2020 28th Iranian Conference on Electrical Engineering (ICEE)

Patch-based Sparse and Convolutional Autoencoders for Anomaly Detection in Hyperspectral Images

Amir Reza Rezvanian
Image processing and Information Analysis Lab, Faculty of Electrical and Computer Engineering
Tarbiat Modares University
Tehran, Iran
a.rezvanian@modares.ac.ir

Maryam Imani
Image processing and Information Analysis Lab, Faculty of Electrical and Computer Engineering
Tarbiat Modares University
Tehran, Iran
maryam.imani@modares.ac.ir

Hassan Ghassemian
Image processing and Information Analysis Lab, Faculty of Electrical and Computer Engineering
Tarbiat Modares University
Tehran, Iran
ghassemi@modares.ac.ir

***Abstract*— Anomaly target detection is one of the major aims of Hyperspectral Image (HSI) processing. Since anomalous pixels Compose a small fraction of the hyperspectral data cube, the use of supervised neural networks presents many complications. The reason is that supervised networks need a large training set to fine-tune the model. In this paper we propose two semi-supervised autoencoder based anomaly detection methods using the reconstruction error of each pixel. The reconstruction error is the mean absolute error between each pixel and its reconstruction by the proposed autoencoder networks. The proposed networks are deep fully-connected sparse autoencoders (SAE) and deep one-dimensional convolutional autoencoders (CAE). In addition, a patch-based anomaly detection method is proposed which takes spatial correlation between neighbouring pixels into account. We use the San Diego airport hyperspectral data to carry out the experiments. The results are compared with some state-of-the-art HSI anomaly detection methods. Quantitative results employing ROC and AUC metrics demonstrate the superior performance of the proposed method compared to several anomaly detectors.**

***Keywords*— anomaly detection, convolutional autoencoder, hyperspectral images, sparse autoencoder**

I. INTRODUCTION

Hyperspectral images obtained by remote sensors capture the reflective properties of objects over hundreds of narrow contiguous spectral bands. Different materials reflect electromagnetic radiation in different ways. Thus, they have different spectral signatures. These unique spectral signatures can then be used to differentiate between pixels of the hyperspectral data cube in order to identify the various materials making up the scene. Target detection algorithms employ spectral signatures to differentiate the target from background pixels. The spectral characteristics of the few target pixels vary greatly from the rest of the pixels. So, target detection may be accomplished by unsupervised anomaly detectors which make no prior assumptions about the target [1-4].

One of the benchmark anomaly detection methods is the Reed-Xiaoli (RX) algorithm [5]. The RX algorithm assumes a multi-variate gaussian distribution for the background clutter, the parameters of which may be estimated either globally or in each different region of the image [6]. The RX algorithm is derived via the Generalized Likelihood Ratio (GLR) test which treats some stochastic parameters as fixed but with unknown values. In [7], the subspace RX (SSRX) algorithm is developed by considering probability distributions for the parameters that are treated as deterministic in the GLR. It is proven that introducing this prior information gained by measurement programs decreases the false alarm rate. In [8] kernel PCA is utilized to move the data onto a higher dimension with the goal of making the difference between background and anomaly pixels more apparent. The component which has the maximum singularity according to kurtosis and skewness is used in anomaly detection. In [9], the hyperspectral data is assumed to be made of three parts: background, noise and anomalies. With this assumption, the task of detecting anomalies becomes a matrix decomposition problem. Most anomaly detection tasks are accomplished by comparing the significant difference between anomalies and features extracted from the background. The low-rank and sparse matrix decomposition (LRaSMD) method avoids the problems of anomaly contamination and calculating the inverse of the covariance matrix by extracting knowledge from both the background and anomalies. In [10], the collaborative representation detector (CRD) is developed and assumes that a linear combination of the pixels surrounding each background sample may be used as a representation of that pixel. In this nonparametric detector, there is no need to assume a specific distribution for the background. Factors such as atmospheric conditions and variations in spectral signatures of materials affect the reflectance properties in hyperspectral data. Tao *et al.* in [11] consider the use of fractional Fourier transform (FrFT) on the original data as a way of dealing with nonstationary noise and combining FrFT and Shannon entropy results in the FrFE-RX method. In this paper two deep autoencoder networks, sparse and convolutional autoencoders are proposed. Anomaly detection is accomplished according to the mean absolute error between each pixel and its respective reconstruction. In addition, the spatial correlation between background pixels is used to develop a patch-based anomaly detection method. The rest of this paper is structured so that section II provides a summary of autoencoders. In section III, the proposed method is explained and the results of the experiments are demonstrated in section IV. We draw relevant conclusions in section V.

