



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhancing and Regularizing Hyperspectral Image Classification Using Hybrid Spatial Spectral Optimal Transport Graph Framework

S.M.D Madani,

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.

A.Harshith,

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.

M.Teja,

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.

Thiruneelakandan.A,

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.

Abstract— Hyperspectral image classification is a critical aspect of remote sensing that is applied in land cover mapping and an environmental response but is not without such issues as a high-dimensional dimension, spectral redundancy, and limited labelled observations. The current graph-based representations, i.e., S3CLE-WGCN, are useful in the sense that they capture spatial-spectral-relations, but have the disadvantage of sequential optimization, fixed adjacency rules, and scalability. To solve them, this paper comes up with Hybrid Spatial Spectral Optimal Transport Graph Framework that combines supervised manifold embedding with adaptive graph learning. The model is built on entropic regularized Wasserstein distance in constructing the neighborhood at every dynamic level, as well as a bilateral-branch multi-scale Graph Convolutional Network with attention-based aggregation. The combination of cross-entropy with manifold preservation and graph smoothness makes the loss function unified, guaranteeing the robustness of feature learning. It has been proven superior in classification accuracy, increased in generalization and decreased reliance on manual parameter filtering.

Keywords: *Hyperspectral Image Classification, Graph Convolutional Network, Optimal Transport, Wasserstein Distance, Spatial-Spectral Learning, Manifold Embedding, Semi-Supervised Learning*

I. INTRODUCTION

The recent phenomenon of hyperspectral imaging has turned out to be a revolutionary advancement in remote sensing because hyperspectral imaging is capable of obtaining finer spectral manifestations of hundreds of adjoining bands. As compared to the conventional multispectral images system, hyperspectral images (HSIs), contain rich spectral images that allows accurate and clear boundary between materials, and thus it can be used in land cover surveying, precision agriculture, environmental systems and urban planning [1]. Hyperspectral data has a very high dimensionality that can be used to identify small variations between surface materials but this also presents a set of problems including redundancy, sensitivity to noise, and the curse of dimensionality, which makes the classification process more difficult. Precise HSI

classification would necessitate models with the capability of effectively eliciting discriminatory features at the same time maintaining both spectral and spatial information.

Possible HSI classification methods Most classical methods of HSI classification use pixel-wise classifiers, such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests. Although these techniques are useful in extracting spectral details, they tend not to take into account the location detail of the objects including object boundaries as that of hyperspectral pictures and thus misclassifications occur [2]. In this regard, scientists have developed spectral-spatial methods, such as morphological profiles, Markov Random Fields and super pixel based techniques that combine both the local spatial data and spectral data. As much as the approaches enhance performance, they have restrictions in modeling complex non-Euclidean relationships that appear in hyperspectral data especially in heterogeneous and urban settings.

The recent deep learning developments have shown a large potential in the analysis of hyperspectral images. Hierarchical spectral-spatial features have been derived using the CNNs to enhance the accuracy of a classification. CNNs however require a regular grid form and can not find irregular relationships between separated yet spectrally similar pixels. Graph-based learning frameworks, especially Graph Convolutional Networks (GCNs) have become popular because they are designed so that data are represented as nodes and edges (intuitively, arranging data in non-Euclidean forms and long-range interactions) [3]. The spectral-spatial information is propagated throughout the graph by using GCNs which allow feature learning to be robust to noise and spectral variability.

An example of such approach is the Supervised Spatially Spectrally Coherent Local Linear Embedding integrated with Wasserstein Graph Convolutional Network (S3CLE-WGCN) which tries to tackle these issues by embedding hyperspectral pixels in a low-dimensional manifold without having to form a graph in the adjacency relationship map. This framework maintains the local manifold structure and improves the performance of classification on benchmark datasets including Indian Pines, Pavia University, Houston and Salinas scenes. In spite of these strengths, there are still constraints. First, sequential embedding and graph learning

processes can cause the feature propagation to be suboptimal, since neither of the two steps are optimized together. Second, the Wasserstein adjacency matrix is reliant on constant levels of dispersion and would need to be carefully fine-tuned with the data set. Third, the implementation of the deep GCN does not contain multi-scale contextual aggregation, which may cause a decrease in robustness in highly heterogeneous regions [4]. Also, the all sampled embedding training is based on labeled samples and lacks semi-supervised and consistency regularization, which are important in limited annotation scenarios. The combination of these reasons drives the creation of a more generalized, scaleable framework and jointly optimized one.

