# Transferred Deep Learning for Anomaly Detection in Hyperspectral Imagery

Wei Li, *Senior Member, IEEE*, Guodong Wu, and Qian Du, *Senior Member, IEEE*

***Abstract*— In this letter, a novel anomaly detection framework with transferred deep convolutional neural network (CNN) is proposed. The framework is designed by considering the following facts: 1) a reference data with labeled samples are utilized, because no prior information is available about the image scene for anomaly detection and 2) pixel pairs are generated to enlarge the sample size, since the advantage of CNN can be realized only if the number of training samples is sufficient. A multilayer CNN is trained by using difference between pixel pairs generated from the reference image scene. Then, for each pixel in the image for anomaly detection, difference between pixel pairs, constructed by combining the center pixel and its surrounding pixels, is classified by the trained CNN with the result of similarity measurement. The detection output is simply generated by averaging these similarity scores. Experimental performance demonstrates that the proposed algorithm outperforms the classic Reed-Xiaoli and the state-of-the-art representation-based detectors, such as sparse representation-based detector (SRD) and collaborative representation-based detector.**

***Index Terms*— Anomaly detection, convolutional neural network (CNN), deep learning, hyperspectral imagery.**

## I. Introduction

HYPERSPECTRAL imagery has been increasingly used for various applications owing to its high spectral resolution, which enables discrimination of different materials [1]. In an image, those pixels (e.g., man-made objects) with significantly different spectral signatures from their neighboring background pixels or the global background pixels are considered as anomalies [2]. In practice, since no prior knowledge of the anomalous spectrum or background information is available, anomaly detection is typically based on statistical characterization for data or background modeling.

The Reed-Xiaoli (RX) algorithm [3], developed under a hypothesis testing where the conditional probability density functions (PDFs) under the two hypotheses (without and with anomaly) are assumed to be Gaussian, has become the benchmark of anomaly detection algorithms [4]. The essence of the RX detector turns out to be the Mahalanobis distance between the pixel under test and the background [5]. Two typical versions of the RX have been categorized: Global RX, which estimates the background statistics of the entire image, and Local RX (LRX), which estimates the background using local statistics. In addition to the classical RX detector, a number of its extensions and other anomaly detection algorithms have also been investigated for hyperspectral data. A time-efficient method has been introduced for anomaly detection in [6]; subspace-projection-based detectors were presented in [7]; random-selection-based detector was described in [8]; the subpixel anomaly detection was discussed in [9]; decision fusion for dual-window-based detector was proposed in [10]; and multiple-window anomaly detection was developed in [11] to capture local spectral variations.

Collaborative representation-based detector (CRD) [12] has been recently proposed. The algorithm was designed to detect anomalies with unknown signatures, where a pixel is claimed to be an anomaly if it cannot be collaboratively represented by background atoms. In CRD, surrounding spatial neighborhoods in the outer region are linearly combined to produce a prediction for the central pixel within a local window, and anomalies can be detected in the residual image, which is obtained by subtracting the predicted background from the original image. A counterpart of CRD is a sparse representation-based detector (SRD) [13], where sparse representation coefficients are solved via an $\ell_1$-norm minimization. The supervised hyperspectral target detection algorithms using sparse representation have been discussed in [14]–[16].

Recently, deep learning-based methods have drawn increasing attention in hyperspectral image analysis [17]–[19]. For example, a deep learning architecture with multilayer stacked autoencoder was proposed to extract high-level features in an unsupervised manner [17]. Deep belief network was employed to extract spectral-spatial features along with logistic regression classification [20]. By adding a regularization term, spatial updated deep autoencoder with the energy function was proposed to extract spectral-spatial information [21]. Furthermore, convolutional neural network (CNN) was employed to exploit deep representation based on spectral signatures in [22], unsupervised sparse features were learned via deep CNN in a greedy layerwise fashion for pixel classification in [23], and CNN was utilized to automatically find spatial-related features at high levels from a subspace after local discriminant embedding [24].

