

Research Article

Deep Convolutional Asymmetric Autoencoder-Based Spatial-Spectral Clustering Network for Hyperspectral Image

Baisen Liu ^{1,2}, Weili Kong ^{1,2} and Yan Wang ¹

¹College of Electrical and Information Engineering, Heilongjiang Institute of Technology, Harbin 150001, China

²College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Weili Kong; kkweil@hrbeu.edu.cn

Received 4 June 2022; Accepted 12 August 2022; Published 13 September 2022

Academic Editor: Kuruva Lakshmana

Copyright © 2022 Baisen Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the complex properties of hyperspectral images (HSI), such as spatial-spectral structure, high dimension, and great spectral variability, HSI clustering is a challenging operation. In this paper, we propose a novel deep convolutional asymmetric autoencoder-based spatial-spectral clustering network (DCAAES²C-Net) which employs a convolutional autoencoder (CAE) and an asymmetric autoencoder to investigate spatial-spectral information. First, we use a CAE to extract spatial-spectral features. Then, we introduce an asymmetric autoencoder between the encoder and decoder of CAE to suppress some non-material-related spatial information in latent feature like shading and texture. By using a collaborative strategy to train the proposed networks, we obtain the representation features in a low dimension. Furthermore, we improve the *k*-means algorithm by using the concept of over-clustering to handle fuzzy representation which is difficult to distinguish the cluster, and utilize it to obtain the final HSI clustering result. The results of the experiments demonstrated that the proposed methodology outperforms other methods on the frequently used hyperspectral image dataset.

1. Introduction

Recent years have witnessed a spurt of progress in remote sensing technology; it has promoted study of hyperspectral remote sensing [1]. HSI is captured by hyperspectral sensors such as hyperspectral imaging spectrometers, which can image regions of interest with nanoscale spectral resolution, gathering rich spectra to capture information about numerous ground objects [2, 3]. It is a 3D cube structure image with tens to hundreds of bands that includes various ground object information and allows meticulous ground object classification using deep networks [4], and it has been widely employed in a variety of industries, such as mineral exploration [5, 6], vegetation monitoring [7, 8], quantitative inversion of physical and biological parameters [9, 10], and military reconnaissance [11, 12]. Deep networks that employ supervised learning, on the other hand, typically require a substantial quantity of labeled data. Unfortunately, sample collection is time consuming, labor intensive, costly, and inefficient in practice, and training samples may be unavailable in some remote and no man's areas, severely limiting

the application capabilities of hyperspectral remote sensing [13]. Thus, to increase the application potential of hyperspectral remote sensing, unsupervised ground object recognition theory and method is necessary to be developed to overcome the limitations of labeled samples and prior information.

Generally, the learning-based clustering methods of HSI include two important elements: clustering algorithm and feature extraction. Among them, the purpose of clustering algorithm is, as a typical unsupervised information analysis technology, it does not rely on any training samples, but only by mining the essential characteristics of the data to achieve the natural division of pixels, which effectively solves the classification problem without prior information. According to the difference of clustering algorithm principle and working mechanism, hyperspectral clustering can be summarized into 9 categories of methods by Zhai et al. [13]: centroid-based methods [14, 15], density-based methods [16, 17], probability-based methods [18], bionics-based methods [19], intelligent computing-based methods [20], graph-based methods [21, 22], subspace clustering

methods [23, 24], deep learning-based methods [25], and hybrid mechanism-based methods [26]. Centroid-based methods and density-based methods are used widely. The method based on clustering center is the first method to be introduced into hyperspectral clustering analysis, and it is also one of the most classic clustering methods.

The purpose of feature extraction is to search a mapping from high-dimensional space to low-dimensional space, so that it can reduce redundant information and preserve crucial information. In the early years, researchers focused on using linear transformations to extract HSI features, such as linear discriminant analysis [27], independent component analysis [28], minimum noise separation transformation [29], and some PCA-based methods [30, 31]; then, some traditional clustering algorithms were applied to achieve clustering results. However, due to the complicated properties of HSI data, the performance is limited [32, 33]. Nowadays, a more advanced HSI clustering method is learning-based method. It is widely used in HSI clustering to tackle the nonlinearity problem, which exceeds the performance of many traditional methods. There are generally two types of learning-based clustering algorithms: spectral-only methods and spatial-spectral methods. The spectral-only methods like automatic fuzzy clustering based on an adaptive multi-objective differential evolution (AFCMDE) [34], scalable graph-based clustering with nonnegative relaxation (SGCNR) [35], and a robust manifold matrix factorization-based method (RMMF) [36], which cluster the HSI pixels by learning the spectral domain feature representations. A specific land-cover class is intuitively represented by an area with multiple pixels, so the center pixel and its neighboring pixels are most likely from the same category. However, the approaches that exclusively use spectral information discard the spatial relationship between neighboring pixels. Thus, some researchers introduced spatial-spectral HSI clustering algorithms to combine spatial and spectral information and get more discriminative features for HSI, based on the spatial-spectral feature representations, clustering methods are used to produce the final clustering result. Lei et al. [24] proposed a deep spatial-spectral subspace clustering network (DS³C-Net) which employed a multiscale autoencoder and self-expressive layers to explore spatial-spectral information and learn the subspace structures then used the spectral clustering to generate the final result. Murphy and Maggioni [37] integrated spectral-spatial diffusion geometry into the diffusion learning algorithm, which has achieved competitive performance and allows analyze the high-dimensional HSI data in a manner that both respects intrinsic pixel geometry in the data and the spatial regularity in the 2D image structure of the pixels. Nalepa et al. [38] used a 3D convolutional autoencoder to extract HSI features and achieved a good result in unsupervised segmentation.

A specific land-cover class in HSI data is generally represented by an area with multiple pixels which have similar spectral characteristics, and thus, how to make better use of spatial information and extract discriminative spatial-spectral features is critical for the HSI clustering task. Most of spatial-spectral methods are based on convolutional networks; thus, the feature will include a large amount of spatial

information such as shadow, texture, and geometric information due to the characteristics of convolutional network. Shen et al. [39] have proved that the actual reflectivity of a substance is only related to its material, and the shadows and textures that generated by the interaction of light and the shape of the substance's surface will interfere with its actual qualities; the information mentioned above is called non-material-related spatial information. Kang et al. [40] have certificated that non-material-related spatial information are meaningless in clustering task and hyperspectral pictures are mainly classified based on the similarity of spectral properties of substances. In addition, their method demonstrates that remove useless spatial information such as shading and texture which not directly related to the material of different objects effectively will obtain an outperformed result. Clustering is unsupervised, unlike the classification; it is very vulnerable to the characteristics of the data itself. Due to this reason, how to effectively remove the information contained in the feature vector is crucial.

