Robust Deep 3D Convolutional Autoencoder for Hyperspectral Unmixing with Hypergraph Learning

Peiyuan Jia, Miao Zhang* and Yi Shen

(*Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, China*)

Abstract：Hyperspectral unmixing aims to acquire pure spectra of distinct substances (endmembers) and fractional abundances from highly mixed pixels. In this paper, a deep unmixing network framework is designed to deal with the noise disturbance. It contains two parts：a three-dimensional convolutional autoencoder (denoising 3D CAE) which recovers data from noised input, and a restrictive non-negative sparse autoencoder (NNSAE) which incorporates a hypergraph regularizer as well as a $l_{2,1}$-norm sparsity constraint to improve the unmixing performance. The deep denoising 3D CAE network was constructed for noisy data retrieval, and had strong capacity of extracting the principle and robust local features in spatial and spectral domains efficiently by training with corrupted data. Furthermore, a part-based nonnegative sparse autoencoder with $l_{2,1}$-norm penalty was concatenated, and a hypergraph regularizer was designed elaborately to represent similarity of neighboring pixels in spatial dimensions. Comparative experiments were conducted on synthetic and real-world data, which both demonstrate the effectiveness and robustness of the proposed network.

Keywords：deep learning；unsupervised unmixing；convolutional autoencoder；hypergraph；hyperspectral data

CLC number：TP753 　　　**Document code**：A　　　**Article ID**：1005-9113(2021)05-0001-08

0　Introduction

The hyperspectral imagery (HSI) is composed of hundreds of narrow bands with high spectral resolution, which captures signatures of different substances measured by surface reflectance. However, the low spatial resolution would lead to the mixed components in HSIs. As a result, unmixing techniques are developed to acquire the spectrum and its spatial distribution of each material from the highly mixed data.

Deep autoencoders are emerging as a popular approach for unsupervised unmixing due to its strong capacity for low-level feature interpretation. The hierarchical structure of the network can learn the latent distribution of the input and extract representative features effectively. Several researches have investigated autoencoder models for linear unmixing. The autoencoder cascade is an instructive work which marginalized denoising autoencoder is concatenated with a nonnegative autoencoder[1]. To handle the problems with outliers, ASNSA adopts local outlier factor to detect the outlier samples[2], while DAEN utilizes the variational autoencoder to perform blind unmixing for HSIs[3]. Endnet develops a novel unmixing scheme with sparsity and smoothness constraints to improve consistency[4]. uDAs incorporates the denoising process into the network by imposing a denoising constraint[5]. However, these network models are usually impacted by noise since the correlation of the spatial and spectral information is not simultaneously taken into consideration.

HSIs exhibit local consistency spatially, and incorporating spatial information can improve the robustness of the unmixing methods. Hypergraph is often taken as an effective manifold structure to represent spatial-spectral relationships among pixels[6-7]. In this paper, a robust 3D convolutional

autoencoder (R3dCAE) unmixing framework was developed with spatial constraint. Here a deep denoising 3-dimensional convolutional autoencoder was designed to cope with the high noise disturbance. Furthermore, a non-negative sparse autoencoder was attached subsequently. To fully exploit the latent spatial contextual information, a hypergraph structure was constructed to preserve local consistency of abundances. $l_{2,1}$-norm regularizer was also imposed to improve the sparsity of the solution. Hidden endmembers and abundances can be obtained properly by training the network with the denoised data.

1　R3dCAE Model for Hyperspectral Unmixing

The proposed R3dCAE model concatenated a denoising 3D convolutional autoencoder (denoising 3D CAE) with a non-negative sparse autoencoder (NNSAE) network (Fig.1). In the proposed method, a set of corrupted data was used to train the denoising 3D CAE to recover original signals. Furthermore, an NNSAE was attached which incorporates the non-negative and sparsity constraints as well as a hypergraph regularizer for accurate estimation of endmembers and abundances.

