Nonlinear Unmixing of Hyperspectral Data via Deep Autoencoder Networks

Mou Wang, *Student Member, IEEE*, Min Zhao, *Student Member, IEEE*, Jie Chen, *Senior Member, IEEE*, and Susanto Rahardja, *Fellow, IEEE*

***Abstract*— Nonlinear spectral unmixing is an important and challenging problem in hyperspectral image processing. Classical nonlinear algorithms are usually derived based on specific assumptions on the nonlinearity. In recent years, deep learning shows its advantage in addressing general nonlinear problems. However, existing ways of using deep neural networks for unmixing are limited and restrictive. In this letter, we develop a novel blind hyperspectral unmixing scheme based on a deep autoencoder network. Both encoder and decoder of the network are carefully designed so that we can conveniently extract estimated endmembers and abundances simultaneously from the nonlinearly mixed data. Because an autoencoder is essentially an unsupervised algorithm, this scheme only relies on the current data and, therefore, does not require additional training. Experimental results validate the proposed scheme and show its superior performance over several existing algorithms.**

***Index Terms*— Autoencoder network, deep learning, hyperspectral imaging, nonlinear spectral unmixing.**

I. Introduction

HYPERSPECTRAL imaging is a continuously growing field of study with a wide range of applications, and it has received considerable attention in the past decade. However, because of the low spatial resolution of imaging devices and the diversity of materials in a scene under evaluation, an observed pixel may contain several materials. Therefore, spectral unmixing is an important task to analyze remotely sensed hyperspectral data, aiming at separating a pixel spectrum into a set of spectral signatures, termed endmembers, and a set of fractional abundances associated with each endmember [1].

In the past decade, spectral unmixing has been extensively studied with classical signal processing and optimization techniques. Endmember identification and abundance estimation can be conducted either in a sequential or in a simultaneous manner [2]. In the pipeline of the sequential technique, endmembers are first determined by using an endmember extraction algorithm, such as N-finder algorithm (N-FINDR) [3], vertex component analysis (VCA) [4] and minimum volume simplex analysis. In this sequel, the problem boils down to determine fractional abundances. Simultaneous techniques are conducted by using nonnegative matrix decomposition (NMF) methods [5], [6] or sparse learning methods [7] to jointly infer the endmembers and evaluate abundances. It is also important to note that the mixture can be classified into linear and nonlinear models. The linear model assumes that the materials are spatially distributed in a checkerboard manner, and high-order interactions of photons can be ignored [8], [9]. Although the linear mixture model is simple and physically interpretable, there are many complex situations in which it may not be appropriate and can be advantageously replaced by a nonlinear one [10], [11].

In recent years, deep learning has shown its superior performance in addressing various nonlinear problems compared to classical methods. Researchers have also investigated to use deep neural networks in hyperspectral image analysis, such as hyperspectral classification [12], [13], and change detection [14]. Only a handful of works studied the unmixing problem with deep neural networks. The works in [15] and [16] perform unmixing based on classifier models; hence, they need a training set with known ground truth, which is often generated by other unmixing approaches. The works in [17]–[22] estimate both endmembers and abundances with variations of autoencoder networks for blind spectral unmixing. However, all these works are restricted to address the linear unmixing problem. For instance, in [22], the decoder uses one linear layer to reproduce the linear mixture model. These solutions are, thus, less attractive, as neural networks are considered more meaningful in solving nonlinear problems, while the linear unmixing can readily be performed with classical methods. In addition, most of the existing autoencoders are proposed with many auxiliary designs in order to satisfy the abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC). For example, many works use a tied-weight structure, which is inappropriate to solve the unmixing problem, as shown in [20].

In this letter, we propose a novel scheme for blind nonlinear unmixing based on a carefully designed deep autoencoder network. Inspired by the widely used postnonlinear model, we provide a new design of the encoder and design a specific deep decoder so that the abundance vectors and endmembers can be extracted in a tractable manner, with a pretraining technique to enhance the performance. Finally, note that unlike [15] and [16], an autoencoder is an unsupervised algorithm so that the estimates are readily obtained once the autoencoder is established, without resorting to extra training data. The experiments with synthetic and real data show that

