IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 1, 2025

Hyperspectral Image Classification Using Diffusion Model

Habiba Banu S
Dept. of Computer Science and Engineering
Mepco Schlenk Engineering College, Sivakasi, India
habibabanu281101_mcs25@mepcoeng.acin

Ruba Soundar K
Associate Professor (Sr.Grade)
Dept. of Computer Science and Engineering
Mepco Schlenk Engineering College, Sivakasi, India
rubasoundar@mepcoeng.ac.in

***Abstract*—** Hyperspectral image (HSI) classification is extensively used in Earth science and is important for remote sensing. Many deep learning techniques have been developed recently for HSI classification; nevertheless, difficulties are frequently encountered with high-dimensional and complex data, making it challenging for relationships between various data elements to be captured. To address this, a novel method, dubbed "SpectralDiff," is proposed, which employs diffusion models for HSI classification. In this approach, noise in the data is repeatedly reduced, creating a clearer representation of the data's structure, thereby facilitating the handling of redundant and high-dimensional data. The framework consists of two major components:

Spectral-Spatial Diffusion Module: The establishment of connections between data samples is facilitated by the spectral-spatial diffusion module, without requiring prior knowledge of the structure. Spatial (position-related) and spectral (color-related) information from the HSI data is extracted. Attention-Based Classification Module: The features gleaned from the diffusion module are then used to classify each pixel in the image. This approach, which emphasizes the connections between multiple samples, enables better classification. Tests conducted on three publicly available datasets demonstrate that SpectralDiff achieves superior performance compared to other state-of-the-art techniques.

Key Words— Diffusion Models, Feature Extraction, Deep generative model, Deep Neural Network (DNN), spectral-spatial diffusion, Hyperspectral Image (HSI) Classification.

I. Introduction

Hyperspectral imaging (HSI) is an advanced technology designed to capture high-resolution spectral information from various objects. Through the integration of spatial and spectral reflectance data, each pixel in an HSI corresponds to a unique spectral curve, offering detailed insights into material differentiation and identification. This technology surpasses human visual capabilities by covering a broader spectral detection range, thereby facilitating a comprehensive understanding of natural phenomena. HSI has found considerable applications across diverse fields, including environmental management, agriculture, ecology, geology, urban planning, and oceanography. One of its most critical applications is the classification of HSI, which assigns pixels to specific land-cover categories such as soil or vegetation. This process is essential to many hyperspectral imaging tasks.

The challenge posed by the high dimensionality of HSI data complicates accurate pixel classification. Given the hundreds of spectral bands and the vast amounts of data involved, identifying relevant features is complex. To address this, several feature extraction methods have been developed to map spectral vectors from high-dimensional space to lower-dimensional feature spaces, including classic statistical techniques like principal component analysis (PCA), minimum noise fraction (MNF), local preserving projection (LPP), linear discriminant analysis (LDA), independent component analysis (ICA), and sparse preserving projection (SPP). However, these methods are limited by the spatial heterogeneity and homogeneity inherent to HSIs, which makes extracting spectral features alone insufficient for optimal utilization.

To improve feature extraction, methods that jointly capture spatial and spectral features have been proposed, such as the extended morphological profile (EMP) and the extended attribute profile (EAP). The introduction and rapid development of deep neural networks (DNNs) have further enhanced the accuracy of HSI classification by automatically learning complex features. DNN-based methods have yielded significant improvements in image classification, segmentation, and object detection, among other areas. Several techniques, including stacked autoencoders, deep fully convolutional networks, and deep prototypical networks, have been applied to HSI classification, effectively leveraging the spatial-spectral nature of hyperspectral data.

Despite these advancements, DNN-based methods struggle to effectively model spectral-spatial relationships across samples. Current approaches primarily rely on graph neural networks (GNNs) to model these relationships, but designing graph structures or neighborhood information adds complexity and introduces subjectivity into the process. GNNs also fail to fully capture the spectral-spatial distribution of HSI data, limiting their perceptiveness toward contextual features.

To address this limitation, a generative framework based on diffusion models, named SpectralDiff, has been proposed. This framework reconstructs the data generation process through iterative denoising, capturing spectral-spatial distribution information more effectively. SpectralDiff consists of two key modules: the spectral-spatial diffusion module and the attention-based classification module. The former employs a Markov process to construct a distribution of hyperspectral cube data, adding

Gaussian noise in the forward process and removing it in the reverse process through a spectral-spatial denoising network. The relationships between samples are constructed using the hidden variables generated in this diffusion process. The attention-based classification module then uses the extracted features to generate per-pixel classification results, improving cross-sample perception and overall classification performance.

The innovative aspect of this approach lies in adopting a generative perspective for constructing sample relationships. By modeling the sample generation process, the proposed framework avoids the need for predefined graph structures or neighborhood information, making it adaptable and independent for future developments. Experimental results demonstrate the superiority of this method compared to existing state-of-the-art techniques, and ablation experiments confirm the effectiveness of the spectral-spatial diffusion features.

II. LITERATURE SURVEY

The study [1] examines the impact of Principal Component Analysis (PCA) on feature discrimination for pattern classification, particularly in hyperspectral data, through both theoretical and experimental approaches. It evaluates PCA as a standalone dimensionality reduction tool and as a preprocessing step in subspace Linear Discriminant Analysis (LDA). The study provides a comprehensive theoretical foundation and empirical validation of the limitations associated with PCA and subspace LDA in hyperspectral target recognition. It highlights the necessity foralternative methods to address small-sample-size issues in high-dimensional spaces. However, the research primarilyfocuses on hyperspectral data, which may limit the generalizability of the findings to other data types. Additionally, the experimental evidence may not encompass all potential use cases or variations in hyperspectral imaging scenarios,potentially overlooking contexts where PCA might still be effective.

This work [2] applies a two-step process combining Minimum Noise Fractions (MNF) for dimensionality reduction with Fast and Adaptive Bidimensional Empirical Mode Decomposition (FABEMD) as a low-pass filter to denoise hyperspectral images. The informative components are then classified using a Support Vector Machine (SVM).

The approach effectively removes noise and enhances classification accuracy, achieving up to 98.14%, while demonstratingstable performance across varying settings and spectral information extraction. However, it faces limitations due to itshigh complexity, necessitating an exhaustive key search to identify the optimal number of components (BIMFs) to remove, along with a lack of automation in selecting the best BIMFs.

