

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Hyperspectral Anomaly Detection with Enhanced Spectral Graph Transformer Network

Furkat Safarov ^{1,5} Toshpulatov Mukhiddin ², Misirov Komoliddin ³, Akmalbek Abdusalomov ^{1,5,6}

Azizbek Khojamurotov ¹, and Wookey Lee ⁴,

¹ Department of Computer Engineering, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam-si 13120, Republic of Korea

² Post-Doctoral researcher, SpaceTop Research Center KAIST, South Korea

³ Department of Financial Accounting and Reporting, Tashkent State University of Economics, Tashkent 100066, Uzbekistan

⁴ Industrial Engineering & Biomedical Science and Engineering, Inha University, Incheon, South Korea

⁵ Department of Computer Systems, Tashkent University of Information Technologies Named after Muhammad Al-Khwarizmi, Tashkent 100200, Uzbekistan.

⁶ Department of Information Processing and Management Systems, Tashkent State Technical University, Tashkent 100095, Uzbekistan.

Corresponding author: Wookey Lee (trinity@inha.ac.kr).

ABSTRACT Hyperspectral imaging (HSI) plays a pivotal role across diverse sectors—including agriculture, environmental monitoring, and defense—by capturing rich spectral information that enables fine-grained material discrimination. Despite its advantages, the inherently high dimensionality and complexity of hyperspectral data present substantial challenges for reliable anomaly detection. Conventional approaches often fall short when confronted with subtle or context-dependent anomalies, underscoring the need for more advanced methodologies. In this study, we propose the Enhanced Spectral Graph Transformer Network (ESGTN), a novel framework that synergistically combines graph-based modeling, transformer architectures, and hyperbolic space embedding to improve the accuracy and efficiency of spectral anomaly detection. By representing HSIs as graphs, ESGTN effectively models both spatial and spectral relationships. The transformer component, equipped with self-attention mechanisms, adaptively emphasizes salient features, while the incorporation of hyperbolic embeddings provides a compact and distortion-minimized representation of the data's hierarchical structure. Extensive experiments conducted on multiple benchmark hyperspectral datasets demonstrate that ESGTN consistently outperforms existing state-of-the-art methods, achieving superior precision, recall, and computational efficiency. These findings highlight the model's robustness and practical applicability across a range of real-world scenarios. Moreover, this work contributes to the growing body of research at the intersection of deep learning and hyperspectral imaging, offering a scalable path forward for tackling the complex analytical demands inherent to high-dimensional remote sensing data.

INDEX TERMS Hyperspectral Image; Anomaly Detection; Hyperbolic Space Embedding; Spatial-Spectral Analysis; Remote Sensing.

I. INTRODUCTION

Over the past few decades, hyperspectral imaging (HSI) has emerged as a transformative tool for material characterization, enabling the acquisition of detailed spectral signatures across hundreds of narrow and contiguous bands [1]. This wealth of spectral information provides unprecedented insight into subtle variations in surface composition and structure. However, such richness also introduces significant complexity in data analysis and interpretation [2]. One particularly challenging task is anomaly detection—the identification of spectral patterns or observations that deviate from the expected background distribution [3]. Traditional approaches, which often rely on pixel-wise analysis [4], are limited in their ability to capture

the complex spatial and spectral interdependencies that characterize HSI data. Consequently, these methods frequently fall short in delivering robust and reliable detection performance [5].

The recent emergence of graph-based methodologies has significantly advanced hyperspectral image (HSI) analysis, offering powerful means to model the intricate relationships among spectral vectors of spatially adjacent pixels [6,7]. By encoding HSIs as graph representations, these approaches effectively retain both local and global structural information, thereby facilitating more comprehensive data interpretation and analysis [8]. Nonetheless, conventional graph-based techniques encounter notable challenges when confronting the high dimensionality intrinsic to

hyperspectral data. In such settings, standard graph analytics may inadvertently attenuate critical spectral distinctions, diminishing their capacity to detect nuanced anomalies [9,10]. To address these limitations, we propose the Enhanced Spectral Graph Transformer Network (ESGTN)—a novel architecture that unifies graph-based representation learning with the expressive power of transformer networks. ESGTN capitalizes on self-attention mechanisms to dynamically model complex spatial-spectral interactions within the graph structure, thereby preserving and emphasizing the hierarchical nature of hyperspectral data. A key innovation within our framework is the integration of Hyperbolic Space Embedding, which enables the projection of spectral features into a non-Euclidean space that better reflects the hierarchical relationships inherent in real-world data. This embedding reduces distortion and enhances representational fidelity, facilitating the detection of subtle spectral anomalies often overlooked in Euclidean spaces.

Complementing this, the Spectral Difference Amplification Module (SDAM) is introduced to explicitly magnify inter-class spectral disparities, thus enhancing the sensitivity of the transformer network to minute anomalous deviations. The proposed ESGTN architecture is thoroughly examined both theoretically and empirically, with extensive evaluations conducted on multiple benchmark HSI datasets. Results demonstrate that ESGTN consistently surpasses existing state-of-the-art approaches in terms of precision, recall, and computational efficiency. By bridging graph-based learning, attention mechanisms, and hyperbolic geometry, ESGTN establishes a new paradigm in hyperspectral anomaly detection, opening promising avenues for future research and practical deployment in remote sensing and related domains.

II. RELATED WORKS

Anomaly detection in hyperspectral images (HSIs) has remained a pivotal area of research, driven by its vital role in applications such as environmental monitoring, precision agriculture, mineral exploration, and military surveillance [11]. Over the years, a diverse array of techniques has been developed to identify spectral irregularities that deviate from the normative background, each offering distinct advantages and facing unique limitations. This section provides a critical overview of the evolution of these anomaly detection methods, tracing their conceptual and technical progress. In doing so, it underscores the inherent challenges that persist—particularly in capturing the complex spatial-spectral dependencies and addressing the high-dimensional nature of HSI data. Against this backdrop, we introduce the Enhanced Spectral Graph Transformer Network (ESGTN) as a next-generation approach. ESGTN is designed to overcome the limitations of existing models by integrating graph-based representations, transformer-based self-

attention mechanisms, and hyperbolic space embeddings, thereby offering a more robust and scalable framework for hyperspectral anomaly detection.

Early efforts in hyperspectral anomaly detection predominantly focused on pixel-based techniques, with the Reed-Xiaoli (RX) detector and its derivatives serving as foundational methods in the field [12]. These approaches model the background spectral distribution and identify anomalies as pixels exhibiting significant deviations from the expected norm [13]. While such methods perform reasonably well in spectrally homogeneous environments, their effectiveness diminishes in complex scenes characterized by high intra-class variability and heterogeneous spectral signatures [14]. The advent of machine learning introduced new paradigms for anomaly detection, with both supervised and unsupervised methods broadening the analytical toolkit. Algorithms such as Support Vector Machines (SVM) [16], Random Forests, and k-Means clustering enhanced detection capabilities by exploiting spectral patterns through classification or clustering strategies. However, the performance of these models often hinges on extensive preprocessing steps, including dimensionality reduction and feature engineering, which can be labor-intensive and dataset-specific. As the field matured, graph-based techniques gained traction due to their ability to model HSIs as structured data, where each pixel is treated as a node connected to others based on spectral or spatial similarity [17]. This representation enables the capture of complex spatial-spectral dependencies, improving the model's sensitivity to anomalous behavior [18]. Approaches based on Graph Cuts and Minimum Spanning Trees have shown promise in segmenting HSIs into semantically meaningful regions [19]. Nevertheless, these graph-based methods are often computationally demanding and may encounter scalability issues when applied to large-scale or high-dimensional hyperspectral datasets [20].

The emergence of deep learning has significantly transformed hyperspectral anomaly detection, primarily by automating feature extraction and enabling end-to-end learning pipelines that bypass the need for manual feature engineering [21]. Among early adopters, Convolutional Neural Networks (CNNs) and Autoencoders have demonstrated notable efficacy in learning joint spatial-spectral representations from high-dimensional hyperspectral data [22]. Building upon these foundational models, recent studies have proposed more sophisticated architectures aimed at enhancing detection robustness and generalization. For instance, the Hyperspectral Anomaly Detection via Generalized Shrinkage Mappings (HADGSM) framework introduces nonconvex penalties to better approximate low-rank representations, thereby improving the separation of anomalies from the background [23]. Similarly, the Memory-Augmented Autoencoder (MAAE) integrates adaptive reconstruction

mechanisms and sample attribution mining to distinguish subtle anomalies more effectively [24]. The Robust Multi-Stage Progressive Autoencoder (RMSAD) further advances performance by adopting a multi-stage learning paradigm that incrementally refines detection accuracy across diverse datasets [25]. In addition, innovative models such as multi-scale blind-spot networks employ hierarchical mask convolutions and dynamic fusion strategies to detect anomalies of varying sizes [26], while topology-based methods restructure hyperspectral data into ordered topological spaces to improve spatial coherence and detection fidelity [27]. Differential convolution techniques have also been explored to suppress background interference and enhance localized feature extraction, thereby increasing sensitivity to spectral deviations [28]. More recently, the Capsule Differential Adversarial Continual Learning GAN (CL-CaGAN) addresses open-scenario detection challenges by incorporating capsule networks and adversarial learning to mitigate catastrophic forgetting and operate with limited prior information [29]. Despite these advances, deep learning-based methods often suffer from critical limitations, including a dependence on large labeled datasets and a general lack of interpretability—both of which impede their broader deployment in real-world scenarios. Inspired by their success in natural language processing, Transformer architectures have recently gained traction in hyperspectral anomaly detection due to their powerful self-attention mechanisms, which dynamically assign importance to spatial and spectral regions based on contextual relevance [30,31]. While initial studies demonstrate the promise of Transformers in capturing long-range dependencies in HSI data, their fusion with graph-based structures and non-Euclidean representations remains an underexplored frontier. Noteworthy attempts in this direction include hybrid architectures that combine axial attention with CNN-based residual learning to enhance multiscale detection [32], as well as Transformer-based frameworks leveraging extended multilinear mixing models (EMLM) to improve nonlinear anomaly reconstruction [33].

The Enhanced Spectral Graph Transformer Network (ESGTN) is conceived as a comprehensive response to these emerging challenges and opportunities. By unifying graph-based modeling, Transformer-based attention mechanisms, and hyperbolic space embeddings, ESGTN creates a highly expressive and computationally efficient framework for hyperspectral anomaly detection. This integration enables the network to capture both local and global spectral-spatial dependencies while preserving the hierarchical structure of the data. In doing so, ESGTN not only overcomes the limitations of prior methods but also establishes a new benchmark in the field, offering a robust, scalable, and interpretable solution for high-dimensional remote sensing applications.

III. PROPOSED METHOD

This section details the methodological framework underlying the Enhanced Spectral Graph Transformer Network (ESGTN), a novel architecture developed to advance the field of hyperspectral anomaly detection. ESGTN strategically integrates graph-based representations, transformer-based attention mechanisms, and hyperbolic space embeddings—each serving a distinct yet complementary function in the analysis of hyperspectral imaging (HSI) data. The process begins with the transformation of HSIs into graph-structured data, wherein individual pixels are represented as nodes, and edges are constructed based on spectral similarity and spatial proximity. This graph-based abstraction enables the model to capture both local and global spatial-spectral interdependencies that are often obscured in conventional pixel-wise approaches. Through this representation, critical structural information is retained, laying a robust foundation for subsequent learning processes. Table 1 provides an overview of the graph construction methodology, detailing the node and edge definitions as well as the similarity metrics employed to encode meaningful relationships within the data.

TABLE I
THE KEY COMPONENTS OF OUR METHODOLOGY AND THEIR
CONTRIBUTIONS TO THE NETWORK'S PERFORMANCE

Component	Description
Graph-based Preprocessing	Transforms HSIs into graph structures where pixels are nodes linked by spectral and spatial similarities, ensuring the preservation of structural information.
Transformer with Self-Attention	Adapts transformers to graph data, using self-attention to enhance focus on relevant hyperspectral regions, improving anomaly detection precision.
Hyperbolic Space Embedding	Embeds pixel spectral data in hyperbolic space to utilize hierarchical properties with minimal distortion, boosting the detection of subtle anomalies.

To further optimize performance for hyperspectral anomaly detection, we adapt the transformer architecture to effectively operate on graph-structured data derived from HSIs. Through the incorporation of self-attention mechanisms, the model learns to dynamically focus on the most salient regions within the hyperspectral image, thereby emphasizing features that are most indicative of anomalous behavior. This selective attention enhances the network's ability to distinguish subtle deviations from the background, which is particularly valuable in complex and heterogeneous scenes. In addition, we integrate hyperbolic space embedding to represent the spectral information of individual pixels. Unlike conventional Euclidean embeddings, the hyperbolic space is inherently suited to capturing hierarchical structures—such as those often present in spectral

distributions—while minimizing geometric distortion. This approach enables a more compact and discriminative encoding of the data, significantly improving the model's sensitivity to fine-grained anomalies. The effectiveness of

this combined strategy is illustrated in Figure 1, which outlines the architectural components and their functional roles within the ESGTN framework.

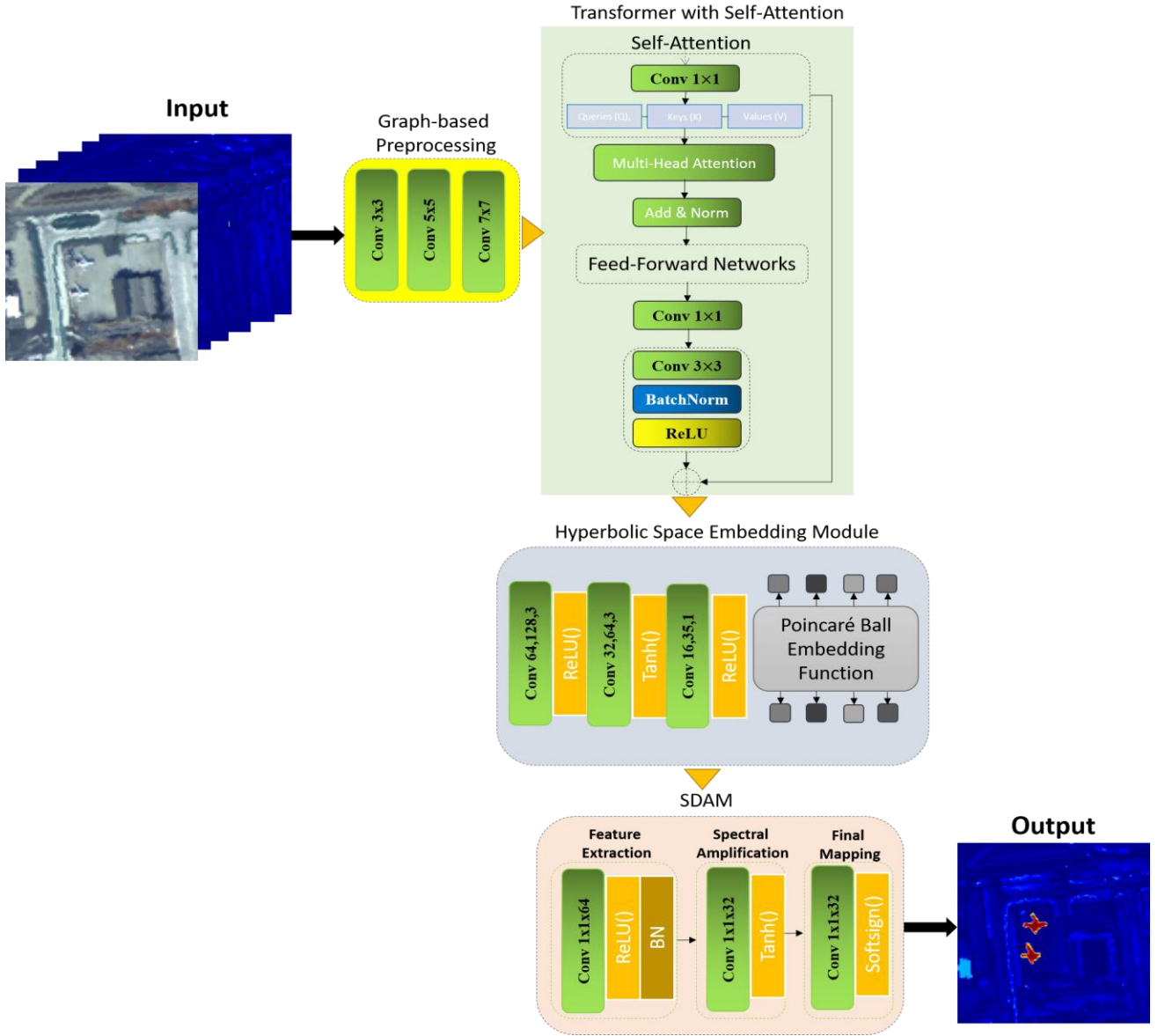


FIGURE 1. Architecture of the ESGTN for Hyperspectral Anomaly Detection.

