ± 6.82 | 28.91 ± 3.78 | 55.53 ± 1.17 | 41.76 ± 6.85 | 33.2 ± 2.58 |
| OMvC | 48.89 ± 8.5 | 45.78 ± 3.96 | 44.44 ± 3.42 | 50.46 ± 9.49 | 46.23 ± 3.41 | 43.49 ± 3.98 | 51.56 ± 8.87 | 48.0 ± 2.98 | 46.67 ± 2.48 |
| MivC | 40.91 ± 3.03 | 35.42 ± 2.07 | 29.27 ± 1.46 | 41.71 ± 4.41 | 34.73 ± 2.93 | 25.55 ± 2.58 | 43.62 ± 3.77 | 38.29 ± 2.03 | 30.78 ± 1.56 |
| MVL_IV | 50.22 ± 5.99 | 45.91 ± 8.49 | 33.87 ± 4.36 | 50.41 ± 6.48 | 46.45 ± 9.68 | 31.78 ± 5.28 | 52.04 ± 5.79 | 48.29 ± 8.52 | 35.51 ± 4.28 |
| MultiNMF | 48.84 ± 4.13 | 44.13 ± 3.71 | 31.27 ± 3.36 | 51.47 ± 4.7 | 52.46 ± 1.28 | 43.56 ± 3.89 | 51.91 ± 3.95 | 49.58 ± 2.6 | 41.11 ± 4.4 |
| DAIMC | 47.78 ± 2.83 | 45.11 ± 4.2 | 36.89 ± 5.18 | 54.76 ± 3.7 | 49.11 ± 3.36 | 38.56 ± 6.46 | 49.78 ± 2.88 | 46.89 ± 4.33 | 38.22 ± 5.86 |
| UEAF | 52.36 ± 2.42 | 49.04 ± 2.66 | 38.98 ± 2.29 | 58.48 ± 1.53 | 53.3 ± 2.2 | 42.8 ± 1.74 | 53.8 ± 2.61 | 50.56 ± 2.6 | 40.93 ± 2.45 |
| AGC_IMC | 54.02 ± 4.25 | 45.33 ± 4.23 | 34.76 ± 0.94 | 59.77 ± 2.15 | 49.66 ± 2.27 | 40.51 ± 1.85 | 55.18 ± 3.75 | 46.84 ± 4.07 | 37.49 ± 1.75 |
| SIIMvC | **54.36 ± 1.22** | **54.64 ± 2.62** | **46.58 ± 4.14** | **60.54 ± 1.24** | **58.07 ± 1.29** | **48.39 ± 3.76** | **56.93 ± 1.75** | **56.76 ± 2.03** | **48.31 ± 4.44** |

*Note*: The experimental results show that the proposed SIIMvC method outperforms all compared methods. The bold values denote the best performances in the compared methods under a specified missing rate.

Abbreviations: ACC, clustering accuracy; AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; NMI, normalised mutual information; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

**TABLE 2** ACC, NMI and Purity of Several Algorithms on the Digit Dataset under missing rate of 10%, 30% and 50% respectively

| | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** |
| BSV | 67.48 ± 3.31 | 55.03 ± 1.15 | 42.27 ± 0.93 | 63.83 ± 1.99 | 49.13 ± 0.49 | 35.03 ± 0.77 | 67.81 ± 3.14 | 55.03 ± 1.15 | 42.33 ± 0.8 |
| Concat | 50.52 ± 2.2 | 38.62 ± 0.79 | 31.28 ± 1.7 | 50.38 ± 0.57 | 38.43 ± 0.43 | 30.19 ± 0.51 | 54.09 ± 1.63 | 40.3 ± 0.61 | 32.41 ± 0.92 |
| GPMvC | 64.13 ± 2.31 | 49.17 ± 2.18 | 36.71 ± 1.87 | 61.46 ± 0.61 | 44.09 ± 1.61 | 30.92 ± 1.23 | 65.34 ± 1.15 | 50.08 ± 1.87 | 37.99 ± 1.28 |
| OMvC | 66.13 ± 5.34 | 51.89 ± 3.18 | 41.73 ± 1.29 | 60.51 ± 4.01 | 46.31 ± 2.57 | 32.38 ± 0.43 | 68.21 ± 4.72 | 54.6 ± 3.11 | 43.39 ± 1.15 |
| MivC | 77.67 ± 4.04 | 56.99 ± 1.76 | 33.2 ± 1.61 | 68.31 ± 2.75 | 49.01 ± 1.75 | 25.73 ± 1.9 | 77.67 ± 4.04 | 57.23 ± 1.66 | 34.15 ± 1.65 |
| MVL_IV | 58.23 ± 11.46 | 69.73 ± 1.78 | 49.72 ± 3.47 | 47.95 ± 11.79 | 57.52 ± 2.01 | 37.64 ± 2.37 | 59.4 ± 10.34 | 69.77 ± 1.79 | 50.34 ± 3.07 |
| MultiNMF | 81.12 ± 3.43 | 61.31 ± 0.67 | 40.56 ± 0.67 | 72.43 ± 2.78 | 68.49 ± 1.04 | 55.45 ± 1.09 | 81.16 ± 3.4 | 69.67 ± 0.73 | 54.78 ± 0.76 |
| DAIMC | 85.26 ± 2.74 | 75.31 ± 6.14 | 56.24 ± 9.87 | 77.18 ± 1.98 | 68.83 ± 4.17 | 48.52 ± 7.19 | 85.26 ± 2.74 | 75.37 ± 6.11 | 56.45 ± 9.89 |
| UEAF | 66.81 ± 2.78 | 36.43 ± 1.68 | 17.92 ± 2.89 | 59.99 ± 0.91 | 31.74 ± 1.77 | 7.46 ± 4.32 | 67.32 ± 2.25 | 36.84 ± 1.49 | 18.3 ± 2.94 |
| AGC_IMC | 83.01 ± 0.65 | **82.64 ± 0.34** | **81.81 ± 2.59** | **88.16 ± 0.33** | **86.75 ± 0.98** | **79.95 ± 1.23** | 87.37 ± 0.46 | **86.76 ± 0.91** | **83.49 ± 1.51** |
| SIIMvC | **87.77 ± 0.41** | 78.12 ± 5.06 | 63.44 ± 1.24 | 79.22 ± 0.35 | 69.84 ± 3.24 | 56.83 ± 1.02 | **87.8 ± 0.38** | 78.17 ± 5.08 | 63.93 ± 0.98 |

