ImMC-CSFL: Imbalanced Multi-view Clustering Algorithm Based on Common-Specific Feature Learning

Xiaocui Li[1,2], Yu Xiao[1], Xinyu Zhang[1,2](✉), Qingyu Shi[1,2], and Xiance Tang[1]

[1] School of Computer Science, Hunan University of Technology and Business, Changsha, China
zhangxinyu247@163.com

[2] Xiangjiang Laboratory, Changsha, China

Abstract. Clustering as one of the main research methods in data mining, with the generation of multi-view data, multi-view clustering has become the research hotspot at present. Many excellent multi-view clustering algorithms have been proposed to solve various practical problems. These algorithms mainly achieve multi-view feature fusion by maximizing the consistency between views. However, in practical applications, multi-view data' initial feature is often imbalanced, resulting in poor performance of existing multi-view clustering algorithms. Additionally, imbalanced multi-view data exhibits significant differences in feature across different views, which better reflects the complementarity of multi-view data. Therefore, it is important to fully extract feature from different views of imbalanced multi-view data. This paper proposes an imbalanced multi-view clustering algorithm based on common specific feature learning, ImMC-CSFL. Two deep networks are used to extract common and specific feature on each view, the GAN network is introduced to maximize the extraction of common feature from multi-view data, and orthogonal constraints are used to maximize the extraction of specific feature from different views. Finally, the learned imbalanced multi-view feature is input for clustering. The experiment result on three different multi-view datasets UCI Digits, BDGP, and CCV showed that our proposed algorithm had better clustering performance, and the effectiveness and robustness were verified through experiment analysis of different modules.

Keywords: imbalanced multi-view data · common-specific information network · deep feature learning network · multi-view clustering

1 Introduction

As one of the main research contents of machine learning, clustering technique has always been a research hotspot. With the rise of multi-view data, originating from different sources or modalities, multi-view clustering has been attracted significant attention from researchers, with a focus on multi-view feature learning. There is not only consistency but also differences between different views, that each view contains the common information or specific information. The key of multi-view feature learning is to comprehensively extract clustering-friendly feature from multi-view data. Multi-view data both

D.-N. Yang et al. (Eds.): PAKDD 2024, LNAI 14645, pp. 221–232, 2024.
https://doi.org/10.1007/978-981-97-2242-6_18

have similarities and disparities among different views. There are numerous excellent multi-view clustering algorithms, mainly including: multi-view spectral clustering algorithm, multi-view subspace clustering algorithm, multi-view clustering methods based on non-negative matrix decomposition, and multi-kernel based multi-view clustering algorithms.

1.1 Motivation

While multi-view clustering algorithms have received considerable attention and displayed good performance, they are exclusively suitable for multi-view data with relatively balanced initial features and perform inadequately for multi-view data with varied quality and imbalanced initial features. The imbalanced multi-view clustering is an urgent problem, which necessitate addressing. The main reason are as follows:

1) For multi-view data with imbalanced initial features, current methods fail to consider the specific information of the different view. It is worth noting that the specific information of the imbalanced multi-view data is particularly rich, which is more beneficial for extracting complementary features;
2) The extraction of the common and specific information of imbalanced multi-view data presents a significant challenge for feature fusion. Feature fusion strategies and clustering methods for the initial feature-imbalanced multi-view dataset are still relatively rare.

1.2 Contribution

To address the aforementioned problem, we propose an imbalanced multi-view clustering algorithm based on common and specific feature learning ImMC-CSFL, a novel approach that effectively integrates the common and specific information of multi-view data in a unified framework. The main contributions of our proposed approach are as follows:

(1) We design a unified framework to integrate common and specific information of imbalanced multi-view data, so that our approach can simultaneously utilize the consistency and complementarity of multi-view data.
(2) ImMC-CSFL incorporates GAN techniques and orthogonal constraints respectively to fully extract the common feature and specific feature of different views by iteratively training the common-specific information learning network and clustering network.
(3) To verify the effectiveness of ImMC-CSFL, extensive experiments were performed on three widely used clustering datasets, UCI Digits, BDGP, and CCV. Experimental results show that our common-specific multi-view feature learning model can more fully extract the feature of imbalanced multi-view data and achieve better clustering results compared with existing mainstream methods.

2 Related Work

At present, there are various excellent multi-view clustering algorithms both domestically and internationally. Based on the different mechanisms used, multi-view clustering is divided into the following categories.

