

Journal Pre-proofs

Morphological-based Attention in Graph Networks for Hyperspectral Anomaly Detection

Maryam Imani, Daniele Cerra

PII: S0273-1177(26)01378-5
DOI: <https://doi.org/10.1016/j.asr.2026.09.068>
Reference: JASR 20160

To appear in: *Advances in Space Research*

Received Date: 6 April 2026
Revised Date: 27 August 2026
Accepted Date: 24 September 2026



Please cite this article as: Imani, M., Cerra, D., Morphological-based Attention in Graph Networks for Hyperspectral Anomaly Detection, *Advances in Space Research* (2026), doi: <https://doi.org/10.1016/j.asr.2026.09.068>

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 COSPAR. Published by Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Morphological-based Attention in Graph Networks for Hyperspectral Anomaly Detection

Maryam Imani^{1,*}, Daniele Cerra²

¹ Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

² Remote Sensing Technology Institute, German Aerospace Center (DLR), Muenchener Strasse 20, Wessling, 82234, Germany

Corresponding author*: maryam.imani@modares.ac.ir

Abstract

Characterizing spectral-spatial features at both local and global scales, and discriminating anomalies from background, are key objectives in hyperspectral anomaly detection. To address these issues, we propose a dual branch network extracting complementary local and global spectral-spatial features. A cross-attention mechanism is applied between the input patch and opening and closing morphological profiles, which convey information about anomalies and background, respectively, to enhance anomaly and background separation. In the first branch, consecutive convolutional layers with skip connections capture multi-level spectral-spatial semantic features. In the second branch, relationships between each spectrum and cluster centroids derived from the full hyperspectral image are modeled as a graph and processed by a graph convolutional network to extract global features. Features from both branches are fused to estimate the background, and the reconstruction residual is used as the anomaly score. Comparisons with several state-of-the-art methods on benchmark hyperspectral datasets demonstrate the proposed method's effectiveness and competitive performance, while maintaining a reasonable computational cost. An ablation study further assesses the contribution of each network component.

Keywords: attention mechanism, morphological operators, graph convolutional network, hyperspectral, anomaly detection.

1. Introduction

Hyperspectral imaging captures rich spectral information for each image element, providing discriminative information that enables the identification of subtle differences between materials. As a result, hyperspectral images are widely used in land cover classification, resource management, change detection, environmental monitoring, mineral exploration, agriculture and target detection (Aldwairi et al., 2025; Lv et al., 2025; Wang et al., 2025a). Unlike hyperspectral target detection (HTD), which typically relies on binary hypothesis testing to identify targets with available spectral information (Imani, 2025), hyperspectral anomaly detection (HAD) is more challenging because it aims to identify unknown targets without prior knowledge of their spectral signatures. In HTD, recent deep learning approaches have attempted to alleviate the limitations caused by sparse target samples and incomplete target priors through adaptive background modeling and feature transfer strategies. For example, a pixel-weighted background reconstruction network with dual-windows blind block (DWBRN) employs pixel-weighted background reconstruction and dual-window blind blocks to suppress target interference during background estimation (Li et al., 2025a). Similarly, a transformer-based cross-domain few-shot learning framework was proposed to address limited target samples and domain shifts in HTD by exploiting transferable feature representations (Feng et al., 2025). However, unlike HTD methods, HAD methods cannot rely on predefined target information and must simultaneously model complex backgrounds and unknown anomalies in an unsupervised manner.

Over the past decades, both classical (Matteoli et al., 2014; Zhou et al., 2016) and modern methods (Wang et al., 2026; Xu et al., 2026) have been proposed for hyperspectral anomaly detection. Classical approaches are mainly based on statistical modeling or dictionary learning, where background representations are constructed from local or global atoms via optimization. While these methods are computationally efficient and perform well in simple and homogeneous scenes, their performance significantly degrades in real-world scenarios with heterogeneous backgrounds, nonlinear spectral variability, and strong similarity between anomalies and background materials.

To overcome these limitations, recent deep learning-based methods have been introduced. Convolutional neural networks enable hierarchical feature learning, while transformer-based models improve long-range dependency modeling. However, most of these approaches rely on large-scale supervision or, in unsupervised settings, depend

heavily on background reconstruction using autoencoders or generative models. This makes anomaly detection challenging, since the separation between anomaly and background is not explicitly guided and often depends on indirect reconstruction errors. As a result, these methods may produce unstable detection performance in complex scenes.

Although several recent studies attempt to enhance discrimination through attention mechanisms or structural priors, two key limitations still remain unresolved: (i) insufficient exploitation of explicit prior knowledge about anomaly and background structures, and (ii) limited integration of local contextual features with global scene-level relationships in a unified framework.

Some detectors rely on morphological operators for contextual feature extraction, either to improve data representation or to refine the obtained anomaly maps. Here, we consider instead morphological opening and closing profiles as features related to anomalies and background, respectively, exploiting them as prior knowledge for unsupervised anomaly detection. Therefore, we compute the similarity between an input patch and the corresponding opening and closing profiles, deriving indicators respectively of anomalies and background, by applying a cross-attention mechanism. This approach not only enriches the feature space by involving contextual features of hyperspectral data, but also enhances the class discrimination by exploring the relationships between indicators of anomalies and background and the input spectra.

The application of convolutional filters to the cross-attention feature maps leads to local feature extraction. In addition to local features, global features are also important to separate anomalies from background (Yang et al., 2025). Graph theory has been recently used to provide a global representation of complex hyperspectral images. Nevertheless, they require a large amount of computational resources when considering a large number of nodes in the graph, as they involve deriving high dimensional adjacency matrices; furthermore, they also usually ignore local information (Ma et al., 2023; Yang et al., 2024). To address this issue, a cluster-based graph convolutional network is introduced in this work, achieving global feature extraction with a relatively low computational burden. Exploiting both local morphological attention-based features and global graph features result in rich local and global spectral-spatial features for hyperspectral anomaly detection.

Effective exploitation of spectral-spatial information at both local and global scales, together with improved discrimination among anomalies and background, are both important for hyperspectral anomaly detection. To address this aspect, a two-branch network is proposed. The first branch uses three inputs: an input patch, the opening morphological profile containing features related to the anomalies, and the closing morphological profile containing low-frequency background spatial variations. By computing the cross-correlation between the input patch and each morphological profile, the network learns to discriminate between anomalies and background, while consecutive convolutional layers extract spectral and spatial features. In the second branch, relationships between the pixel under test and clusters derived from the full hyperspectral image are modelled as a graph, which is processed by a graph convolutional network (GCN) to capture global information. The local spectral-spatial features and global graph features are finally fused within an end-to-end network to reconstruct the background. The most relevant contributions of this work are the following:

1. A morphological element-wise cross-attention mechanism is proposed for hyperspectral anomaly detection. Unlike existing morphological approaches that mainly use opening and closing profiles as preprocessing or postprocessing operators, the proposed method treats these profiles as complementary prior representations of anomalies and background, respectively, and explicitly models their relationships with the input patch through cross-attention to enhance feature discrimination.
2. A lightweight cluster-based graph convolutional network is introduced to capture global contextual information. Instead of constructing graphs from all image pixels, which leads to large adjacency matrices and high computational complexity, the proposed approach builds the graph from cluster centroids, significantly reducing computational cost while preserving global spectral-spatial relationships.
3. A dual-branch feature fusion framework is developed to jointly exploit local morphological attention features and global graph features. The local branch extracts discriminative spectral-spatial representations from morphological cross-attention maps, while the global branch models long-range contextual dependencies using the cluster-based GCN. Their complementary information is fused end-to-end for background reconstruction and anomaly detection.

2. Literature review

2.1. Statistical and Distribution-Based Methods

Earliest anomaly detectors, such as Reed-Xiaoli (RX) (Reed and Yu, 1990), rely on the statistical distribution of the background to measure the distance of the pixel under test from the background. RX assumes a Gaussian

distribution for the background, and relies on the Mahalanobis distance to compute the anomaly score. However, complex background data may not be well represented by a Gaussian distribution. Different versions of RX such as kernel RX (Kwon and Nasrabadi, 2005), weighted RX (WRX) (Guo et al., 2014), and median-mean line RX (MMLRX) (Imani, 2017) have been proposed to boost the performance of RX. The two-step generalized likelihood ratio test (2S-GLRT) (Liu et al., 2022) estimates covariance in a two-step likelihood framework. The kernel isolation forest-based anomaly detector (KIFD) (Li et al., 2020) assumes that anomalies are more easily isolated than background pixels, due to their distinctive spectral features. Therefore, an isolation forest is constructed using a set of random samples from the hyperspectral image; subsequently, pixels which are more easily isolated in the kernel space are identified as anomalies. In addition, KIFD uses a recursive detection strategy jointly exploiting local and global information in the image. Despite improvements, these methods remain sensitive to distribution mismatch and limited modeling capacity in complex scenes.

2.2. Representation-Based Methods

Other anomaly detectors are dictionary-based, for example low-rank representation-based, sparse representation (SR), and collaborative representation (CR) methods. In SR models, an overcomplete dictionary can represent background pixels but fails to represent anomalies, with the reconstruction residual used as the anomaly score (Xu et al., 2020). The weighted low-rank sparse dictionary learning model combines low-rank representation with sparse representation to effectively represent complex background data (Li and Yuan, 2024). The CR-based detector (CRD) (Li and Du, 2015), along with its variants, improve the detection performance with a lower complexity by relaxing the problem to a l_2 -norm optimization.

Different versions of the CR-based detectors have been suggested (Vafadar and Ghassemian, 2018; Imani, 2024), such as the local Summation Unsupervised Nearest Regularized Subspace with an Outlier Removal Anomaly Detector (LSUNRSORAD) and the Local Summation Anomaly Detection based on CR and Inverse Distance Weight (LSAD-CR-IDW) (Tan et al., 2019). The approximation accuracy of the linear representation in CRD decreases as the number of heterogeneous pixels in the background increases. To address this issue, the CRD with background purification and saliency weight (CRDBPSW) method introduces an outlier removal method step to purify the background data (Hou et al., 2022). A discriminative coefficient analysis-based CR method is introduced for hyperspectral anomaly detection in (Yang et al., 2024).

Although representation-based methods are not dependent on the statistical distribution of data, they still cannot handle anomaly detection in spectrally complex hyperspectral images. Background reconstruction is an approach widely used for anomaly detection in this case, which is efficient if a model adaptable to the background features can be found. The anomalies can then be detected by reconstructing the background and suppressing its contribution from the original image. Some classic methods also rely on background reconstruction, such as RX exploiting background statistics, or CRD using a local dictionary-based background reconstruction.

2.3. Band Selection Methods

Background reconstruction-based anomaly detectors benefit from spectral band selection, as not all bands contribute equally to anomaly separability in hyperspectral data. However, the lack of prior knowledge or labeled samples makes unsupervised band selection challenging. To address this, self-supervised band selection (SSBS) (Wang and Hu, 2025) identifies a compact set of discriminative bands that enhance anomaly-background separation. It introduces a feature diversity constraint and evaluates subset collaboration to improve both diversity and representativeness of selected bands.

Residual-driven band selection (RDBS) (Shang et al., 2022) uses density peak clustering to estimate anomaly/background priors and derives band prioritization criteria for selecting representative anomaly-sensitive bands. The spectral-spatial out-of-distribution-based method (HADUBS) (Zhang et al., 2024c) further evaluates band importance using out-of-distribution characteristics and enhances robustness via a mutual-information-based local outlier factor.

More recently, DBS-BRN (Liu et al., 2026) integrates band selection into the reconstruction process using a learnable Top-k mechanism with response-guided scoring, enabling adaptive optimization of spectral subsets during anomaly detection. Overall, these methods demonstrate that adaptive band selection improves reconstruction-based anomaly detection by reducing spectral redundancy and enhancing discriminability, although most approaches still decouple band selection from spatial feature learning.

2.4. Background Modeling and Reconstruction Methods

Recent approaches have further improved background modeling by explicitly exploiting spectral decomposition, spatial structure, and anomaly–background differences. Zhao et al. (2024) proposed HELWC, which combines empirical mode decomposition (EMD) with local weighted contrast. EMD is first employed to reduce spectral noise and estimate a spectral reference, after which Jensen–Shannon divergence and local weighted contrast are used to measure spectral differences and enhance anomaly responses. This strategy improves background suppression while reducing the influence of spectral variability.

More recently, You et al. (2026) introduced S3CRAD, which integrates superpixel-guided background inpainting with spatial–spectral constrained representation. The superpixel-guided strategy preserves anomaly boundaries during background estimation, while multi-feature fusion and spatial–spectral weighting constrain the representation coefficients to better distinguish anomalies from their reconstructed background. The final anomaly score is obtained from the residual between the original and reconstructed pixels.

Wu et al. (2026) proposed DSFNet, a dual-stream spatial–spectral fusion network that separately enhances spectral and spatial representations before fusing them for background reconstruction. Its local spectral feature enhancement module emphasizes anomaly–background spectral differences, while a multi-branch convolution module captures multi-scale spatial structures. The reconstructed background is subsequently compared with the original hyperspectral image using Mahalanobis distance to obtain the anomaly map. These methods demonstrate that effective anomaly detection increasingly depends on accurate background characterization and on explicitly modeling complementary spectral and spatial information. However, background reconstruction may still suffer when anomalous structures are partially incorporated into the estimated background or when the learned representation does not explicitly exploit anomaly- and background-related priors.

2.5. Deep Learning-Based Methods

Recently, deep learning is being increasingly utilized for hyperspectral representation, thanks to its capability of performing a hierarchical deep feature extraction, which can be useful for anomaly detection in images with complex background. A light convolutional neural network (CNN) based on residual learning is proposed for hyperspectral anomaly detection in (Zhang et al., 2024a). The Fourier conditional mask-based mixed attention (FCMMA) introduced a Fourier conditional mask-based mixed attention mechanism to suppress anomalous high-frequency components while preserving background information in the frequency domain (Wang et al., 2025b).

Commonly used networks for hyperspectral anomaly detection are autoencoders (AEs) and generative adversarial networks (GANs). A dual-discriminative conditional GAN enhanced by an optimization algorithm is suggested in (Shanmugam and Amali, 2024), producing hyperspectral images closer to the original ones. Convolutional AEs and GANs do not take into account bands intercorrelation during background reconstruction. To mitigate inter-channel correlation, an attention-based GAN is introduced in (Bahmanpour et al., 2025) by applying both skip connections and attention mechanisms on both convolutional channels and spectral channels.

To enhance the hyperspectral anomaly detection using AEs, a pixel-associated AE (PAAE) is introduced in (Xiang et al., 2024), which computes global pixel similarities by exploiting a superpixel distance estimation, resulting in the construction of background and anomaly dictionaries. Self-exploration advantage of dictionary representation is integrated with the ability to explore details from complex scenes using AEs in (Hu et al., 2025), relying on an attention constraint low-rank and sparse AE model.

Due to background spectral variability, complex background samples may be poorly reconstructed, leading to high false alarm rates or missed detections. To solve this issue, the use of two cascaded AEs with an adaptive pixel-level attention network (CAPNet) is proposed in (Zhao et al., 2025a). As another cascaded AEs, we can refer to the AE unmixing network for hyperspectral anomaly detection at sub-pixel level for targets which result mixed with the background (Li et al., 2025b). As most AEs do not consider the correlation of spectral, spatial and structural information both locally and globally, a local-to-global spatial-structural-spectral joint reconstruction framework (S3S-LTGRMA) is introduced in (Chen et al., 2025), which integrates multi-head attention, graph regularization, and superpixel modeling for joint spectral–spatial–structural learning.

More recently, Huo et al. (2026) proposed a dual-stream background modeling network with anomaly suppression (DSBMNet) to address both anomaly leakage and inaccurate background reconstruction. DSBMNet employs a multi-directional masked convolution module to suppress the reconstruction of anomalous regions, while its local and global

branches exploit attention-driven multi-scale grouped convolution and spatial-spectral attention to model fine-grained and global background characteristics, respectively. These complementary mechanisms improve the purity of reconstructed backgrounds and reduce both missed detections and false alarms.

