Multi-Level Structural Contrastive Subspace Clustering Network

Peipei Zhang, Wenjie Zhu, *Member, IEEE,* and Wei Qi Yan, *Senior Member, IEEE*

***Abstract*—Deep subspace clustering methods based on autoencoder (AE) have achieved impressive performance in various applications. However, these methods often place excessive reliance on the AE framework, which focuses primarily on pixel-level reconstruction while overlooking the structural information inherent in the data. To overcome this limitation, we propose a novel approach called the Multi-level Structural Contrastive Subspace Clustering Network (MSCSCN). Unlike traditional AE-based methods, MSCSCN departs from the AE paradigm and introduces multi-level contrastive prediction to improve feature learning. Specifically, MSCSCN integrates multi-level features from both original and augmented data within a self-expression learning process, enhancing the learned pairwise affinities. Additionally, we propose a structural contrastive loss, which strengthens cluster boundary discrimination by effectively utilizing pairwise affinities and structural information. Our experimental results on several benchmark datasets demonstrate that MSCSCN outperforms competitive deep subspace clustering methods, highlighting its superior capability in improving clustering performance and capturing the underlying structural information within the data.**

***Index Terms*—Deep subspace clustering, Self-expression matrix, Multi-level feature, Structural information.**

I. Introduction

SUBSPACE clustering aims to discover the intrinsic structure of the data in an unsupervised manner, segmenting data points by effectively representing high-dimensional data as a union of low-dimensional subspaces, thus addressing the challenges associated with high dimensionality. Subspace clustering has been successfully applied in various areas, including motion segmentation [1, 2], face clustering [3, 4], and bioinformatics [5, 6].

Subspace clustering methods leverage the self-expression property through sparse [7, 8, 9, 10, 11] or low-rank [12, 13, 14, 15] constraints to learn pairwise affinities. These methods assume that data points lie within linear subspaces, where points within the same subspace can be linearly represented by each other. However, this assumption often fails in real-world data. To address this, recent deep subspace clustering methods [16, 17, 18, 19, 20] utilize deep neural networks as nonlinear mapping functions to better handle nonlinearity, achieving significantly improved clustering performance. One pioneering approach, the Deep Subspace Clustering Network (DSC-Net) [21], integrates an AE with a self-expression layer, preserving intermediate representation information while modeling the self-expression property. Since then, numerous variants of DSC-Net have emerged, addressing various aspects such as adversarial training [17, 22], regularization constraints [19, 20, 23], and self-supervised learning [18, 24]. As research advances, there has been growing attention on exploring additional attributes of the self-expression matrix, such as structural refinement [25], the double self-expression learning technique [26], and more.

Although DSC-Net and its variants have showed remarkable clustering performance, they face some limitations: 1) Most approaches focus primarily on learning high-level features and the self-expression matrix, while neglecting the learning of low-level features and the structural information inherent in the data. 2) AE-based representation learning may not align with the objectives of subspace clustering, which prioritizes discovering discriminative features for cluster boundary identification rather than pixel-level image reconstruction.

To address these issues, we propose a novel deep subspace clustering framework, the Multi-level Structural Contrastive Subspace Clustering Network (MSCSCN). Moving away from the traditional AE architecture, our approach emphasizes the use of structural information to learn clustering-guided representations. Specifically, we abandon the AE paradigm and instead train the network with a multi-level contrastive loss. Both low-level and high-level features are incorporated into the self-expression learning process to strengthen subspace-preserving properties. To further enhance cluster boundary discrimination, we introduce a structural contrastive loss that captures clustering-guided embeddings by leveraging structured information through pairwise affinities. The key contributions are summarized as follows:

- We propose a novel deep subspace clustering framework, MSCSCN, which departs from the AE paradigm by integrating multi-level contrastive learning with both raw and augmented data into the self-expression learning process, thereby enhancing subspace-preserving properties.
- We introduce a structural contrastive loss to capture clustering-guided embeddings, improving cluster boundary discrimination by leveraging structural information and pairwise affinities.