The methodology [3] focuses on the automatic construction of Extended Attribute Profiles (EAP) based on the standard deviation attribute, utilizing statistics from training samples to guide filter parameters. This approach combines unsupervised and supervised feature reduction techniques, such as PCA, KPCA, NWFE, and DBFE, for hyperspectral image classification. The proposed MNF-FABEMD process effectively eliminates noise and enhancesclassification accuracy, achieving up to 98.14% overall accuracy while maintaining stable performance without the need for parameter tuning. However, the two-step process adds complexity, as determining the optimal number of BIMFs to remove currently necessitates an exhaustive search, indicating a potential area for automation.

The paper [4] introduces a Geometric-Spectral Reconstruction Learning (GSRL) method for open-set classification of remote sensing images, leveraging hyperspectral (HSI) and LiDAR data. This method consists of two primary modules: a geometric-spectral reconstruction module that learns and reconstructs geometric-spectral features, and a geometric-spectral open-set adaptation module that employs extreme value analysis to identify unknown classes by comparing reconstructed and original feature matrices. The GSRL method effectively distinguishes between known and unknown classes, thereby enhancing its practical application in real-world scenarios. However, it faces challengesin complex environments with highly diverse unknown classes, and its effectiveness may be limited when extreme value analysis is inadequate for accurately differentiating known from unknown classes.

The authors [5] enhanced Denoising Diffusion Probabilistic Models (DDPMs) by learning reverse process variancesthrough a reparameterization technique and a hybrid learning objective that combines the Variational Lower Bound (VLB) with the simplified objective proposed by Ho et al. (2020). This approach effectively reduces gradient noise, improves log-likelihood, and facilitates faster sampling. Notably, it significantly reduces the number of sampling steps, making the models more practical for real-world applications, while also enhancing mode coverage comparedto Generative Adversarial Networks (GANs), as evidenced by higher recall in precision-recall metrics. However, the method requires substantial computational resources for training, especially on high-diversity datasets like ImageNet,and incorporating learned variances and hybrid objectives can introduce additional complexity in training, particularly concerning gradient noise

The method [6] uses Adaptive Spatial Pyramid Constraint (ASPC) for classifying hyperspectral images (HSIs) by first measuring image complexity through edge detection. It then segments the image into various scales using a spatial pyramid structure. For classification, the method applies different loss functions to labeled and unlabeled regions, enhancing accuracy even when only a few labeled samples are available. The ASPC approach improves model performance by leveraging spatial-spectral correlations and adapts segmentation scales based on image complexity, ensuring efficient utilization of spatial information. However, its effectiveness is contingent upon the accuracy of edge detection for complexity evaluation, which may vary depending on image quality.

This work [7] introduces SpectralFormer, a transformer-based backbone network designed for hyperspectral image (HSI) classification. The model incorporates two innovative modules: Groupwise Spectral Embedding (GSE), which learns local spectral representations, and Cross-layer Adaptive Fusion (CAF), which transfers memory-like information between layers. Extensive experiments on three HSI datasets demonstrate that SpectralFormer significantly improves classification accuracy, achieving around 10% higher overall accuracy (OA) compared to CNNs, RNNs, and traditional transformers. It is also flexible for both pixel-wise and patch-wise inputs. However, transformers may still struggle to capture fine local context as effectively as CNNs, and the cross-layer skip connections may introduce added complexity without fully addressing the loss of local context.

The Context-Aware Dynamic Graph Convolutional Network (CAD-GCN) [8] for hyperspectral image (HSI) classification models pixel relationships through graph-based structures, enabling it to capture long-range connections between image regions and dynamically update these connections to enhance classification accuracy. This approach allows CAD-GCN to effectively capture contextual relations, which is particularly beneficial in complex or inhomogeneous regions, leading to improved classification performance. However, the dynamic graph updating process increases computational complexity, potentially resulting in longer training times.

The Spectral Spatial Graph Attention Network (SSGAT) method [9] for hyperspectral image (HSI) classification constructs a graph where each HSI sample is a node, with neighboring nodes connected based on spectral and spatial information. By computing attention between nodes, SSGAT aggregates essential features, enhancing classification accuracy. Leveraging semi-supervised learning, it combines labeled and unlabeled data, effectively addressing scenarios with limited labeled samples. Tested on public datasets, SSGAT demonstrates superior performance compared to other methods. While it improves classification by combining spectral and spatial information, the method's reliance on constructing large neighborhood graphs may increase computational complexity, and its performance can be sensitive to the neighborhood size.

This methodology [10] introduces a Modified Locality-preserving Projection (MLPP) for hyperspectral image classification, which adaptively selects a varying number of nearest neighbors for each data point to maximize the distance between non-nearest neighbors, enhancing class separability. A parameter-free weighted graph is utilized to preserve local structures while boosting class discrimination, increasing robustness and flexibility without user-defined parameters However, the adaptive neighbor selection may add computational complexity, particularly for large datasets, and the method's testing on only two datasets limits its generalizability assessment.

Summary of Literature:

1. **Dimensionality Reduction with PCA and LDA**: Studies have explored Principal Component Analysis (PCA) as a dimensionality reduction technique, both as a standalone tool and as a preprocessing step for Linear Discriminant Analysis (LDA). Although effective in reducing dimensions, PCA struggles with small-sample-size issues in high- dimensional data typical of HSI, highlighting the need for alternative approachesoise Reduction and Feature Extraction: A combination of Minimum Noise Fractions (MNF) and Fast Adaptive Bidimensional Empirical Mode Decomposition (FABEMD) has proven effective for denoising HSIs, achieving a high classification accuracy. However, this two-step process is computationally complex and lacks automation in selecting optimal components fornoise removal .

*2.***Reconstruction:** For open-set classification, a method called Geometric-Spectral Reconstruction Learning (GSRL) combines HSI and LiDAR data, effectively identifying unknown classes. However, the method struggles with diverseunknown classes, especially when extreme value analysis fails to differentiate well between known and unknown classes .