Recent deep networks have increasingly attempted to explicitly decouple anomaly enhancement from background reconstruction. Zhao et al. (2026) proposed TBCNet, a twin-branch collaborative network that jointly performs anomaly enhancement and background suppression using complementary spectral and spatial information. A center-oriented spectral enhancement Mamba block is introduced to strengthen anomaly-sensitive spectral representations by aggregating contextual information around the center pixel. In parallel, a weighted-mask-constrained local spatial convolution suppresses anomalous interference during background reconstruction, thereby producing cleaner background representations. The outputs of the two branches are subsequently integrated through a cross-guided spatial-spectral fusion module. Although this design provides effective complementary modeling of anomaly and background information, it relies on relatively sophisticated sequence modeling and spatial reconstruction modules.

2.6. Morphological Methods

Morphological approaches exploit spatial context via filtering and profile operations. In (Imani, 2018), the potential of the morphological filters for contextual feature extraction is exploited to provide an appropriate representation in each principal component of the hyperspectral image. Subsequently, a morphological profile and attribute filter algorithm is introduced in (Andika et al., 2020). At first, an appropriate band containing anomaly information is selected based on entropy and histogram counts. Then, the morphological profile is used to suppress the contribution of the background in the selected band. After that, the attribute filter is used to filter the false anomalous pixels. The discrimination ability of autoencoders is combined with spatial information conveyed by morphological operators in (Feng and Zhang, 2021). More recent approaches include intrinsic decomposition with morphological filtering (Jiao et al., 2025).

Recent studies have also explored morphological and local contextual information together with spatial-spectral feature representation. Zhao et al. (2025b) proposed HLSC-SSGF, which combines local sub-block contrast with spatial-spectral gradient features for hyperspectral anomaly detection. The method employs local sub-block ratio-multiply contrast to enhance low-contrast anomalies and suppress background responses, while spatial-spectral gradient features are used to distinguish anomalous structures from complex background edges. A complementary ratio-difference contrast mechanism further exploits gradient-angle differences between targets and their surrounding background. This combination enables HLSC-SSGF to exploit local contrast, spatial structure, spectral information, and gradient characteristics for improved anomaly discrimination. Nevertheless, its detection mechanism remains primarily based on handcrafted contrast and gradient measures rather than jointly learning complementary local and global representations in an end-to-end framework.

2.7. Graph-Based Methods

Graph-based methods model hyperspectral images as graphs to capture global dependencies. Early hybrid approaches (Zhang et al., 2024b) combine graph Fourier transform with autoencoders (GFT-AE) but lack full spatial modeling. The Graph Attention Network-Beta Wavelet Graph Neural Network-based (GAN-BWGNN) hyperspectral anomaly detection method is introduced in (A et al., 2025), relying on graphs, attention mechanism, and beta wavelets to identify anomalies. Each pixel is here considered as a node in the graph, with edges indicating pixel correlations based on spectral features and derived by applying the K-nearest neighbor algorithm. Here, the Graph Attention Network component dynamically prioritizes relevant spatial characteristics using an adaptive attention mechanism. The beta wavelets also identify spectral anomalies by leveraging the right-shifted spectral energy phenomenon where the presence of anomalies causes a right shift in spectral energy. A similar approach is followed in (Ding et al., 2025).

Category	Representative Methods	Key Idea	Main Limitation
Statistical	RX, KRX, WRX, MMLRX, 2S-GLRT, KIFD	Background statistical modeling or isolation-based detection	Strong distribution assumptions, poor robustness in heterogeneous scenes

Representation-based	SR, CRD, LSUNRSORAD, LSAD-CR-IDW, CRDBPSW	Sparse/linear reconstruction using local dictionaries	Limited modeling capacity for complex backgrounds
Band Selection	SSBS, RDBS, HADUBS, DBS-BRN	Select or learn discriminative spectral bands for reconstruction-based detection	Often decoupled from spatial learning or full spectral-spatial interaction
Background Modeling & Reconstruction	HELWC, S3CRAD, DSFNet	Spectral decomposition, background inpainting/reconstruction, and complementary spatial-spectral modeling	Possible anomaly leakage into background estimation and dependence on reconstruction quality
Deep Learning	CNN, FCMMA, GAN, AE, PAAE, CAPNet, S3S-LTGRMA, DSBMNet, TBCNet	Hierarchical feature learning, anomaly suppression, and reconstruction-based background modeling	Implicit anomaly modeling, high computational cost
Morphological and Local Contextual Methods	Morphological profiles, attribute filters, HLSC-SSGF, hybrid AE methods	Exploit spatial context through morphological operators, local contrast, and spatial-spectral gradient features	Often rely on predefined/handcrafted contextual representations rather than jointly learning complementary anomaly-background priors
Graph-Based	GFT-AE, GAN-BWGNN, ESGTN	Global spectral-spatial modeling via graph structures	High computational complexity, weak local-global fusion

Table 1. Taxonomy and comparison of hyperspectral anomaly detection methods.

As individual pixels are considered as graph nodes, computations may be unfeasible in larger scenes. Moreover, although global features are well extracted through graph processing and attention mechanisms, local features are not subsequently explicitly combined for the final anomaly detections. The Enhanced Spectral Graph Transformer Network (ESGTN), suggested in (Furkat Eshpulatovich et al., 2025), synergistically combines hyperbolic space embedding, transformer architectures and graph-based modeling. Both spectral and spatial relationships are modeled by graphs, with a transformer with self-attention mechanisms highlighting salient features, and a distortion-minimized representation of the data is provided by hyperbolic embeddings.

Table 1 summarizes representative hyperspectral anomaly detection methods across different categories, highlighting their key ideas and main limitations.

2.8. Research Gap and Motivation

Despite the significant progress achieved by existing hyperspectral anomaly detectors, several important challenges remain unresolved. Statistical and representation-based methods are generally limited by strong assumptions regarding

background distribution or reconstruction quality, making them less effective in heterogeneous scenes. Recent deep learning approaches have substantially improved feature representation through CNNs, autoencoders, transformers, and attention mechanisms. Nevertheless, most of these methods primarily rely on implicit background reconstruction or feature learning, without explicitly incorporating prior information that characterizes anomalies and background.

Morphological and contextual methods have demonstrated their effectiveness in exploiting spatial information; however, existing approaches mainly use morphological profiles, contrast measures, or gradient features as predefined representations or detection/refinement mechanisms. The complementary roles of opening and closing profiles as anomaly- and background-related priors have not been explicitly modeled through a learnable cross-attention mechanism.

Similarly, graph-based methods are effective in capturing global spectral-spatial relationships, but they commonly construct graphs using all image pixels as graph nodes, resulting in high computational complexity, redundant graph connections, and limited scalability to large hyperspectral scenes. Moreover, local morphological information and global graph representations are often learned independently, preventing effective exploitation of their complementary characteristics.

Therefore, there remains a clear research gap in developing an unsupervised anomaly detector that simultaneously (i) explicitly exploits the complementary information provided by morphological opening and closing profiles during feature learning, (ii) efficiently captures global contextual relationships with reduced computational complexity, and (iii) jointly integrates local morphological features and global graph representations within a unified end-to-end framework. To address these limitations, this work proposes a dual-branch network that combines morphological cross-attention with a lightweight cluster-based graph convolutional network for robust hyperspectral anomaly detection.

3. Proposed network

We propose a two-branch network for hyperspectral anomaly detection, which we refer to as Opening and Closing-based Attention integrated with Graph features (OCAG), presented in different configurations. The network is composed of two branches, capturing respectively local and global features. Subsequently, local and global features are analyzed jointly to represent the background. Finally, the residual term is associated to an anomaly score. Fig. 1 shows the block diagram of the proposed OCAG network, where part (a) represents the local branch, part (b) the global branch, and part (c) the background estimation. The first branch computes cross-correlations between the input patch and its associated opening and closing profiles, by an element-wise cross-attention (CA) mechanism extracting local features, while the second branch builds a clustering-based graph and relies on a graph convolutional network (GCN) to extract global features.

In the first branch, for each input patch, opening and closing profiles are first derived using three convolutional layers followed by a ReLU. Then, the CAs between the input patch and the openings (CA_o) and the closings (CA_c) are computed, and their concatenation with the processed input patch is fed to convolutional layers with skip connections to extract $\frac{n_b}{2}$ features, where n_b is the number of spectral bands. In the second branch, a feature matrix $\mathbf{X}$ and adjacency matrix $\mathbf{A}$ associated with the input patch are computed based on relationships between the central pixel and the clusters' centroids of the hyperspectral image, yielding global features. Similarly to the processing of the morphological profiles, $\mathbf{X}$ and $\mathbf{A}$ are fed to a GCN to extract $\frac{n_b}{2}$ features. Finally, the $\frac{n_b}{2}$ local features extracted from the first branch are fused with the $\frac{n_b}{2}$ global features extracted from the second branch to provide the background estimate $\hat{\mathbf{y}}$ of the central pixel $\mathbf{y}$. Details of the proposed network are described in the following sections.

3.1. Local branch

Applying the opening operator better isolates relevant objects, while applying the closing operator fills gaps on the contours, removes small holes, and smooths the image in dark regions. Hence, opening profiles contain details and variational information of anomalous targets, while closing profiles are driven by homogenous information in the background.

We extract the morphological profiles from the three principal components (PCs) with highest associated eigenvalues, after rotation of the hyperspectral image in the principal component space. Although the first three principal components capture the dominant spectral variability and are therefore used to construct the morphological profiles, low-variance components may still contain subtle anomaly-related spectral information. Investigating the

contribution of additional principal components remains an important direction for future study. A morphological profile is derived from the repeated application of the opening and closing operators by reconstruction (Gonzalez and Woods, 2018). For each PC of the hyperspectral image, the opening profile is defined by:

$$\Pi_\gamma(x) = \{\Pi_{\gamma\lambda} : \Pi_{\gamma\lambda} = \gamma_\lambda^*(x), \forall \lambda \in [0, n_f]\} \quad (1)$$

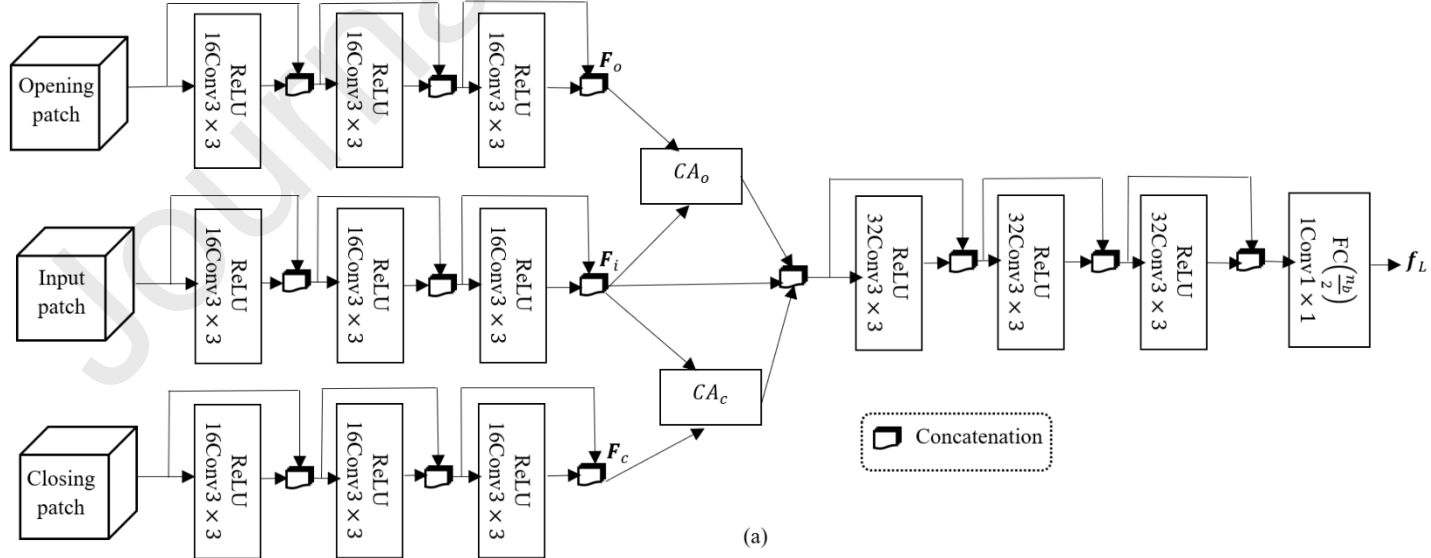
where n_f is the number of considered structuring elements (SEs), which determine the morphological profile dimensionality, λ is the size of chosen SE, and γ_λ^* indicates the morphological opening operator by reconstruction. Similarly, the closing profile is defined as follows:

$$\Pi_\varphi(x) = \{\Pi_{\varphi\lambda} : \Pi_{\varphi\lambda} = \varphi_\lambda^*(x), \forall \lambda \in [0, n_f]\} \quad (2)$$

where $\varphi_\lambda^*(x)$ represents the morphological closing operator by reconstruction. By applying n_f opening and closing operators by reconstruction in each PC using a disk-shaped SE with radius of $r = 1:n_f$, an opening profile with dimensionality $3n_f$ and a closing profile with dimensionality of $3n_f$ are associated to the three PCs. Disk-shaped SEs are widely used in morphological profile construction in hyperspectral image processing (Benediktsson et al., 2005; Imani and Ghassemian, 2019). Here, our choice was motivated by the isotropic spatial property of these SEs, which makes them particularly suitable for these tasks, where anomalies may appear in arbitrary orientations. Unlike square or diamond elements, which may introduce respectively axis-aligned bias and emphasize diagonal directions, the disk provides rotation-invariant spatial filtering, thus avoiding directional bias. Furthermore, disk SEs provide smoother area growth as their radius increases, resulting in more stable multiscale profile responses compared to the discrete growth pattern of square elements.

Each of the opening and closing profiles, along with the input patches, is individually processed through three convolutional layers, each containing 16 filters with a kernel size of 3×3 (16Conv3 $\times$ 3) for local feature extraction. The output of each layer is concatenated with its input through skip connection to allow the network to learn both low-level and high-level features. The feature maps generated from the input patch are denoted by $F_i \in \mathcal{R}^{p \times p \times d}$, while those derived by the opening and closing profiles are indicated as $F_o \in \mathcal{R}^{p \times p \times d}$ and $F_c \in \mathcal{R}^{p \times p \times d}$, respectively, where $p \times p$ is the input patch size and $d = 16$ is the dimensionality. To find the relationships between the input patch and an opening profile, the opening cross-attention (CA_o) is computed between F_i and F_o .

Unlike conventional cross-attention, where the query is typically generated from the target feature representation and the key/value are obtained from another modality or feature stream, the proposed cross-attention is designed to explicitly evaluate the consistency between the morphological prior and the input hyperspectral patch. Specifically, the input patch is used to generate both the key and value representations, preserving the original spectral-spatial information to be refined, while the opening or closing morphological profile generates the query representation,



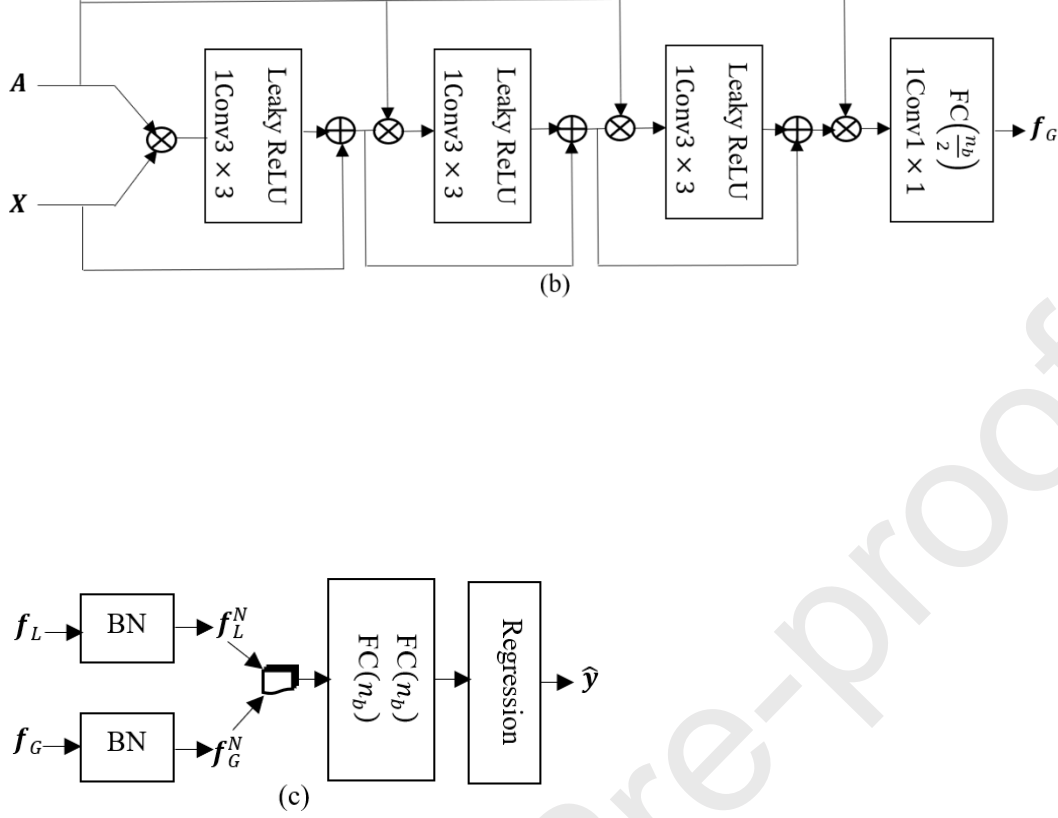


Fig. 1. Block diagram of the proposed OCAG network. (a): local branch, (b): global branch, (c): background estimation.

acting as a morphology-guided prior. Consequently, the attention weights are determined by measuring the similarity between the morphological profile and the input patch, allowing the network to selectively emphasize input features that are consistent with the structural characteristics represented by the opening or closing profile. This design differs fundamentally from standard attention mechanisms, whose objective is to model generic feature dependencies, whereas the proposed module incorporates task-specific prior knowledge into the attention process for anomaly-background discrimination.

