Stacked Graph Fusion Denoising Autoencoder for Hyperspectral Anomaly Detection

Yongshan Zhang, *Member, IEEE*, Yijiang Li, Xinxin Wang, Xinwei Jiang, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*— Anomaly detection for hyperspectral images (HSIs) is a challenging problem to distinguish a few anomalous pixels from a majority of background pixels. Most existing methods cannot simultaneously explore both structural and spatial information from global and local perspectives. In this letter, we propose a stacked graph fusion denoising autoencoder (SGFDAE) for hyperspectral anomaly detection. Specifically, the global and local graphs are constructed from an HSI to explore potential structural and spatial information. With the designed graph fusion strategy, an advanced graph denoising autoencoder with deep architecture is developed in a hierarchical manner. To achieve better reconstruction and detection, a greedy layerwise unsupervised pretraining strategy is presented for network training. Experiments show that SGFDAE achieves 97.17%, 98.43%, and 98.90% detection accuracies by averaging the results of the datasets from three different scenes and outperforms the state-of-the-art methods.**

***Index Terms*— Anomaly detection, denoising autoencoder, graph neural network, hyperspectral imagery.**

I. Introduction

HYPERSPECTRAL images (HSIs) are always represented by 3-order tensor and provide abundant spatial and spectral information for Earth observation [1]. As an essential application for HSIs, anomaly detection is to identify very limited anomalous pixels that significantly deviate from dominant background pixels [2]. It is a challenging research hotspot and earns much attention from researchers.

To uncover anomalous pixels, many effective methods have been proposed [3]. As an important category, shallow learning methods come from traditional machine learning that can be roughly divided into statistics-based and representation-based methods. Based on the Gaussian distribution assumption, statistics-based methods perform detection with the distance measure, e.g., the Reed–Xiaoli (RX) detector [4] and its variants [5]. Representation-based methods distinguish background pixels and anomalies based on the constructed dictionary or decomposed sparse matrix, e.g., collaborative representation-based detectors [6], sparse representation-based detectors [7], and tensor-based detectors [8].

Differently, deep learning models are based on neural networks to detect anomalies. Self-supervised learning-based networks have been recently proposed to detect anomalies for HSIs. For example, pixel-shuffle downsampling blind-spot reconstruction network [9], and blind-spot self-supervised learning network [10] perform detection based on self-supervised learning technique and blind-spot architecture. For HSIs, global and local information can be a complementarity for each other to promote detection. Jiang et al. [11] proposed a low-rank embedded network to learn global features for detection without considering local characteristics. Furthermore, Lu et al. [12] detected anomalies based on a manifold-constrained autoencoder with both global and local reconstruction errors. They exploit global and local information at feature level without exploring the complex relationships among pixels. Graph neural network can be a good detector due to its superior ability of exploring the potential relationships between normal and anomalous pixels [13], [14]. To achieve better detection, Fan et al. [13] and Li et al. [14] presented advanced graph autoencoders by exploring pixel relationships with local spatial information in graph structures. However, they do not pay attention to global information in graph construction, limiting their performance enhancement.

Motivated by these observations, we propose an SGFDAE for hyperspectral anomaly detection, as shown in Fig. 1. Based on the global and local graphs, an advanced graph autoencoder is developed with graph fusion and random corruption. A greedy layerwise unsupervised pretraining strategy is presented to train the deep architecture. The advantage of SGFDAE is that structural information and spatial consistency are exploited from global and local perspectives to promote detection.

The contributions of this letter are outlined as follows.

1) We propose an SGFDAE detector by considering structural and spatial information from global and local perspectives. The optimal network is learned with a greedy layerwise pretraining strategy for better detection.
2) The local graph is constructed with the guidance of diverse homogenous regions. It can be treated as the complementarity of the global graph to provide additional useful information for further learning.
3) A graph fusion strategy is devised to fuse the global and local graphs into a unified graph for the proposed

Manuscript received 3 April 2024; revised 17 May 2024; accepted 6 June 2024. Date of publication 19 June 2024; date of current version 8 July 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62106241 and in part by the Knowledge Innovation Program of Wuhan-Shuguang under Grant 2023010201020335. *(Corresponding author: Xinwei Jiang.)*

Yongshan Zhang, Yijiang Li, and Xinwei Jiang are with the School of Computer Science, China University of Geosciences, Wuhan 430074, China (e-mail: yszhang.cug@gmail.com; ysjxw@hotmail.com).