3.Transformer-Based : SpectralFormer is a transformer-based model that incorporates modules like Groupwise Spectral Embedding (GSE) and Cross-layer Adaptive Fusion (CAF) for local spectral representation and memory-like cross-layer information transfer. Despite achieving higher classification accuracy than CNNs and traditional transformers, it still faces limitations in capturing fine local context .

4.Graph-Based Networks: Con Dynamic Graph Convolutional Network (CAD-GCN) and Spectral Spatial Graph Attention Network (SSGAT) employ graph-based methods to capture spatial and spectral relationships between pixels. CAD-GCN dynamically updates connections for enhanced contextual relations, while SSGAT uses semi-supervised learning to combine labeled and unlabeled data. Both methods improve classification accuracy, though they are computationally intensive due to the need for large neighborhood graphs and dynamic updates .

5.Locality-Preserving Projection: Areserving Projection (MLPP) adaptively selects nearest neighbors for each datapoint, enhancing class separability. It shows improved robustness and flexibility but adds complexity due to adaptiveneighbor selection and has limited testing across datasets .

Best Spectral-Spatial Denoising Network : Among the denois, networks that combine **spectral and spatial information** like the **Spectral Spatial Graph Attention Network (SSGAT)** demonstrate strong performance by utilizing semi-supervised learning, which benefits situations with limited labeled data. SSGAT's use of attention mechanisms enhances feature aggregation from both spectral and spatial domains, effectively balancing denoising and classification accuracy. Additionally, **CAD-GCN** and **SpectralFormer** are strong candidates, given their effective handling of spatial context and spectral representation, though they involve high computational demands. Ultimately,SSGAT could be considered a top choice for spectral-spatial denoising in HSI classification when allows.

III. PROPOSED SYSTEM

Hyperspectral image (HSI) classification has become an essential tool for various remote sensing applications in Earth sciences, including environmental monitoring, agriculture, and urban planning. The classification process involves assigning each pixel within an HSI to a specific land-cover class based on its spectral and spatial features. However,the high dimensionality and redundancy inherent in HSI data present challenges for conventional classification methods. To address these

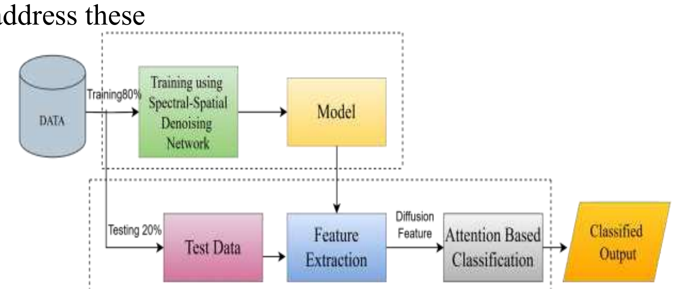


Fig : Poposed System

challenges, we propose a novel system called "SpectralDiff," which uses diffusion modelsto enhance the accuracy and robustness of HSI classification. This system comprises two primary modules: the Spectral-Spatial Diffusion Module and the Attention-Based Classification Module.

1. Data Input and Preprocessing

The system begins with the **Data Input** stage, where raw HSI data is collected. Hyperspectral data typically includeshundreds of contiguous spectral bands for each pixel, capturing detailed information about the materials present in each spatial location. However, the high dimensionality of HSI data introduces redundancy and noise, making direct classification challenging. Therefore, preprocessing steps like normalization and data augmentation are applied to prepare the data for further processing. This preprocessing ensures that the model can handle the variations in lighting,sensor conditions, and other environmental factors.

2. Spectral-Spatial Diffusion Module

The Spectral-Spatial Diffusion Module is the core of the SpectralDiff approach and addresses one of the main limitations of traditional HSI classification methods: capturing spatial-spectral relationships without relying on predefined graph structures. Unlike graph-based methods, which depend on manually designed structures that introduce complexity and subjectivity, the Spectral-Spatial Diffusion Module adopts a generative framework based on diffusion models. This approach is inspired by recent advances in deep generative models, which can effectively model complex data distributions through iterative denoising.

- **Diffusion Process**: The diffusion model operates through a forward and reverse process. In the forward process, Gaussian noise is incrementally added to the HSI data, simulating a random walk away from the original data distribution. This noisy data serves as an input to the reverse reverse diffusion process, where a **spectral-spatial denoising network** removes the noise step-by-step. Each denoising step reveals increasingly refined features, gradually reconstructing the original data's spectral-spatial structure.
- **Feature Extraction Through Denoising**: As the diffusion model denoises the data, it captures underlying structures in the HSI data, which correspond to spatial and spectral relationships between different pixels. These relationships are represented by hidden variables generated during the diffusion process, forming a basis for effective feature extraction.

By leveraging these denoised features, the Spectral-Spatial Diffusion Module highlights both spectral characteristics (such as material composition) and spatial context (such as texture and spatial patterns) in the data.

- The output of this module is a set of spectral-spatial features that significantly reduce the data's dimensionality whilepreserving essential information. This data transformation helps alleviate the challenges posed by the high dimensionality of HSI data, resulting in features that are both informative and compact.

3. Feature Extraction

Following the diffusion-based denoising process, the **Feature Extraction** stage further refines the spectral-spatial features for effective classification. Feature extraction techniques in HSI data aim to transform raw high-dimensionaldata into a lower-dimensional feature space that retains meaningful characteristics for classification.

- **Dimensionality Reduction and Spectral-Spatial Representation**: Dimensionality reduction methods, such as principal component analysis (PCA) or autoencoder-based methods, can be used here to condense the features extracted by the diffusion process. This helps eliminate any residual redundancy in the data, ensuringthat only the most significant spectral and spatial information is retained.
- **Enhanced Spatial Context**: This stage ensures that the features maintain a balance between spectral precision and spatial context, addressing the limitation of traditional pixel-based approaches that overlook spatial correlations.

By the end of this step, each pixel is represented by a set of spectral-spatial features that are well-suited for classification tasks, providing a robust foundation for the next module.

4. Attention-Based Classification Module

The **Attention-Based Classification Module** is designed to classify each pixel within the HSI by focusing on the most relevant features extracted from the diffusion module. Traditional classification models often struggle to prioritize important features, especially when handling high-dimensional HSI data. The attention mechanism addresses this limitation by selectively weighting features that contribute most to accurate classification.