This work was supported in part by the Natural Science Foundation of Zhejiang Province of China under Grant LY24F030005, in part by National Key Research and Development Program of China under Grant 2021YFC3340402, and in part by the Fundamental Research Funds for the Provincial Universities of Zhejiang under Grant 2022YW40. (Corresponding author: Wenjie Zhu.)

Peipei Zhang and Wenjie Zhu are with the College of Information Engineering, China Jiliang University, Hangzhou 310018, China (e-mail: zhwj@cjlu.edu.cn).

Wei Qi Yan is with the School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand (e-mail: weiqi.yan@aut.ac.nz).

II. METHODOLOGY

In this section, we present the MSCSCN framework, as illustrated in Fig. 1. The framework comprises three modules: multi-level feature extraction, low-rank self-expression learning, and multi-level structural contrastive prediction, each of which is explained in detail below.

A. Multi-Level Feature Extraction

To effectively extract local features while preserving spatial locality, we employ a multi-level convolutional neural network consisting of several convolutional layers. Let $f(\cdot)$ represent the convolutional encoder shared by both the original and augmented data, with a total of L layers. $f_l(\cdot)$ denotes the network at layer l. Given the data points $X = \{x_i \mid i = 1, 2, \cdots, N\} \in \mathbb{R}^{d\times N}$, where d is the feature dimension and N is the number of samples, we apply a random transformation to generate augmented data $\hat{X} = \{\hat{x}_i \mid i = 1, 2, \cdots, N\} \in \mathbb{R}^{d\times N}$. Consequently, we obtain multi-level features $Z_l = f_l(X) \in \mathbb{R}^{d_l\times N}$ and $\hat{Z}_l = f_l(\hat{X}) \in \mathbb{R}^{d_l\times N}$, where d_l represents the feature dimension at layer l. We define Z_1 to Z_{L-1} as the low-level features and Z_L as the high-level features. By leveraging both of them, our approach ensures a more comprehensive feature learning process that benefits subspace clustering.

B. Low-Rank Self-Expression Learning

Deep subspace clustering methods leverage the self-expression property, which posits that a data point within a subspace can be represented as a linear combination of other points in the same subspace. To improve the model's ability to capture the global structure of the data, we employ a multi-level low-rank self-expression loss, defined as follows:

$$\mathcal{L}_{\text{SE}} = \gamma \sum_{l=1}^{L} \left\| \left[Z_l; \hat{Z}_l\right] - \left[Z_l; \hat{Z}_l\right] SS^\top \right\|_F^2 + \lambda \left\| SS^\top \right\|_F^2, \tag{1}$$

where $\left[Z_l; \hat{Z}_l\right] \in \mathbb{R}^{2d_l\times N}$, $S \in \mathbb{R}^{N\times m}$, and m is the parameter that controls the rank of C. The self-expression matrix is represented as a symmetric matrix of the form $C = SS^\top$. Given that the rank of S is constrained by the smaller of the two dimensions m and N, we set m to be significantly smaller than N. As a result, the rank of C is limited, and the number of parameters is greatly reduced. Based on the observation that the core structure of the data remains consistent despite minor transformations, we assert that data correlations should remain invariant under such transformations, and the manifold structure should be preserved. Notably, in Eq. (1), both the original and augmented data contribute to the construction of the similarity graph, enhancing the learned pairwise affinities.