To this end, the opening value V_o , opening key K_o , and opening query Q_o components are generated by applying 16 convolutional filters with a kernel size of 1×1 as follows:

$$V_o = 16\text{Conv}1 \times 1 (F_i) \quad (3)$$

$$K_o = 16\text{Conv}1 \times 1 (F_i) \quad (4)$$

$$Q_o = 16\text{Conv}1 \times 1 (F_o) \quad (5)$$

While $V_o \in \mathcal{R}^{p \times p \times d}$ and $K_o \in \mathcal{R}^{p \times p \times d}$ are computed from F_i , $Q_o \in \mathcal{R}^{p \times p \times d}$ is computed from F_o . Then, the opening element-wise cross-attention is computed by (Imani and Lim, 2025):

$$CA_o = \text{softmax}\left(\frac{K_o \odot Q_o}{\sqrt{d}}\right) \odot V_o \quad (6)$$

where $\odot$ denotes the elementwise multiplication, the softmax operation is then applied along the feature/channel dimension independently at each spatial location, and $CA_o \in \mathcal{R}^{p \times p \times d}$ quantifies the cross-correlations between the input patch and the opening profile. Similarly, the element-wise closing cross-attention (CA_c) is obtained with the closing value V_c , closing key K_c and closing query Q_c components computed as:

$$V_c = 16\text{Conv1} \times 1(F_i) \quad (7)$$

$$K_c = 16\text{Conv1} \times 1(F_i) \quad (8)$$

$$Q_c = 16\text{Conv1} \times 1(F_c) \quad (9)$$

where $V_c \in \mathcal{R}^{p \times p \times d}$ and $K_c \in \mathcal{R}^{p \times p \times d}$ are obtained from the input patch features F_i , and $Q_c \in \mathcal{R}^{p \times p \times d}$ from the closing profile features F_c . Subsequently, $CA_c \in \mathcal{R}^{p \times p \times d}$ is calculated as follows:

$$CA_c = \text{softmax}\left(\frac{K_c \odot Q_c}{\sqrt{d}}\right) \odot V_c \quad (10)$$

The proposed attention mechanism serves a different purpose from standard self-attention or transformer-based cross-attention. Instead of learning correlations solely among latent feature representations, it uses morphology-derived queries to guide the selection of informative spectral-spatial features from the input patch. Since the opening profile and the closing profile provide complementary structural information, the two cross-attention modules capture complementary relationships between the input patch and morphology-derived representations. The resulting feature maps therefore contain more discriminative information for separating anomalies from the background before subsequent convolutional feature extraction. Then, F_i , CA_o and CA_c features are concatenated along the depth:

$$F_{cat} = \text{cat}(F_i; CA_o; CA_c) \quad (11)$$

where $\text{cat}(\cdot; \cdot; \cdot)$ denotes the concatenation operation and $F_{cat} \in \mathcal{R}^{p \times p \times 3d}$ are the concatenated features additionally processed by three convolutional layers consisting of $32\text{Conv3} \times 3$ filters followed by a ReLU with residual connections. Finally, to extract the local features, a convolutional filter with kernel size of 1×1 ($1\text{conv1} \times 1$) is used for dimensionality reduction in the depth direction, and a fully connected (FC) layer with $\frac{n_b}{2}$ units is used to provide $\frac{n_b}{2}$ local features, denoted by $f_L \in \mathcal{R}^{\frac{n_b}{2} \times 1}$.

3.2. Global branch

To extract the global information from the general background, the hyperspectral image is clustered into m clusters by using the k-means algorithm (Rishch et al., 2022). Then, a graph is constructed for the central pixel in the input patch, consisting of m nodes, with each edge representing the relationship between the central pixel of the input patch and the m clusters' centroids. Associated with each input patch, a graph with related feature matrix of $X \in \mathcal{R}^{m \times n_b}$ and adjacency matrix of $A \in \mathcal{R}^{m \times m}$ is constructed, where the elements of the feature matrix are computed as:

$$X(j,:) = \mu_j \odot y; j = 1, \dots, m \quad (12)$$

where $y \in \mathcal{R}^{n_b \times 1}$ is the pixel under test in the center of the input patch, μ_j is the center of cluster j and $\odot$ is the elementwise multiplication. To quantify the dissimilarity between each pair of nodes, the Mahalanobis distance is employed (Ghosh et al., 2025). Specifically, the Mahalanobis distance between the i th and j th nodes is computed as

$$A(i,j) = (x_i - x_j)^T C^{-1} (x_i - x_j) \quad (13)$$

where $x_i \in \mathcal{R}^{n_b \times 1}$ and $x_j \in \mathcal{R}^{n_b \times 1}$ are the i th and j th row of the feature matrix, and $C \in \mathcal{R}^{n_b \times n_b}$ is the covariance matrix of all pixels in the image. The covariance matrix is computed once for the entire image and shared across all local patches, providing a consistent global statistical reference. This choice also avoids the instability of estimating covariance matrices independently from individual patches, where the number of available samples is limited. The

resulting matrix $\mathbf{A}$ therefore represents pairwise dissimilarities, where larger values indicate greater statistical differences between nodes. This distance-based relational matrix is subsequently used by the graph convolutional branch to propagate information among the graph nodes. The resulting distance-based graph is processed by the GCN to extract global feature representations.

The computed $\mathbf{X}$ and $\mathbf{A}$ are provided as inputs to the GCN for extracting graph features (Imani and Cerra, 2025). As illustrated in Fig. 1(b), the relational matrix $\mathbf{A}$ and feature matrix $\mathbf{X}$ are first multiplied through matrix multiplication to perform graph-based feature propagation. The resulting feature representation is then passed through a 3×3 convolution, which serves as a learnable projection for transforming the propagated feature representation, rather than modeling spatial relationships among the k-means cluster centroids. A ReLU activation is subsequently applied, and the resulting representation is added to the input feature matrix through a residual connection. The output is then passed to the next graph-convolutional layer. This operation is repeated three times. Finally, a filter and a fully connected layer with $\frac{n_b}{2}$ units are employed for feature reduction and generation of the global feature representation $\mathbf{f}_G \in \mathcal{R}^{\frac{n_b}{2} \times 1}$.

Since the graph nodes are generated from k-means cluster centroids, the resulting graph can potentially be affected by the initialization of the clustering procedure. To ensure reproducibility, a fixed random seed is employed during k-means clustering in all experiments. Nevertheless, this fixed initialization does not fully characterize the robustness of the graph representation to alternative cluster initializations. A systematic investigation of graph stability across multiple k-means initializations is therefore left for future work.

3.3. Background estimation

The local features $\mathbf{f}_L$ and global features $\mathbf{f}_G$ extracted respectively from the first and second branches are then combined. To this end, each of them is normalized by a batch normalization (BN) layer as follows:

$$\mathbf{f}_L^N = \text{BN}(\mathbf{f}_L) \quad (14)$$

$$\mathbf{f}_G^N = \text{BN}(\mathbf{f}_G) \quad (15)$$

Subsequently, their concatenation provides the fused final features:

$$\mathbf{f} = \text{cat}(\mathbf{f}_L^N; \mathbf{f}_G^N) \quad (16)$$

Here, according to Fig. 1(c), $\mathbf{f} \in \mathcal{R}^{n_b \times 1}$ is fed to two FC layers with n_b units and a regression layer to achieve the background estimate of the central pixel $\mathbf{y}$ denoted as $\hat{\mathbf{y}}$. The residual term is finally computed as anomaly score as follows:

$$r = |\ln(\|\mathbf{y} - \hat{\mathbf{y}}\|_1)| \quad (17)$$

where $|\cdot|$ denotes the absolute value of the input, $\ln$ is the natural logarithm and $\|\cdot\|_1$ denotes the L_1 -norm, measuring the aggregate spectral reconstruction discrepancy between the original pixel and its reconstructed background. The logarithmic transformation is subsequently applied to compress the dynamic range of the reconstruction residual and reduce the dominance of exceptionally large residuals. The absolute operator ensures that the resulting anomaly score is nonnegative. The proposed OCAG algorithm is given in Algorithm 1.

Algorithm 1. OCAG

Input: hyperspectral image, free parameters: the input patch size ($p \times p$), the number of morphological features (n_f), and the number of clusters (m)

Output: detection map

1. Generate the morphological opening profiles using (1).
2. Generate the morphological closing profiles using (2).
3. Consider a patch centered at the pixel under test in the hyperspectral image, and each of the corresponding opening and closing profiles.
4. Generate the input feature maps $\mathbf{F}_i \in \mathcal{R}^{p \times p \times d}$ from the input patch by applying the convolutional filters according to Fig. 1(a).

5. Generate the opening feature maps $F_o \in \mathcal{R}^{p \times p \times d}$ from the associated patch of the opening profiles by applying the convolutional filters according to Fig. 1(a).
6. Generate the closing feature maps $F_c \in \mathcal{R}^{p \times p \times d}$ from the associated patch of the closing profiles by applying the convolutional filters according to Fig. 1(a).
7. Compute the opening cross-attention CA_o using (6).
8. Compute the closing cross-attention CA_c using (10).
9. Fuse F_i , CA_o and CA_c features to provide F_{cat} according to (11).
10. Process F_{cat} using the convolutional and FC layers according to Fig. 1(a) to obtain the local features f_L .
11. Provide the feature matrix $X \in \mathcal{R}^{m \times n_b}$ according to (12).
12. Provide the adjacency matrix $A \in \mathcal{R}^{m \times m}$ according to (13).
13. Feed X and A to the graph convolutional network according to Fig. 1(b) to provide the global features f_G in output.
14. Normalize f_L in step 10 and f_G in step 13 using (14) and (15), respectively; then fuse them using (16) to provide the final features f .
15. Feed f to the shallow FC network according to Fig. 1(c) to achieve the background estimate of the central pixel y denoted as $\hat{y}$.

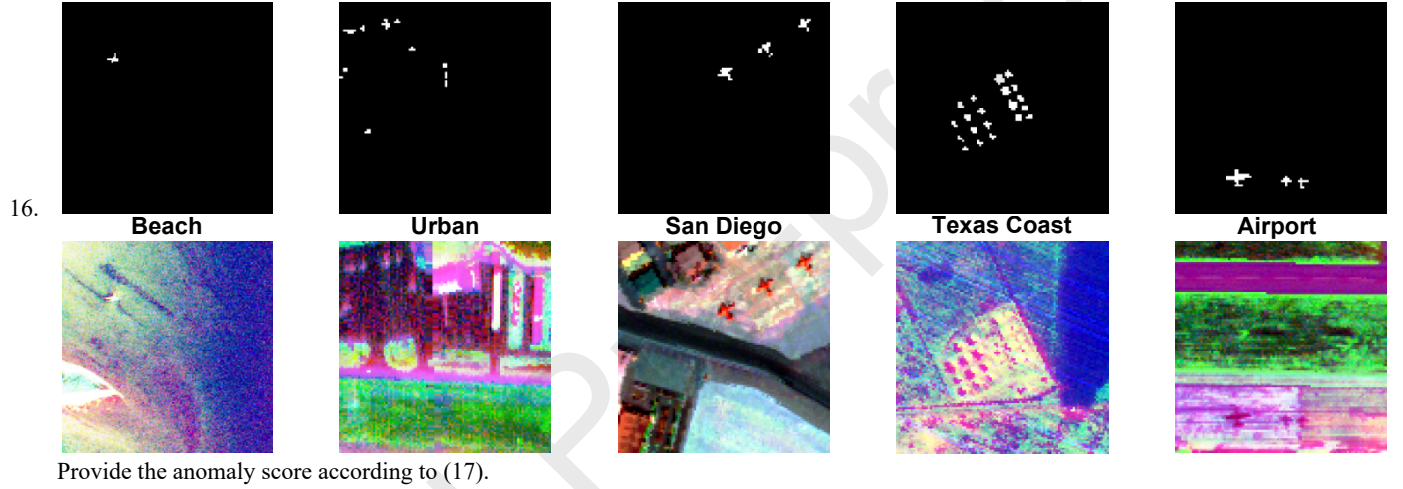


Fig. 2. Pseudo color-images and related anomaly maps in the first and second row, respectively.

Table 2. Detection results for different number of clusters (m).

Metric	2	5	10	15	20
AUC(P_D, P_F)	0.9859	0.9889	0.9874	0.9871	0.9831
AUC(P_D, τ)	0.8060	0.8194	0.7848	0.8079	0.7503
AUC(P_F, τ)	0.1270	0.1707	0.1674	0.1508	0.1643
AUC(TD)	1.7919	1.8083	1.7722	1.7950	1.7334
AUC(BS)	0.8589	0.8182	0.8200	0.8363	0.8188
Training Time (s)	529.41	558.23	721.31	764.16	834.98
Prediction time (s)	54.97	57.18	70.97	75.59	78.33
FLOPs (M)	0.32	0.46	0.69	0.93	1.16

Table 3. Detection results for different values of patch size (p).

Metric	5	7	9	11
AUC(PD,PF)	0.9856	0.9874	0.9871	0.9830
AUC(PD, τ)	0.7538	0.7848	0.7796	0.7059

AUC(PF, τ)	0.1679	0.1674	0.1721	0.1404
AUC(TD)	1.7394	1.7722	1.7667	1.6889
AUC(BS)	0.8177	0.8200	0.8150	0.8426
Training Time (s)	522.96	721.31	1127.17	1619.10
Prediction Time (s)	47.44	70.97	107.91	149.62
FLOPs (M)	0.69	0.69	0.69	0.69

Table 4. Detection results for different number of morphological filters (n_f).

Metric	10	25	50	100
AUC(PD,PF)	0.9876	0.9898	0.9874	0.9872
AUC(PD, τ)	0.7724	0.7942	0.7848	0.7578
AUC(PF, τ)	0.1429	0.1530	0.1674	0.1270
AUC(TD)	1.7600	1.7840	1.7722	1.7450
AUC(BS)	0.8447	0.8368	0.8200	0.8602
Training Time (s)	643.43	655.42	721.31	1108.48
Prediction Time (s)	59.78	65.42	70.97	85.66
FLOPs (M)	0.69	0.69	0.69	0.69

4. Experiments

4.1. Datasets and evaluation measures

Five benchmark hyperspectral images acquired in the USA by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) are used in the experiments reported in this paper. The Beach image having a size of 150×150 pixels was acquired over Cat Island, and has 188 spectral bands with a spatial resolution of 17.2m. The Urban image with spatial dimensions of 100×100 was acquired over Gainesville, and contains 191 spectral bands with a spatial resolution of 3.5m. The San Diego image has a size of 100×100 pixels and 189 spectral bands, with a spatial resolution of 3.5m. The Texas Coast image has a size of 100×100 pixels, and consists of 207 spectral bands with spatial resolution of 17.2m. Finally, the Airport image, with spatial dimensions of 100×100 pixels, was acquired over Gulfport, and contains 191 spectral bands with a spatial resolution of 3.4m. The pseudo-color images of the hyperspectral datasets and the corresponding anomaly maps are shown in Fig. 2.