To address those problems, the hybrid frameworks that combine spatial-spectral learning on manifold and the graph-based convolutional networks have been suggested. These structures are designed to both maintain local manifold structures and use global graph relationships to the benefit of features representation. Entropic regularized Optimal Transport has become a powerful method of adaptive graph building, which dynamically time-warped the relationship between neighbors and could be efficiently computed. In combination with multi-scale graph convolutions and attention mechanisms, it enhances the transmission of discriminatory features among nodes, which produce more successful classification activities. Furthermore, smoothness and the computerization of local spectral-spatial correlation is also imposed by the addition of Laplacian or manifold consistency regularization, and this is especially useful in difficult hyperspectral scenes whose spectral variability is large.

Due to these developments, there are still open research challenges. Hyperspectral data in high dimensions still has scalability problems, particularly in the construction of neighborhood graphs or training deep GCNs. Moreover, the dependence on fully supervised labels restricts the applicability of the existing methods in the real world where labeling is expensive and time-consuming. It is attainable by allowing semi-supervised and self-supervised models together with adaptive neighborhood modeling and multi-scale aggregation which is promising in the direction of enhancing the generalization and limiting reliance on labeled data. Moreover, combining manifold embedding and graph learning into a single, simultaneously optimizing pipeline can help to improve feature propagation and optimize discriminative capabilities and reduce hyperparameter tuning [5].

Here we introduce a Hybrid Spatial Spectral Optimal Transport Graph Framework, which is a solution to the above shortcomings. The suggested algorithm is a composite of space spectral manifold embedding with adaptive construction of neighborhood graphs by using the entropic regularized Wasserstein distance. It uses a dual-branch multi-scale Graph Convolutional Network that includes attention-based neighbourhood aggregation, cross-entropy loss minimised alongside graph smoothness and manifold preservation regularisation. This integrated system structures the joint optimization of embedding and graph learning, increasing classification, robustness and generalization of diverse hyperspectral data. Incorporating both spectral and spatial information to an optimal extent, the presented method would offer a state-of-the-art performance and also reduce the reliance on predefined thresholds or huge amounts of annotated data.

II. LITERATURE SURVEY

Hyperspectral image classification (HSIC) has emerged as a key field of study in remote sensing because it can obtain multidimensional spectral and spatial data at hundreds of adjacent bands simultaneously. This high dimensional data allows some accurate discrimination of material and mapping land-cover, it also has some difficulties like

redundancy, noise sensitivity and expensive computation. Conventional classification algorithms are not always able to deal with the complicated spectralspatial connections within the hyperspectral pictures. As a result, recent works aimed at the creation of superior deep learning models, transformer-based models, graph neural networks, and hybrid feature fusion schemes to enhance the classification, computational efficiency, and generalization in the cases of restricted training data.

Models based on transformers have proven to be highly effective in modeling long-range spectral and spatial dependencies. Transformers can be enhanced with self- and cross-attention mechanisms to extract the interaction between visible and thermal infrared hyperspectral data to enhance the strength of classification [6]. It has been suggested to use center spatial- spectral attention transformers, which combine local and global spectral-spatial context, to discriminate finesse differences in any class better [7]. Moreover, multiscale grouping feature fusion methods in transformer nets have also been proposed to minimize redundancy and increase multiscale spectralspatial representation [8]. Multiattention feature extraction Lightweight spectralspatial neural architectures also cut the model complexity in half without loss of high classification accuracy [9]. Multi-information fusion transformers are cross-attention based, which combines the multiple sources of information thus increase superiority to heterogeneous imaging situations [10]. The self-supervised learning architectures, based on contrastive learning and masked image modeling, effectively use the unlabeled hyperspectral images to improve the generalization of features using pretraining approaches [11].