- **Self-Attention Mechanism**: The attention mechanism emphasizes significant features by assigning higher weights to them, enhancing the model's focus on critical spatial-spectral relationships. This is particularly useful in hyperspectral data, where subtle variations in spectral values can signify important differences in material composition.
- **Cross-Sample Perception**: The attention module enables cross-sample perception by dynamically adjusting feature weights based on the context of neighboring pixels. This context-aware classification approach enhances the model's ability to identify land-cover classes with high precision, even when neighboring pixels have similar spectral signatures but belong to different classes.

The attention-based classifier outputs per-pixel classification results, where each pixel is assigned a category label. This pixel-wise classification allows for precise mapping of different land-cover types in the hyperspectral image, supporting applications that require detailed spatial analysis, such as environmental monitoring and urban planning.

5. Classified Hyperspectral Image Output

The **Classified Hyperspectral Image Output** is the final result of the SpectralDiff framework. The system producesan HSI where each pixel is classified into categories like vegetation, soil, or urban area. This output enables practitioners to visualize and analyze the distribution of various land-cover types across the imaged area.

The proposed SpectralDiff system offers several advantages over traditional HSI classification methods:

- **Adaptability**: Unlike graph-based methods, SpectralDiff does not rely on manually designed graphstructures, making it more flexible and adaptable to various datasets.
- **Improved Accuracy**: By leveraging a generative model with an attention-based classifier, SpectralDiff captures complex spatial-spectral relationships, leading to more accurate classifications.
- **Reduced Complexity**: The diffusion process simplifies the feature extraction process by automatically modeling sample relationships, reducing the need for complex, predefined structures

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 1, 2025

IV Experimental Settings

1. Datasets:
   Three hyperspectral datasets were used to evaluate the algorithm:
   - **Indian Pines (IP)**: Captured in Indiana, USA, with 145 × 145 pixels and 200 spectralbands.
   - **Pavia University (PU)**: Collected in Italy, with 610 × 340 pixels and 103 spectral bands.
   - **Salinas (SA)**: Collected in California, USA, with 512 × 217 pixels and 224 spectral bands.

2. **Evaluation Metrics**:The model's effectiveness was measured using Overall Accuracy (OA), Average Accuracy (AA), Kappa coefficient (κ), and classification accuracy for each land-cover category.

Overall Accuracy (OA):
It's a measure of how often the model makes the correct prediction.

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Population (TP + TN + FP + FN)}}$$

True Positives (TP): Correctly predicted positive instances.
*True Negatives (TN)***:** Correctly predicted negative instances.
False Positives (FP): Incorrectly predicted positive instances (also called Type I error).
False Negatives (FN): Incorrectly predicted negative instances (also called Type II error).

Kappa Coefficient(k):

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

po: The observed accuracy **pe:** The expected accuracy.

Average Accuracy (AA):
This is particularly useful when the dataset is imbalanced, i.e., when some classes have more samples than others

$$\text{AA} = \frac{1}{C} \sum_{i=1}^{C} \frac{\text{Correct predictions for class } i}{\text{Total samples in class } i}$$

Sum the accuracies of all classes and divide by the number of classes.

- C is the number of classes.
- The fraction for each class *i* represents the accuracy within that specific class (i.e., how well the model classified samples belonging to that class).

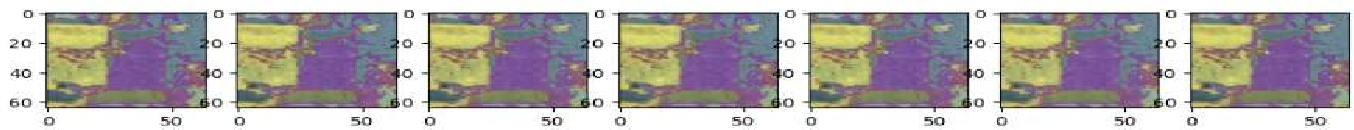
Hyperspectral images often involve many different land cover types, materials, or objects that may not be evenly distributed, average accuracy ensures that smaller or minority classes are not overshadowed by larger classes, which can happen when using overall accuracy.
Average accuracy gives equal importance to all classes, making it a fairer metric in many cases.

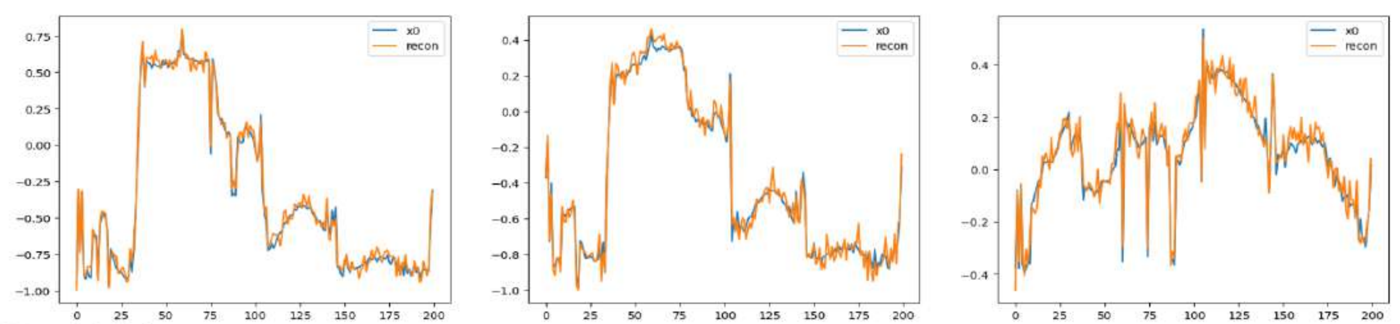
3. **Training Setup**:The model was trained on a high-performance computing system with an AMD EPYC CPUand dual NVIDIA RTX 3090 GPUs.The diffusion model used the Adam optimizer, a learning rate of 1e-4, and a batch size of 256.The classification model used the Adam optimizer with a learning rate of 1e-3 and abatch size of 64. The model converged in under 50 epochs for all datasets.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 1, 2025

