Classification using Enhanced Spectral and Spatial Transformer with Grasshopper Optimization

Pilligundla Niharika
Department of Computer Science and Artificial Intelligence, SR University, Telangana, India
niharika4jan@gmail.com (corresponding author)

Shanker Chandre
Department of Computer Science & Artificial intelligence, SR University, Warangal, Telangana, India
shankar.chandre@gmail.com

Received: 6 November 2024 | Revised: 4 December 2024, 11 December 2024, and 17 December 2024 | Accepted: 20 December 2024

 DOI: https://doi.org/10.48084/etasr.9517

ABSTRACT

Hyperspectral image (HSI) classification plays a crucial role in remote sensing, allowing the identification of various land cover types. Traditionally, Convolutional Neural Networks (CNNs) have been widely used for this purpose. However, they often face challenges related to high training parameter requirements and limited capacity for feature extraction, affecting their overall effectiveness. To overcome these challenges, this study proposes a novel approach integrating the Enhanced Deep Spectral and Spatial Transformer (EDSST) with Grasshopper Optimization (GHO). EDSST leverages transformer architecture to perform advanced spectral and spatial feature extraction, effectively mitigating the limitations of CNNs. This method improves feature abstraction and classification performance by reducing the number of training parameters while implementing a self-focusing mechanism. This approach incorporates a Classification Head (CH) with an orthogonal softmax activation function to accurately classify hyperspectral images. The proposed method was rigorously evaluated using the Salinas dataset, a benchmark in HSI classification research. The results show substantial improvements over existing techniques, achieving an accuracy of 99.5472%, precision of 99.5574%, recall of 99.5267%, and an F score of 99.6145%. These findings not only demonstrate the effectiveness of the proposed method in HSI classification but also highlight its efficiency and robustness, offering a promising solution for future applications in remote sensing and environmental monitoring.

Keywords-convolutional neural networks; enhanced deep spectral and spatial transformer; grasshopper optimization; classification head; hyperspectral images

I. INTRODUCTION

Hyperspectral Imaging (HSI) is a technique for obtaining the electromagnetic spectrum from the visible to the infrared wavelength ranges through remote sensing [1]. Hundreds of narrow spectral bands are available from a specific region on the Earth's surface with HSI sensors [2]. HSIs consist of pixels corresponding to wavelength-specific spectral reflectances [3]. As HSIs are capable of distinguishing subtle spectral variations, they have found widespread use in a wide range of fields. Due to advances in imaging spectrometry, hyperspectral sensors can capture the reflectance intensity of a given scene at a higher spatial and spectral resolution than ever before. HSIs are obtained by simultaneously capturing both spatial features and continuous detection spectra of different objects [4, 5].

HIS characterization is important for applications ranging from environmental monitoring and agriculture to mineralogy and conservation, allowing detailed spectral analysis for accurate characterization and detection in various domains [6]. HSI classification, which assigns each pixel vector to a specific class, is one of the main analysis tasks and has received a lot of attention from researchers. Several traditional methods, such as the Support Vector Machine (SVM) and K-Nearest Neighbor (KNN), have been proposed for this task [7]. However, these approaches ignore correlations between pixels on spatial axes, wasting spatial information. The high dimensionality of HSIs with hundreds of spectral bands leads to the curse of dimensionality, complicating data processing and analysis [8]. Obtaining labeled samples for training is often expensive and time-consuming, resulting in insufficient data for supervised learning.

II. LITERATURE SURVEY

Many studies have proposed HSI detection and classification methods. In [9], SpectralFormer was introduced to enhance the HSI classification by addressing the limitations

of CNNs in representing spectral signatures. SpectralFormer uses transformers to capture local spectral sequence info, yielding group-wise embeddings applicable to pixels and patches. This method achieved state-of-the-art results without convolutional or recurrent units. However, classic transformer-based networks such as ViT struggle with detailed spectral discrepancies and layer-to-layer memory transmission, impacting performance on hyperspectral data. In [10], an HSI classification method was proposed to address the computational challenges of traditional CNNs while maintaining high performance by introducing a combination of the ghost-module architecture with a CNN-based classifier. As a result, computational costs were significantly reduced and satisfactory results were achieved. As a drawback, deep learning-based HSI classification methods require large amounts of labeled data and can result in overfitting.