Multi-view Spectral Clustering. It integrates multi-view data using graph fusion and employs spectral clustering for segmentation. Huang et al. introduced MvSCN, emphasizing intra-view invariance and inter-view consistency [1],. Zhu et al. proposed OMSC to address the limitations of two-step methods [2]. Yin et al. introduced a one-step method based on CSNE [3]. Jia et al. developed MVSC for tensor low-rank representations, focusing on intra-view and inter-view relationships [4] El Hajjar et al. presented CNESE, incorporating non-negative embedding [5].

Multi-view Subspace Clustering. It explores consistent subspaces within multi-view data to cluster similar data types. Gao et al. introduced MVSC in 2015, clustering subspace features from each view simultaneously [6]. Brbic et al. proposed low-rank sparse subspace multi-view clustering by constructing a shared affinity matrix to learn a unified subspace representation [7]. Zhang et al. addressed subspace recognition issues with flexible multi-view representation learning [8], further proposing LMSC to extract latent complementary information from multiple views [9]. Kang et al. introduced LMVSC, a large-scale multi-view subspace clustering algorithm with linear time complexity [10].

Multi-view Nonnegative Matrix Factorization Clustering. It employs non-negative matrix factorization to decompose the multi-view feature matrix into an indicator matrix and a base matrix, forming a multi-view shared indicator matrix[11]. In 2018, Zhang et al. introduced clustering analysis based on multi-view matrix decomposition, leveraging the local structure of samples [12]. Mekthanavanh et al. developed a multi-view social network video clustering model using non-negative matrix decomposition to create a shared consistent matrix from the latent feature matrix [13]. Nie et al. proposed FMVBKM for fast bilateral K-means multi-view clustering, introducing fast multi-view matrix triple decomposition [14]. Liu et al. addressed the issue of assigning equal weights to views in multi-view NMF algorithms with WM-NMF, which assigns weights to views to reduce the impact of unimportant views [15].

Multi-view Clustering Based on Multiple Kernels. It achieves clustering in a higher dimensional feature space by using a kernel function to map the sample features into this space. Liu et al. proposed a matrix-induced regularized multi-kernel k-means clustering MKKM [16], which reduces the selection of redundant kernels and enhances the diversity of kernels. Based on this, Liu et al. proposed a multicore clustering based on the subspace partitioning of nearest-neighbor kernels [17]and a missing multicore learning algorithm AMKL [18]. Sun et al. proposed MKLR-RMSC [19], a robust multi-view subspace clustering method using multi-kernel low-rank representations to extract unique and complementary view-specific information.

3 Imbalanced Multi-view Clustering Algorithm Based on Common-Specific Feature Learning (ImMC-CSFL)

The complementarity between different views of multi-view data can be reflected in the two aspects: 1) Feature consistency between different views, i.e. different views contain the consistency information; 2) Feature specificity between different views, i.e. each view contains its own differential information. To fully extract the consistency information and specificity information of multi-view data, and to leverage the complementarity of multi-view data, we proposed an imbalanced multi-view clustering algorithm based on common-specific feature learning. As shown in Fig. 1, the framework includes a deep feature extraction module, a common information learning module based on a GCN, and a specific information learning module on differential loss.