V. RESULTS AND DISCUSSION

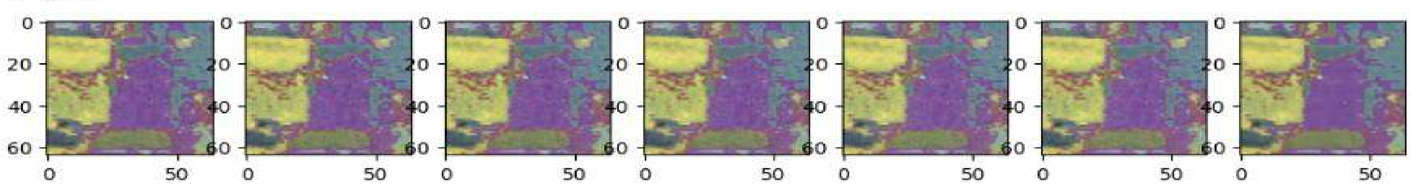
```
ori data load shape is (145, 145, 200) (145, 145)
[data] load data shape data=(145, 145, 200), label=(145, 145)
[data light split] data patches shape data=(9, 64, 64, 200), label=(9,)
------[data] after transpose train, test------
X.shape= (9, 1, 200, 64, 64)
Y.shape= (9,)
model path is  ./save_model/IP_diffusion/unet3d_10.pkl
load model done. model_path=./save_feature/Indian_unet3d_patch64_without_downsample_kernal5_fix_10/save_feature
<ipython-input-8-9004860b107a>:225: FutureWarning: You are using `torch.load` with `weights_only=False` (the current default value)
  model.load_state_dict(torch.load(model_path, map_location=device))
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=5, index=0 done.
save full matrix done. t=5, index=0, shape=(145, 145, 6400)
save matrix t=5, index=1 done.
save full matrix done. t=5, index=1, shape=(145, 145, 6400)
save matrix t=5, index=2 done.
save full matrix done. t=5, index=2, shape=(145, 145, 6400)
--- 5 ---
```





```
feature extract t=5 done.
```

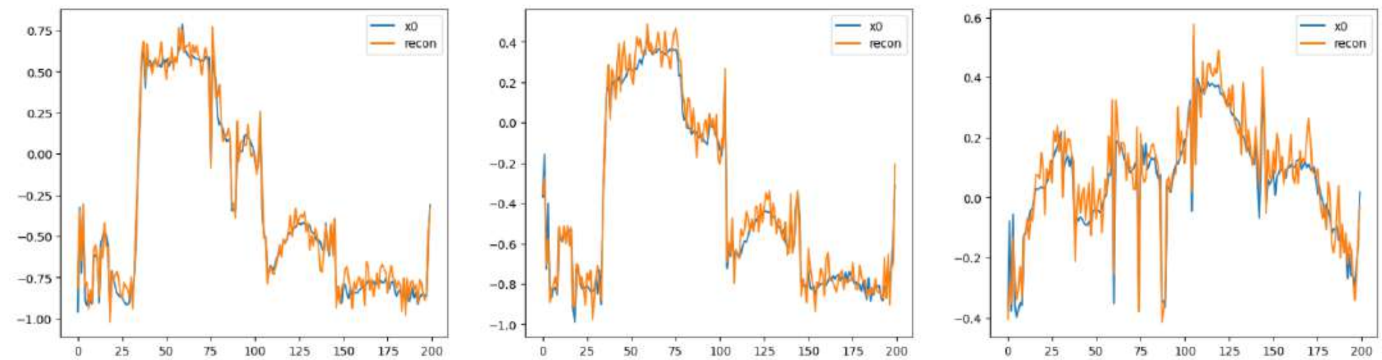
```
feature extract t=5 done.
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=10, index=0 done.
save full matrix done. t=10, index=0, shape=(145, 145, 6400)
save matrix t=10, index=1 done.
save full matrix done. t=10, index=1, shape=(145, 145, 6400)
save matrix t=10, index=2 done.
save full matrix done. t=10, index=2, shape=(145, 145, 6400)
--- 10 ---
```



International Journal of Information Technology & Computer Engineering

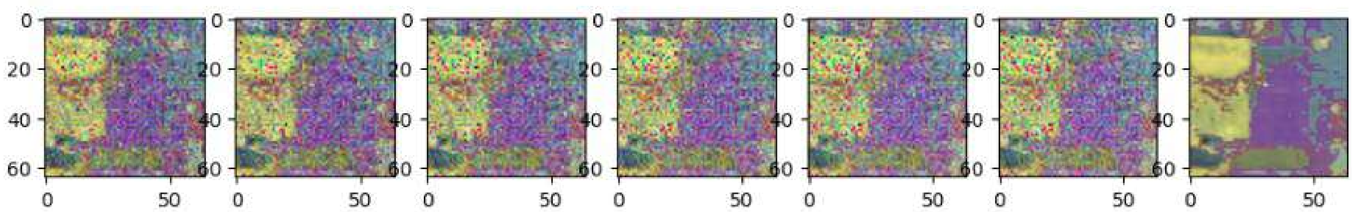
ISSN 2347–3657

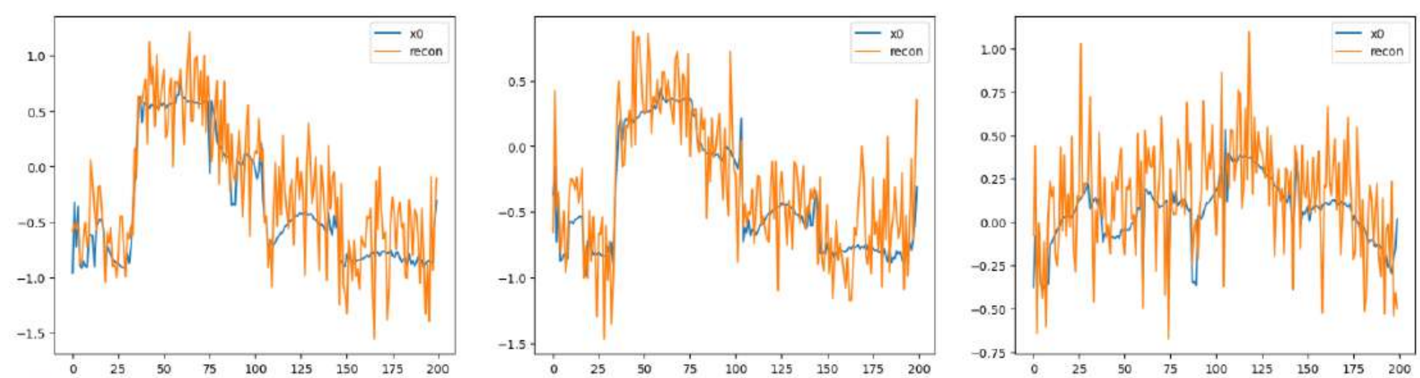
Volume 13, Issue 1, 2025



feature extract t=10 done.

```
feature extract t=10 done.
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=50, index=0 done.
save full matrix done. t=50, index=0, shape=(145, 145, 6400)
save matrix t=50, index=1 done.
save full matrix done. t=50, index=1, shape=(145, 145, 6400)
save matrix t=50, index=2 done.
save full matrix done. t=50, index=2, shape=(145, 145, 6400)
--- 50 ---
```