In this letter, an anomaly detection framework based on deep learning is proposed. As far as we know, it is the first time that deep CNN is applied for hyperspectral image anomaly detection. In the proposed CNN-based detection

Manuscript received August 29, 2016; revised November 8, 2016 and January 1, 2017; accepted January 22, 2017. Date of publication March 10, 2017; date of current version April 20, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant NSFC-91638201 and Grant 61571033 and in part by the Higher Education and High-Quality and World-Class Universities under Grant PY201619.

W. Li and G. Wu are with the College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China (e-mail: liwei089@ieee.org).

Q. Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: du@ece.msstate.edu).

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.2017.2657818

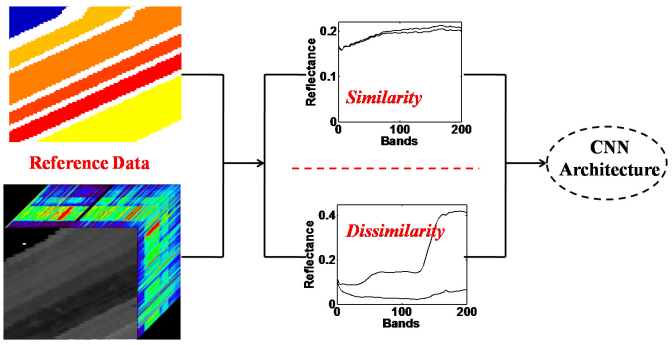



Fig. 1. Illustration of selecting similar or dissimilar pixel pairs from a reference data.

(denoted as CNND), a reference image scene (from the same sensor) with labeled samples is needed, from which two types of pixel pairs are selected—a pair of samples from the same class (on behalf of similarity) and a pair of samples from different classes (on behalf of dissimilarity). Each difference between paired samples is then fed into a deep CNN, whose architecture is well designed. For a testing pixel, neighboring differences between pixel pairs constructed using its surroundings are classified by the trained CNN with the results of similarity measurement. The detection output is simply generated by averaging these similarity results. It is expected that CNN architecture extracts discriminative features from these two types of pixel pairs, which can quantitatively measure similarity or dissimilarity between samples in one pair. Thus, during the testing process, the proposed CNND is able to recognize the center pixel as an anomaly if the one is dissimilar to its surroundings, and vice versa.

## II. Proposed Anomaly Detection Framework

The proposed detection framework mainly includes three steps: learning a deep CNN architecture using pixel pairs selected from a reference data, measuring the similarity between the testing pixel and its surroundings by the trained CNN, and averaging these similarity values as the final detection output.

### A. CNN-Based Anomaly Detection

First of all, in order to learn a multilayer CNN, a reference data with ground truth are employed and difference between pixel pairs is generated. There are two types of pixel pairs, as shown in Fig. 1; that is, a "similarity" class (the label is denoted as 0) includes pairs of samples from the same class, and a "dissimilarity" class (the label is denoted as 1) includes pairs of samples from different classes. A sufficient number of differences between two-class pixel pairs are then fed into the designed CNN architecture.

After obtaining the input samples, a deep CNN architecture is used to extract features of differences between pixel pairs, as shown in Fig. 2. CNNs represent feed-forward neural networks, which consist of various combinations of convolutional layers, average-pooling layer, and exploit local correlation patterns by enforcing a local connectivity pattern between neurons of adjacent layers. The designed CNN architecture contains 16 learnable convolutional layers and rectified linear unit layers after each convolutional layer. The convolutional layers with stride two are used to reduce the spectral dimensionality. Once the spectral dimension is achieved to a desired value (e.g., 7), a average-pooling tensor is fed into fully connected (FC) layer. To be used for classification, the chain of the CNN architecture ends in an FC network with sigmoid and label information (layer).