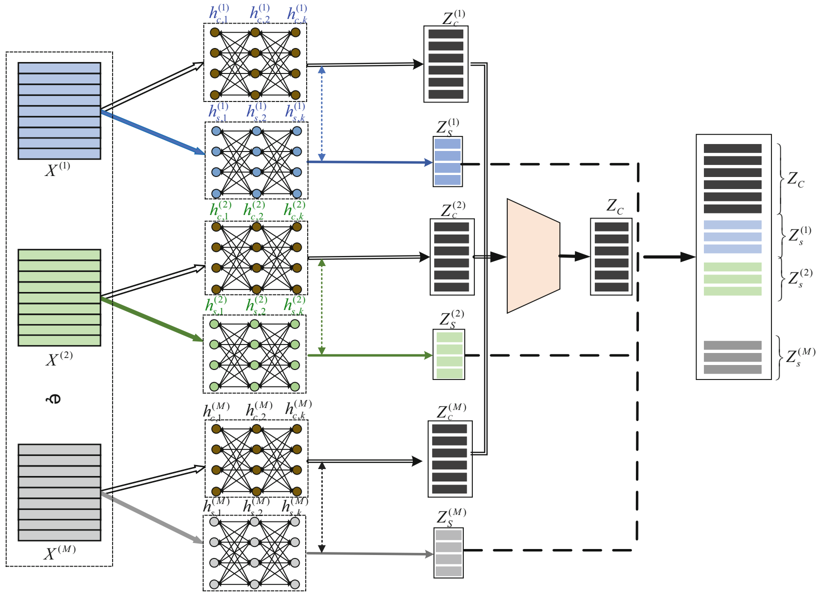

Fig. 1. The framework of imbalanced multi-view clustering algorithm based on common-specific feature learning.

3.1 Deep Feature Extraction Module

The feature extraction module based on deep convolutional networks contains two sub-networks: i.e., the common information extraction sub-network and the specific information extraction sub-network, which are used to extract the common feature across all views and the specific feature of each view. For a multi-view dataset $X = \{X^{(1)}, X^{(2)}, \ldots, X^{(M)}\}$ $\mathrm{X} = \{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(\mathrm{M})}\}$ $\mathrm{X} = \{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(\mathrm{M})}\}$ $\mathrm{X} = \{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(\mathrm{M})}\}$ $\mathrm{X} = \{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(\mathrm{M})}\}$, where M represents the number of views, $X^{(m)} \in \mathbb{R}^{d^m \times N}$, d^m is the dimensionality of samples in the *m-th* view, and N denotes the total number of samples. Each view is inputted into a deep learning network that connects two separate deep learning networks with various fully connected layers. The deep feature extraction module can fully extract the useful information of multi-view data, which is valuable for subsequent feature learning.

Assuming that both the common information extraction sub-network and the specific information extraction sub-network in each view consist of $n + 1$ fully connected layers, with each $k - th$ layer containing p_{sk} units, $k \in [0, n]$. The output of sample x from the *k-th* layer of the common information network in the *m-th* view can be calculated as follows:

$$f_{ck}^{m}(x) = h_{ck}^{m} = \varphi(W_{ck}^{m}h_{ck}^{m-1} + b_{ck}^{m}) \tag{1}$$

where $W_{ck}^{m} \in \mathbb{R}^{p_{ck} \times p_{c(k-1)}}$ and $b_{ck}^{m} \in \mathbb{R}^{p_{ck}}$ indicate the weight matrix and bias vector of the *k-th* layer in the common information extraction sub-network, respectively. φ is a non-linear activation function, commonly include *sigmoid* and *tanh*.

Simultaneously, the output of sample x from the *k-th* layer of the specific information net-work in the *m-th* view can be calculated as follows:

$$f_{sk}^{m}(x) = h_{sk}^{m} = \varphi(W_{sk}^{m}h_{sk}^{m-1} + b_{sk}^{m}) \tag{2}$$

where $W_{sk}^{m} \in \mathbb{R}^{p_{sk} \times p_{s(k-1)}}$ and $b_{sk}^{m} \in \mathbb{R}^{p_{sk}}$, represent the weight matrix and bias vector of the *k-th* layer in the specific information extraction sub-network, respectively.

Therefore, for the *i-th* sample x_i^m in the *m-th* view, we can separately obtain corresponding common information and specific information, denoted as $h_{c,i}^{m}$ and $h_{s,i}^{m}$.

$$h_{c,i}^{m} = f_{cn}^{m}(x_i^m)$$

$$h_{s,i}^{m} = f_{sn}^{m}(x_i^m) \tag{3}$$

3.2 Common Information Learning Module

By inputting the multi-view data into common information feature extraction sub-networks, we can obtain common information of the same sample from different views. To maximize the common information extraction from different views, Generative Adversarial Network (GAN) technology is utilized in our framework. Figure 2 shows the structure of the common information learning module. Differ from traditional GAN, we consider the deep common feature extraction network on each view as a generator G, therefore there will be M generators, and input the common feature from the M generators into a for M-categorical discriminator D. The goal of G in this module is to generate feature with similar distributions on different views, so that the discriminator D struggles to determine the feature coming from which view. On the contrary, the goal of D is to determine which view the incoming feature come from through adversarial training. The aim of adversarial learning is to ensure that the extracted common information from different views is as similar as possible. Essentially, this model strives to maximize the extraction of the common information from different views. Through the above analysis, the loss of this module is as follows:

