Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-7, Issue-2, February 2025
DOI: 10.53469/jrse.2025.07(02).10

Enhancing Multiview Subspace Clustering with L_(1,2) Regularization and Self-Labeling Supervision

Qinghao Han[1], Shenglei Pei[1,2,*], Lin Tan[1]

[1]School of Intelligence Science and Engineering, Qinghai Minzu University, Xining, 810007
[2]National Demonstration Center for Experimental Communication Engineering Education,
Qinghai Minzu University, Xining, 810007
**Correspondence Author*

Abstract: ***In recent years, multiview subspace clustering has gained widespread attention due to its ability to effectively integrate complementary information from multiple views, revealing the underlying structure in high-dimensional data. However, existing methods still face challenges in handling complex data scenarios due to their limited representation power. Among these methods, the Multiview Deep Subspace Clustering Network (MvDSCN) has improved clustering performance to some extent by embedding multiview relationships into the feature learning and self-representation stages through the design of a diversity network (Dnet) and a universality network (Unet). However, we observe that the shared representation learned by MvDSCN lacks sufficient discriminative power, which negatively impacts the quality of the self-representation matrix. Furthermore, due to the limitations of its unsupervised learning strategy, the model struggles to effectively leverage latent label information to guide feature learning, thus constraining the improvement in clustering performance. To address these issues, we propose a novel multiview subspace clustering method, L12SL-MvSC, based on L_(1,2)regularization and self-labeling supervision. First, we apply regularization to the self-representation coefficient matrix to select discriminative sample relationships. Then, we introduce a self-labeling supervision strategy, which generates pseudo-labels to assist network training, further enhancing the quality of self-representation learning and clustering performance. Experimental results on benchmark datasets demonstrate the effectiveness of the proposed method.***

Keywords: Multi-view Clustering, Subspace clustering, Self-labeling supervision, Self-expression learning**.**

1. Introduction

In the real world, data is usually described by multiple modalities and feature hierarchies. These modalities may include images, audio, text, etc., and each modality contains multiple feature descriptors. For example, images may include features such as color, texture, and shape, while audio may include spectral features, temporal features, etc. These data from different sources and dimensions together constitute multi-view data [1], which can comprehensively represent the same object or event from multiple perspectives. Due to its rich diversity and information complementarity, multi-view data has significant advantages in many application scenarios. In order to explore the information of multiple views for different tasks, Multi-view Learning (MVL) methods have been widely studied [2]. MVL aims to jointly learn data representations from multiple perspectives in order to explore their shared structures and potential relationships, and has become an important research field [3].

In the study of MVL, Multi-view Clustering (MVC) is an important unsupervised learning method that aims to directly utilize the complementary information from multiple views and automatically mine the underlying clustering structure shared by multiple views without labels [4]. Although existing MVC methods have achieved good performance in many practical tasks, they still face some challenges when dealing with high-dimensional complex data. Many methods directly perform clustering based on original features and fail to effectively mine the potential low-dimensional structure of the data. High-dimensional data often contains a lot of redundant information and noise. Direct clustering in the original feature space will affect the accuracy of the clustering results. Therefore, how to discover the low-dimensional structure of data through dimensionality reduction or subspace learning becomes the key to improving clustering quality.

In order to address these challenges, the Multi-view Subspace Clustering (MVSC) method [5-7] has gradually become a research focus and has received widespread attention. The MVSC method can better reveal the structural characteristics of the data by learning the potential low-dimensional subspace representation of each view and exploring the consistency relationship between views. Compared with traditional MVC, MVSC can not only process high-dimensional data more effectively, but also capture the potential low-dimensional structure of the data, thereby improving the clustering performance. Although the existing multi-view subspace clustering methods have achieved good results in many applications [8-10], their performance still has certain shortcomings when facing complex data, which is mainly reflected in the following two aspects. First, many methods adopt a two-stage strategy, that is, first extracting features from the data and then learning the association matrix. This separate processing method lacks a close connection between feature learning and subspace clustering tasks, resulting in the failure to fully utilize the potential multi-view relationship. Second, existing methods usually ignore end-to-end joint optimization and fail to learn hierarchical representations in the feature learning process, which limits the adaptability and clustering performance of the method on multi-view data.