Xinxin Wang and Yicong Zhou are with the Department of Computer and Information Science, University of Macau, Macau, China (e-mail: yicongzhou@um.edu.mo).

Digital Object Identifier 10.1109/LGRS.2024.3416454

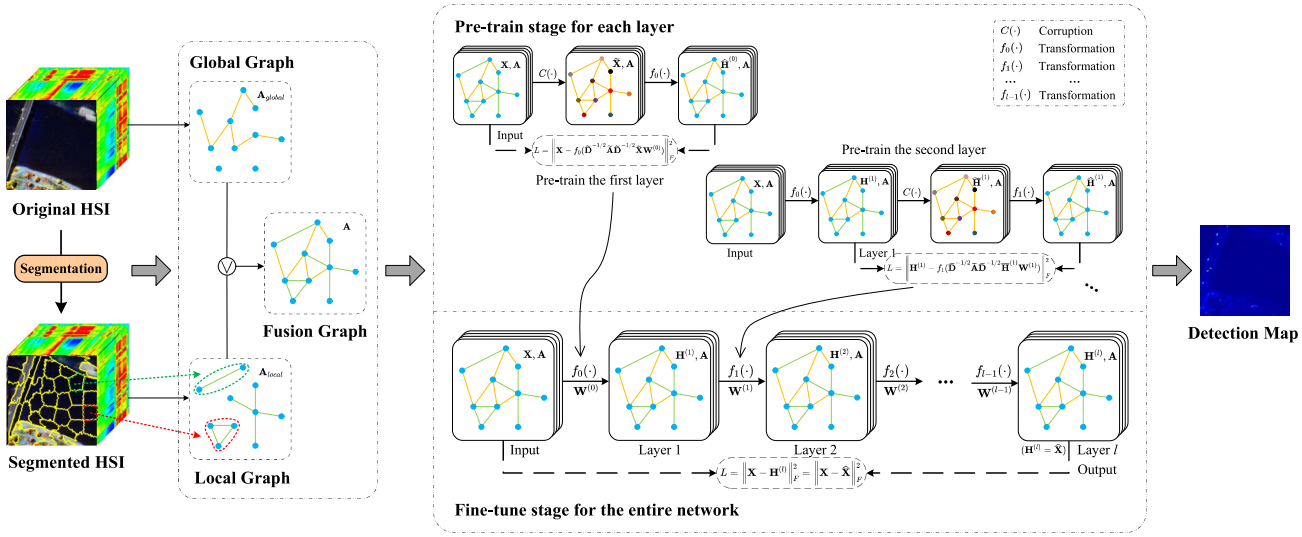

Fig. 1. Overview of the proposed SDGDAE detector. Given an HSI, the global and local graphs are constructed to explore the relationships among all pixels and the pixels within the same local region, respectively. Fusing the global and local graphs, a deep graph denoising autoencoder is developed and trained with a greedy layerwise pretraining strategy for anomaly detection.

detector. It provides global context information and local spatial information in an individual graph structure.

II. Proposed Method

A. Graph Construction

To explore structural information, we construct graph structures from both global and local perspectives that can provide complementary information for each other.

1) Global Graph Construction: It is common to measure the neighborhood relationship between two pixels within k nearest neighbors of each other based on the spectral similarity [15]. Assume that the HSI cube with $W \times H$ pixels and B bands is denoted as $\mathcal{X} \in \mathbb{R}^{W\times H\times B}$, we reshape it into a matrix $\mathbf{X} \in \mathbb{R}^{N\times B}$ $(N = W \times H)$. To construct the graph, we treat each pixel as a node and the relationship between two pixels as an edge. Given any two pixels $\mathbf{x}_i$ and $\mathbf{x}_j$, we define the global graph $\mathbf{A}_{\text{global}} \in \mathbb{R}^{N\times N}$ by traversing the whole HSI to record structural information as

$$A_{\text{global}(ij)} = \begin{cases} 1, & \mathbf{x}_i \in \mathcal{N}_k(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where $\mathcal{N}_k(\mathbf{x}_i)$ denotes the set of k nearest neighboring pixels of $\mathbf{x}_i$. Specifically, two pixels with neighborhood relationship should be connected by an edge if they belong to the k nearest neighboring pixel set of each other. Otherwise, there is no connecting edge between them.