$$L_c = \underset{G}{min}\underset{D}{max}(\sum\nolimits_{i=1}^{N}\sum\nolimits_{m=1}^{M} l_i^m logD(G^m(x_i^m))) \tag{4}$$

where, G^m indicates the generator (common information extraction network) on the *m-th* view, $G^m(x_i^m)$ indicates the feature generated form sample x_i^m generated by generator G^m, and l_i^m indicates the real label for sample x_i^m. The output of $D(G^m(x_i^m))$ is the probability that the generated sample comes from view m:

$$P_i^m = D(G^m(x_i^m)) \tag{5}$$

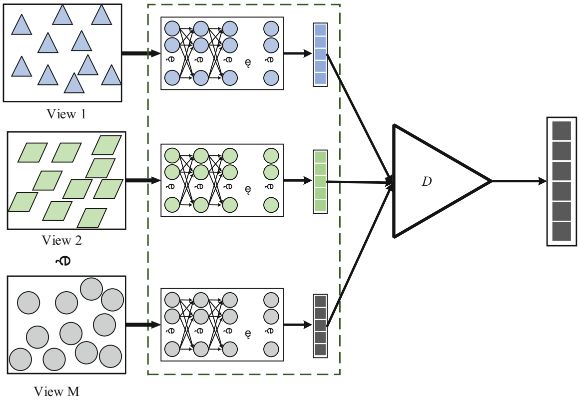

Fig. 2. Common Information Learning Module.

3.3 Specific Information Learning Module

To maximize the extraction of specific information, we minimize the correlation between specific information and common information. In our approach an orthogonal constraint is applied between the specific and common information within each view. The specific information learning network is shown in Fig. 3.

For the *i-th* sample x_i^m in the m-th view, it is simultaneously input into both the common information extraction sub-network and the specific information extraction sub-network. This results in obtaining the common information feature vector for the sample, denoted as $h_{c,i}^m$, and the specific information feature vector, denoted as $h_{s,i}^m$. Therefore, by incorporating an orthogonal constraint, the loss function is as follows:

$$L_s = \sum\nolimits_{i=1}^{N} \|(h_{c,i}^m)^T h_{s,i}^m\|^2 \tag{6}$$

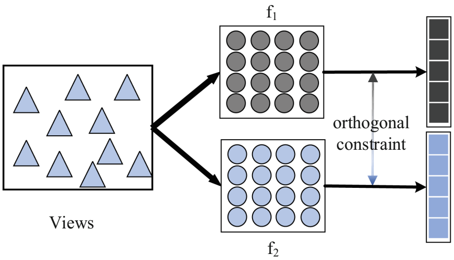

Fig. 3. Specific Information Learning Module.

3.4 Deep Multi-view Clustering Based on Common-Specific Feature Learning

Through the common-specific multi-view feature learning network, the common feature and specific feature can be obtained from each view. For the i-th sample x_i^m in the m-th view, take $h^m_{c,i}$ and $h^m_{s,i}$ as the extracted common and specific feature on view m. Then, by combining the common and specific feature vectors extracted from all views, the common-specific feature h_i for sample i is obtained as follows:

$$h_i = [(h_{c,i})^T, (h^1_{s,i})^T, (h^2_{s,i})^T, \ldots, (h^M_{s,i})^T]^T \tag{7}$$

where $h_{c,i}$ represents the common feature of all views, i.e., multi-view common feature, it can be computed as follows:

$$hh_{c,i} = \frac{1}{M}\sum\nolimits_{m=1}^{M} h^m_{c,i} \tag{8}$$

Then input the multi-view common-specific feature into the clustering network. Through the iterative training of the common-specific feature learning network and the clustering network, as shown in Fig. 4, to learn a positive clustering structure based on common-specific feature.

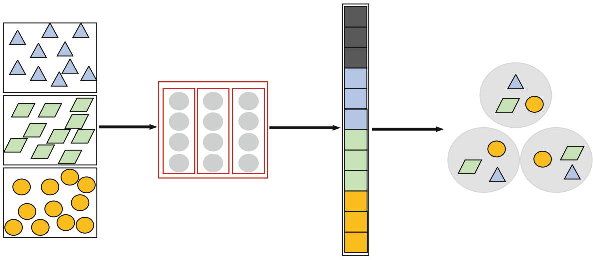

Fig. 4. The Deep Multi-View Clustering based on Common-Specific Feature Learning.