To address these issues, Multiview Deep Subspace Clustering Networks (MvDSCN) [11] was proposed as an effective improvement framework. MvDSCN introduces a diversity network (Dnet) and a universality network (Unet), and utilizes the complementarity between multi-view data in the feature learning and self-representation stages to better capture the shared and personalized features in the data, thereby improving the clustering performance. However, despite the excellent performance of MvDSCN, we still find that it has the following shortcomings: First, the shared representation learned by MvDSCN lacks sufficient discriminability in some cases, especially when faced with complex data. Second, MvDSCN still relies on unsupervised learning strategies and fails to effectively use label information for precise feature guidance, which further affects the accuracy and stability of clustering.

To solve the above problems, we proposed an improved method based on $\ell_{1,2}$ regularization and self-label supervision strategy. Specifically, this paper introduced $\ell_{1,2}$ -norm regularization in the MvDSCN framework. By constraining the learning process of the self-expression matrix, the model's ability to distinguish samples of different categories is effectively improved, especially on high-dimensional complex data. In addition, this paper designs a self-label supervision strategy that combines spectral clustering to generate pseudo-labels. The generated pseudo-labels are used to further train the network, fully explore the potential structural information of multi-view data, and improve the clustering performance of the model in an unsupervised setting. The main contributions of this paper are summarized as follows:

1) We propose an enhanced multi-view subspace clustering algorithm based on $\ell_{1,2}$ -norm regularization and self-labeling supervision, termed L12SL-MvSC. This method aims to improve the discriminative capability of shared representations and guide network training through latent label information, thereby enhancing clustering performance.

2) We introduce $\ell_{1,2}$-norm regularization into the learning processes of Dnet and Unet, imposing $\ell_{1,2}$-norm constraints on the self-expression matrix to select the most representative sample relationships, which better supports the subsequent spectral clustering process.

3) We adopt a self-label supervision strategy and realize a two-way feedback mechanism between self-expression learning and clustering performance through pseudo-label generation, which improves the clustering adaptability of multi-view data.

2. Related Work

In recent years, most representative algorithms for multi-view subspace clustering (MVSC) have optimized self-representation learning by designing specific regularization strategies to better explore the consistency and diversity between different views [12]. For instance, the Latent Multi-view Subspace Clustering (LMSC) method [6] facilitates clustering by leveraging the latent representations of data points, while simultaneously exploring complementary information derived from various views. The Consistent and Specific Multi-View Subspace Clustering (CSMSC) [13] method learns multi-view self-representation by sharing consistent representation and specific representation, aiming to capture the commonalities and differences between views. Adversarial Multiview Clustering Networks With Adaptive Fusion (AMvC) [22] adopts an adversarial deep learning framework, extracts latent features with the help of a multi-view encoder, and generates reconstructed samples of each view through a generator to enhance the consistency of latent representations between views. This method optimizes view weights through an adaptive fusion layer and combines $\ell_{1,2}$-norm regularization to improve the discriminative ability of shared representations. In addition, the Multiview Subspace Clustering via Low-Rank Symmetric Affinity Graph (LSGMC) [14] method aims to improve clustering performance by combining low-rank consistency and symmetric affinity graphs to make full use of the consistency and angle information in multi-view data. Enhanced Latent Multi-view Subspace Clustering (ELMSC) [15] improves the way of learning multi-view latent representations by constructing an enhanced data matrix. This method stacks the data matrices of different views to the block diagonal position of the enhanced matrix to better mine the complementary information between multiple views. At the same time, consistency information is captured by the non-block diagonal parts based on the similarity between views. Multi-view Subspace Clustering for Learning Joint Representation via Low-rank Sparse Representation (MSCLR) [16] combines low-rank sparse representation under the framework of consistency and specificity representation to mine the shared structure between views. This method uses ℓ_1 norm and frobenius norm to promote a sparser representation while preserving the geometric structure and ensuring the grouping effect. In addition to optimizing self-representation learning through regularization strategies to mine the consistency and diversity between views, researchers have further focused on the applicability of multi-view subspace clustering algorithms in large-scale data scenarios, conducted in-depth explorations around their scalability, and proposed a variety of optimization methods [17-20]. Overall, The process of most multi-view subspace clustering algorithms can usually be divided into two main steps. First, researchers design specific regularization techniques to learn a self-representative affinity matrix from multi-view data, aiming to capture the shared information and potential structure between different views. The core of this step is to reveal the intrinsic characteristics of the data by maximizing the consistency between different views while maintaining the diversity of each view. Secondly, the obtained affinity matrix is usually input into the traditional spectral clustering algorithm for the final clustering analysis. Spectral clustering methods group data points by calculating the eigenvectors of the affinity matrix to complete the clustering task.