*Note*: The experimental results show that the proposed SIIMvC method outperforms the compared methods in most cases and is comparable to the AGC_IMC. The bold values denote the best performances in the compared methods under a specified missing rate.

Abbreviations: ACC, clustering accuracy; AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; NMI, normalised mutual information; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

**TABLE 3** ACC, NMI and Purity of Several Algorithms on the 3Sources Dataset under missing rate of 10%, 30% and 50% respectively

| | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** |
| BSV | 53.38 ± 2.5 | 46.9 ± 3.07 | 38.72 ± 1.88 | 43.02 ± 2.09 | 34.2 ± 2.22 | 21.9 ± 1.5 | 65.55 ± 2.1 | 59.18 ± 1.53 | 49.95 ± 1.92 |
| Concat | 56.36 ± 1.47 | 50.56 ± 3.28 | 43.91 ± 3.98 | 49.34 ± 1.07 | 38.53 ± 1.91 | 25.47 ± 3.5 | 70.11 ± 0.94 | 63.47 ± 2.23 | 54.45 ± 3.41 |
| GPMvC | 48.79 ± 3.09 | 45.6 ± 5.78 | 36.67 ± 6.73 | 43.26 ± 4.03 | 36.41 ± 3.21 | 23.57 ± 7.33 | 65.34 ± 3.11 | 59.61 ± 3.25 | 49.83 ± 5.91 |
| OMvC | 47.57 ± 4.85 | 40.24 ± 3.42 | 37.99 ± 4.67 | 34.15 ± 2.77 | 25.03 ± 2.33 | 15.86 ± 6.99 | 56.8 ± 2.37 | 49.47 ± 2.53 | 44.85 ± 5.25 |
| MivC | 55.44 ± 7.17 | 48.85 ± 4.24 | 34.31 ± 3.65 | 53.82 ± 5.15 | 41.18 ± 3.47 | 23.37 ± 6.49 | 73.38 ± 5.5 | 62.62 ± 2.81 | 50.06 ± 4.86 |
| MVL_IV | 57.57 ± 3.82 | 52.25 ± 4.92 | 38.32 ± 3.46 | 56.97 ± 1.81 | 40.68 ± 0.74 | 22.74 ± 3.91 | 76.69 ± 3.65 | 65.91 ± 2.55 | 54.77 ± 2.66 |
| MultiNMF | 44.56 ± 3.49 | 40.9 ± 0.48 | 35.11 ± 0.85 | 31.39 ± 2.18 | 27.19 ± 2.45 | 18.71 ± 1.27 | 55.54 ± 1.53 | 52.64 ± 0.87 | 45.76 ± 1.49 |
| DAIMC | 50.41 ± 4.49 | 53.49 ± 4.36 | 47.34 ± 7.4 | 47.52 ± 5.33 | 42.74 ± 8.15 | 34.6 ± 9.03 | 67.81 ± 5.15 | 62.25 ± 5.54 | 60.0 ± 7.02 |
| UEAF | 53.37 ± 4.53 | 61.3 ± 5.0 | 51.48 ± 5.49 | 54.61 ± 1.82 | 53.32 ± 3.64 | 42.77 ± 1.85 | 72.9 ± 3.44 | 71.36 ± 3.41 | 64.02 ± 2.31 |
| AGC_IMC | **76.33 ± 0.94** | **70.06 ± 2.86** | **55.81 ± 5.0** | **70.04 ± 2.09** | **60.75 ± 1.77** | 43.04 ± 3.94 | **83.91 ± 0.77** | **80.36 ± 1.79** | 67.63 ± 1.62 |
| SIIMvC | 57.49 ± 2.26 | 60.88 ± 5.11 | 54.58 ± 1.67 | 58.76 ± 3.22 | 54.29 ± 1.69 | **46.88 ± 1.98** | 77.73 ± 1.95 | 76.91 ± 1.26 | **69.1 ± 1.95** |

*Note*: The experimental results show that the proposed SIIMvC method outperforms the compared methods in most cases and is comparable to the AGC_IMC. The bold values denote the best performances in the compared methods under a specified missing rate.

Abbreviations: ACC, clustering accuracy; AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; NMI, normalised mutual information; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

(1) It can be found that the proposed SIIMvC outperforms most of the compared state-of-the-art incomplete MvC methods on all datasets on the ACC, the NMI and the purity, including the BSV, the Concat, the GPMvC, the OMvC, the MivC, the MVL_IV, the MultiNMP, and the DAIMC. The proposed SIIMvC also outperforms the UEAF on most of experimental cases except on the 3Source dataset and BBCSprot dataset when the instance missing rate is 30%. Compared to the AGC_IMC, the proposed SIIMvC shows advantages on datasets of the BUAA and the BDGP, and comparable performance on datasets of the Digit dataset, the 3Sources dataset, and the BBCSport dataset. On the other side, the worst performances are found that the MivC method in BUAA dataset, the Concat method in Digit dataset, the MultiNMF method in 3Source dataset and the BSV method in BBCSport dataset.

