Hyperspectral Anomaly Detection Using Robust Principal Component Analysis with Autoencoding Adversarial Network

Atsuya Emoto and Ryo Matsuoka
The University of Kitakyushu,
Faculty of Environmental Engineering,
1-1 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, Japan.

***Abstract*—This paper introduces an unsupervised anomaly detection method for hyperspectral (HS) images, integrating robust principal component analysis (RPCA) and a deep autoencoding adversarial network (AEAN). The proposed method is solved using a plug-and-play alternating direction method of multipliers (PnP-ADMM). This method uses AEAN as a proximity operator to promote low-rankness in reconstructed spectral data, effectively distinguishing anomalous pixels from the background. Moreover, the performance of anomaly detection is enhanced by using the anomaly map estimated by the proposed method as weights for the WLRX detector. Experimental results show that our method outperforms some existing methods in detection accuracy.**

I. Introduction

Hyperspectral (HS) images, with several hundred spectral bands, are widely used in earth surface classification, crop health assessment, environmental monitoring, and mineral exploration. Detecting anomalies, which are small areas with unique spectral characteristics, poses a challenge and often relies on background modeling or statistical methods. Various approaches have been proposed, including statistical methods, subspace methods, linear mixture models, and deep neural networks [1]–[10].

A representative example is the Reed-Xiaoli (RX) anomaly detector [3], which detects anomalous pixels in HS data. The RX detector calculates the anomaly score of each pixel based on the Mahalanobis distance between the pixel and the background, with higher scores indicating a higher probability of being an anomaly. The spectral mean and covariance of the HS data characterize the background. Many advanced methods based on RX have been proposed, e.g., [1], [2].

Some deep learning-based methods have been proposed for HS anomaly detection [4]–[7], often using autoencoder (AE) models. These methods train AEs with pixels from roughly estimated background regions, resulting in low reconstruction errors for background regions and high errors for anomalous regions. Anomalous pixels are detected based on these errors. Adversarial learning between original and reconstructed spectra improves accuracy, termed a deep autoencoding adversarial network (AEAN) [7]. Combining AEAN with the local weighted RX (WLRX) detector [2] further enhances detection accuracy. However, accurate background estimation is challenging, and training with inaccurate estimates can reduce detection accuracy.

In HS anomaly detection, the observed tensor is assumed to comprise a low-rank background and a sparse anomaly component and is decomposed using robust principal component analysis (RPCA) [8]–[10]. RPCA can be applied to the 2D matrix, utilizing a nuclear norm to enhance the spectral low-rankness. The 3D approach implements a tensor nuclear norm (TNN) to improve low-rank properties in both spatial and spectral domains. However, in complex scenes, TNN often leads to checkerboard artifacts that diminish the accuracy of anomaly detection.

In this paper, we propose a novel HS anomaly detection technique that integrates RPCA and AEAN. We employ the AEAN model to extract the spectral characteristics of the background regions from the observed data. Furthermore, to improve the separation accuracy of background and anomaly components and to enable robust anomaly detection, we incorporate the AEAN model as a regularization for the background component in the RPCA optimization problem. The proposed problem can be solved using the plug-and-play alternating direction method of multipliers (PnP-ADMM). Experiments show that the proposed method can more accurately detect anomaly pixels than some conventional methods.

II. Preliminaries

In this paper, bold-faced lowercase and uppercase letters indicate vectors and matrices, respectively. Real-valued N-dimensional vector spaces denotes by $\mathbb{R}^N$. We define the $N \times M$ real-valued matrix as $\mathbb{R}^{N\times M}$. Three-dimensional tensor is denoted by $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and its $\{i,j,k\}$-th element is denoted by $[\boldsymbol{\mathcal{X}}]_{i,j,k}$.