Through the above analysis, the total loss of the imbalanced multi-view clustering algorithm based on common-specific feature learning is designed as follows:

$$L = L_c + \lambda_1 L_s + \lambda_2 L_{clu} \tag{9}$$

where λ_1 and λ_2 are balancing factors used to adjust the weights of each part of the loss in the overall objective function. L_{clu} represents the clustering loss and is computed using the following formula:

$$L_{clu} = \sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1}^{K} p_{ij} log\frac{p_{ij}}{q_{ij}} \tag{10}$$

where K is the number of clusters, q_{ij} represents the soft assignment probability of sample i belonging to cluster j, and p_{ij} represents the target probability of sample i belonging to cluster $j.q_{ij}$ and p_{ij} are calculated as follows:

$$q_{ij} = \frac{(1 + \left\| h_i - \mu_j \right\|^2/\alpha)^{-\frac{\alpha+1}{2}}}{\sum_{j'=1}^{K} (1 + \left\| h_i - \mu_{j'} \right\|^2/\alpha)^{-\frac{\alpha+1}{2}})} \tag{11}$$

where, u_j indicates the center of cluster j, which can be obtained through clustering algorithms such as k-means applied to the common-specific feature extracted from the samples. α is a parameter variable, in our framework its value is set as 1, following reference [20].

$$p_{ij} = \frac{q_{ij}^2 / f_j}{\sum_{j'=1}^{K} q_{ij}^2 / f_{j'}} \tag{12}$$

where f_j is the sum of soft assignment probabilities of all samples belonging to cluster j, specifically:

$$f_j = \sum_{i=1}^{N} q_i^j \tag{13}$$

4 Experiment

4.1 Experimental Datasets and Evaluation Criteria

In order to validate the effectiveness of our proposed method, the experiments were taken in three multi-view datasets, UCI Digits [21], BDGP [22], and CCV [23] Dataset.

UCI Digits: The dataset contains 10 classes of handwritten digits, each with 200 different digits, for a total of 2000 data [25] that contains 6 feature sets.

BDGP: The dataset contains 2500 drosophila embryo data categorized into 5 classes. Each image consists of 1750-D visual vectors and 79-D textual feature vectors, i.e., the dataset contains 2 modalities.

CCV: The dataset consists of 9317 video segments collected on YouTube and contains 20 different semantic categories.

Evaluation Criteria: In the experiments clustering accuracy (ACC), normalized mutual information (NMI), and clustering purity (Purity) were selected as evaluation criteria.

4.2 Methods of Comparison

In order to fully evaluate algorithm performance, some excellent multi-view clustering methods were collected for comparison. Specifically,

(1) Traditional multi-view clustering methods: BestView [24], ConSC [25], RMSC, MVSC, CSMSC [26], MCNDCL [27].
(2) Multi-view clustering methods based on deep learning: DCCA [28], DMSC [29], DAMC [20].

4.3 Experimental Results

Table 1 shows the results of our method and the compared methods on the UCI Digits. From the results, we can see that the ImMC-CSFL can achieve the best performance on most of the criteria. Different from the BestView and ConSC method which is a single view-based clustering, the ImMC-CSFL effectively utilizes the advantages of multi-view data. The performance of ImMC-CSFL is also significantly improved compared to traditional multi-view methods. The ImMC-CSFL can learn more discriminative and clustering-friendly feature through deep learning techniques.

Table 1. The clustering performance of ImMC-CSFL and compared methods on UCI

Methods	ACC	NMI	Purity
BestView	68.2	66.3	69.9
ConSC	82.8	80.2	83.1
RMSC	86.3	78.0	90.4
MVSC	81.8	85.9	80.2
CSMSC	79.8	76.4	81.2
MCNDCL	90.3	84.6	90.7
DCCA	81.4	78.1	81.4
DMSC	91.6	85.5	91.6
DAMC	<u>96.5</u>	<u>93.2</u>	<u>96.5</u>
ImMC-CSFL	**97.8**	**95.6**	**98.2**