2) Local Graph Construction: Spatial information of local regions reflecting the land-cover distribution is important for HSI analysis. To localize s homogenous regions, we first perform principal component analysis (PCA) [1] on $\mathbf{X}$ to obtain the first principal component I_f with major information, and then conduct the simple linear iterative clustering (SLIC) method [16] on I_f according to the texture complexity and spatial consistency. Mapping the segmentation labels to original HSI, we can obtain s nonoverlapping superpixels with all different bands as

$$\mathcal{X} = \{P_u\}_{u=1}^{s} \tag{2}$$

where P_u is the uth superpixel containing pixels with similar spectral signatures and neighboring spatial distribution.

To construct the graph with local spatial consistency, we also treat each pixel as a node and discover the relationship between two pixels as an edge according to the segmentation results. Given any two pixels $\mathbf{x}_i \in P_u$ and $\mathbf{x}_j \in P_v$, we define the local graph $\mathbf{A}_{\text{local}} \in \mathbb{R}^{N\times N}$ based on each superpixel to explore spatial consistency as

$$A_{\text{local}(ij)} = \begin{cases} 1, & u = v \text{ and } \left(\mathbf{x}_i \in \mathcal{N}_k^u(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_k^u(\mathbf{x}_i)\right) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where $\mathcal{N}_k^u(\mathbf{x}_i)$ represents the set of k nearest neighboring pixels of $\mathbf{x}_i$ in the uth superpixel. Different from (1), the pixelwise relationships are explored within the same superpixel.

B. Stacked Graph Fusion Denoising Autoencoder Detector

With $\mathbf{A}_{\text{global}}$ and $\mathbf{A}_{\text{local}}$, we design a graph fusion strategy to obtain a unified fusion graph and develop an SGFDAE detector with greedy layerwise unsupervised pretraining strategy.

1) Graph Fusion Strategy: The construction of $\mathbf{A}_{\text{global}}$ depends on the nearest spectral similarity of pairwise pixels by traversing the entire image without considering local spatial information. In contrast, the construction of $\mathbf{A}_{\text{local}}$ relies on the relationship of pairwise pixels within each homogenous region but ignores the situation that two highly related pixels are spatially located far away. It has been verified that local spatial information within homogenous regions is beneficial for HSI analysis. When using both the global and local graphs, they can provide complementary information for each other to enhance detection performance.

Thus, we design a graph fusion strategy to fuse $\mathbf{A}_{\text{global}}$ and $\mathbf{A}_{\text{local}}$ into a unified fusion graph $\mathbf{A} \in \mathbb{R}^{N\times N}$ as

$$\mathbf{A} = \mathbf{A}_{\text{global}} \vee \mathbf{A}_{\text{local}} \tag{4}$$

where $\vee$ denotes the OR logical operation. In (4), $\mathbf{A}$ is obtained by performing OR logical operation on elements in the same

position of $\mathbf{A}_{\text{global}}$ and $\mathbf{A}_{\text{local}}$. With this fusion strategy, $\mathbf{A}$ shows the graph structure of pixels with both global and local information.

2) Network Architecture: Compared with traditional networks [17], graph networks are good at learning informative representations of samples/nodes based on their neighborhood information [13]. With the graph structures as prior knowledge, graph networks are more able to learn global and local features.