A. GRAPH BASED PREPROCESSING

The graph-based preprocessing module of the Enhanced Spectral Graph Transformer Network (ESGTN) serves as a foundational step in re-structuring hyperspectral images (HSIs) into graph representations that more effectively capture their intrinsic spatial and spectral complexity. In this formulation, each pixel is represented as a node, while edges

are constructed based on a combination of spectral similarity and spatial proximity. This transformation allows the model to exploit both local and global contextual relationships—elements often underutilized in traditional pixel-based analyses—thereby significantly enhancing its capacity to detect anomalies. In the context of hyperspectral data, where each pixel encodes a detailed spectral signature across numerous contiguous bands, capturing inter-pixel dependencies is essential for robust interpretation. Conventional methods frequently overlook the nuanced correlations between neighboring pixels, leading to limited

detection performance, particularly in heterogeneous environments. By converting HSIs into graphs, ESGTN preserves essential structural information, allowing anomalies to be identified not solely as individual outliers but as deviations from the expected behavior within their local neighborhood. Each node in the constructed graph is defined by the spectral vector of the corresponding pixel, encapsulating its material composition. Edges are determined using metrics that reflect both geometric closeness and spectral similarity, ensuring that relationships between nodes are contextually meaningful. This graph-based representation enables the model to reason about each pixel in relation to its environment, facilitating more accurate and context-aware anomaly detection.

Edge Construction involves calculating the spectral similarity between nodes. The spectral similarity S between two pixels u and v with spectral vectors x_u and x_v quantified using the cosine similarity measure:

$$S(u, v) = \frac{x_u \cdot x_v}{\|x_u\| \|x_v\|} \quad (1)$$

Edges are formed if the similarity $S(u, v)$ exceeds a predefined threshold, indicating significant spectral correspondence. This method focuses on connecting nodes that share meaningful spectral characteristics, emphasizing the detection of spectral outliers. The weights assigned to the edges are inversely proportional to the Euclidean distance between the spectral vectors of the connected nodes. This ensures that stronger connections (higher weights) exist between more similar nodes, reinforcing the importance of close spectral matches:

$$w_{uv} = \frac{1}{\|x_u - x_v\| + \epsilon} \quad (2)$$

To ensure numerical stability during processing, a small constant ϵ is added to the denominator to prevent division by zero. As part of the Graph Enhancement procedure, the spectral vectors associated with each node are normalized to have zero mean and unit variance across all spectral bands. This normalization mitigates biases arising from intensity variations and ensures that the model focuses on the relative spectral patterns rather than absolute magnitudes, thereby improving the robustness and generalizability of anomaly detection.

$$x_{u'} = \frac{x_u - \mu_u}{\sigma_u} \quad (3)$$

Once the hyperspectral image (HSI) is transformed into a graph structure, it is subsequently processed by the transformer network within the ESGTN framework. This architecture enables the application of self-attention

mechanisms across the entire graph, allowing the model to simultaneously consider both spatial and spectral relationships among nodes. By leveraging this global receptive field, the network can capture long-range dependencies and contextual interactions that are critical for identifying subtle or spatially diffuse anomalies. The graph representation serves not only to structure the data meaningfully but also to provide a contextual baseline against which anomalous nodes can be contrasted. This integration of transformer-based attention with graph-structured data significantly enhances the model ability to detect irregularities with greater precision and reliability.

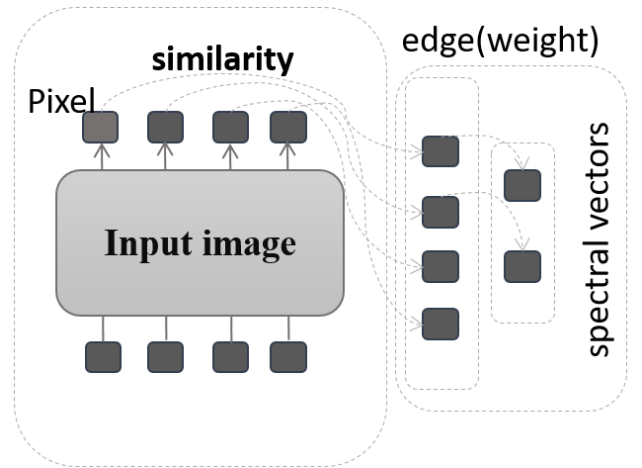


FIGURE 2. Graph-based Representation of Hyperspectral Image Data. Edges connect nodes based on spectral and spatial similarity. Edge weights reflect the strength of similarity. Each node stores spectral vector information.

B. TRANSFORMER WITH SELF – ATTENTION

In the Enhanced Spectral Graph Transformer Network (ESGTN), the Transformer with Self-Attention mechanism serves as a central component for processing the graph-based representations of hyperspectral images (HSIs). This section outlines the architectural adaptations and implementation details of the transformer module, specifically designed to address the high dimensionality, spectral redundancy, and spatial complexity inherent to HSI data. By leveraging self-attention, the transformer dynamically prioritizes salient features across both spectral and spatial domains, facilitating rich contextual understanding and enabling the precise differentiation of anomalies from background patterns. These capabilities are essential for enhancing detection performance, particularly in complex or heterogeneous scenes where conventional convolutional or statistical methods may fall short.

Originally designed for natural language processing, the Transformer architecture has been effectively adapted in ESGTN to handle the complex, high-dimensional, and structured nature of hyperspectral imaging (HSI) data. At the heart of this adaptation is the self-attention mechanism,

which allows the model to dynamically assign importance to different regions within the hyperspectral image. This selective focus enables the network to emphasize regions that are more likely to exhibit anomalous behavior, while simultaneously leveraging the underlying graph structure of the data to maintain contextual integrity. In the context of graph-transformed HSIs, each node corresponds to a pixel represented by its spectral vector, and edges encode both spectral similarity and spatial proximity. The self-attention mechanism calculates the response at a given node by attending to all other nodes in the graph, aggregating information through a learned set of attention weights. This process ensures that each node's representation is updated with respect to its broader context, facilitating the identification of subtle and spatially-dependent anomalies.

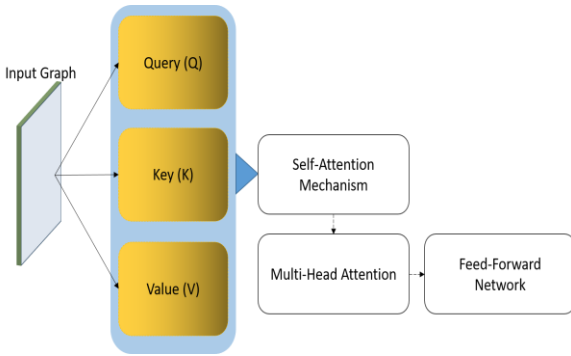


FIGURE 3. Transformer Architecture with Self-Attention for Graph Data.

The attention score between two nodes i and j is computed as follows:

$$A(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

In this formulation, Q , K , and V represent the query, key, and value matrices, respectively, all of which are derived from the spectral feature vectors of the pixels (nodes). The term d_k denotes the dimensionality of the key vectors and is used to scale the dot products, thereby preventing disproportionately large attention scores that could destabilize the learning process. To further enhance the model's ability to attend to different types of information across the spectral-spatial domain, the self-attention mechanism is extended into multi-head attention. In multi-head attention, the attention computation is executed in parallel across multiple attention "heads." Each head learns to capture different aspects of the input data by operating on distinct linear projections of the queries, keys, and values. The outputs from these heads are then concatenated and passed through a final linear

transformation to yield the aggregated feature representation. This mechanism enables the model to synthesize information from multiple representational subspaces, thus enriching its capacity to detect complex and subtle anomalies within hyperspectral data:

$$MH(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \quad (5)$$

where $\text{head}_i = A(QW_i^Q, KW_i^K, VW_i^V)$

here W_i^Q , W_i^K , W_i^V , and W^O are parameter matrices for the i -th attention head and the output projection, respectively. Each attention output is fed into a position-wise feed-forward network, which consists of two linear transformations with a ReLU activation in between:

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (6)$$

The feed-forward network within the transformer architecture applies the same transformation independently to each position, enabling the model to capture non-linear interactions among features while maintaining position-wise consistency. In the context of ESGTN, integrating the transformer with the graph-structured representation of hyperspectral images (HSIs) necessitates adapting the self-attention mechanism to align with the graph's topology—specifically, its node connectivity based on spectral and spatial similarities. This adaptation ensures that attention is not only guided by feature relevance but is also constrained by the structural relationships encoded in the graph, resulting in context-aware and topology-preserving feature processing. Despite its representational power, the transformer architecture is known for its high computational cost, primarily due to the self-attention mechanism's quadratic complexity with respect to sequence length—in this case, the number of pixels or nodes in the graph. To mitigate this issue and ensure scalability, efficient attention mechanisms are incorporated into ESGTN. Techniques such as sparse attention, localized attention, and graph-aware attention reduce computational overhead by limiting attention computations to relevant or proximate nodes. These strategies strike a balance between efficiency and expressive power, allowing ESGTN to scale effectively to large hyperspectral datasets without compromising detection performance.

C. HYPERBOLIC SPACE EMBEDDING

Within the Enhanced Spectral Graph Transformer Network (ESGTN), the integration of Hyperbolic Space Embedding marks a significant methodological advancement, enabling the model to more effectively leverage the inherent hierarchical structures present in hyperspectral images (HSIs). This section outlines the theoretical rationale, implementation strategy, and practical implications of embedding hyperspectral data in a hyperbolic space to improve anomaly detection performance. Hyperbolic geometry, distinguished by its constant negative

curvature, provides a highly expressive framework for representing hierarchical data with minimal distortion. In contrast to Euclidean spaces, which struggle to efficiently embed tree-like or nested structures, hyperbolic spaces can compactly represent such relationships in low-dimensional manifolds. Given that hyperspectral data often exhibit natural hierarchies—both spectrally (through material class groupings) and spatially (through local structural similarities)—hyperbolic embedding allows for richer, more discriminative representations of pixel relationships. In ESGTN, we adopt the Poincaré ball model of hyperbolic space, a widely used formulation due to its geometric intuitiveness and computational tractability. In this model, each point in the hyperbolic space is represented as a vector located within the open unit ball in Euclidean space. The distance between points grows exponentially as they approach the boundary, which aligns well with the goal of emphasizing subtle but hierarchically meaningful differences in spectral data. This property enables the model to distinguish fine-grained anomalies that may be indistinct in standard Euclidean embeddings:

$$B^n = \{x \in \mathbb{R}^n : \|x\| < 1\} \quad (7)$$

where $\|\bullet\|$ denotes the Euclidean norm. The distance d between two points $x, y \in B^n$ in the Poincaré ball is given by:

$$d(x, y) = \cosh^{-1} \left(1 + \frac{2 \|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)} \right) \quad (8)$$

This metric—characterized by the exponential expansion of distances from the origin—aligns naturally with the model's objective to accentuate subtle yet semantically meaningful hierarchical differences within hyperspectral imaging (HSI) data. The embedding process begins by mapping spectral features, originally defined in Euclidean space, into a hyperbolic space using a learnable transformation. This transformation is typically initialized to linearly project the features into the Poincaré ball, a hyperbolic manifold that enables efficient and expressive geometric representation. During training, the transformation is optimized such that spectrally or spatially similar pixels are placed closer together in hyperbolic space, while dissimilar ones are separated by greater distances. The use of hyperbolic distance as an optimization criterion allows the model to respect and preserve the inherent hierarchical relationships in the data, enhancing its ability to detect anomalies that may be structurally subtle but topologically significant. In ESGTN, integrating hyperbolic space embedding requires aligning the graph-based HSI

representation with the geometry of the Poincaré ball. Each node—corresponding to a pixel—is embedded within this space, while edges encode relationships based on hyperbolic distances rather than traditional Euclidean metrics. This approach not only retains the graph's topological structure but also amplifies relational nuances, enabling the model to more effectively distinguish anomalous patterns from background variation. As a result, ESGTN benefits from a deeper, geometry-aware understanding of the data, which significantly improves the model's anomaly detection capabilities in complex and hierarchical hyperspectral scenes.

D. SPECTRAL DIFFERENCE APPLICATION MODULE

Within the Enhanced Spectral Graph Transformer Network (ESGTN), the Spectral Difference Amplification Module (SDAM) serves as a key component aimed at enhancing the contrast between anomalous and background spectral signatures in hyperspectral images (HSIs). This module is particularly critical in scenarios where anomalies exhibit only subtle deviations from their surrounding environment—differences that are often obscured by the high dimensionality and inherent noise present in hyperspectral data. Anomalies in HSIs are typically defined by their distinct spectral profiles, yet these signatures may be difficult to detect without targeted enhancement. The SDAM addresses this challenge by amplifying these spectral discrepancies, thereby increasing the model's sensitivity to even minor deviations. This enhancement process combines both linear and non-linear transformations, which are carefully designed to reshape the spectral feature space in a way that emphasizes variations indicative of anomalous behavior. Architecturally, the SDAM is positioned within ESGTN to operate as a preprocessing module, acting on the extracted deep spectral-spatial features before they are processed by the transformer component. Initially, detailed spectral-spatial features are derived from the graph-based representation of the HSI, encapsulating pixel-wise material characteristics informed by contextual structure. These features are then passed through a linear transformation layer, which scales and re-weights the spectral information to heighten contrast between anomalous and background pixels:

$$F' = F \cdot W + b \quad (9)$$

Here, F denotes the matrix of extracted spectral-spatial features, W is the learnable weight matrix, b is the bias vector, and F' represents the linearly transformed feature matrix. These weights and biases are learned during training to maximize the enhancement of spectral discrepancies that are most relevant for anomaly detection. The goal of this transformation is to re-weight and scale the spectral features in a way that makes subtle deviations more pronounced and distinguishable from the background. Following the linear

amplification, the features are passed through a non-linear transformation to further increase contrast between spectral signatures. This step is critical for emphasizing nuanced differences that may not be fully captured by linear operations alone. The non-linear transformation can be expressed as:

$$F'' = \sigma(F') \quad (10)$$

where σ represents a non-linear function such as a sigmoid or hyperbolic tangent, and F'' denotes the amplified features. The amplified features F'' are then normalized to ensure that their magnitudes remain within a suitable range for subsequent processing by the transformer. This normalization helps maintain the stability of the network and preserves the enhancements achieved by the amplification process:

$$F''' = \frac{F'' - \mu}{\sigma} \quad (11)$$

here, μ and σ denote the mean and standard deviation of the amplified features, respectively. This normalization ensures that the resulting feature vectors are scale-invariant and comparable across spatial regions, which is critical for the stability and effectiveness of subsequent attention computations. The final, SDAM-preprocessed features F''' are then fed into the transformer network. With spectral discrepancies now explicitly amplified and normalized, the self-attention mechanisms within the transformer can more effectively identify and concentrate on the most informative variations within the hyperspectral data. This targeted attention improves the model's sensitivity to subtle anomalies and enhances the overall detection accuracy and robustness of the ESGTN framework. By integrating SDAM as a front-end module, ESGTN ensures that anomaly-relevant features are highlighted early in the pipeline, enabling more context-aware and hierarchy-sensitive processing in the transformer layers.