C. Multi-Level Structural Contrastive Prediction

Data augmentation is applied to the data, resulting in $2N$ points. For each data point x_i, we select points with high-confidence affinities to form positive pairs, while the remaining $2N - 2$ points are treated as negative pairs. These points are then passed through two stacked MLP layers with ReLU activations, which nonlinearly project the representations. The

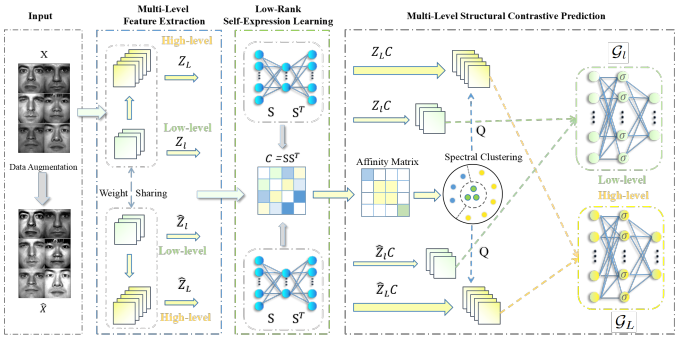


Fig. 1: An overview of the proposed MSCSCN framework, including multi-level feature extraction, low-rank self-expression learning, and multi-level structural contrastive prediction.

features extracted by each layer correspond to a projection head, leading to a set of multi-level projection heads $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_L\}$, where $\mathcal{G}_l(\cdot) \cdots \mathcal{G}_{L-1}(\cdot)$ represent the low-level projection heads, and $\mathcal{G}_L(\cdot)$ serves as the high-level projection head.

To incorporate each feature layer into self-expression learning, we apply self-expression mapping to obtain the representations Z_lC, which are then projected through $\mathcal{G}$. This results in $H_l = \left[h_1^l, h_2^l, \cdots h_N^l\right] = \mathcal{G}_l(Z_lC) \in \mathbb{R}^{D_l\times N}$ and $\hat{H}_l = \left[\hat{h}_1^l, \hat{h}_2^l, \cdots \hat{h}_N^l\right] = \mathcal{G}_l(\hat{Z}_lC) \in \mathbb{R}^{D_l\times N}$, where D_l stands for the dimension of $\mathcal{G}_l(\cdot)$, $l = 1. \cdots, L$. For clarity, we define the probability $P_l(x_j \mid x_i)$ as the likelihood that x_j lies within the same subspace as the given data point x_i:

$$P_l(x_j \mid x_i) = \frac{sim\left(h_i^l, h_j^l\right)}{\sum_{k=1}^{N} sim\left(h_i^l, h_k^l\right) + sim\left(h_i^l, \hat{h}_k^l\right)}. \tag{2}$$

In Eq. (2), $sim(a, b)$ represents the pairwise similarity, i.e., $sim(a, b) = \exp\left(\frac{a^\top b}{\|a\|\|b\|}/\sigma\right)$, where a, b denote two feature vectors and σ represents the temperature factor.

Consequently, the multi-level contrastive loss $\mathcal{L}_{\text{MC}}$ can be formulated as the sum of the negative log-likelihoods for all data points in the dataset. Let $\ell_l(x_i, x_j) = -\log P_l(x_j \mid x_i)$, we have

$$\mathcal{L}_{\text{MC}} = \frac{1}{2N} \sum_{l=1}^{L} \sum_{i=1}^{N} \left(\ell_l(x_i, \hat{x}_i) + \ell_l(\hat{x}_i, x_i)\right). \tag{3}$$

Since previous methods often overlook the geometric structure among samples, their ability to capture data point similarities is constrained. To address this issue, we propose a structural contrastive loss, $\mathcal{L}_{\text{SC}}$, leveraging high-level features:

$$\mathcal{L}_{\text{SC}} = \frac{1}{4|Q|} \sum_{\substack{i\neq j \\ Q_{ij}=1}}^{N} \begin{Bmatrix} \ell_L(x_i, x_j) + \ell_L(x_i, \hat{x}_j) \\ +\ell_L(\hat{x}_i, x_j) + \ell_L(\hat{x}_i, \hat{x}_j) \end{Bmatrix}. \tag{4}$$

In Eq. (4), we apply the spectral clustering algorithm to obtain the pseudo-labels from the previous iteration and construct the structural matrix $Q \in \{0, 1\}^{N\times N}$, where $|Q|$ represents the total number of ones. $Q_{ij} = 1$ if x_i and x_j belong to the same subspace, in which case they are treated as a positive pair, otherwise, $Q_{ij} = 0$. The structural matrix Q quantifies

the similarities between data samples, and is updated every T epochs in our method.