A. Plug-and-Play Alternating Direction Method of Multipliers (PnP-ADMM)

Alternating Direction Method of Multipliers (ADMM) [11], [12] is a proximal splitting algorithm that can treat convex optimization problems of the form

$$\min_{\mathbf{x}\in\mathbb{R}^{N_1},\ \mathbf{z}\in\mathbb{R}^{N_2}} F(\mathbf{x}) + G(\mathbf{z}) \ \text{ s.t. } \ \mathbf{z} = \mathbf{L}\mathbf{x}, \tag{1}$$

where F and G are usually assumed to be a quadratic and proximable function, respectively, and $\mathbf{L} \in \mathbb{R}^{N_2 \times N_1}$ is a matrix with full-column rank. For any $\mathbf{x}^{(0)} \in \mathbb{R}^{N_1}, \mathbf{z}^{(0)} \in \mathbb{R}^{N_2}, \mathbf{b}^{(0)} \in \mathbb{R}^{N_2}$ and $\rho > 0$, the ADMM algorithm is given by

$$\begin{cases} \mathbf{x}^{(t+1)} = \arg\min_{\mathbf{x}} \left\{ F(\mathbf{x}) + \frac{\rho}{2} \|\mathbf{z}^{(t)} - \mathbf{L}\mathbf{x} - \mathbf{b}^{(t)}\|_2^2 \right\}, \\ \mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \left\{ G(\mathbf{z}) + \frac{\rho}{2} \|\mathbf{z} - \mathbf{L}\mathbf{x}^{(t+1)} - \mathbf{b}^{(t)}\|_2^2 \right\}, \\ \quad = \mathrm{prox}_{1/\rho G}(\mathbf{L}\mathbf{x}^{(t+1)} + \mathbf{b}^{(t)}) \\ \mathbf{b}^{(t+1)} = \mathbf{b}^{(t)} + \mathbf{L}\mathbf{x}^{(t+1)} - \mathbf{z}^{(t+1)}, \end{cases} \tag{2}$$

where the superscript (t) denotes the iteration number. The update of $\mathbf{z}$ reduces to the proximity operator of the function $G(\cdot)$[1]. The sequence generated by (2) quickly converges to an optimal solution of (1).

In PnP-ADMM [14], [15], the solution of the sub-problem w.r.t. $\mathbf{z}$ (assuming $\mathbf{L}$ is the identity matrix) is replaced by an off-the-shelf noise removal algorithm, to yield

$$\mathbf{z}^{(t+1)} = \mathcal{D}_\sigma \left(\mathbf{x}^{(t+1)} + \mathbf{b}^{(t)} \right), \tag{3}$$

where $\mathcal{D}_\sigma$ denotes the Gaussian denoiser, and σ is the standard deviation of the assumed additive white Gaussian noise (AWGN). Auto-encoder-based networks are often used as denoiser [15].

B. Robust Principal Component Analysis

RPCA decomposes observational data, represented as a two- or three-dimensional tensor, into a low-rank tensor and a sparse tensor. When applied to HS images, RPCA enhances the low-rank structure of the tensor while preserving spatial correlation.

Let $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ (n_1 and n_2 are the height and width of the HS image, and n_3 is the number of the spectral bands) be an observed HS image. The standard RPCA optimization problem [8], [9] is defined by

$$\min_{\boldsymbol{\mathcal{L}}, \boldsymbol{\mathcal{S}}} \|\boldsymbol{\mathcal{L}}\|_* + \lambda \|\boldsymbol{\mathcal{S}}\|_1 \quad \text{s.t.} \quad \boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}, \tag{4}$$

where $\boldsymbol{\mathcal{L}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ are the background and anomaly components, respectively, and $\lambda > 0$ is a balancing weight to control the two terms. The first term is the regularization term of $\boldsymbol{\mathcal{L}}$ based on the nuclear norm, which promotes the low-rankness of $\boldsymbol{\mathcal{L}}$. The second term is the regularization term of $\boldsymbol{\mathcal{S}}$ based on the ℓ_1-norm, which promotes sparsity of $\boldsymbol{\mathcal{S}}$. Note that RPCA assumes $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}$ as the observation model.

III. Proposed Method

The proposed method aims to achieve high-precision and robust anomaly detection in HS images through three steps: training the AEAN model using denoised HS images by singular value decomposition (SVD), solving the optimization problem based on RPCA with AEAN using PnP-ADMM, and generating the anomaly map from the estimated anomaly component by combining the WLRX detector [1], [2].