For each dataset, 25% of the total pixels are randomly selected to train the network in an unsupervised manner as a regression task, while 25% of the data is selected for validation and the remainder is used as test. This implies that the pixel labels are considered unknown. In each training sample, the patch considered around the training sample, its opening and closing patches, and its associated feature matrix and adjacency matrix, are fed as input to the network, and the feature vector of the central pixel is used as the reconstructed background in the output.

To assess the detection accuracy of the anomaly detectors, we employ a 3D ROC curve, which is a function of detection probability P_D , false alarm probability P_F , and a threshold τ .

To comprehensively evaluate hyperspectral anomaly detection performance, we adopt the 3D ROC analysis. Three component AUC values are first computed from the corresponding 2D ROC curves, namely $AUC(P_D, P_F)$, $AUC(P_D, \tau)$, and $AUC(P_F, \tau)$. Following the standard evaluation protocol, the target detectability (TD) and background suppressibility (BS) are quantified by

$$AUC(TD) = AUC(P_D, P_F) + AUC(P_D, \tau) \quad (18)$$

$$AUC(BS) = AUC(P_D, P_F) - AUC(P_F, \tau) \quad (19)$$

A larger AUC(TD) (range: [0,2]) indicates stronger target detectability, whereas a larger AUC(BS) (range: [-1,1]) indicates better background suppression.

The experiments are implemented in MATLAB R2022b on a workstation with a Core i5 processor and 32GB RAM. The networks are trained by an Adam optimizer, with a learning rate of 0.001, a mini-batch size of 128, and 50 epochs. The training and validation RMSE and loss curves of the OCAG network during 20 epochs are shown in Fig. 3. Since the same network architecture, loss function, optimizer, and training settings are used for all datasets, the convergence behavior is illustrated only for the representative San Diego dataset. Similar convergence trends were observed for the remaining datasets.

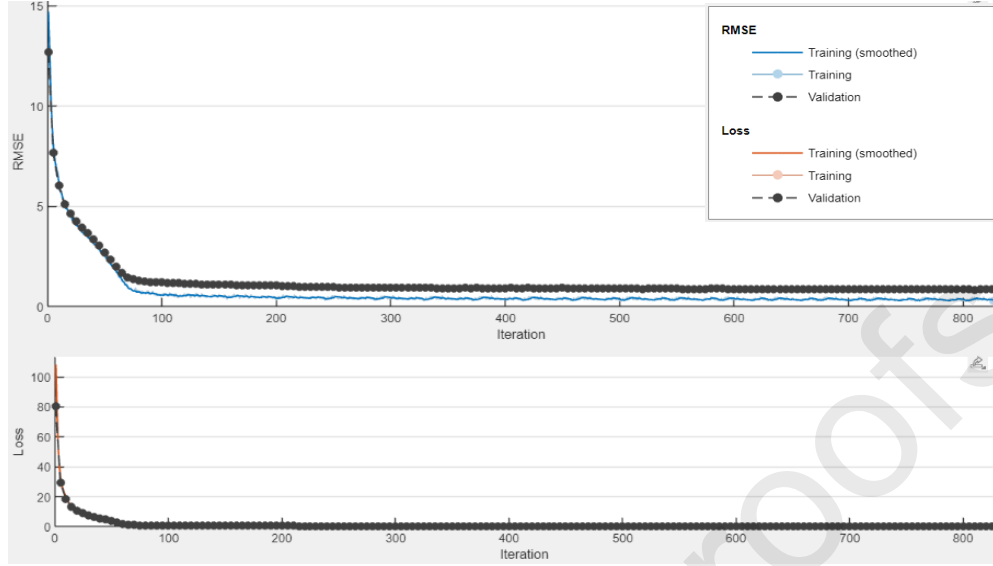


Fig. 3. The convergence behavior of the OCAG network.

4.2. Parameter settings

There are three free parameters in the proposed anomaly detector: the input patch size ($p \times p$), the number of morphological features (n_f), i.e., depth of opening and closing profiles, and the number of clusters (m). The impact of these parameters is discussed in this section. Tables 2-4 show the anomaly detection results (area under curve (AUC) values and computation time) obtained by varying each parameter while keeping the other two constant. The best value in each row of the tables is indicated in bold. The specific results are related to the San Diego dataset, which is considered representative here. Values of $p = 7$, $m = 10$ and $n_f = 50$ are used as fixed parameters, chosen empirically as the best compromise between detection accuracy and computational burden.

In Table 2 m is set to 2, 5, 10, 15, or 20 clusters. Usually, in an anomaly detection task, data are grouped into two classes: a majority background class, and a minority anomaly class. Here, the background is ideally a homogeneous region with small variations, but in complex images it may be formed by multiple clusters, resulting in a multi-modal distribution. The value $m = 2$ is considered to represent a very simple case of two clusters related to anomalies and background. By considering the presence of different materials and increasing to $m = 5$, the $AUC(P_D, P_F)$ improves, while results degrade for values of m larger than 5. Setting $m = 5$ clusters leads also to the highest $AUC(P_D, \tau)$ values, while the lowest $AUC(P_F, \tau)$ is obtained with $m = 2$, since representing the background with just 2 clusters cannot capture its variability. Increasing the number of clusters from $m = 2$ to $m = 5$ improves the target detectability from $AUC(TD) = 1.7919$ to 1.8083, which is the highest value among all tested settings. Although $m = 2$ achieves the highest background suppressibility ($AUC(BS) = 0.8589$), its lower target detectability indicates that an overly simplified background model cannot adequately represent the variability of complex scenes. Increasing m beyond 5 does not further improve the detection performance while introducing additional computational cost. Therefore, $m = 5$ is selected for the subsequent experiments as the best compromise between target detectability, background suppressibility, and computational efficiency.

In Table 3, we set $m = 10$ and $n_f = 50$, and the patch size ranges from 5×5 to 11×11 , with a step size of 2. The two best AUC values are obtained by setting $p = 7$. Here, increasing the patch size from 5×5 to 7×7 results in an improvement of $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$, which degrade after further increasing the patch size. This is expected, as contextual features from the local neighborhood can be derived using a minimum patch size, but may become redundant or less relevant to the central pixel in larger patches. The results show that increasing the patch size from (5×5) to (7×7) improves the target detectability, reaching the highest value of $AUC(TD) = 1.7722$. The best AUC (P_F, τ) value is instead obtained using $p = 11$. Nevertheless, a 11×11 patch leads to the lowest $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ values. Although the largest patch (11×11) achieves the highest background suppressibility

(AUC(BS)=0.8426), its target detectability decreases considerably and both the training and prediction times increase substantially. Consequently, a patch size of (7×7) is adopted.

In Table 4, n_f is set to 10, 25, 50, and 100 to assess the impact of the number of morphological feature maps. By increasing n_f from 10 to 25, $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ improve, and they slightly degrade for larger values. Moreover, increasing n_f from 10 to 25 improves the target detectability, yielding the highest value of $AUC(TD)=1.7840$. Although the lowest $AUC(P_F, \tau)$ and the highest background suppressibility ($AUC(BS)=0.8602$) are obtained by setting $n_f = 100$, the $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ values are at their lowest here. It also incurs the highest computational cost without improving target detectability. The configuration $n_f=50$ achieves target detectability and background suppressibility that remain very close to the best-performing settings while providing richer multiscale morphological representations than $n_f=25$. Since the proposed detector is intended to operate on hyperspectral images with different scene complexities, $n_f=50$ is selected as a more robust configuration for the subsequent experiments.

Overall, increasing p , m , or n_f leads to higher computational cost in terms of both training and prediction time. Based on the above analysis, the final parameter configuration adopted throughout the remaining experiments is $p=7$, $m=5$, and $n_f=50$. It should be noted that the parameter sensitivity study intentionally fixes two parameters while varying the third to independently evaluate its influence. Consequently, Table 3 uses $m=10$ and $n_f=50$ when analyzing the effect of the patch size, whereas the final experimental configuration is determined after jointly considering the results of all three parameter studies.

4.3. Ablation study

The contribution of the different modules of the proposed method is assessed in this section. The following settings are compared:

- Residual convolutional network (RCN): three convolutional layers ($32\text{conv}3 \times 3$) followed by ReLU and residual connections, as used in the local branch, are applied to the input patch for spectral and contextual feature extraction, and the final part of the network is composed by two FC layers with n_b units and a regression layer.
- Opening based attention (OA): the opening cross-attention (CA_o), concatenated with the processed input patch (F_i), is used in the local branch for local feature extraction, with the extracted features fed to the output block consisting of two FC and one regression layer.
- Closing based attention (CA): the closing cross-attention (CA_c), concatenated with the processed input patch (F_i), is used as local features and fed into the output block.
- Opening and closing based attention (OCA): this case uses the complete local branch, which is a combination of OA and CA.
- Graph convolutional network (GCN): this case only uses the global branch, where X and A inputs are fed into the graph convolutional network for global feature extraction, and the extracted features are given to the output block for background estimation.
- OA and GCN (OAG): a combination of OA and GCN, where the opening-based cross-attention in the local branch is combined with the graph features in the global branch and given to the output block.
- CA and GCN (CAG): a combination of CA and GCN, where the closing-based cross-attention features from the local branch are combined with the graph features from the global branch, and used for regression in the output.
- Opening and closing based attention integrated with graph features (OCAG): this is the complete workflow as shown in Fig. 1, where both opening cross-attention and closing cross-attention in the local branch, and also the graph features obtained from GCN, are combined for background estimation.

Specifically:

- RCN evaluates only the residual convolutional feature extraction without any morphological or graph-based modeling.
- OA introduces opening-based cross-attention to enhance anomaly-sensitive local features.
- CA replaces OA with closing-based cross-attention to emphasize background-dominant representations.
- OCA combines both opening and closing cross-attention, forming the complete local branch.
- GCN evaluates only the global graph-based representation without local morphological information.
- OAG and CAG integrate local (OA or CA) and global (GCN) branches.

- OCAG represents the full proposed framework, combining both opening/closing local attention with global graph features.

Detection results of the ablation study for all datasets are reported in Table 5. For each dataset, in each row, the **first rank** is in bold, the second rank is underlined, and the *third rank* is in italic. The associated ROC curves and detection maps beside the anomaly maps used as references are shown in Fig. 4 and Fig. 5, respectively.

For the Beach dataset, OA and CA achieve the highest $AUC(P_D, P_F)$, indicating that local morphological attention alone is highly effective in separating anomalies from background in relatively structured scenes. From the $AUC(P_D, \tau)$ point of view, GCN, OCAG and CAG rank first to third, respectively, and the best $AUC(P_F, \tau)$ value is associated with OCAG. When evaluated using $AUC(TD)$, OCAG and CAG achieve the best performance, demonstrating that integrating closing-based features with global graph information significantly improves overall detectability. From the background suppression perspective, OCAG achieves the best $AUC(BS)$, confirming that the full integration of opening, closing, and graph representations provides the most robust modeling of background variability. In contrast, RCN consistently yields the lowest performance across most of metrics, confirming that convolutional features alone are insufficient for complex hyperspectral anomaly detection.

For the Urban dataset, CAG ranks first according to $AUC(P_D, P_F)$, and second in terms of $AUC(P_D, \tau)$, and has therefore the overall best performance. OAG ranks first according to $AUC(P_D, \tau)$ and second in terms of $AUC(P_D, P_F)$. In terms of both measures, OCAG ranks third. The best $AUC(P_F, \tau)$ is achieved by OCA, while the worst one is obtained by RCN. In the Urban scene, OAG achieves the highest $AUC(TD)$, indicating that combining opening-based local features with global graph modeling is particularly effective in highly heterogeneous environments. Meanwhile, OCA achieves the best $AUC(BS)$, showing that integration of opening and closing-based modeling is more effective for suppressing urban background clutter. OCAG achieves competitive but slightly lower $AUC(TD)$ and $AUC(BS)$, indicating that while full fusion improves stability, some redundancy between local and global representations may slightly reduce peak performance in this dataset. RCN again provides the weakest performance, confirming the necessity of both morphological and graph-based modeling.

For the San Diego dataset, CAG achieves the best performance in terms of $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$, while CA ranks second. The best $AUC(P_F, \tau)$ value is obtained by GCN. CAG achieves the highest $AUC(TD)$, while GCN achieves the best $AUC(BS)$. This suggests that global graph modeling is particularly effective in suppressing structured background noise in this scene. OCAG remains competitive in both metrics, demonstrating stable performance when both local and global information are jointly exploited.

For the Texas Coast dataset, OA, CA and their combination, i.e., OCA achieve the same $AUC(P_D, P_F)$ values rank first, while GCN ranks second. In terms of $AUC(P_D, \tau)$ and $AUC(P_F, \tau)$, OAG and OCAG rank first, respectively, while in terms of both measures CAG ranks second. OAG achieves the highest $AUC(TD)$, indicating that opening-based local attention combined with global modeling is particularly effective in highly heterogeneous coastal environments. In contrast, OCAG achieves the best $AUC(BS)$, confirming that the full model is more effective in suppressing complex background distributions.

According to the obtained results for the Airport dataset, the best $AUC(P_D, P_F)$ values are obtained by OCAG and CAG, respectively. The best $AUC(P_D, \tau)$ values are obtained by OAG and CAG, respectively. Finally, the best $AUC(P_F, \tau)$ values are achieved by OCA and OCAG, respectively. OAG achieves the highest $AUC(TD)$, while OCAG achieves the best $AUC(BS)$. This indicates that opening-based local features combined with global context are particularly effective for structured man-made environments, whereas full OCAG fusion enhances background suppression.

Overall, the ablation results demonstrate that:

- Opening-based attention (OA) enhances anomaly sensitivity.
- Closing-based attention (CA) improves background modeling.
- GCN provides strong global contextual regularization.
- OAG and CAG represent effective hybrid configurations.
- OCAG consistently provides the most balanced trade-off between $AUC(TD)$ and $AUC(BS)$ across datasets, even if it is not always the best-performing method in each individual metric.

This confirms that hyperspectral anomaly detection benefits from complementary integration of local morphological cues and global graph structure, rather than relying on a single representation.

To further clarify the independent contribution of each key component, we explicitly analyze the role of the cross-attention mechanism, morphological profiles, and feature fusion strategy within the proposed framework. The ablation variants reported in Table 5 correspond directly to these components, enabling a controlled evaluation of their impact on detection performance.

The results show that the morphological profile representation (OA and CA) plays a fundamental role in improving local discriminative ability. Both opening- and closing-based representations consistently outperform the baseline RCN across all datasets, confirming that incorporating morphological filtering significantly enhances structural characterization of background and anomaly regions.

In addition, the cross-attention mechanism further improves feature selectivity by adaptively weighting informative spatial-spectral responses. This is evident from the consistent improvement of OCA over OA and CA individually, indicating that combining opening and closing responses within an attention-driven framework provides more robust local feature modeling than using either component alone.

Furthermore, the integration of the global graph convolutional branch (GCN) introduces complementary contextual information that is not captured by local morphological descriptors. The performance gains observed in OAG and CAG compared to their purely local counterparts demonstrate that graph-based modeling effectively enhances global consistency and improves anomaly-background separability.

Finally, the full feature fusion strategy in OCAG, which combines opening, closing, and graph-based representations, provides a unified multi-level representation. Although OCAG does not always outperform all simplified variants on every dataset or metric, it consistently achieves strong and stable performance, confirming that the fusion strategy enhances robustness across heterogeneous scenes.

The ablation results indicate that the complete OCAG model is not always superior to its simplified variants (OA, CA, OAG, or CAG) in terms of individual metrics on all datasets. This behavior can be explained by the complementary and sometimes competing roles of local morphological features and global graph representations.

Specifically, configurations such as OA and CA focus on highly specialized representations: OA emphasizes anomaly-sensitive structures derived from opening profiles, while CA emphasizes background-dominant structures derived from closing profiles. In certain datasets with relatively homogeneous or well-structured backgrounds, these specialized representations may be sufficient to achieve strong performance in either AUC(TD) or AUC(BS), without requiring full fusion.