3. Proposed Method

In this section, we will introduce the proposed L12SL-MvSC method in detail, and its overall framework is shown in Figure 1. Drawing on and improving the network architecture of MvDSCN, our framework consists of two sub-networks, Dnet and Unet, and a self-label supervision module. Among them, Dnet aims to extract the self-representation information unique to each view to fully capture the diversity characteristics in multi-view data; Unet is responsible for learning the self-representation matrix shared across views to characterize the common structure of the data. The self-label supervision module further uses the generated pseudo-labels to optimize the latent representation and enhance the consistency and robustness of the clustering. This design enables the modules to work together to explore and fuse the potential characteristics of multi-view data from multiple angles.

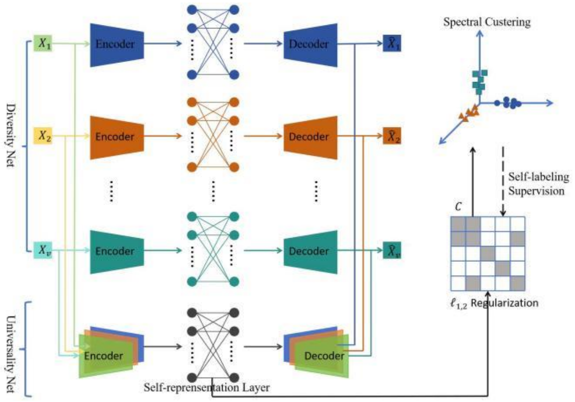

Figure 1: Framework of the proposed L12SL-MvSC method. $\hat{X}_i$ includes the data $\hat{X}_i^d$ and $\hat{X}_i^u$, decoded by the specific decoder of the i-th view and the shared decoder, respectively.

3.1 Dnet

The Dnet aims to capture the unique characteristics of each view in multi-view data. By focusing on view-specific self-representation learning, it ensures the diversity between different views and fully utilizes the potential information of multi-view data. To this end, Dnet designs an independent encoder-decoder structure for each view, where the encoder extracts the latent features of the input view and the decoder uses these features to reconstruct the input data. This design not only ensures the integrity of the unique information of each view, but also effectively reduces the interference of shared or redundant features. At the same time, drawing on the work of reference [11], Dnet uses a convolutional autoencoder, whose convolutional layer has fewer parameters and stronger learning ability, and can extract more detailed features while improving computational efficiency.

In order to improve the lack of discriminability of latent features and improve the quality of the self-expressive matrix, we introduce $\ell_{1,2}$-norm regularization constraint on the self-expressive matrix. This regularization has been proven to be effective in previous work [21,22], which can effectively highlight the key information unique to each view and reduce the interference of redundant features, thereby improving the accuracy of view feature expression.

On this basis, the objective function of Dnet combines reconstruction loss, self-expression loss and $\ell_{1,2}$-norm regularization constraints, and is defined as follows:

$$\mathcal{L}_D = \sum_{i=1}^{v} \left\| X_i - \hat{X}_i^d \right\|_F^2 + \lambda_1 \left\| Z_i^d - Z_i^d C_i \right\|_F^2 + \lambda_2 \| C_i \|_{1,2}. \quad (1)$$

Among them, X_i represents the i-th view of the input data, v represents the total number of views, $\hat{X}_i^d$ represents the data decoded by the specific decoder of the i-th view, Z_i^d signifies the latent feature representation obtained through the specific encoder of the i-th view, and C_i represents the self-expression matrix of the i-th view.