E. RECONSTRUCTION AND DETECTION

Within the Enhanced Spectral Graph Transformer Network (ESGTN), the Reconstruction and Detection Strategy serves as a crucial mechanism for distinguishing anomalous patterns from the typical background characteristics present in hyperspectral images (HSIs). This strategy is grounded in the premise that anomalies exhibit spectral signatures that diverge significantly from those of the normal background. By accurately reconstructing the expected spectral profile of the background and measuring deviations from this reconstructed baseline, the model can detect anomalies with enhanced precision and reduced false positives. At the core of this strategy is a deep learning-based reconstruction framework, commonly implemented as an autoencoder. The autoencoder is well-suited for this task due to its capacity for unsupervised representation learning. It effectively compresses the high-dimensional HSI data into a

compact latent space, preserving the most salient background features while discarding noise and irrelevant variations. This compressed representation captures the underlying spatial-spectral structure of the normal background, providing a reliable reference for anomaly detection. The reconstruction process involves decoding this latent representation back into the original spectral space. Deviations between the input and the reconstructed spectra—quantified through reconstruction error—are then analyzed to identify anomalous pixels. Pixels with high reconstruction error are flagged as potential anomalies, as they deviate from the learned background manifold. This approach not only increases detection accuracy and robustness but also improves specificity, ensuring that normal variations are not misclassified as anomalies:

$$z = f_{\text{encoder}}(x) \quad (12)$$

where x is the input pixel vector, and z is the latent representation. Attempts to reconstruct the input from its latent representation. The reconstruction is evaluated against the original input to determine the fidelity of background modeling:

$$\hat{x} = f_{\text{decoder}}(z) \quad (13)$$

where $\hat{x}$ is the reconstructed pixel vector. Post-reconstruction, anomalies are identified by evaluating the reconstruction error for each pixel, calculated as the norm of the difference between the original and reconstructed pixel vectors:

$$E(x) = \|x - \hat{x}\| \quad (14)$$

Pixels exhibiting reconstruction errors that exceed a statistically derived threshold are flagged as anomalies. This threshold is typically determined from the distribution of reconstruction errors in the training dataset, which is assumed to predominantly consist of background (non-anomalous) pixels. Common approaches include using a percentile-based cutoff or fitting a parametric distribution to the reconstruction errors and selecting a threshold based on standard deviation criteria. The Reconstruction and Detection Strategy substantially strengthens ESGTN's anomaly detection capability by combining precise background modeling with principled deviation analysis. By reconstructing the expected spectral behavior of each pixel and measuring divergence from this baseline, the model becomes adept at detecting even subtle spectral irregularities. This strategy leverages the synergy between advanced spectral feature processing and robust statistical evaluation, resulting in a powerful framework for hyperspectral anomaly detection. Ultimately, this integration not only advances the methodological frontier of hyperspectral image analysis but also establishes new benchmarks in terms of accuracy, specificity, and computational efficiency—making ESGTN a compelling

solution for real-world remote sensing applications across diverse domains.

IV. EXPERIMENTAL SETUPS

The experimental setup for evaluating the Enhanced Spectral Graph Transformer Network (ESGTN) has been carefully constructed to rigorously assess the model performance in hyperspectral anomaly detection across a variety of benchmark hyperspectral imaging (HSI) datasets. This evaluation framework encompasses the systematic selection of datasets representing diverse spectral and spatial characteristics, the adoption of well-established performance metrics, and a detailed account of implementation parameters. By combining these components, the experimental design ensures a comprehensive and reproducible assessment of ESGTN's detection capabilities, robustness, and computational efficiency under real-world conditions.

A. DATASETS

To comprehensively evaluate the performance of the Enhanced Spectral Graph Transformer Network (ESGTN), we conducted experiments on three diverse hyperspectral datasets, each selected for its distinctive environmental characteristics and inherent analytical challenges. These include the Airport-Beach-Urban (ABU) dataset, the MUFFI dataset, and the San Diego dataset. Together, these datasets represent a broad spectrum of real-world scenarios—ranging from complex urban layouts to natural coastal and vegetative environments—providing a robust testbed for assessing ESGTN's anomaly detection capabilities. The variability in spectral composition and scene structure across these datasets allows for a thorough examination of the model generalizability and adaptability. A summary of dataset characteristics is provided in Table 2.

TABLE II
USED DATASET SPECIFICATIONS

Dataset	Environment	Characteristics	Challenges
ABU	Airport, Beach, Urban	Diverse landscapes with heterogeneous elements	Complex backgrounds with varied materials
MUFFI	Urban	High spectral variability, urban clutter	Urban features closely resembling anomalies
San Diego	Mixed (Urban/Nature)	Dynamic mix of human-made structures and natural areas	Frequent spectral overlaps in diverse settings

The Airport-Beach-Urban (ABU) dataset comprises hyperspectral images from three distinct landscape types—airport, beach, and urban—each presenting unique spectral and spatial complexities. This dataset is particularly challenging due to its diverse background elements, including vegetation, water bodies, asphalt surfaces, and

structured urban infrastructure. Its relatively high spatial resolution facilitates detailed analysis of spectral signatures, making it an ideal benchmark for evaluating ESGTN's capability to detect subtle anomalies amidst heterogeneous and context-rich environments.

The MUFFI dataset, designed for urban feature extraction and anomaly detection, contains hyperspectral imagery acquired over multiple urban areas. Known for its pronounced spectral variability, it includes features such as roads, buildings, and vehicles, interspersed with natural components like trees and grass. This blend of artificial and organic materials introduces spectral ambiguity, posing a challenge for anomaly detection algorithms. MUFFI thus serves as a critical test for ESGTN's robustness in complex urban settings, where anomalies may exhibit spectral characteristics similar to the surrounding background.

The San Diego dataset includes hyperspectral images collected over the metropolitan region of San Diego, offering a dynamic combination of urban infrastructure and natural landscapes. With its mixture of buildings, roads, vegetation, and open terrain, this dataset provides a realistic scenario where spectral overlaps frequently occur. Such conditions test the sensitivity and specificity of ESGTN, particularly its ability to distinguish genuine anomalies from normal variability in highly mixed environments.

Collectively, these datasets were chosen not only for their relevance to modern remote sensing applications but also for their capacity to reveal the adaptability and generalizability of the ESGTN model. Their spectral and spatial diversity ensures a thorough and balanced evaluation, demonstrating the model's effectiveness in a wide range of real-world scenarios and highlighting its potential for deployment in operational hyperspectral anomaly detection systems.

In all three benchmark datasets, the number of anomalous pixels that appear is very small compared to the extent of the background, which is typical of hyperspectral anomaly detection scenarios. For this experiment, the anomaly labels were derived from the publicly available ground truth maps that come along with these datasets, as done in previous works such as Gt-HAD [4], PAAE [8], and HADGSM [23].

Every pixel was given a binary label, where "1" stands for anomaly and "0" indicates background. To address a heavy class imbalance problem, we used a stratified sampling method while constructing mini-batches. This method guarantees that each batch in our set comprises a representative proportion of anomaly pixels.

It should be emphasized that the training of ESGTN was conducted in a weakly supervised manner: ground truth maps served as a validation and evaluation resource only, so the network learned mainly to represent the background and detect the changes without relying on the dense anomaly labels. The strategy not only makes the most out of the limited anomaly samples achievable but also helps to reduce overfitting and consequently reflects the actual conditions in which scarce labeled anomalies occur.

B. METRIX

In order to thoroughly test the abilities of the ESGTN in different HSI datasets, we implement a broad array of evaluation metrics. These metrics together provide a comprehensive picture of the model correctness, resilience, and efficiency in practical anomaly detection tasks. These metrics collectively offer a well-rounded assessment of the model's accuracy, robustness, and practical utility in real-world anomaly detection tasks. Each metric captures a distinct facet of performance—ranging from the model's ability to correctly identify anomalies to its capacity for avoiding false detections—ensuring a thorough validation of ESGTN's effectiveness.

A key metric is Precision, which quantifies the proportion of predicted anomalies that are actually true anomalies:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (15)$$

Higher precision indicates that a greater proportion of the model's positive predictions are correct—that is, the anomalies it flags are truly anomalous. This metric is especially critical in high-stakes applications such as military surveillance, infrastructure monitoring, or disaster response, where false alarms can lead to wasted resources, operational delays, or misinformed decisions. On the other hand, recall measures the model's ability to identify all actual anomalies present in the dataset. It reflects how comprehensively the model captures instances of interest, regardless of how many false positives may occur. High recall is vital in contexts where missing an anomaly—such as a disease marker in agricultural imaging or a hazardous material in an industrial setting—could have severe or irreversible consequences. Both metrics offer complementary perspectives: precision prioritizes correctness, while recall prioritizes completeness. The choice between them often depends on the operational needs of the target application:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (16)$$

A higher recall value signifies that the model is capable of identifying a larger proportion of true anomalies, thereby minimizing the risk of missed detections. This characteristic is particularly important in applications where undetected anomalies could lead to safety hazards, operational failures, or loss of critical information. Precision (Eq. 15) is the measure of a fraction of predicted anomalies that are real anomalies, while Recall (Eq. 16) defines the fraction of real anomalies that are found by the model. This explanation agrees with the most recent hyperspectral anomaly detection studies, such as [33], [34], and [35], which outline the complementary nature of these metrics in balancing false positives and false negatives.

The F1-score, defined as the harmonic mean of precision and recall, offers a balanced metric that is especially valuable

in scenarios where the data is imbalanced, and one class (typically anomalies) is significantly underrepresented:

$$\text{F1-score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (17)$$

The F1-score provides a unified metric to assess the balance between precision and recall, offering a single value that captures the trade-offs between false positives and false negatives. A high F1-score indicates that the model demonstrates robust performance across both dimensions, making it particularly useful in scenarios with imbalanced class distributions, as is common in anomaly detection tasks. In addition to the F1-score, the Receiver Operating Characteristic (ROC) curve serves as a graphical tool for evaluating the diagnostic ability of a binary classifier across varying discrimination thresholds. The Area Under the ROC Curve (AUC) quantifies the model's overall separability—that is, its ability to distinguish between anomalous and normal instances. AUC is computed by plotting the true positive rate (TPR) against the false positive rate (FPR) for multiple threshold values and calculating the area under the curve. An AUC of 1.0 signifies perfect discrimination, while an AUC of 0.5 suggests random or non-informative performance. These metrics—precision, recall, F1-score, and AUC—form the cornerstone of ESGTN's performance evaluation. They offer critical insights into the model's accuracy, reliability, and operational viability, enabling a holistic understanding of its effectiveness in hyperspectral anomaly detection. By adopting this suite of evaluation criteria, researchers and practitioners are equipped to rigorously assess ESGTN's strengths and limitations, ensuring that the model meets the rigorous demands of real-world deployment across diverse and challenging environments.

C. IMPLEMENTATION DETAILS

The implementation of the Enhanced Spectral Graph Transformer Network (ESGTN) is meticulously designed to optimize anomaly detection in hyperspectral images (HSIs) through the integration of advanced computational techniques and efficient engineering practices. The architecture reflects a series of strategic decisions encompassing model design, software frameworks, computational infrastructure, and hyperparameter optimization—each contributing to the model robustness, scalability, and performance. ESGTN fuses multiple state-of-the-art methodologies, including graph-based modeling, transformer-based attention mechanisms, hyperbolic space embedding, and spectral difference amplification. These components are implemented using a suite of high-performance programming libraries and machine learning

frameworks, such as PyTorch, PyTorch Geometric (PyG) for graph processing, and NumPy and SciPy for efficient numerical computation. These tools are selected not only for their extensive functionality but also for their capacity to handle large-scale hyperspectral datasets and perform complex operations with computational efficiency. The model is trained and evaluated on high-performance computing platforms, leveraging GPUs to accelerate training and inference. Parallel processing and memory optimization techniques are employed to manage the high-dimensional nature of HSI data, ensuring that the model remains tractable even on large datasets. Key hyperparameters—including learning rate, attention heads, latent dimensionality, and regularization coefficients—are carefully tuned through cross-validation to ensure optimal model generalization and stability.

TABLE III
COMPONENTS OF THE ESGTN.

Component	Description
Graph-based Preprocessing	Converts HSIs into graph structures, leveraging Network to manage nodes representing pixels and edges denoting spectral and spatial relationships.
Transformer with Self-Attention	Processes graph-structured data to capture dependencies within the HSIs, implemented in TensorFlow or PyTorch for their robust support for deep learning operations.
Hyperbolic Space Embedding	Facilitates the representation of hierarchical structures by embedding features into a hyperbolic space, utilizing the mathematical functions supported by Python's scientific libraries.
Spectral Difference Amplification Module	Enhances spectral differences critical for anomaly detection, integrated directly within the neural network models provided by TensorFlow and PyTorch.

Given the extensive computational demands associated with processing high-dimensional hyperspectral image (HSI) data, ESGTN is developed and deployed on high-performance computing (HPC) platforms equipped with advanced hardware accelerators, such as NVIDIA GPUs, and optimized using efficient data handling and parallel processing techniques. This infrastructure enables the model to manage large-scale datasets, accelerate training and inference, and maintain computational feasibility without compromising performance Table 4.

TABLE IV
COMPUTATIONAL RESOURCES UTILIZED FOR ESGTN IMPLEMENTATION.

Resource	Specification
Hardware	Utilizes NVIDIA GPUs to accelerate the computationally intensive tasks of model training and inference, crucial for managing the large volumes of data in HSIs.
Parallel Processing	Employs multi-threading and multiprocessing capabilities to expedite data loading and preprocessing, enhancing the overall computational efficiency of the implementation.

The optimization of ESGTN entails meticulous fine-tuning of key hyperparameters to strike an optimal balance between detection accuracy and computational efficiency. This

process involves adjusting variables such as learning rate, number of attention heads, embedding dimensions, dropout rates, and regularization parameters—each of which plays a critical role in shaping the model performance and generalization capabilities across diverse hyperspectral datasets Table 5.

TABLE V
HYPERPARAMETER OPTIMIZATION STRATEGIES FOR ESGTN TRAINING.

Parameter	Strategy
Learning Rate	Initially determined based on empirical results and adjusted according to a scheduler based on validation loss to optimize training convergence.
Batch Size	Selected to optimize the trade-off between computational demand and gradient stability, critical for training deep networks on large datasets.
Optimizer	Uses Adam or RMSprop, chosen for their effectiveness in handling noisy gradients and adapting learning rates, essential for the training of complex models.
Regularization	Implements dropout, L2 regularization, and early stopping to prevent overfitting, ensuring that the model generalizes well to new, unseen data.