Fig. 3 shows the flowchart of the CNN-based anomaly detection in hyperspectral imagery with dual windows. The center pixel is represented as $T$; in this example, the inner window is $3 \times 3$ and the outer window is $5 \times 5$. Thus, there are 16 samples (i.e., $\{T, 1\}, \{T, 2\}, \ldots, \{T, 16\}$) fed into CNN. The output of the network is an $16 \times 1$ vector with similarity scores (i.e., $r_1, r_2, \ldots, r_{16}$), and the final detection output is simply generated by averaging these values, $D(T) = (1/16) \sum_{i=1}^{16} r_i$. Note that, since the label of "similarity" class is denoted as 0 and the one of the "dissimilarity" class is denoted as 1, the range of $r_i$ is between 0 and 1. The detection output $D(T)$ is compared with the prescribed threshold $\eta$—if $D(T) > \eta$, the pixel $T$ is an anomaly; otherwise normal pixel.

### B. Analysis on the Proposed CNND

In the CNND framework, the reference map is important, since it guarantees a sufficient number of differences of pixel pairs to learn the parameter space of the CNN. As shown in Fig. 2, the total number of trainable parameters in the designed architecture can be approximately 40 000. If the employed reference data have $C$ classes with $m$ labeled samples per class, the number of pixel pairs, representing "similarity" class, can be calculated using all combinations, $N = C * P(m, 2)/2 = C * (m!/((m-2)!2!))$. Suppose $C = 16$ and $m = 200$, the value $N = 318\,400$, which is much larger than 40 000. Moreover, for "dissimilarity" class, even more pixel pairs can be selected. It is worth mentioning that the hyperspectral sensor collecting the reference data is the same one capturing the data for anomaly detection; in this way, the transferred CNN can make similarity assessment accurately.

Similar to the traditional LRX or CRD [12], dual-window strategy is also employed in the proposed CNND framework. We usually assume that the anomalous pixels are rarely present and all (or most) of the samples inside the outer window and outside the inner window represent normal pixels. One of the situations may appear as follows.

1) If the testing pixel $T$ is normal and none of the surroundings is anomalous either, the value $D(T)$ will be very close to 0 and CNND views $T$ as normal.
2) If the testing pixel $T$ is normal and one or two samples of surroundings are anomalous, the value $D(T)$ after averaging is still relatively nearly to 0 and CNND views $T$ as normal.
3) If the testing pixel $T$ is an anomalous pixel and none of surroundings is anomalous, the value $D(T)$ will be very close to 1 and CNND views $T$ as anomaly.
4) If the testing pixel $T$ is an anomalous pixel and one or two samples of surroundings are anomalous, the value $D(T)$ after averaging is still nearly to 1 and CNND views $T$ as anomaly.