Unlike other contrastive subspace clustering methods [26, 27], our method builds positive and negative pairs via data augmentation and the structural matrix. Instead of directly operating on the self-expression matrix, our loss is applied to features and does not rely on an AE-based framework.

By incorporating Eq. (3), the framework leverages the complementary strengths of both low-level and high-level features, thereby enhancing its ability to capture high-quality feature representations. Eq. (4) enables the framework to identify intra-cluster structure, thereby facilitating the extraction of clustering-guided embeddings and subspace learning.

D. Overall Objective

The overall objective consists of the multi-level structural contrastive loss and the self-expression loss is as:

$$\mathcal{L} = \alpha \mathcal{L}_{\text{MC}} + \mathcal{L}_{\text{SC}} + \mathcal{L}_{\text{SE}}, \tag{5}$$

where α is the trade-off parameter. Our training strategy consists of two phases. First, we initialize $Q = I$, and pretrain the network using Eqs. (1) and (3). We save the final clustering assignment to initialize Q. Then we fine-tune the framework by Eq. (5). For clarity, the procedure of our framework is specified in Algorithm 1.

Algorithm 1 Procedure for MSCSCN framework

Input: Input data X, augmentation strategies, tradeoff parameters λ, α, γ, temperature factor $\sigma = 0.1$, maximum iteration M, T.
Initialization: $Q = I$.
Pretrain the network via Eqs. (1) and (3).
for $Iter$ = 1 to M **do**
  Fine-tune the network via Eq. (5).
  if $Iter\%T == 0$ **then**
    Run spectral clustering to update Q.
  end if
end for
Output: data segementation.

III. Experiments

A. Experiment Setting

We evaluate our framework through extensive experiments on benchmark datasets, including ORL[1], EYaleB[2], COIL20[3], MNIST[4], UMIST[5] and COIL100[6]. Table I provides their detailed descriptions. Then, we compare our method with benchmark subspace clustering approaches including SSC [28], ENSC [29], KSSC [30], SSC-OMP [31], EDSC [32], LRR [33], LRSC [12], DSC-Net [21], AE+SSC [21], DASC [17], SSConSCN [18], AASSC-Net [25], LDLRSC [20], DSESC [26] and EDS-SC[23]. As shown in Table II, consistent with previous deep subspace clustering methods, we adopt the same network architecture as their encoders to ensure fair comparisons with prior approaches. However, for COIL20 and COIL100, a single-layer encoder is unsuitable for our method. Therefore, we employ a multi-layer architecture. In all experiments, the learning rate is set to 1.0 $\times 10^{-3}$, and the temperature factor is set to 0.1. We use Adam as the optimizer and set $T \in \{3, 4, 5, 6\}$ for all experiments. The output dimension of the projection head is 64 for ORL, UMIST, and MNIST, 128 for COIL20 and COIL100, and 256 for EYaleB, with consistent settings across different levels within each dataset. Additionally, we select the parameters at intervals of one order of magnitude, ranging from 10^{-4} to 10^{2}. The specific hyperparameters are listed in the final column of Table II. We explore various data augmentation techniques for our framework, including random horizontal flipping across all six datasets, and dataset-specific augmentations: contrast and random Gaussian blur for ORL, brightness and random Gaussian blur for UMIST, brightness and contrast for EYaleB and MNIST, and random rotation for COIL20 and COIL100.