**TABLE 4** ACC, NMI and Purity of Several Algorithms on the BBCSport Dataset under missing rate of 10%, 30% and 50% respectively

| | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** |
| BSV | 55.52 ± 6.97 | 47.76 ± 4.82 | 38.79 ± 2.73 | 31.79 ± 6.68 | 27.61 ± 8.28 | 15.69 ± 2.48 | 60.17 ± 5.53 | 54.66 ± 7.36 | 43.28 ± 1.97 |
| Concat | 64.76 ± 1.42 | 57.1 ± 1.84 | 42.69 ± 3.19 | 48.99 ± 1.69 | 37.41 ± 1.4 | 19.01 ± 2.88 | 72.41 ± 1.45 | 63.29 ± 1.71 | 48.81 ± 3.57 |
| GPMvC | 58.1 ± 5.83 | 46.6 ± 4.5 | 44.74 ± 2.29 | 43.33 ± 6.34 | 27.75 ± 6.16 | 21.61 ± 3.55 | 66.69 ± 5.45 | 53.21 ± 5.28 | 51.97 ± 4.59 |
| OMvC | 57.41 ± 4.9 | 50.69 ± 5.36 | 37.24 ± 1.42 | 34.08 ± 5.35 | 27.58 ± 9.01 | 12.79 ± 3.13 | 57.93 ± 4.28 | 51.55 ± 5.89 | 39.14 ± 1.31 |
| MivC | 69.84 ± 2.23 | 54.4 ± 4.06 | 34.41 ± 2.24 | 59.75 ± 5.44 | 40.93 ± 6.92 | 14.71 ± 3.37 | 78.43 ± 2.34 | 63.4 ± 5.45 | 42.1 ± 3.24 |
| MVL_IV | 63.55 ± 8.14 | 40.0 ± 4.86 | 33.24 ± 3.6 | 43.24 ± 6.74 | 16.91 ± 4.82 | 7.72 ± 0.94 | 69.17 ± 5.71 | 46.88 ± 2.78 | 39.16 ± 2.95 |
| MultiNMF | 73.83 ± 1.71 | 60.59 ± 4.26 | 50.91 ± 2.08 | 59.6 ± 0.94 | 58.88 ± 3.32 | 44.54 ± 1.83 | 83.38 ± 0.19 | 72.91 ± 2.44 | 62.12 ± 1.84 |
| DAIMC | 64.66 ± 7.08 | 53.97 ± 7.07 | 51.72 ± 9.24 | 54.39 ± 5.33 | 42.19 ± 4.93 | 31.05 ± 11.58 | 74.31 ± 5.89 | 65.52 ± 6.54 | 57.41 ± 10.08 |
| UEAF | 75.21 ± 4.89 | **76.53 ± 4.36** | 57.09 ± 4.82 | 66.56 ± 6.74 | 58.8 ± 6.14 | 41.78 ± 5.79 | 84.83 ± 4.13 | 77.57 ± 4.08 | 64.05 ± 3.31 |
| AGC_IMC | **77.59 ± 3.45** | 75.26 ± 4.37 | 57.97 ± 4.79 | **71.68 ± 4.14** | 61.89 ± 5.51 | 39.63 ± 5.04 | **88.28 ± 3.69** | 81.31 ± 3.97 | 66.38 ± 4.6 |
| SIIMvC | 78.07 ± 2.18 | 74.45 ± 2.22 | **67.78 ± 4.95** | 68.58 ± 1.35 | **65.72 ± 3.81** | **52.84 ± 6.91** | 86.71 ± 0.5 | **84.81 ± 2.88** | **77.69 ± 5.79** |

*Note*: The experimental results show that the proposed SIIMvC method outperforms the compared methods in most cases and is comparable to the AGC_IMC. The bold values denote the best performances in the compared methods under a specified missing rate.

Abbreviations: ACC, clustering accuracy; AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; NMI, normalised mutual information; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

**TABLE 5** ACC, NMI and Purity of Several Algorithms on the BDGP Dataset under missing rate of 10%, 30% and 50% respectively

| | ACC (%) | | | NMI (%) | | | Purity (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** | **0.1** | **0.3** | **0.5** |
| BSV | 41.62 ± 0.62 | 36.82 ± 0.55 | 33.9 ± 2.08 | 24.69 ± 0.31 | 19.21 ± 0.85 | 13.96 ± 0.71 | 42.68 ± 0.3 | 38.18 ± 0.68 | 34.82 ± 2.2 |
| Concat | 43.22 ± 0.55 | 39.85 ± 0.46 | 30.92 ± 1.08 | 22.32 ± 1.09 | 19.32 ± 0.89 | 9.38 ± 0.89 | 44.05 ± 0.55 | 41.4 ± 0.58 | 32.4 ± 0.92 |
| GPMvC | 43.17 ± 3.28 | 37.77 ± 1.62 | 33.31 ± 1.35 | 17.96 ± 2.48 | 13.15 ± 0.83 | 8.05 ± 1.77 | 44.9 ± 2.42 | 39.23 ± 0.93 | 34.52 ± 1.23 |
| OMvC | 39.38 ± 1.32 | 42.7 ± 5.25 | 36.48 ± 3.36 | 11.87 ± 0.6 | 19.68 ± 4.4 | 9.98 ± 2.65 | 40.59 ± 0.62 | 43.9 ± 4.64 | 37.14 ± 3.05 |
| MivC | 24.92 ± 0.94 | 25.34 ± 0.86 | 25.16 ± 0.88 | 3.5 ± 0.48 | 3.99 ± 1.06 | 3.34 ± 1.07 | 25.46 ± 0.46 | 25.86 ± 0.78 | 26.57 ± 0.86 |
| MVL_IV | 39.38 ± 3.06 | 36.1 ± 2.9 | 30.48 ± 1.3 | 13.57 ± 1.31 | 10.25 ± 1.41 | 6.41 ± 0.8 | 40.27 ± 2.5 | 37.03 ± 2.5 | 31.73 ± 1.72 |
| MultiNMF | 23.72 ± 0.09 | 23.74 ± 0.04 | 23.36 ± 0.18 | 2.14 ± 0.04 | 3.05 ± 0.04 | 4.31 ± 0.05 | 24.48 ± 0.07 | 25.02 ± 0.02 | 25.22 ± 0.07 |
| DAIMC | 42.55 ± 2.98 | 34.58 ± 3.61 | 26.17 ± 2.65 | 18.64 ± 3.43 | 10.75 ± 2.95 | 3.03 ± 2.7 | 42.94 ± 2.48 | 35.56 ± 3.33 | 26.58 ± 3.05 |
| UEAF | 45.22 ± 1.06 | 38.27 ± 2.01 | 31.93 ± 0.72 | 18.81 ± 1.26 | 13.27 ± 1.24 | 7.96 ± 0.63 | 45.23 ± 1.06 | 39.08 ± 1.75 | 33.54 ± 1.02 |
| AGC_IMC | 33.75 ± 1.68 | 33.45 ± 2.47 | 29.96 ± 1.04 | 11.13 ± 0.97 | 10.12 ± 1.23 | 5.79 ± 1.44 | 34.8 ± 1.47 | 34.8 ± 2.52 | 30.92 ± 0.65 |
| SIIMvC | **49.42 ± 0.65** | **43.46 ± 1.58** | **42.08 ± 3.31** | **26.83 ± 0.61** | **21.81 ± 1.33** | **16.09 ± 3.27** | **50.73 ± 0.46** | **44.79 ± 0.84** | **42.33 ± 2.81** |