In this letter, we concentrate on investigating spatial-spectral information from pixel patches and using the concept of over-clustering to improve k -means algorithm. Our key contributions include the following:

- (1) We propose a novel deep convolutional asymmetric autoencoder-based spatial-spectral clustering network (DCAAES²C-Net) to extract the discriminative spatial-spectral features
- (2) An asymmetric autoencoder is introduced to suppress non-material-related spatial information in feature representations which generated by CAE
- (3) We improve the k -means algorithm by using the concept of over-clustering to handle fuzzy representation which is difficult to distinguish the cluster

The remainder of this paper is organized as follows. Section 2 presents the proposed DCAAES²C-Net for unsupervised spatial-spectral feature learning. Section 3 reports and discusses experimental results over three benchmark hyperspectral datasets. Finally, conclusions are drawn in Section 4.

2. Method

2.1. DCAAES²C-Net. Simple autoencoder is a three-layer feed-forward fully connected network. Units in the previous layers are connected to all units in the next layer. The size of the input layer and output layer is equal to the input size. According to universal approximation theorem, deepening the depth of the network can provide more advantages. Thus, the deep autoencoder typically used to learn feature representations. We show this network in Figure 1.

The fundamental structure of the convolutional autoencoder (CAE) is extended by altering the fully linked layers to convolution layers. The input layer and output layer sizes are the same as in the standard autoencoder, while the decoder network changes to convolution layers and the decoder

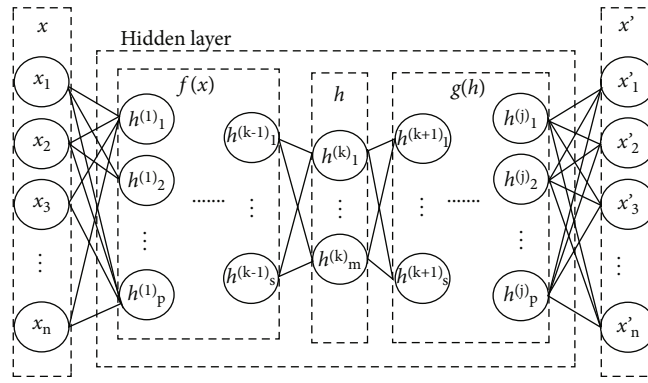


FIGURE 1: Architecture of deep autoencoder.

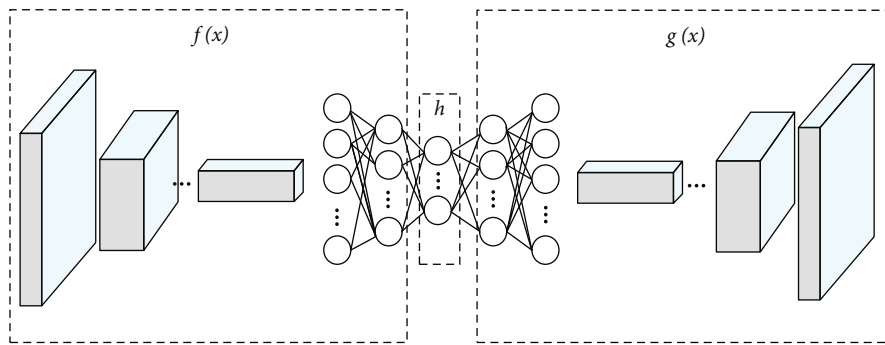


FIGURE 2: Architecture of convolutional autoencoder.

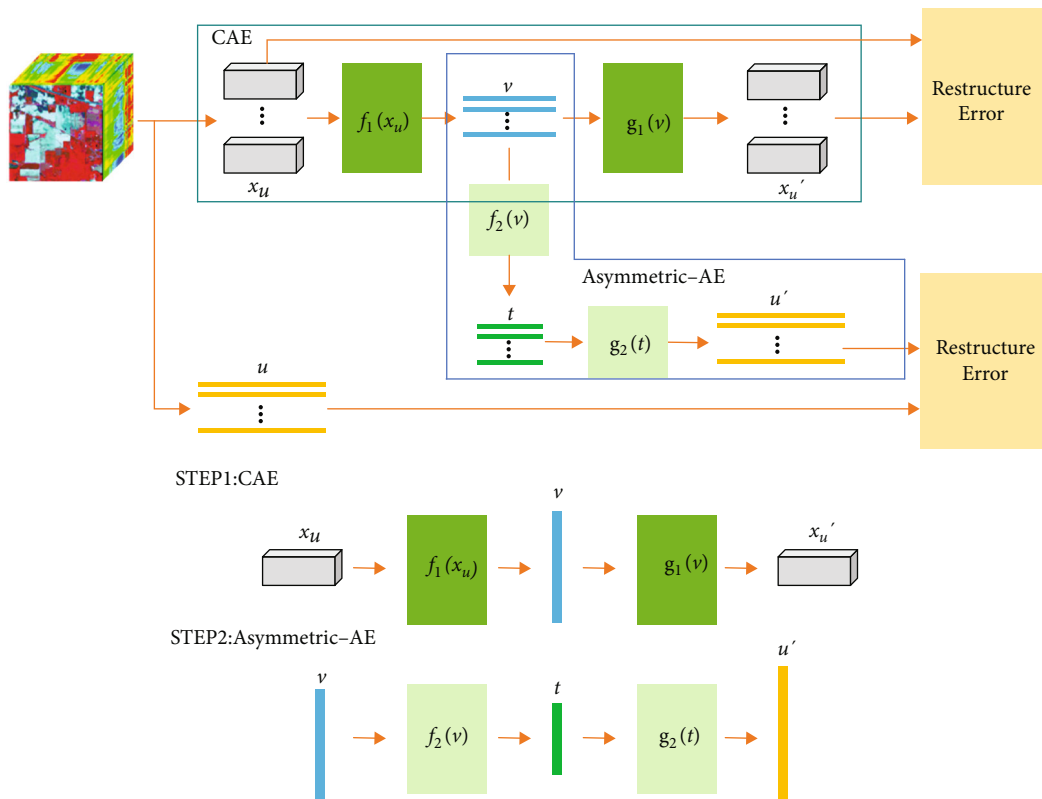


FIGURE 3: Architecture of proposed network.