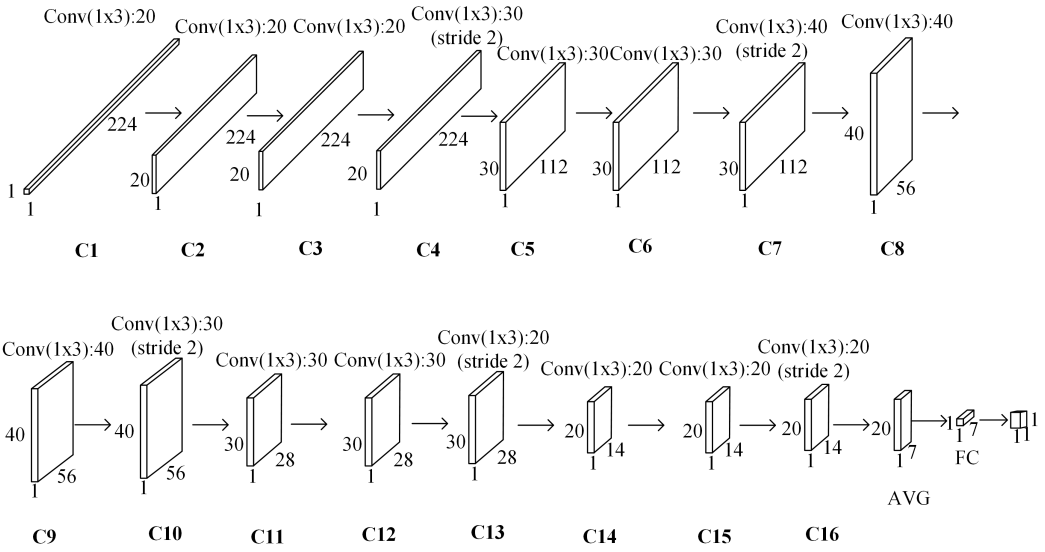



Fig. 2. Flowchart of the proposed deep CNN architecture with convolutional layers and average-pooling layers. Input represents the difference between pixel pairs obtained from available training samples (suppose $d = 224$).

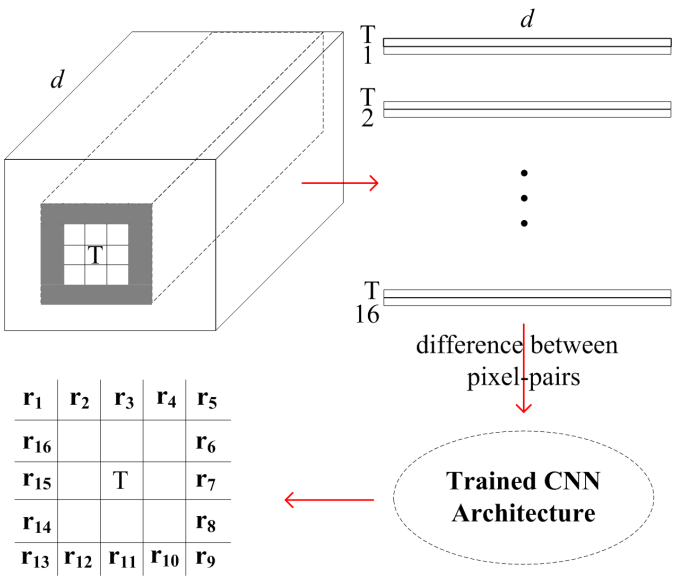


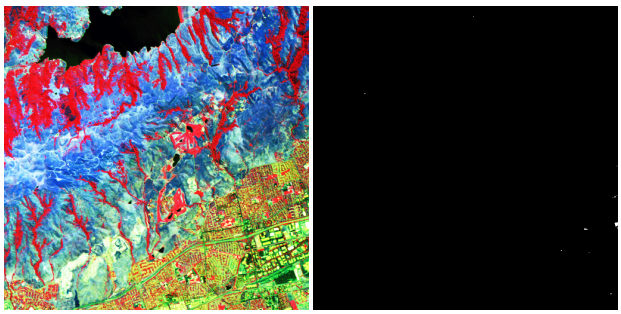

Fig. 3. Flowchart illustration of the CNN-based anomaly detection in hyperspectral imagery.

## III. Experimental Results

We investigate the performance of the proposed CNND, and compare it with the classic LRX and the state-of-the-art representation-based detectors, i.e., SRD and CRD. For detection tasks, receiver-operating-characteristic (ROC) [25] curve is usually employed for quantitative evaluation. For ROC generation, a detection output map is normalized to [0, 1], and the threshold is gradually changed from 0 to 1; after comparing with the known ground-truth map, the resulting probability of false alarm (denoted as $P_f$) and probability of detection (denoted as $P_d$) are plotted.