*Note*: The experimental results show that the proposed SIIMvC method outperforms all compared methods. The bold values denote the best performances in the compared methods under a specified missing rate.

Abbreviations: ACC, clustering accuracy; AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BDGP, Berkeley Drosophila Genome Project; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; NMI, normalised mutual information; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

(2) Among the compared methods, the UEAF method also achieves excellent clustering performances among compared methods in most cases. The incomplete data is combined with the error matrix to perform clustering in a unified framework. Thus, we can see that instances completion plays a significant role in the performance of the clustering. The proposed SIIMvC achieves better clustering performances than the UEAF method since it exploits the correlation between existing instances and missing instances. For instance, while 30% and 50% of instances were removed, the proposed SIIMvC achieves 3.76% and 5.53% improvements than the UEAF.

(3) As shown in Tables 1–5, it is clear that the proposed SIIMvC performs better than most of the methods when

the missing rates of the instances are growing. It is noteworthy that while the missing rate of the instance rises in each view, the clustering performance of all methods decrease. For example, in the BUAA dataset, while 10% of instances were removed, the proposed SIIMvC performs 2.00% higher ACC (54.36% vs. 52.36%) than the UEAF. When the instance missing rate grows to 30%, the proposed SIIMvC performs 5.60% higher (54.64% vs. 49.04%) than the UEAF. And even when the instance missing rate grows to 50%, the proposed SIIMvC performs 7.60% higher (46.58% vs. 38.98%) than the UEAF.

We also evaluate the performance of the proposed SIIMvC under different cases of missing rates. Figure 3 shows that the performance of classification accuracy degrades while the missing rate gets bigger from 0.1 to 0.7 under all experimental datasets. It is reasonable that the classification performances grow down while more data is missing since the common representation shared by all views is hard to learn.

## 4.4 | Computational complexity analysis

To our knowledge, eigenvalue decomposition, singular value decomposition, and inverse operation consume the most computational cost. As shown in Algorithm 1, only two inverse operation processes in ***step*1** and ***step*6** can be found. Specifically, the inverse operation in ***step*6** can be initialised in advance. The complexity of inverse operation in ***step*1** is $O(n^3)$ in the loop. The eigenvalue decomposition process in ***step*3** costs $O(n^3)$ and the singular value decomposition in ***step*5** consumes $O(m_v c^2)$. Therefore, our algorithm cost totally $O(\tau(2n^3 + \sum_{v=1}^{v_n} m_v c^2))$, where $\tau$ is the iteration number.

Table 6 lists the computational complexity of all the compared methods and the proposed SIIMvC. From the table, we can find that the computational complexity of our algorithm is equal to that of UEAF [18]. In other words, we get better results but do not increase computational complexity.

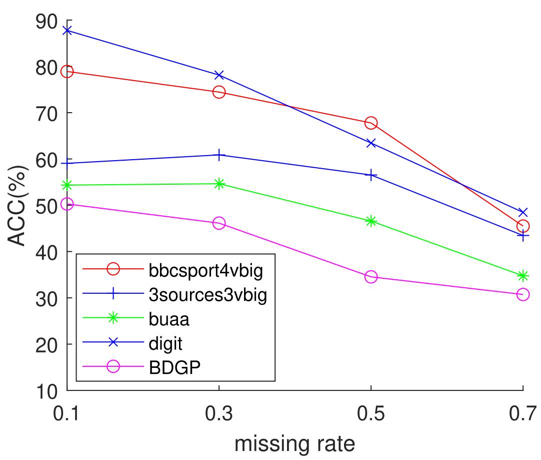



**FIGURE 3** The performances of classification accuracy of the proposed self-inferring multi-view clustering on all experimental dataset degrade while the missing rate is getting bigger from 0.1 to 0.7.