1.1　Denoising Autoencoder

The denoising autoencoder attempts to recover the clean data from the noisy input using the coupled encoder-decoder network. The encoder projects the data x into low-dimensional embedding $z = f_{\theta}(x) = \sigma(W_e x + b_e)$, and the decoder retrieves the original data $\hat{x} = g_{\theta'}(z) = \sigma(W_d z + b_d)$ conversely in an unsupervised manner. $\theta = \{W_e, b_e\}$ and $\theta' = \{W_d, b_d\}$ denote the trainable parameters, respectively, and $\sigma(\cdot)$ is the activation function. The denoising autoencoder generates a group of corrupted samples $\tilde{x}_i, i = 1, 2, \cdots, n$ with a stochastic mapping $\tilde{x} \sim q_D(\tilde{x} | x)$, and recovers the original signal as

$$\min_{\theta,\theta'} J(\theta,\theta') = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \| g_{\theta'}(f_{\theta}(\tilde{x}_i)) - x_i \|_2^2 \quad (1)$$

Here masking noise of factor ν was adopted in the sample process, i.e., a fraction ν of the elements of x was randomly chosen to be 0. The denoising autoencoder can be interpreted as to define and learn a low-dimensional manifold for high dimensional input and capture the main variations in the data[8]. Representative high-order features can be extracted by nonlinear data transformation.
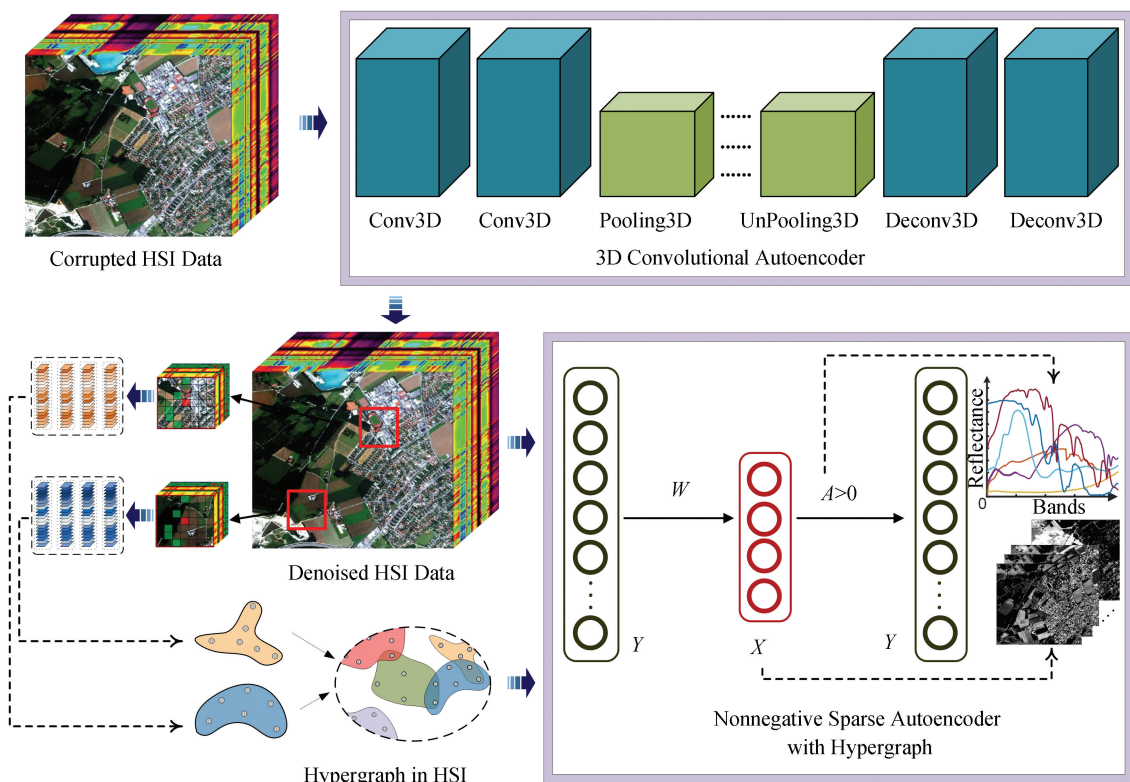

Fig.1　The block diagram of the proposed deep network for HSI unmixing

1.2　Deep 3D Denoising Convolutional Autoencoder

To reduce the negative impact by noise disturbance, here a denoising 3D convolutional autoencoder was constructed, and the volumetric representation of the corrupted data was fed into the network directly.

The HSI data is a third order tensor containing multiple channels to store spectral signatures, and the high spectral and spatial correlation makes it difficult for quantitative analysis of HSIs. Therefore, here a 3D convolution network capable of exploiting the spectral-spatial correlation in HSI was constructed. Principle and robust features from spatial context and spectral signature can be extracted automatically using 3D convolution layers and 3D pooling layers[9].