Table 2. The clustering performance of ImMC-CSFL and compared methods on BDGP

Methods	ACC	NMI	Purity
BestView	94.0	89.4	94.2
ConSC	58.4	38.4	58.4
RMSC	60.2	56.3	60.2
MVSC	68.2	56.9	68.4
CSMSC	94.9	84.9	94.8
MCNDCL	88.5	85.7	86.5
DCCA	57.8	40.9	57.8
DMSC	68.1	50.6	73.8
DAMC	<u>98.2</u>	<u>94.6</u>	**98.2**
ImMC-CSFL	**98.5**	**95.8**	<u>97.6</u>

Table 2 shows the evaluation results of all competing methods on the BDGP dataset. It shows that most of the multi-view methods underperform than BestView, our method

ImMC-CSFL is higher than the second best. In addition, the performance of the ImMC-CSFL method is significantly higher than BestView, which indicates that ImMC-CSFL is able to fully integrate the imbalanced multi-view data to improve the clustering performance.

Table 3 shows the experimental results of ImMC-CSFL and other compared methods on the CCV dataset. From the results, it is obvious that all the methods give unsatisfactory results on this dataset. That is because the quality of the samples on each view on this dataset is not high (as can be reflected from the results of BestView). This leads to unsatisfactory performance even after fusing multiple views, and improvement for this aspect will be a direction for subsequent research. Although all the methods underperform on this dataset, the ImMC-CSFL can also alleviate this phenomenon to some extent.

Table 3. The clustering performance of ImMC-CSFL and compared methods on CCV dataset

Methods	ACC	NMI	Purity
BestView	19.5	17.6	22.0
ConSC	10.6	8.6	10.8
RMSC	21.6	18.0	24.1
MVSC	19.3	15.2	21.0
CSMSC	23.9	18.7	27.8
MCNDCL	24.2	19.3	25.6
DCCA	20.7	15.9	21.9
DMSC	17.5	13.5	25.1
DAMC	25.6	22.5	**28.6**
ImMC-CSFL	**28.7**	**25.4**	28.5

5 Summary

In this paper, we propose an imbalanced multi-view clustering algorithm based on common-specific feature learning, called ImMC-CSFL. ImMC-CSFL first extracts the common information feature vector and the specific information feature vector on each view through two deep feature extraction networks, respectively. Then, through GAN technology, maximize common feature extracted from different views. Orthogonal constraints are introduced to minimize the correlation between common feature and specific feature. Finally, input the learned common specific feature into the clustering network for iterative training. Extensive experimental results and validation analyses on three public datasets show that the proposed ImMC-CSFL has better performance than the existing mainstream methods.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (No. 62206092), the Natural Science Foundation of Hunan Province (No. 2023JJ40236, 2023JJ40239 and 2022JJ40129), the Research Foundation of Education Bureau of Hunan Province (No. 21B0582, 21B0565 and 21B0572), the Open Project of Xiangjiang Laboratory (No. 22XJ03012,22XJ03014,22XJ03022).