Manuscript received November 30, 2018; revised January 28, 2019; accepted February 15, 2019. The work of J. Chen was supported in part by NSFC under Grant 61671382 and Grant 61811530283, in part by NSF of Shenzhen under Grant JCYJ20170301555315873, and in part by 111 Project under Grant B18041. *(Mou Wang and Min Zhao contributed equally to this work.) (Corresponding author: Jie Chen.)*

The authors are with the School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China, also with the Key Laboratory of Ocean Acoustics and Sensing, Ministry of Industry and Information Technology, Xi'an 710072, China, and also with the Research and Development Institute, Northwestern Polytechnical University, Shenzhen 518057, China (e-mail: dr.jie.chen@ieee.org).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2019.2900733

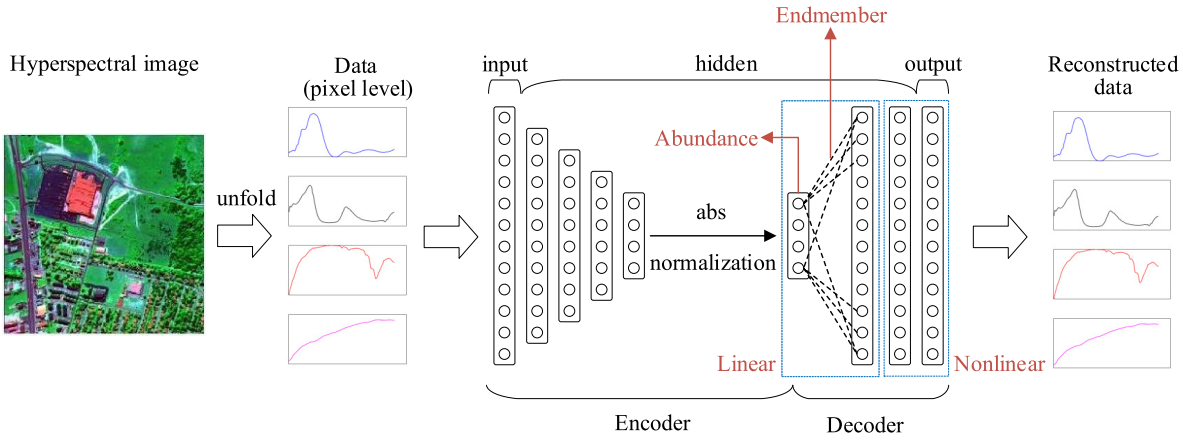

Fig. 1. Schematic of the proposed system.

our proposed scheme exhibit superior performance compared to typical and state-of-the-art algorithms.

II. Data Model

Let column vector $\mathbf{x} \in \mathbb{R}^B$ denote an observed pixel consisting of B spectral bands, and $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_R]$ denote the $(B \times R)$ endmember matrix, with each column $\mathbf{m}_i$ being a spectral signature. Let $\mathbf{a} = [a_1, a_2, \ldots, a_R]^\top \in \mathbb{R}^R$ be the abundance vector associated with the pixel.

The linear mixing model assumes that any observed pixel is a linear combination of the endmembers, weighted by the fractional abundances, that is,

$$\mathbf{x} = \mathbf{M}\mathbf{a} + \mathbf{n} \tag{1}$$

with $\mathbf{n} \in \mathbb{R}^B$ being a noise vector. Considering the physical interpretation of the fractional abundances, it is often suggested that abundance vector $\mathbf{a}$ satisfies ANC and ASC given by

$$a_i \geq 0, \quad \forall i \in \{1, \ldots, R\} \tag{2}$$

$$\sum_{i=1}^{R} a_i = 1. \tag{3}$$

There are, however, situations, involving multiple scattering effects, in which model (1) may be inappropriate and could be advantageously replaced by a nonlinear one. Consider the general mixing mechanism

$$\mathbf{x} = \Phi(\mathbf{M}, \mathbf{a}) + \mathbf{n} \tag{4}$$

where Φ is an implicit function that defines nonlinear interactions between the endmembers in matrix $\mathbf{M}$ parameterized by $\mathbf{a}$. In order to extract the abundance and endmembers in a tractable manner, we consider the following postnonlinear structure:

$$\mathbf{x} = \Psi(\mathbf{M}\mathbf{a}) + \mathbf{n} \tag{5}$$

where Ψ represents the nonlinear function applied to the linear transform $\mathbf{Ma}$. This post nonlinear relation is widely used, several models [23]–[25], including the physic-based Hapke model [26], can be considered as its specific cases. Note that although we focus on this postnonlinear form, our scheme does not rely on any specific model.