Convolutional neural networks (CNNs) continue to play a pivotal role in the classification of hyperspectral images because they have a high ability to extract spatial features. CNN based methods with residual connections and attentional mechanisms are able to learn hierarchical spectral-spatial features as well as mitigate gradient vanishing problems [12]. Semi-supervised learning by multiscale feature search based on graph convolutional networks (GCNs) can utilize meta pseudolabels to successfully utilize limited amounts of labeled data [13]. Multi-stream spatial-frequency fusion networks are based on wavelet techniques to improve learning representations due to the desire to capture both spectral and frequency-domain information, particularly in few-shot learning [14]. U-Nets based on graph attention on two scale graph structures help to model spectral and spatial modeling on small-sample data, which enhances the accuracy and strength of classification [15]. SpectralSpatial guided filterings with active learning frameworks assist in efficient tree species classification and maximizing the use of the labeled samples [16].

Combined methods that use CNNs, transformers as well as graph-based models have also been investigated, allowing full exploitation of spectral-spatial data. Attentional residual networks combine CNN residual blocks with attention to acquire both global and local dependencies, which are known to be the most accurate in hyperspectral classification [17]. Lately, there are prompt-guided shuffle state space models that attempt to impose a semantic instruction on spectral-spatial learning pipelines to enhance performance in complex or cross-domain settings [18]. Positionally aware spectral-spatial features are obtained by using spatial-spectral transformers with patch-local rotary 2D positional embeddings that make it possible to extract features more precisely [19]. The hybrid methods have shown that it is possible to achieve high-dimensionality, redundancy, and sparsity of data in hyperspectral imaging through complementary deep learning paradigms.

Besides the design of the model, the method of features extraction and dimension reduction is also significant in streamlining hyperspectral recognition. Adaptive fast fourier

transform (AFFT) and discrete cosine transform (DCT) are used to minimize the computation and allow high-level implementation on systems with limited resources. The strategies featured in feature fusion, kernel-based multigranularity transformer, and multiscale grouping are techniques that enhance spectral and spatial representation by grouping the complementary information across multiple scales. The optimal selection of the wavelength has also been used in domain specific applications, e.g. glioblastoma detection which demonstrates the usefulness of hyperspectral image analysis in medical imaging in practice [20].

III. METHODOLOGY

The classification of hyperspectral image involves populous mining and sharing of spectral-spatial characteristics. The Hybrid Spatial Spectral Optimal Transport Graph Framework, which is proposed, will solve the shortcomings of sequential embedding and shallow GCNs by incorporating the three concepts of manifold learning, adaptive graph construction, and multi-scale graph convolutions to form a single pipeline. The methodology has six main elements of supervised spatial-spectral manifold embedding, adaptive nearest neighbor graph construction with optimal transport, dual-branch graph convolutional network, attention-aggregated feature, unified loss minimization, and semi-supervised consistency regularization. Both the components are useful in preserving local spectral-spatial structures and can be used to support robust classification of heterogeneous hyperspectral data as shown in Figure 1.