In [11], PLG-KELM (PCA-LBP-optimized Kernel Extreme Learning Machine) was proposed for HSI classification, employing PCA for spectral dimensionality reduction and LBP for spatial texture extraction, and optimizing KELM with the gray wolf optimization algorithm. This method used PCA to reduce spectral dimensionality and LBP to extract local texture features from HSIs. Despite enhancing the classification model by optimizing KELM parameters, a limitation was its lower operational efficiency due to computational complexity, potentially restricting real-time applications in resource-constrained environments. In [12], an end-to-end Fully Convolutional Segmentation Network (FCSN) was introduced for HSI classification, which labeled all pixels within HSI cubes simultaneously and showed superior generalization to spatial land-cover distributions compared to cropped cube-based CNN methods. However, a limitation was the difficulty of simulating realistic spatial land-cover distributions with the proposed HSI cube generation method and the inability of FCSN as a supervised method to utilize unlabeled pixels.

In [13], a hybrid deep learning method was used for HIS classification, combining a 3D fast learning block with a 2D CNN to extract spectral and spatial features. This method addressed challenges including noise, limited labeled samples, overfitting, and insufficient feature extraction, enhanced by optimization techniques such as batch normalization, dropout, and L2 regularization. Experimental results on the Salinas, University Pavia, and Indian Pines datasets showed that although 3D CNNs with similar layer setups outperformed traditional 2D CNNs, the hybrid model's performance was contingent on the availability of sufficient training samples, potentially limiting its complexity reduction and feature extraction optimization benefits for HSI classification in cases of inadequate data. In [14], a Rotation-Invariant Attention Network (RIAN) was introduced for HSI classification, mitigating the sensitivity issues of traditional methods with 3D convolutions on rotated HSIs. RIAN integrated a Center Spectral Attention (CSpeA) module to focus on central pixel spectral properties, reducing interference. Additionally, a Rectified Spatial Attention (RSpaA) module extracted rotation-invariant spectral-spatial features from HSI patches. Despite its strengths, RSpaA's limitation was its reliance on pixel similarity for spatial aggregation, potentially limiting its ability to effectively capture textural information from HSIs.

III. PROPOSED HSI CLASSIFICATION

HSIs were obtained from publicly available datasets, followed by data normalization, feature extraction using EDSST, and classification with optimized hyperparameters through Grasshoppers Hyperparameter Optimization (GHO), as shown in Figure 1. Data normalization ensured consistency and quality. Then, EDSST was used for advanced feature extraction and classification, combining spectral and spatial feature extraction capabilities with a CH that incorporates an orthogonal softmax activation function. This hybrid approach leverages the transformer architecture to extract and analyze spectral and spatial features more effectively than traditional CNNs, which often require more training parameters and may struggle with abstract feature extraction. Following feature extraction, the CH categorizes the images into classes such as vegetables, bare soil, and vineyards. To further improve classification accuracy, the EDSST hyperparameters were optimized using the GHO algorithm. The performance of the proposed method was evaluated using the Salinas dataset and analyzed for classification accuracy. The entire workflow was implemented and tested using Python to ensure robust evaluation and performance measurement. Figure 1 shows the flow diagram of the proposed method for HSI classification.

A. Preprocessing

Normalization is commonly used in ML and computer vision to analyze features in the Salinas dataset. Normalization standardizes and manages the data for consistency and is important for the hyperspectral characteristics of images for multiple reasons. Normalization standardizes pixel values in HSIs. Images can differ in lighting, camera quality, and other factors, affecting pixel values. Normalized pixel values eliminate scale differences, improve data comparability, and prevent the dominance of features with larger numerical values in the analysis. Machine learning methods, particularly deep learning models, may converge slowly or fail if input data exhibit large numerical variability. Normalization speeds up and stabilizes training, helping the model learn from data.