In the proposed network, the encoder was constructed with 6 convolutional layers and 3 maxpooling layers, and the decoder was designed with 6 deconvolution layers and 3 unpooling layers inversely. ReLU was chosen as the activation function[10]. Table 1 shows the structure of proposed R3dCAE network for the input synthetic data $Y \in \mathbb{R}^{26 \times 26 \times 188}$.

Table 1　Structure of proposed R3dCAE network for hyperspectral unmixing

Network	Layer	Filter	Units	Channel
	Input	–	(26,26,188)	–
	3D conv	(3,3,5)	(26,26,188)	16
	3D conv	(3,3,5)	(24,24,184)	16
	maxpool	–	(12,12,92)	–
	3D conv	(3,3,3)	(10,10,90)	32
	3D conv	(3,3,3)	(10,10,90)	32
	maxpool	–	(5,5,45)	–
	3D conv	(2,2,2)	(4,4,44)	96
	3D conv	(2,2,2)	(4,4,44)	96
3D CAE	maxpool	–	(2,2,22)	–
	maxunpool	–	(4,4,44)	–
	3D deconv	(2,2,2)	(4,4,44)	96
	3D deconv	(2,2,2)	(5,5,45)	96
	maxunpool	–	(10,10,90)	–
	3D deconv	(3,3,3)	(10,10,90)	32
	3D deconv	(3,3,3)	(12,12,92)	32
	maxunpool	–	(24,24,184)	–
	3D deconv	(3,3,5)	(26,26,188)	16
	3D deconv	(3,3,5)	(26,26,188)	16
	Output	–	(26,26,188)	–
	Input	–	188	–
NNSAE	Hidden	–	5	–
	Output	–	188	–

1.3　NNSAE with Hypergraph Structure

In the linear model, HSI data can be expressed as

$$Y = AX + N \qquad (2)$$

where $Y \in \mathbb{R}^{L \times N}$ is the observation matrix containing N pixels with L bands, $A = [a_1, \cdots, a_m]$ is the endmember matrix with p signatures, $X = [x_1, \cdots, x_N]^T$ is the abundance matrix, and N is the noise matrix. The abundance non-negativity constraint (ANC) $X \geqslant 0$, and the abundance sum-to-one constraint (ASC) $\mathbf{1}_p^T X = \mathbf{1}_N^T$, are imposed for physical meaning. Operator " $\geqslant$ " denotes element greater relationship.

Autoencoders are unsupervised network models to transform data into low-dimension representations and learn efficient features. Generally, a simple three-layer autoencoder learns the latent representations by the encoder and then recovers the input by the decoder. Consequently, the unmixing problem can be expressed by the autoencoder model. In the unmixing autoencoder, the latent representation X can be acquired from the input Y in the encoding part $X = f(Y) = \sigma(W \cdot Y)$, and the number of hidden units is specified as the number of endmembers. In the decoding part $g(X) = AX$, the weight matrix A is regarded as the set of endmembers, and the output of the encoder X from hidden layer serves as the abundances associated with the endmembers.

Meanwhile, in the proposed work, a hypergraph structure was constructed to exploit spatial contextual information. A hypergraph is an extended undirected graph where the edges can hold any number of vertices. Let vertex set $V = \{v_1, v_2, \ldots, v_N\}$ be a finite set with N elements. The set $E = \{e_1, e_2, \ldots, e_M\}$ is the edge set with M edges defined on V, and each edge e_j is a subset of V. Define $G = (V, E)$ as a hypergraph on V such that

$$\bigcup_{j=1}^{M} e_j = V, e_j \neq \varnothing, j = 1, 2, \cdots, M \qquad (3)$$

The edge e_j in a hypergraph $G = (V, E)$ can be associated with a positive weight $w(e_j)$. Denote W_s as the diagonal weight matrix with diagonal entries $w(e_j), j = 1, 2, \cdots, M$. The degree of vertex v is the summed weights of associated edges, defined as $d(v) = \Sigma_{\{e \in E | v \in e\}} w(e)$, while the degree of edge e is the cardinality of the set e, defined as $\delta(e) = |e|$. The laplacian matrix is defined as[11]

$$L_H = D_v - H W_s (D_e)^{-1} H^T \qquad (4)$$

where H is the incidence matrix, D_v and D_e are the diagonal matrices with the diagonal entries $(D_v)_{ii} = d(v_i)$, $(D_e)_{jj} = \delta(e_j)$, respectively, and W_s is the

weight matrix. The incidence matrix $H \in \mathbb{R}^{|V| \times |E|}$ represents incidence relationship between a vertex e_j and an edge v_i. The entry $h_{ij} \in H$ is set to 1 if $v_i \in e_j$, 0 otherwise.