978-1-7281-7296-5/20/$31.00 ©2020 IEEE

II. A Brief Overview of Autoencoders

An autoencoder is a neural network composed of two parts. The first part called encoder, tries to compress the input $X \in \mathbb{R}^D$, where D represents the spectral resolution of the image, into a latent space $h \in \mathbb{R}^H$ with smaller dimensions than the input space. The second part called decoder, tries to reconstruct the input using the latent representation in a way that minimizes the cost function in (1).

$$\min_{W,\beta} \frac{1}{D}\sum_{i=1}^{D} \mathcal{L}(W,\beta;x_i,\hat{x}_i) + \frac{\lambda}{2}\|W\|_2^2 . \quad (1)$$

$\hat{x}$ is the reconstruction of the input pixel x by the autoencoder. W is the matrix of weights and β is the bias vector for the autoencoder network. The first term is the reconstruction error of $\hat{x}$ computed as the mean absolute error between x and $\hat{x}$. The second term is a regularization term that prevents the network from overfitting by incentivizing it to learn smaller weights. By using the encoder output and connecting it to another autoencoder, deep autoencoders may be formed. Each added layer in a deep architecture leads the network to learn more effective and abstract representations of the original data. In this paper two deep autoencoder models are proposed for anomaly detection: 1. Sparse autoencoder and 2. Convolutional autoencoder.

I. Proposed Autoencoders

One category of anomaly detectors is deviation-based methods and the use of autoencoders for detecting anomalous data falls into this category. In such methods the data is first compressed using a dimensionality reduction technique. Then, it is reconstructed by using the most influential components. Thus, the difference that is present in the reconstructed data point can be used to judge that instance of data as either anomalous or normal.

In order to detect anomalies in a hyperspectral scene, the network is trained only on background samples. Once the training is done and the network has been fine-tuned, the background samples along with anomalous data are used to test the model. This is a semi-supervised method for detecting anomalies. The network is trained via supervised learning and tested in an unsupervised manner. Since the network is trained only on background samples and learns to represent only normal pixels, the introduction of anomalous samples in the testing stage causes much higher reconstruction errors for anomaly pixels. Thus, the reconstruction error is used to judge if a pixel is anomalous [12]. If the error is higher than the chosen threshold, the pixel is labeled as an anomaly.

Two autoencoder networks have been proposed for anomaly detection. In addition, we propose a patch-based autoencoder method for anomaly detection integrated with fully-connected and convolutional layers.

A. Sparse Autoencoder

Sparse autoencoders are formed by adding a sparsity penalty, normally in the form of L1 regularization or Kullback-Leibler (KL) divergence, to the activations of the neurons in the hidden layers. This causes the latent representation to be a sparse vector which has most of its elements constrained close to zero. The cost function of such autoencoders may be formulated as (2).

Table 1. Design parameters of the sparse autoencoder.