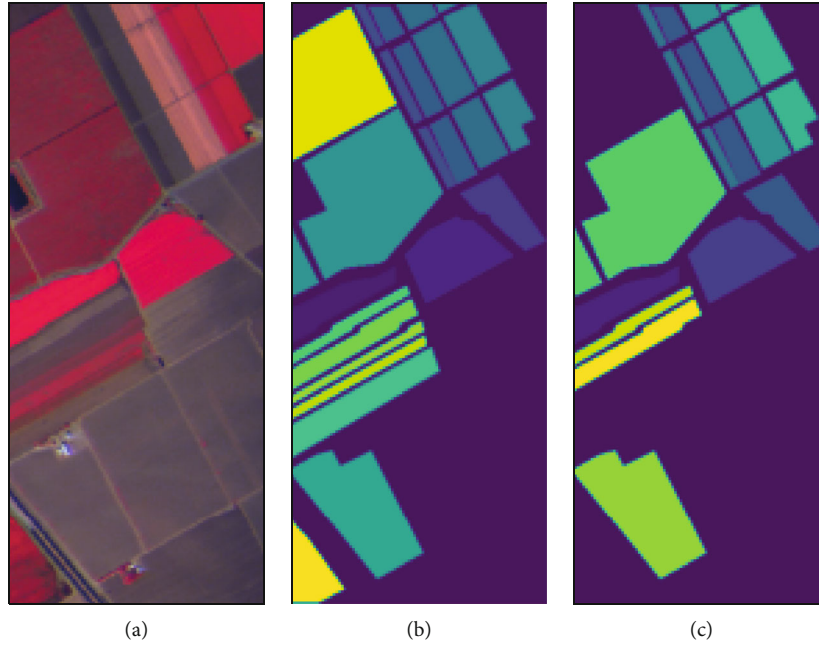
Over-clustering k -means**Input:** HSI feature vectors, n , k .**1:** Set k according goal clusters, n is 2~3 times of k **2: repeat****3:** Do K -means cluster with n clusters**4:** Keep the top k clusters with the most elements, the other clusters are set to the background cluster.**5:** Observe the clustering result and adjust n .**6: until** Remove noisy class effectively**7: output:** HSI clustering resultALGORITHM 1: Flow chart of over-clustering K -means algorithm.

FIGURE 4: (a) is the false-color image of Salinas hyperspectral image, (b) is the original real ground object of Salinas hyperspectral image, and (c) is the real ground object map of Salinas hyperspectral image after the combination of ground object types.

TABLE 1: Network parameter settings of CAE.

Layers	Kernel	Output	Channel
Input layer	/	5*5	204
Conv layer	(3, 3) * 256	3*3	256
Conv layer	(3, 3) * 512	1*1	512
Dense layer	/	512	/
Dense layer	/	256	/
Dense layer	/	128	/
Dense layer	/	256	/
Dense layer	/	512	/
Reshape layer	/	1*1	512
Transpose layer	(3, 3) ^T * 512	3*3	512
Transpose layer	(3, 3) ^T * 256	5*5	256
Output layer	(3, 3) * 204	5*5	204

TABLE 2: Network parameter settings of asymmetric-AE.

Layers	Kernel	Output	Channel
Input layer	/	128	/
Dense layer	/	64	/
Dense layer	/	32	/
Dense layer	/	10	/
Dense layer	/	32	/
Dense layer	/	64	/
Dense layer	/	128	/
Dropout layer	/	128	/
Output layer	/	204	/

network changes to transposed convolutional layers. We show this network in Figure 2.

In this paper, we first employ a CAE to learn the spatial-spectral information, and the input data of network is changed from spectral vector to patch. In general, it is considered

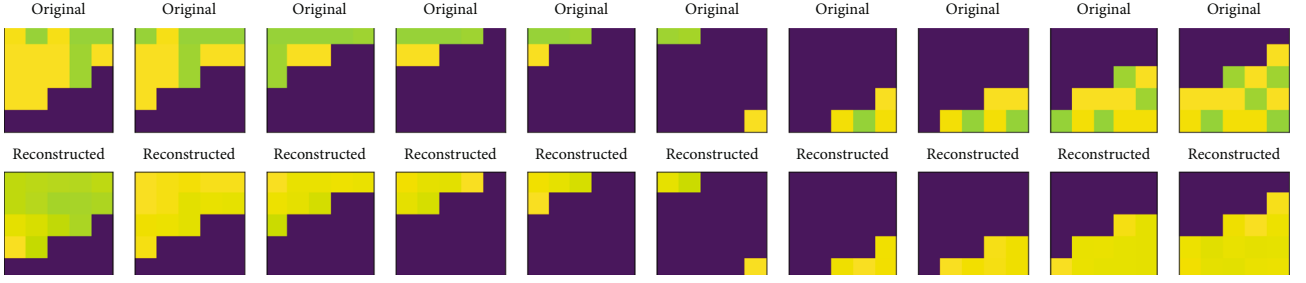


FIGURE 5: The original image and reconstructed image.

that the center pixel and its neighborhood contain correlation information, and the main purpose of introducing spatial information is to use the correlation of the center pixel and its neighborhood to enhance the features of the center pixel [41]. However, because the convolutional network is very sensitive to picture geometric characteristics, some non-material-related spatial information is also incorporated in the feature vector, such as the shape of edges, textures, and shadows. For example, when we use CAE to learn the information of the point which is near the edge of land, the feature vectors will contain the information of the edge's shape which will be the principal component of the feature vectors. It will lead the feature vectors of the edge points cannot contain the information which can describe its material correctly. In the following, we call this kind of non-material-related spatial information spatial noise. To address this problem, we designed an asymmetric autoencoder stack on the pretrained CAE to suppress the spatial noise in the output of CAE, as shown in Figure 3. In this process, since the hidden layer compresses the output of CAE and decompresses it to the corresponding spectrum, the spatial noise components will be dropout.