Similarly, CAG and OAG reduce model complexity by excluding one of the two local representations (opening or closing). This reduction can sometimes improve performance in specific datasets by limiting feature redundancy and avoiding interference between opening- and closing-based cues. In particular, CAG can perform competitively when closing-based information is more informative for background modeling, while OAG may be more effective in scenes where anomaly-related structures are better captured by opening profiles. Closing-based representations already capture strong background modeling capability in many scenes.

In contrast, the complete OCAG model integrates both opening and closing cross-attention together with graph-based global features, which provides a more expressive representation but may also introduce feature redundancy or partial over-smoothing in certain scenes, leading to slightly lower performance in specific metrics or datasets.

It is worth noting that the simplified CAG configuration achieves comparable or even superior performance to the complete OCAG model on several datasets and evaluation metrics. This observation suggests that the additional opening branch is not uniformly beneficial across all hyperspectral scenes. The effectiveness of the opening profile depends on the structural characteristics of the anomalies and the surrounding background. Morphological opening primarily emphasizes small bright structures by suppressing fine background details, whereas morphological closing better preserves homogeneous background regions and fills small gaps. For datasets in which background modeling is the dominant factor for anomaly separation, the closing-based representation may provide sufficient discriminative information, explaining the strong performance of CAG. Nevertheless, the opening branch provides complementary information that becomes valuable in scenes containing small, isolated, or structurally distinctive anomalies. Rather than consistently improving every evaluation metric on every dataset, its main contribution is to increase the robustness and generalization capability of the detector across diverse imaging conditions. This complementary behavior is reflected by the stable performance of OCAG across all benchmark datasets, even though CAG may achieve the best result on individual datasets or specific evaluation measures.

However, OCAG consistently maintains stable and competitive performance across all datasets and both evaluation criteria (AUC(TD) and AUC(BS)), without significant degradation in any scenario. While OAG or CAG may achieve the best performance in specific datasets or metrics, their performance is less consistent across all test conditions. Therefore, rather than outperforming all variants in every case, the proposed OCAG framework is best interpreted as a robust unified model that balances opening-based anomaly sensitivity, closing-based background modeling, and global graph structure representation.

The number of learnable parameters is reported in Table 6. As expected, GCN and RCN exhibit the lowest complexity, while OCAG has the highest number of parameters due to its dual-branch architecture. However, all proposed variants remain significantly more lightweight than typical deep learning-based hyperspectral models.

Training and prediction times are reported in Table 7. GCN is consistently the fastest method due to its simpler architecture, while OCAG is computationally more expensive because of cross-attention, morphological profile

Table 5. Detection results of the ablation study for different datasets.

Beach								
	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
$AUC(P_{D_i}, P_F)$	0.9581	0.9942	<u>0.9941</u>	0.9926	0.9590	0.9926	0.9926	<u>0.9935</u>
$AUC(P_{D_i}, \tau)$	0.6979	0.7918	0.7540	0.7624	0.8602	0.8183	<u>0.8484</u>	<u>0.8492</u>
$AUC(P_F, \tau)$	0.1897	0.1308	0.1278	<u>0.1220</u>	0.4359	<u>0.1221</u>	0.1261	0.1182
AUC(TD)	1.6560	1.7860	1.7481	1.7550	<u>1.8192</u>	1.8109	1.8410	1.8427
AUC(BS)	0.7684	0.8634	0.8663	0.8706	0.5231	<u>0.8705</u>	0.8665	0.8753
Urban								
	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
$AUC(P_{D_i}, P_F)$	0.8493	0.9460	0.9568	0.9491	0.9610	<u>0.9826</u>	0.9830	<u>0.9779</u>
$AUC(P_{D_i}, \tau)$	0.4525	0.5768	0.6515	0.5729	0.6649	0.7841	0.7354	<u>0.6845</u>
$AUC(P_F, \tau)$	0.3165	<u>0.2645</u>	0.2749	0.2527	<u>0.2712</u>	0.2978	0.2995	0.2876
AUC(TD)	1.3018	1.5228	1.6083	1.5220	1.6259	1.7667	1.7184	<u>1.6624</u>
AUC(BS)	0.5328	0.6815	0.6819	0.6964	<u>0.6898</u>	0.6848	0.6835	<u>0.6903</u>
San Diego								
	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
$AUC(P_{D_i}, P_F)$	0.9852	0.9898	<u>0.9908</u>	<u>0.9902</u>	0.9844	0.9894	0.9918	0.9889
$AUC(P_{D_i}, \tau)$	0.7988	0.7916	<u>0.8325</u>	0.8101	0.7757	0.7702	0.8423	<u>0.8194</u>
$AUC(P_F, \tau)$	0.1753	0.1701	0.1733	0.1698	0.1203	<u>0.1687</u>	<u>0.1625</u>	0.1707
AUC(TD)	1.7840	1.7814	<u>1.8233</u>	1.8003	1.7601	1.7596	1.8341	<u>1.8083</u>
AUC(BS)	0.8099	0.8197	0.8175	0.8204	0.8641	<u>0.8207</u>	<u>0.8293</u>	0.8182
Texas Coast								
	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
$AUC(P_{D_i}, P_F)$	0.8350	0.9952	0.9952	0.9952	<u>0.9908</u>	0.9826	<u>0.9857</u>	0.9844
$AUC(P_{D_i}, \tau)$	0.6852	0.7860	0.7991	0.7967	<u>0.8650</u>	0.8991	<u>0.8697</u>	0.7941
$AUC(P_F, \tau)$	0.3615	0.3501	0.3912	0.4077	<u>0.3481</u>	0.3569	<u>0.3473</u>	0.3293
AUC(TD)	1.5202	1.7812	1.7943	1.7919	1.8558	1.8817	<u>1.8554</u>	1.7785
AUC(BS)	0.4735	<u>0.6451</u>	0.6040	0.5875	<u>0.6427</u>	0.6257	0.6384	0.6551
Airport								
	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
$AUC(P_{D_i}, P_F)$	0.9833	0.9676	0.9844	<u>0.9884</u>	0.9636	0.9873	<u>0.9905</u>	0.9930
$AUC(P_{D_i}, \tau)$	0.7146	0.6378	<u>0.7648</u>	0.7599	0.6992	0.8574	<u>0.8463</u>	0.7602
$AUC(P_F, \tau)$	<u>0.3664</u>	0.4003	0.4530	0.3474	0.4616	0.5211	0.4768	<u>0.3488</u>
AUC(TD)	1.6979	1.6054	1.7492	1.7483	1.6628	1.8447	1.8368	<u>1.7532</u>
AUC(BS)	<u>0.6169</u>	0.5673	0.5314	<u>0.6410</u>	0.5020	0.4662	0.5137	0.6442

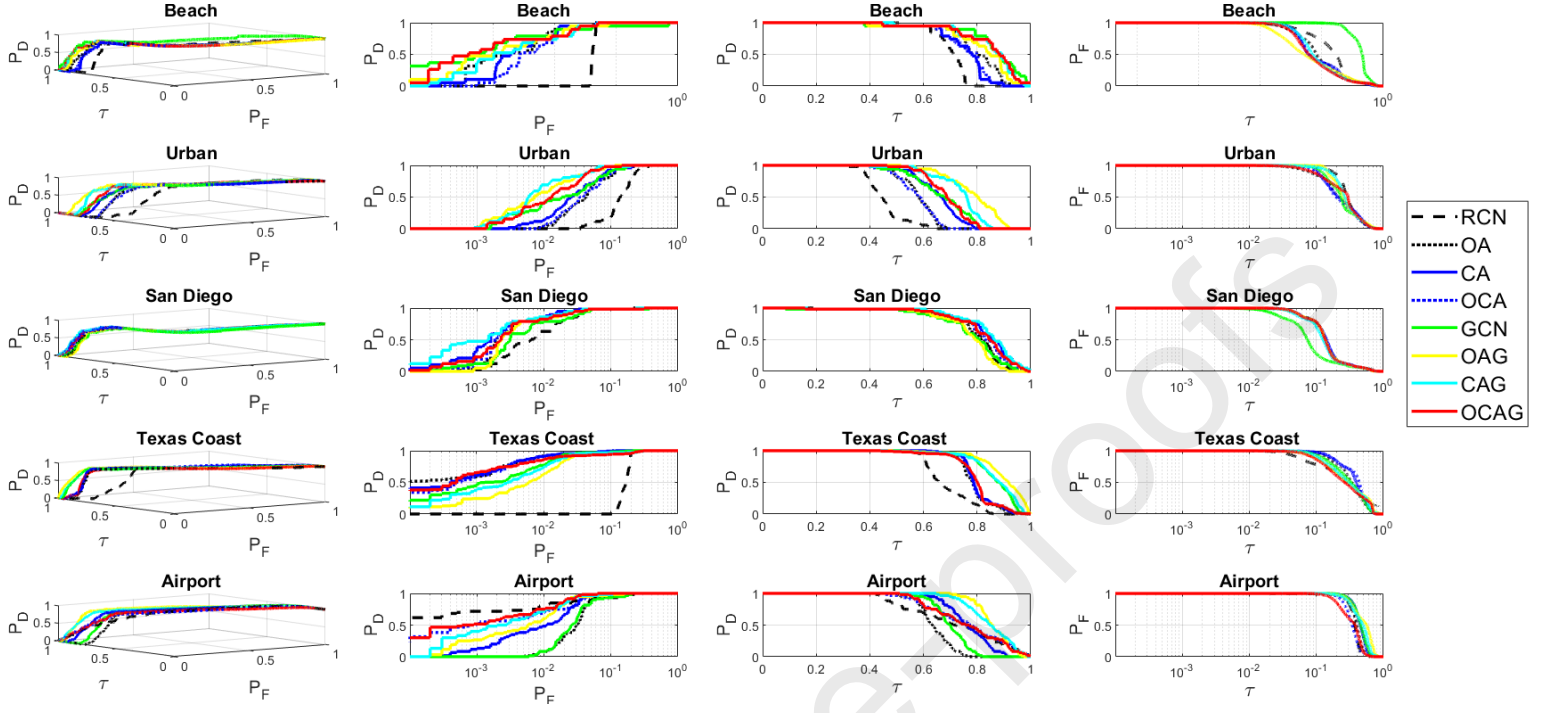


Fig. 4. ROC curves of the ablation study for different datasets.

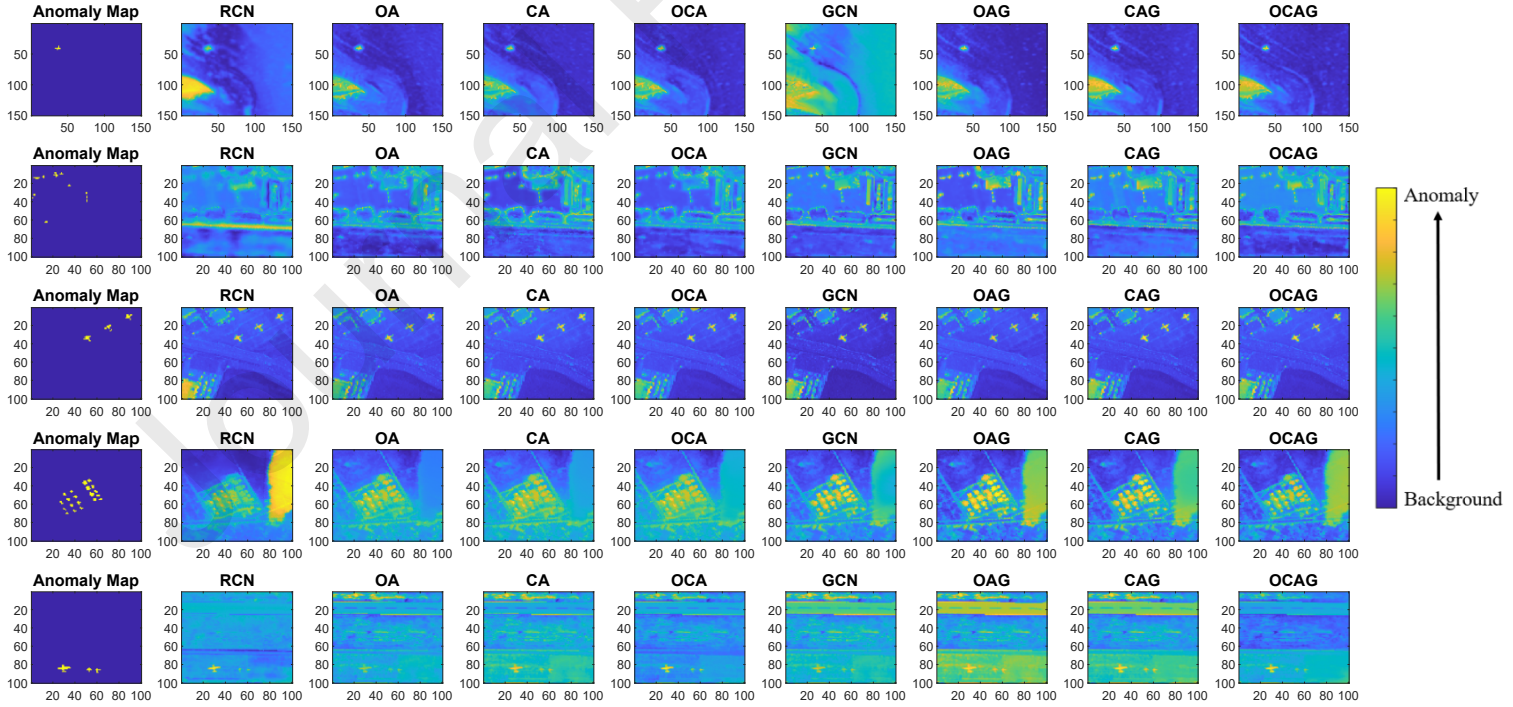


Fig. 5. Detection maps of the ablation study for different datasets.

Table 6. Number of learnable parameters for all different configurations.

Table 7. Training time and prediction time for all different configurations.

Dataset	RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
Beach	235.3k	474.8k	474.8k	571.1k	142k	581.1k	581.1k	677.5k
Urban	239.2k	480.7k	480.7k	577.1k	147.1k	591k	591k	687.5k
Sandiego	236.6k	443.2k	443.2k	506k	144.1k	551.2k	551.2k	614k
Texas Coast	260.3k	511.9k	511.9k	608.9k	172.7k	641.4k	641.4k	738.4k
Airport	239.2k	480.7k	480.7k	577.1k	147.1k	591k	591k	687.5k

extraction, and graph construction. Nevertheless, OCAG still operates within a practical runtime (on the order of minutes per dataset), making it applicable for real-world hyperspectral anomaly detection scenarios.

Dataset		RCN	OA	CA	OCA	GCN	OAG	CAG	OCAG
Beach	Training time (s)	1336.84	1900.93	2140.97	3209.57	349.40	2597.57	2593.33	3301.09
	Prediction time (s)	83.99	105.57	132.40	128.69	19.77	120.62	116.25	133.05
Urban	Training time (s)	327.34	431.43	486.77	635.36	94.89	542.76	556.07	711.91
	Prediction time (s)	39.06	51.81	52.17	61.07	10.58	53.78	57.72	63.71
Sandiego	Training time (s)	323.94	353.35	413.86	483.87	88.42	483.66	627.43	558.23
	Prediction time (s)	41.21	47.77	48.55	55.69	10.63	54.09	52.76	57.18
Texas Coast	Training time (s)	318.01	436.68	497.35	611.61	91.91	563.31	570.76	849.77
	Prediction time (s)	43.35	54.69	55.96	65.60	10.22	60.04	61.39	67.79
Airport	Training time (s)	366.15	412.58	471.21	591.51	88.74	537.74	540.47	631.61
	Prediction time (s)	41.58	53.67	54.02	63.36	10.44	58.59	59.05	66.25

4.4. Comparison with other anomaly detectors

The proposed methods (OCAG and CAG) are compared with state-of-the-art anomaly detectors in this section, namely RX (Reed and Yu, 1990), Kernel RX (Kwon and Nasrabadi, 2005), 2S-GLRT (Liu et al., 2022), KIFD (Li et al., 2020), CRD (Li and Du, 2015), LSUNRSORAD (Tan et al., 2019), LSAD-CR-IDW (Tan et al., 2019), and CRDBPSW (Hou et al., 2022). In addition to the AUC values, the prediction time required to produce the full detection map is reported. The comparison results across all datasets are reported in Table 8, and the associated ROC curves are shown in Fig. 6. The detection maps are shown in Fig. 7, and the corresponding binary detection maps obtained using a confidence level of 99% in Fig. 8. For the proposed framework, two variants are evaluated: CAG (closing + graph features) and OCAG (opening + closing + graph features). Since different datasets exhibit different background characteristics (e.g., homogeneous vs. highly structured scenes), we follow a dataset-adaptive selection strategy consistent with validation performance. Specifically, for each dataset, we select the variant (CAG or OCAG) that achieves the best performance on a validation subset and report it in Table 8. This ensures that the comparison reflects the best achievable performance of the proposed framework while maintaining a consistent evaluation protocol across datasets and baseline methods. Although different configurations of the proposed model are reported across datasets, this does not introduce unfair advantage, since both variants share the same backbone design, training strategy, and parameter settings, and differ only in the inclusion of opening-based attention. Moreover, the comparison remains fair because all competing methods are evaluated under fixed configurations, whereas the proposed method reports the best-performing variant among two closely related ablations of the same architecture, rather than introducing additional external information or training data.