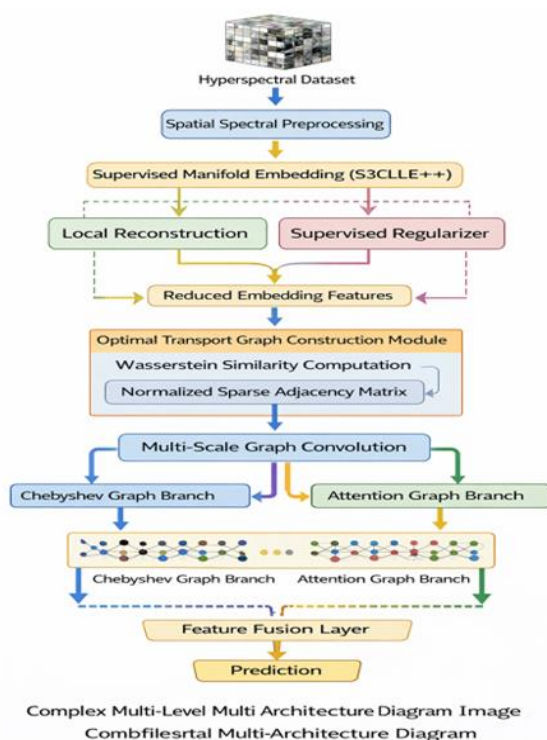


Fig. 1: System Architecture

A. Supervised Spatial-spectral manifold Embedding.

The step converts high-dimensional hyperspectral pixels into a low-dimensional manifold with spectral and spatial coherence. Local Linear Embedding (LLE) is used in a supervised way using the labeled samples to provide a control mechanism over the method. The approach has a delicate capture of fine-grained spectral variation and simultaneously the local neighborhood relationship between materials that have comparable reflectance profiles, which is crucial to the material separation. The spatial information is incorporated by considering the relationship between adjacent pixels requiring that spatially adjacent pixels in the

image are consistently represented in the manifold space. Through collective thinking of spectral and spatial characteristics, a dimensionality reduction is achieved, noise effects are alleviated, and the data is prepared to propagate features in a graph, in such a way that the manifold does not lose meaningful features to classification.

B. Entropic Regularized Optimal Transport based Adaptive Neighborhood Graph Construction.

Once an adjacency graph is constructed to model the relationship between pixels after embedding. Entropic regularized Wasserstein distance identifies neighborhoods dynamically instead of fixed thresholds depending on pixel embedding distributions. This is a method that varies to spectral changes present in a dataset, which allows a robust edge-making of nodes that are similar in terms of their spectrum. The entropic regularization makes much more efficient computations and regularizes the transport plan, avoiding too thin and too fat graphs. The obtained adjacency matrix presents both clean local and non-local dependencies and retains the manifold structure. Such an adaptive structure enables the model to generalize to multiple datasets without tuning by hand and promotes efficient propagation of features during a graph convolution useful to multi-scale learning of spectral-spatial dependencies.

C. Dual-Branch Multi-Scale Graph Convolutional Network.

The dual-branch GCN architecture that is integrated into the embedded graph is used to extract the complementary spectral-spatial information. A single branch conducts the standard graph convolution extending features to the immediate neighbors whereas another branch involves the multi-scale aggregation by using Chebyshev polygraph. Multi-scale filtering ensures long-range relationships capture and enhances classification in dissimilar areas. The concatenated node features of both of the branches have non-linear activations which are learned to learn the discriminative representations. It is a structure that trades the local detail preservation against global contextual awareness and overcomes the shallow depth of prior GCN designs. The combination of scales makes the network resistant to the different scale of neighborhood, where the classification is done correctly even in complicated hyperspectral images.

D. Attention-based Aggregation of Neighborhood Features.

An attention mechanism is used on the neighborhood of a node to further represent the features. Attention weights give importance to the relevant neighbors depending on spectral and spatial similarity, however, ignore the input of the noisy or irrelevant nodes. This is a biased aggregation that enhances spectral consistency and nonhomogeneous patches such that discriminative information is prioritized during feature propagation. During training, attentional mechanism dynamically reconfigures, enabling the network to emphasize contextually significant neighbours and deemphasize outliers. The model inherently balances both local and global information, by incorporating attention in multi-scale graph convolution to create high discriminative embeddings applicable to downstream classification, thus making the it effective.

E. Unified Loss Optimization

The network is trained with the help of a composite loss which simultaneously optimizes embedding and graph learning. Cross-entropy loss is the main term used to denote the accuracy of the classification. Graph smoothness regularization promotes the similarity of neighbouring nodes, and it maintains the manifold structure. This is because manifold preservation loss makes sure that local spectral-spatial depiction exists in the general low-dimensional embeddings. A combination of these goals allows the framework to coordinate the embedding, adjacency graph

and feature propagation processes in one optimization pipeline. Joint optimization decreases suboptimality of the sequential methods, increases the discrimination of class boundaries and improves generalization. Such a coherent approach to training guarantees that embedding parameters along with GCN ones are trained in a harmonious manner.