TABLE I
Summary of Datasets

Datasets	ORL	EYaleB	COIL20	MNIST	UMIST	COIL100
Number of samples	400	2432	1440	1000	480	7200
Number of classes	40	38	20	10	20	100
Image size	28×28	48×42	32×32	32×32	32×32	32×32

TABLE II
Encoder and Hyperparameter Settings of Datasets

Datasets	Kernel Size	Channels	Hyperparameters $\{\alpha, \lambda, \gamma\}$
ORL	(3×3, 3×3, 3×3)	(5, 3, 3)	$\{10, 10^{-1}, 10^{-3}\}$
EYaleB	(5×5, 3×3, 3×3)	(10, 20, 30)	$\{10^{-2}, 10^{-2}, 10^{-4}\}$
COIL20	(3×3, 3×3)	(5, 10)	$\{10, 10^{-3}, 10^{-4}\}$
MNIST	(5×5, 3×3, 3×3)	(10, 20, 30)	$\{10^{-1}, 10^{-3}, 10^{-1}\}$
UMIST	(5×5, 3×3, 3×3)	(15, 10, 5)	$\{10^{-1}, 10^{-3}, 10^{-1}\}$
COIL100	(3×3, 3×3)	(20, 30)	$\{10, 10^{-1}, 10^{-3}\}$

B. Experiment Results and Discussion

Table III summarizes the compared results in terms of ACC and NMI across various datasets using different algorithms, leading to the following key observations and conclusions:

- Traditional subspace clustering methods exhibit suboptimal performance, underscoring the significant advantages of deep representations. For instance, on the EYaleB dataset, our proposed method outperforms the best conventional approach by a notable margin, achieving a 14.48% improvement in ACC and a 12.92% improvement in NMI. This highlights the importance of leveraging deep learning techniques for subspace clustering tasks.
- The incorporation of structural information plays a crucial role in improving clustering performance, as demonstrated by AASSC-Net. However, our method leverages this structural information more effectively, leading to superior results. On the MNIST dataset, our approach outperforms AASSC-Net by 8.10% in ACC and 11.82% in NMI, showing that the integration of structural cues significantly enhances the clustering process.
- Our proposed MSCSCN method consistently achieves the best performance across all datasets. For example, on the UMIST dataset, our method improves ACC by

[1]http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[2]http://vision.ucsd.edu/extyaleb/CroppedYaleBZip/CroppedYale.zip
[3]http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php
[4]http://yann.lecun.com/exdb/mnist
[5]http://eprints.lincoln.ac.uk/id/eprint/16081/
[6]http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php

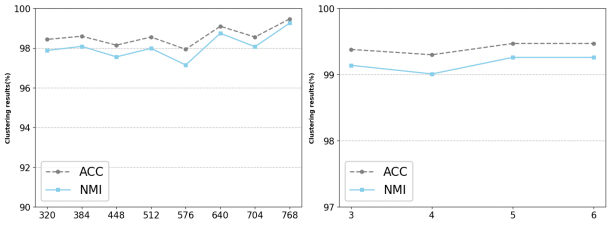
TABLE III
CLUSTERING RESULTS IN TERMS OF ACC AND NMI ON ORL, EYALEB, COIL20, MNIST, UMIST AND COIL100 DATASETS