In HSIs, specific characteristics of wide spectral ranges may have higher pixel intensities than others. When variations are not standardized, the framework concentrates too much on certain areas, biasing predictions. Normalization reduces biases. Normalizing input data within the range [0, 1] improves neural network activation functions, ensuring that they are within the required range. This study normalized vegetables, bare soil, and vineyard image data characteristics using min-max scaling. Each data point was divided by the range obtained by subtracting the lowest value from the greatest:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

This is advantageous when there is a significant variation in pixel intensities across distinct images since it ensures consistency in the dataset.

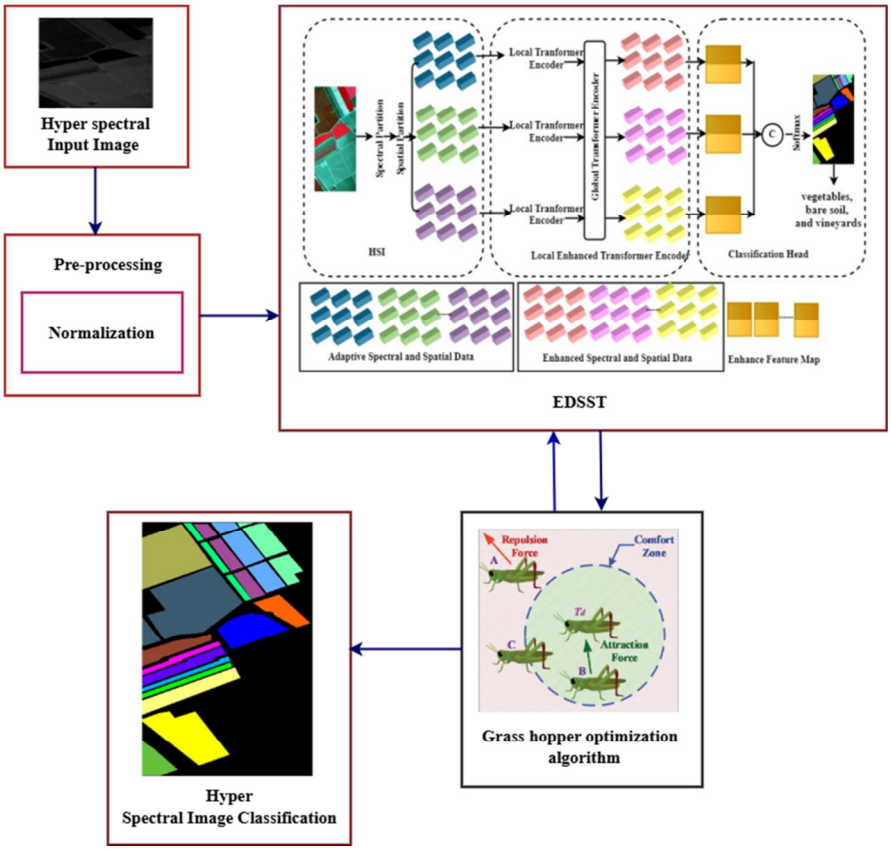

Fig. 1. Flow diagram of the proposed HIS classification method.

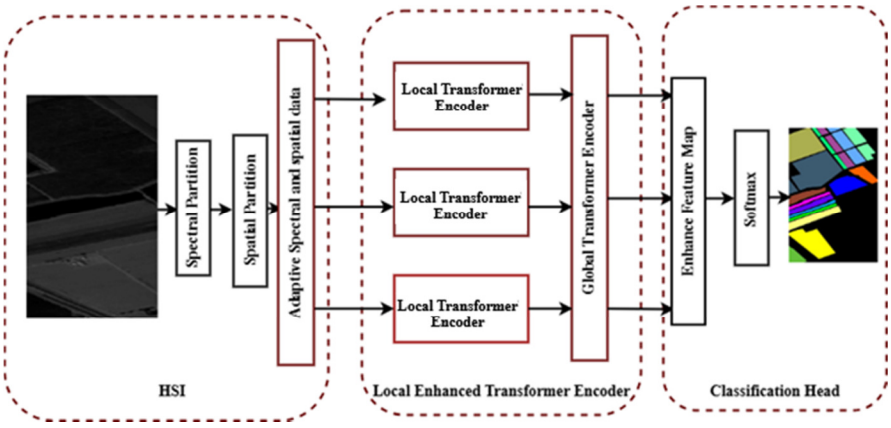

Fig. 2. General structure of EDSST.