For the beach dataset, considering $AUC(P_D, P_F)$ values, CRDBPSW has the best performance, with OCAG closely following. In terms of $AUC(P_D, \tau)$, OCAG and KIFD outperform other methods by a far margin. In terms of $AUC(P_F, \tau)$, RX, CRDBPSW and 2S-GLRT yield the best results. Note that the local RX is implemented here. In terms of running time, CRDBPSW is the fastest method with a significant margin from its competitors. In terms of $AUC(TD)$, OCAG achieves the most balanced performance among all methods. The higher $AUC(TD)$ confirms that OCAG provides strong anomaly enhancement by jointly exploiting local morphological attention and global structural information. However, methods such as CRDBPSW show lower $AUC(TD)$ despite high $AUC(P_D, P_F)$, indicating weaker joint separability of anomalies when threshold variation is considered.

According to the obtained results for the Urban dataset, the proposed CAG outperforms all other methods by a far margin in terms of $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ values. The lowest $AUC(P_F, \tau)$ value is obtained by RX, which however is the worst method in terms of the two other AUC measures, and in terms of speed, CRDBPSW is the fastest method. For the Urban dataset, CAG demonstrates strong performance in $AUC(TD)$ confirming that closing-based features combined with graph modeling are particularly effective in structured man-made environments and indicate enhanced sensitivity to compact urban anomalies. In contrast, 2S-GLRT and RX exhibit significantly lower $AUC(TD)$, reflecting limited capability to maintain detection consistency across thresholds.

For the San Diego dataset, in terms of $AUC(P_D, P_F)$, CRDBPSW and CAG rank respectively first and second. Considering $AUC(P_D, \tau)$, the proposed CAG provides the best results by a far margin, while RX, 2S-GLRT and CRDBPSW provide the lowest $AUC(P_F, \tau)$ values. KIFD and CRDBPSW are the fastest methods, while the proposed CAG is implemented in 52.76 seconds. Regarding $AUC(TD)$, CAG achieves the strongest anomaly detectability, indicating that closing-based representations are highly effective for capturing spatially coherent anomalies in this dataset. However, CRDBPSW shows stronger $AUC(BS)$, suggesting better suppression of background fluctuations. This indicates a trade-off between detectability and background suppression among different methods, whereas the proposed model maintains a relatively more balanced performance.

According to the detection results for the Texas Coast dataset, the proposed CAG ranks first in terms of the first two AUC measures. RX, CRDBPSW and 2S-GLRT provide the slowest $AUC(P_F, \tau)$; and CRDBPSW runs faster than all methods. The most accurate detection map with the lowest false alarm rate is achieved by CAG. In terms of $AUC(TD)$, CAG clearly outperforms all competing methods, confirming its strong ability to jointly enhance anomaly responses across different threshold levels. Compared with classical methods such as RX and 2S-GLRT, the improvement in $AUC(TD)$ demonstrates significantly better global separability between anomalies and background structures.

Table 8. Detection results of different methods for all datasets. For the proposed framework, the best-performing configuration between CAG and OCAG (selected based on validation performance on each dataset) is reported to ensure optimal but fair comparison under identical evaluation protocol.

Beach									
	RX	Kernel RX	2S-GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR-IDW	CRDBPSW	OCAG
$AUC(P_D, P_F)$	0.9543	0.9009	0.9663	0.9900	0.9761	0.9903	<i>0.9930</i>	0.9948	<u>0.9935</u>
$AUC(P_D, \tau)$	0.2778	<i>0.6650</i>	0.0628	<u>0.8381</u>	0.4402	0.5561	0.4323	0.2225	0.8492
$AUC(P_F, \tau)$	0.0036	0.1429	<i>0.0090</i>	0.0830	0.0370	0.0145	0.0165	<u>0.0037</u>	0.1182
AUC(TD)	1.2321	<i>1.5659</i>	1.0291	<u>1.8281</u>	1.4163	1.5464	1.4253	1.2173	1.8427
AUC(BS)	0.9507	0.7580	0.9573	<u>0.9070</u>	0.9391	<i>0.9758</i>	<u>0.9765</u>	0.9911	0.8753
Time (seconds)	<u>126.74</u>	3992.44	<i>127.04</i>	152.30	145.51	802.38	721.67	4.59	133.05
Urban									
	RX	Kernel RX	2S-GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR-IDW	CRDBPSW	CAG
$AUC(P_D, P_F)$	0.8780	0.8875	0.7506	0.9014	0.9110	<u>0.9615</u>	<i>0.9277</i>	0.8912	0.9830
$AUC(P_D, \tau)$	0.0166	<u>0.4247</u>	0.0758	0.3365	0.1153	<i>0.3763</i>	0.2963	0.1240	0.7354
$AUC(P_F, \tau)$	0.0021	0.1369	<u>0.0174</u>	0.1297	0.0333	0.0578	<u>0.0639</u>	<i>0.0235</i>	0.2995
AUC(TD)	0.8946	<i>1.3122</i>	0.8264	1.2379	1.0263	<u>1.3378</u>	1.2240	1.0152	1.7184
AUC(BS)	<i>0.8759</i>	0.7506	0.7332	0.7717	<u>0.8777</u>	0.9037	0.8638	0.8677	0.6835
Time (seconds)	21.52	137.49	18.15	28.66	18.42	<i>12.06</i>	<u>11.48</u>	2.11	63.71
San Diego									
	RX	Kernel RX	2S-GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR-IDW	CRDBPSW	CAG
$AUC(P_D, P_F)$	0.9854	0.9466	0.9828	0.9799	<i>0.9902</i>	0.9793	0.9539	0.9966	<u>0.9918</u>
$AUC(P_D, \tau)$	0.2849	<u>0.7187</u>	0.2030	0.5487	0.6058	0.4904	0.2138	<i>0.6316</i>	0.8423
$AUC(P_F, \tau)$	0.0017	0.2402	<u>0.0064</u>	0.1119	0.0548	0.0711	0.0534	<i>0.0076</i>	0.1625
AUC(TD)	1.2703	<u>1.6653</u>	1.1858	<i>1.5286</i>	1.5960	1.4697	1.1677	<i>1.6282</i>	1.8341
AUC(BS)	<u>0.9837</u>	0.7064	<i>0.9764</i>	0.8680	0.9354	0.9082	0.9005	0.9890	0.8293
Time (seconds)	67.92	334.88	70.40	36.24	<i>46.79</i>	342.34	203.04	<u>38.62</u>	52.76
Texas Coast									
	RX	Kernel RX	2S-GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR-IDW	CRDBPSW	CAG
$AUC(P_D, P_F)$	0.9000	0.9421	0.9081	0.8700	0.9505	<u>0.9721</u>	<i>0.9704</i>	0.9618	0.9857
$AUC(P_D, \tau)$	0.0255	<u>0.6721</u>	0.0852	<i>0.3530</i>	0.0600	0.0622	0.2182	0.0496	0.8697
$AUC(P_F, \tau)$	1.0496e-04	0.1878	<i>0.0019</i>	0.1252	0.0057	0.0020	0.0202	<u>0.0015</u>	0.3473
AUC(TD)	0.9255	1.6142	0.9933	<i>1.2230</i>	1.0105	1.0343	1.1886	1.0114	1.8554
AUC(BS)	0.8999	0.7543	0.9062	0.7448	0.9448	0.9701	<i>0.9502</i>	<u>0.9603</u>	0.6384
Time (seconds)	26.19	301.60	27.93	32.69	24.41	<i>12.66</i>	<u>12.13</u>	4.02	61.39
Airport									
	RX	Kernel RX	2S-GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR-IDW	CRDBPSW	OCAG
$AUC(P_D, P_F)$	0.9695	0.9492	0.8898	0.9798	<i>0.9821</i>	<u>0.9821</u>	<u>0.9927</u>	0.9791	0.9930
$AUC(P_D, \tau)$	0.1338	<i>0.4413</i>	0.2486	0.4207	0.1470	<u>0.4767</u>	0.3706	0.1540	0.7602
$AUC(P_F, \tau)$	0.0065	0.1085	<i>0.0189</i>	0.1343	0.0336	0.0545	0.0534	0.0136	0.3488
AUC(TD)	1.1033	1.3905	1.1384	<i>1.4005</i>	1.1291	<u>1.4588</u>	1.3633	1.1331	1.7532
AUC(BS)	<u>0.9630</u>	0.8407	0.8709	0.8455	<i>0.9485</i>	0.9276	0.9393	0.9655	0.6442
Time (seconds)	<i>42.60</i>	520.41	295.47	<u>30.70</u>	43.15	443.02	431.62	10.88	66.25

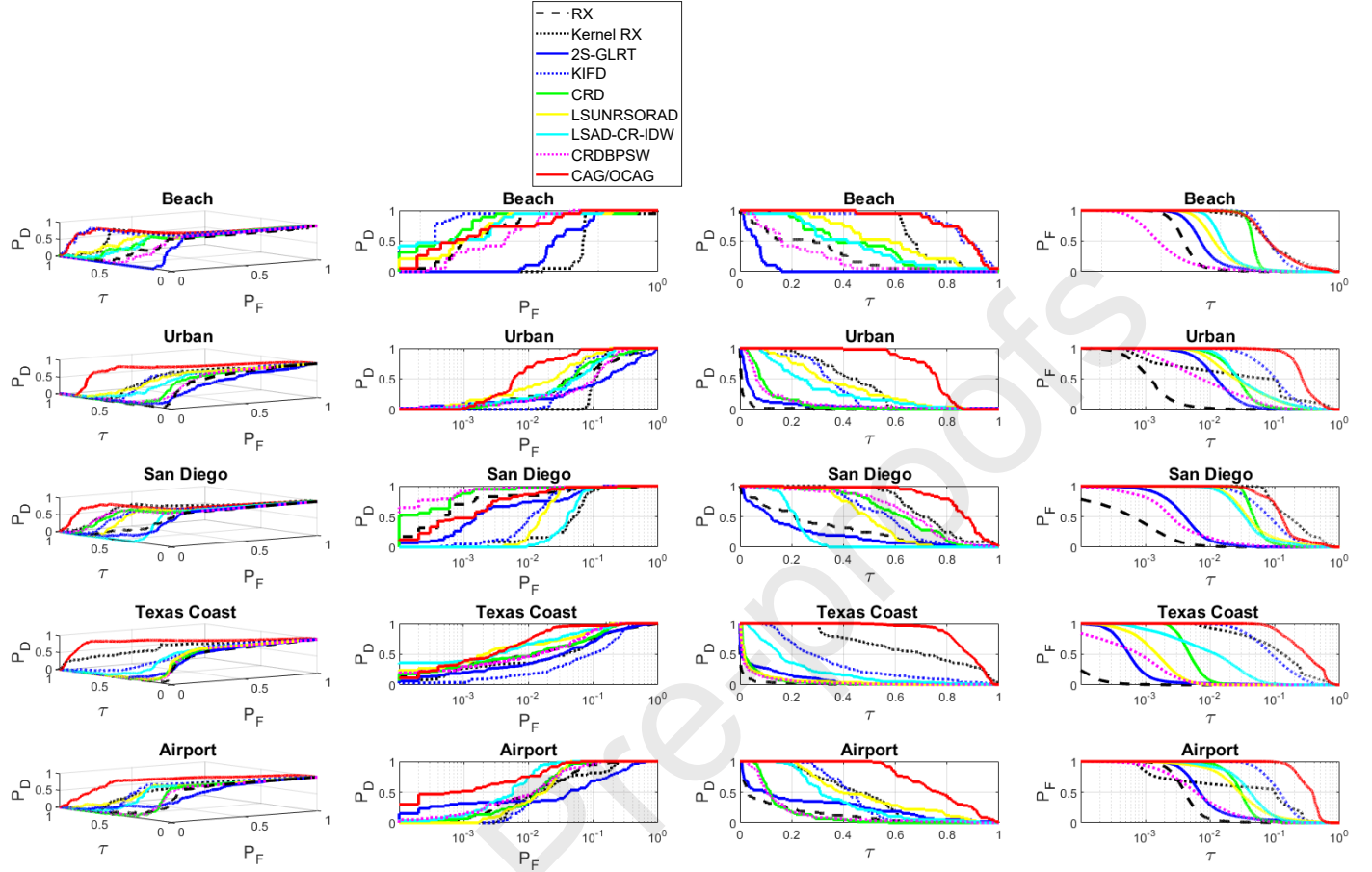


Fig. 6. ROC curves of different methods for all datasets.

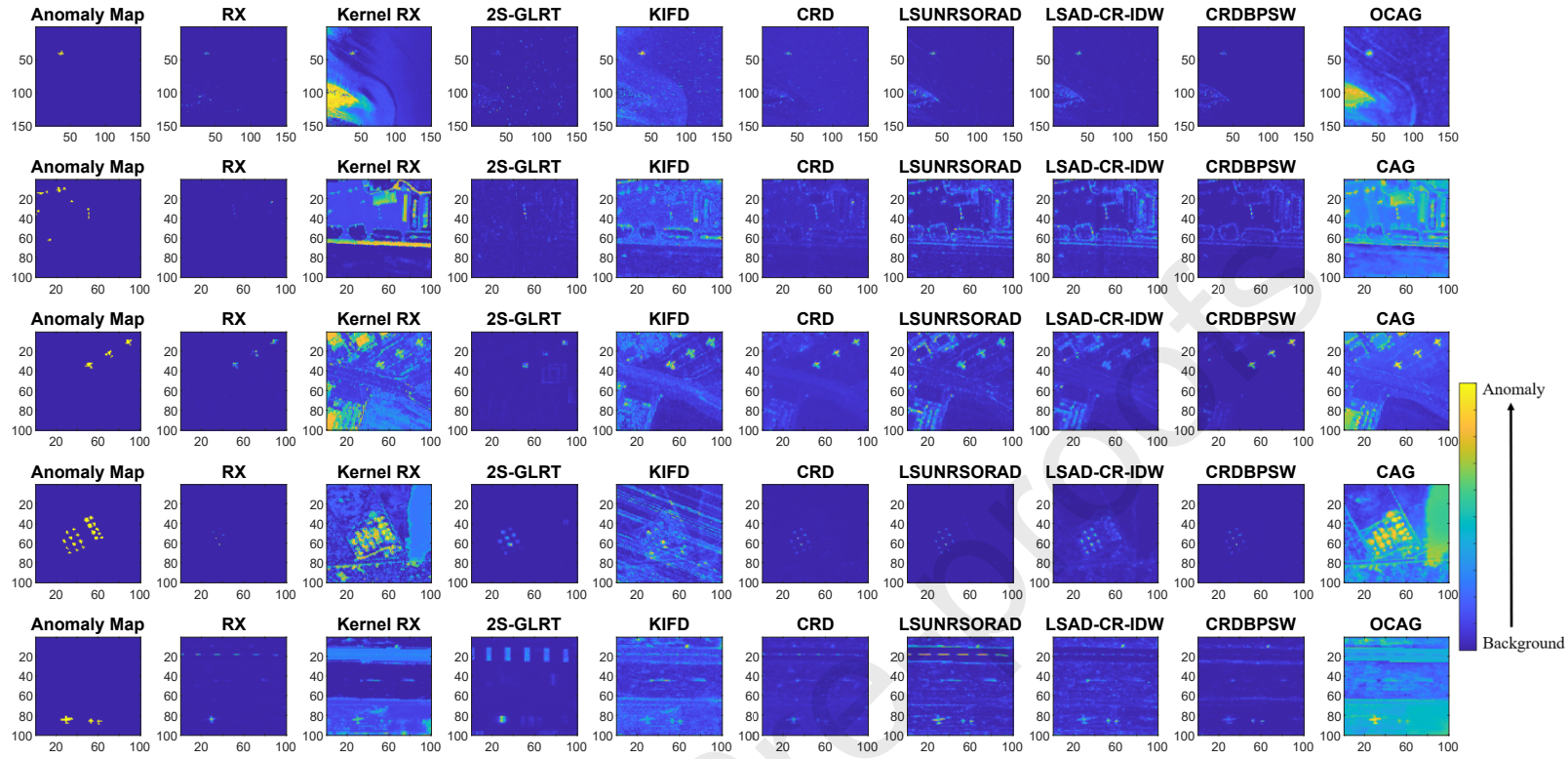


Fig. 7. Color detection maps of different methods for all datasets.