Layer	Number of Neurons	Activation
Input	189	–
1	90	ReLU
2	40	ReLU
3	25	ReLU
4	40	ReLU
5	90	ReLU
6	189	ReLU
Total Number of Trainable Parameters		43,694

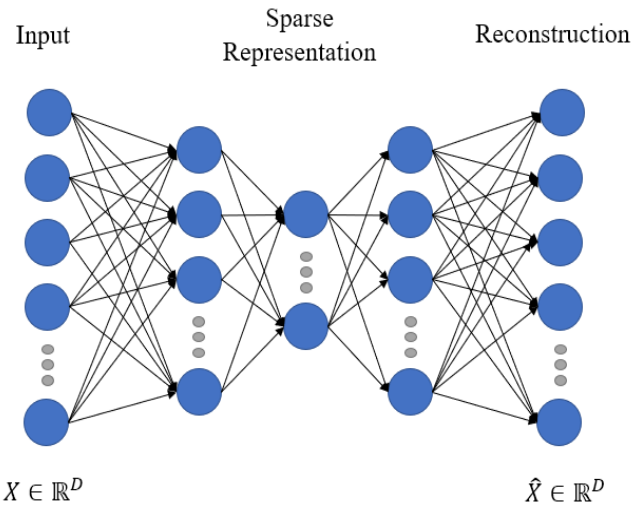

Fig. 1. Architecture of the sparse autoencoder.

$$\min_{W,\beta} \mathcal{L}(W,\beta;x,\hat{x}) + \frac{\lambda_1}{2}\|W\|_2^2 + \lambda_2 \sum_{i=1}^{K} KL(\rho||\hat{\rho}_m) . \quad (2)$$

The third term adds the sparsity penalty. ρ is a small sparsity parameter and $\hat{\rho}_m = 1/N \sum_{i=1}^{N} h_m^{(i)}$, $m=1,\ldots,K$ is the average sparseness of the mth neuron in the ith hidden layer and the KL divergence forces $\hat{\rho}_m$ to be close to ρ. The parameters λ_1 and λ_2 may be changed to put more emphasis on regularization or sparsity as needed [15]. Fig. 1 shows the architecture of the fully-connected sparse autoencoder (SAE) network. The input to the network is the spectrum of a pixel. Table. 1 shows the design parameters for the autoencoder networks used in this paper.

B. Convolutional Autoencoder

In addition to the fully-connected sparse autoencoder, a convolutional autoencoder is used in this paper, since convolutional neural networks have proven very effective in feature extraction [13-15]. Each stage of the convolutional autoencoder is composed of a convolution layer followed by a pooling layer. In this paper one-dimensional convolution is performed over the spectrum of each pixel in the hyperspectral data and the output of the convolutional operator at position x of the jth feature map in the ith layer is shown in (3).

$$v_{ij}^{x} = tanh\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1} w_{ijm}^{p}\, v_{(i-1)m}^{x+p}\right). \quad (3)$$

Table 2. Design parameters of the convolutional autoencoder.

Layer	Operator Size	Number of Feature Maps	Layer Output Shape
Convolution	16	15	(189, 15)
Max-Pooling	3	15	(63, 15)
Convolution	16	10	(63, 10)
Max-Pooling	3	10	(21, 10)
Convolution	16	5	(21, 5)
Convolution	16	5	(21, 5)
Up Sampling	3	5	(63, 5)
Convolution	16	10	(63, 10)
Convolution	16	15	(63, 15)
Up Sampling	3	15	(189, 15)
Convolution	16	1	(189,1)

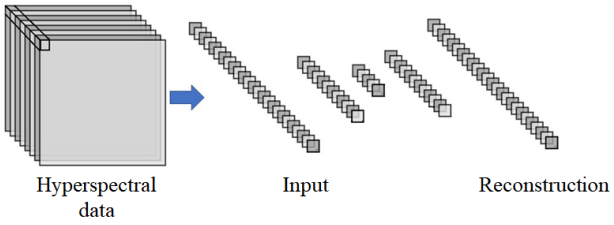

Fig. 2. Architecture of the convolutional autoencoder.

m indexes the feature maps in the previous layer, w_{ijm}^{p} is the weight of position p connected to the mth feature map, P_i is the length of the convolutional kernel and b_{ij} is the bias of the jth feature map in the ith layer. Fig. 2 shows the architecture of the Convolutional autoencoder (CAE) network. Table. 2 shows the design parameters for the convolutional autoencoder networks used in this paper.