III. Proposed Framework

The structure of an autoencoder network consists of two building blocks, namely, an encoder and a decoder. Encoder f_E compresses the input $\mathbf{x}$ into a low-dimensional representation $\mathbf{v}$, i.e., $\mathbf{v} = f_E(\mathbf{x})$. Then, decoder f_D uncompresses the encoded vector to reconstruct the original data $\hat{\mathbf{x}}$ as faithfully as possible, i.e., $\hat{\mathbf{x}} = f_D(\mathbf{v})$. The network is trained to minimize a loss $\mathcal{L}(\mathbf{x}, \hat{\mathbf{x}})$ between the input $\mathbf{x}$ and the reconstructed $\hat{\mathbf{x}}$. Both the encoder and decoder can either be shallow or deep, but generally, it is believed that deep networks possess a superior modeling capability.

The system scheme proposed in this letter is shown in Fig. 1. We propose to design a decoder with two parts, namely, a linear part and a nonlinear part. The weights of the linear part are coded in the matrix $\mathbf{W}$, and the input–output relationship governed by the nonlinear part is denoted by $\Psi'(\cdot)$. Denoting the compressed vector from encoder by $\mathbf{v}$, we then have

$$\hat{\mathbf{x}} = f_D(\mathbf{v}) = \Psi'(\mathbf{W}\mathbf{v}) + \mathbf{n}. \tag{6}$$

Once the entire network is properly trained, the autoencoder network can effectively encode and reconstruct the hyperspectral data. It is important to note that (6) matches the model described by (5), we can interpret the output $\mathbf{v}$ of the encoder network as the estimate of the abundance vector $\mathbf{a}$, and weights $\mathbf{W}$ of the linear part in the decoder as the extracted endmember matrix $\mathbf{M}$, i.e., the unmixing is performed by

$$\text{Abundance estimation:} \quad \hat{\mathbf{a}} \leftarrow \mathbf{v} \tag{7}$$

$$\text{Endmember extraction:} \quad \hat{\mathbf{M}} \leftarrow \mathbf{W}. \tag{8}$$

The nonlinear part $\Psi'(\cdot)$ represents the nonlinear interactions among the endmembers. We shall now elaborate the design of the encoder and decoder in the following sections.

A. Encoder

In this letter, a regular deep network is used as the encoder with its structure reported in the upper part of Table I. No further specific constraints are imposed on this encoder in order to fully use the capacity of the network. The four fully connected layers gradually narrow down from the input layer to deeper layers. The node number of the last hidden layer is equal to the dimension of the endmembers. In Table I, the dimensions of the input layer and the last hidden layer are

TABLE I
STRUCTURE OF THE PROPOSED AUTOENCODER NETWORK

	Layers		Activation function	units
Encoder	Input layer		-	224
	Hidden layer		ϕ	128
	Hidden layer		ϕ	64
	Hidden layer		ϕ	16
	Hidden layer		-	4
	abs + normalization			
Decoder	Linear part	Hidden layer	-	224
	Nonlinear part	Hidden layer	ϕ	224
		Output layer	-	224

φ can be ReLU, LReLU, and Sigmoid. The unit numbers of each layer depend on experiment data, for example, the settings of network in this table correspond to experiment on synthetic data in our work.