3.2 Unet

The Unet aims to capture the shared characteristics of multi-view data to learn a universal self-representation matrix that can characterize the consistency between views. To achieve this goal, Unet is used as a core module in the overall framework to build a consistent representation across views by jointly learning latent features from different views. Specifically, Unet adopts the same convolutional encoder-decoder architecture as the Dnet network to characterize the global pattern of multi-view data by extracting shared latent features. The self-representation matrices of different views are aligned with the shared common self-representation matrix through a universal regularization method to ensure the consistency of each view in the same latent space.

In order to achieve the above goals, the optimization process of Unet is constrained by the following objective function:

$$\mathcal{L}_U = \sum_{i=1}^{v} \left\| X_i - \hat{X}_i^u \right\|_F^2 + +\lambda_1 \| Z_i^u - Z_i^u C \|_F^2 + \lambda_2 \| C \|_{1,2} + \lambda_3 \| C - C_i \|_F^2. \quad (2)$$

Among them, $\hat{X}_i^u$ represents the data decoded by the consistent decoder of the i-th view, Z_i^u represents the potential feature representation obtained by the consistent encoder of the i-th view, and C represents the common self-expressive matrix shared by all views. The first term is the reconstruction loss, the second term is the self-expressive loss, and the third term is the constraint on the shared self-expressive matrix. In addition, the fourth term is the consistency constraint, which aims to ensure the consistency between the self-expressive matrix of each view C_i and the shared self-expressive matrix C. The hyperparameters λ_1, λ_2 and λ_3 are used to balance the trade-offs among the respective loss terms, thereby controlling the equilibrium between different objectives.

On this basis, in order to fully exploit the complementary information from multiple views, we introduce diversity regularization consistent with MvDSCN. Specifically, the form of diversity regularization is as follows:

$$\mathcal{L}_H = \sum_{ij} HSIC(Y, C_i, C_j). \quad (3)$$

Here, C_i and C_j denote the self-expression matrices of the i-th and j-th views, respectively. $Y \in C_i \times C_j$. $HSIC(\cdot)$ is used to measure the high-order correlation between variables. The detailed calculation method can be found in reference [11]. By introducing diversity regularization into the objective function, we can penalize the dependency between views, thereby more effectively capturing the complementary information between views.

3.2 Self-labeling Supervision Module

The self-labeling supervision module aims to further improve the quality of the self-expressive matrix by generating pseudo labels, thereby enhancing the consistency and robustness of clustering. Inspired by [23], after obtaining a consistent self-expressive matrix, pseudo labels can be generated through spectral clustering. In this paper, with the help of the universal self-expressive matrix captured by Unet, we obtain the latent label information through spectral clustering as feedback for subsequent optimization to guide the robust learning of the self-expressive matrix.

In order to effectively integrate the self-labeling supervision module into the optimization process, the objective function is defined as follows:

$$\mathcal{L}_S = tr(Q^T L Q). \quad (4)$$

Here, L is the Laplacian matrix of C, $Q \in R^{n\times k}$ is a cluster indicator matrix containing n samples and k clusters, representing the pseudo-labels generated by C, and $tr(\cdot)$ denotes the trace operation of the matrix.