## 4.5 | Parameter complexity analysis

Furthermore, we analyse the sensitivity of all parameters related to the ACC. There are totally two basic parameters in the proposed SIIMvC, including $\lambda_1$, $\lambda_2$. $\lambda_1$ is the parameter contributing to feature graph learning and self-inferring. $\lambda_2$ controls the similarity graph $S$ to keep c components. Following the exponential rule in refs. [18, 40] of parameters, we evaluate the influence of these parameters on the experimental datasets. For different datasets, there is a specific combination of parameters to obtain the best result. For each dataset, we fix parameter $k = 7$ first, then conduct experiments with different combinations of $\lambda_1$ and $\lambda_2$. Note that the range of parameters is from $10^{(-5)}$ to $10^{(5)}$. As a demonstration, Table 7 shows the parameter settings of all the compared methods and the proposed SIIMvC on the BUAA dataset while the missing rate is 0.3. In Figure 4, we can find that the proposed SIIMvC is robust on classification accuracy to $\lambda_2$ in the range from $10^{(-5)}$ to $10^{(5)}$, and the $\lambda_1$ in the optimal range of $[10^{(-5)}, 10^{(-1)}]$ on all five datasets in the experiment.

## 4.6 | Component analysis

To vertify the effectiveness of the consensus learning and similarity graph learning, we compare the proposed SIIMvC with two models in all datasets. (1) **Consensus Learning**: We only use consensus learning term (9) in model, including self-inferring strategy. (2) **Similarity Graph Learning**: We remove consensus learning term (9) from model.

**TABLE 6** Computational complexity of other methods and our SIIMvC

| Methods | Computational complexity |
|---|---|
| BSV | $O(t\sum_{v=1}^{v_n} m_v^2 n)$ |
| Concat | $O\left(tn\left(\sum_{v=1}^{v_n} m_v\right)^2\right)$ |
| GPMvC | $O(\tau(\sum_{v=1}^{v_n} \max(m_v, n_v)n_v c + tnc^2))$ |
| OMvC | $O\left(\tau\left(T\sum_{v=1}^{v_n} m_v cn + tnc^2\right)\right)$ |
| MivC | $O\left(\tau\left(T\sum_{v=1}^{v_n} m_v cn + tnc^2\right)\right)$ |
| MultiNMF | $O\ (\tau(vm_v nc))$ |
| DAIMC | $O\left(\tau\left(Tndm_{\max} + v_n m_{\max}^3\right) + tnc^2\right)$ |
| UEAF | $O\left(\tau\left(2n^3 + \sum_{v=1}^{v_n} m_v c^2\right)\right)$ |
| AGC_IMC | $O(\tau(cn^2))$ |
| SIIMvC | $O\left(\tau\left(2n^3 + \sum_{v=1}^{v_n} m_v c^2\right)\right)$ |

*Note*: '$\tau$' denotes the number of iterations loop. '$m_{\max}$' denotes the largest number of dimensionality among all views. 't' is the number of k-means iteration. 'c' is the number of classes.

Abbreviations: AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

Experimental results of the proposed SIIMvC and two compared models are shown in Figure 5. It can be found that the proposed SIIMvC achieves better results than the other two models in most cases, which indicate that integrating the consensus learning term is beneficial to the clustering performance.

**TABLE 7** Parameters of some methods (methods with no parameter are not shown in this table) on BUAA dataset while missing rate is 0.3

| Methods | Parameters setting |
|---|---|
| BSV | None |
| GPMvC | alphas = [1e-2 1e-2], beta = 10 |
| OMvC | alpha = [1e-2 1e-2], beta = [1e-7 1e-7] tol = 1e-5 |
| MivC | alpha = [1e-3 1e-3], beta = [1e-4 1e-4] |
| MVL_IV | gamaUp = 0.7, lambdaUp = 1.1, sigma = 1e-2 |
| MultiNMF | alpha = [1e-2, 1e-2] |
| DAIMC | alpha = 1e1, beta = 1 |
| UEAF | lambda1 = 1e-2, lambda2 = 1e-1, lambda3 = 1e-1 |
| AGC_IMC | lambda1 = 1e1, lambda4 = 1e-2 $r$ = 3 |
| SIIMvC | lambda1 = 1e-4, lambda4 = 1e-3 |

Abbreviations: AGC_IMC, adaptive graph completion-based incomplete multi-view clustering; BSV, best single view; DAIMC, doubly aligned incomplete multi-view clustering; GPMvC, partial multi-view clustering using graph regularized nonnegative matrix factorization; MivC, multi-incomplete-view clustering; MultiNMF, multi-view clustering via joint nonnegative matrix factorization; MVL_IV, multi-view learning with incomplete views; OMvC, online multi-view clustering algorithm; SIIMvC, self-inferring multi-view clustering; UEAF, unified embedding alignment framework.

## 4.7 | Convergence analysis

We also analyse the convergence of the proposed SIIMvC. In detail, we record the objective value while increasing the iteration number, which was used to evaluate convergence performance. Figure 6 shows the convergence curves of the proposed SIIMvC on all experimental datasets under missing rate of 0.3. It is clear that the objective values go down quickly to a convergence value in the first five iterations. It is obvious that the proposed SIIMvC takes no more than 10 iterations to perform convergence, which indicates a great convergence performance.

# 5 | CONCLUSION

In this paper, we propose a novel incomplete MvC algorithm named SIIMvC, which exploits the existing instances to infer missing instances. In the proposed SIIMvC, an improved consensus learning method and an improved similarity graph learning method are unified for consensus representation learning. Compared to the most relevant UEAF methods, the proposed SIIMvC takes the same amount of computational complexity but fewer calculated parameters. Furthermore, the proposed SIIMvC can take the local structure of the multi-view data into account for better consensus representation learning. Extensive experiments on five popular datasets show that the proposed SIIMvC can achieve promising performance in incomplete multi-view clustering. Also, the robustness on no mater the parameter of algorithm or the change of the missing rate is verified in the experiments. In the future, we will explore further the