$f_1(x)$ represents the coding map of CAE, $g_1(x)$ represents the decoding map of CAE, x_u is the input data of the convolution network, and u is the original spectral information of the central pixel of x_u . For CAE, suppose there is a mapping for any ε , equation (2) holds. For v can completely reconstruct x_u , it is believed that the feature vector v can represent the information contained in x_u .

$$v = f_1(x_u), \quad (1)$$

$$P(|g_1(v) - x_u| < \varepsilon) = 1. \quad (2)$$

The overall correlation between u and x_u can be measured using the reciprocal of the Euclidean distance between the spectral vector u at the target point and the feature vector u' at the network output, as shown in equation (4). The information in v that indicates weak correlation with u is discarded by an asymmetric self-encoder, and a dropout layer is set in the decoder part to guarantee that the spatial information is not discarded in its entirety. The encoding mapping of the asymmetric depth self-encoder is denoted by $f_2(x)$, and the decoding mapping is denoted by $g_2(x)$. Maximizing the correlation between u and x_u is then equiv-

alent to maximizing the correlation between u and u' , as shown in equation (6), when $d(u', u) \rightarrow 0$.

$$u \in x_u, u \in R^c, u' \in R^c, \quad (3)$$

$$\text{correlation}(u, u') \sim \frac{1}{d(u', u)}, \quad (4)$$

$$t = f_2(v), \quad (5)$$

$$P(|g_2(t) - u| < \varepsilon) = 1, \quad (6)$$

In Asymmetric-AE, the input data is v , and the reconstruction object is the original spectral information of the target pixel. Its main function is to enhance the spectral information of the target pixel contained in v and suppress the spatial information that is irrelevant to the target pixel, that is, spatial noise.

2.2. Over-Clustering K-Means. K -means clustering algorithm is a clustering method based on clustering center. This method is sensitive to outlier noise points, and this kind of outlier noise point will destroy the stability of clustering and exert great impact on clustering accuracy. The method which is based on feature density is not sensitive to such noise points such as DBSCAN clustering algorithm [42]. According to a priori, the distribution of noise vector in feature space is sparse. Therefore, when DBSCAN is used to cluster hyperspectral data, the noise vector will not be divided into the final clustering results. Based on this idea, we improve the K -means clustering algorithm, so that the K -means clustering algorithm also has the characteristics as DBSCAN to separate the noise vector in sparse region. The improved method is shown in Algorithm 1. The cost function of the improved k -means algorithm is shown in equation (7), where k is the number of goal clusters, n is a hyper-parameter and it always bigger than k .

$$E = \sum_{i=1}^n \sum_{x \in C_i} \|x - \mu_i\|_2^2 = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 + \sum_{i=k+1}^n \sum_{x \in C_i} \|x - \mu_i\|_2^2, \quad (7)$$

The noise points are divided into background clusters, and the cost function E can be equivalent to E' , where μ_0

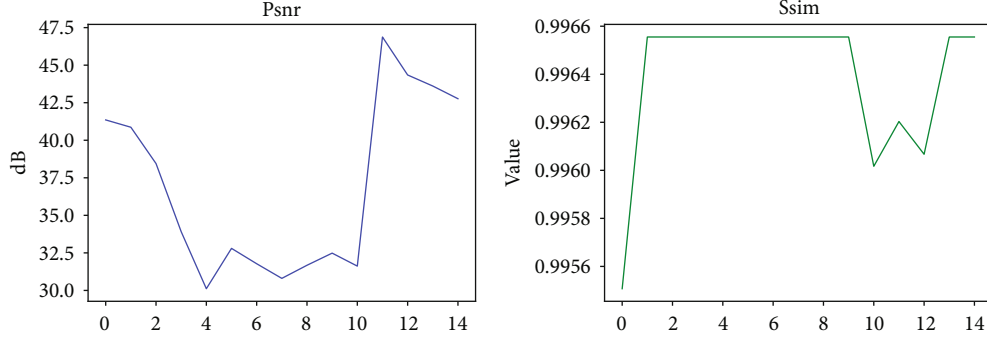


FIGURE 6: PSNR and SSIM of original images and reconstructed images.

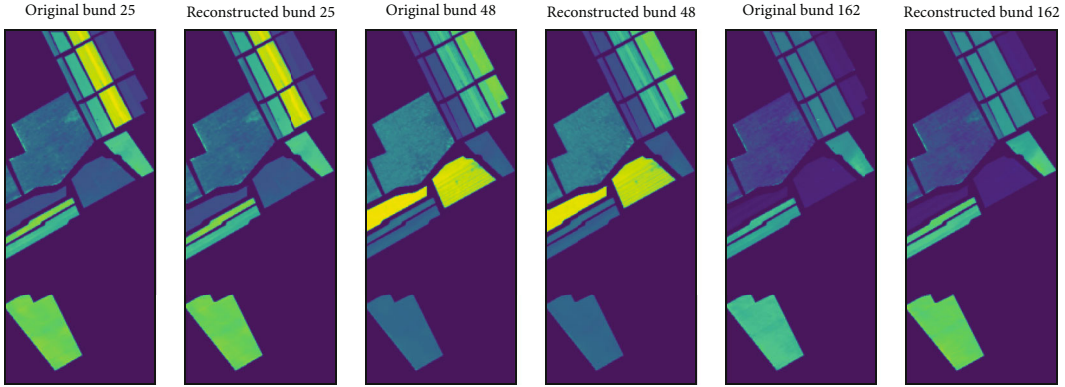


FIGURE 7: The original image and reconstructed image of bund 25, 48, and 162 (random selected).

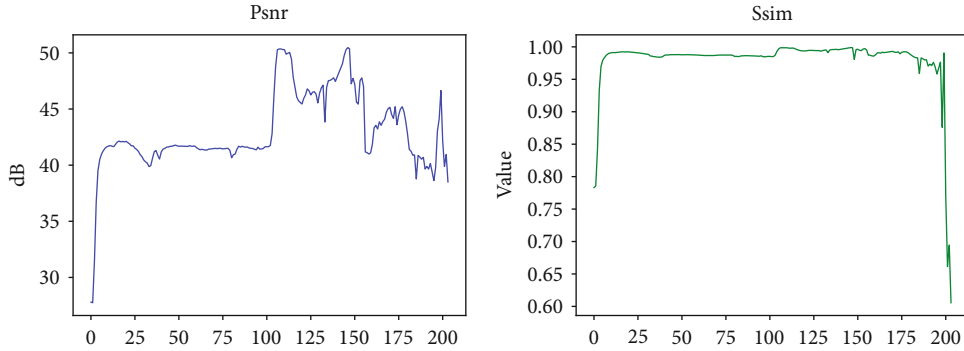


FIGURE 8: PSNR and SSIM of each bund.

represents the mean vector of the background cluster. According to a priori, the values of all elements in the background cluster are 0, that is, $\mu_0 = \mathbf{0}$. The distance within the noise cluster is generally larger than that within the non-noise cluster. Therefore, the following formula can be derived.

$$\sum_{i=k+1}^n \sum_{x \in C_i} \|x - \mu_i\|_2^2 = \sum_{i=k+1}^n \sum_{x \in C_i} \|0 - \mu_0\|_2^2 = 0, \quad (8)$$

$$E' = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 + 0, \quad (9)$$

TABLE 3: Comparison of other typical autoencoder models.