To improve the generalization and robustness, we design an advanced graph autoencoder by introducing random feature corruption and using the fusion graph for HSI reconstruction and detection. The denoising process can be regarded as a regularization for the network to achieve better reconstruction. Suppose that $\mathbf{X}$ is randomly corrupted, we can get its corrupted matrix $\widetilde{\mathbf{X}} \in \mathbb{R}^{N\times B}$. Specifically, each element in $\mathbf{X}$ may be corrupted by setting as 0 with a probability p $(0 < p < 1)$. To reconstruct $\mathbf{X}$ with $\mathbf{A}$ from $\widetilde{\mathbf{X}}$, we formulate the graph fusion denoising autoencoder with only one layer as

$$\begin{aligned} L(\mathbf{W}) &= \left\|\mathbf{X} - f\left(\widetilde{\mathbf{X}}, \mathbf{A}\right)\right\|_F^2 = \left\|\mathbf{X} - \widehat{\mathbf{X}}\right\|_F^2 \\ &= \left\|\mathbf{X} - f\left(\widetilde{\mathbf{D}}^{-1/2}\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-1/2}\widetilde{\mathbf{X}}\mathbf{W}\right)\right\|_F^2 \end{aligned} \tag{5}$$

where $f(\cdot)$ is the activation function, $\mathbf{W} \in \mathbb{R}^{B\times B}$ is the weight matrix, and $\widehat{\mathbf{X}} = f(\widetilde{\mathbf{X}}, \mathbf{A})$ is the reconstruction of $\mathbf{X}$ learned from $\widetilde{\mathbf{X}}$. Besides, $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$ is the fusion graph matrix with self-loops and $\widetilde{\mathbf{D}}$ is the degree matrix with diagonal elements $\widetilde{D}_{ii} = \sum_{j=1}^{N} \widetilde{A}_{ij}$.

Equation (5) can be used as a building block to construct a deep architecture by stacking it layer by layer. For example, we take $\mathbf{X}$ as the input to learn $\mathbf{H}^{(1)}$ as the output for the first layer. Then, we use $\mathbf{H}^{(1)}$ as the input to learn $\mathbf{H}^{(2)}$ as the output of the second layer. Given l layers, we regard the feature representation of the lth layer $\mathbf{H}^{(l)}$ as the output of the whole network that is learned from the $(l-1)$th layer. Thus, we can obtain the proposed SGFDAE with deep architecture for anomaly detection. Since the proposed detector uses neighborhood information to learn the reconstructed pixels, it can well-reconstruct background pixels while anomalies are not and achieve superior detection, because the connected pairwise pixels in the global and local graphs belong to background pixels with higher probability than anomalous pixels. The underlying reason for this situation is that the fixed size of neighborhoods for constructing the global and local graphs may introduce improper connections between background pixels and anomalous pixels due to the very limited size of anomalous pixels and a large proportion of background pixels.

3) Network Training: Here, we adopt a greedy layerwise unsupervised pretraining strategy to learn the optimal network. In this strategy, pretraining is to learn the predefined weights as network initialization instead of using randomly generated weights, while fine-tuning is to learn to optimal weights based on pretrained initialization.

For the pretraining stage, network weights of each layer are determined by the reconstruction of a single-layer graph denoising autoencoder. Given $\mathbf{H}^{(i)}$ as the input for the ith layer, we first get its corrupted matrix $\widetilde{\mathbf{H}}^{(i)}$ and then learn its reconstruction $\widehat{\mathbf{H}}^{(i)}$ with $\mathbf{A}$ from $\mathbf{H}^{(i)}$ via the transformation $f_i(\cdot)$ as

$$L\left(\mathbf{W}^{(i)}\right) = \left\|\mathbf{H}^{(i)} - f_i\left(\widetilde{\mathbf{D}}^{-1/2}\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-1/2}\widetilde{\mathbf{H}}^{(i)}\mathbf{W}^{(i)}\right)\right\|_F^2 \tag{6}$$

where $\mathbf{W}^{(i)} \in \mathbb{R}^{B\times B}$ is the pretrained weight matrix for the ith layer. In (6), we hope $\widehat{\mathbf{H}}^{(i)} = f_i(\widetilde{\mathbf{H}}, \mathbf{A})$ could be well-reconstructed from $\widehat{\mathbf{H}}^{(i)} = \mathbf{H}^{(i)}$ as possible, where $\widehat{\mathbf{H}}^{(0)} = \mathbf{X}$. After the reconstruction, we can get $\mathbf{W}^{(i)}$ to learn the feature representation for the next layer as $\mathbf{H}^{(i+1)} = f_i(\mathbf{H}^{(i)}, \mathbf{A})$. Network pretraining can be achieved in such a manner layer by layer. Finally, we can obtain $\mathbf{H}^{(l)} = f_{l-1}(\mathbf{H}^{(l-1)}, \mathbf{A})$ as the output of the whole network with l layers.