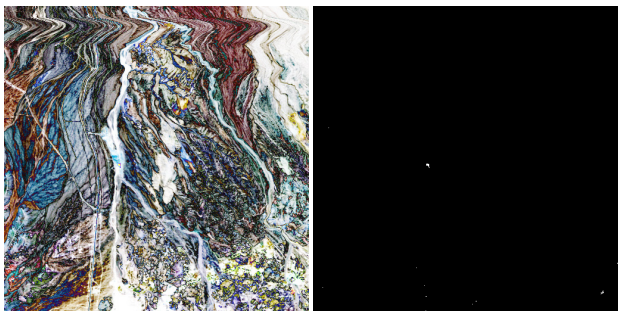
### A. Hyperspectral Data

The two testing experimental data were obtained from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor[1] covering the Moffett Field, California, and Cuprite, Nevada, respectively. These two scenes both consist of 512 × 512 pixels [as shown in Figs. 4(a) and 5(a)] with 224 bands spanning the wavelength interval 0.4 to 2.5 $\mu$m. The spatial resolution is approximately 20 m. Here, we implement Global RX to find the anomalous pixels in the scene [12], where 59 anomalies are identified when the threshold is set



(a) (b)

Fig. 4. (a) Pseudocolor image of Moffett Field scene (using bands 60, 40, and 20). (b) Ground-truth map of 59 anomalous pixels.



(a) (b)

Fig. 5. (a) Pseudocolor image of Cuprite scene (using bands 60, 40, and 20). (b) Ground-truth map of 46 anomalous pixels.

[1] http://aviris.jpl.nasa.gov/

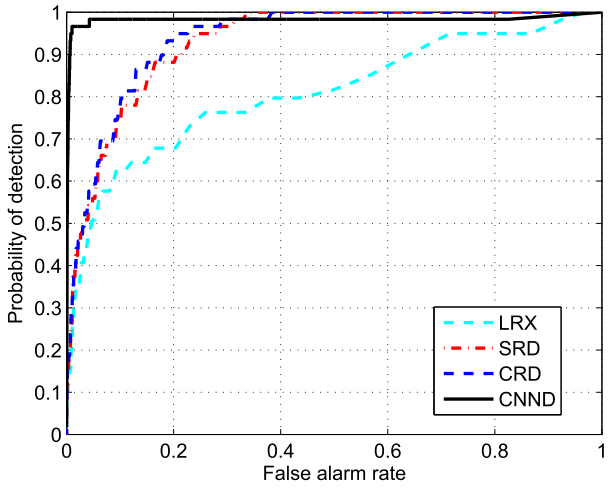



Fig. 6. Detection performance of different anomaly detectors using the Moffett Field data.

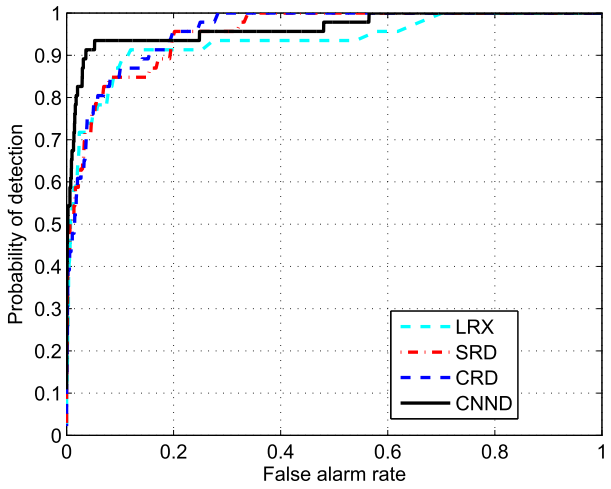



Fig. 7. Detection performance of different anomaly detectors using the Cuprite data.

TABLE I
AUC (%) PERFORMANCE OF DETECTORS FOR VARIOUS SIZES OF WINDOWS ($w_{\text{IN}}, w_{\text{OUT}}$) USING THE MOFFETT FIELD DATA

| | (7,9) | (5,7) | (3,7) | (3,5) |
|---|---|---|---|---|
| CNND | 98.28 | 98.20 | 98.26 | 98.14 |
| CRD | 93.74 | 89.31 | 90.79 | 82.37 |
| SRD | 93.04 | 89.14 | 90.37 | 80.58 |
| LRX | 80.78 | 76.30 | 66.12 | 58.28 |

to 0.5 in Fig. 4(b) and 46 anomalies in Fig. 5(b), which are used as ground truth for local detectors.

