RESEARCH ARTICLE OPEN ACCESS

Hyperspectral Image Classification Based on Multilevel Spectral-spatial Transformer Network

Bikkili Alekya Himabindu*, U.Vishnu Priya**,G.Siva Poojitha***,J.Bhargavi****

*(Electronics and Communication Engineering, Santhiram Engineering College Nandyal
Email:alekhya.ece@srecnandyal.edu.in)
**(Electronics and Communication Engineering, Santhiram Engineering College Nandyal
Email:20X51a04c1@srecnandyal.edu.in)
***(Electronics and Communication Engineering, Santhiram Engineering College Nandyal
Email:20X51a0486@srecnandyal.edu.in)
****(Electronics and Communication Engineering, Santhiram Engineering College Nandyal
Email:20X51a0490@srecnandyal.edu.in)

--************************----------------------------------

Abstract:

Deep learning techniques have found widespread application in hyperspectral image classification (HSIC), with Convolutional Neural Networks (CNNs) emerging as a dominant model in recent years. Despite the significant progress made by CNN-based approaches, they still encounter several challenges, including limited utilization of long-range information, constrained receptive fields, and high computational demands. To address these challenges, this study introduces a multi-level spectral-spatial transformer network (MSTNet) tailored for HSIC within an image-based classification framework. The proposed network leverages a transformer encoder to learn feature representations and integrates multi-level features through a decoder to yield classification outcomes. Experimental results on two real hyperspectral datasets demonstrate the effectiveness and superiority of the proposed method.

Keywords **—Hyperspectral image, MSTNet,MFAD,TE**

--************************----------------------------------

I. INTRODUCTION

Hyperspectral satellite imagery has revolutionized the way we observe and analyze the Earth's surface. Unlike traditional satellite images, which capture data across just a few broad spectral bands, hyperspectral imagery provides a wealth of information by capturing data across hundreds or even thousands of narrow, contiguous bands. This rich dataset allows for detailed analysis of the Earth's surface and its features. Each pixel in a hyperspectral image contains a spectrum, which represents the intensity of reflected or emitted light across different wavelengths. By analyzing these spectra, scientists and analysts can identify and characterize materials based on their unique spectral signatures. This capability enables hyperspectral imagery to differentiate between different types of vegetation, minerals, soil types, water bodies, and human-made materials with great precision.

One of the key advantages of hyperspectral imagery is its high spectral resolution. This allows for the detection of subtle differences in surface materials that may not be visible in traditional satellite images. For example, hyperspectral imagery can distinguish between healthy and stressed vegetation, identify specific mineral compositions in rocks, and detect pollutants in water bodies.

Hyperspectral satellite imagery finds applications in a wide range of fields, including agriculture, environmental monitoring, mineral exploration, urban planning, disaster management, and defense intelligence. In agriculture, hyperspectral imagery can be used to monitor crop health, optimize irrigation, and detect pest infestations. In environmental monitoring, it can track changes in land cover, monitor deforestation, and assess water quality.

Mineral exploration companies use hyperspectral imagery to identify potential mineral deposits based on the spectral signatures of rocks and soils. Urban planners utilize it to map land use and monitor urban sprawl. During disasters, such as wildfires or oil spills, hyperspectral imagery can provide valuable information for emergency response efforts.

The development of hyperspectral satellite technology has been driven by advancements in sensor technology, data processing techniques, and satellite platforms. Modern hyperspectral sensors are capable of capturing data with high spatial, spectral, and temporal resolutions, allowing for detailed and frequent monitoring of the Earth's surface.

II. SURVEY

The integration of DL methods with attention mechanisms and innovative network architectures holds promise for advancing HSI classification accuracy and efficiency. Continued research in this direction is essential to unlock the full potential of DL for HSI analysis and its diverse applications.

Hyperspectral imaging (HSI) stands out for its ability to capture rich spatial and spectral information across hundreds of adjacent bands, offering high-resolution and high-dimensional data. Its applications span diverse domains, including sea ice detection, ecosystem monitoring, vegetation analysis, and classification tasks. Recently, HSI classification has garnered significant attention in both research and industrial sectors, propelled by the growing demand for accurate analysis and interpretation of hyperspectral data.