[1]The proximity operator [13] is a key tool of proximal splitting techniques. Let $\mathbf{x} \in \mathbb{R}^N$ be an input vector. For any $\gamma > 0$, the proximity operator of f over $\mathbb{R}^N$ is defined by $\mathrm{prox}_{\gamma f}(\mathbf{x}) := \arg\min_{\mathbf{y} \in \mathbb{R}^N} f(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{x} - \mathbf{y}\|^2$.

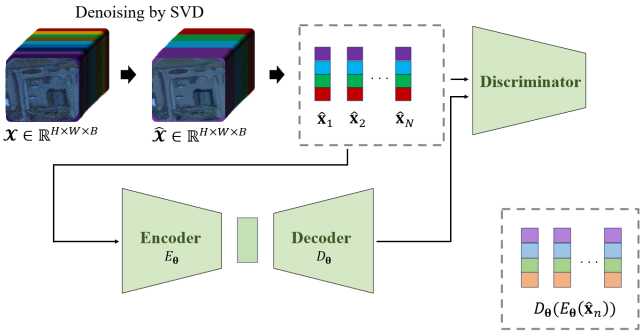

Fig. 1. The 1D-AEAN model and its training.

TABLE I
THE PROPOSED 1D-AEAN ARCHITECTURE. PARAMETER NOTATION: KERNEL SIZE × NUMBER OF INPUT CHANNELS × NUMBER OF OUTPUT CHANNELS

Nets	Layer	Parameters	BN	Activation
Encoder	Conv1D	11 × 1 × 5	YES	LReLU
	Conv1D	9 × 5 × 10	YES	LReLU
	Conv1D	7 × 10 × 15	YES	LReLU
	Conv1D	5 × 15 × 20	YES	LReLU
Decoder	Conv1DTranspose	5 × 20 × 15	YES	LReLU
	Conv1DTranspose	7 × 15 × 10	YES	LReLU
	Conv1DTranspose	9 × 10 × 5	YES	LReLU
	Conv1DTranspose	11 × 5 × 1	NO	Sigmoid
Discriminator	Conv1D	11 × 1 × 5	YES	LReLU
	Conv1D	9 × 5 × 10	YES	LReLU
	Conv1D	7 × 10 × 15	YES	LReLU
	Conv1D	5 × 15 × 20	YES	LReLU
	Dense	20 × 1	NO	Sigmoid

A. Training the AEAN Model

We train the 1D model of the AEAN (1D-AEAN) to learn an encoder-decoder structure for reconstructing background components from HS images. The conventional method [7] uses the modified kernel RX algorithm [3] to eliminate anomaly pixels and identify background pixels for training. However, RX is noise-sensitive and computationally expensive. To address this, we preprocess using SVD to remove noise and anomalous pixels, generating training data (Fig. 1).

Let an observed HS image $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{H \times W \times B}$ be rearranged into a 2D matrix $\mathbf{X} \in \mathbb{R}^{N \times B}$, where $N = HW$. The matrix $\mathbf{X}$ is denoised using SVD by retaining the top 90% of the singular components. The process involves centering the data, performing SVD, retaining the significant singular values, and reconstructing the matrix. This denoised HS data is then used to train the 1D-AEAN model.

Table I summarizes the details of the architecture of the proposed 1D-AEAN model.

B. Solving the proposed optimization problem

Next, we describe the proposed optimization problem based on RPCA with AEAN. In the proposed method, we enhance the accuracy of estimating the sparse tensor $\boldsymbol{\mathcal{S}}$, which represents the anomalous regions, by modeling the low-rank nature of the background regions using a low-dimensional model (low-dimensional tensor $\boldsymbol{\mathcal{L}}$) based on the 1D-AEAN model.

To adopt the 1D-AEAN model as the proximity operator in the PnP-ADMM framework, we define $\mathcal{R}_{\text{AEAN}}$ as the regularization term in the minimization problem. We formulate the proposed optimization problem as follows.

$$\min_{\boldsymbol{\mathcal{L}},\boldsymbol{\mathcal{S}}} \mathcal{R}_{\text{AEAN}}(\boldsymbol{\mathcal{L}}) + \lambda \|\boldsymbol{\mathcal{S}}\|_{2,1}^{(\mathcal{G})} \quad \text{s.t. } \boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}, \tag{5}$$

where $\lambda > 0$ is a balancing weight of two terms.