In order to represent the relationship among pixels, a k-uniform hypergraph structure was constructed in the HSI with each edge containing K vertices. The spectrum $y_i \in Y$ was deemed as one vertex v_i in set V. The edge e_i with respect to v_i was generated by including v_i and its K neighbors in the square local region with side length d. For simplicity, K was set equal to d. The attached weight which described the similarity of pixels in spectral domains was calculated by

$$w(e_i) = \sum_{y_j \in e_i} \exp\left(-\frac{\| y_i - y_j \|_2^2}{\sigma^2} \right) \qquad (5)$$

where $\exp(\cdot)$ denotes the exponential function, and $\sigma = \sum_{i=1}^{N} \sum_{y_j \in e_i} \| y_j - y_i \| / KN$ denotes the average distance of all pixels in a local area. A larger weight was attached to the edge if the inclusive pixels had higher similarities. A convex hypergraph regularizer was designed to hold the abundance consistency among pixels in the same edge as

$$J_{HG}(X) = \sum_{e \in E} \sum_{|i,j| \in e} \frac{w(e)}{\delta(e)} \| x_i - x_j \|_F^2 = \text{Tr}(X L_H X^{\text{T}}) \qquad (6)$$

The network learned the weights and representations that reconstructed the data by minimizing the reconstruction error as well as the hypergraph regularizer. Furthermore, a $l_{2,1}$-norm penalty was imposed on W to encourage small number of nonzero lines of W and promote row sparsity[12]. Therefore, the cost function was defined as

$$J(A, W) = \frac{1}{2} \| A\sigma(WY) - Y \|_2^2 + \lambda_{HG} \text{Tr}(\sigma(WY) L_H \sigma(WY)^{\text{T}}) + \gamma \| W \|_{2,1} \qquad (7)$$

In the implementation the block-coordinate descent scheme was adopted for optimization. Each time one matrix was updated while other variables were fixed. The updating rule for matrix A is given as

$$A = \max(A - \alpha \nabla A, 0) \qquad (8)$$
$$\nabla A = (AX - Y)X^{\text{T}}, \ X = \sigma(WY) \qquad (9)$$

The learning rate α was determined following the Armijo rule[5]. W was estimated by back propagation as

$$W^{k+1} = W^k - \beta \nabla W^k \qquad (10)$$

where

$$\nabla W^k = (A^{\text{T}}(AX - Y) + 2\lambda_{HG} X L_H) \cdot \sigma^{-1}(Y) + \gamma \text{diag}((W.^2 * \mathbf{1}_L).^{1/2}) W \qquad (11)$$

In the training process, the matrices of A and X were initialized by VCA[13] and FCLS[14] algorithms and iteratively updated until convergence is approached.

2 Experimental Results and Analysis

In this section, experiments were designed on both synthetic and real-world data to illustrate the unmixing performance. Parameters for the proposed network were set as follows: $\nu = 0.3, \alpha = 1, \beta = 10^{-3}$, $\gamma = \lambda_{HG} = 10^{-6}, K = 5$. Results of MVSA[15], RCoNMF[16], MVCNMF[17], SGSNMF[18], uDAS[5] were also given for comparative purposes. For quantitative comparison, unmixing results were evaluated with spectral angle distance (SAD) in angles and root-mean-square error (RMSE).

2.1 Unmixing Experiments on Synthetic Data

The synthetic image data containing 26×26 pixels was linearly generated from 5 endmembers selected from signatures in USGS spectral library[19]. The image space was partitioned into several blocks containing pure pixels, and a low-pass filter was applied to construct the abundances. Then the abundances were normalized to satisfy sum-to-one constraint. After generating the synthetic data with endmembers and abundances, the scene was finally contaminated with i.i.d. Gaussian noise.

Table 2 and Table 3 show the unmixing performance of different algorithms for synthetic data, given in mean SAD and RMSE values, respectively, with 30 parallel tests. From Table 2, it can be observed that the proposed R3dCAE achieved lower spectral discrepancy for endmember extraction under different noise disturbances. In the R3dCAE framework, noises can be removed from the observed spectra, so that the non-negative autoencoder was able to obtain endmembers and abundances with higher accuracy by incorporating the sparsity constraint and the hypergraph regularizer. uDAS can have a comparable performance for endmember abstraction as it adopted the denoising constraint in the network. Unsatisfactory results by methods of MVSA, RCoNMF, MVCNMF, and SGSNMF at low SNR indicate that they were

highly sensitive to the noise. On the other hand, the results in Table 3 show that R3dCAE yielded lower RMSE values in abundance estimation, which also demonstrate its superiority over comparative methods in hyperspectral unmixing tasks.