The reference data[2] were also collected by the AVIRIS sensor for an area over Salinas Valley, California. The image comprises $512 \times 217$ pixels with a spatial resolution of 3.7 m with 224 bands. It mainly contains vegetables, bare soils, and vineyard fields. There are 16 classes and hundreds of labeled samples per class.

[2]http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

TABLE II
EXECUTION TIME (IN SECONDS) OF DIFFERENT ANOMALY DETECTORS USING TWO EXPERIMENTAL DATA

| Moffett Field | CNND | SRD | CRD | LRX |
|---|---|---|---|---|
| Time | 483 | 89 | 1356 | 1306 |
| Cuprite | CNND | SRD | CRD | LRX |
| Time | 496 | 96 | 1304 | 1321 |

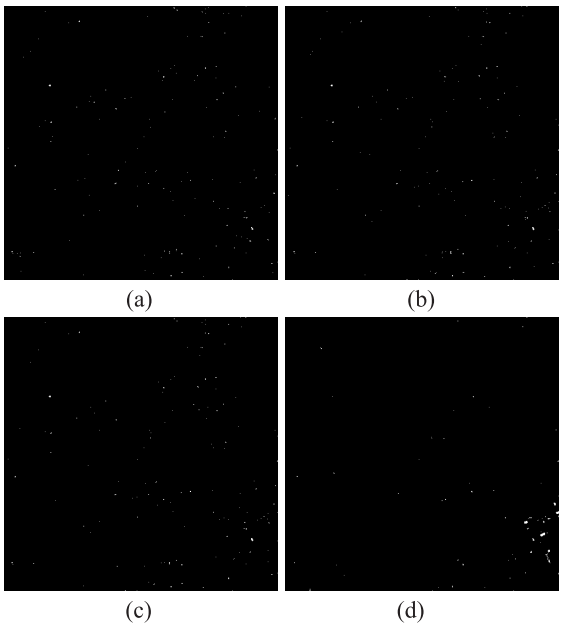



Fig. 8. Detection maps for the Moffett Field data when $P_f$ is fixed to 0.05. (a) LRX: $P_d = 0.5085$. (b) SRD: $P_d = 0.5593$. (c) CRD: $P_d = 0.5932$. (d) CNND: $P_d = 0.9831$.

### B. Detection Performance

Parameters need to be tuned for the proposed CNND. In addition to the weights that can be automatically learned during training, there are several other important parameters in the designed CNN architecture, such as learning rate and window size. Learning rate determines the convergence speed in the procedure of back propagation, which can significantly affect the training performance. According to our empirically study, we initially set the learning rate as 0.001. Furthermore, the dual-window sizes during testing also affect the final detection performance for all the dual-window-based detectors. Table I demonstrates the effect of various sizes of windows ($w_{\text{in}}, w_{\text{out}}$) using the area under ROC curve (denoted as AUC) using the Moffett Field data. It is apparent that the performance of CNND is relatively stable and keeps the best when compared with other detectors. In our following experiments, we chosen window size $(w_{\text{in}}, w_{\text{out}}) = (7, 9)$.

With the best parameters, Figs. 6 and 7 show the ROC performance of different anomaly detectors using two experimental data. From the results, it is obvious that the proposed CNND is always superior to LRX, SRD, and CRD, especially for low $P_f$. Fig. 8 further shows the detection maps when $P_f$ is fixed to a small value (e.g., 0.05) and $P_d$ is the maximum using the Moffett Field data. The proposed CNND still performs the best with the largest $P_d$ (i.e., 0.9831), which is consistent with the results in Fig. 6.