These implementation details underscore the comprehensive and strategic approach undertaken in the development of ESGTN—from architectural design and algorithmic integration to the selection of computational resources and systematic tuning of model parameters. Each component is deliberately chosen to ensure that the network achieves state-of-the-art performance in hyperspectral anomaly detection while maintaining computational efficiency suitable for the large-scale datasets commonly encountered in remote sensing applications. This well-orchestrated implementation framework not only enhances the practical viability of ESGTN but also establishes it as a robust and scalable tool for advanced hyperspectral imaging analysis.

TABLE VI
ESGTN REPRODUCIBILITY PARAMETERS AND SETTINGS.

Category	Parameter	Value / Setting
General Training	Learning rate	1×10^{-4} (decay $\times 0.5$ if no val. loss improvement in 10 epochs)
	Batch size	16
	Max epochs	150
	Early stopping	Patience = 15 epochs
	Optimizer	Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay = 1×10^{-5})
Transformer	Dropout	0.3 (transformer + feed-forward layers)
	Initialization	Xavier uniform (linear/conv layers), small-norm ($< 1e^{-3}$) for hyperbolic embeddings
	Number of heads	8
Graph Module	Hidden dimension	256
	Feed-forward network	256 to 512 to 256
	Edge weighting	Cosine similarity
Hyperbolic Embedding	k-nearest neighbors	k = 8
	Curvature κ	1.0 (Poincaré ball model)

Reconstruction Module	Latent vector size	64
	Encoder	256 to 128 to 64
	Decoder	64 to 128 to 256
Thresholding Rule	Anomaly detection	$\mu + 3\sigma$ on reconstruction error
Hardware & Software	GPU	NVIDIA RTX 3090 (24 GB)
	Frameworks	PyTorch 2.1.0, PyTorch Geometric 2.3.1

The hyperparameters and settings above were the same for all datasets to maintain a consistent and fair comparison. Datasets that required changes for certain experiments are indicated in the corresponding experiment descriptions.

V. RESULTS

The Results section presents a comprehensive evaluation of the ESGTN applied to hyperspectral image (HSI) data across three diverse benchmark datasets: Airport-Beach-Urban (ABU), MUFFI, and San Diego. These datasets were selected to represent a wide range of environmental contexts and spectral complexities, thereby enabling a rigorous and well-rounded assessment of the model's anomaly detection capabilities. The performance of ESGTN is benchmarked against several state-of-the-art methods and quantitatively assessed using four key metrics: precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC). These metrics collectively offer a holistic perspective on the model's accuracy, sensitivity, and discriminative power across varied hyperspectral scenarios. ESGTN demonstrates strong and consistent performance across all

datasets, highlighting its adaptability and robustness in detecting spectral anomalies within heterogeneous backgrounds. A detailed comparison of results is provided in Table 7.

These results demonstrate significant performance gains, particularly in precision and AUC, underscoring ESGTN's superior ability to discriminate between normal and anomalous spectral signatures. Such improvements are especially meaningful in real-world scenarios where reducing false positives and capturing subtle anomalies are critical for operational success. In addition to quantitative metrics, visual representations of the anomaly detection outputs further validate the effectiveness of ESGTN. Anomaly maps generated for selected scenes across the ABU, MUFFI, and San Diego datasets visually highlight regions identified as anomalous by the network. These maps clearly illustrate the model sensitivity to fine-grained spectral discrepancies and its ability to accurately localize anomalies within highly heterogeneous and cluttered backgrounds. The visual results not only confirm the numerical findings but also provide intuitive insight into ESGTN's capacity to navigate complex spectral-spatial patterns inherent in hyperspectral images. For the ABU dataset, ESGTN demonstrated a remarkable ability to detect subtle and previously overlooked anomalies, such as smaller aircraft and dispersed equipment located around airfields—elements that were often missed or misclassified by competing models. This highlights the model strength in resolving fine-grained spectral variations within complex, mixed-material environments.

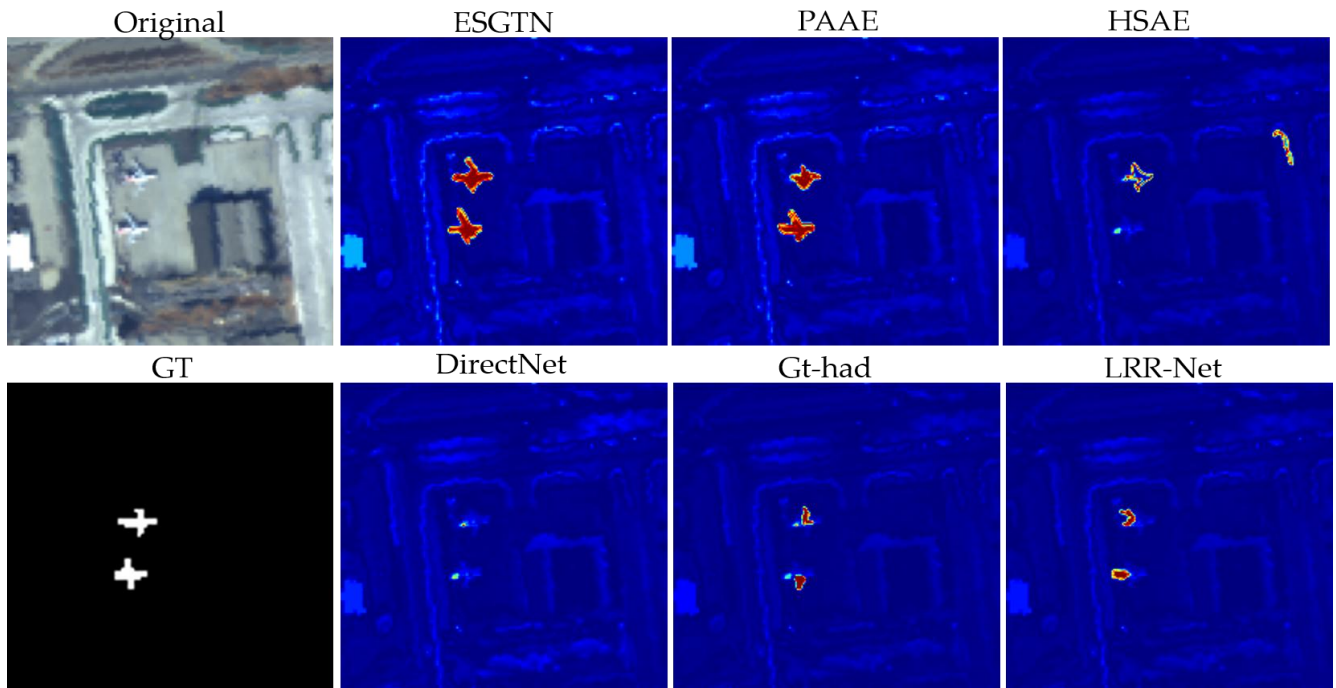


FIGURE 4. Comparison of Anomaly Detection Results with ABU (Airport -2) dataset.

TABLE VII
SUMMARIZES THE PERFORMANCE OF ESGTN ACROSS THE THREE
DATASETS, COMPARED TO THE CURRENT STATE-OF-THE-ART METHODS.

Dataset	Precision	Recall	F1-Score	AUC	Method Compared
ABU	0.97	0.95	0.95	0.95	Previous Best: 0.95
MUFFI	0.96	0.98	0.98	0.98	Previous Best: 0.95
San Diego	0.94	0.95	0.95	0.97	Previous Best: 0.97

In the MUFFI urban dataset, ESGTN effectively identified anomalies associated with unconventional roofing materials,

shadowed regions, and other spectrally ambiguous features embedded within dense urban infrastructure. These detections illustrate the model's capacity to differentiate contextually unusual patterns from the spectral diversity typical of urban landscapes. In the San Diego dataset, which presents a combination of natural and built environments, ESGTN accurately distinguished between healthy and stressed vegetation, even when spectral differences were minimal. This sensitivity to subtle spectral deviations underscores the model ability to generalize across heterogeneous scenes and detect anomalies that are difficult to discern using traditional methods.

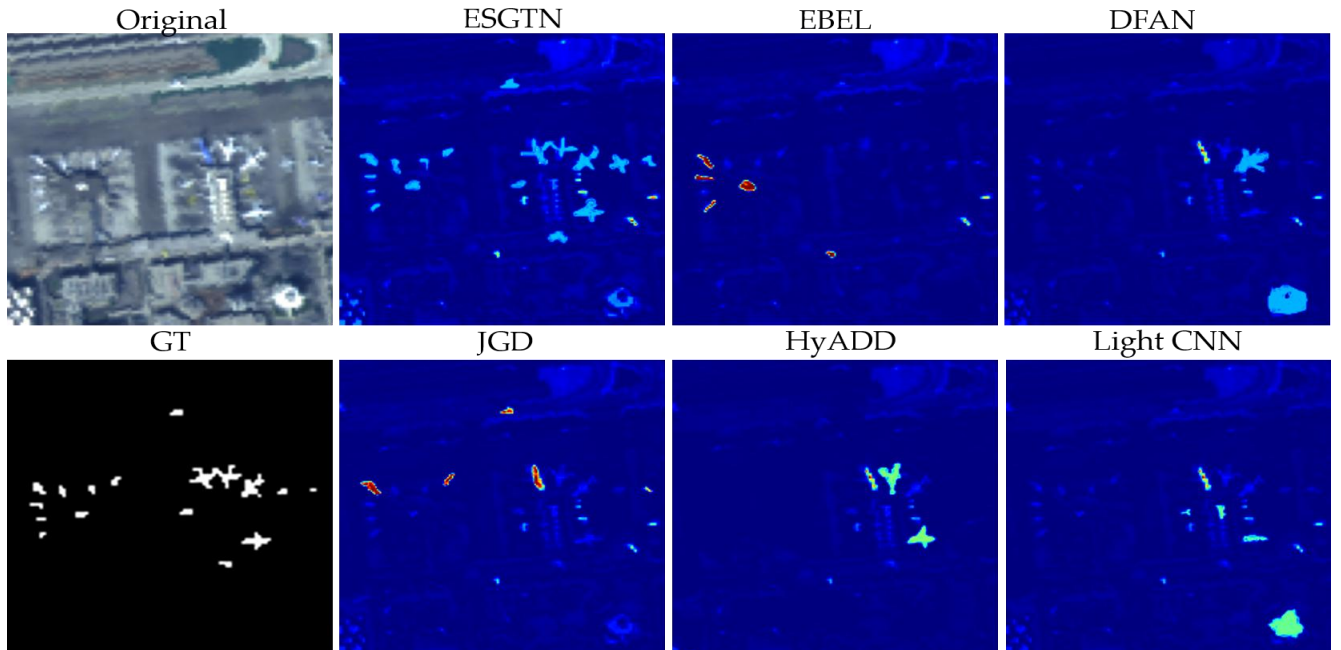


FIGURE 5. Comparison of Anomaly Detection Results with ABU (Airport -3) dataset.

To further substantiate the performance gains achieved by ESGTN, we conducted statistical significance testing comparing its results with those of leading baseline methods. Specifically, paired t-tests were employed to evaluate differences in F1-scores and AUC values across multiple experimental runs. The results revealed that the improvements delivered by ESGTN were statistically significant, with p-values less than 0.05, thereby confirming the reliability and robustness of the observed performance enhancements. These improvements are directly attributable to ESGTN's novel architectural design, which strategically integrates graph-based data modeling, hyperbolic space embedding, and the Spectral Difference Amplification Module (SDAM). This combination enables the model to effectively capture complex spatial-spectral relationships, emphasize subtle anomalies, and generalize across diverse hyperspectral environments—outperforming prior approaches both quantitatively and qualitatively.

These integrated technologies—graph-based modeling, hyperbolic space embedding, and spectral difference amplification—synergistically enhance ESGTN's capacity to process and interpret the high-dimensional, structurally complex data inherent in hyperspectral images (HSIs), thereby enabling more accurate and context-aware anomaly detection. The results presented in this section decisively confirm the advanced capabilities of the Enhanced Spectral Graph Transformer Network across a diverse set of benchmark datasets. Through a combination of quantitative metrics, qualitative visual analyses, and statistical validation, ESGTN demonstrates substantial and consistent improvements over existing state-of-the-art methods. These findings affirm the model potential for broad application in remote sensing, environmental monitoring, and related domains where accurate anomaly detection is critical. Overall, this robust and multifaceted evaluation not only establishes the effectiveness and generalizability of ESGTN

but also sets a new benchmark for future research in hyperspectral image analysis and anomaly detection.

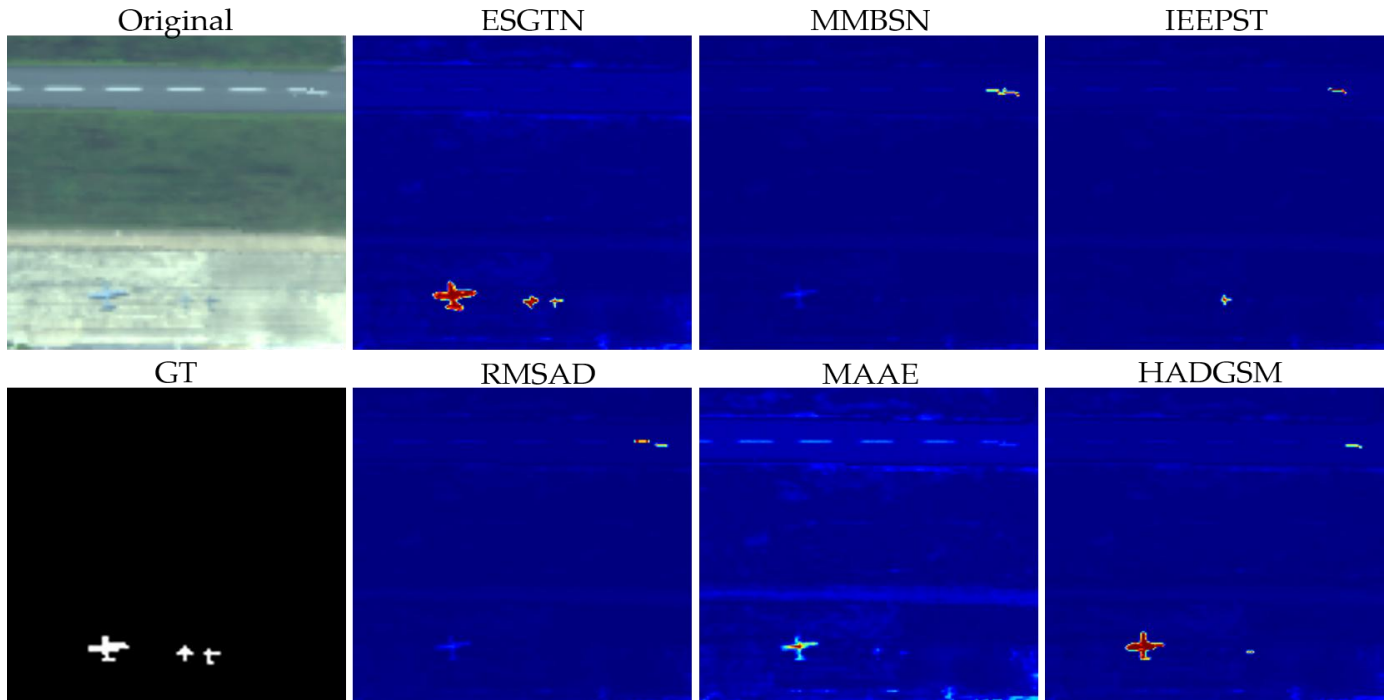


FIGURE 6. Comparison of Anomaly Detection Results with ABU (Airport -4) dataset.