For the fine-tuning stage, the pretrained $\{\mathbf{W}^{(i)}\}_{i=0}^{(l-1)}$ are used as initialization. Taking $\widehat{\mathbf{X}} = \mathbf{H}^{(l)}$ as the deep reconstruction of $\mathbf{X}$, we formulate the loss function of SGFDAE as

$$L\left(\left\{\mathbf{W}^{(i)}\right\}_{i=0}^{(l-1)}\right) = \left\|\mathbf{X} - \mathbf{H}^{(l)}\right\|_F^2 = \left\|\mathbf{X} - \widehat{\mathbf{X}}\right\|_F^2. \tag{7}$$

Equation (7) describes the reconstruction error between the input $\mathbf{X}$ and it final deep reconstruction $\mathbf{H}^{(l)}$. To achieve the optimal reconstruction, we adopt the backpropagation algorithm to fine-tune all the weight matrices. Compared with random initialization, the pretrained initialization is more easily fine-tuned to reach the optimal state when training network.

C. Anomaly Detection

After network training, we measure the difference between the original pixel $\mathbf{x}_i \in \mathbf{X}$ and the reconstructed pixel $\widehat{\mathbf{x}}_i \in \widehat{\mathbf{X}}$ to detect anomalous pixels from an HSI. The degree of $\mathbf{x}_i$ to be anomalous can be quantified by

$$d_i = \left\|\mathbf{x}_i - \widehat{\mathbf{x}}_i\right\|_2 \tag{8}$$

where d_i is the anomalous degree of $\mathbf{x}_i$. The higher value of d_i indicates the higher probability of $\mathbf{x}_i$ being an anomaly. Potential anomalies can be identified with a predefined threshold.

III. Experiments

A. Experimental Setup

1) Datasets: We use 13 HSI datasets for anomaly detection.[1] There are four images for the airport scene, four images for the beach scene, and five images for the urban scene. These datasets are with the size of 100×100 or 150×150.

2) Compared Methods: To verify the effectiveness of our SGFDAE, we use several state-of-the-art methods for comparison. The four learning-based methods are RX [4], FrFE [18], GTVLRR [7], and PTA [8], while the five network-based methods are RGAE [13], LREN [11], Auto-AD [19], PDB-SNet [9], and GT-HAD [20].

3) Evaluation Metrics: To measure the detection performance, we adopt three well-accepted metrics for evaluation, i.e., receiver operating characteristic (ROC) curve, area under ROC curve (AUC) value, and separability map between anomaly and background.

[1] http://xudongkang.weebly.com/data-sets.html

TABLE I

COMPARISON OF AUC VALUES ACROSS DIFFERENT ANOMALY DETECTION METHODS ON HSI DATASETS. THE OPTIMAL RESULTS ARE SHOWN IN BOLD