References

1. Huang, Z., Zhou, J., Peng, X., et al.: Multi-view Spectral Clustering Network. IJCAI, pp. 2563–2569 (2019)
2. Zhu, X., Zhang, S., He, W., et al.: One-step multi-view spectral clustering. IEEE Trans. Knowl. Data Eng.Knowl. Data Eng. **31**(10), 2022–2034 (2019)
3. Yin, H., Hu, W., Li, F., et al.: One-step multi-view spectral clustering by learning common and specific nonnegative embeddings. Int. J. Mach. Learn. Cybern.Cybern. **12**(7), 2121–2134 (2021)
4. Jia, Y., Liu, H., Hou, J., et al.: Multi-view spectral clustering tailored tensor low-rank representation. IEEE Trans. Circuits Syst. Video Technol. **31**(12), 4784–4797 (2021)
5. El Hajjar, S., Dornaika, F., Abdallahde, F., et al.: Multi-view spectral clustering via constrained nonnegative embedding. Inf. Fusion **78**, 209–217 (2022)
6. Gao, H., Nie, mF., Li, X., et al.: Multi-view subspace clustering. In: ICCV 2015, pp. 4238–4246 (2015)
7. Brbic, M., Kopriva, I.: Multi-view low-rank sparse subspace clustering. Pattern Recognit. **73**, 247–258 (2018)
8. Li, R., Zhang, C., Hu, Q., et al.: Flexible multi-view representation learning for subspace clustering. IJCAI 2019, pp. 2916–2922 (2019)
9. Zhang, C., Hu, Q., Fu, H., et al.: Generalized latent multi-view subspace clustering. IEEE Trans. Pattern Anal. Mach. Intell.Intell. **42**(1), 86–99 (2020)
10. Kang, Z., Zhou, W., Zhao, Z., et al.: Large-scale multi-view subspace clustering in linear time. In: AAAI 2020, pp. 4412–4419 (2020)
11. Liu, J., Wang, C., Gao, J., et al.: Multi-view clustering via joint nonnegative matrix factorization. In: SDM 2013, 252–260 (2013)
12. Zhang, Y., Kong, X.W., Wang, Z.F., et al.: Cluster analysis based on multi-view matrix decomposition. J. Autom. **2018**(44).12, 2160–2169 (2018)
13. Mekthanavanh, V., Li, T., Meng, H., et al.: Social web video clustering based on multi-view clustering via nonnegative matrix factorization. Int. J. Mach. Learn. Cybern.Cybern. **10**(10), 2779–2790 (2019)
14. Nie, F., Shi, S., Li, X.: Auto-weighted multi-view co-clustering via fast matrix factorization. Pattern Recogn.Recogn. **102**, 107207 (2020)
15. Liu, S.S., Lin, L.: Integrative clustering of multi-view data by nonnegative matrix factorization. ArXiv, abs/2110.13240 (2021)
16. Liu, X., Dou, Y., Yin, J., et al.: Multiple kernel k-means clustering with matrix-induced regularization. In: AAAI 2016, pp. 1888–1894
17. Zhou, S., Liu, X., Li, M., et al.: Multiple kernel clustering with neighbor-kernel subspace segmentation. IEEE Trans. Neural Networks Learn. Syst. **31**(4), 1351–1362 (2020)
18. Liu, X., Wang, L., Zhu, X., et al.: Absent Multiple Kernel Learning Algorithms. IEEE Trans. Pattern Anal. Mach. Intell.Intell. **42**(6), 1303–1316 (2020)
19. Zhang, X., Ren, Z., Sun, H., et al.: Multiple kernel low-rank representation-based robust multi-view subspace clustering. Inf. Sci. **551**, 324–340 (2021)

20. Li, Z., Wang, Q., Tao, Z., Gao, Q., Yang, Z.: Deep adversarial multi-view clustering network. In: Twenty-Eighth International Joint Conference on Artificial Intelligence, Macao, China, pp. 2952–2958 (2019)
21. Rai, N., Negi, S., Chaudhury, S., Deshmukh, O.: Partial multi-view clustering using graph regularized NMF. In: 23rd International Conference on Pattern Recognition, Cancún, Mexico, pp. 2192–2197 (2016)
22. Cai, X., Wang, H., Huang, H., Ding, C.H.Q.: Joint stage recognition and anatomical annotation of drosophila gene expression patterns. Bioinform. **28**(12), 16–24 (2012)
23. Jiang, Y., Ye, G., Chang, S., Ellis, D.P.W., Loui, A.C.: Consumer video understanding: a benchmark database and an evaluation of human and machine performance. In: 1st International Conference on Multimedia Retrieval, Trento, Italy, pp. 1–8 (2011)
24. Ng, Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. In: Advances in Neural Information Processing Systems, Vancouver, Canada, pp. 849–856 (2001)
25. Kumar, A., Rai, P., III, H.D.: Co-regularized multi-view spectral clustering. In: Annual Conference on Neural Information Processing Systems, Granada, Spain, pp. 1413–1421 (2011)
26. Luo, S., Zhang, C., Zhang, W., Cao, X.: Consistent and specific multi-view subspace clustering. In: Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, pp. 3730–3737 (2018)
27. Li, X., Zhou, K., Li, C., et al.: Multi-view clustering via neighbor domain correlation learning. Neural Comput. Applic.Applic. **33**, 3403–3415 (2021). https://doi.org/10.1007/s00521-020-05185-y
28. Andrew, G., Arora, R., Bilmes, J.A., Livescu, K.: Deep canonical correlation analysis. In: 30th International Conference on Machine Learning, Atlanta, GA, USA, vol. 28 of JMLR Workshop and Conference Proceedings, pp. 1247–1255 (2013)
29. Abavisani, M., Patel, V.M.: Deep multimodal subspace clustering networks. J. Sel. Topics Signal Processing **12**(6), 1601–1614 (2018)