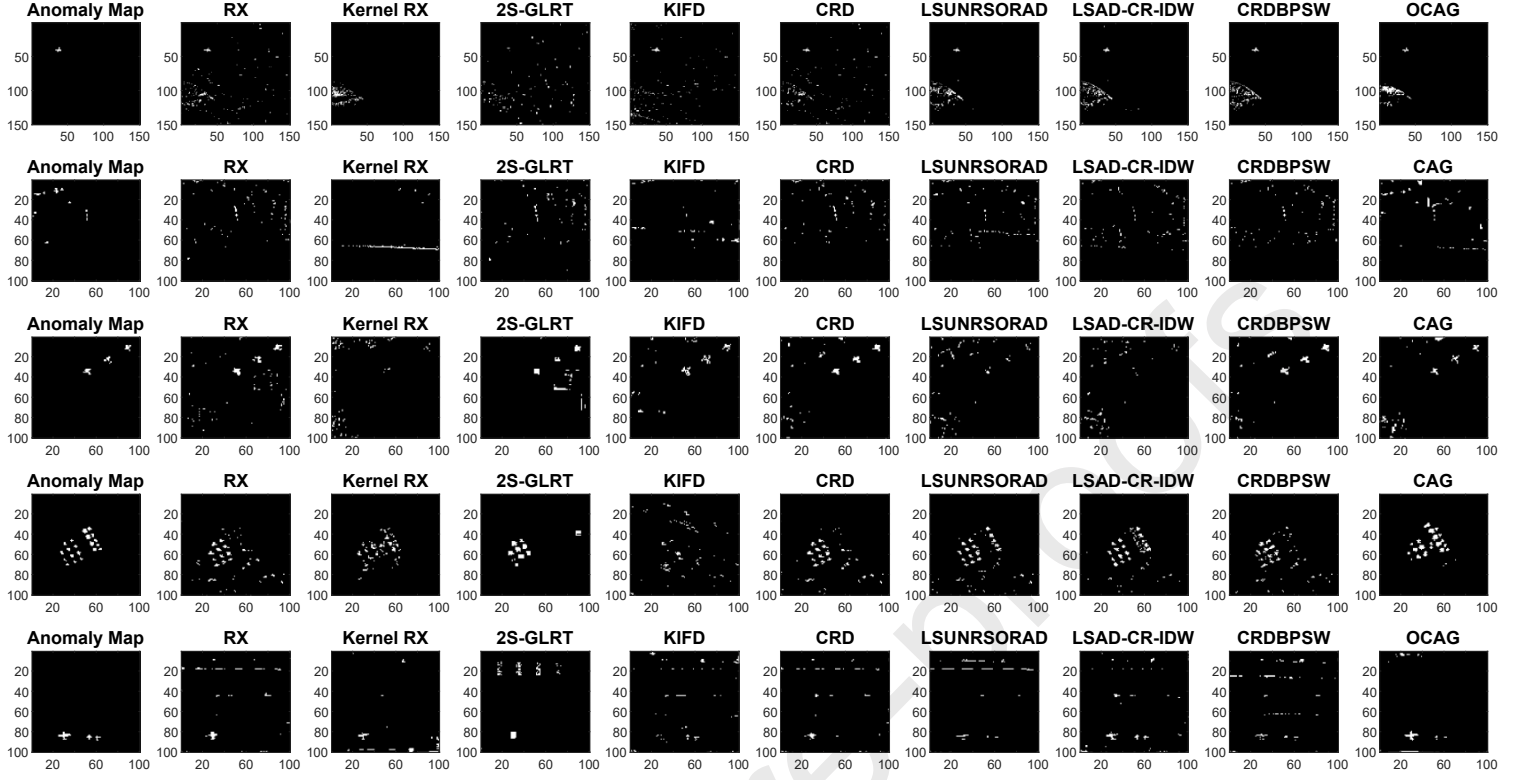


Fig. 8. Binary detection maps of different methods for all datasets.

Table 9. Average AUC values across all datasets.

	RX	Kernel RX	2S- GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR- IDW	CRDBPSW	Proposed (CAG/OCAG)
$AUC(P_D, P_F)$	0.9374	0.9253	0.8995	0.9442	0.9620	0.9771	0.9675	0.9647	0.9894
$AUC(P_D, \tau)$	0.1477	0.5844	0.1351	0.4994	0.2737	0.3923	0.3062	0.2363	0.8114
$AUC(P_F, \tau)$	0.0028	0.1633	0.0107	0.1168	0.0329	0.0400	0.0415	0.0100	0.2553

Table 10. Standard deviation of AUC values across all datasets.

	RX	Kernel RX	2S- GLRT	KIFD	CRD	LSUNRSORAD	LSAD-CR- IDW	CRDBPSW	Proposed (CAG/OCAG)
$AUC(P_D, P_F)$	0.0413	0.0258	0.0821	0.0489	0.0287	0.0097	0.0247	0.0388	0.0042
$AUC(P_D, \tau)$	0.1167	0.1251	0.0758	0.1851	0.2120	0.1748	0.0854	0.2053	0.0533
$AUC(P_F, \tau)$	0.0022	0.0461	0.0065	0.0185	0.0157	0.0268	0.0193	0.0079	0.0965

According to the detection results for the Airport dataset, OCAG ranks first in terms of the first two AUC measures, with a large margin in the case of $AUC(P_D, \tau)$. As in previous datasets, the best value of $AUC(P_F, \tau)$ is obtained by RX, CRDBPSW and 2S-GLRT, and the fastest method is CRDBPSW, which requires 11 seconds. Moreover, Kernel RX, LSUNRSORAD and LSAD-CR-IDW are the slowest methods, requiring 8.7 minutes, 7.4 minutes and 7.2 minutes, respectively. However, the proposed method can provide the detection map in about one minute. Although this is about 6 times slower than CRDBPSW, it is approximately 6.7 to 7.8 times faster than Kernel RX, LSUNRSORAD, and LSAD-CR-IDW. Visual inspection suggests that the most accurate detection map with the lowest false alarm rate is achieved by OCAG. For the Airport dataset, OCAG achieves the best performance in $AUC(TD)$, confirming improved detectability of small and sparse anomalies typical in airport scenes.

Overall, across all datasets, the proposed methods (OCAG and CAG) rank first or second in terms of $AUC(P_D, P_F)$, and with a significant margin rank first in terms of $AUC(P_D, \tau)$. Moreover, they provide the maps with the lowest false alarm rates, while requiring a tolerable amount of computational resources (from less than one minute to about two minutes per dataset). The proposed method demonstrates robust and competitive performance, achieving either the best or near-best results across multiple datasets according to $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ measures. Overall, across all datasets, $AUC(TD)$ consistently highlights the superiority of the proposed OCAG and CAG in enhancing anomaly separability. However, the incorporation of morphological attention features and graph-based global information also increases the score variation of some background pixels, resulting in higher $AUC(P_F, \tau)$ and consequently lower $AUC(BS)$ values. Therefore, the proposed method favors anomaly-background discrimination and detection sensitivity rather than maximizing background suppression alone.

The performance differences among the compared methods can be explained by their underlying modeling strategies. Statistical detectors such as RX, Kernel RX, and 2S-GLRT rely primarily on background distribution assumptions. Although these methods are computationally efficient and generally achieve low $AUC(P_F, \tau)$ values, they are less effective in heterogeneous scenes where the background cannot be accurately represented by a single statistical model. Consequently, their $AUC(P_D, \tau)$ values remain relatively low, indicating limited capability to progressively separate anomalies from complex backgrounds.

Representation-based methods, including CRD, LSUNRSORAD, LSAD-CR-IDW, and CRDBPSW, improve robustness by reconstructing the background from neighboring pixels rather than relying solely on statistical assumptions. Their performance is generally superior to classical statistical detectors because collaborative representation better captures local spectral characteristics. However, when neighboring regions contain spectrally heterogeneous materials or anomaly-like structures, the quality of the reconstructed background deteriorates, leading to reduced anomaly-background separability and limiting the achievable detection performance.

The proposed method differs from existing approaches by integrating two complementary sources of information. First, the opening and closing morphological profiles provide explicit prior cues related to anomaly structures and background structures, respectively, through the proposed cross-attention mechanism. Second, the graph convolutional branch introduces global contextual relationships that cannot be captured by local convolutions alone. The fusion of these complementary local and global features results in stronger discrimination between anomalies and background, which explains the consistently higher $AUC(P_D, \tau)$ values obtained across different datasets.

Fig. 6 presents both the 3D ROC representation and three corresponding 2D ROC curves, namely $AUC(P_D, P_F)$, $AUC(P_D, \tau)$, and $AUC(P_F, \tau)$, providing a comprehensive evaluation of detection performance from different perspectives. The 3D ROC curve illustrates the joint behavior of detection probability, false alarm probability, and decision threshold, while the 2D projections allow a more detailed analysis of pairwise relationships between these variables.

In particular, the $AUC(P_D, P_F)$ curves reflect the overall separability between anomalies and background samples. As also reported in Table 8, the proposed OCAG/CAG method achieves consistently higher curves, indicating improved ranking of anomaly scores across all operating points. The $AUC(P_D, \tau)$ curves evaluate how detection probability evolves with respect to the decision threshold. The proposed method shows a faster and more stable increase in P_D as τ varies, which confirms its stronger sensitivity to anomaly responses and is consistent with its higher values in Table 8. Conversely, the $AUC(P_F, \tau)$ curves describe how false alarms evolve across threshold changes. The proposed method maintains a controlled growth of P_F , indicating robustness in background suppression, although some classical methods achieve lower absolute values due to overly conservative detection behavior.

The average AUC values across all datasets, together with their standard deviations, are reported in Tables 9 and 10, respectively. The proposed method achieves the highest mean $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$ values while exhibiting the smallest standard deviations for both metrics, demonstrating excellent detection accuracy, strong target discrimination, and consistent performance across heterogeneous hyperspectral scenes. According to the $AUC(P_F, \tau)$ metric, the RX detector achieves the lowest average value, indicating the strongest background suppression capability among the compared methods.

The results indicate that the proposed detector provides the best overall detection performance, as reflected by its highest $AUC(P_D, P_F)$. Moreover, it achieves a substantially higher $AUC(P_D, \tau)$ than the competing methods, indicating that anomalies are detected more effectively over a wide range of decision thresholds. Although its average $AUC(P_F, \tau)$ is higher than that of several conventional methods, this metric should be interpreted together with $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$. The proposed method intentionally improves anomaly sensitivity, resulting in a moderate increase in false alarms while providing a much larger improvement in detection capability. Consequently, it achieves the best overall balance between detection accuracy and target detectability across all evaluated datasets.

The proposed method is compared with several advanced anomaly detectors for the San Diego dataset in Table 11. According to $AUC(P_D, P_F)$, the best performance is obtained by PAEE (Pixel-Associated AutoEncoder), which considers global pixel similarity. PAEE simultaneously exploits the advantages of background and anomaly dictionary construction using superpixel distance estimation, subsequent autoencoders, and a two hidden layers feature similarity constraint network. The proposed CAG follows, along with CAPNet, which consists of two cascaded autoencoders with an adaptive pixel-level attention module. In terms of $AUC(P_D, \tau)$, the proposed CAG and OCAG detectors rank respectively first and second by a large margin with respect to other detectors, with PAEE and CAPNet resulting third and fourth, respectively. Compared with recent deep-learning methods, although PAEE achieves the highest $AUC(P_D, P_F)$, indicating excellent overall ranking capability, its considerably lower $AUC(P_D, \tau)$ suggests that the anomaly scores increase more slowly with decreasing threshold than those of the proposed method. CAPNet, SSBS, and S3S-LTGRMA mainly rely on reconstruction- or representation-based learning, whereas the proposed detector explicitly

incorporates complementary morphological priors together with global graph information. As a result, CAG achieves the highest $AUC(P_D, \tau)$, demonstrating the strongest anomaly–background separability across decision thresholds. In terms of $AUC(P_F, \tau)$, the best value is obtained by S3S-LTGRMA, which utilizes five types of information: local, global, spectral, spatial, and structural. CAPNet and SSBS rank second and third, respectively. Although some competing methods achieve lower $AUC(P_F, \tau)$ values, CAG maintains a tolerable level of background suppression while requiring substantially less computational time than PAAE and CAPNet.

Table 11. Comparison results with some state-of-the-art methods for San Diego dataset.

	CAG (proposed)	OCAG (proposed)	PAAE (Xiang et al., 2024)	CAPNet (Zhao et al., 2025a)	SSBS (Wang and Hu, 2025)	S3S-LTGRMA (Chen et al., 2025)
AUC (P_D, P_F)	<u>0.9918</u>	0.9889	0.9971	<i>0.9916</i>	0.9764	0.9903
$AUC(P_D, \tau)$	0.8423	0.8194	<i>0.6283</i>	0.3814	0.1007	0.1854
$AUC(P_F, \tau)$	0.1625	0.1707	0.0347	<u>0.0144</u>	<i>0.0254</i>	0.0092
AUC(TD)	1.8341	<u>1.8083</u>	<i>1.6254</i>	1.3730	1.0771	1.1757
AUC(BS)	0.8293	0.8182	<i>0.9624</i>	0.9772	0.9510	0.9811
Time	<u>52.76</u>	<i>57.18</i>	325.37	345.13	---	35.12

In terms of the derived metric AUC(TD), which combines detection accuracy and sensitivity to threshold variation, the proposed CAG achieves the highest value, followed closely by OCAG. This indicates that the proposed models provide the most consistent increase in detection probability as the decision threshold varies, reflecting strong overall separability between anomaly and background responses. In contrast, recent deep learning-based methods such as CAPNet, SSBS, and S3S-LTGRMA yield noticeably lower AUC(TD) values, suggesting weaker stability of detection performance across operating points. Regarding AUC(BS), which evaluates background suppression capability, the best performance is achieved by S3S-LTGRMA and CAPNet, followed closely by PAAE. These methods benefit from strong feature reconstruction and multi-source representation learning, which enhances background modeling. In comparison, the proposed CAG and OCAG show slightly lower AUC(BS) values, indicating that they prioritize anomaly sensitivity and detection strength rather than aggressive background suppression. However, their values remain competitive and consistent with their strong performance in $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$, demonstrating a balanced trade-off between detection sensitivity and background discrimination.

From the computational point of view, S3S-LTGRMA, CAG, and OCAG are the fastest methods, while CAPNet and PAAE are the slowest. Overall, the proposed CAG detector exhibits relatively high $AUC(P_D, P_F)$, and the highest value of $AUC(P_D, \tau)$, and its reasonable computation cost makes it a strong candidate for hyperspectral anomaly detection. The CAG network for the San Diego dataset contains 551.2k learnable parameters, which is approximately one order of magnitude smaller than those required by other networks such as CAPNet (4600k parameters).

5. Conclusion

This paper presented the Opening and Closing based Attention integrated with Graph features (OCAG) network for hyperspectral anomaly detection. The proposed framework explicitly exploits opening and closing morphological profiles as complementary indicators of anomalies and background, respectively, and introduces a dedicated cross-attention mechanism to model their relationships with the input patch. In parallel, a lightweight cluster-based graph convolutional network captures global spectral-spatial contextual information with a substantially lower computational burden than conventional pixel-wise graph construction. By combining local morphological attention features with global graph representations, the proposed network effectively reconstructs the background while enhancing the discrimination of anomalous targets.