Scenes	Images	Traditional learning-based methods				Network-based methods					
		RX	FrFE	GTVLRR	PTA	RGAE	LREN	Auto-AD	PDBSNet	GT-HAD	SGFDAE
Airport	Airport(1)	0.8221	0.8865	0.9009	0.8743	0.8856	0.9179	0.8711	0.9544	**0.9657**	0.9492
	Airport(2)	0.8404	0.9544	0.8935	0.9095	0.9553	0.9536	0.8872	0.9486	0.9524	**0.9806**
	Airport(3)	0.9288	0.9267	0.9141	0.9165	0.9562	0.9592	0.9201	0.9505	0.9605	**0.9612**
	Airport(4)	0.9526	0.9789	0.9776	0.9954	0.9723	0.9902	0.8537	0.9807	0.9947	**0.9957**
	Average	0.8860	0.9366	0.9215	0.9239	0.9424	0.9552	0.8830	0.9585	0.9683	**0.9717**
Beach	Beach(1)	0.9807	0.9808	0.9782	0.9569	0.9352	0.9536	0.9461	0.9539	0.9653	**0.9837**
	Beach(2)	0.9106	0.8277	0.9273	0.8961	0.9721	**0.9739**	0.9057	0.8991	0.8934	0.9671
	Beach(3)	**0.9998**	0.9996	0.9889	0.9464	0.9922	0.9965	0.9307	0.9629	0.9685	0.9983
	Beach(4)	0.9538	0.8556	0.9813	0.9312	0.9027	0.9785	0.9746	0.9539	0.9757	**0.9883**
	Average	0.9612	0.9159	0.9689	0.9327	0.9506	0.9756	0.9393	0.9424	0.9507	**0.9843**
Urban	Urban(1)	0.9907	0.9892	0.9027	0.9803	0.9895	0.9790	0.9778	0.9851	0.9804	**0.9974**
	Urban(2)	0.9946	0.9956	0.8556	0.9973	0.9976	0.9960	0.9955	0.9987	**0.9992**	0.9983
	Urban(3)	0.9513	0.9626	0.9355	0.9703	0.9735	0.9545	0.9665	0.9709	0.9575	**0.9812**
	Urban(4)	0.9887	0.9790	0.9269	0.9960	0.9971	0.9905	0.9972	0.9902	0.9661	**0.9988**
	Urban(5)	0.9692	0.9639	0.9368	0.9568	0.9597	0.9567	0.9620	0.9593	0.9486	**0.9695**
	Average	0.9789	0.9781	0.9115	0.9801	0.9835	0.9753	0.9798	0.9808	0.9703	**0.9890**

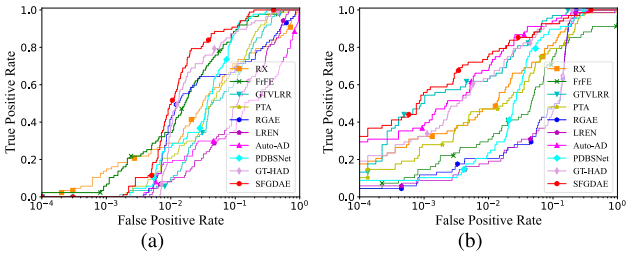

Fig. 2. ROC curves of different methods. (a) Airport(2). (b) Beach(4).

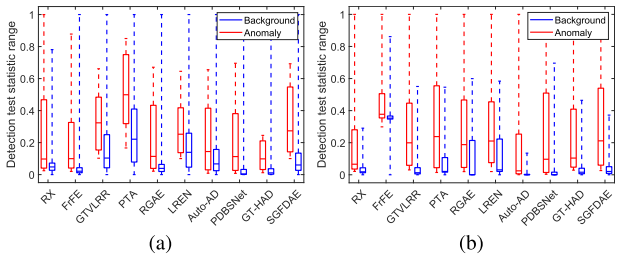

Fig. 3. Separability maps of different methods. (a) Airport(2). (b) Beach(4).

4) Parameter Settings: There are four hyperparameters in SGFDAE. In our setting, k and s are both chosen from {50, 100, 150, 300}, l varies from {2, 3, 5}, and p is tuned from {0, 0.1, 0.2, 0.3}. For other compared methods, the hyperparameters are set as suggested in original papers. The reported results are obtained by averaging ten repeated trials with the optimal parameter setting.

B. Results and Analyses

1) Detection Performance: The AUC results of different detection methods are shown in Table I. Except for the Airport(1), Beach(2), and Beach(3) datasets, SGFDAE consistently shows superior performance than the competitors. Compared with the superior methods (PDBSNet and GT-HAD), SGFDAE still obtains at least 4.19% and 3.36% average enhancements on the airport scene datasets. The superiority of SGFDAE over other competitors demonstrates its capability of anomaly detection by integrating the effective graph fusion strategy into the deep denosing architecture to simultaneously explore structural and spatial information in global and local perspectives.