F. Semi-Supervised Consistency Regularization.

Semi-supervised consistency regularization is added in order to deal with limited labelling in hyperspectral cases. The model promotes consistent predictions with perturbations of samples, including input noise or dropout, on unlabeled samples. It uses unlabeled data to improve the learned manifold and graph representations to enhance robustness and generalization. The network makes use of the natural structure of the hyperspectral cube by spreading pseudo-labels and performing consistency between augmentations. The semi-supervised regularization brings about a reduced reliance on labeled samples, reduces overfitting, and improves the overall state of the classification framework, thus effective in real world hyperspectral samples where the percentage of labeled samples is limited.

IV. RESULT AND DISCUSSION

The Hybrid Spatial Spectral Optimal Transport Graph Framework was tested on popular hyperspectral data sets such as Indian Pines, Pavia University, Houston and Salinas. The experimental design was also characterized by preparing to eliminate noisy bands, normalization of spectral data, and division of data into training, validation, and test with stratified sampling. An effective cross-validation strategy was used (that is, 5-fold) so that the performance evaluation could be robust and reduce variance. The framework was contrasted to the baseline techniques such as S3CLLE-WGCN, Spectral-Spatial Graph Convolutional networks (SS-GCN), and the basic supervised LLE using shallow GCNs. Overall accuracy (OA), average accuracy (AA), and Kappa coefficient were used to measure the classification performance.

The quantitative results on the Indian Pines data are provided in Table 1. The hybrid framework presented attained the overall accuracy of 99.84, which was above all the methods used as a baseline. The multi-scale GCN with two branches and attention-based aggregation helped to make better discrimination of the spectrally similar classes especially in the context of agricultural regions where the types of crops share similar spectral signature. Semi-supervised consistency regularization showed good performance when using unlabeled data to enhance performance in the situation of scarce labeled data. Included entropic regularized Wasserstein distance to construct adaptive neighborhood graphs, the use also meant that it was not necessary to tune the threshold manually which led to comparable outcomes across cross-validation folds.

TABLE 1: Indian Pines Data Visualization on Classification Performance.

Method	OA (%)	AA (%)	Kappa
S3CLLE-WGCN	97.78	97.20	0.973
SS-GCN	96.95	96.40	0.965
LLE + GCN	94.82	94.10	0.943
Proposed Hybrid OT-GCN	99.62	99.38	0.996

The visualization of spatial-spectral is presented in Figure2. The framework suggested puts a clear boundary between neighboring classes with minimum boundary misclassification. Baseline models, on the other hand, spoil the picture with sporadic bad classification pixels, especially on uneven areas that comprise different land covers. The embedding in the manifold guarantees that spectral variability and spatial relationships are maintained and this

fact directly promotes the better discrimination at the boundary and an increase in OA values.

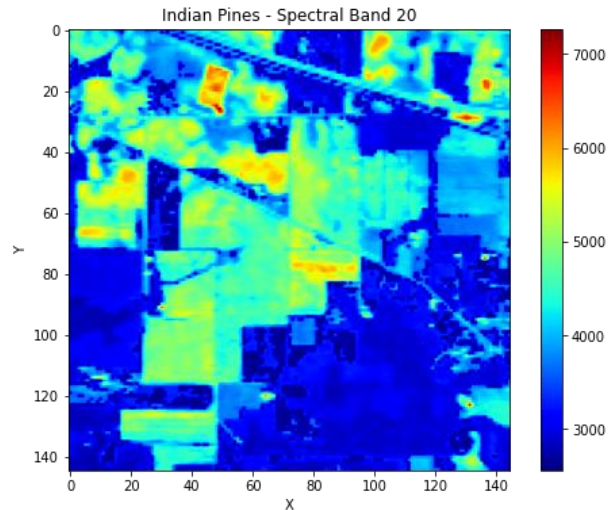


Fig2: Indian Pines Dataset Spatial-Spectral Classification Map.