The second term is the group ℓ_1-norm of $\boldsymbol{\mathcal{S}}$, defined as

$$\|\boldsymbol{\mathcal{S}}\|_{2,1}^{(\mathcal{G})} := \left\| \left\{ \|\boldsymbol{\mathcal{S}}^{(\mathcal{G}_1)}\|_2, \ldots, \|\boldsymbol{\mathcal{S}}^{(\mathcal{G}_N)}\|_2 \right\} \right\|_1, \tag{6}$$

where $\{\mathcal{G}_1, \ldots, \mathcal{G}_N\} \in \mathcal{G}$ are the sets of indices of the one-dimensional vectors in the spectral direction, and $\mathcal{G}_n$ has the index sets of the spectral vector of the n-th pixel.

To find a solution of (5), we employ the PnP-ADMM algorithm described in Section II-A. The minimization problem (5) is not directly applicable to PnP-ADMM. Thus, we reformulate it as

$$\min_{\boldsymbol{\mathcal{L}},\boldsymbol{\mathcal{S}}} \mathcal{R}_{\text{AEAN}}(\boldsymbol{\mathcal{L}}) + \lambda \|\boldsymbol{\mathcal{S}}\|_{2,1}^{(\mathcal{G})} + \iota_{\{\boldsymbol{\mathcal{X}}\}}(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}), \tag{7}$$

where $\iota_{\{\boldsymbol{\mathcal{X}}\}}(\cdot)$ is the indicator function[2] of $\{\boldsymbol{\mathcal{X}}\}$. This function guarantees the equal constraint $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}$ on optimization. Furthermore, by introducing auxiliary variables $\boldsymbol{\mathcal{Z}}_1, \boldsymbol{\mathcal{Z}}_2$, and $\boldsymbol{\mathcal{Z}}_3$, we rewrite the minimization problem into the following equivalent expression:

$$\begin{aligned} \min_{\boldsymbol{\mathcal{L}},\boldsymbol{\mathcal{S}},\boldsymbol{\mathcal{Z}}_i (i=1,2,3)} & \mathcal{R}_{\text{AEAN}}(\boldsymbol{\mathcal{Z}}_1) + \lambda \|\boldsymbol{\mathcal{Z}}_2\|_{2,1}^{(\mathcal{G})} + \iota_{\{\boldsymbol{\mathcal{X}}\}}(\boldsymbol{\mathcal{Z}}_3), \\ & \text{s.t. } \boldsymbol{\mathcal{Z}}_1 = \boldsymbol{\mathcal{L}}, \boldsymbol{\mathcal{Z}}_2 = \boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{Z}}_3 = \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}}. \end{aligned} \tag{8}$$

The minimization problem (8) can be applied to PnP-ADMM[3].

C. Generating Anomaly Maps

In the WLRX-based detector [2], assuming that background pixels follow a multivariate Gaussian distribution, we use the estimated $\boldsymbol{\mathcal{S}}$ as weights in WLRX to calculate the final anomaly map. The proposed RPCA with AEAN separates the spectrum of anomalous regions into a sparse tensor, assigning high weights to anomalous pixels. To enhance robustness, we use the inverse of the result of the closing operation on $\sum_k |[\boldsymbol{\mathcal{S}}]_{i,j,k}|$ as weights $\omega_i (i = 1, \ldots, N)$.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbb{R}^B$ denote the sub-vectors corresponding to each pixel of the input HS image $\boldsymbol{\mathcal{X}}$. We calculate the weighted mean vector $\mathbf{m}$ as $\mathbf{m} = \sum_n \hat{\omega}_n \mathbf{x}_n$ and the weighted covariance matrix $\mathbf{C}$ as $\mathbf{C} = \sum_{n=1}^N \hat{\omega}_n (\mathbf{x}_n - \mathbf{m})(\mathbf{x}_n - \mathbf{m})^\top$, where $\hat{\omega}_n$ is the n-th normalized weights. The anomaly score for $n = 1, \ldots, N$ is calculated by $AD(\mathbf{x}_n) = (\mathbf{x}_n - \mathbf{m})^\top \mathbf{C}^{-1} (\mathbf{x}_n - \mathbf{m})$.