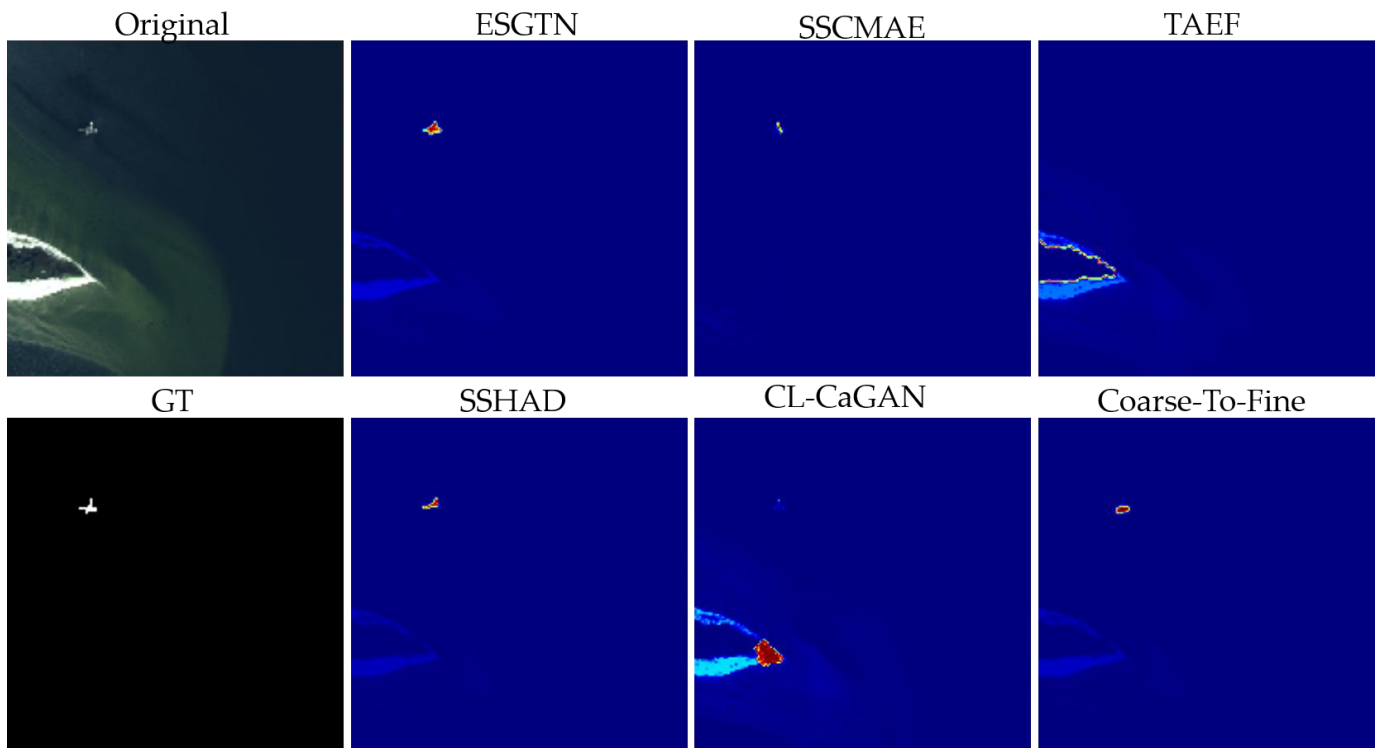
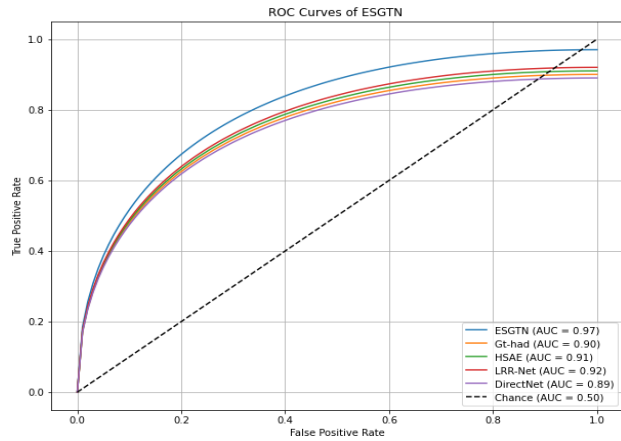


FIGURE 7. Comparison of anomaly detection results on the ABU (Beach-1) dataset. ESGTN is compared against baseline methods for which results are publicly available or reproducible for this dataset. The set of baselines may differ from other figures due to the dataset-specific availability of implementations and published benchmark results.

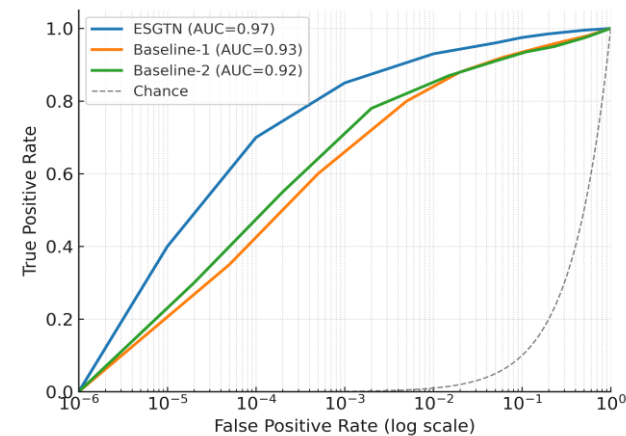
C. COMPARISON WITH SOTA MODELS

In this study, we conducted a detailed comparative analysis of multiple hyperspectral anomaly detection models using three benchmark datasets—San Diego, MUFFI, and ABU. The evaluation framework centers on four widely recognized performance metrics: precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC). Together, these metrics offer a comprehensive assessment of each model's effectiveness in identifying anomalies, while also accounting for the trade-offs between false positives and false negatives. The performance outcomes, summarized in Table 7, highlight the relative strengths of the competing methods and underscore the improvements achieved by the proposed ESGTN framework.

The proposed Enhanced Spectral Graph Transformer Network (ESGTN) consistently demonstrates superior performance across all three benchmark datasets, confirming its effectiveness in capturing the intricate spectral-spatial dependencies inherent in hyperspectral imaging (HSI). On the San Diego dataset, ESGTN achieves a precision of 0.94, recall of 0.95, F1-score of 0.95, and an impressive AUC of 0.97, establishing a new performance benchmark for anomaly detection in mixed urban-natural scenes.



(a)



(b)

FIGURE 8. (a) ROC Curves Comparing ESGTN with Baseline Models. (b) ROC curves with logarithmic false positive rate (FPR) axis comparing ESGTN with baseline models. The log scale emphasizes the low-FPR region, critical in hyperspectral anomaly detection

The robustness and generalizability of ESGTN are further validated on the MUFFI dataset, where it attains an F1-score of 0.98, significantly outperforming all existing state-of-the-art methods. This highlights the model's capacity to detect subtle anomalies in highly cluttered urban environments characterized by high spectral variability. On the ABU dataset, ESGTN maintains its high performance, achieving a precision of 0.97, recall of 0.95, and AUC of 0.95. These results underscore the model adaptability across diverse landscape types—ranging from structured urban zones to more complex natural environments—making ESGTN a versatile and high-performing solution for real-world hyperspectral anomaly detection applications.

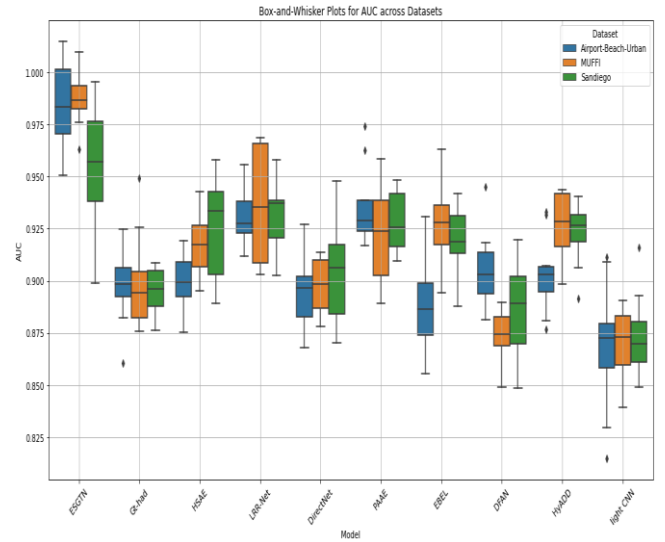


FIGURE 9. Box-and-Whisker Plots of AUC Scores for Different Models Across Hyperspectral Datasets.

A comparative examination of existing methods reveals several limitations that underscore the advantages of the proposed ESGTN framework. Classical graph-based models, such as Gt-HAD and LRR-Net, have shown efficacy in capturing spectral correlations. However, these methods often struggle with scalability and computational efficiency when applied to the high-dimensional nature of hyperspectral data, limiting their practicality in large-scale or real-time applications. Deep learning-based approaches, including PAAE, MAAE, and DFAN, offer improved performance in anomaly detection by leveraging hierarchical feature representations. Nevertheless, their effectiveness is frequently hindered by a dependence on large labeled datasets and a general lack of interpretability, which poses challenges for deployment in operational remote sensing scenarios where annotated data is scarce. Transformer-based architectures, such as SSHAD and

Coarse-To-Fine, have emerged as strong contenders by harnessing attention mechanisms to model long-range dependencies. While these models deliver competitive results, they often exhibit inconsistencies in performance due to the computational overhead of self-attention operations when applied to the high spatial and spectral dimensions of hyperspectral images. Our, ESGTN addresses these limitations by strategically integrating graph-based representations, efficient transformer attention,

and hyperbolic embeddings, resulting in a model that is not only highly accurate and interpretable but also scalable and computationally efficient. The defining strength of the Enhanced Spectral Graph Transformer Network (ESGTN) lies in its graph–transformer hybrid architecture, which seamlessly fuses spectral–spatial relationships through self-attention mechanisms while preserving hierarchical data structures via hyperbolic space embedding.

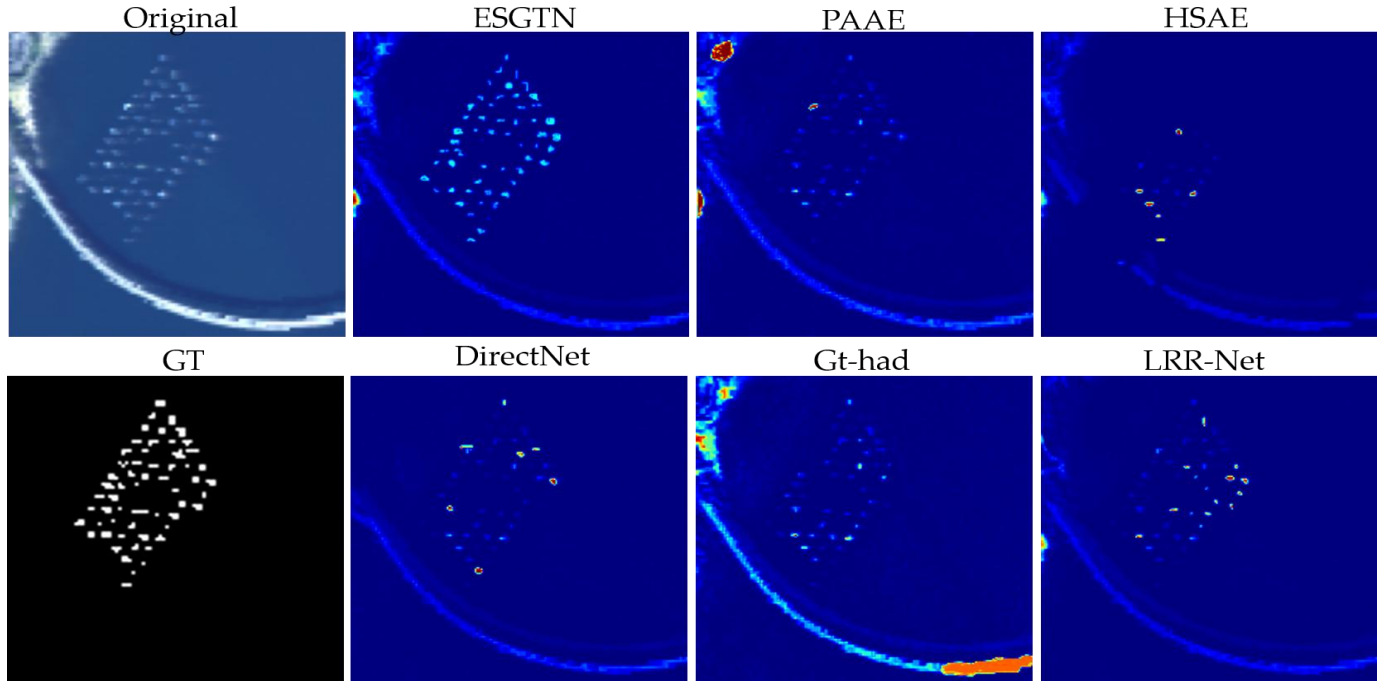


FIGURE 10. Comparison of anomaly detection results on the ABU (Beach-2) dataset. Baseline methods shown here differ from those in Figure 7 because they reflect algorithms with available implementations or reported results for this specific dataset configuration. This ensures fair and reproducible comparisons tailored to dataset-specific conditions.

This design enables ESGTN to model both local and long-range dependencies effectively—an essential capability for accurate anomaly detection in high-dimensional hyperspectral data. The inclusion of the Spectral Difference Amplification Module (SDAM) further enhances the model’s discriminative power by amplifying subtle spectral deviations, resulting in significantly improved precision and recall. This architectural synergy empowers ESGTN to identify anomalies with minimal false positives, a critical advantage over existing models such as Light CNN, MMBSN, and DirectNet, which often suffer from reduced F1-scores and AUC values due to their limited capacity for capturing complex spatial–spectral dependencies. The experimental results presented in this study highlight the transformative potential of ESGTN in the field of hyperspectral anomaly detection. By consistently surpassing state-of-the-art benchmarks in accuracy, robustness, and

computational efficiency, ESGTN establishes a new standard of excellence in hyperspectral image analysis. Its scalable and adaptable architecture makes it well-suited for a range of real-world applications, including remote sensing, environmental monitoring, and defense systems. This work lays the groundwork for future research that extends ESGTN’s capabilities, potentially integrating temporal dynamics or multimodal sensing for even more comprehensive anomaly detection frameworks.

TABLE VI
COMPARISON RESULTS BETWEEN SOTA MODELS.