Datasets	ORL		EYaleB		COIL20		MNIST		UMIST		COIL100	
Methods	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
LRSC [12]	67.75	80.82	67.43	77.19	74.16	84.52	51.40	55.76	67.29	74.98	49.33	58.10
LRR [33]	81.00	86.03	84.99	86.36	81.18	87.47	69.79	76.30	53.86	56.32	40.18	47.21
SSC [28]	67.00	82.06	62.54	86.04	86.31	88.92	45.30	47.09	69.04	74.89	55.10	58.41
EnSC [29]	69.25	82.20	61.14	82.43	87.60	89.52	49.83	54.95	69.31	75.69	57.32	59.24
EDSC [32]	70.38	77.99	84.99	86.36	83.71	88.28	56.50	57.52	69.37	75.22	61.87	67.51
AE+SSC [21]	74.50	88.24	74.75	77.64	87.11	89.90	48.40	53.37	70.42	75.15	56.07	58.71
SSC-OMP [31]	71.00	79.52	73.72	78.03	64.10	74.12	34.00	32.72	64.38	70.68	32.71	—
KSSC [30]	71.43	80.70	69.21	73.59	70.87	82.43	52.20	56.23	65.31	73.77	52.82	60.47
DSCNet [21]	86.00	90.34	97.33	97.03	94.86	94.08	75.00	73.19	73.12	76.62	69.04	70.15
DASC [17]	88.25	93.15	98.56	98.01	96.39	96.86	80.40	78.00	76.88	80.42	72.15	72.86
SSConSCN [18]	89.50	—	98.48	—	97.86	—	—	—	—	—	73.33	—
AASSC-Net [25]	90.75	94.31	—	—	98.40	98.29	84.60	76.09	83.54	89.02	—	—
LDLRSC [20]	87.50	93.25	97.86	97.11	97.78	97.58	77.70	69.01	93.33	94.63	—	—
DSESC [26]	88.75	92.20	97.94	97.25	98.33	98.10	—	—	82.50	89.10	—	—
EDS-SC [23]	89.75	—	98.00	97.76	97.99	98.10	—	—	—	—	—	—
MSCSCN (Ours)	**91.50**	**94.57**	**99.47**	**99.26**	**99.03**	**99.15**	**92.70**	**87.91**	**95.00**	**95.95**	**75.38**	**83.08**

TABLE IV
ABLATION STUDIES ON ALL DATASETS

Datasets	ORL		EYaleB		COIL20		MNIST		UMIST		COIL100	
Methods	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
$\mathcal{L}_{\text{MC}}+\mathcal{L}_{\text{SE}}$	89.25	94.19	96.67	95.78	95.53	98.12	80.40	72.57	89.38	94.42	71.57	81.49
$\mathcal{L}_{\text{SC}}+\mathcal{L}_{\text{SE}}$	89.50	94.28	98.68	98.21	92.78	96.70	92.50	87.77	93.13	95.81	69.92	82.48
$\mathcal{L}_{\text{MC}}+\mathcal{L}_{\text{SC}}+\mathcal{L}_{\text{SE}}$	**91.50**	**94.57**	**99.47**	**99.26**	**99.03**	**99.15**	**92.70**	**87.91**	**95.00**	**95.95**	**75.38**	**83.08**

1.67% and NMI by 1.32%, setting a new benchmark for subspace clustering tasks. These results demonstrate the effectiveness of our approach, underscoring its ability to outperform other competitive deep subspace clustering methods.

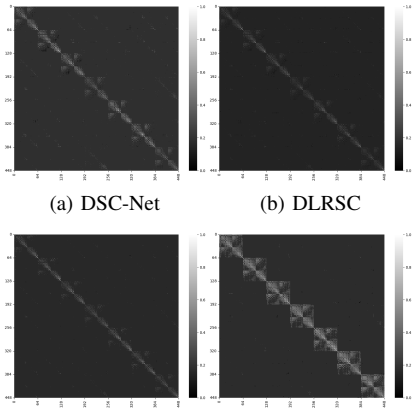


(a) m (b) T

Fig. 2: Clustering results in terms of ACC and NMI vs. different choices of parameter m and T on the EYaleB dataset.

C. Ablation Study

To assess the impact of various components within our framework, we conduct an ablation study and present the results in Table IV. Since our proposed framework integrates both multi-level contrastive loss and high-level structural contrastive loss based on the self-expression model, we denote "$\mathcal{L}_{\text{MC}}$ + $\mathcal{L}_{\text{SE}}$" as the training configuration excluding $\mathcal{L}_{\text{SC}}$, while "$\mathcal{L}_{\text{SC}}$ + $\mathcal{L}_{\text{SE}}$" corresponds to the configuration where $\alpha = 0$. The results highlight the importance of incorporating both structured information and multi-level features in subspace clustering. Specifically, removing $\mathcal{L}_{\text{SC}}$ forces the framework to treat each data point as belonging to a unique subspace, thus restricting the exploration of both intra-sample and inter-cluster similarities. Additionally, excluding $\mathcal{L}_{\text{MC}}$ weakens the framework's ability to capture high-quality features, as relying solely on high-level features for the structural contrastive loss neglects the contribution of low-level features.