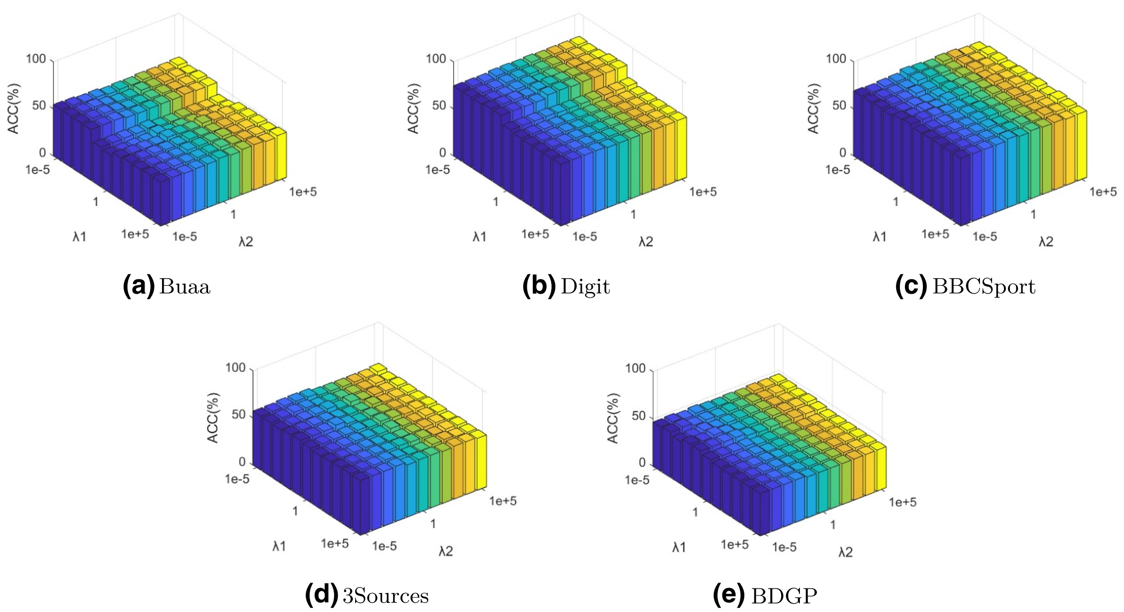



**FIGURE 4** Classification accuracy (%) on different combinations of parameter $\lambda1$ and parameter $\lambda2$ for all experimental datasets. Results show that self-inferring multi-view clustering is robust to parameter $\lambda2$, and the optimal parameter $\lambda1$ is in the range of $[10^{-5}, 10^{-1}]$. (a) Buaa, (b) Digit, (c) BBCSport, (d) 3Sources and (e) BDGP. BDGP, Berkeley Drosophila Genome Project

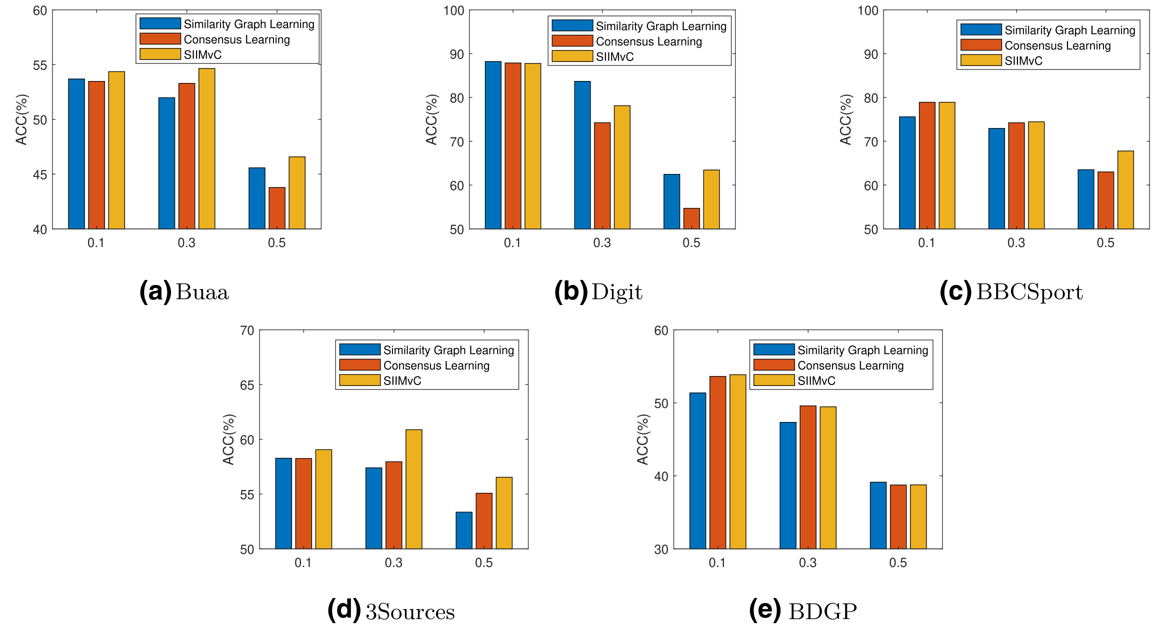



**FIGURE 5** Classification accuracy (%) of similarity graph learning component, consensus learning component, and the proposed SIIMvC on all experimental datasets while the missing rate is 10%, 30% and 50% respectively. The results show that the proposed SIIMvC outperforms the similarity graph learning component and consensus learning component, specially when the missing rate is 50%. (a) Buaa, (b) digit, (c) BBCSport, (d) 3Sources and (e) BDGP. BDGP, Berkeley Drosophila Genome Project; SIIMvC, self-inferring multi-view clustering