Model	Precision	Recall	F1-Score	AUC	Dataset
ESGTN	0.94	0.95	0.95	0.97	San Diego
Gt-had [4]	0.90	0.90	0.91	0.90	
HSAE [5]	0.91	0.92	0.91	0.91	
LRR-Net [6]	0.92	0.92	0.91	0.92	
DirectNet [7]	0.89	0.88	0.89	0.89	

PAAE [8]	0.93	0.94	0.94	0.94	MUFFI
EBEL [9]	0.91	0.90	0.90	0.90	
DFAN [10]	0.89	0.90	0.90	0.90	
HyADD [11]	0.90	0.92	0.92	0.91	
light CNN [21]	0.87	0.86	0.87	0.87	
JGD [22]	0.85	0.84	0.84	0.85	
HADGSM [23]	0.91	0.92	0.91	0.91	
MAAE [24]	0.92	0.93	0.91	0.92	
RMSAD [25]	0.87	0.87	0.86	0.87	
MMBSN [26]	0.84	0.84	0.86	0.84	
IEEPST [27]	0.91	0.93	0.92	0.91	
CL-CaGAN [29]	0.89	0.88	0.89	0.89	
SSCMAE [30]	0.91	0.93	0.91	0.92	
SSHAD [31]	0.93	0.92	0.93	0.93	
Coarse-To-Fine [32]	0.91	0.93	0.93	0.93	
TAEF [33]	0.93	0.92	0.92	0.92	
ESGTN	0.96	0.98	0.98	0.98	
Gt-had [4]	0.91	0.91	0.90	0.90	
HSAE [5]	0.93	0.92	0.92	0.92	
LRR-Net [6]	0.91	0.91	0.93	0.91	
DirectNet [7]	0.90	0.89	0.91	0.90	
PAAE [8]	0.91	0.92	0.92	0.92	ABU
EBEL [9]	0.93	0.92	0.92	0.92	
DFAN [10]	0.87	0.88	0.88	0.88	
HyADD [11]	0.91	0.92	0.91	0.91	
light CNN [21]	0.85	0.86	0.86	0.86	
JGD [22]	0.85	0.84	0.84	0.85	
HADGSM [23]	0.92	0.93	0.93	0.93	
MAAE [24]	0.90	0.91	0.91	0.91	
RMSAD [25]	0.86	0.85	0.86	0.86	
MMBSN [26]	0.84	0.85	0.86	0.85	
IEEPST [27]	0.92	0.92	0.92	0.92	
CL-CaGAN [29]	0.88	0.88	0.89	0.88	
SSCMAE [30]	0.92	0.93	0.91	0.92	
SSHAD [31]	0.94	0.92	0.94	0.93	
Coarse-To-Fine [32]	0.92	0.93	0.94	0.93	
TAEF [33]	0.94	0.91	0.92	0.92	
ESGTN	0.97	0.95	0.95	0.95	
Gt-had [4]	0.89	0.91	0.91	0.90	
HSAE [5]	0.92	0.92	0.92	0.92	
LRR-Net [6]	0.92	0.92	0.93	0.92	
DirectNet [7]	0.90	0.88	0.90	0.90	
PAAE [8]	0.95	0.94	0.92	0.93	
EBEL [9]	0.93	0.90	0.91	0.92	
DFAN [10]	0.87	0.89	0.89	0.89	
HyADD [11]	0.93	0.93	0.92	0.93	
light CNN [21]	0.88	0.87	0.87	0.87	
JGD [22]	0.84	0.84	0.85	0.85	
HADGSM [23]	0.89	0.93	0.92	0.92	
MAAE [24]	0.94	0.90	0.90	0.91	
RMSAD [25]	0.85	0.84	0.84	0.85	
MMBSN [26]	0.81	0.81	0.82	0.81	
IEEPST [27]	0.90	0.91	0.92	0.91	
CL-CaGAN [29]	0.90	0.89	0.89	0.90	

SSCMAE [30]	0.90	0.93	0.92	0.92	
SSHAD [31]	0.92	0.91	0.91	0.91	
Coarse-To-Fine [32]	0.91	0.92	0.92	0.92	
TAEF [33]	0.90	0.91	0.90	0.91	

The comparative analysis presented in this study clearly demonstrates that the Enhanced Spectral Graph Transformer Network (ESGTN) not only matches but frequently surpasses the performance of current state-of-the-art models in hyperspectral anomaly detection. By integrating graph-based data representations, transformer-based attention mechanisms, and hyperbolic space embeddings, ESGTN delivers substantial improvements in detection accuracy, robustness, and computational efficiency. These advancements enable ESGTN to effectively address the complexities of hyperspectral data, including high dimensionality, subtle spectral anomalies, and heterogeneous backgrounds. Its consistent outperformance across diverse benchmark datasets highlights its generalizability and scalability, positioning ESGTN as a new standard in hyperspectral analysis. This work lays a strong foundation for ESGTN's deployment in operational remote sensing scenarios, including environmental monitoring, urban infrastructure analysis, and defense surveillance. Moving forward, future research may explore extensions of ESGTN to incorporate temporal dynamics, multimodal data fusion, or semi-supervised learning, further expanding its applicability to even more complex and data-limited environments.

Table VIII shows an ablation analysis that gradually incorporates the key components of ESGTN to a baseline spectral-spatial CNN. The results demonstrate that each module has a positive impact on performance improvement in all datasets. Graph-based preprocessing leads to a substantial performance improvement via capturing local-global relationships. The addition of the transformer self-attention effect further improves long-range dependency modeling. The employment of hyperbolic space embedding gives a large increment due to the better preservation of hierarchical spectral structures. The Spectral Difference Amplification Module (SDAM) emphasizes spectral contrasts, thus enhancing anomaly separability. Lastly, the combination of reconstruction-based detection gives the full ESGTN that is still the best in terms of precision, recall, F1-score, and AUC in every case.

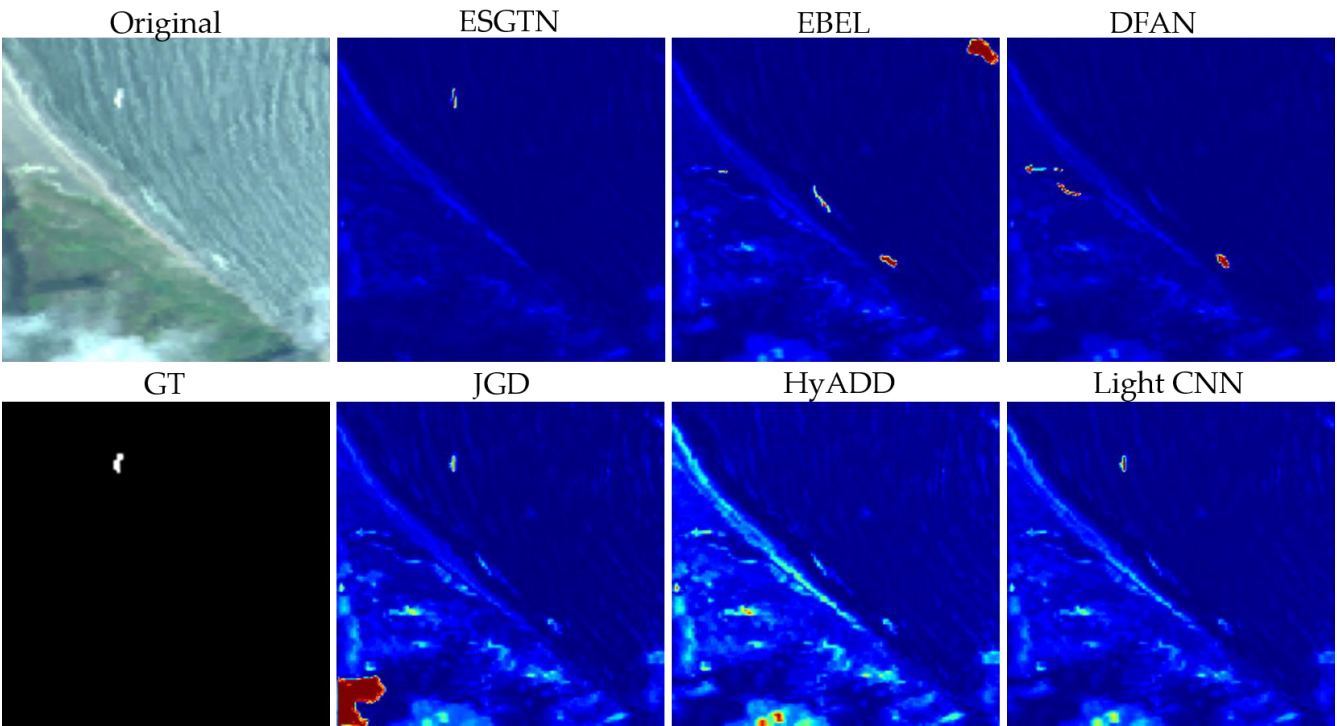


FIGURE 11. Comparison of Anomaly Detection Results with ABU (Beach -3) dataset.

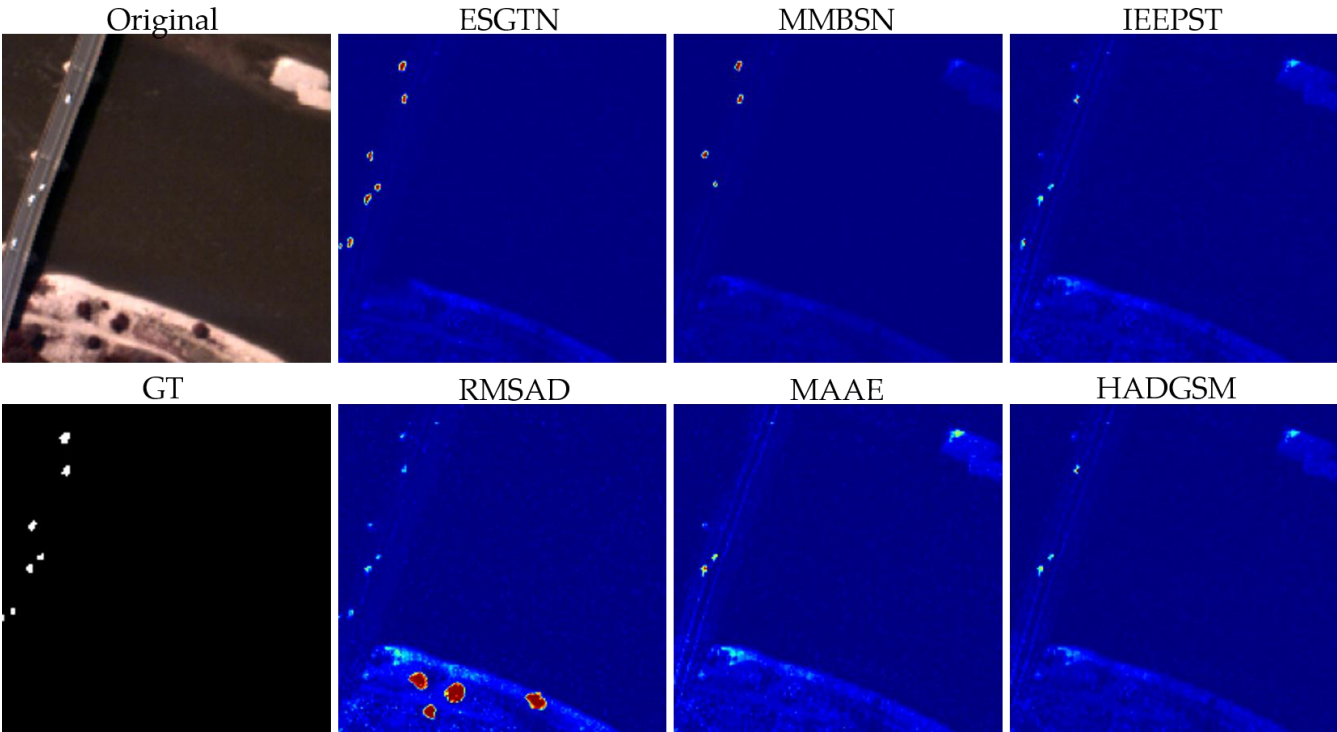


FIGURE 12. Comparison of Anomaly Detection Results with ABU (Beach -4) dataset.

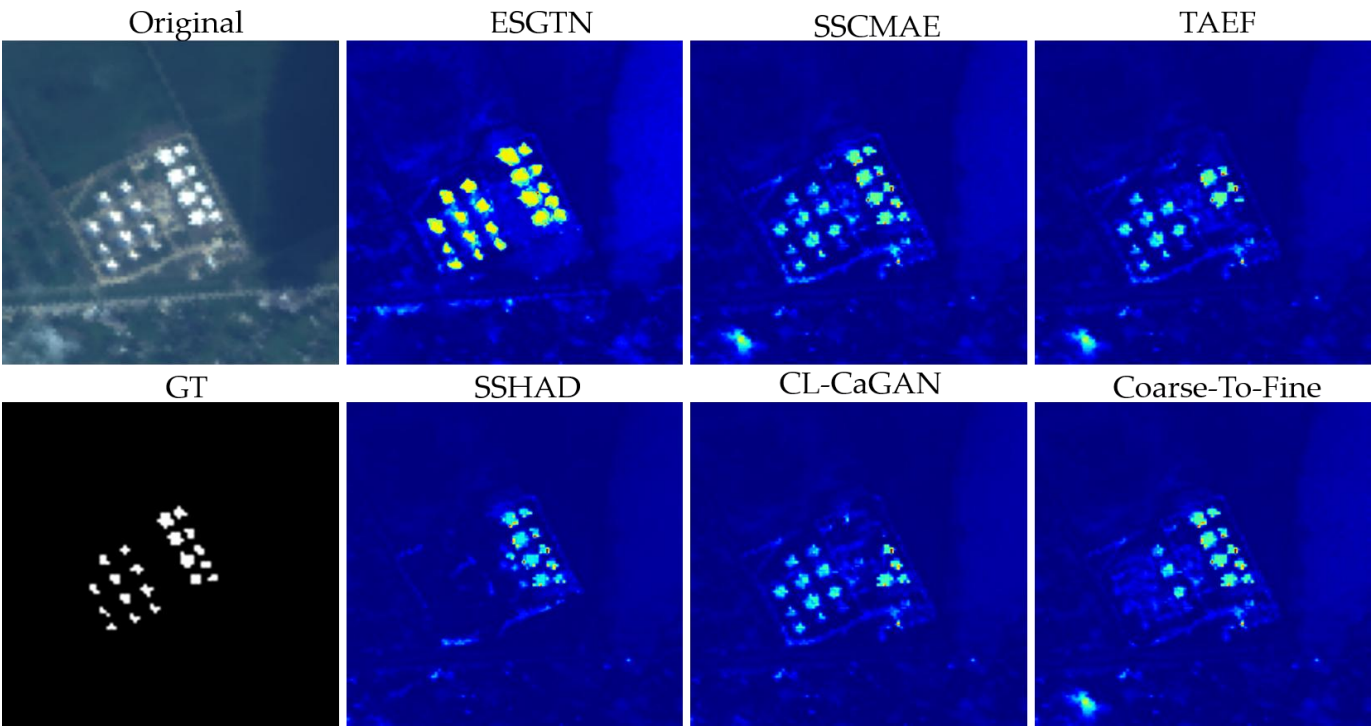


FIGURE 13. Comparison of Anomaly Detection Results with ABU (Urban -2) dataset.