However, the task of classifying hyperspectral images is inherently complex. The abundance of wavebands poses a challenge for classification models, particularly in achieving high accuracy with limited training samples. Traditional methods often rely on manual feature extraction and hyperparameter tuning, necessitating expert knowledge and time-consuming adjustments. Machine learning techniques, such as logistic regression, AdaBoost, and support vector machines, have been employed for image classification but may struggle to effectively handle the high-dimensional and rich spectral information present in hyperspectral images.

In contrast, deep learning approaches offer promise by automatically learning robust and discriminative features in a data-driven manner. They have demonstrated superior performance in hyperspectral image classification, outperforming traditional methods in accuracy and efficiency. However, challenges persist in this domain, including class imbalance within hyperspectral image datasets and the need for effective spatial-spectral feature extraction.

Addressing these challenges requires innovative solutions. One major issue is the imbalanced distribution of classes within hyperspectral image datasets, which can bias classification models towards majority classes and impact overall accuracy. Additionally, the high dimensionality of hyperspectral data complicates feature extraction, necessitating strategies to capture both spatial and spectral features effectively.

To tackle these challenges, this paper proposes a novel deep-learning-based model specifically designed to address class imbalance in HSI classification. The model leverages a combination of 1D_2D convolutional networks for spatial-spectral feature extraction, autoencoders, and generative adversarial networks (GANs) for generating synthetic images of minority classes to rebalance the dataset. Finally, a 2D convolutional network is employed for image classification on the balanced dataset.

In summary, the contributions of this paper lie in proposing a comprehensive solution to the challenges of class imbalance and effective feature extraction in hyperspectral image classification. By leveraging deep learning techniques and innovative model architectures, the proposed approach aims to improve classification accuracy and efficiency in handling hyperspectral data, thereby advancing the capabilities of HSI analysis for various applications.

III. PROPOSED SYSTEM

The proposed image-based classification framework offers several advantages over the traditional patch-based method. It facilitates a faster reasoning process by operating directly on the entire image and conserves computational resources by eliminating the need for iterative patch processing during both training and testing phases. Additionally, the image-based approach enhances performance and versatility across different datasets and semantic segmentation models, making it a promising solution for HSIC tasks.

3.1 Multi-level Spectral-Spatial Transformer Network

The multi-level spectral-spatial transformer network (MSTNet) represents a novel architecture tailored for hyperspectral image (HSI) classification tasks, aiming to harness both spectral and spatial information efficiently. At the core of MSTNet lie two pivotal modules: the Transformer Encode (TE) and the Multi-Level Feature Aggregation Decoder (MFAD). These modules collaborate to process and integrate spectral and spatial features comprehensively. The TE module assumes the responsibility of executing Multi-Head Self-Attention (MSA) learning on the input data, enabling it to capture intricate relationships and dependencies among spectral bands effectively. By transforming the three-dimensional feature map into a two-dimensional sequence, the TE module facilitates MSA learning, thus enhancing the network's ability to encode spectral information efficiently. On the other hand, the MFAD module plays a pivotal role in aggregating features learned by the TE module across multiple levels. This aggregation process culminates in the generation of a multi-level fusion feature map, which encapsulates both spectral and spatial information in a cohesive manner.

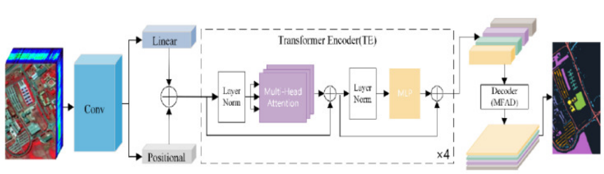

Figure -1:Proposed system

3.2 Transformer Encoder

Transformers, originally introduced for natural language processing tasks, have gained significant traction in various fields, including computer vision. In hyperspectral image classification, where both spectral and spatial information is crucial, transformer models offer a promising avenue for extracting meaningful features and improving classification accuracy. At the core of transformer models is the self-attention mechanism, which enables the model to capture dependencies between different elements in the input sequence. In the context of hyperspectral images, this means that the model can learn to attend to relevant spectral bands and spatial locations when making classification decisions.