TABLE II
RMSE COMPARISON

	SNR=10dB		
	linear	bilinear	PNMM
MVC-NMF	0.0754 ± 0.0193	0.0936 ± 0.0281	0.0664 ± 0.0142
LinearAE	0.0790 ± 0.0216	0.2678 ± 0.2203	0.2643 ± 0.2143
VCA-K-Hype	0.0857 ± 0.0063	0.1031 ± 0.0097	0.0452 ± 0.0014
VCA-NDU	0.0828 ± 0.0226	0.0942 ± 0.0321	0.0464 ± 0.0059
N-FINDR-MLM	0.1266 ± 0.0445	0.1210 ± 0.0422	0.0504 ± 0.0058
Proposed	$\mathbf{0.0564 \pm 0.0011}$	$\mathbf{0.0621 \pm 0.0010}$	$\mathbf{0.0364 \pm 0.0016}$
	SNR=20dB		
	linear	bilinear	PNMM
MVC-NMF	0.0261 ± 0.0024	0.0892 ± 0.0179	0.0380 ± 0.0043
LinearAE	0.0670 ± 0.0092	0.2752 ± 0.2274	0.2763 ± 0.2340
VCA-K-Hype	0.0515 ± 0.0021	0.0594 ± 0.0025	0.0443 ± 0.0013
VCA-NDU	0.0608 ± 0.0118	0.0670 ± 0.0181	0.0495 ± 0.0066
N-FINDR-MLM	0.0671 ± 0.0088	0.0708 ± 0.0088	0.0520 ± 0.0058
Proposed	$\mathbf{0.0241 \pm 0.0016}$	$\mathbf{0.0427 \pm 0.0012}$	$\mathbf{0.0373 \pm 0.0011}$
	SNR=30dB		
	linear	bilinear	PNMM
MVC-NMF	$\mathbf{0.0104 \pm 0.0003}$	0.0942 ± 0.0144	0.0376 ± 0.0028
LinearAE	0.0607 ± 0.00063	0.2751 ± 0.2289	0.2767 ± 0.2357
VCA-K-Hype	0.0422 ± 0.0014	0.0698 ± 0.0041	0.0443 ± 0.0014
VCA-NDU	0.0535 ± 0.0089	0.0728 ± 0.00290	0.0476 ± 0.0060
N-FINDR-MLM	0.0595 ± 0.0029	0.0695 ± 0.0053	0.0497 ± 0.0057
Proposed	0.0211 ± 0.0014	$\mathbf{0.0427 \pm 0.0015}$	$\mathbf{0.0372 \pm 0.0010}$

Boldface numbers denote the lowest RMSEs

224 and 4, respectively, as we use data with $B = 224$ bands and $R = 4$ endmembers in the experiments; however, this structure can be adopted to any values of B and R. Except for the last hidden layer, the first three layers adopt activation functions ϕ, which can be rectified linear unit (ReLU), Leaky ReLU (LReLU), and Sigmoid.

The ANC and ASC constraints on $\mathbf{a}$ should be carefully addressed. In order to meet the ANC, the works in [20] and [22] use a threshold to enforce the vector to be nonnegative, and the work in [19] uses a nonnegative autoencoder to guarantee the ANC over the whole network. The former strategy deactivates a large number of nodes in the network, and the capability of network is thus not fully utilized. In addition, we find that it leads to oversparse abundance estimation, even if the true abundance vector is not sparse at all. The latter strategy imposes strong constraints on the network and makes it difficult to design the network. To ensure the abundance vector to meet these constraints, we propose to take the absolute value of the output of the encoder network, and then normalize this nonnegative vector by the sum of its entries. This operation can be written as

$$a_i = \frac{|a_i|}{\sum_{i=1}^{R} |a_i|}. \tag{9}$$

The structure of the proposed encoder is significantly simpler than that in [22]; however, it is more effective.

B. Decoder

A deep network is also adopted for the decoder in the proposed scheme, with the structure reported in the lower part of Table I. As mentioned, our proposed decoder consists of a linear part and a nonlinear part. The linear part of the decoder represents the linear model for spectral unmixing, and the associated weights are used to extract the endmembers. We initialize the weights with the endmember candidates obtained from the VCA algorithm. In addition, similar to [20], we do not use the bias in the linear part of the decoder.

The nonlinear part of decoder represents the nonlinear iterations among the endmembers. However, it is reasonable to assume that the linear model is a rough approximation of the mixture, and the effect of the nonlinear part can be considered as a fluctuation, sharing the similar principle as the model in [11]. Considering this, we pretrain the nonlinear part of the decoder using the same observed pixel as the input and output pair. After the pretraining phase, we fine-tune the decoder network when training the whole network with Adam. Different learning rates are used to train the model, for example, the learning rate of the encoder was set to 1×10^{-4}, and the learning rate of the nonlinear part of the decoder was set to 1×10^{-5} in experiments.

IV. EXPERIMENTS

In this section, experiments are conducted to implement the proposed unmixing scheme with both synthetic and real image data and compared the obtained results with some typical and state-of-the-art unmixing methods.

The accuracy of abundance estimation is measured via the root-mean-square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{NR} \sum_{i=1}^{N} \|\mathbf{a}_i - \hat{\mathbf{a}}_i\|^2} \tag{10}$$

where N represents the number of pixels, $\mathbf{a}_i$ is the true abundance, and $\hat{\mathbf{a}}_i$ is its estimate of the ith pixel.