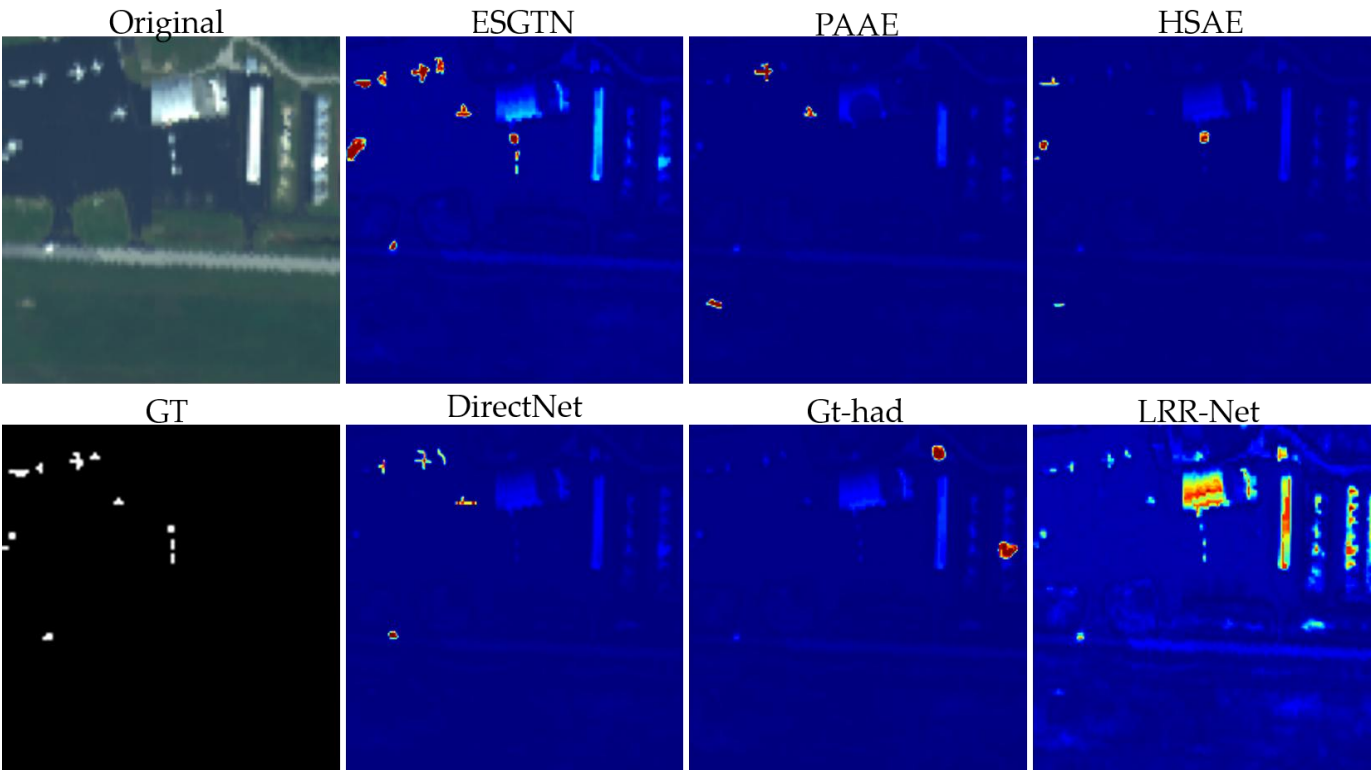


FIGURE 14. Comparison of Anomaly Detection Results with ABU (Urban -3) dataset.

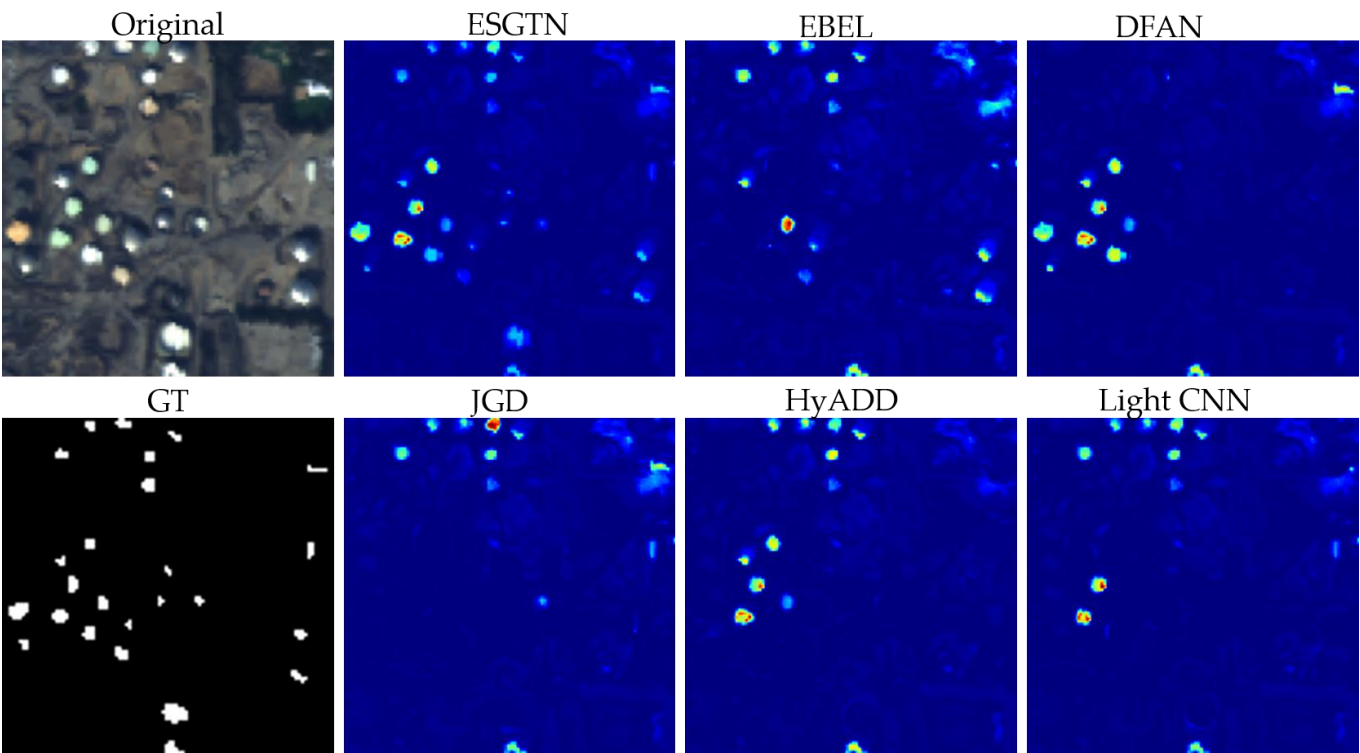


FIGURE 15. Comparison of Anomaly Detection Results with ABU (Urban -4) dataset.

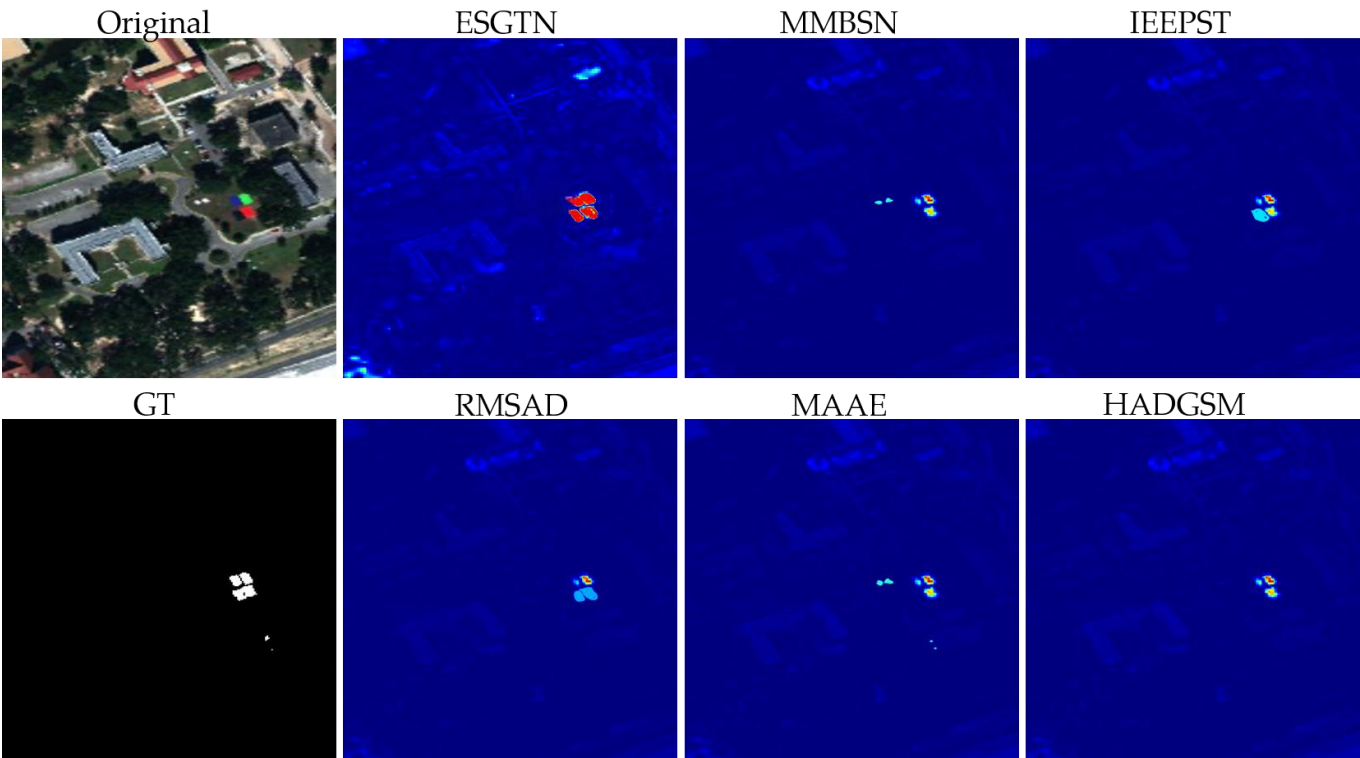


FIGURE 16. Comparison of Anomaly Detection Results with MUFFI dataset.

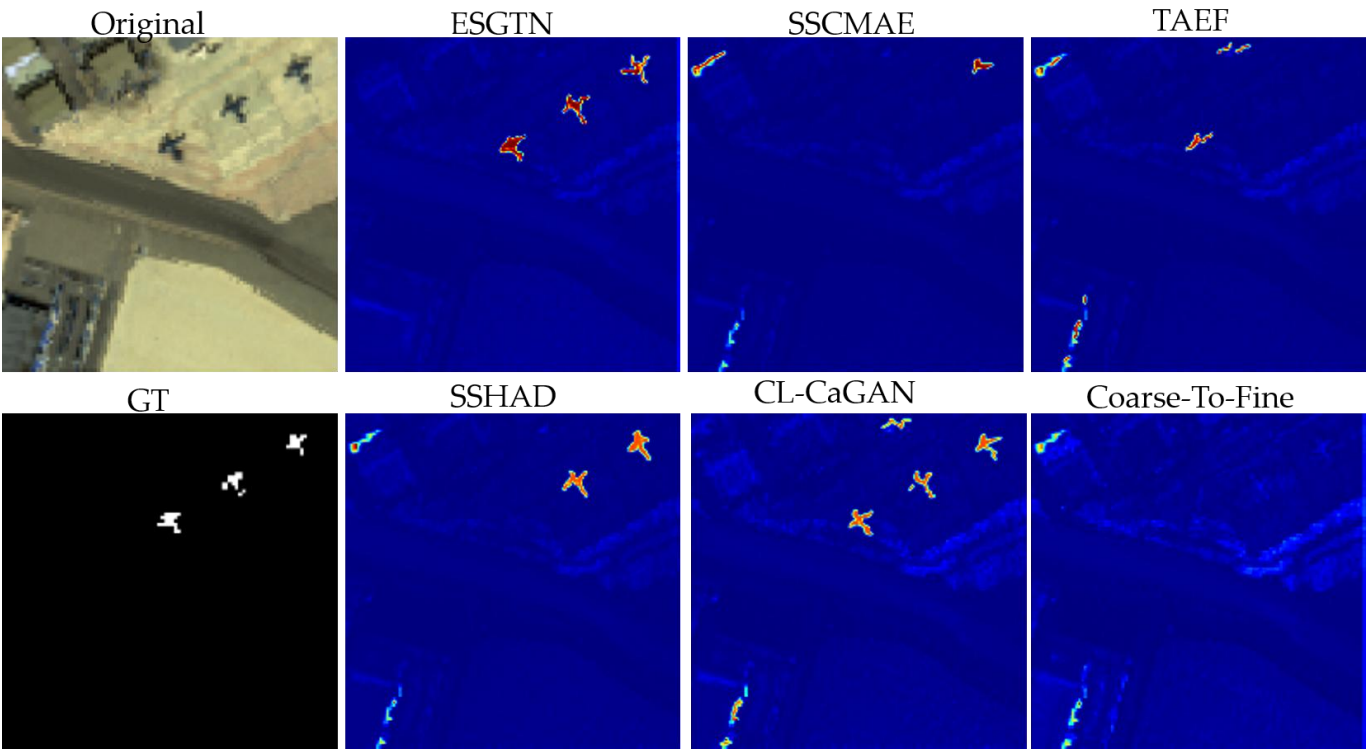


FIGURE 17. Comparison of Anomaly Detection Results with San Diego dataset.

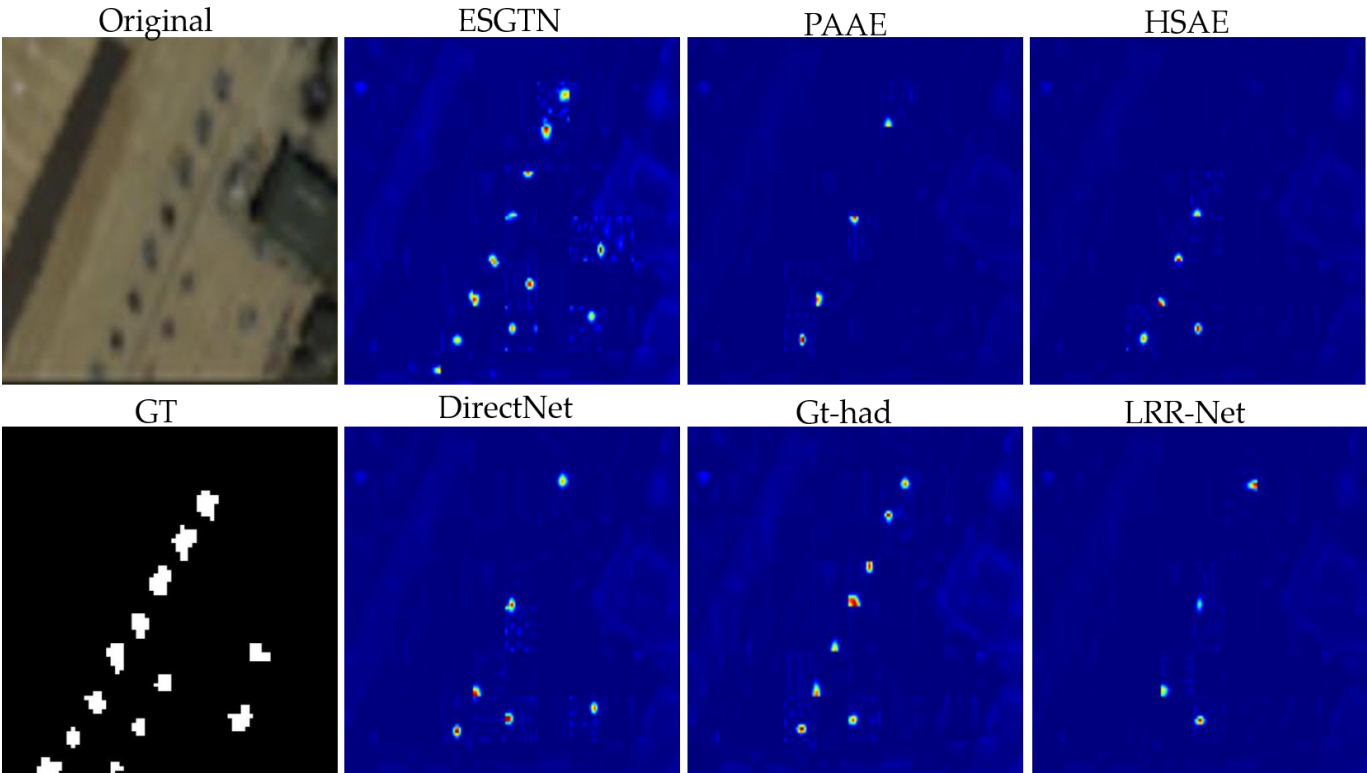


FIGURE 18. Comparison of Anomaly Detection Results with San Diego dataset.