For illustrative purposes, Fig.2 shows the comparison between the ground truth signatures and endmembers obtained by proposed R3dCAE framework when SNR = 20dB, which reveals a good match for endmember extraction. Fig.3 shows a graphical comparison of estimated abundances by different unmixing methods. Abundance maps by R3dCAE exhibit spatial consistency and contain less noisy pixels due to the introduction of hypergraph structure.

Table 2　SAD（in angles）comparison at different noise levels

SNR（dB）	MVSA	RCoNMF	MVCNMF	SGSNMF	uDAS	R3dCAE
5	51.71±5.42	44.55±1.67	45.89±11.54	14.43±0.08	15.39±2.42	4.91±0.83
10	18.94±0.00	17.22±0.32	10.54±3.44	8.24±0.53	4.98±0.88	2.51±0.39
15	16.09±0.15	12.28±0.17	7.28±2.36	4.45±0.30	2.81±0.14	2.05±0.07
20	6.17±0.00	5.17±0.05	2.90±0.00	2.98±0.41	1.11±0.05	1.09±0.05
25	5.12±0.00	3.89±0.01	2.88±0.00	2.58±0.53	0.85±0.03	0.99±0.04
30	2.08±0.00	1.49±0.01	3.42±0.01	2.42±0.47	1.64±0.01	0.48±0.02

Table 3　RMSE comparison at different noise levels

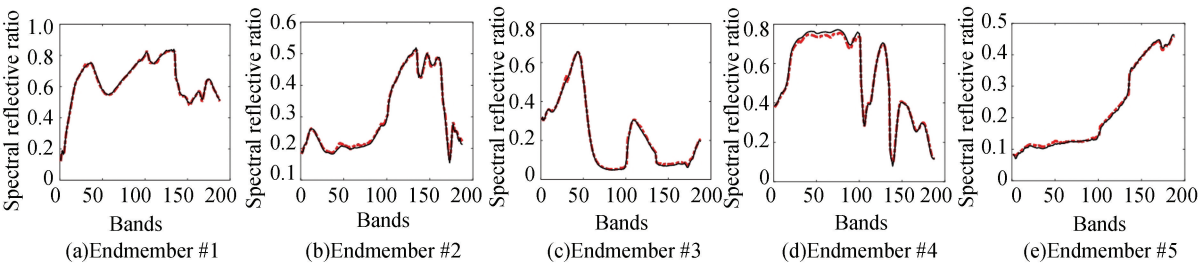| SNR（dB）	MVSA	RCoNMF	MVCNMF	SGSNMF	uDAS	R3dCAE
5	$1.07\pm0.11\times10^{-1}$	$1.07\pm0.12\times10^{-1}$	$1.93\pm0.28\times10^{-1}$	$4.30\pm0.09\times10^{-2}$	$1.24\pm0.20\times10^{-1}$	$1.50\pm0.49\times10^{-2}$
10	$3.37\pm0.00\times10^{-2}$	$2.63\pm0.05\times10^{-2}$	$1.54\pm0.07\times10^{-2}$	$1.50\pm0.15\times10^{-2}$	$1.37\pm0.67\times10^{-2}$	$2.58\pm0.48\times10^{-3}$
15	$1.99\pm0.02\times10^{-2}$	$1.43\pm0.01\times10^{-2}$	$3.31\pm0.47\times10^{-3}$	$4.81\pm1.01\times10^{-3}$	$5.42\pm0.24\times10^{-3}$	$1.35\pm0.10\times10^{-3}$
20	$8.16\pm0.00\times10^{-3}$	$4.45\pm0.05\times10^{-3}$	$1.84\pm0.00\times10^{-3}$	$1.36\pm0.17\times10^{-3}$	$1.33\pm0.04\times10^{-3}$	$5.21\pm0.12\times10^{-4}$
25	$3.81\pm0.00\times10^{-3}$	$2.32\pm0.00\times10^{-3}$	$1.67\pm0.10\times10^{-3}$	$1.22\pm0.14\times10^{-3}$	$7.60\pm0.27\times10^{-3}$	$3.38\pm0.19\times10^{-4}$
30	$7.50\pm0.00\times10^{-4}$	$3.93\pm0.00\times10^{-4}$	$2.60\pm0.23\times10^{-3}$	$4.81\pm1.59\times10^{-3}$	$1.13\pm0.01\times10^{-4}$	$\mathbf{7.88\pm0.64\times10^{-5}}$