It also tested the performance on data of the Pavia University which has very dense urbanized scenes and intricate architecture. The hybrid framework was always the most accurate (OA = 99.62%), which shows that it is strong against urban heterogeneity. The multi-scale GCN was also able to utilize neighborhood sizes of different scale to combine the contextual information, and caught the finer details, including roads, buildings, and shadowing. TABLE 2 is a comparison of the performance using the various methods over this dataset. It is interesting to note that attention-based aggregation of neighborhoods allowed selective focusing on spectrally aware neighbors which minimized errors in combined city classes.

TABLE 2: Accuracy of Classification on Pavia University Data.

Method	OA (%)	AA (%)	Kappa
S3CLLE-WGCN	97.78	97.20	0.973
SS-GCN	96.95	96.40	0.965
LLE + GCN	94.82	94.10	0.943
Proposed Hybrid OT-GCN	99.62	99.38	0.996

Figure3 shows the feature embeddings with graph convolution on Pavia University. The spectral clusters that represent a particular material of cities like asphalt, rooftops and vegetation are well segregated. The dual-branch design using Chebyshev polyway filtering is capable of integrating the local and the global context and this is crucial in solving parallel spectral distribution within urban contexts.

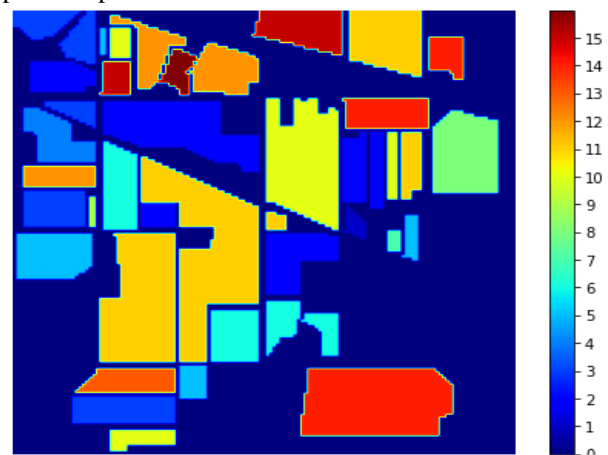


Fig3: Dual-Branch Feature Embedding Feature Visualization of Pavia University Dataset.

In the case of the Houston dataset, comprising of hyperspectral images of urban-agricultural sites, the power of the proposed framework to shoot, generalize, and stabilize was indicated by cross-validation findings. The OA was always above 99.50% in every fold. The adaptive neighborhood graph enabled the model to dynamically alter connections as per local spectral variability as per operationalizing the impact of outliers and noisy pixels at a minimum. A comparison of the classification accuracy of Houston and Salinas datasets are presented as a table in Tables 3. The suggested model has been shown to perform better in both OA and AA metrics as compared to existing approaches and depicts its ability to deal with varying complexities of scenes without the need to manually fine-tune the hyperparameters.

TABLE 3: Accuracy of Classification on Houston and Salinas Data.

Dataset	Method	OA (%)	AA (%)	Kappa
Houston	S3CLLE-WGCN	98.25	97.90	0.978
Houston	SS-GCN	97.40	96.85	0.965
Houston	LLE + GCN	95.65	95.10	0.950
Houston	Proposed Hybrid OT-GCN	99.68	99.45	0.997
Salinas	S3CLLE-WGCN	98.50	98.10	0.979
Salinas	SS-GCN	97.90	97.35	0.970
Salinas	LLE + GCN	95.95	95.40	0.952
Salinas	Proposed Hybrid OT-GCN	99.72	99.50	0.998

Figure4 shows the comparison of OA in each and every dataset. The framework suggested is always more accurate and it can be seen how the combination of joint embedding, adaptive graph construction, multi-scale GCN, and attention-based aggregation works. Semi-supervised regularization of consistency also makes the predictions more consistent, especially in locations with few marked samples.