In order to effectively integrate these modules for joint optimization, we combine formulas (1), (2), (3) and (4) to derive the overall objective function, which combines the loss terms of Dnet, Unet and self-label supervision modules. The specific definition is as follows:

$$\begin{aligned}\mathcal{L}_{all} = \textstyle\sum_{i=1}^{v} \left\|X_i - \hat{X}_i^d\right\|_F^2 + \left\|X_i - \hat{X}_i^u\right\|_F^2 + \lambda_1\left(\left\|Z_i^d - Z_i^d C_i\right\|_F^2 + \|Z_i^u - Z_i^u C\|_F^2\right) + \\ \textstyle\sum_{i=1}^{v} \lambda_2\left(\|C_i\|_{1,2} + \|C\|_{1,2}\right) + \lambda_3\|C - C_i\|_F^2 + \\ \lambda_4 \textstyle\sum_{ij} HSIC\left(Y, C_i, C_j\right) + \lambda_5 tr(Q^T L Q), s.t. Q^T Q = I.\end{aligned} \quad (5)$$

4. Experiments

In this section, we experimentally evaluate the performance and advantages of the proposed L12SL-MvSC method on the multi-view subspace clustering task.

4.1 Experiment Setup

Datasets: To evaluate the clustering effect of the L12SL-MvSC algorithm, two multi-view public benchmark datasets and a real-world RGB-D Object dataset [23] are used in the experiment. Table 1 shows the detailed information of each dataset.

1) ORL: This dataset consists of images from 40 individuals, with each individual represented by 10 images. The images were captured under varying lighting conditions and include different facial expressions and details.

2) Still DB: This dataset contains 467 images from six action categories. Each image is described by three types of features: SIFT Bag-of-Words (BoWs), color SIFT BoWs, and shape context BoWs.

3) RGB-D Object: This dataset includes RGB and depth images of 300 objects grouped into 51 categories. For the experiments, 50 categories were randomly selected, with each category containing 10 samples. All images were resized to 64×64 pixels for consistency.

Table 1: Statistics of the three benchmark multi-view datasets.

Dataset	Samples	Class	View	Dimension
ORL	400	40	3	4096/3304/6750
Still DB	467	6	3	200/200/200
RGB-D Object	500	50	2	4096/4096

Comparison Methods: In our experiments, we selected several representative single-view and multi-view methods as baselines for comparison to comprehensively evaluate the effectiveness of the proposed method.

1) BestSV [24]. This method is based on single view data and uses the standard spectral clustering algorithm to achieve the best clustering performance on a single view.

2) LRR [25]. This method uses low-rank representation to restore the subspace structure of the data and obtains a more stable self-expression matrix through global low-rank constraints.

3) LMSC [6]. This method integrates multi-view information through latent space mapping and learns a shared self-expression matrix in a common latent subspace, thereby improving the clustering effect.

4) SURE [26]. This method adopts a contrastive learning strategy, taking available sample pairs as positive samples and randomly selecting some cross-view samples as negative samples. At the same time, it introduces noise-resistant contrast loss to alleviate the impact of erroneous negative samples caused by random sampling.

5) MvDSCN [11]. This method adopts a deep self-expressive learning framework, which consists of two sub-networks, Dnet and Unet. Dnet learns a view-specific self-expressive matrix, while Unet learns a self-expressive matrix shared by all views to simultaneously capture the personalized information and global structure of multiple views.

Evaluation Metrics: Following [11], we adopt four widely used metrics to assess clustering performance: accuracy (ACC), normalized mutual information (NMI), adjusted Rand index (ARI), and F-measure. To reduce the influence of randomness, each method is executed 15 times, and the average performance is reported.

4.2 Clustering Result

Tables 2-4 report the clustering results of each method on three benchmark datasets. It can be seen that our method achieves the best performance among almost all the compared methods. For ease of comparison, the best results for each indicator are highlighted in bold in the table. Specifically, the performance of multi-view methods is significantly improved compared with single-view methods, mainly due to the fact that multi-view methods can effectively utilize the complementary information between views, which cannot be fully captured by the limitation of single-view methods. In addition, for the traditional multi-view method LMSC, our method significantly surpasses the performance of LMSC through a deep learning framework. Although LMSC effectively integrates multi-view information, it relies on predefined handcrafted features and fails to fully exploit the potential of deep learning models. Our L12SL-MvSC method significantly outperforms LMSC on all datasets, indicating that deep learning is better able to capture latent structures and improve clustering performance when dealing with complex datasets. Comparisons with state-of-the-art methods, such as SURE and MvDSCN, further highlight the improvements achieved by our approach. Although both SURE and MvDSCN adopt the deep learning framework for MVC, our method achieves further improvement in performance by introducing key innovations. For example, on the RGB-D Object dataset, compared with MvDSCN, our method improves ACC, NMI, ARI and F-measure by 2.4%, 0.7%, 1.6% and 1.1% respectively. MvDSCN introduces a complex network structure, embeds multi-view relations into feature learning through two sub-networks, and achieves good performance. However, the unsupervised nature and lack of discriminative feature learning limit its further optimization. On the ORL dataset, although MvDSCN slightly outperforms L12SL-MvSC in NMI and F-measure, the improvements of our method on all metrics such as accuracy and NMI are more balanced, which indicates that our method exhibits more stable and reliable performance on different datasets.