Model	PSNR/dB	SSIM
AE	22.786322	0.9605432
CAE	40.830353	0.9831473
Proposed model	41.673317	0.9751903

$$0 < \sum_{i=k+1}^n \sum_{x \in C_i} \|x - \mu_i\|_2^2, \quad (10)$$

$$E' < E. \quad (11)$$

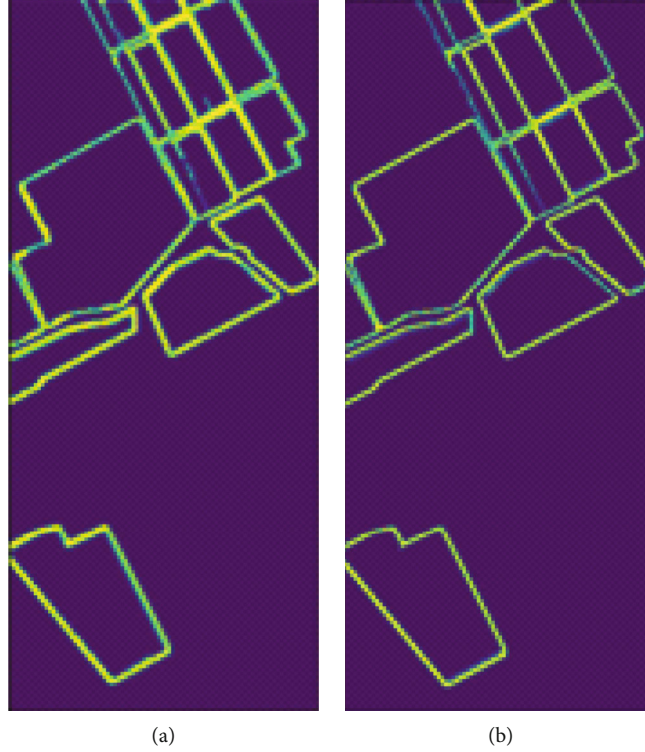


FIGURE 9: DBSCAN cluster noisy label map of CAE (a) and proposed model (b).

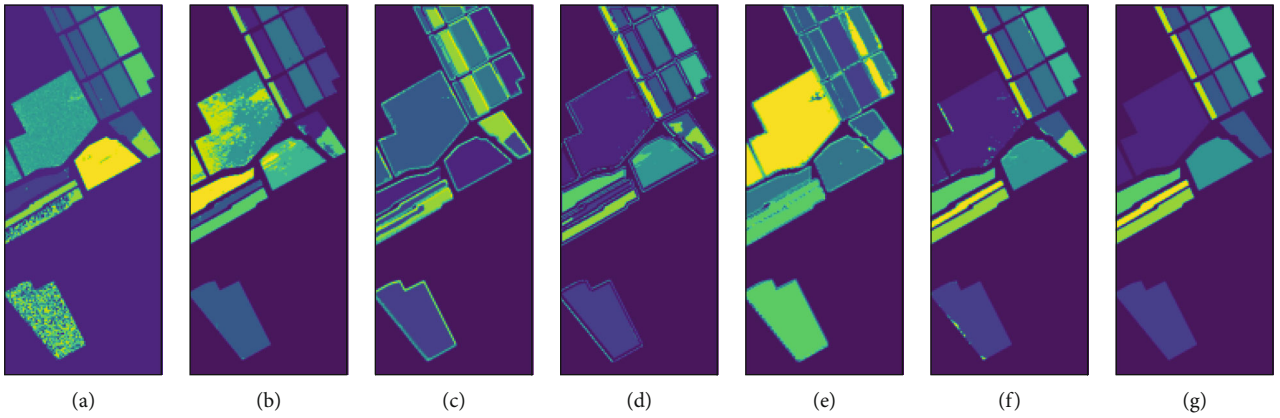


FIGURE 10: The result contradistinction of proposed clustering method and other typical methods; (a) PCA+K-Means, (b) AE+K-Means, (c) CAE+K-Means, (d) CAE+over-clustering K -Means (e) DCAAES²+K-means, (f) DCAAES²C-Net, and (g) ground truth.

It can be seen that by dividing the noise points into background clusters, the cost function is further optimized while removing the noise points, and the clustering performance is better.

3. Experimental Results

3.1. Experiments Setting and Dataset. In this paper, all the experiments are carried out using a PC equipped with InterCorei7-10700K CPU and a single GPU of GeForce RTX 3070.

Salinas scene is used as the experimental data in this paper. The size of the image is 512×217 , and it contains

16 classes. We select 10 categories for research. Set the data labels of the 10th, 13th, and 16th categories to zero. The 3rd and 5th categories are collectively referred to as one category. No changes will be made to categories 1, 2, 4, 6, 7, 8, 9, 11, and 12, as shown in Figure 4.

We first normalize the hyperspectral data by min-max normalization. In order to prevent the mutual influence between the various bands, in this paper, when the min-max normalization is performed, each band is normalized separately. To more accurately evaluate the effectiveness of the feature extraction network, we use the sampling method of Bootstrap Sample to extract the training set, and the elements that do not appear in the training set are used as the test set [43].

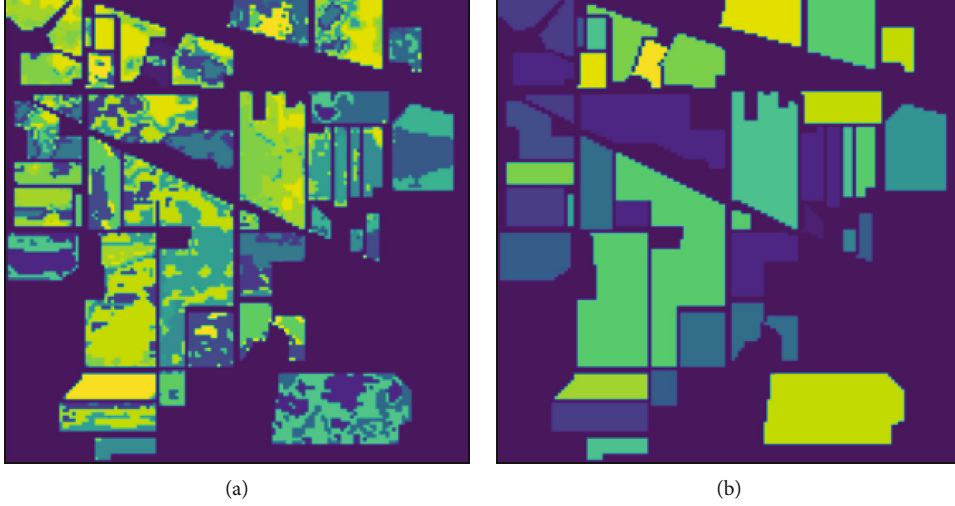


FIGURE 11: (a) The clustering result of Indian pines. (b) Ground truth.