Fig. 2 presents the ROC curves of different methods to investigate the detection performance with different thresholds. Compared with the competitors, SGFDAE always acquires higher values of higher true positive rate (TPR) with different values of false-positive rate (FPR) or obtains a higher TPR even when FPR is extremely low. To survey the ability of separating anomalies from background pixels, we show the anomaly–background separability maps in Fig. 3. The anomaly and background boxes of SGFDAE are completely or better separated, while those of other methods are overlapped more or less. This indicates that SGFDAE shows the superior anomaly–background separation ability by highlighting anomalies and suppressing background pixels. In summary, SGFDAE is a good anomaly detector for HSIs.

TABLE II

COMPARISON OF RUNNING TIME ACROSS DIFFERENT METHODS ON HSI DATASETS FROM THE AIRPORT SCENE (MEASURED BY SECONDS)

Methods	Airport(1)	Airport(2)	Airport(3)	Airport(4)
RX	0.8221	0.8404	0.9288	0.9526
FrFE	12.36	11.68	15.78	16.45
GTVLRR	137.56	167.39	146.32	129.98
PTA	46.38	18.76	29.33	27.38
RGAE	122.45	149.56	125.36	209.09
LREN	109.78	124.25	136.05	122.37
Auto-AD	198.35	206.57	289.65	294.82
PDBSNet	494.42	526.68	551.78	512.32
GT-HAD	50.91	50.08	52.62	51.23
SGFDAE	90.46	52.51	46.33	84.32

2) Running Time Study: Table II shows the running time of all the compared methods on the airport scene datasets. It is clear that RX requires the minimum runtime because of its simple calculation and noniterative property. Among the network-based methods, SGFDAE requires the second shortest runtime behind GT-HAD, showing its high efficiency due to the simple network design.

3) Ablation Study: To investigate the effect of the graph fusion strategy, we develop two variants of SGFDAE for ablation study. The one is SGGDAE integrating with only the global graph $\mathbf{A}_{\text{global}}$, and the other is SLGDAE embedded with only the local graph $\mathbf{A}_{\text{local}}$. As shown in Table III, SGFDAE

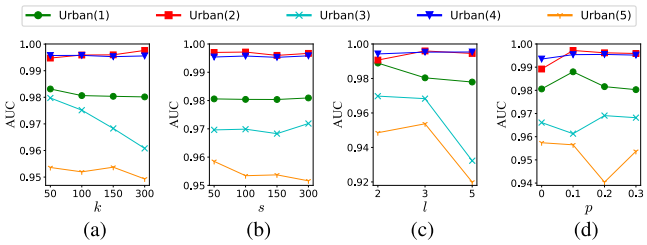

Fig. 4. Effect of different parameters for SGFDAE on HSI datasets from the urban scene. (a) Effect of k. (b) Effect of s. (c) Effect of l. (d) Effect of p.

TABLE III

COMPARISON OF AUC VALUES ACROSS SGFDAE AND ITS VARIANTS ON HSI DATASETS. THE OPTIMAL RESULTS ARE SHOWN IN BOLD

Scenes	Images	SGGDAE	SLGDAE	SGFDAE
Airport	Airport(1)	0.8913	0.9469	**0.9492**
	Airport(2)	0.9720	0.9754	**0.9806**
	Airport(3)	0.9611	0.9541	**0.9612**
	Airport(4)	0.9853	**0.9968**	0.9957
	Average	0.9524	0.9683	**0.9717**
Beach	Beach(1)	0.9683	0.9836	**0.9837**
	Beach(2)	0.9554	0.9337	**0.9671**
	Beach(3)	**0.9988**	0.9986	0.9983
	Beach(4)	0.9715	0.9786	**0.9883**
	Average	0.9735	0.9737	**0.9843**
Urban	Urban(1)	0.9911	0.9856	**0.9974**
	Urban(2)	0.9989	**0.9990**	0.9983
	Urban(3)	0.9801	0.9754	**0.9812**
	Urban(4)	0.9959	0.9965	**0.9988**
	Urban(5)	0.9584	0.9614	**0.9695**
	Average	0.9849	0.9836	**0.9890**

shows superior performance than the two variants in most cases. SGFDAE obtains 6.50% improvement compared with SGGDAE on the Airport(1) dataset. Similar improvements can also be found on other datasets. This demonstrates the advantage of the designed graph fusion strategy to treat global and local information as the complementary materials in a geometric structure for superior anomaly detection.