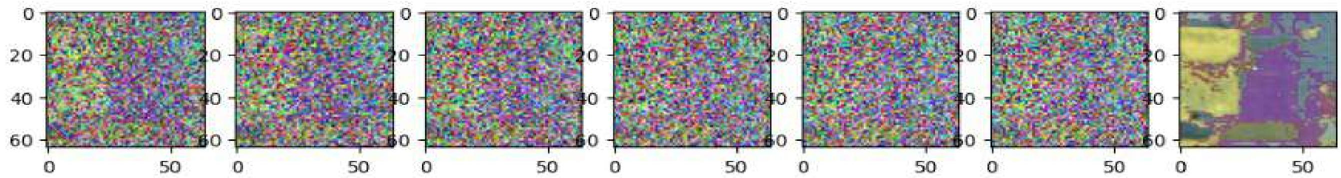


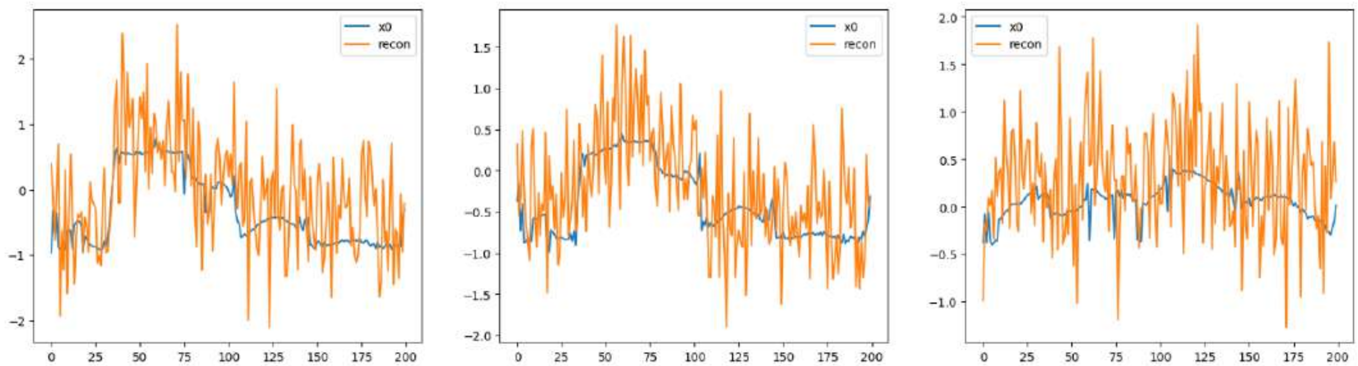


feature extract t=50 done.

ISSN 2347–3657
Volume 13, Issue 1, 2025

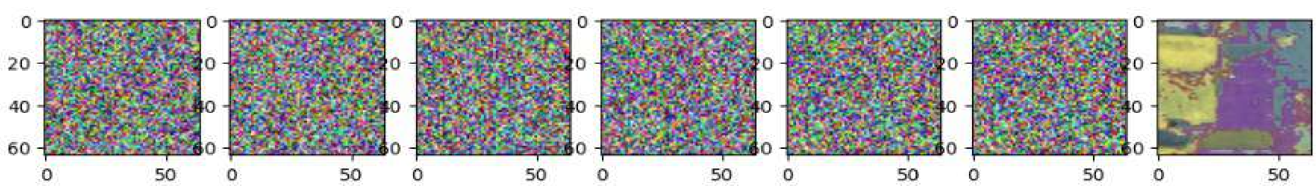
```
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=100, index=0 done.
save full matrix done. t=100, index=0, shape=(145, 145, 6400)
save matrix t=100, index=1 done.
save full matrix done. t=100, index=1, shape=(145, 145, 6400)
save matrix t=100, index=2 done.
save full matrix done. t=100, index=2, shape=(145, 145, 6400)
--- 100 ---
```

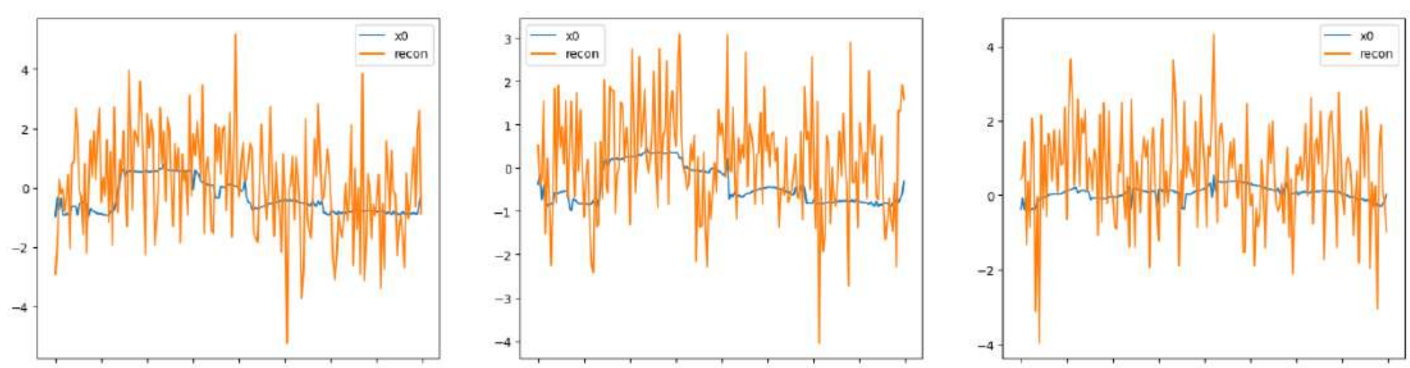





```
feature extract t=100 done.
```

```
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=200, index=0 done.
save full matrix done. t=200, index=0, shape=(145, 145, 6400)
save matrix t=200, index=1 done.
save full matrix done. t=200, index=1, shape=(145, 145, 6400)
save matrix t=200, index=2 done.
save full matrix done. t=200, index=2, shape=(145, 145, 6400)
--- 200 ---
```