[2]Let $\mathbf{x} \in \mathbb{R}^N$ be an input vector. For a given non-empty closed convex set $\mathcal{C}$, the indicator function of $\mathcal{C}$ is defined by $\iota_C(\mathbf{x})$ which returns 0 if $\mathbf{x} \in \mathcal{C}$, and $+\infty$ otherwise.

[3]We omit the details of the proposed PnP-ADMM algorithm due to space limitations.

TABLE II
AUCs OF EACH METHOD (THE HIGHEST AUC IS HIGHLIGHTED IN BOLD AND THE SECOND HIGHEST IN UNDERLINE.

Scene	LRX [1]	WLRX [2]	RPCA [8]	1D-AEAN [7]	Ours
Airport1	0.9563	0.9310	0.9575	<u>0.9598</u>	**0.9717**
Airport2	0.9623	0.9638	<u>0.9828</u>	0.9692	**0.9831**
Airport3	0.8280	0.9028	0.9147	<u>0.9252</u>	**0.9468**
Airport4	0.9821	0.9337	**0.9999**	0.9851	<u>0.9919</u>
Beach1	0.9751	<u>0.9853</u>	0.9760	0.9793	**0.9882**
Beach2	0.9443	0.9693	0.9457	<u>0.9804</u>	**0.9824**
Beach3	0.9984	0.9999	0.9957	**1.0000**	**1.0000**
Beach4	0.9708	**0.9910**	0.9612	0.9794	<u>0.9839</u>
Urban1	0.9912	0.9899	0.9771	<u>0.9929</u>	**0.9932**
Urban2	0.9548	**0.9971**	0.9649	0.9793	<u>0.9907</u>
Urban3	0.9791	0.9342	<u>0.9897</u>	0.9818	**0.9910**
Urban4	0.9822	0.9930	0.9943	<u>0.9960</u>	**0.9967**
Urban5	0.9573	0.9635	<u>0.9791</u>	0.9660	**0.9795**
Average	0.9601	0.9657	0.9722	<u>0.9765</u>	**0.9845**

IV. Experimental Results

The proposed anomaly detection method was evaluated using ROC and AUC metrics on the ABU (airport, beach, urban) dataset. Noisy short-wavelength bands were removed, image intensity was normalized to $[0, 1]$, and the spectral domain was resized to 80 bands using bicubic interpolation.

We compared the proposed method with local-RX (LRX) [1], WLRX [2], and RPCA [8]. In addition, to verify the effectiveness of RPCA with AEAN in our method, we also compared it with the 1D-AEAN model. For each input HS image, we trained the 1D-AEAN model as described in Section III-A, computed the reconstruction error, and generated the anomaly map from the error as described in Section III-C. This approach can be seen as a slightly modified version of the 1D-AEAN model presented in [7]. We termed it "1D-AEAN".

Table II shows the results of the AUC scores for the proposed and conventional methods. The proposed method consistently outperforms conventional methods in most scenes and secures high ranks in several specific scenes. While LRX shows a lower average performance, WLRX and RPCA perform competitively. Additionally, 1D-AEAN achieves notable results in specific scenes, including a perfect score in one and a high average performance across methods.

Figure 2 shows binary anomaly maps generated by our method and conventional methods to maximize the F-score against the ground truth. The proposed method results in fewer false detections and higher accuracy across all scenes. RX-based methods and RPCA show lower accuracy and higher false detections of non-anomalous pixels. 1D-AEAN has slightly more false detections and lower accuracy than the proposed method.