4) Parameter Sensitivity Study: To study the parameter sensitivity of SGFDAE, we alternately vary one parameter and fix others as default values. The variation in AUC on the urban scene datasets is reported in Fig. 4. It is noted that the AUC values are varying with different parameter settings. Thus, it is necessary to allocate appropriate values for the four parameters to pursue better detection performance.

IV. CONCLUSION

In this letter, we proposed an SGFDAE for hyperspectral anomaly detection. To explore geometric structures, the global graph is constructed by traversing the whole image and the local graph with spatial consistency is constructed by traversing each superpixel individually. With the global and local graphs, a graph fusion strategy is designed to generate a unified fusion graph. Based on the fusion graph, SGFDAE is designed and trained with a greedy layerwise unsupervised pretraining strategy for reconstruction and anomaly detection. Experiments on HSI datasets have demonstrated the superiority of SGFDAE.

REFERENCES

[1] Y. Zhang, G. Jiang, Z. Cai, and Y. Zhou, "Bipartite graph-based projected clustering with local region guidance for hyperspectral imagery," *IEEE Trans. Multimedia*, early access, doi: 10.1109/TMM.2024.3394975.

[2] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022.

[3] Y. Xu, L. Zhang, B. Du, and L. Zhang, "Hyperspectral anomaly detection based on machine learning: An overview," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3351–3364, 2022.

[4] 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, Aug. 1990.

[5] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397, Feb. 2005.

[6] 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.

[7] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, Jan. 2020.

[8] L. Li, W. Li, Y. Qu, C. Zhao, R. Tao, and Q. Du, "Prior-based tensor approximation for anomaly detection in hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1037–1050, Mar. 2022.

[9] D. Wang, L. Zhuang, L. Gao, X. Sun, M. Huang, and A. J. Plaza, "PDB-SNet: Pixel-shuffle downsampling blind-spot reconstruction network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5511914.

[10] L. Gao, D. Wang, L. Zhuang, X. Sun, M. Huang, and A. Plaza, "BS3LNet: A new blind-spot self-supervised learning network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5504218.

[11] K. Jiang, W. Xie, J. Lei, T. Jiang, and Y. Li, "LREN: Low-rank embedded network for sample-free hyperspectral anomaly detection," in *Proc. AAAI Conf. Artif. Intell.*, May 2021, vol. 35, no. 5, pp. 4139–4146.

[12] X. Lu, W. Zhang, and J. Huang, "Exploiting embedding manifold of autoencoders for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1527–1537, Nov. 2020.

[13] G. Fan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "Hyperspectral anomaly detection with robust graph autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511314.

[14] K. Li et al., "Spectral difference guided graph attention autoencoder for hyperspectral anomaly detection," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–17, 2023.

[15] X. Chen, Y. Zhang, X. Feng, X. Jiang, and Z. Cai, "Spectral-spatial superpixel anchor graph-based clustering for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.

[16] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[17] Y. Wu, X. Chen, X. Huang, K. Song, and D. Zhang, "Unsupervised distribution-aware keypoints generation from 3D point clouds," *Neural Netw.*, vol. 173, May 2024, Art. no. 106158.

[18] R. Tao, X. Zhao, W. Li, H.-C. Li, and Q. Du, "Hyperspectral anomaly detection by fractional Fourier entropy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4920–4929, Dec. 2019.

[19] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Auto-AD: Autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5503314.

[20] J. Lian, L. Wang, H. Sun, and H. Huang, "GT-HAD: Gated transformer for hyperspectral anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, doi: 10.1109/TNNLS.2024.3355166.