B. Enhanced Deep Spectral and Spatial Transform (EDSST)

EDSST was used for feature extraction and classification, which is a hybrid of DSST with a CH. CH is usually an orthogonal softmax activation function. EDSST is used to extract spectral and spatial features, and CH is used to classify the image. To begin, images were divided into nonoverlapping patches using tokenization, which makes the patches resemble tokens in Natural Language Processing (NLP). Then, the pure transformer encoder was applied to the tokens. Then, classification was performed using the classification token that matches the representation of the global image. Figure 2 shows the overall framework of EDSST.

An HSI module includes a learnable super-pixel sectional submodule. More specifically, the first high-resolution HSI information $H \in R^{He \times Wi \times B}$ is initially input to create a two-layer to reduce redundant images and produce low-grade discriminating features $D \in R^{He \times Wi \times C}$, where He, Wi, B, and C stand for the hidden feature's height, width, band number, and measurement, respectively. In the shallow CNN, each convolutional block consists of a Rectified Linear Unit (ReLU)

activation function, a 1×1 convolutional layer, and a Batch Normalized (BN) layer.

Next, a group block is used, comprising the ReLU activation function, the BN layer, and the 1×1 group convolutional layer to remove portions of the spectral dimension's low-level features. The process of spectral partitioning, which splits the input channels and filters into groups causing each group to do a separate convolution operation and generate the following set of patches, can be effectively substituted with the group convolutional block:

$$D \to D_{spe} = [D^1, D^2, \ldots, D^S] \quad (2)$$

where D_{spe} denotes the collection of patches produced through the convolution of the group process, $D^i \in R^{He \times Wi \times C_s}$ refers to the spectral-spatial characteristics of the i^{th} group, $i = 1, 2, \ldots, S$, and C_s is the feature dimension of each group given by $C_s = C/S$. The proposed HSI module can produce spectral-spatial tokens that are indicative of the HSIs and fully account for its image-spectrum merging structure and image redundancy property.

1) Improved Transformer Encoder

The novel transformer encoder with local enhancement consists of two transformer encoders stacked in serial order: a local and a global. This is in contrast to the original transformer encoder. The term implies that although the former is intended to gather local spectral-spatial information among spatially adjacent tokens in each group, the latter is utilized to maintain the ability to capture global linkages among tokens from the spectral dimension. Figure 3 shows a local transformer encoder.

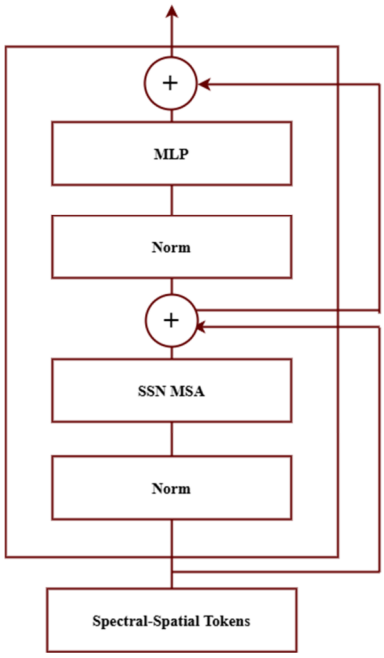

Fig. 3. Local transformer encoder.

2) Local Transformer Encoder

One of the most important components of the original transformer encoder is the Multihead Self-Attention (MSA) layer, which is highly influential in constructing long-range interactions between tokens. However, the original global MSA layer fails to take into account the important local information in the spatial dimension, degrading execution, particularly when constrained. Furthermore, the original global MSA layer might introduce interference with the semantic modeling procedure by taking into account a few tokens that are not relevant. The goal was to overcome these constraints by creating a new local transformer encoder on the Spatial Nearest Neighbor (SNN) MSA layer. Based on the SNN MSA layer, the local transformer encoder is designed to increase the similarity between spatial surrounding tokens in each group:

$$A_{p1}^{'i} = A_{p1-1}^{i} + MSA_{snn}(LN(A_{p1-1}^{i})) \quad (3)$$

$$A_{p1}^{i} = A_{p1}^{'i} + MLP(LN(A_{p1}^{'i})) \quad (4)$$

where $A_{p1-1}^{i} = \left[A_{p1-1}^{i,1}, A_{p1-1}^{i,2}, \ldots, A_{p1-1}^{i,M}\right]$ are the $(p_1 - 1)$ adaptive layer spectral-spatial tokens, $p_1 = 1, 2, \ldots, P_1$, P_1 is the encoder of the local transformer depth, and MSA_{snn} represents the SNN MSA layer.

3) Global Transformer Encoder

The original MSA layer and the SNN MSA layer complement each other in modeling and apply a normal global transformer encoder after the local transformers to extract features from local to global relationships. Tokens from every group acquired by the locally improved transformer encoder are put together into $A_{P1} = \left[A_{P1}^{1}, A_{P1}^{2}, \ldots, A_{p1}^{S}\right]$ and redesigned to $B \in R^{M \times S \times C_s}$. Afterward, they are received by the global transformer encoder. The encoder for the global transformer is still capable of capturing the following global relationships between tokens from the spectral dimension:

$$X_{p2}^{'} = X_{p2-1} + MSA(LN(B_{p2_{-1}})) \quad (5)$$

$$X_{p2} = X_{p2}^{'} + MLP(LN(B_{p2}^{'})) \quad (6)$$

where X_{p2-1} are the $(p_2 - 1)$ layer's token embeddings, $p_2 = 1, 2, \ldots, P_2$, and P_2 is the global transformer encoder's depth.

Despite sharing the same network architecture, the original and the global transformer encoder differ slightly in terms of building dependencies. The global transformer encoder is used to simulate long-range dependencies between tokens in different spectral groups, whereas the original transformer encoder captures long-range spatial interactions between tokens in each group.

4) Classification Head (CH)

All the improved tokens are used for classification rather than settling on a single extra classification token. For this reason, a series of improved spectral-spatial tokens X_{p2} is converted into the row-normalized pixel-super-pixel correlations in the image-like feature representations V. This can be calculated using basic matrix multiplication:

$$V = \tilde{R}B_{P2} \quad (7)$$

The output features of each group are concatenated and fed into the softmax classifier, which consists of a linear projection layer and a softmax activation layer, along with the class probabilities of each pixel in the HSI. The GHO algorithm is used to optimize the hyperparameters of EDSST to enhance classification accuracy by efficiently exploring the parameter space and finding optimal configurations to strike a balance between exploration and exploitation to improve the model's performance and generalization.

5) Hyperparameter Optimization Using GHO

Hyperparameter optimization is the procedure of determining the best mix of EDSST hyperparameter settings to optimize performance on data in a reasonable quantity. This process is essential to EDSST's capacity for precise result detection. Most of this input uses the hyperparameters' default values. GHO [15] was used to optimize hyperparameters. Table I shows the hyperparameter initialization range.

TABLE I. HYPERPARAMETER INITIALIZATION RANGES

Hyper-parameters	Range
Number of Layers (NL)	1 to 4
Number of Attention Heads (NA)	1 to 8
Hidden Size (HS)	256
Dropout Rate (DR)	0.1
Learning Rate (LR)	1e-4
Batch Size (BS)	32
Epochs (E)	1 to 50

The GHO algorithm's step-by-step procedure is as follows:

1. Initialization: Selecting the ideal hyperparameter is the primary goal of this approach. Initially, define the top and lower bounds of the problem, the dimensionality D of the variables, the maximum number of iterations, and the Grasshopper size N. Each solution represented by the Grasshopper consists of hyperparameters, including NL, NA, HS, DR, LR, BS, and E. At first, a random selection is made. The first solution format is given by:

$$P_N = \{S_1, S_2, \dots, S_N\} \quad (8)$$

X_N is the N^{th} solution or Grasshopper's position.