3.3 Multi-Level Feature Aggregation Decoder

The Multi-Level Feature Aggregation Decoder (MFAD) serves as a pivotal component in completing the pixel-level segmentation process within the multi-level spectral-spatial transformer network (MSTNet). Once the feature extraction phase is completed in the Transformer Encoder (TE), the features need to be reshaped from the 2D embedding shape to the 3D feature map to facilitate pixel-level segmentation. To enhance the representation capability of the decoder, the concept of multi-level feature aggregation is introduced.

IV RESULT AND DISCUSSION

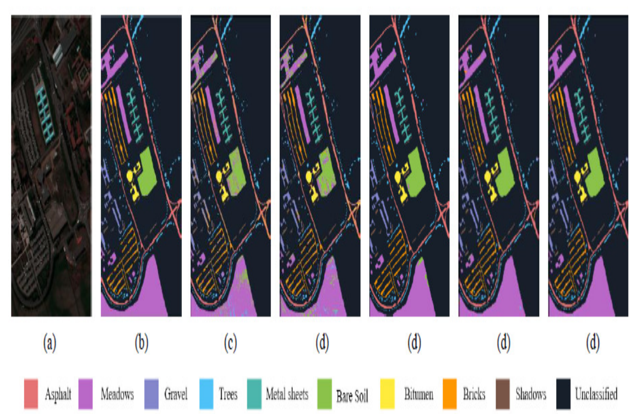

Figure -2:Output for a satellite capture image

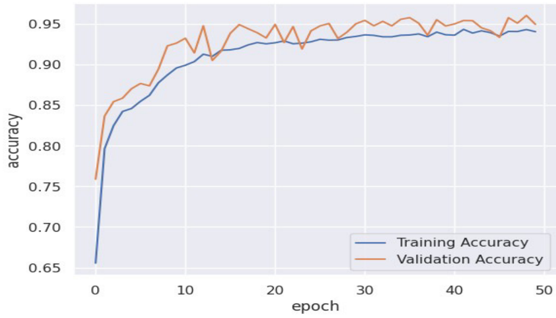

Figure 3:Epoch Vs Accuracy

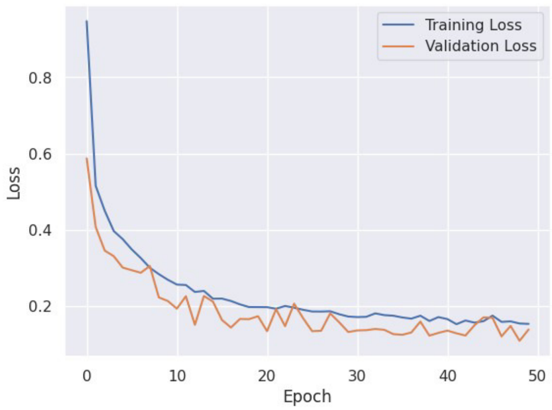

Figure -4:Epoch Vs Loss

Moreover, the integration of an image-based classification framework introduces efficiencies in the classification process, particularly in terms of time efficiency. By adopting this framework, MSTNet streamlines the classification process, making it more computationally efficient without compromising on accuracy. This enhancement is crucial for applications where timely analysis of hyperspectral data is essential.

Experimental evaluations conducted on two real hyperspectral datasets demonstrate the superior performance of MSTNet in terms of classification accuracy, time efficiency, and robustness. The model achieves outstanding results across various evaluation metrics, showcasing its effectiveness in accurately classifying hyperspectral images while maintaining computational efficiency and robustness to variations in the input data.

IV. CONCLUSIONS

In conclusion, the proposed MSTNet represents a significant advancement in hyperspectral image classification, offering a powerful solution that combines advanced techniques with practical considerations. By leveraging self-attention mechanisms and an image-based classification framework, MSTNet achieves superior performance in classification accuracy, time efficiency, and robustness, making it a valuable tool for various applications in fields such as agriculture, environmental monitoring, and remote sensing.