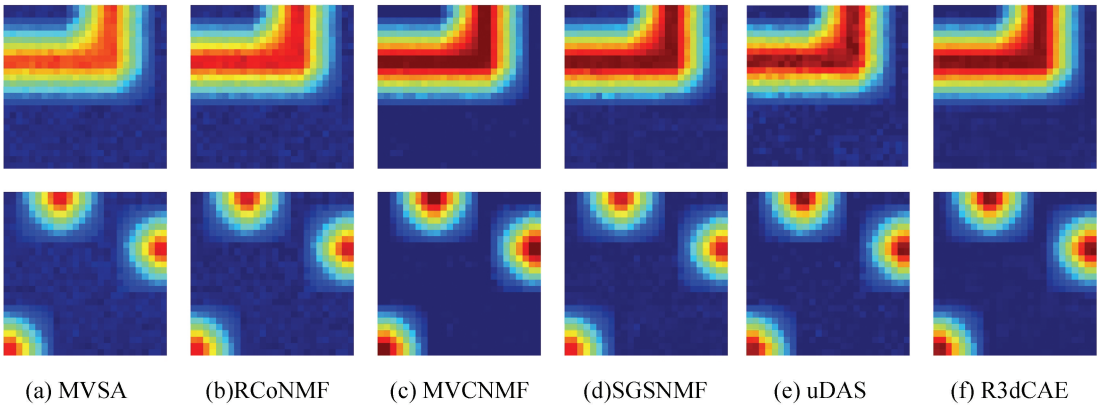Fig.2　Endmembers by the proposed R3dCAE（black），compared with ground truth signatures（red）with SNR = 20 dB

Fig.3　Estimated abundance maps of Endmember #1（Top Row）and #3（Bottom Row）on synthetic data by the proposed R3dCAE, compared with methods of MVSA, RCoNMF, MVCNMF, SGSNMF, and uDAS with SNR = 20 dB

In summary，the experimental results reveal that proposed R3dCAE outperformed other unmixing techniques. This is because the deep denoising network architecture can hierarchically extract robust and effective local features in spectral and spatial dimensions with 3D operations，and reconstruct high-quality denoised data. Additionally，incorporating spatial information by hypergraph structure in the NNSAE improves the unmixing performances. The $l_{2,1}$-norm regularizer was also introduced for sparsity enhancement of solution. However，it is worth noting that the proposed R3dCAE method does not rely on a specific model for noise generation. The Gaussian process can model the major part of the additive noise，including the thermal noise and quantization noise in HSI data[20]，and the experiments reveal the effectiveness of proposed framework for images with Gaussian noise. For HSIs contaminated by signal dependent noise and pattern noise，R3dCAE model may yield inferior results compared with methods by modeling the signal noise strictly. In addition，a proper parameter λ_{HG} should be selected elaborately to achieve the optimal unmixing performance. The presence of outliers would also bring negative impact on accurate endmembers and abundances estimation.

2.2　Unmixing Experiments on Real Data

The well-known AVIRIS Cuprite dataset was adopted to validate the unmixing performance of the proposed algorithm in this experiment. The scene consisted of 224 spectral bands ranging from 0.4 μm to 2.5 μm，with nominal spectral resolution of 10 nm. Here a 250×191 pixel subset of the scene was utilized as it contains representative mineral spectral signatures for unmixing. Low SNR and water vapor absorption bands were removed and the remaining bands were employed to study the unmixing performances of different methods.

To compare the capacity for endmember extraction on real Cuprite data，Table 4 shows the SAD values of endmembers obtained by considered unmixing algorithms on different minerals. It is observed that the proposed unmixing autoencoder network R3dCAE yielded lower spectral discrepancy overall，which suggests that it has a strong capacity for extracting endmembers in the mixed hyperspectral data. Fig. 4 also illustrates the fractional distribution of abundances acquired by R3dCAE，and the abundance maps exhibit good discriminability and present condensation appearance for different materials. On the other hand，methods of MVC-NMF and SGSNMF can achieve good unmixing performances，while the unsatisfactory results by MVSA and RCoNMF indicate that those methods are insufficient to identify all endmembers present in the data.