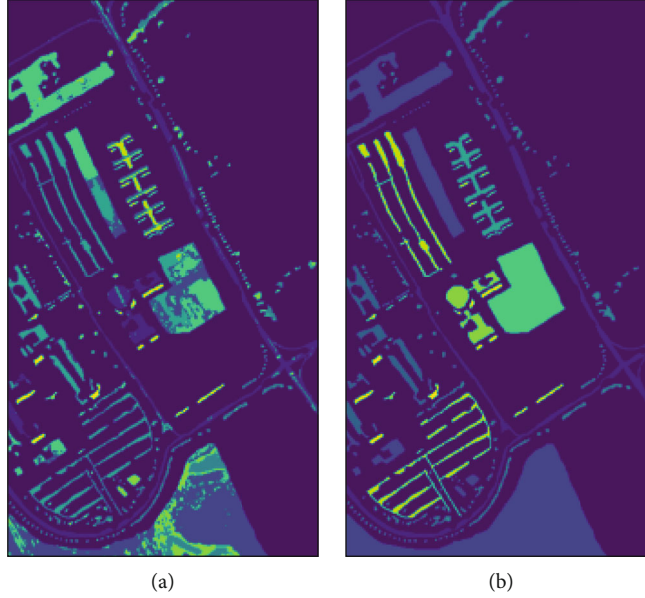


FIGURE 12: (a) The clustering label of PaviuaU. (b) Ground truth.

3.2. Training Model. The model built and trained according to the structure is shown in Tables 1 and 2. Epoch of CAE is set to 300 and batch size is set to 256. Epoch of Asymmetric-AE is set to 1000 and batch size is set to 1000. The gradient optimization function uses Adam, the loss function uses MSE, and the activation function uses Relu.

3.3. Evaluation. This paper uses PSNR and SSIM as the evaluation criteria of the similarity between the input image and the reconstructed image [44, 45]. For the CAE, 10 input images are randomly selected from the test set for reconstruction. The original image and its reconstructed image are shown in Figure 5. Calculate the PSNR and SSIM of these sub-images and their reconstructed images, as shown in Figure 6. The ten sets of PSNR values we get are the lowest 31 dB and the highest 46 dB; the lowest SSIM value is 0.967

TABLE 4: The evaluation contradistinction of proposed clustering method and other typical methods.

Method	RI	FMI	AMI	DB
(a)	0.9905	0.9888	0.9094	1.3924
(b)	0.9920	0.9907	0.9271	0.7089
(c)	0.9363	0.9236	0.7299	0.9643
(d)	0.8954	0.8871	0.7351	0.9045
(e)	0.8992	0.8759	0.7808	0.8140
(f)	0.9960	0.9954	0.9536	0.8359

Note: (a) PCA + K-Means, (b) AE + K-Means, (c) CAE + K-Means, (d) CAE + Over-clustering K-Means (e) DCAAES² + K-means, (f) DCAAES²C-Net, (g) Ground truth.

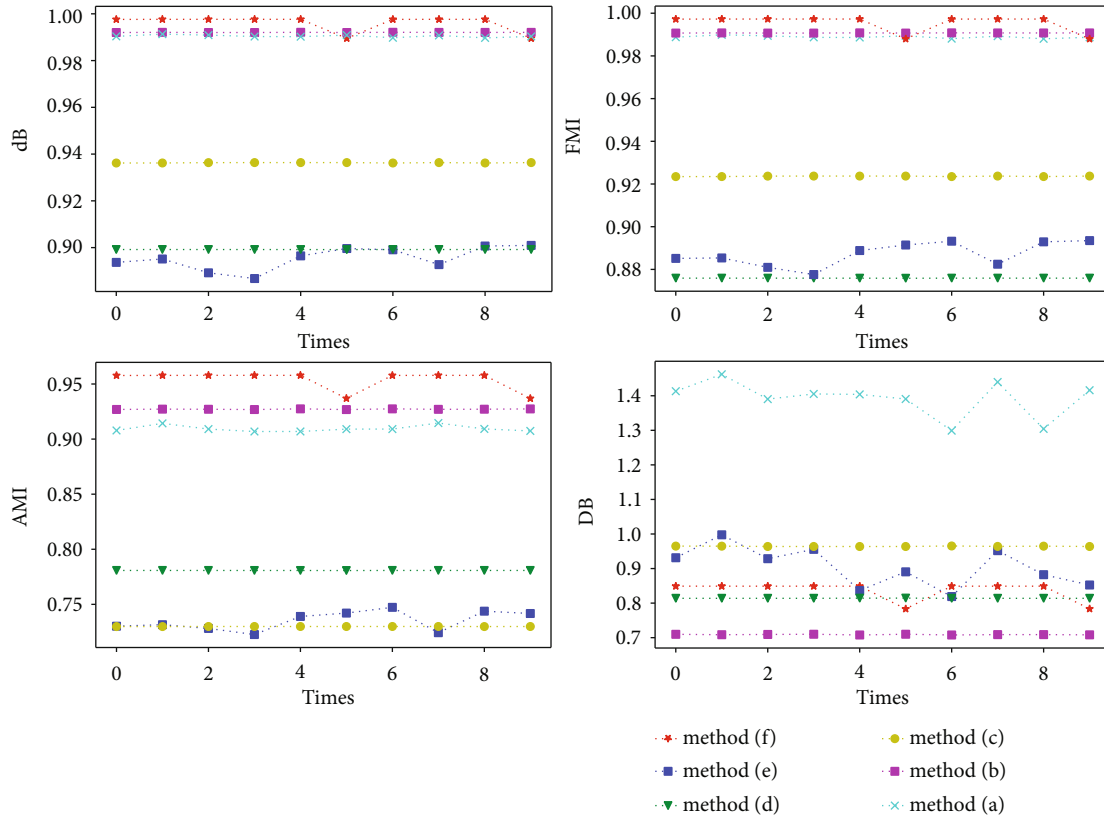


FIGURE 13: The curve of average evaluation score.

and the highest is 0.994. It can be seen CAE can learn the information in the sub-image and compress it into a feature vector with a smaller dimension.