Two widely used metrics, namely, spectral angle distance (SAD) and spectral information divergence (SID), are used to evaluate the similarity between a true endmember $\mathbf{m}$ and its estimate $\widehat{\mathbf{m}}$, respectively, given by

$$\text{SAD} = \cos^{-1}\left(\frac{\mathbf{m}^\top \widehat{\mathbf{m}}}{\|\mathbf{m}\| \|\widehat{\mathbf{m}}\|}\right) \tag{11}$$

and

$$\text{SID}(\mathbf{m}|\widehat{\mathbf{m}}) = \sum_j \mathbf{p}_j \log\left(\frac{\mathbf{p}_j}{\widehat{\mathbf{p}}_j}\right) \tag{12}$$

where $\mathbf{p} = (\mathbf{m}/\mathbf{1}^\top \mathbf{m})$ and $\widehat{\mathbf{p}} = (\widehat{\mathbf{m}}/\mathbf{1}^\top \widehat{\mathbf{m}})$.

A. Unmixing Performance on Synthetic Data

The synthetic data are generated by the linear, bilinear, and postnonlinear mixing models (PNMM, [24]) with several

TABLE III
SAD AND SID COMPARISON

		SNR=10dB			SNR=20dB			SNR=30dB		
		linear	bilinear	PNMM	linear	bilinear	PNMM	linear	bilinear	PNMM
SAD	MVC-NMF	**2.6789**	**4.0459**	5.9447	1.5529	3.0656	**4.7072**	0.6804	**3.0534**	**4.4064**
	LinearAE	4.5966	9.6726	8.0556	3.1852	4.7581	5.7560	4.2749	6.2164	6.9985
	VCA-K-Hype/NDU	3.2942	4.6773	**5.0810**	1.4615	2.4438	5.0424	0.6747	3.3831	4.9690
	N-FINDR-MLM	18.0984	18.0489	5.1694	5.9772	6.2460	5.0186	2.0760	3.1794	5.1652
	Proposed	3.4071	4.7289	5.4456	**1.3341**	**2.3112**	5.4361	**0.4415**	3.1273	5.3542
SID	MVC-NMF	**0.0041**	0.0226	0.0211	0.0016	0.0129	0.0113	0.0005	0.0094	**0.0090**
	LinearAE	0.0475	0.0550	0.0906	0.0079	0.0469	0.0299	0.0426	0.0605	0.0488
	VCA-K-Hype/NDU	0.0079	**0.0124**	**0.0140**	0.0014	0.0045	0.0133	0.0005	0.0087	0.0130
	N-FINDR-MLM	0.2103	0.2028	0.2124	0.0291	0.0346	**0.0112**	0.0036	0.0089	0.0123
	Proposed	0.0090	0.0125	0.0163	**0.0012**	**0.0037**	0.0158	**0.0001**	**0.0074**	0.0154

Boldface numbers denote the lowest SADs and SIDs

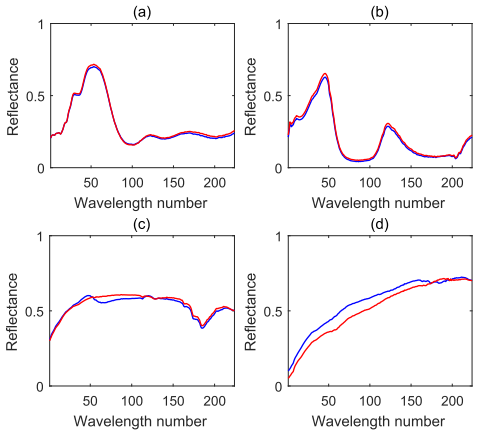

Fig. 2. Illustration of extracted endmembers. (a)–(b) Four endmembers. Red curves: ground truth. Blue curves: extracted endmembers (SNR = 20 dB, bilinear model case).

endmembers signatures from the United States Geological Survey (USGS) digital spectral library. These spectra consist of 224 contiguous bands. Four pure material spectra are selected as endmembers, and the abundance fractions are generated from Dirichlet distribution. A total number of 300 000 pixels are generated to evaluate the performance. Additive zero-mean Gaussian noise is added to the data with the signal-to-noise ratio (SNR) set to 10, 20 and 30 dB.