Table 4　Results obtained by different unmixing algorithms for AVIRIS Cuprite dataset

Category	MVSA	RCoNMF	MVC-NMF	SGSNMF	uDAs	R3dCAE
Alunite	14.1463	16.6959	11.6773	18.1827	7.8019	6.1215
Andradite	33.1832	24.1872	6.3295	3.7974	7.6458	5.3372
Buddingtonite	11.4212	24.1960	5.5772	7.6853	4.3761	7.2741
Dumortierite	26.8598	32.3027	7.6590	8.5687	8.4989	6.9441
Kaolinite-1	11.6595	21.1149	5.7043	4.9133	7.0262	6.3960
Kaolinite-2	7.4447	29.9843	5.7167	6.2882	5.1785	3.4954
Muscovite	15.4013	52.7701	7.3229	9.2725	7.6617	5.9764
Montmorillonite	8.5258	6.1461	5.2050	3.7206	4.7154	6.3705
Nontronite	10.7342	18.2927	5.2006	5.1781	6.0138	5.0948
Pyrope	23.6987	13.6850	5.6561	3.6352	6.4316	5.1538
Sphene	46.8019	9.2949	4.6725	6.5439	7.6966	4.5265
Chalcedony	4.4446	12.0125	9.0900	7.8414	9.3906	4.7248
Mean	17.8601	21.7235	6.6509	7.1356	6.8698	5.6179

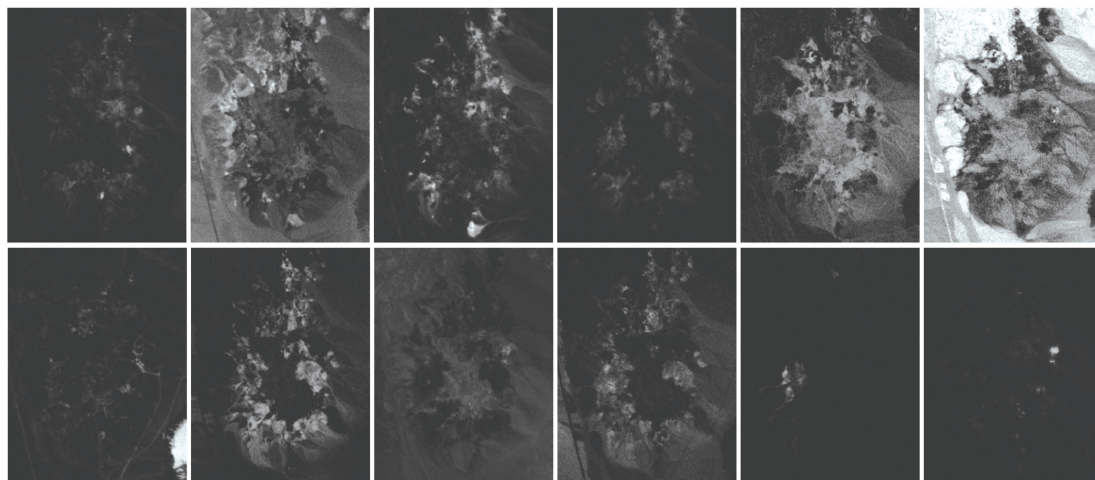**Fig. 4 Estimated abundance maps of the AVIRIS Cuprite data by the proposed framework（Top row：Alunite，Andradite，Buddingtonite，Dumortierite，Kaolinite-1，and Kaolinite-2；Bottom row：Muscovite，Montmorillonite，Nontronite，Pyrope，Sphene，and Chalcedony）**

3 Conclusions

In hyperspectral unmixing problems，the noise disturbance would bring negative impact on the estimation of endmembers and abundances in practical application. To cope with high noise disturbance in HSI data，a robust 3D convolutional autoencoder unmixing framework was proposed for hyperspectral unmixing problem in this paper. A denoising 3-dimensional convolutional autoencoder was designed to retrieve original spectral signatures，and a restrictive non-negative sparse autoencoder with spatial constraint was followed for endmember extraction and abundance estimation. The encoding and decoding processes by 3D convolutional operations make it possible to extract joint spectral-spatial information hierarchically for clean data retrieval. Then the non-negative autoencoder was attached subsequently with sparsity constraint. Furthermore， the spatial-contextual information was exploited to improve the robustness of the network and the hypergraph structure was constructed in the unmixing scheme to preserve the spatial consistency of abundances. Comparative experiments were conducted on synthetic data and real data，which reveals the superiority of proposed network architecture for hyperspectral unmixing.