Asymmetric-AE is trained base the feature vectors generated by the CAE. Since the output of the Asymmetric-AE is the spectral vector, the similarity between the original image and its reconstructed image is analyzed directly. After calculation, the PSNR and SSIM of the original image and reconstructed image are about 41.67db and 0.983, respectively. The reconstructed image and the original image of the Asymmetric-AE are shown in Figure 7. The PSNR and SSIM of each band are shown in Figure 8. It can see that PSNR and SSIM still maintain a high value in general, and the original image can be reconstructed well after feature vector reduction again.

The comparison of feature extraction effectiveness between the proposed model and other typical autoencoder models is shown in Table 3.

Figure 9 is the label graph of noise points in DBSCAN clustering of CAE and the model proposed in this paper (domain parameter is $(\epsilon, \text{MinPts}) = (3, 200)$).

The number of noise points in the clustering results of CAE is 9997, and the number of noise points in the clustering results of the proposed model is 6996. It can be seen from the comparison that the features extracted in this paper have better performance in clustering performance, but there are still some spatial noises. Therefore, the improved k -means clustering algorithm is used to further reduce the influence of noise points on clustering accuracy in this paper.

TABLE 5: The inference runtime in different datasets.

Dataset	Salinas	PaviuaU	Indian pines
Runtime(s)	6.96	2.88	2.54

In order to scientifically evaluate the performance of the clustering method proposed in this paper, FMI, RI, AMI, and DB are used as the evaluation indicators of the clustering performance. Among them, FMI, RI, and AMI are external indicators that require ground truth as a reference standard [46–48]. The higher the score, the better the clustering performance. DB is an internal index [49], which represents the average similarity between clusters, and the ratio of the flat distance within a cluster to the distance between clusters is used as the evaluation criterion for similarity. 0 is the lowest value, and the lower the value, the better the clustering effect. Due to the randomness of K -means clustering algorithm [50], 10 clustering experiments are carried out in each experiment, and the clustering result with the highest evaluation index is taken as the final output. The comparison of clustering results between our method and other typical clustering methods on Slinas dataset is shown in Figure 10. The clustering results of Indian pines and PaviuaU are shown in Figures 11 and 12, respectively. After 10 experiments, the average evaluation score is shown in Table 4 and the curve is shown in Figure 13, which indicate that the clustering performance of the proposed feature extraction model combined with the improved k -means

clustering algorithm is significantly better than that of other typical models (the number of clustering centers of the improved k -means clustering algorithm is 33, and the number of real clusters is 11).

The inference runtime of our method in different datasets is shown in Table 5. Due to our net compresses the information into a 10-dimension vector, thus computational cost is greatly reduced.

4. Conclusion

In this letter, we propose a novel DCAAES²C-Net, which explore spatial-spectral information by using an asymmetric autoencoder to suppress spatial noise information component in feature. Besides, we use the concept of over-clustering to improve the k -means algorithm to reduce the influence of fuzzy feature. Finally, the improved clustering algorithm is applied based on the output of autoencoder network to obtain the HSI clustering result. Experimental results on Salinas scene demonstrate the effectiveness of the proposed method. The RI index of the clustering results is 0.9960, which improves 0.4%~6.3%; the FMI index is 0.9954, which improves 0.5%~7.8%; the AMI index is 0.9536, which improves 2.8%~30.6%; and the DB index is 0.8359, which decreases -17.9%~39.9%.

In the future, we will focus on exploring the spatial-spectral from multiscale patch, combining the proposed method with one-stage clustering method and try to utilize this method to improve the performance of semi-supervised classification task.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by the Natural Science Foundation of Heilongjiang Province for Key Projects, China (Grant no. ZD2021F004) and the Postdoctoral Scientific Research Developmental Fund of Heilongjiang Province, China (Grant no. LBH-Q18110).

References

- [1] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1923–1938, 2019.
- [2] I. Numata, D. A. Roberts, O. A. Chadwick, J. P. Schimel, L. S. Galvão, and J. V. Soares, "Evaluation of hyperspectral data for pasture estimate in the Brazilian Amazon using field and imaging spectrometers," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1569–1583, 2008.
- [3] L. Gaddis, J. Anderson, K. Becker et al., "An overview of the integrated software for imaging spectrometers (ISIS)," in *Lunar and Planetary Science Conference*, vol. 28, p. 387, Houston, 1997.
- [4] S. Mei, J. Ji, Y. Geng, Z. Zhang, X. Li, and Q. Du, "Unsupervised spatial-spectral feature learning by 3D convolutional autoencoder for hyperspectral classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6808–6820, 2019.
- [5] F. A. Kruse, J. W. Boardman, and J. F. Huntington, "Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1388–1400, 2003.
- [6] L. Tusa, L. Andreani, M. Khodadadzadeh et al., "Mineral mapping and vein detection in hyperspectral drill-core scans: application to porphyry-type mineralization," *Minerals*, vol. 9, no. 2, p. 122, 2019.
- [7] U. Bradter, J. O'Connell, W. E. Kunin, C. W. Boffey, R. J. Ellis, and T. G. Benton, "Classifying grass-dominated habitats from remotely sensed data: the influence of spectral resolution, acquisition time and the vegetation classification system on accuracy and thematic resolution," *Science of the Total Environment*, vol. 711, article 134584, 2020.
- [8] H. Zhang, J. Kang, X. Xu, and L. Zhang, "Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: a case study of Yi'an County, Heilongjiang province, China," *Computers and Electronics in Agriculture*, vol. 176, article 105618, 2020.
- [9] R. Darvishzadeh, C. Atzberger, A. Skidmore, and M. Schlerf, "Mapping grassland leaf area index with airborne hyperspectral imagery: a comparison study of statistical approaches and inversion of radiative transfer models," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. 894–906, 2011.
- [10] B. Kong, H. Yu, R. Du, and Q. Wang, "Quantitative estimation of biomass of alpine grasslands using hyperspectral remote sensing," *Rangeland Ecology & Management*, vol. 72, no. 2, pp. 336–346, 2019.
- [11] K. C. Tiwari, M. K. Arora, and D. Singh, "An assessment of independent component analysis for detection of military targets from hyperspectral images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 5, pp. 730–740, 2011.
- [12] M. Shimoni, R. Haelterman, and C. Perneel, "Hyperspectral imaging for military and security applications: combining myriad processing and sensing techniques," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 101–117, 2019.
- [13] H. Zhai, H. Zhang, P. Li, and L. Zhang, "Hyperspectral image clustering: current achievements and future lines," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 4, pp. 35–67, 2021.
- [14] S. A. El Rahman, "Hyperspectral imaging classification using ISODATA algorithm: Big data challenge," in *2015 Fifth International Conference on e-Learning (econf)*, pp. 247–250, Manama, Bahrain, 2015.
- [15] J. M. Haut, M. Paoletti, J. Plaza, and A. Plaza, "Cloud implementation of the K-means algorithm for hyperspectral image analysis," *The Journal of Supercomputing*, vol. 73, no. 1, pp. 514–529, 2017.