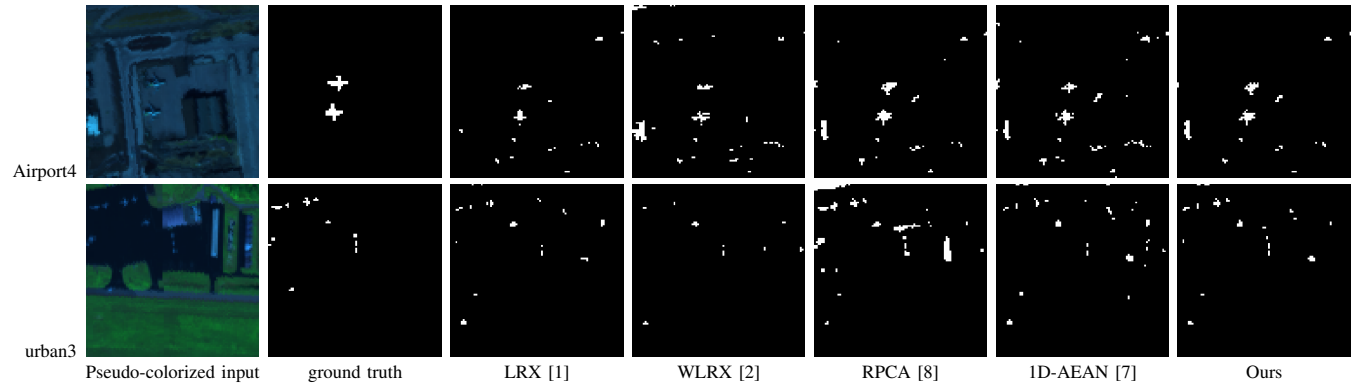

Fig. 2. Anomaly detection results obtained by binarizing the anomaly maps to maximize the F-score.

V. Conclusion

We introduced the unsupervised anomaly detection method for HS images. This method combines RPCA and AEAN and is solved using PnP-ADMM. We utilize a pre-trained autoencoder as a proximity operator in RPCA to encourage low rankness in the reconstructed spectral data, effectively detecting anomalous pixels from the background. The experimental results confirmed the effectiveness of our method compared to some conventional approaches.

Acknowledgment

This work was partially supported by the FAIS Research and Development Project Support Program, "Satellite Data Utilization Demonstration and New Technology Development Project," and JSPS KAKENHI Grant Number 24K15075.

References

[1] S. Küçük and S. E. Yüksel, "Comparison of rx-based anomaly detectors on synthetic and real hyperspectral data," in *Proc. Workshop Hyperspect. Image Sig. Process.: Evol. Remote Sens. (WHISPERS)*, IEEE, 2015, pp. 1–4.

[2] Q. Guo, B. Zhang, Q. Ran, L. Gao, J. Li, and A. Plaza, "Weighted-rxd and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2351–2366, 2014.

[3] C. Zhao, X.-F. Yao, and Y. Yan, "Modified kernel RX algorithm based on background purification and inverse-of-matrix-free calculation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 544–548, 2017.

[4] W. Xie, J. Lei, B. Liu, Y. Li, and X. Jia, "Spectral constraint adversarial autoencoders approach to feature representation in hyperspectral anomaly detection," *Neural Networks*, vol. 119, pp. 222–234, 2019.

[5] W. Xie, B. Liu, Y. Li, J. Lei, and Q. Du, "Autoencoder and adversarial-learning-based semisupervised background estimation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5416–5427, 2020.

[6] Z. Zhao and B. Sun, "Hyperspectral anomaly detection via memory-augmented autoencoders," *CAAI Transactions on Intelligence Technology*, vol. 8, no. 4, pp. 1274–1287, 2023.

[7] S. Arisoy, N. M. Nasrabadi, and K. Kayabol, "Unsupervised pixel-wise hyperspectral anomaly detection via autoencoding adversarial networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2021.

[8] W. Sun, C. Liu, J. Li, Y. M. Lai, and W. Li, "Low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery," *Journal of Applied Remote Sensing*, vol. 8, no. 1, pp. 083 641–083 641, 2014.

[9] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Int. Conf. Comput. vis. pattern recognit. (CVPR)*, 2016, pp. 5249–5257.

[10] M. Mishima, I. Kobayashi, and R. Matsuoka, "Hyperspectral anomaly detection based on local-tensor-nuclear-norm," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. Symp. (IGARSS)*, 2023, pp. 2157–2160.

[11] D. Gabay and B. Mercier, "A dual algorithm for the solution of non linear variational problems via finite element approximation," *Comput. Math. App.*, vol. 2, no. 1, pp. 17–40, 1976.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[13] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *C. R. Acad. Sci. Paris*, vol. 255, pp. 2897–2899, 1962.

[14] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play admm for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 84–98, 2016.

[15] Z. Meng, X. Yuan, and S. Jalali, "Deep unfolding for snapshot compressive imaging," *International Journal of Computer Vision*, vol. 131, no. 11, pp. 2933–2958, 2023.