C. Patch-based Autoencoder

Since the background is composed of many different materials with different spectral signatures, the use of a single autoencoder network to find a compressed representation for the whole scene leads to high reconstruction errors and a less discriminative anomaly score. Additionally, the small training set leads to the risk of overfitting the network. The pixels in close proximity are more likely to have similar spectral signatures. So, trying to find a different latent representation for different parts of the scene is a more effective way for calculating the anomaly score. In order to consider the spatial correlation of spectral signatures between neighbouring pixels, it is proposed to break down the original hyperspectral cube into smaller patches. Consider a hyperspectral cube with samples $X = \{x_{ij}\}_{i,j=1}^{n_i,n_j}$ in $\mathbb{R}^D$, where D represents the spectral resolution of the image and i and j index the spatial domain. This cube is broken down into smaller cubes, with samples $X = \{x_{pq}\}_{p,q=1}^{n_p,n_q}$ in $\mathbb{R}^D$. n_p and n_q are smaller than n_i and n_j, respectively. Fig. 3 demonstrates the similarity between the spectral signatures of pixels in the same patch. The background samples in each patch are fed into the same autoencoder network to find an appropriate representation for each part of the data. Fig. 4 shows the process of patch based sparse autoencoder (PSAE) anomaly detection and patch based convolutional autoencoder (PCAE) anomaly detection.

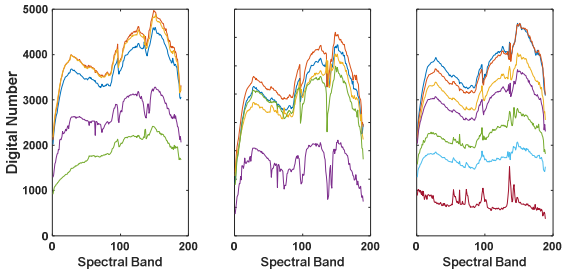

Fig. 3. Spectral signatures of some pixels in three different patches.

Once training is done, anomaly and background pixels from each patch are used to test the network trained for that patch. Pixels that cause a larger reconstruction error are labelled as anomalies.