(a) DSC-Net (b) DLRSC

(c) DSESC (d) MSCSCN(Ours)

Fig. 3: Visualization of the self-expression matrix using different methods on the EYaleB dataset.

D. Parameter and Visualization Analysis

We evaluate the sensitivity of MSCSCN to the parameters m and T through experiments and visualize a subset of C on the EYaleB dataset to demonstrate its self-expression capability. Here, m controls the rank of C and is configured within the range of $5 \times K$ to $12 \times K$, where $K = 64$ for the EYaleB dataset. Meanwhile, T determines the update frequency of Q. As shown in Fig. 2, the results reveal that MSCSCN consistently achieves low clustering errors across a wide range of m and T, highlighting its stability with respect to parameter variations. Additionally, as shown in Fig. 3, MSCSCN demonstrates a more distinct block-diagonal structure, fewer inter-cluster connection errors, and stronger intra-cluster connections compared to DSC-Net, DLRSC, and DSESC approaches.

IV. CONCLUSION

In this letter, we propose a multi-level structural contrastive model for deep subspace clustering. Unlike previous AE-based methods that focus on pixel-level reconstruction, our approach emphasizes data geometry through structural contrastive loss, enhancing cluster boundary discrimination. By integrating both low-level and high-level features into the self-expression learning process, the framework improves the preservation of subspace representations. Our experiments on benchmark datasets demonstrate the model's superiority over competing subspace clustering methods.

REFERENCES

[1] G. Xia, H. Sun, L. Feng, G. Zhang, and Y. Liu, "Human motion segmentation via robust kernel sparse subspace clustering," *IEEE Trans. Image Process*, vol. 27, no. 1, pp. 135–150, 2017.

[2] Y. Lochman, C. Olsson, and C. Zach, "Learned trajectory embedding for subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2024, pp. 19 092–19 102.

[3] L. Cao, L. Shi, J. Wang, Z. Yang, and B. Chen, "Robust subspace clustering by logarithmic hyperbolic cosine function," *IEEE Signal Process. Lett.*, vol. 30, pp. 508–512, 2023.

[4] Y. Qiu and P. Hao, "Self-supervised deep subspace clustering network for faces in videos," *The Vis. Comput.*, vol. 37, no. 8, pp. 2253–2261, 2021.

[5] H. Wang, J. Zhao, C. Zheng, and Y. Su, "scDSSC: Deep sparse subspace clustering for scRNA-seq data," *PLOS Comput. Biol.*, vol. 18, no. 12, e1010772, 2022.

[6] Z. Wang, H. Wang, J. Zhao, J. Xia, and C. Zheng, "scVSC: Deep variational subspace clustering for single-cell transcriptome data," *IEEE-ACM Trans. Comput. Biol. Bioinform.*, 2024.

[7] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, 2013.

[8] Y. Chen, C.-G. Li, and C. You, "Stochastic sparse subspace clustering," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 4155–4164.

[9] C.-G. Li, C. You, and R. Vidal, "On geometric analysis of affine sparse subspace clustering," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1520–1533, 2018.

[10] Y. Xu, P. Hu, J. Dai, N. Yan, and J. Wang, "Sparseness and correntropy-based block diagonal representation for robust subspace clustering," *IEEE Signal Process. Lett.*, vol. 31, pp. 1154–1158, 2024.

[11] L. Cao, L. Shi, J. Wang, Z. Yang, and B. Chen, "Robust subspace clustering by logarithmic hyperbolic cosine function," *IEEE Signal Process. Lett.*, vol. 30, pp. 508–512, 2023.