REFERENCES

[1] Zhang, M.; Li, W.; Du, Q. Diverse Region-Based CNN for Hyperspectral Image Classification. *IEEE Trans. Image Process.* **2018**, *27*, 2623–2634.
[2] Ghamisi, P.; Plaza, J.; Chen, Y.; Li, J.; Plaza, A.J. Advanced Spectral Classifiers for Hyperspectral Images: A review. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–32.
[3] Li, J.; Du, Q.; Li, Y.; Li, W. Hyperspectral Image Classification with Imbalanced Data Based on Orthogonal Complement Subspace Projection. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3838–3851.
[4] Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep Learning for Hyperspectral Image Classification: An Overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709.
[5] Wambugu, N.; Chen, Y.; Xiao, Z.; Tan, K.; Wei, M.; Liu, X.; Li, J. Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102603.
[6] Gao, H.; Chen, Z.; Xu, F. Adaptive spectral-spatial feature fusion network for hyperspectral image classification using limited training samples. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102687.
[7] Xu, Y.; Du, B.; Zhang, L. Beyond the Patchwise Classification: Spectral-Spatial Fully Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Big Data* **2019**, *6*, 492–506.
[8] Naji, H.A.H.; Xue, Q.; Lyu, N.; Duan, X.; Li, T. Risk Levels Classification of Near-Crashes in Naturalistic Driving Data. *Sustainability* **2022**, *14*, 6032.
[9] Naji, H.A.H.; Xue, Q.; Zhu, H.; Li, T. Forecasting Taxi Demands Using Generative Adversarial Networks with Multi-Source Data. *Appl. Sci.* **2021**, *11*, 9675.
[10] Jia, S.; Jiang, S.; Lin, Z.; Li, N.; Xu, M.; Yu, S. A survey: Deep learning for hyperspectral image classification with few labeled samples. *Neurocomputing* **2021**, *448*, 179–204.
[11] Khan, S.; Rahmani, H.; Shah, S.A.A.; Bennamoun, M. A Guide to Convolutional Neural Networks for Computer Vision. *Synth. Lect. Comput. Vis.* **2018**, *8*, 1–207.
[12] He, N.; Paoletti, M.E.; Haut, J.N.M.; Fang, L.; Li, S.; Plaza, A.; Plaza, J. Feature Extraction with Multiscale Covariance Maps for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 755–769.
[13] Chen, Y.; Zhu, K.; Zhu, L.; He, X.; Ghamisi, P.; Benediktsson, J.A. Automatic Design of Convolutional Neural Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7048–7066.
[14] Chen, Y.; Zhu, L.; Ghamisi, P.; Jia, X.; Li, G.; Tang, L. Hyperspectral Images Classification with Gabor Filtering and Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2355–2359.
[15] Zhang, X.; Wang, Y.; Zhang, N.; Xu, D.; Luo, H.; Chen, B.; Ben, G. Spectral–Spatial Fractal Residual Convolutional Neural Network with Data Balance Augmentation for Hyperspectral Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10473–10487.
[16] Gao, H.; Zhang, Y.; Chen, Z.; Li, C. A Multiscale Dual-Branch Feature Fusion and Attention Network for Hyperspectral Images Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8180–8192.

[17] Seydgar, M.; Naeini, A.A.; Zhang, M.; Li, W.; Satari, M. 3-D Convolution-Recurrent Networks for Spectral-Spatial Classification of Hyperspectral Images. *Remote Sens.* **2019**, *11*, 883.

[18] Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251.

[19] Hamida, A.B.; Benoit, A.; Lambert, P.; Amar, C.B. 3-D deep learning approach for remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4420–4434.

[20] Al-Alimi, D.; Cai, Z.; Al-Qaness, M.A.; Alawamy, E.A.; Alalimi, A. ETR: Enhancing transformation reduction for reducing dimensionality and classification complexity in hyperspectral images. *Expert Syst. Appl.* **2023**, *213*, 118971.