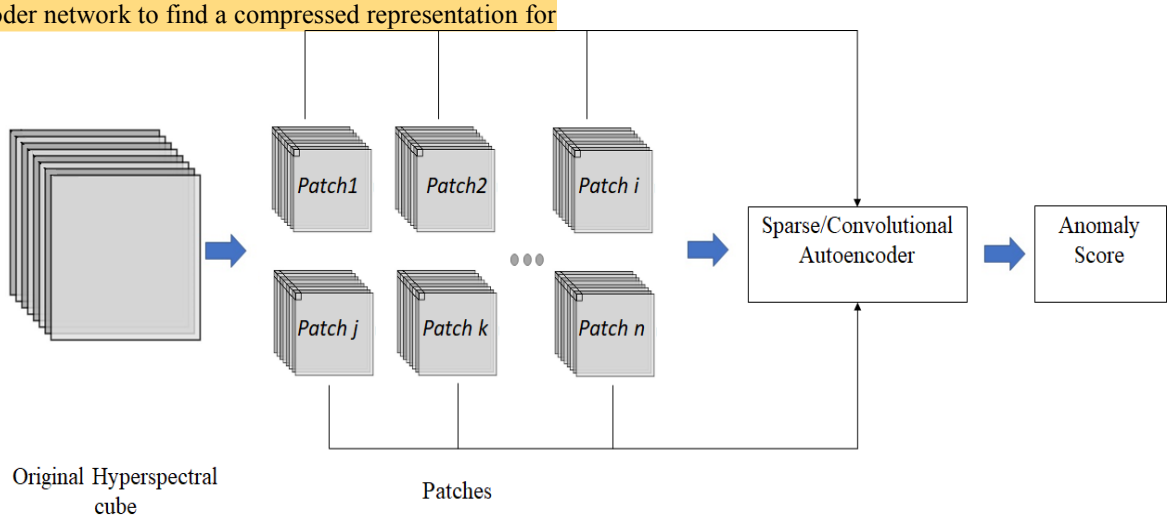

Fig. 4. The proposed patch-based anomaly detection method.

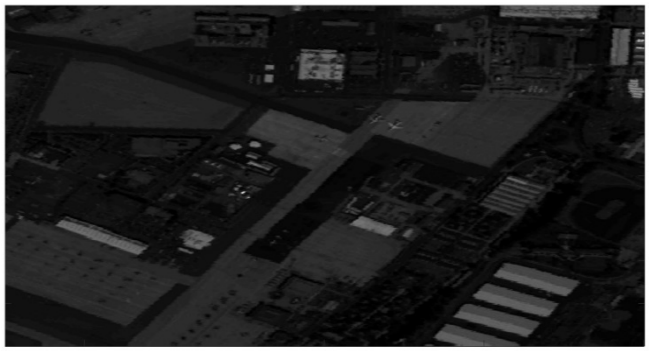

Fig. 5. Band 19 of San Diego hyperspectral data.

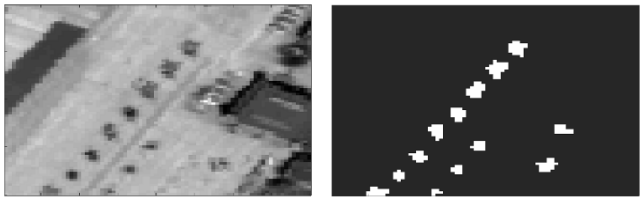

Fig. 6. San Diego hyperspectral data. (a) The 80×80 patch used in the experiments. (b) Reference map.

By breaking down the original image into smaller patches, we preserve the local structure of the image. The use of spatial information along with spectral signatures, improves discrimination between background and anomaly pixels.

II. Experimental Results

We use the San Diego airport hyperspectral data collected by the AVIRIS sensor to carry out the experiments. The original image has a 400×400 spatial domain along with 224 spectral bands as shown in fig. 5. After removing parts of the spectra that have low SNR and have been affected by water absorption, an 80×80×189 patch of the original hyperspectral cube demonstrated in fig. 6 is used for the experiments. The 13 aircrafts in the scene are regarded as the anomaly target pixels. The performances of the proposed methods SAE and CAE are compared with global RX, Local RX, digital wavelet transform (DWT-RX) [16], derivative RX (Deriv-RX) [17], discrete Fourier transform (DFT-RX) [18] and FrFE-RX. The 2-D receiver operating characteristic (ROC) curve showing correctly identified anomalies against false alarm cases at various increasingly larger threshold settings and the area under the 2-D ROC curve are used to quantitatively measure the performance of the proposed methods.

Fig. 7 show the ROC curve of various anomaly detection methods in comparison with SAE and CAE. A detector with a ROC curve near the upper left corner of the figure correctly identifies more anomalous pixels and has a smaller false positive rate and thus, a better performance. It is shown that the proposed CAE based anomaly detection almost always has the best performance, followed by Deriv-RX and the proposed SAE based anomaly detection. In the next experiment, we examine the effects of breaking down the hyperspectral data to smaller patches. In this way, the spatial correlation between pixels are considered. In order to do this, the original 80×80×189 hyperspectral image is broken down to different numbers of patches. Fig. 8 shows how the ROC