- [16] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.
- [17] Y. Chen, S. Ma, X. Chen, and P. Ghamisi, "Hyperspectral data clustering based on density analysis ensemble," *Remote sensing letters*, vol. 8, no. 2, pp. 194–203, 2017.
- [18] N. Acito, G. Corsini, and M. Diani, "An unsupervised algorithm for hyperspectral image segmentation based on the Gaussian mixture model," in *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, vol. 6, pp. 3745–3747, Toulouse, France, July 2003.
- [19] A. Baraldi and F. Parmiggiani, "A neural network for unsupervised categorization of multivalued input patterns: an application to satellite image clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 305–316, 1995.
- [20] S. Das, A. Abraham, and A. Konar, "Automatic clustering using an improved differential evolution algorithm," *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, pp. 218–237, 2008.
- [21] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.
- [22] W. Liu, S. Li, X. Lin, Y. Wu, and R. Ji, "Spectral-spatial co-clustering of hyperspectral image data based on bipartite graph," *Multimedia Systems*, vol. 22, no. 3, pp. 355–366, 2016.
- [23] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectral-spatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3672–3684, 2016.
- [24] J. Lei, X. Li, B. Peng, L. Fang, N. Ling, and Q. Huang, "Deep spatial-spectral subspace clustering for hyperspectral image," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, pp. 2686–2697, 2021.
- [25] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, "Towards k-means-friendly spaces: simultaneous deep learning and clustering," in *International conference on machine learning*, pp. 3861–3870, Sydney, Australia, July 2017.
- [26] V. E. Neaogoe and V. Chirila-Berbentea, "Improved Gaussian mixture model with expectation-maximization for clustering of remote sensing imagery," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3063–3065, Beijing, China, 2016.
- [27] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 5, pp. 894–898, 2011.
- [28] D. Hongtao, H. Qi, X. Wang, R. Ramanath, and W. E. Snyder, "Band selection using independent component analysis for hyperspectral image processing," in *32nd Applied Imagery Pattern Recognition Workshop, 2003. Proceedings*, pp. 93–98, Washington, DC, USA, 2003.
- [29] L. Gao, B. Zhao, X. Jia, W. Liao, and B. Zhang, "Optimized kernel minimum noise fraction transformation for hyperspectral image classification," *Remote Sensing*, vol. 9, no. 6, p. 548, 2017.
- [30] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai, and L. Wang, "Super-PCA: a superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4581–4593, 2018.
- [31] J. Zabalza, J. Ren, M. Yang et al., "Novel folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 112–122, 2014.
- [32] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE Journal of Selected topics in applied earth observations and remote sensing*, vol. 7, no. 6, pp. 2094–2107, 2014.
- [33] P. Zhong, Z. Gong, S. Li, and C.-B. Schonlieb, "Learning to diversify deep belief networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3516–3530, 2017.
- [34] Y. Zhong, S. Zhang, and L. Zhang, "Automatic fuzzy clustering based on adaptive multi-objective differential evolution for remote sensing imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 5, pp. 2290–2301, 2013.
- [35] R. Wang, F. Nie, Z. Wang, F. He, and X. Li, "Scalable graph-based clustering with nonnegative relaxation for large hyperspectral image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7352–7364, 2019.
- [36] L. Zhang, L. Zhang, B. Du, J. You, and D. Tao, "Hyperspectral image unsupervised classification by robust manifold matrix factorization," *Information Sciences*, vol. 485, pp. 154–169, 2019.
- [37] J. M. Murphy and M. Maggioni, "Spectral-spatial diffusion geometry for hyperspectral image clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1243–1247, 2020.
- [38] J. Nalepa, M. Myller, Y. Imai, K.-I. Honda, T. Takeda, and M. Antoniak, "Unsupervised segmentation of hyperspectral images using 3-D convolutional autoencoders," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 11, pp. 1948–1952, 2020.
- [39] J. Shen, X. Yang, X. Li, and Y. Jia, "Intrinsic image decomposition using optimization and user scribbles," *IEEE Transactions on Cybernetics*, vol. 43, no. 2, pp. 425–436, 2013.
- [40] X. Kang, S. Li, L. Fang, and J. A. Benediktsson, "Intrinsic image decomposition for feature extraction of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2241–2253, 2015.
- [41] Q. Wang, Q. Li, and X. Li, "Hyperspectral band selection via optimal neighborhood reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8465–8476, 2020.
- [42] H. Bäcklund, A. Hedblom, and N. Neijman, "A density-based spatial clustering of application with noise," *Data Mining TNM033*, vol. 33, pp. 11–30, 2011.
- [43] Y. Xu and R. Goodacre, "On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning," *Journal of Analysis and Testing*, vol. 2, no. 3, pp. 249–262, 2018.
- [44] G. H. Chen, C. L. Yang, L. M. Po, and S. L. Xie, "Edge-based structural similarity for image quality assessment," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 2, Toulouse, France, 2006.
- [45] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *2010 20th international conference on pattern recognition*, Istanbul, Turkey, 2010.

- [46] E. B. Fowkles and C. L. Mallows, "A method for comparing two hierarchical clusterings: rejoinder," *Journal of the American Statistical Association*, vol. 78, no. 383, p. 584, 1983.
- [47] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985.
- [48] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance," *The Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.
- [49] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.
- [50] J. J. López, J. A. Aguado, F. Martín, F. Muñoz, A. Rodríguez, and J. E. Ruiz, "Hopfield- $_K_$ -Means clustering algorithm: a proposal for the segmentation of electricity customers," *Electric Power Systems Research*, vol. 81, no. 2, pp. 716–724, 2011.