References

［1］Guo R，Wang W，Qi H R. Hyperspectral image unmixing using autoencoder cascade. Proceedings of the 2015 7th Workshop on Hyperspectral Image and Signal Processing：Evolution in Remote Sensing（WHISPERS）. Piscataway：IEEE，2015. 1-4. DOI：10.1109/whispers.2015.8075378.

［2］Su Y C，Marinoni A，Li J，et al. Nonnegative sparse autoencoder for robust endmember extraction from remotely sensed hyperspectral images. Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium（IGARSS）. Piscataway：IEEE，2017. 205-208. DOI：10.1109/igarss.2017.8126930.

［3］Su Y C，Li J，Plaza A，et al. Deep Auto-Encoder Network for Hyperspectral Image Unmixing. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium（IGARSS）. Piscataway：IEEE，2018. 6400-6403. DOI：10.1109/igarss.2018.8519571.

［4］Ozkan S，Kaya B，Akar G B. Endnet：sparse autoencoder network for endmember extraction and hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing，2019，57（1）：482-496. DOI：10.1109/tgrs.2018.2856929.

［5］Qu Y，Qi H R. uDAS：an untied denoising autoencoder with sparsity for spectral unmixing. IEEE Transactions on Geoscience and Remote Sensing，2019，57（3）：1698-1712. DOI：10.1109/tgrs.2018.2868690.

［6］Jia P Y，Zhang M，Shen Y. Hypergraph learning and reweighted l1-norm minimization for hyperspectral unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing，2019，12（6）：1898-1904. DOI：10.1109/jstars.2019.2916058.

［7］Wang W H，Qian Y T，Tang Y Y. Hypergraph-regularized sparse NMF for hyperspectral unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing，2016，9（2）：681-694. DOI：10.1109/jstars.2015.2508448.

［8］Vincent P，Larochelle H，Bengio Y，et al. Extracting and composing robust features with denoising autoencoders.

Proceedings of the 25th International Conference on Machine Learning (ICML 2008). Eqaila, Kuwait：ACM, 2008. 1096−1103. DOI：10.1145/1390156.1390294.

[9] Ji J Y, Mei S H, Hou J H, et al. Learning sensor-specific features for hyperspectral images via 3-dimensional convolutional autoencoder. Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Piscataway：IEEE, 2017. 1820−1823. DOI：10.1109/igarss.2017.8127329.

[10] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS). Ft Lauderdale, 2011. 315−323. DOI：10.1.1.208.6449.

[11] Zhang Z H, Bai L, Liang Y H, et al. Joint hypergraph learning and sparse regression for feature selection. Pattern Recognition, 2017, 63：291−309. DOI：10.1016/j.patcog.2016.06.009.

[12] Iordache M-D, Bioucas-Dias J M, Plaza A. Collaborative sparse regression for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1)：341−354. DOI：10.1109/tgrs.2013.2240001.

[13] Nascimento J M P, Dias J M B. Vertex component analysis：a fast algorithm to unmix hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(4)：898−910. DOI：10.1109/tgrs.2005.844293.

[14] Heinz D C, Chang C I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 2001, 39(3)：529−545. DOI：10.1109/36.911111.

[15] Li J, Agathos A, Zaharie D, et al. Minimum volume simplex analysis：a fast algorithm for linear hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(9)：5067−5082. DOI：10.1109/tgrs.2015.2417162.

[16] Li J, Bioucas-Dias J M, Plaza A, et al. Robust collaborative nonnegative matrix factorization for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(10)：6076−6090. DOI：10.1109/tgrs.2016.2580702.

[17] Miao L D, Qi H R. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing, 2007,45(3)：765−777. DOI：10.1109/tgrs.2006.888466.

[18] Wang X Y, Zhong Y F, Zhang L P, et al. Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(11)：6287−6304. DOI：10.1109/tgrs.2017.2724944.

[19] Bioucas-Dias J M, Plaza A, Dobigeon N, et al. Hyperspectral unmixing overview：geometrical, statistical, and sparse regression-based approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2012, 5(2)：354−379. DOI：10.1109/JSTARS.2012.2194696.

[20] Landgrebe D A, Malaret E. Noise in remote-sensing systems：the effect on classification error. IEEE Transactions on Geoscience and Remote Sensing, 1986, GE−24(2)：294−300. DOI：10.1109/tgrs.1986.289648.