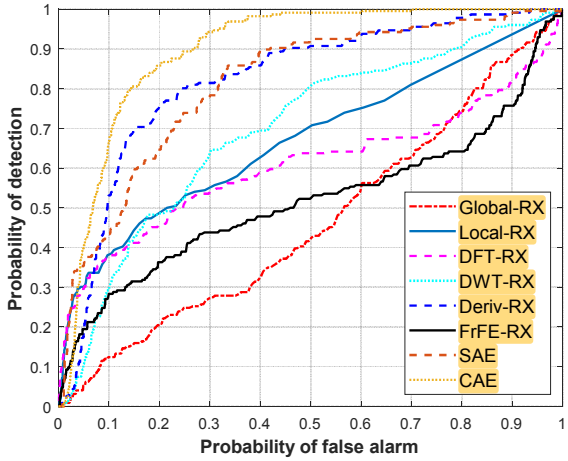

Fig. 7. ROC curves of various anomaly detectors using the San Diego hyperspectral data.

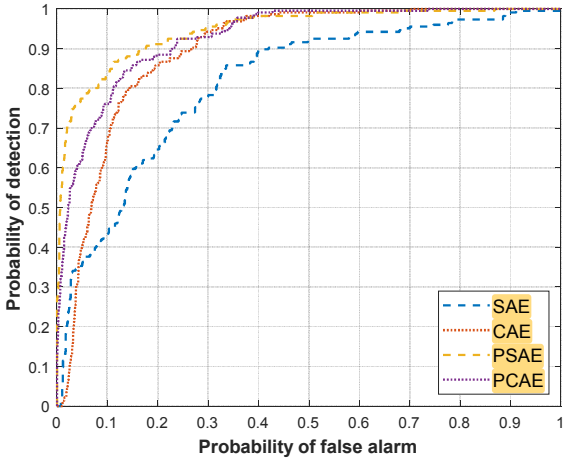

Fig. 8. Changes in ROC curves of SAE anomaly detection and CAE anomaly detection after breaking the hyperspectral data into smaller parts.

curves of the proposed SAE anomaly detection and CAE anomaly detection methods were affected by this change. It is demonstrated that the patch-based SAE (PSAE) and patch-based CAE (PCAE) methods have a better performance than their counterparts which use the entire hyperspectral image. Table. 3 shows the AUC score of the various Anomaly detection methods and the time each method requires. It is evident that the detectors designed by breaking down the hyperspectral image into smaller patches have a higher AUC and the best chance of detecting anomalies with fewer false positive cases. As seen in table. 3, autoencoder based anomaly detection methods are slower but achieve much better performances. The time required to train the network increases with the number of patches. Fig. 9 shows how the number of patches affect AUC. The improvement caused by using patches is insignificant after dividing the image into 16 patches.

III. Conclusions

In this paper, we proposed two semi-supervised autoencoder models for hyperspectral anomaly detection based on reconstruction error. The networks designed for this

Table 3. AUC scores of various anomaly detectors and the time each algorithm needs.

Method	Global-RX	Local-RX	DFT-RX	DWT-RX	Deriv_RX	FrFE-RX	SAE	CAE	PSAE	PCAE
AUC	0.4632	0.6709	0.6061	0.6973	0.8204	0.5175	0.8142	0.8962	0.9470	0.9293
Time (sec)	4.6	62	0.49	11.8	0.3	31.3	508	249	1650	831