[12] R. Vidal and P. Favaro, "Low rank subspace clustering (LRSC)," *Pattern Recognit. Lett.*, vol. 43, pp. 47–61, 2014.

[13] X. Zhu, S. Zhang, Y. Li, J. Zhang, L. Yang, and Y. Fang, "Low-rank sparse subspace for spectral clustering," *IEEE Trans. Knowledge Data Eng.*, vol. 31, no. 8, pp. 1532–1543, 2018.

[14] K. Kang, C. Chen, and C. Peng, "Consensus low-rank multiview subspace clustering with cross-view diversity preserving," *IEEE Signal Process. Lett.*, vol. 30, pp. 1512–1516, 2023.

[15] F. Wang, C. Chen, and C. Peng, "Essential low-rank sample learning for group-aware subspace clustering," *IEEE Signal Process. Lett.*, vol. 30, pp. 1537–1541, 2023.

[16] X. Peng, J. Feng, J. T. Zhou, Y. Lei, and S. Yan, "Deep subspace clustering," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 12, pp. 5509–5521, 2020.

[17] P. Zhou, Y. Hou, and J. Feng, "Deep adversarial subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1596–1604.

[18] J. Zhang *et al.*, "Self-supervised convolutional subspace clustering network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5473–5482.

[19] M. Kheirandishfard, F. Zohrizadeh, and F. Kamangar, "Deep low-rank subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2020, pp. 864–865.

[20] Y. Chen, L. Cheng, Z. Hua, and S. Yi, "Laplacian regularized deep low-rank subspace clustering network," *Appl. Intell.*, vol. 53, no. 19, pp. 22 282–22 296, 2023.

[21] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, "Deep subspace clustering networks," *Proc. Adv. Neural Inf. Process. Syst.*, 2017.

[22] Z. Yu, Z. Zhang, W. Cao, C. Liu, C. P. Chen, and H.-S. Wong, "Gan-based enhanced deep subspace clustering networks," *IEEE Trans. Knowledge Data Eng.*, vol. 34, no. 7, pp. 3267–3281, 2020.

[23] Q. Wang, X. Chen, Y. Li, and Y. Lin, "Elastic deep sparse self-representation subspace clustering network," *Neural Process. Lett.*, vol. 56, p. 58, 2024.

[24] C. Chen, H. Lu, H. Wei, and X. Geng, "Deep subspace image clustering network with self-expression and self-supervision," *Appl. Intell.*, vol. 53, no. 4, pp. 4859–4873, 2023.

[25] Z. Peng, H. Liu, Y. Jia, and J. Hou, "Adaptive attribute and structure subspace clustering network," *IEEE Trans. Image Process.*, vol. 31, pp. 3430–3439, 2022.

[26] L. Zhao, Y. Ma, S. Chen, and J. Zhou, "Deep double self-expressive subspace clustering," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2023, pp. 1–5.

[27] L. Fu *et al.*, "Subspace-contrastive multi-view clustering," *ACM Trans. Knowl. Discov. Data.*, vol. 18, no. 9, pp. 1–35, 2024.

[28] E. E. R. Vidal *et al.*, "Sparse subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 6, 2009, pp. 2790–2797.

[29] C. You, C.-G. Li, D. P. Robinson, and R. Vidal, "Oracle based active set algorithm for scalable elastic net subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3928–3937.

[30] V. M. Patel and R. Vidal, "Kernel sparse subspace clustering," in *Proc. IEEE Int. Conf. Image Process.*, IEEE, 2014, pp. 2849–2853.

[31] C. You, D. Robinson, and R. Vidal, "Scalable sparse subspace clustering by orthogonal matching pursuit," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3918–3927.

[32] P. Ji, M. Salzmann, and H. Li, "Efficient dense subspace clustering," in *Proc. IEEE Wint. Conf. on Appl. Comput. Vis.*, 2014, pp. 461–468.

[33] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 663–670.