The ablation study validates the effectiveness of the proposed architecture by demonstrating the complementary contributions of the morphological attention mechanism and the graph branch. In particular, the results indicate that, for several benchmark datasets, combining the closing morphological profile with graph features alone already provides highly accurate anomaly detection, highlighting the dominant role of background characterization in complex hyperspectral scenes. Nevertheless, the complete OCAG architecture provides the most consistent performance across different datasets by jointly exploiting complementary anomaly- and background-related information.

Experimental results on five benchmark hyperspectral datasets demonstrate the effectiveness and robustness of the proposed approach. Compared with conventional statistical and representation-based detectors, the proposed method consistently achieves superior overall detection performance. Comparisons with recent deep-learning-based methods

on the San Diego dataset further show that the proposed detector attains the highest $AUC(P_D, \tau)$, indicating stronger target detection capability over a wide range of decision thresholds, while maintaining competitive $AUC(P_D, P_F)$. The average performance across all datasets also confirms the robustness of the proposed method, which achieves the highest mean $AUC(P_D, P_F)$ and $AUC(P_D, \tau)$, together with low performance variation across different hyperspectral scenes. In addition, the proposed networks remain lightweight across all benchmark datasets, requiring approximately 551k–641k learnable parameters for CAG and 614k–738k learnable parameters for OCAG, while exhibiting a reasonable computational cost. This makes them efficient alternatives to larger reconstruction-based networks.

Despite these promising results, the proposed method has some limitations. The opening and closing morphological profiles are generated using fixed structuring elements with predefined shapes and sizes. Consequently, they may not adequately represent anomalies with arbitrary geometries or highly heterogeneous backgrounds, limiting their adaptability to diverse scenes.

Future work will therefore focus on developing adaptive morphological representations with learnable or data-driven structuring elements capable of automatically adjusting their shape and size according to image content. In addition, morphological attribute filters, which operate on connected image components according to geometric properties such as area, elongation, compactness, and shape, represent a promising alternative to conventional structuring-element-based operators. Integrating these adaptive morphological representations with the proposed attention framework, together with more advanced graph learning strategies, may further improve the robustness and generalization capability of hyperspectral anomaly detection in complex environments.

References

- A, R., Q. Gao, X. Zhang, W. Feng, S.K. Ali, 2025. Hyperspectral anomaly detection leveraging spatial attention and right-shifted spectral energy, *PLoS One*, 20 (9), e0330640.
- AlDwairi, M., F. Abdul Gafoor, M. Alciabahy, A. R. Alsaleh, A. AlMahdouri, S. Alyounis, D. E. AlMomani, T. Zhang, M. R. AlShehhi, 2025. Hyperspectral spectroscopy and machine learning for non-destructive monitoring of leaf water and chlorophyll content under greenhouse lighting conditions, *Environmental Technology & Innovation*, 40, 104432.
- Andika, F., M. Rizkinia, M. Okuda, 2020. A Hyperspectral Anomaly Detection Algorithm Based on Morphological Profile and Attribute Filter with Band Selection and Automatic Determination of Maximum Area, *Remote Sensing*, 12 (20), 3387.
- Bahmanpour, M., D. Abbasi-Moghadam, A. Sharifi, 2025. Attention-GANomaly: unsupervised hyperspectral anomaly detection via attention based GANs, *Infrared Physics & Technology*, 150, 106017.
- Benediktsson, J., J. Aever Palmason, and J. R. Sveinsson, 2005. Classification of Hyperspectral Data from Urban Areas Based on Extended Morphological Profiles, *IEEE Transactions on Geoscience and Remote Sensing*, 43 (3), 480–491.
- Chen, X., Y. Cui, S. Gao, L. Wang, C. Zhao, T. Luo, 2025. S3S-LTGRMA: Local-to-global based spatial-structural-spectral joint reconstruction with residual multi-head autoencoder for hyperspectral anomaly detection, *Infrared Physics & Technology*, 145, 105704.
- Ding, Y., H. Yan, J. He, *et al.*, 2025. A novel graph convolution and frequency domain filtering approach for hyperspectral anomaly detection, *Complex Intell. Syst.*, 11, 119.
- Feng, J., L. Zhang, 2021. Hyperspectral anomaly detection based on autoencoder and spatial morphology extraction, *J. Appl. Rem. Sens.*, 15 (3) 038507.
- Feng, S., *et al.*, 2025. Transformer-Based Cross-Domain Few-Shot Learning for Hyperspectral Target Detection, *IEEE Transactions on Geoscience and Remote Sensing*, 63, 1-16, Art no. 5501716.
- Furkat Eshpulatovich, S., M. Toshpulatov, M. Komoliddin Mamasabirovich, A. Abdusalomov, A. Khojamurotov, W. Lee, 2025. Hyperspectral Anomaly Detection With Enhanced Spectral Graph Transformer Network, *IEEE Access*, 13, 170554-170576.
- Ghosh, A., A. K. Ghosh, R. SahaRay, S. Sarkar, 2025. Classification using global and local Mahalanobis distances, *Journal of Multivariate Analysis*, 207, 105417.
- Gonzalez, R.C., R.E. Woods, 2018. Digital Image Processing. 4th Edition, Pearson Education, New York.
- Guo, Q., B. Zhang, Q. Ran, L. Gao, J. Li, A. Plaza, 2014. Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 7 (6), 2351–2366.
- Hou, Z., W. Li, R. Tao, *et al.*, 2022. Collaborative representation with background purification and saliency weight for hyperspectral anomaly detection, *Sci. China Inf. Sci.*, 65, 112305.
- Hu, X., Z. Li, L. Luo, H. R. Karimi, D. Zhang, 2025. Dictionary trained attention constrained low rank and sparse autoencoder for hyperspectral anomaly detection, *Neural Networks*, 181, 106797.
- Huo, Y., S. Wang, C. Wang, M. Zhang, H. Wang, 2026. Dual-stream background modeling network with anomaly suppression for hyperspectral anomaly detection, *International Journal of Applied Earth Observation and Geoinformation*, 148, 105233.

- Imani, M., 2017. RX Anomaly Detector with Rectified Background, *IEEE Geoscience and Remote Sensing Letters*, 14 (8), 1313-1317.
- Imani, M., 2018. Anomaly detection using morphology-based collaborative representation in hyperspectral imagery, *European Journal of Remote Sensing*, 51 (1), 457-471.
- Imani, M., H. Ghassemian, 2019. Morphology-based structure-preserving projection for spectral-spatial feature extraction and classification of hyperspectral data, *IET Image Processing*, 13 (2), 270-279.
- Imani, M., 2024. Collaborative Representation Based Unsupervised CNN for Hyperspectral Anomaly Detection, *Infrared Physics and Technology*, 141, 105498.
- Imani, M., 2025. Discriminant analysis-based attention network for hyperspectral target detection, *Optics & Laser Technology*, 182, Part B, 112208.
- Imani, M., S. Lim, 2025. An efficient eigenvalue decomposition-based attention network for polarimetric synthetic aperture radar image classification, *Advances in Space Research*, 76 (5), 2795-2812.
- Imani, M., D. Cerra, 2025. Triple Graph Convolutional Network for Hyperspectral Image Feature Fusion and Classification, *Remote Sensing*, 17 (9), 1623.
- Jiao, J., L. Xiao and C. Wang, 2025. Hyperspectral Anomaly Detection Based on Intrinsic Image Decomposition and Background Subtraction, *IEEE Access*, 13, 15723-15738.
- Kwon, H., N. M. Nasrabadi, 2005. Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery, *IEEE Trans. Geosci. Remote Sens.*, 43 (2), 388-397.
- Li, W., Q. Du, 2015. Collaborative representation for hyperspectral anomaly detection, *IEEE Trans. Geosci. Remote Sens.*, 53 (3), 1463-1474.
- Li, Z., S. Feng, C. Zhao, X. Du, W. Li and R. Tao, 2025a. Pixel-Weighted Background Reconstruction Network with Dual-Windows Blind Block for Hyperspectral Target Detection, *Chinese Journal of Electronics*, 34 (6), 1730-1745.
- Li, K., Y. Wang, Q. Ling, Y. Cai, Y. Qin, 2025b. A cascaded autoencoder unmixing network for Hyperspectral anomaly detection, *International Journal of Applied Earth Observation and Geoinformation*, 136, 104405.
- Li, X., Y. Yuan, 2024. Hyperspectral anomaly detection based on weighted low-rank sparse dictionary learning, *Neurocomputing*, 610, 128572.
- Li, S., K. Zhang, P. Duan, X. Kang, 2020. Hyperspectral Anomaly Detection With Kernel Isolation Forest, *IEEE Transactions on Geoscience and Remote Sensing*, 58 (1), 319-329.
- Liu, J., Z. Hou, W. Li, R. Tao, D. Orlando, H. Li, 2022. Multipixel Anomaly Detection With Unknown Patterns for Hyperspectral Imagery, *IEEE Transactions on Neural Networks and Learning Systems*, 33 (10), 5557-5567.
- Liu, H., K. Liu and L. Wang, 2026. Hyperspectral Anomaly Detection Method Based on Joint Optimization of Dynamic Band Selection and Background Reconstruction, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 19, 19410-19429.
- Lv, Z., M. Zhang, W. Sun, T. Lei, J. A. Benediktsson, T. Liu, 2025. Land cover change detection with hyperspectral remote sensing images: A survey, *Information Fusion*, 123, 103257.
- Matteoli, S., M. Diani and J. Theiler, 2014. An Overview of Background Modeling for Detection of Targets and Anomalies in Hyperspectral Remotely Sensed Imagery, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (6), 2317-2336.
- Ma, Y., S. Cai and J. Zhou, 2023. Adaptive Reference-Related Graph Embedding for Hyperspectral Anomaly Detection, *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1-14, Art no. 5504514.
- Reed, I. S., X. Yu, 1990. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution, *IEEE Trans. Acoust., Speech Signal Process.*, 38 (10), 1760-1770.
- Risheh, A., P. Tavakolian, A. Melnikov, A. Mandelis, 2022. Infrared computer vision in non-destructive imaging: Sharp delineation of subsurface defect boundaries in enhanced truncated correlation photothermal coherence tomography images using K-means clustering, *NDT & E International*, 125, 102568.
- Shang, X., M. Song, Y. Wang and H. Yu, 2022. Residual-Driven Band Selection for Hyperspectral Anomaly Detection, *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5, Art no. 6004805.
- Shanmugam, P., S. A. Miruna Joe Amali, 2024. Dual-discriminator conditional generative adversarial network optimized with hybrid manta ray foraging optimization and volcano eruption algorithm for hyperspectral anomaly detection, *Expert Systems with Applications*, 238, Part C, 122058.
- Tan, K., Z. Hou, F. Wu, Q. Du, Y. Chen, 2019. Anomaly Detection for Hyperspectral Imagery Based on the Regularized Subspace Method and Collaborative Representation, *Remote Sensing*, 11 (11), 1318.
- Vafadar, M., H. Ghassemian, 2018. Anomaly Detection of Hyperspectral Imagery Using Modified Collaborative Representation, *IEEE Geoscience and Remote Sensing Letters*, 15 (4), 577-581.
- Wang, Y., Y. Hou, Z. Li, W. Ren, F. Jiang, F. Li, 2025a. HSCRN: A hyperspectral image classification network based on dynamic context and feature refinement, *Optics & Laser Technology*, 192, Part A, 113395.
- Wang, X. *et al.*, 2025b. FCMMMA: Fourier Conditional Mask-Based Mixed Attention Method for Hyperspectral Anomaly Detection, *IEEE Transactions on Geoscience and Remote Sensing*, 63, 1-12, Art no. 5507012.
- Wang, R., J. Hu, 2025. Fewer interpretable bands via self-supervision for hyperspectral anomaly detection, *Neurocomputing*, 616, 128964.

- Wang M., et al., 2026. OT-AD: Optimal Transport-Guided Transformer for Hyperspectral Anomaly Detection, *IEEE Transactions on Geoscience and Remote Sensing*, 64, 1-18, Art no. 5502318.
- Wu, L., M. You, D. Zhao, X. Xu, Y. Yao, A. P. Vijayakumar, K. Cheng, P. Xiang, 2026. DSFNet: Dual-stream spatial-spectral fusion network for hyperspectral anomaly detection, *Digital Signal Processing*, 175, 106019.
- Xiang, P., S. Ali, J. Zhang, S. K. Jung, H. Zhou, 2024. Pixel-associated autoencoder for hyperspectral anomaly detection, *International Journal of Applied Earth Observation and Geoinformation*, 129, 103816.
- Xu, Y., B. Du, L. Zhang, S. Chang, 2020. A Low-Rank and Sparse Matrix Decomposition- Based Dictionary Reconstruction and Anomaly Extraction Framework for Hyperspectral Anomaly Detection, *IEEE Geoscience and Remote Sensing Letters*, 17 (7), 1248-1252.
- Xu, S., Y. Liu, Y. Zhang and H. Gao, 2026. Lightweight KAN Convolution Spectral–Spatial Network With Purification Window for Hyperspectral Anomaly Detection, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 19, 6892-6906.
- Yang, Y., Y. Li, D. Liang and T. Guo, 2025. Interactive Global-Local Anomaly Fusion Network for Hyperspectral Anomaly Detection, 2025 2nd International Conference on Intelligent Perception and Pattern Recognition (IPPR), Chongqing, China, pp. 191-194.
- Yang, X., B. Tu, Q. Li, J. Li and A. Plaza, 2024. Graph Evolution-Based Vertex Extraction for Hyperspectral Anomaly Detection, *IEEE Transactions on Neural Networks and Learning Systems*, 35 (12), 17372-17386.
- Yang, Y., Q. Yang, S. Song, D. Liu, J. Zhang, 2024. Discriminative coefficient analysis-based collaborative representation with enhanced Background-Anomaly separation for hyperspectral anomaly detection, *Infrared Physics & Technology*, 141, 105476.
- You, M., Y. Yao, D. Zhao, Z. Zhao, P. V. Arun, Y. Wang, H. Zhou, R. Chi, 2026. S3CRAD: Superpixel-guided background inpainting and spatial-spectral constrained representation for hyperspectral anomaly detection, *Optics and Lasers in Engineering*, 201, 109657.
- Zhang, L., F. Lin, B. Fu, 2024b. A joint model based on graph and deep learning for hyperspectral anomaly detection, *Infrared Physics & Technology*, 139, 105335.
- Zhang, H., H. Sun, X. Sun, H. Gao, L. Gao and B. Zhang, 2024c. Spectral–Spatial Out-of-Distribution-Based Unsupervised Band Selection Method for Hyperspectral Anomaly Detection, *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1-15, Art no. 5538915.
- Zhang, J., P. Xiang, J. Shi, X. Teng, D. Zhao, H. Zhou, H. Li, J. Song, 2024a. A light CNN based on residual learning and background estimation for hyperspectral anomaly detection, *International Journal of Applied Earth Observation and Geoinformation*, 132, 104069.
- Zhao D., et al., 2024. Hyperspectral Anomaly Detection Based on Empirical Mode Decomposition and Local Weighted Contrast, *IEEE Sensors Journal*, 24 (20), 33847-33861.
- Zhao, Z., J. Song, M. You, P. Xiang, D. Zhao, J. Zhang, H. Zhou, D. Wang, X. Wang, 2025a. Hyperspectral anomaly detection via cascaded convolutional autoencoders with adaptive pixel-level attention, *Expert Systems with Applications*, 279, 127366.
- Zhao, D., X. Xu, M. You, P.V. Arun, Z. Zhao, J. Ren, L. Wu, H. Zhou, 2025b. Local Sub-Block Contrast and Spatial–Spectral Gradient Feature Fusion for Hyperspectral Anomaly Detection. *Remote Sens.*, 17, 695.
- Zhao, D., M. You, P. Xiang, J. Hu, Y. Asano, X. Yu, C.-C. Hsu, H. Zhou, J. Ren, 2026. TBCNet: Twin-branch collaborative network for hyperspectral anomaly detection, *Pattern Recognition*, 180, Part F, 114647.
- Zhou, J., C. Kwan, B. Ayhan and M. T. Eismann, 2016. A Novel Cluster Kernel RX Algorithm for Anomaly and Change Detection Using Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 54 (11), 6497-6504.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.