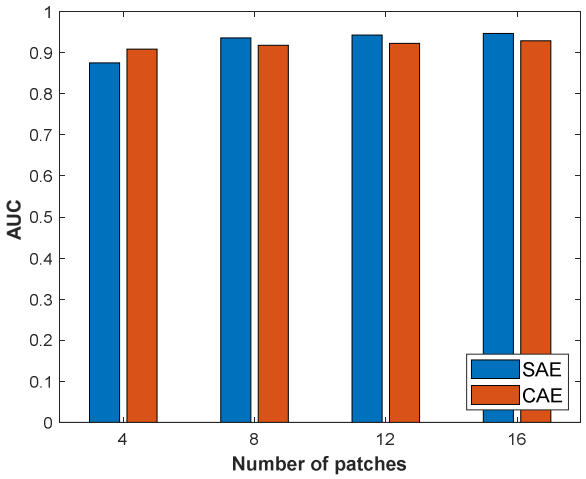

Fig. 9. The effect of number of patches on AUC

purpose were deep sparse and deep convolutional autoencoders. Furthermore, we take advantage of the spatial correlation between neighbouring pixels, by breaking down the data into smaller patches and thus limiting the variations in the spectral signatures each autoencoder network needs to compress. Experiments on a hyperspectral data set proved the superior performance of the proposed methods in comparison with global and local RX, DFT-RX, DWT-RX, Deriv-RX, and FrFE-RX was proven via the quantitative measures of ROC curve and the area under the ROC curve.

REFERENCES

[1] Nasser M. Nasrabadi, "Hyperspectral Target Detection: An Overview of Current and Future Challenges", *IEEE Sig. Proc. Mag.*, vol. 31, no. 1, pp. 34-44, Jan 2014.

[2] M. Imani, "RX Anomaly Detector with Rectified Background", *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 8, pp. 1313-1317, August 2017.

[3] D. Manolakis, G. Shaw, "Detection algorithms for hyperspectral imaging applications", *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29-43, Jan 2002.

[4] M. Imani, "Hyperspectral anomaly detection using differential image", *IET Image Processing*, vol. 12, no. 5, pp. 801-809, January 2018.

[5] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.

[6] J. M. Molero, E. M. Garzón, I. García, A. Plaza, "Analysis and Optimizations of Global and Local Versions of the RX Algorithm for Anomaly Detection in Hyperspectral Data", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 801-814, April 2013.

[7] A. P. Schaum, "Joint subspace detection of hyperspectral targets", *IEEE Aerospace Conference Proceedings*, vol. 3, pp. 1818–1824 Mar. 2004.

[8] Y. Gu, Y. Liu, Y. Zhang, "A Selective Kernel PCA Algorithm for Anomaly Detection in Hyperspectral Imagery", *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, May 2006.

[9] Y. Zhang, B. Du, L. Zhang, S. Wang, "A Low-Rank and Sparse Matrix Decomposition-Based Mahalanobis Distance Method for Hyperspectral Anomaly Detection", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1376-1389, March 2016.

[10] W. Li, Q. Du, "Collaborative Representation for Hyperspectral Anomaly Detection", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1463-1474, March 2015.

[11] R. Tao, X. Zhao, W. Li, H. C. Li, Q. Du, "Hyperspectral Anomaly Detection by Fractional Fourier Entropy", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Early Access)*, pp. 1-10, September 2019.

[12] J. An, S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability", *Special Lecture on IE*, vol. 2, no. 1, December 2015.

[13] Y. Shi, J. Lei, Y. Yin, K. Cao, Y. Li, C. Chang, "Discriminative Feature Learning with Distance Constrained Stacked Sparse Autoencoder for Hyperspectral Target Detection", *IEEE Geoscience and Remote Sensing Letters*, vol. 19, no. 9, pp. 1462-1466, September 2019.

[14] S. Mei, J. Ji, Y. Geng, Z. Zhang, X. Li, 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, September 2019.

[15] Y. Chen, H. Jiang, C. Li, X. Jia, P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232-6251, October 2016.

[16] Y. Tang, Y. Lu, and H. Yuan, "Hyperspectral image classification based on three-dimensional scattering wavelet transform,", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2467–2480, May 2015.

[17] D. Liu L. Han, "Spectral curve shape matching using derivatives in hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 4, pp. 504–508, Apr. 2017.

[18] C. Candan, M. A. Kutay, and H. Ozaktas, "The discrete fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1329–1337, May 2000.