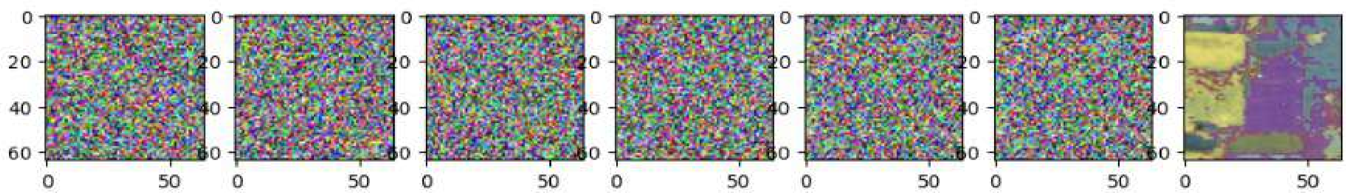


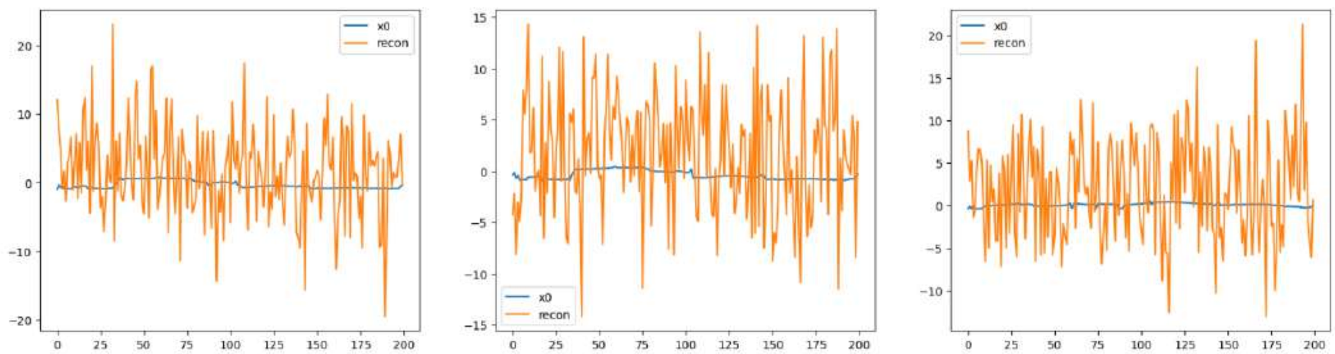



```
feature extract t=200 done.
```

ISSN 2347–3657
Volume 13, Issue 1, 2025

```
(9, 256, 25, 64, 64)
(9, 128, 50, 64, 64)
(9, 64, 100, 64, 64)
save matrix t=400, index=0 done.
save full matrix done. t=400, index=0, shape=(145, 145, 6400)
save matrix t=400, index=1 done.
save full matrix done. t=400, index=1, shape=(145, 145, 6400)
save matrix t=400, index=2 done.
save full matrix done. t=400, index=2, shape=(145, 145, 6400)
--- 400 ---
```






```
feature extract t=400 done.
done.
```

Performance Analysis

1. **Algorithm Comparison**:The proposed model was compared with other algorithms (e.g., CNN1D, CNN2D,SSRN, SSFTT) and showed improved performance, especially in balancing spatial and spectral data.Traditional models like CNN1D and CNN2D performed less effectively due to limited feature extraction, while advanced models like SSRN and SSFTT improved spatial-spectral classification.
2. **Classification Performance**:The model achieved better results on the IP and SA datasets in terms of OA, AA, and κ, showing balanced performance across different land-cover categories.On the PU dataset, the model performed slightly weaker in balancing classifications across subcategories but still maintained high accuracy overall.
3. **Sample Size Impact**:Increasing the sample size led to better accuracy across all models. The proposed model outperformed others even with smaller sample sizes.
4. **Qualitative Results**:Visual results showed that the proposed model produced classification maps with lowernoise and better alignment with the ground truth compared to other models.

Model Analysis

1. **Ablation Study**:Using diffusion features as input improved accuracy over using raw features, demonstratingthe model's effectiveness in capturing spectral-spatial relationships.
2. **Diffusion Model Analysis**:The diffusion model successfully reconstructed spectral curves from noisy inputs,showing that it embedded relevant spectral information.Experiments showed that smaller timestamps (less denoising steps) and the first layer index yielded optimal classification performance.
3. **Inference Time**:The proposed two-stage algorithm had longer inference times compared to simpler modelsbut remained comparable to complex CNN and Transformer-based algorithms.

VI CONCLUSION

In this study, we introduce a groundbreaking approach for hyperspectral imaging (HSI) analysis, leveraging a generative perspective to construct the spectral-spatial distribution of HSI data. This novel method, termed **SpectralDiff**, fundamentally redefines the way relationships between samples are established in hyperspectral datasets. By focusing on the spectral-spatial diffusion process, SpectralDiff adapts to the intrinsic structure of data without requiring prior knowledge of graph structures or neighborhood information. This adaptability allows it to construct inter-sample relationships dynamically, making it a robust tool for analyzing complex HSI data.

The cornerstone of SpectralDiff lies in its ability to capture the data distribution and contextual information embedded within HSI. Unlike traditional methods that rely heavily on predefined graph structures or static spatial relationships, SpectralDiff leverages the generative paradigm to enable a cross-sample perception mechanism. This mechanism facilitates the diffusion of information across spectral and spatial dimensions, effectively modeling the interdependence among samples and providing a more comprehensive representation of the HSI data manifold.