Experimental results demonstrate the effectiveness of L12SL-MvSC in multi-view subspace clustering tasks. Our method demonstrates excellent performance on multiple benchmark datasets, mainly due to two key innovations. First, by introducing $\ell_{1,2}$ -norm regularization constraints, the quality of the self-expression matrix is enhanced, which helps the model to distinguish samples of different categories more effectively. Second, the self-label supervision strategy that combines pseudo-label generation with multi-view relationship learning enables the model to fully explore the potential structural information of the data.

Table 12: Clustering performance on ORL dataset.

Method	ACC	NMI	ARI	F-measure
BestSV	0.777	0.903	0.738	0.711
LRR	0.773	0.895	0.724	0.731
LMSC	0.819	0.931	0.769	0.758
SURE	0.843	0.925	0.791	0.788
MvDSCN	0.870	**0.943**	0.819	**0.834**
L12SL-MvSC	**0.883**	0.934	**0.821**	0.825

Table 3: Clustering performance on Still DB dataset.

Method	ACC	NMI	ARI	F-measure
BestSV	0.297	0.104	0.063	0.221
LRR	0.306	0.109	0.066	0.260
LMSC	0.328	0.137	0.088	0.269
SURE	0.363	0.227	0.142	0.296
MvDSCN	0.377	0.245	0.169	0.320
L12SL-MvSC	**0.405**	**0.519**	**0.272**	**0.437**

Table 4: Clustering performance on RGB-D Object dataset.

Method	ACC	NMI	ARI	F-measure
BestSV	0.278	0.554	0.106	0.125
LRR	0.299	0.589	0.143	0.156
LMSC	0.335	0.593	0.151	0.167
SURE	0.373	0.622	0.202	0.228
MvDSCN	0.388	0.639	0.210	0.225
L12SL-MvSC	**0.412**	**0.646**	**0.226**	**0.236**

5. Conclusion

This paper proposes an enhanced multi-view subspace clustering method (L12SL-MvSC) integrating $\ell_{1,2}$ -norm regularization and self-labeling supervision strategies, designed to optimize the quality of the self-expression matrix and significantly improve clustering performance. By introducing $\ell_{1,2}$ -norm regularization constraints, we can effectively select the most representative sample relations, thereby improving the quality of the self-expression matrix and enhancing the discriminative ability of feature learning. At the same time, combined with the self-label supervision strategy for pseudo-label generation, we implement a two-way feedback mechanism between self-expression learning and clustering performance, further improving the adaptability of multi-view data in clustering tasks. Experimental results show that L12SL-MvSC outperforms existing multi-view subspace clustering methods on multiple benchmark datasets, especially showing stronger clustering capabilities when processing high-dimensional complex data. Despite significant performance gains, there is still room for improvement. Future research can explore how to further optimize the self-label generation strategy to improve the efficiency and robustness of the model when processing large-scale datasets. In addition, combining other supervised learning methods and optimization techniques may further improve the clustering accuracy of the model.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant Nos. 62266035.