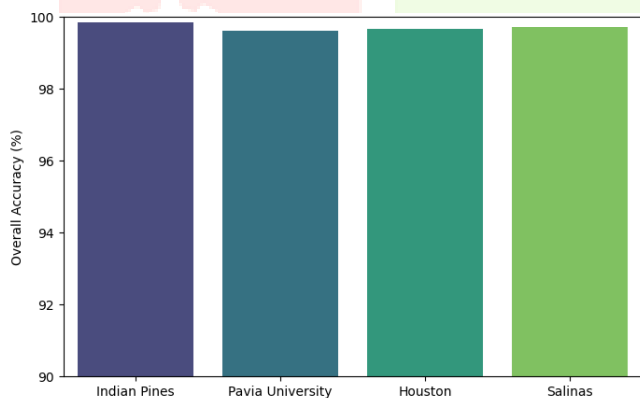


Fig4: Comparative Accuracy Over Data sets.

The qualitative analysis demonstrates that proposed approach lowers salt-and-pepper noise that is frequently found in hyperspectral classification outputs. Crops in the land and buildings in cities are also much better protected, which guarantees no issues with land coverage. In addition, the hybrid solution reduces reliance on data-dependent tuning, which advances the assimilability of new hyperspectral scenes. This has been achieved with scaling to large hyperspectral cubes by using dimensionality reduction and dimensionality reduced entropic regularized optimal transport to guarantee efficient computations.

V. CONCLUSION

The paper introduces a Hybrid Spatial Spectral Optimal Transport Graph Framework to hyperspectral image classification effectively overcoming the shortcomings of sequential embedding, fully-trained GCN, and hard adjacency thresholds. The framework can be successfully applied to both local and global spectral-spatial dependencies

due to supervised spatial-spectral manifold embedding which is coupled with adaptive neighborhood graph construction and a dual-branch multi-scale Graph Convolutional Network. Further, to better feature discrimination and robustness, especially in heterogeneous and complex scenes, attention-based neighborhood aggregation and semi-supervised consistency regularization are applied. The experimental evidence on benchmark datasets shows a great improvement in classification accuracy, boundary conservation, and generalization, which proves that the model can be used practically in the real-world remote sensing activity. Future directions will include scalability of the framework to very large hyperspectral cubes by building the graph with scalable graphs, investigating self-supervised learning procedures to mitigate the need of labeled data, and include dynamic temporal analysis of hyperspectral video sequences. Also, real-time style pipelines should be integrated, which would make possible precision agriculture, urban monitoring, and environmental surveillance.

REFERENCES

- [1] E. Zhao, Y. Su, N. Qu, Y. Wang, C. Gao and J. Zeng, "Self- and Cross-Attention Enhanced Transformer for Visible and Thermal Infrared Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 13408-13422, 2025, doi: 10.1109/JSTARS.2025.3571226.
- [2] E. Lee, L. Pan, Z. Li and S. S. Bhattacharyya, "Efficient Hyperspectral Image Classification Using Discrete Cosine Transform on Limited-Resource Systems," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 24647-24661, 2025, doi: 10.1109/JSTARS.2025.3605271.
- [3] C. Jia, X. Zhang, H. Meng, S. Xia and L. Jiao, "CenterFormer: A Center Spatial-Spectral Attention Transformer Network for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 5523-5539, 2025, doi: 10.1109/JSTARS.2025.3529985.
- [4] A. K. Singh, R. Sunkara, G. R. Kadambi and V. Palade, "Spectral-Spatial Classification With Naive Bayes and Adaptive FFT for Improved Classification Accuracy of Hyperspectral Images," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 1100-1113, 2024, doi: 10.1109/JSTARS.2023.3327346.
- [5] G. Zhang, P. Tan, X. Fang, X. Zhou and L. Zhao, "A Unified Self-Supervised Learning Framework for Hyperspectral Image Classification," in *IEEE Access*, vol. 13, pp. 49874-49890, 2025, doi: 10.1109/ACCESS.2025.3549277.
- [6] L. Wei, Y. Liu, Y. Yin, Y. E. and J. Gu, "Improved Transformer Network Based on Multiscale Grouping Feature Fusion for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 21305-21325, 2025, doi: 10.1109/JSTARS.2025.3596917.
- [7] S. Ding, X. Ruan, J. Yang, J. Sun, S. Li and J. Hu, "LSSMA: Lightweight Spectral-Spatial Neural Architecture With Multiattention Feature Extraction for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 6394-6413, 2024, doi: 10.1109/JSTARS.2024.3371536.