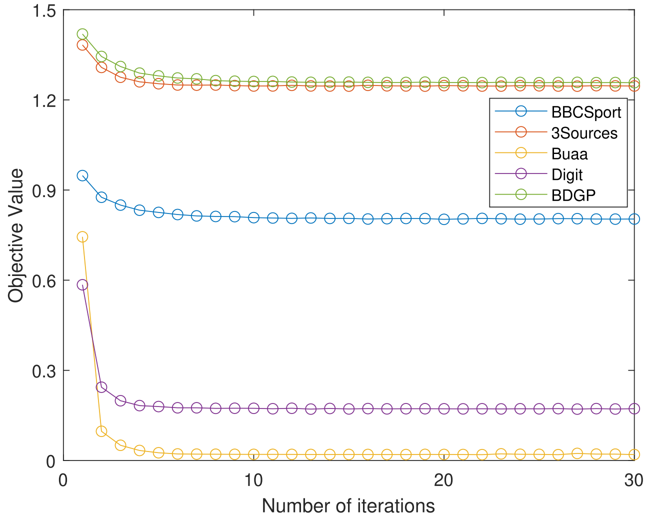



**FIGURE 6** Convergence curves of the proposed self-inferring multi-view clustering method on all experimental datasets under missing rate of 0.3. Results show that the objective values decrease quickly to a convergence value in the first five iterations.

potential of the theoretical framework in more incomplete multi-view learning tasks, for example, multimedia retrieval.

## AUTHOR CONTRIBUTIONS

**Junjun Fan**: Formal analysis; Writing – original draft. **ZeQi Ma**: Formal analysis; Methodology; Writing – review & editing. **Jiajun Wen**: Writing – review & editing. **Zhihui Lai**: Methodology; Writing – review & editing. **Weicheng Xie**: Writing – review & editing. **Wai Keung Wong**: Writing – review & editing.

## ACKNOWLEDGEMENTS

This work was supported in part by the Guangdong Science and Technology Plan under Grant 2020A1515110602, 2020A1515010707, in part by the National Natural Science Foundation of China under Grant 61703283, in part by the Guangdong Basic and Applied Basic Research Foundation 2021A1515011318, 2017A030310067, in part by the Shenzhen Municipal Science and Technology Innovation Council under the Grant JCYJ20190808113411274, in part by the Shenzhen Visual Object Detection and Recognition Key Laboratory Open Project HITSZ20220287, in part by the Overseas High-Caliber Professional in Shenzhen under Project 20190629729C, in part by the High-Level Professional in Shenzhen under Project 20190716892H, in part by the Research Foundation for Postdoctor Worked in Shenzhen under Project 707-0001300148 and 707-0001310414, in part by the National Engineering Laboratory for Big Data System Computing Technology, in part by the Guangdong Laboratory of Artificial-Intelligence and Cyber-Economics (SZ), in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society, in part by the Scientific Research Foundation of Shenzhen University under Project 2019049, Project 860-000002110328 and Project 827-000526, and in part by the Laboratory for Artificial Intelligence in Design (Project Code: RP3-3) under the InnoHK Research Clusters, Hong Kong SAR Government.

## CONFLICT OF INTEREST

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

## DATA AVAILABILITY STATEMENT

All datasets used in the submitted paper are available publicly on these websites and papers: (1) Digit dataset: https://archive.ics.uci.edu/ml/datasets/Multiple+Features. (2) 3Sources dataset: http://erdos.ucd.ie/datasets/3sources.html. (3) BUAA-visnir face dataset (BUAA): https://github.com/hdzhao/IMG/tree/master/data. (4) BBCSport dataset: https://github.com/GPMVCDummy/GPMVC/tree/master/partialMV/PVC/recreateResults/data. (5) BGDP dataset https://fruitfly.org/sequence/human-datasets.html.

## ORCID

*Junjun Fan* https://orcid.org/0000-0001-5306-7258
*ZeQi Ma* https://orcid.org/0000-0001-5449-7196
*Jiajun Wen* https://orcid.org/0000-0002-1688-2541