References

[1] Si X, Yin Q, Zhao X, et al. Consistent and diverse multi-view subspace clustering with structure constraint[J]. Pattern Recognition, 2022, 121: 108196.
[2] Wen J, Zhang Z, Fei L, et al. A survey on incomplete multiview clustering[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 53(2): 1136-1149.
[3] Yan X, Hu S, Mao Y, et al. Deep Multi-view Learning methods: A review[J]. Neurocomputing, 2021, 448: 106-129.
[4] Zhou L, Du G, Lü K, et al. A Survey and an Empirical Evaluation of Multi-view Clustering Approaches[J]. ACM Computing Surveys, 2024, 56(7): 1-38.
[5] Gao H, Nie F, Li X, et al. Multi-view subspace clustering[C]. Proceedings of the IEEE international conference on computer vision. 2015: 4238-4246.
[6] Zhang C, Hu Q, Fu H, et al. Latent multi-view subspace clustering[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4279-4287.
[7] Zhang C, Fu H, Hu Q, et al. Generalized latent multi-view subspace clustering[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 42(1): 86-99.
[8] Chen P, Liu L, Ma Z, et al. Smoothed multi-view subspace clustering[C]. Neural Computing for Advanced Applications: Second International Conference, 2021: 128-140.
[9] Liu Q, Lu Z, Huan W, et al. Adaptive weighted multi-view subspace clustering method for recognizing urban functions from multi-source social sensing data[J]. Geo-spatial Information Science, 2024: 1-25.

[10] Fraj M, HajKacem M A B, Essoussi N. Multi-view subspace text clustering[J]. Journal of Intelligent Information Systems, 2024: 1-24.
[11] Zhu P, Yao X, Wang Y, et al. Multiview Deep Subspace Clustering Networks[J]. IEEE Transactions on Cybernetics, 2024.
[12] Si X, Yin Q, Zhao X, et al. Consistent and diverse multi-view subspace clustering with structure constraint[J]. Pattern Recognition, 2022, 121: 108196.
[13] Luo S, Zhang C, Zhang W, et al. Consistent and specific multi-view subspace clustering[C]. Proceedings of the AAAI conference on artificial intelligence. 2018, 32(1).
[14] Lan W, Yang T, Chen Q, et al. Multiview subspace clustering via low-rank symmetric affinity graph[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023.
[15] Shi L, Cao L, Wang J, et al. Enhanced latent multi-view subspace clustering[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024.
[16] Khan G A, Hu J, Li T, et al. Multi-view subspace clustering for learning joint representation via low-rank sparse representation[J]. Applied Intelligence, 2023, 53(19): 22511-22530.
[17] Kang Z, Zhou W, Zhao Z, et al. Large-scale multi-view subspace clustering in linear time[C]. Proceedings of the AAAI conference on artificial intelligence. 2020, 34(04): 4412-4419.
[18] Sun M, Zhang P, Wang S, et al. Scalable multi-view subspace clustering with unified anchors[C]. Proceedings of the 29th ACM international conference on multimedia. 2021: 3528-3536.
[19] Chen M S, Wang C D, Huang D, et al. Efficient orthogonal multi-view subspace clustering[C]. Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining. 2022: 127-135.
[20] Liu S, Wang S, Zhang P, et al. Efficient one-pass multi-view subspace clustering with consensus anchors[C]. Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(7): 7576-7584.
[21] Ming D, Ding C. Robust flexible feature selection via exclusive L21 regularization[C]. Proceedings of the 28th international joint conference on artificial intelligence. 2019: 3158-3164.
[22] Wang Q, Tao Z, **a W, et al. Adversarial multiview clustering networks with adaptive fusion[J]. IEEE transactions on neural networks and learning systems, 2022, 34(10): 7635-7647.
[23] Lai K, Bo L, Ren X, et al. A large-scale hierarchical multi-view rgb-d object dataset[C]. 2011 IEEE international conference on robotics and automation. IEEE, 2011: 1817-1824.\
[24] Ng A, Jordan M, Weiss Y. On spectral clustering: Analysis and an algorithm[J]. Advances in neural information processing systems, 2001, 14.
[25] Liu G, Lin Z, Yan S, et al. Robust recovery of subspace structures by low-rank representation[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 171-184.
[26] Yang M, Li Y, Hu P, et al. Robust Multi-view Clustering with incomplete information[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 1055-1069.