$$S_i = \{NL, NA, HS, DR, LR, BS, E\}_i \quad (9)$$

2. Calculating fitness: Following initialization, each solution's fitness is evaluated using the GHO technique. In this case, the classification accuracy is defined for fitness purposes. The solution with the highest fitness value is deemed to be the most effective. The fitness function is given by:

$$Fitness = Max\left(\frac{TP+TN}{TP+TN+FP+FN}\right) \quad (10)$$

3. The nymph and adult phases make up their life cycle. The nymph phase is distinguished by sluggish motions and short steps, whereas the adult phase is characterized by long-range and sharp movements. The nymph and adult motions comprise the phases of GOA's intensity and diversity:

$$S_i = G_i + P_i + Y_i \quad (11)$$

where S_i denotes the location of the i^{th} grasshopper, G_i is grasshoppers' social engagement with one another, P_i indicates the grasshopper's i^{th} gravitational force, and Y_i is the advection of wind. Interpersonal communication G_i is described as follows:

$$G_i = \sum_{\substack{j=1 \\ j\neq 1}}^{N} g(x_{ij})\hat{x}_{ij} \quad (12)$$

where N represents the amount of grasshoppers, $x_{ij} = |S_j - S_i|$ specifies a distance in Euclidean terms between the j^{th} and i^{th} grasshoppers, and $\hat{d}_{ij} = \frac{S_j - S_i}{x_{ij}}$ is a unit vector representing the social forces created from the i^{th} to the j^{th} grasshopper.

When there is no force involved and two grasshoppers are apart, there is neither attraction nor repulsion. Comfort zone is the term for this region. The force of gravity P_i is determined by the subsequent formula:

$$P_i = -p\hat{e}_p \quad (13)$$

where p stands for gravitational constant and $\hat{e}_p$ denotes a vector in units pointing towards the center of the earth. The advection of wind Y_i is determined by:

$$Y_i = u\hat{e}_w \quad (14)$$

The swarm mechanism does not converge at a target point because the grasshoppers quickly reach their comfort zone. An improved iteration of this formula is provided as:

$$S_i^x = w\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} w\frac{ub_x - lb_x}{2} g\left(\left|S_j^x - S_i^x\right|\right)\frac{S_j - S_i}{x_{ij}}\right) + \hat{T}_x \quad (15)$$

where ub_x and lb_x indicate the top and lower boundaries within the x^{th} dimension, correspondingly. $\hat{T}_x$ indicates the greatest solutions discovered along the x spatial dimension. Note that g is comparable to g in (1). P is equivalent to 0 and Y is always in the direction of the ideal solution $\hat{T}_x$. As a result, it offers an excellent mix of intensification and diversity. The variable w_2 is employed to decrease the grasshoppers' comfort, attraction, and repulsion zones in proportion to the number of repeats. w_1 and w_2 are seen as a single parameter, given by:

$$w = w_{max} - \frac{t(w_{max} - w_{min})}{t_{max}} \quad (16)$$

where w_{max} and w_{min} represent the maximum and minimum values of w, respectively, t is the current iteration and t_{max} is the maximum number of iterations.

4. Termination condition: The process is repeated until optimal hyperparameter selection is made. The EDSST receives the chosen hyperparameter value.

IV. RESULTS AND DISCUSSION

The tests were performed on a PC equipped with an Intel Core i5 4570s CPU at 2.90 GHz and 8 GB of RAM using Python on Windows 64-bit. The experimental configuration included two data centers with four hosts and a total RAM of 8 GB. The host had a bandwidth of 2800 Mbps.

A. Dataset Description

AVIRIS [16] captured images over the California Salinas Valley. This dataset has 16 distinct land cover classes. Over a 400-2500 nm range, the dataset has a 512×217 pixel size and 3.7 m spatial resolution. There are 224 spectral reflectance bands in this scenery.

B. Experimental Results

The experimental results, including accuracy, precision, recall, F score, and accuracy vs. loss, show the performance of the proposed method compared to FCN, GAN, and LSTM. The proposed method achieved superior performance metrics across these evaluations, indicating better overall performance. Figure 4 illustrates the training and testing accuracy and loss over 500 epochs for various models. The proposed EDSST model showed a high accuracy that stabilizes near 1.0, while its loss decreases significantly, indicating effective learning. In contrast, ResNets and CNNs exhibited lower accuracy and higher loss values, suggesting they were less effective in this context. Overall, the proposed EDSST outperformed the other models in both accuracy and loss reduction.