Table I shows the configurations used in this experiment. We set the learning rate as 1×10^{-4} and used the Adam optimizer to train the network. The batch size was set to 512, and the number of the training epoch was set to 50. Our proposed method is compared with the following: 1) a widely used blind unmixing method minimum volume constrained NMF (MVC-NMF); 2) a recently proposed deep autoencoder network based on the linear unmixing method (LinearAE) [22]; 3) endmember extraction with VCA and abundance estimation with K-Hype [11]; 4) VCA and abundance estimation with nonlinear neighbor and band dependent unmixing (NDU) [27]; and 5) endmember extraction with N-FINDR and abundance estimation with multilinear mixing model (MLM) [25]. Each experiment is conducted multiple times, and the means and variances of the results are recorded. RMSE, SAD, and SID of different unmixing methods are reported in Tables II and III for different mixing models under different SNRs. Fig. 2 illustrates the comparison between the obtained endmembers and the references in the USGS spectral library.

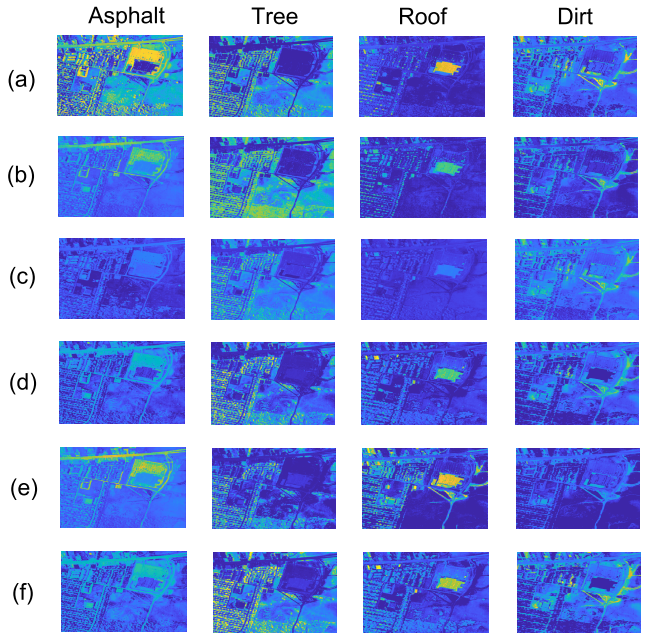

Fig. 3. Estimated abundance maps of four endmembers of (a) proposed method, (b) LinearAE, (c) MVC-NMF, (d) VCA-K-Hype, (e) N-FINDR-MLM, and (f) VCA-NDU.

These results show that our proposed method achieves the best abundance estimation performance and sufficiently good endmember estimation performance with both linear and nonlinear mixture models. Note that when the nonlinearity is not extremely high in the data, the linear methods for endmembers extraction can still provide sufficiently good results, especially when extra regularization such as the volume constraint is added. In addition, the proposed method is robust to noise, benefiting from the fact that the low-dimensional vector generated from encoder maintains the main information and gets rid of redundant information and noise. Therefore, no specific denoising layer as that in [21] is used. Our method is based on a nonlinear mixture model, but MVC-NMF and LinearAE are derived with the linear mixture model. Thus, the performance of the proposed method is better than that of MVC-NMF and LinearAE in all nonlinear cases. Both K-Hype and our method are the nonlinear unmixing algorithm. However, K-Hype uses the kernel-based method, and the selection of the kernel and its parameters highly affects its performance. The proposed method builds a model by learning from data, and therefore, we avoid the issues in the kernel selection.

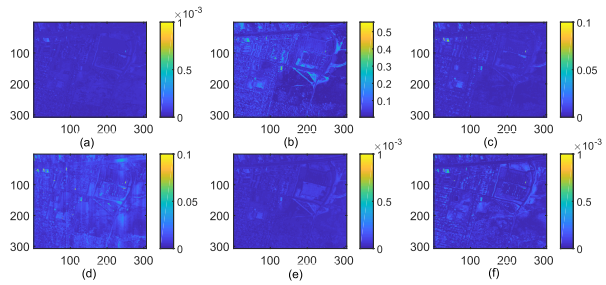

Fig. 4. Reconstruction error maps of (a) proposed method (RE = 0.0055), (b) LinearAE (RE = 0.0697), (c) MVC-NMF (RE = 0.0081), (d) VCA-K-Hype (RE = 0.0104), (e) N-FINDR-MLM (RE = 0.0115), and (f) VCA-NDU (RE = 0.0155).