One of the key advantages of SpectralDiff is its generative approach, which captures the underlying spectral-spatial features more effectively than state-of-the-art techniques. By modeling the spectral-spatial diffusion process, SpectralDiff adapts to the inherent structure of data and uncovers nuanced relationships that static models often overlook. Experimental results validate the superiority of SpectralDiff, demonstrating that it consistently outperforms contemporary methods across various HSI classification benchmarks. The ability to integrate spectral and spatial information seamlessly leads to significant performance gains, making it a valuable tool for advancing HSI research and applications.

Looking ahead, the potential for diffusion models to address challenges such as out-of-distribution (OOD) generalization and detection within HSI is particularly promising. Diffusion models, by virtue of their generative framework, excel at capturing underlying data manifolds through iterative diffusion processes. This enables strong generalization capabilities to unseen examples outside the training distribution, a critical requirement for real-world HSI applications where training data may not fully encompass all possible scenarios.

Future research in this domain could explore how diffusion models like SpectralDiff can be extended to tackle OOD detection problems in HSI. Experimental findings already suggest that diffusion models exhibit strong detection performance for OOD samples, indicating their potential to transform anomaly detection and generalization tasks in HSI.

As diffusion-based methods continue to evolve, they are poised to make substantial contributions to the fields of OOD generalization and detection. With ongoing advancements, diffusion models are expected to unlock exciting opportunities in analyzing complex HSI data, enabling applications in areas such as environmental monitoring, agriculture, urban planning, and beyond.

In conclusion, SpectralDiff represents a significant leap forward in hyperspectral imaging analysis by leveraging a generative perspective to model spectral-spatial diffusion processes. Its ability to adaptively construct inter-sample relationships and capture contextual information positions it as a powerful tool for addressing contemporary challenges in HSI. Furthermore, the exploration of diffusion models for OOD generalization and detection holds immense potential, paving the way for robust, scalable solutions in the analysis of high-dimensional hyperspectral data.

VII. REFERENCES

1. S. Prasad and L. M. Bruce, "Limitations of Principal Components Analysis for Hyperspectral Target Recognition," in *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 625-629, Oct. 2022, doi: 10.1109/LGRS.2008.2001282
2. M. -D. Yang, K. -S. Huang, Y. F. Yang, L. -Y. Lu, Z. -Y. Feng and H. P. Tsai, "Hyperspectral Image Classification Using Fast and Adaptive Bidimensional Empirical Mode Decomposition With Minimum Noise Fraction," in *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1950-1954, Dec. 2023, doi: 10.1109/LGRS.2016.2618930
3. P. R. Marpu, M. Pedergnana, M. Dalla Mura, J. A. Benediktsson and L. Bruzzone, "Automatic Generation of Standard Deviation Attribute Profiles for Spectral–Spatial Classification of Remote Sensing Data," in *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 2, pp. 293-297, March 2023, doi: 10.1109/LGRS.2012.2203784.
4. L. Fang, D. Zhu, J. Yue, B. Zhang and M. He, "Geometric-Spectral Reconstruction Learning for Multi- Source Open-Set Classification With Hyperspectral and LiDAR Data," in *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 10, pp. 1892-1895, October 2022, doi: 10.1109/JAS.2022.105893.
5. Alex Nichol, Prafulla Dhariwal, “Improved Denoising Diffusion Probabilistic Models” in arxiv >cs> vol 2 Feb 18 2022
6. J. Yue, D. Zhu, L. Fang, P. Ghamisi and Y. Wang, "Adaptive Spatial Pyramid Constraint for Hyperspectral Image

Classification With Limited Training Samples," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022
7. D. Hong et al., “SpectralFormer: Rethinking hyperspectral image classification with transformers,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5518615
8. S. Wan, C. Gong, P. Zhong, S. Pan, G. Li and J. Yang, "Hyperspectral Image Classification With Context-Aware Dynamic Graph Convolutional Network," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 597-612, Jan. 2021
9. Z. Zhao, H. Wang and X. Yu, "Spectral–Spatial Graph Attention Network for Semisupervised Hyperspectral Image Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022, Art no. 5503905, doi: 10.1109/LGRS.2021.3059509.
10. Y. Zhai *et al.*, "A Modified Locality-Preserving Projection Approach for Hyperspectral Image Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 8, pp. 1059-1063, Aug. 2023,doi: 10.1109/LGRS.2016.2564993.
11. J. Yue, L. Fang, and M. He, “Spectral–spatial latent reconstructionfor open-set hyperspectral image classification,” IEEE Trans. Image Process., vol. 31, pp. 5227–5241, 2022.
12. Saharia et al., “Photorealistic text-to-image diffusion models with deep language understanding,” in Proc. NIPS, vol.35, pp. 36479–36494, 2022.
13. M. Zhang et al., “A survey on graph diffusion models: Generative AI in science for molecule, protein and material,” 2023, arXiv:2304.01565
14. Jing, G. Corso, J. Chang, R. Barzilay, and T. Jaakkola, “Torsional diffusion for molecular conformer generation,” 2022, arXiv:2206.01729
15. J. Yue, L. Fang, S. Xia, Y. Deng, and J. Ma, “Dif-fusion: Towards high color fidelity in infrared and visibleimage fusion with diffusion models,”2023, arXiv:2301.08072.
16. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, “RePaint: Inpainting using denoising diffusion probabilistic models,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2022, pp. 11451–11461.
17. M. Daniels, T. Maunu, and P. Hand, “Score-based generative neural networks for large-scale optimal transport,” in Proc. NIPS, vol. 34, pp. 12955–12965, 2021.
18. R. S. Zimmermann, L. Schott, Y. Song, B. A. Dunn, and D. A. Klindt,“Score-based generative classifiers,” 2021, arXiv:2110.00473.
19. B. Liu, A. Yu, X. Yu, R. Wang, K. Gao, and W. Guo, “Deep Multiview learning for hyperspectral image classification,” IEEE Trans. Geosci.Remote Sens., vol. 59, no. 9, pp. 7758–7772, Sep. 2021.
20. L. Sun, G. Zhao, Y. Zheng, and Z. Wu, “Spectral–spatial feature tokenization transformer for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 60,Art. no. 5522214,2022