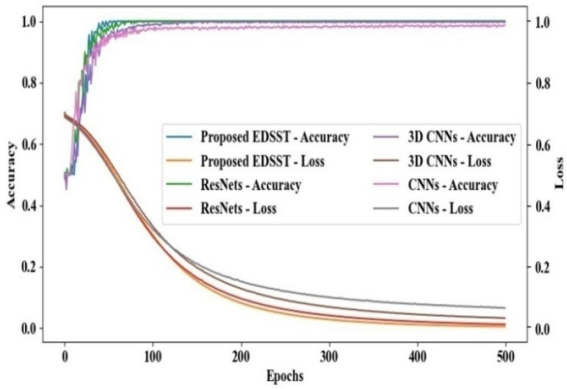

Fig. 4. Training and testing accuracy and loss.

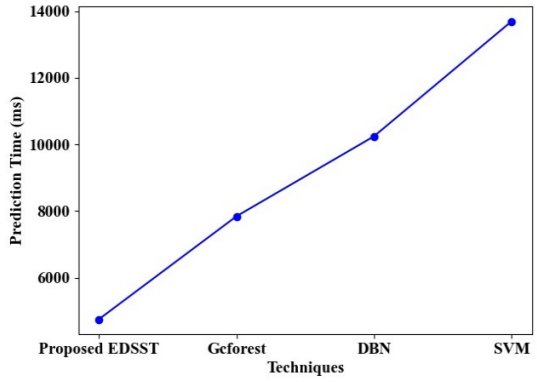

Fig. 5. Prediction time comparison compared to existing methods.

Figure 5 compares the prediction times of the EDSST, Gcforest, DBN, and SVM methods. The proposed method achieved the shortest forecast time. SVM, on the other hand, had the longest prediction time, which makes it less appropriate for applications that need speed. Overall, the pattern shows that the prediction time tends to grow along with model complexity.

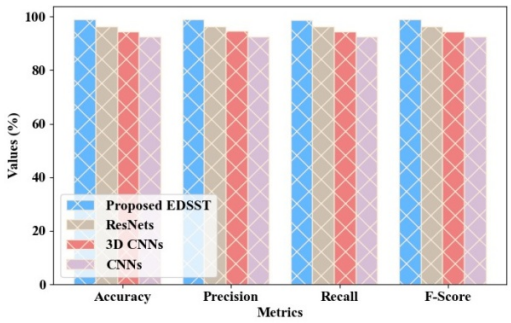

Fig. 6. Evaluation metrics.

Figure 6 presents a comparison of the proposed EDSST model against ResNets, 3D CNNs, and CNNs across four key metrics: accuracy, precision, recall, and F-score. The proposed EDSST consistently outperformed the other models in all metrics, indicating its superior classification capability. Each metric reflects the model's effectiveness in correctly identifying positive instances, with the EDSST achieving the highest values.

V. CONCLUSION

This paper introduced a novel approach to HSI classification by combining EDSST with GHO. EDSST leverages a transformer architecture for advanced spectral and spatial feature extraction, addressing the limitations of CNNs by reducing training parameter requirements and enhancing feature abstraction through a self-focusing mechanism. The integration of a CH with an orthogonal softmax activation function further enhances classification accuracy for HSIs across diverse categories. The proposed model achieved an accuracy of 99.5472%, precision of 99.5574%, recall of 99.5267%, and an F score of 99.6145% on the Salinas dataset. These results highlight the proposed method's significant efficiency and accuracy improvements, making it a robust solution for HSI classification.

REFERENCES

[1] S. E. Qian, "Hyperspectral Satellites, Evolution, and Development History," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7032–7056, 2021, https://doi.org/10.1109/JSTARS.2021.3090256.

[2] S. L. Ustin and E. M. Middleton, "Current and near-term advances in Earth observation for ecological applications," *Ecological Processes*, vol. 10, no. 1, Jan. 2021, Art. no. 1, https://doi.org/10.1186/s13717-020-00255-4.