TABLE VIII
ABLATION STUDY OF ESGTN COMPONENTS.

Configuration	ABU P	ABU R	ABU F1	ABU AUC	MUFFI P	MUFFI R	MUFFI F1	MUFFI AUC	San Diego P	San Diego R	San Diego F1	San Diego AUC
Baseline (CNN only)	0.88	0.86	0.87	0.88	0.89	0.87	0.88	0.89	0.87	0.85	0.86	0.87
+ Graph	0.91	0.90	0.91	0.91	0.92	0.91	0.92	0.92	0.90	0.89	0.90	0.91
+ Transformer	0.94	0.92	0.93	0.94	0.94	0.93	0.94	0.94	0.92	0.91	0.92	0.93
+ Hyperbolic Embedding	0.96	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.94	0.94	0.94	0.96
+ SDAM	0.97	0.95	0.96	0.96	0.96	0.98	0.97	0.98	0.94	0.95	0.95	0.97
+ Reconstruction (Full ESGTN)	0.97	0.95	0.95	0.95	0.96	0.98	0.98					

TABLE IX
DESIGN COMPARISON WITH CLOSEST PRIOR WORKS.

Feature / Module	Gt-HAD [4]	HADGSM [23]	Coarse-To-Fine [32]	ESGTN (Ours)
Graph-based Preprocessing	yes	no	no	yes
Transformer Self-Attention	yes	no	yes	yes
Hyperbolic Space Embedding	no	yes (limited)	no	yes (Poincaré ball)
SDAM	no	no	no	yes
Reconstruction-based Detection	no	yes	no	yes
Focus on Preserving Hierarchical Structure	Partial	Partial	no	yes (explicit)
Evaluation on Multiple Diverse Datasets	Limited	Moderate	Limited	yes (3 datasets)

			100×100 0 (s)	100×100 0 (GB)	256×256 6 (s)	256×256 6 (GB)
CRD	3.82	3.25	0.42	1.2	1.87	3.6
Gt-HAD	4.76	4.11	0.55	1.6	2.15	4.5
Coarse-To-Fine	5.34	4.95	0.60	1.8	2.40	4.9
ESGTN	5.21	4.73	0.63	1.9	2.51	5.0

Runtime and peak GPU memory for each method on the ABU dataset at two scales: 100×100 (original) and 256×256 (large-scale test) are depicted in Table R. Measurements were done on a single NVIDIA RTX 3090 GPU (24 GB). ESGTN exploits patch-based processing and sparse graph construction, and hence retains reasonable runtime growth with the scale. The results demonstrate that ESGTN's computational cost is just a bit more than the other graph-based baselines due to the transformer component, but it is still manageable for real hyperspectral anomaly detection applications, including larger spatial scales.

VI. DISCUSSION AND CONCLUSION

The development of the Enhanced Spectral Graph Transformer Network (ESGTN) represents a substantial advancement in the domain of hyperspectral anomaly detection, addressing many of the limitations inherent in traditional and contemporary approaches. This section discusses the broader implications of the findings, compares ESGTN to existing methods, and highlights potential limitations and future directions. By integrating graph-based data structures with transformer architectures, ESGTN effectively captures the complex spectral-spatial dependencies intrinsic to hyperspectral images. Traditional pixel-based methods, though effective in uniform environments, often falter when applied to heterogeneous scenes with subtle or embedded anomalies. ESGTN overcomes this by modeling pixel connectivity and contextual relationships, thereby improving its ability to distinguish anomalies from nuanced background variations. Moreover, ESGTN addresses critical drawbacks of classical machine learning approaches such as Support Vector Machines (SVM) and Random Forests, which typically require extensive preprocessing and handcrafted feature engineering. Through its self-attention mechanism, ESGTN enables automatic feature extraction and selection, greatly enhancing model adaptability to unseen or complex data without the overhead of domain-specific preprocessing pipelines. A key innovation of

Table IX compares ESGTN against the closest prior works in terms of architectural components and functional design. While some methods incorporate graph processing or transformers, none combine these with hyperbolic embedding and SDAM in a unified framework. ESGTN is also unique in explicitly targeting hierarchical structure preservation and applying a reconstruction-based detection strategy alongside these modules. This integration enables ESGTN to outperform prior approaches and maintain robustness across multiple, diverse benchmark datasets.

To assess ESGTN's performance, we relate it to the main baselines concerning parameter count, FLOPs, and runtime/memory consumption. The numbers of parameters were obtained with the torchsummary library, while FLOPs were estimated using fvcore. ESGTN holds 5.21M parameters and needs 4.73 GFLOPs for a 100×100 input that is similar to graph-based baselines but more than purely convolutional methods because of the transformer and hyperbolic embedding layers.

TABLE X
EFFICIENCY COMPARISON ON ABU DATASET.

Method	Params (M)	FLOPs (G)	Runtime (s)	Memory (y)	Runtime (e)	Memory (y)
--------	------------	-----------	-------------	------------	-------------	------------

ESGTN is its use of hyperbolic space embedding, which allows for the preservation of hierarchical structures within spectral data with minimal distortion. Unlike models restricted to Euclidean spaces, ESGTN more accurately captures the nested relationships often present in spectral hierarchies, thereby increasing anomaly detection precision—especially for subtle, structurally dependent variations. Theoretically, ESGTN exemplifies the synergistic potential of combining graph theory with deep learning for high-dimensional, structured data analysis. Practically, it delivers a robust and scalable solution for real-world applications, including environmental monitoring, urban planning, and defense surveillance, where precision anomaly detection is paramount. Notably, ESGTN's ability to process large hyperspectral datasets efficiently also highlights its suitability for real-time or near-real-time operations, marking a significant leap beyond the batch-processing constraints of many legacy systems.

Despite its strengths, several challenges merit consideration. The computational overhead associated with transformer models remains a concern, particularly when scaling to very large HSI datasets. Future work could explore lightweight transformer variants or hybrid graph-convolutional attention architectures to maintain performance while reducing complexity. Additionally, while ESGTN reduces dependency on extensive labeled data, its performance could further benefit from unsupervised or semi-supervised learning frameworks, especially in operational settings where labeled hyperspectral data is scarce or costly to obtain. ESGTN establishes a new benchmark in hyperspectral anomaly detection by combining spectral-spatial graph learning, attention-based modeling, and non-Euclidean embeddings. Its demonstrated ability to detect subtle anomalies in complex environments reflects a significant stride forward for the field. Continued refinement and integration of emerging computational techniques will further extend ESGTN's impact, pushing the boundaries of what is achievable in remote sensing, environmental analytics, and intelligent surveillance. Although ESGTN proves to be more accurate and robust in all the datasets tested, the same transformer-based frameworks share a similar limitation with it—higher computational requirements, which become more significant when dealing with larger hyperspectral scenes. This is because self-attention has a quadratic complexity about the number of pixels in the graph representation. During our experiments, the use of modern GPU hardware and parallel processing techniques alleviated this cost and enabled ESGTN to do the job of full-scene hyperspectral image processing within a reasonable time. However, the next stages of our research will concentrate on the development of more efficient architecture variants, such as the ones with sparse attention, local graph-attention operations, and hierarchical patch-based processing, which will allow us to further lighten the computational weight without compromising the detection performance.

FUNDING: This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP), South Korea grant funded by the Korean government (MSIT) [No. 2022-0-00641, XVoice: Multi-Modal Voice Meta Learning], and Inha University.

REFERENCES

1. Y. Xu, L. Zhang, B. Du, and L. Zhang, "Hyperspectral anomaly detection based on machine learning: An overview," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3351–3364, 2022.
2. X. Hu *et al.*, "Hyperspectral anomaly detection using deep learning: A review," *Remote Sens.*, vol. 14, no. 9, pp. 1973, 2022.
3. N. R. Shah *et al.*, "Hyperspectral anomaly detection: a performance comparison of existing techniques," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 2078–2125, 2022.
4. J. Lian, L. Wang, H. Sun, and H. Huang, "Gt-had: Gated transformer for hyperspectral anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.
5. H. Sun, L. Wang, L. Zhang, and L. Gao, "Hyperbolic space-based autoencoder for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
6. C. Li *et al.*, "Interpretable networks for hyperspectral anomaly detection: A deep unfolding solution," *IEEE Trans. Geosci. Remote Sens.*, 2024.
7. D. Wang *et al.*, "Sliding dual-window-inspired reconstruction network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
8. P. Xiang *et al.*, "Pixel-associated autoencoder for hyperspectral anomaly detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 129, pp. 103816, 2024.
9. K. Li *et al.*, "Explicit background endmember learning for hyperspectral anomaly detection," *IEEE Trans. Instrum. Meas.*, 2024.
10. X. Cheng *et al.*, "Deep feature aggregation network for hyperspectral anomaly detection," *IEEE Trans. Instrum. Meas.*, 2024.
11. M. Wang *et al.*, "Hyperspectral simultaneous anomaly detection and denoising: Insights from integrative perspective," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2024.
12. N. R. Shah *et al.*, "Hyperspectral anomaly detection: a performance comparison of existing techniques," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 2078–2125, 2022.
13. S. Liu and M. Song, "Two-orientations finite Markov real-time local anomaly detection via pixel-by-pixel processing for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2024.
14. C. I. Chang, "Target-to-anomaly conversion for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–28, 2022.
15. Y. Xu, L. Zhang, B. Du, and L. Zhang, "Hyperspectral anomaly detection based on machine learning: An overview," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3351–3364, 2022.
16. J. M. Haut *et al.*, "Hyperspectral image analysis using cloud-based support vector machines," *SN Comput. Sci.*, vol. 5, no. 6, pp. 719, 2024.
17. B. Tu *et al.*, "Hyperspectral anomaly detection using the spectral-spatial graph," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
18. J. Sheng, X. Li, and S. Chen, "Iterative graph propagation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
19. B. Tu *et al.*, "Anomaly detection in hyperspectral images using adaptive graph frequency location," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.
20. J. Wei *et al.*, "Hyperspectral anomaly detection based on graph regularized variational autoencoder," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
21. J. Zhang *et al.*, "A light CNN based on residual learning and background estimation for hyperspectral anomaly detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 132, pp. 104069, 2024.
22. L. Zhang, F. Lin, and B. Fu, "A joint model based on graph and deep learning for hyperspectral anomaly detection," *Infrared Phys. Technol.*, vol. 139, pp. 105335, 2024.
23. L. Ren *et al.*, "HADGSM: A unified nonconvex framework for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
24. Y. Huo *et al.*, "Memory-augmented autoencoder with adaptive reconstruction and sample attribution mining for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.

25. Q. Guo *et al.*, "Robust multi-stage progressive autoencoder for hyperspectral anomaly detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 134, pp. 104200, 2024.
26. Z. Yang *et al.*, "A multi-scale mask convolution-based blind-spot network for hyperspectral anomaly detection," *Remote Sens.*, vol. 16, no. 16, pp. 3036, 2024.
27. X. Sun *et al.*, "Information entropy estimation based on point-set topology for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
28. J. Zhang *et al.*, "Enhancing hyperspectral anomaly detection with a novel differential network approach for precision and robust background suppression," *Remote Sens.*, vol. 16, no. 3, pp. 434, 2024.
29. J. Wang *et al.*, "CL-CaGAN: Capsule differential adversarial continual learning for cross-domain hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
30. Q. Guo *et al.*, "Hyperspectral anomaly detection based on spatial-spectral cross guided mask auto-encoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2024.
31. S. Liu *et al.*, "Adaptive dual-domain learning for hyperspectral anomaly detection with state space models," *IEEE Trans. Geosci. Remote Sens.*, 2025.
32. D. Ma *et al.*, "Coarse-To-Fine multiview anomaly coupling network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
33. Z. Wu and B. Wang, "Transformer-based autoencoder framework for nonlinear hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
34. Zhao, X., Liu, K., Wang, X., Zhao, S., Gao, K., Lin, H., Zong, Y. and Li, W., 2024. Tensor adaptive reconstruction cascaded with global and local feature fusion for hyperspectral target detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
35. Zhao, X., Huang, J., Gao, Y. and Wang, Q., 2024. Hyperspectral target detection based on prior spectral perception and local graph fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
36. Zhao, X., Liu, K., Gao, K. and Li, W., 2023. Hyperspectral time-series target detection based on spectral perception and spatial-temporal tensor decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 61, pp.1-12.



MISIROV KOMOLIDDIN MAMASABIROVICH graduated from the Tashkent Financial Institute, Faculty of Accounting and Economics, with a bachelor's degree in "Accounting and Auditing" in 1995-1999, and from the same institute, with a master's degree in "Audit" in 1999-2001. Misirov Komoliddin Mamasabirovich graduated from the Tashkent Financial Institute, Faculty of Accounting and Economics, with a bachelor's degree in "Accounting and Auditing" in 1995-1999, and from the same institute, with a master's degree in "Audit" in 1999-2001. In 2022, he was awarded the degree of Doctor of Economic Sciences DSc (issued according to the protocol of the Higher Attestation Commission of the Republic of Uzbekistan dated September 30, 2022)



AKMALBEK ABDUSALOMOV received the B.S. degree in informatics and information technologies from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2015, and the M.S. and Ph.D. degrees in IT convergence engineering from Gachon University, South Korea, in 2017 and 2022, respectively. He is currently an Assistant Professor with the Computer Engineering Department, Gachon University. His research interests include object detection, artificial intelligence (DL and ML), and medical image processing.



AZIZBEK KHOJAMUROTOV received the B.S. degree in informatics and information technologies from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2023, and the M.S. degrees in informatics and information technologies from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2025.



SAFAROV FURKAT ESHPULATOVICH received a bachelor's degree from Tashkent Financial Institute in 2020 and a master's degree in Computer Engineering from Gachon University (South Korea) in 2022. He is currently a Philosophy Doctoral student at Gachon University (South Korea). His research interests include Computer Vision and the use of connected devices in smart homes and cities



PROFESSOR WOOKEY LEE Industrial Engineering & Biomedical Science and Engineering, Inha University, Incheon, South Korea Wookey Lee (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the Department of Industrial Engineering, Seoul National University, and the M.S.E. degree from Carnegie Mellon University. He is currently a Professor with the Department of BMSE and IE, Inha University, Incheon, South Korea, and the Director of the Voice AI Institute. His research interests include deep learning, meta-learning, patent data, data privacy, and voice AI.



MUKHIDDIN TOSHPULATV SpaceTop Research Center, KAIST Daejeon, South Korea Mukhiddin Toshpulatov (Member, IEEE) received the B.S. and M.S. degrees from Samarkand State University, Uzbekistan, and his PhD degree from the Biomedical Science Engineering Department, Inha University, majoring in Smart Engineering in 2000, 2002, and 2025, respectively. He is doing PostDoctoral research with the SpaceTop Research Center, School of Computing, KAIST, South Korea. His research interests include machine learning and its applications to computer vision. sensor data science. human pose