B. Unmixing Performance on Real Image Scene

This section illustrates the unmixing results of the proposed method on real hyperspectral data. The widely used Urban dataset is used in our experiment, with the spatial size of the image being 307×307 pixels, and spectral dimension being 210 wavelengths ranged from 400 to 2500 nm. In our experiment, 162 channels are used where the channels affected by water absorption and atmospheric environment are removed. Six endmembers are involved in these data, including "asphalt," "grass," "tree," "roof," "metal," and "dirt".

In this experiment, the same network setting as given in Table I is used except that the unit number 224 is changed to 162 and the unit number of the last hidden layer of the encoder is changed from 4 to 6 so that these numbers are corresponding to the spectral dimension and endmember number. We set the learning rate to 10^{-4} and also use Adam optimizer to train the network. The batch size is set to 512, and the number of epoch is set to 50. Fig. 3 shows the estimated abundance maps of four prominent minerals. We observe that the abundance maps provide a clearer indication of different materials. Our method also leads to the lowest reconstruction error among these algorithms, with reconstruction error information provided in Fig. 4.

V. Conclusion

In this letter, we proposed a novel scheme for nonlinear hyperspectral unmixing based on a specifically designed deep autoencoder network. The encoder compressed the hyperspectral data into a low-dimensional vector associated with the abundance vector. The decoder uncompressed the encoded vector to reconstruct the input. Endmembers could be extracted from the weights of the linear component of the decoder. The proposed method was tested with both synthetic and real data, and experiment results confirmed the superior performance of our scheme.

References

[1] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[3] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–276, Oct. 1999.

[4] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[5] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[6] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.

[7] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett*, vol. 4, no. 3, pp. 446–450, Jul. 2007.

[8] D. C. Heize and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[9] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.

[10] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[11] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, Jan. 2013.

[12] Q. Wang, X. He, and X. Li, "Locality and structure regularized low rank representation for hyperspectral image classificationn," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, pp. 911–923, Feb. 2019.

[13] Q. Wang, S. Liu, J. Chanussot, and X. Li, "Scene classification with recurrent attention of VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1155–1167, Feb. 2019.

[14] Q. Wang, Z. Yuan, and X. G. Li, "Getnet: A general end-to-end twodimensional CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2019.

[15] G. A. Licciardi and F. D. Frate, "Pixel unmixing in hyperspectral data by means of neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4163–4172, Nov. 2011.

[16] X. Zhang, Y. Sun, J. Zhang, P. Wu, and L. Jiao, "Hyperspectral unmixing via deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1755–1759, Nov. 2018.

[17] Y. Su, A. Marinoni, J. Li, A. Plaza, and P. Gamba, "Nonnegative sparse autoencoder for robust endmember extraction from remotely sensed hyperspectral images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*. Jul. 2017, pp. 205–208.

[18] Y. Qu, R. Guo, and H. Qi, "Spectral unmixing through part-based nonnegative constraint denoising autoencoder," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 209–212.

[19] Y. Su, A. Marinoni, J. Li, J. Plaza, and P. Gamba, "Stacked nonnegative sparse autoencoders for robust hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett*, vol. 15, no. 9, pp. 1427–1431, Sep. 2018.

[20] Y. Qu and H. Qi, "uDAS: An untied denoising autoencoder with sparsity for spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1698–1712, Mar. 2019.

[21] R. Guo, W. Wang, and H. Qi, "Hyperspectral image unmixing using autoencoder cascade," in *Proc. 7th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jun. 2015, pp. 1–4.

[22] B. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Hyperspectral unmixing using a neural network autoencoder," *IEEE Access*, vol. 6, pp. 25646–25656, 2018.

[23] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2807–2820, Oct. 1999.

[24] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, Jun. 2014.

[25] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 240–251, Jan. 2016.

[26] B. Hapke, "Bidirectional reflectance spectroscopy: 1. Theory," *J. Geophys. Res.*, vol. 86, no. B4, pp. 3039–3054, Apr. 1981.

[27] R. Ammanouil, A. Ferrari, C. Richard, and S. Mathieu, "Nonlinear unmixing of hyperspectral data with vector-valued kernel functions," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 340–354, Jan. 2017.