## REFERENCES

1. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. In: Proceedings of the Eleventh Annual Conference on Computational Learning Theory, pp. 92–100 (1998). https://doi.org/10.1145/279943.279962
2. Greene, D., Cunningham, P.: A matrix factorization approach for integrating multiple data views. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, vol. 5781, pp. 423–438. Springer, Berlin (2009). https://doi.org/10.1007/978-3-642-04180-8_45
3. Xu, C., Tao, D., Xu, C.: A survey on multi-view learning. arXiv preprint arXiv:1304.5634 (2013)
4. Chao, G., Sun, S., Bi, J.: A survey on multi-view clustering. arXiv preprint arXiv:1712.06246, abs/1712.06246 (2017)
5. Li, Y., Yang, M., Zhang, Z.: A survey of multi-view representation learning. IEEE Trans. Knowl. Data Eng. 31(10), 1863–1883 (2018). https://doi.org/10.1109/TKDE.2018.2872063
6. Peng, X., et al.: Comic: multi-view clustering without parameter selection. In: International Conference on Machine Learning, pp. 5092–5101. PMLR (2019)
7. Tzortzis, G., Likas, A.: Kernel-based weighted multi-view clustering. In: 2012 IEEE 12th International Conference on Data Mining, pp. 675–684 (2012). https://doi.org/10.1109/ICDM.2012.43
8. Tang, W., Lu, Z., Dhillon, I.S.: Clustering with multiple graphs. In: 2009 Ninth IEEE International Conference on Data Mining, pp. 1016–1021. IEEE, Washington (2009). https://doi.org/10.1109/ICDM.2009.125
9. Xu, C., Tao, D., Xu, C.: Multi-view intact space learning. IEEE Trans. Pattern Anal. Mach. Intell. 37(12), 2531–2544 (2015). https://doi.org/10.1109/TPAMI.2015.2417578
10. Wong, W.K., et al.: Clustering structure-induced robust multi-view graph recovery. IEEE Trans. Circ. Syst. Video Technol. 30(10), 3584–3597 (2019). https://doi.org/10.1109/TCSVT.2019.2945202
11. De Sa, V.R.: Spectral clustering with two views. In: ICML Workshop on Learning with Multiple Views, pp. 20–27 (2005)
12. Guo, Y.: Convex subspace representation learning from multi-view data. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 27, pp. 387–393 (2013)
13. Rai, P., et al.: Multiview clustering with incomplete views. In: Proceedings of the NIPS Workshop on Machine Learning for Social Computing. Citeseer (2010)
14. Li, S.-Y., Jiang, Y., Zhou, Z.-H.: Partial multi-view clustering. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 28 (2014)
15. Shao, W., He, L., Yu, P.S.: 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, vol. 9284, pp. 318–334. Springer, Berlin (2015). https://doi.org/10.1007/978-3-319-23528-8_20
16. Shao, W., et al.: Online multi-view clustering with incomplete views. In: 2016 IEEE International Conference on Big Data (Big Data), pp. 1012–1017. IEEE, Washington (2016). https://doi.org/10.1109/BigData.2016.7840701
17. Hu, M., Chen, S.: Doubly aligned incomplete multi-view clustering. arXiv preprint arXiv:1903.02785 (2019)
18. Wen, J., et al.: Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 5393–5400 (2019). https://doi.org/10.1609/aaai.v33i01.33015393
19. Wen, J., et al.: DIMC-Net: Deep Incomplete Multi-View Clustering Network, pp. 3753–3761. Association for Computing Machinery, New York (2020). ISBN 9781450379885. https://doi.org/10.1145/3394171.3413807
20. Lin, Y., et al.: Completer: incomplete multi-view clustering via contrastive prediction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11174–11183 (2021)
21. Wen, J., et al.: Adaptive graph completion based incomplete multi-view clustering. IEEE Trans. Multimed. 23, 2493–2504 (2021). https://doi.org/10.1109/TMM.2020.3013408
22. Xiao, Y., et al.: Auto-weighted sample-level fusion with anchors for incomplete multi-view clustering. Pattern Recogn. 130, 108772 (2022). https://doi.org/10.1016/j.patcog.2022.108772
23. Yang, M., et al.: Robust multi-view clustering with incomplete information. IEEE Trans. Pattern Anal. Mach. Intell, 1 (2022). https://doi.org/10.1109/TPAMI.2022.3155499
24. Nie, F., Cai, G., Li, X.: Multi-view clustering and semi-supervised classification with adaptive neighbours. In: Thirty-first AAAI Conference on Artificial Intelligence, pp. 2408–2414. AAAI Press, Palo Alto (2017)
25. Liang, N., et al.: Incomplete multi-view clustering with cross-view feature transformation. IEEE Trans. Artif. Intell. 3(5), 749–762 (2022). https://doi.org/10.1109/TAI.2021.3139573
26. Mohar, B., et al.: The Laplacian spectrum of graphs. Graph Theor. Comb. Appl. 2(871-898), 12 (1991)
27. Zhan, K., et al.: Multiview consensus graph clustering. IEEE Trans. Image Process. 28(3), 1261–1270 (2018). https://doi.org/10.1109/TIP.2018.2877335
28. Fan, K.: On a theorem of Weyl concerning eigenvalues of linear transformations I. Proc. Natl. Acad. Sci. USA 35(11), 652–655 (1949). https://doi.org/10.1073/pnas.35.11.652
29. Nie, F., et al.: The constrained laplacian rank algorithm for graph-based clustering. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30, pp. 1969–1976. AAAI Press, Palo Alto (2016)
30. Zou, H., Hastie, T., Tibshirani, R.: Sparse principal component analysis. J. Comput. Graph Stat. 15(2), 265–286 (2006). https://doi.org/10.1198/106186006x113430
31. Huang, D., Sun, J., Wang, Y.: The BUAA-visnir face database instructions. Technical Report IRIP-TR-12-FR-001, School Comput. Sci. Eng., Beihang Univ., Beijing, p. 3 (2012)
32. Hand Written Dataset. https://archive.ics.uci.edu/ml/datasets/Multiple+Features. Accessed 16 Mar 2021. [Online]
33. 3Sources Dataset. http://erdos.ucd.ie/datasets/3sources.html. Accessed 19 Mar 2021. [Online]
34. Greene, D., Cunningham, P.: Practical solutions to the problem of diagonal dominance in kernel document clustering. In: Proceedings of the 23rd International Conference on Machine Learning, vol. 148, pp. 377–384. ACM, New York (2006). https://doi.org/10.1145/1143844.1143892
35. BDGP Dataset. https://fruitfly.org/sequence/human-datasets.html. Accessed 28 June 2014. [Online]
36. Zhao, H., Liu, H., Fu, Y.: Incomplete multi-modal visual data grouping. In: IJCAI, pp. 2392–2398. IJCAI/AAAI Press, Palo Alto (2016)

37. Rai, N., et al.: Partial multi-view clustering using graph regularized NMF. In: 2016 23rd International Conference on Pattern Recognition (ICPR), pp. 2192–2197. IEEE, Washington (2016). https://doi.org/10.1109/ICPR.2016.7899961
38. Zhao, H., Ding, Z., Fu, Y.: Multi-view clustering via deep matrix factorization. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
39. Liu, J., et al.: Multi-view clustering via joint nonnegative matrix factorization. In: Proceedings of the 2013 SIAM International Conference on Data Mining, pp. 252–260. SIAM, Philadelphia (2013). https://doi.org/10.1137/1.9781611972832.28
40. Zhao, S., et al.: Intrinsic and complete structure learning based incomplete multiview clustering. IEEE Trans. Multimed. 1 (2021). https://doi.org/10.1109/TMM.2021.3138638

**How to cite this article:** Fan, J., et al.: Self-inferring incomplete multi-view clustering. IET Comput. Vis. 1–15 (2022). https://doi.org/10.1049/cvi2.12152