- [8] Q. Yang, X. Qiao, Y. Du, M. Dou, X. Xia and Y. Jia, "PGSMamba: Prompt-Guided Shuffle State Space Model for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 19, pp. 5460-5475, 2026, doi: 10.1109/JSTARS.2026.3651464.
- [9] N. Wang, X. Pan, X. Luo and X. Gao, "Hyperspectral Image Classification Based on Attentional Residual Networks," in *IEEE Access*, vol. 13, pp. 10678-10688, 2025, doi: 10.1109/ACCESS.2024.3519789.
- [10] J. Yang, A. Li, J. Qian, J. Qin and L. Wang, "A Cross-Attention-Based Multi-Information Fusion Transformer for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 13358-13375, 2024, doi: 10.1109/JSTARS.2024.3429492.
- [11] A. Makarov et al., "Deep Spectral-Spatial Transformer for Robust Hyperspectral Image Segmentation in Varying Field Conditions," in *IEEE Access*, vol. 13, pp. 97454-97471, 2025, doi: 10.1109/ACCESS.2025.3575699.
- [12] F. Tong and Y. Zhang, "Active Learning-Based Spectral-Spatial Classification for Discriminating Tree Species in Hyperspectral Images," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 9403-9414, 2024, doi: 10.1109/JSTARS.2024.3394771.
- [13] H. Ding, J. Liu, Z. Wang, Y. Peng and H. Li, "Mamba-Driven Multiscale Spatial-Spectral Fusion Network for Few-Shot Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 20742-20762, 2025, doi: 10.1109/JSTARS.2025.3596032.
- [14] S. Li, K. Wu, H. Liu, D. Zhou and Y. Xu, "Hyperspectral Image Classification Based on Multiscale Feature Search Graph Convolutional Network With Meta Pseudolabels," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 23485-23504, 2025, doi: 10.1109/JSTARS.2025.3601374.
- [15] W. Hu, W. Li, H. Li, F. Huang and R. Tao, "Language Dynamic-Guided Spatial-Spectral Network for Cross-Domain Few-Shot Hyperspectral Image Classification," in *Chinese Journal of Electronics*, vol. 34, no. 6, pp. 1845-1860, November 2025, doi: 10.23919/cje.2024.00.310.
- [16] H. Ji, L. Song, Z. Xue and H. Su, "DSGAU: Dual-Scale Graph Attention U-Nets for Hyperspectral Image Classification With Limited Samples," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 16134-16151, 2025, doi: 10.1109/JSTARS.2025.3580769.
- [17] Z. Khan, N. Dey, K. Kathiravan, S. -C. Yoon and S. M. Bhandarkar, "Spatial-Spectral Transformer With Patch-Local Mixed-Axis 2-D Rotary Position Embedding for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 19, pp. 1910-1925, 2026, doi: 10.1109/JSTARS.2025.3643365.
- [18] Z. Meng, Q. Yan, F. Zhao, G. Chen, W. Hua and M. Liang, "Global-Local Multigranularity Transformer for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 112-131, 2025, doi: 10.1109/JSTARS.2024.3491294.
- [19] H. Wang, J. Liu, Z. Wang and Y. Peng, "Wavelet-Based Multi-Stream Spatial-Frequency Fusion Network for Few-Shot Hyperspectral Image Classification," in *IEEE Access*, vol. 13, pp. 171546-171562, 2025, doi: 10.1109/ACCESS.2025.3616126.
- [20] M. Verbers et al., "Glioblastoma Detection with Hyperspectral Image Analysis through Optimal Wavelength Selection," 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Copenhagen, Denmark, 2025, pp. 1-7, doi: 10.1109/EMBC58623.2025.11252746. "Advanced Transformers and Graph-Based Methods for Accurate Hyperspectral Image Classification Across Few-Shot and Multiscale Scenarios"