[3] P. Mishra, A. Karami, A. Nordon, D. N. Rutledge, and J. M. Roger, "Automatic de-noising of close-range hyperspectral images with a wavelength-specific shearlet-based image noise reduction method," *Sensors and Actuators B: Chemical*, vol. 281, pp. 1034–1044, Feb. 2019, https://doi.org/10.1016/j.snb.2018.11.034.

[4] R. Pillay, J. Y. Hardeberg, and S. George, "Hyperspectral imaging of art: Acquisition and calibration workflows," *Journal of the American Institute for Conservation*, vol. 58, no. 1–2, pp. 3–15, Apr. 2019, https://doi.org/10.1080/01971360.2018.1549919.

[5] M. Zhu, L. Jiao, F. Liu, S. Yang, and J. Wang, "Residual Spectral–Spatial Attention Network for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 449–462, Jan. 2021, https://doi.org/10.1109/TGRS.2020.2994057.

[6] S. Xu, M. Wang, X. Shi, Q. Yu, and Z. Zhang, "Integrating hyperspectral imaging with machine learning techniques for the high-

resolution mapping of soil nitrogen fractions in soil profiles," *Science of The Total Environment*, vol. 754, Feb. 2021, Art. no. 142135, https://doi.org/10.1016/j.scitotenv.2020.142135.

[7] F. Xie and Y. Xu, "An efficient regularized K-nearest neighbor structural twin support vector machine," *Applied Intelligence*, vol. 49, no. 12, pp. 4258–4275, Dec. 2019, https://doi.org/10.1007/s10489-019-01505-5.

[8] S. Swain, A. Banerjee, M. Bandyopadhyay, and S. C. Satapathy, "Dimensionality Reduction and Classification in Hyperspectral Images Using Deep Learning," in *Machine Learning Approaches for Urban Computing*, M. Bandyopadhyay, M. Rout, and S. Chandra Satapathy, Eds. Singapore: Springer, 2021, pp. 113–140.

[9] D. Hong *et al.*, "SpectralFormer: Rethinking Hyperspectral Image Classification With Transformers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022, https://doi.org/10.1109/TGRS.2021.3130716.

[10] M. E. Paoletti, J. M. Haut, N. S. Pereira, J. Plaza, and A. Plaza, "Ghostnet for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 12, pp. 10378–10393, Sep. 2021, https://doi.org/10.1109/TGRS.2021.3050257.

[11] H. Chen, F. Miao, Y. Chen, Y. Xiong, and T. Chen, "A Hyperspectral Image Classification Method Using Multifeature Vectors and Optimized KELM," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2781–2795, 2021, https://doi.org/10.1109/JSTARS.2021.3059451.

[12] H. Sun, X. Zheng, and X. Lu, "A Supervised Segmentation Network for Hyperspectral Image Classification," *IEEE Transactions on Image Processing*, vol. 30, pp. 2810–2825, 2021, https://doi.org/10.1109/TIP.2021.3055613.

[13] S. Ghaderizadeh, D. Abbasi-Moghadam, A. Sharifi, N. Zhao, and A. Tariq, "Hyperspectral Image Classification Using a Hybrid 3D-2D Convolutional Neural Networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7570–7588, 2021, https://doi.org/10.1109/JSTARS.2021.3099118.

[14] X. Zheng, H. Sun, X. Lu, and W. Xie, "Rotation-Invariant Attention Network for Hyperspectral Image Classification," *IEEE Transactions on Image Processing*, vol. 31, pp. 4251–4265, 2022, https://doi.org/10.1109/TIP.2022.3177322.

[15] Y. Meraihi, A. B. Gabis, S. Mirjalili, and A. Ramdane-Cherif, "Grasshopper Optimization Algorithm: Theory, Variants, and Applications," *IEEE Access*, vol. 9, pp. 50001–50024, 2021, https://doi.org/10.1109/ACCESS.2021.3067597.

[16] "AVIRIS - Airborne Visible / Infrared Imaging Spectrometer - Data." https://aviris.jpl.nasa.gov/data/get_aviris_data.html.