Computational complexity is also investigated. Here, all experiments are carried out using softwares on an Intel(R) Core(TM) i7-3770 CPU machine with 8 GB of RAM. All the programs of the proposed CNND are implemented using Python language and TensorFlow[3] library. TensorFlow is an open source software library for numerical computation using data flow graphs. For CRD, SRD, and LRX, all the programs are carried out using MATLAB. The execution time (*in seconds*) for the experimental data is listed in Table II. Note that the proposed CNND is faster than CRD and LRX while slower than SRD. For SRD, the $\ell_1$-norm minimization is implemented in related toolbox[4] using the MEX function, which calls C program in MATLAB; otherwise, the computational cost should be much higher than CRD using the $\ell_2$-norm minimization.

## IV. Conclusion

In this letter, we proposed a transferred deep CNN-based strategy for anomaly detection in hyperspectral imagery. In the proposed CNND, a reference data set was employed to train the CNN with pixel pairs generated from labeled samples. For each pixel under testing, difference between pixel pairs, constructed by combining with its surrounding pixels, was classified by the trained CNN with the results of similarity measurement. The final detection was achieved through a voting process. Experimental results demonstrated that the proposed CNND outperformed the existing LRX and the state-of-the-art SRD and CRD.

Currently, based on the preliminary study, the reference data are preferred to be from the same sensor as the testing data; for example, both reference and testing data are from the AVIRIS sensor in our experiments. The main reason is that "similarity" or "dissimilarity" pixel-pair features learned by the deep CNN architecture facilitate the similarity measurement effectively between data obtained from the same sensor. How to generalize such transfer learning to different sensors will be investigated in the further research.

## References


[1] L. Zhang, Q. Zhang, L. Zhang, D. Tao, X. Huang, and B. Du, "Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding," *Pattern Recognit.*, vol. 48, no. 10, pp. 3102–3112, Oct. 2015.

[2] B. Du and L. Zhang, "A discriminative metric learning based anomaly detection method," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6844–6857, Nov. 2014.

[3] 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.

[4] J. M. Molero, E. M. Garzón, I. García, and A. Plaza, "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 801–814, Apr. 2013.

[5] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1314–1325, Jun. 2002.

[6] O. Duran and M. Petrou, "A time-efficient method for anomaly detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3894–3904, Dec. 2007.

[7] L. Ma, M. M. Crawford, and J. Tian, "Local manifold learning-based $k$-nearest-neighbor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4099–4109, Nov. 2010.

[8] B. Du and L. Zhang, "Random-selection-based anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1578–1589, May 2011.

[9] S. Khazai, A. Safari, B. Mojaradi, and S. Homayouni, "An approach for subpixel anomaly detection in hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 769–778, Apr. 2013.

[10] W. Li and Q. Du, "Decision fusion for dual-window-based hyperspectral anomaly detector," *J. Appl. Remote Sens.*, vol. 9, no. 1, p. 097297, Feb. 2015.

[11] W. M. Liu and C. I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 644–658, Apr. 2013.

[12] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463–1474, Mar. 2015.

[13] J. Li, H. Zhang, L. Zhang, and L. Ma, "Hyperspectral anomaly detection by the use of background joint sparse representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2523–2533, Jun. 2015.

[14] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 676–680, Jul. 2011.

[15] W. Li, Q. Du, and B. Zhang, "Combined sparse and collaborative representation for hyperspectral target detection," *Pattern Recognit.*, vol. 48, no. 12, pp. 3904–3916, Dec. 2015.

[16] Y. Zhang, B. Du, and L. Zhang, "A sparse representation-based binary hypothesis model for target detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1346–1354, Mar. 2015.

[17] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[18] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.

[19] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[20] Y. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, Jun. 2015.

[21] X. Ma, H. Wang, and J. Geng, "Spectral–spatial classification of hyperspectral image based on deep auto-encoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4073–4085, Sep. 2016.

[22] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sensors*, vol. 2015, Jan. 2015, Art. no. 258619.

[23] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1349–1362, Mar. 2016.

[24] W. Zhao and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, Aug. 2016.

[25] J. A. Hanley and B. J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology*, vol. 148, no. 3, pp. 839–843, Sep. 1983.


[3] http://tensorflow.org/

[4] http://spams-devel.gforge.inria.fr/