Vol. 08, No. 3 (2026) 1623-1632, doi: 10.24874/PES08.03.017

Proceedings on Engineering Sciences

www.pesjournal.net

EFFICIENT SPATIAL–SPECTRAL REPRESENTATION LEARNING FOR THE CLASSIFICATION OF HYPERSPECTRAL IMAGES USING A 2D CNN FRAMEWORK

Abhishek Gupta
Pankaj Agarwal
Pramod Kumar Soni[1]

Received 11.06.2026
Revised 25.07.2026
Accepted 30.08.2026

Keywords:

Hyperspectral Image, CNN, Landcover, Deep Learning, Environment, Classification.

A B S T R A C T

Recent progress in remote sensing technologies has facilitated to capture the information in hundreds of spectral bands called hyperspectral imagery (HSI). HSI contains a huge amount of information and can facilitate precise discrimination of material and land cover types and can be used in a wide range of applications. Although information extraction from HSI imagery is a tedious task due to its high dimensions and intermixing of pixels in information in different spectral bands. To mitigate all these issues in this work, a computationally efficient CNN framework, ESSL, is designed that effectively extracts spatial–spectral representations from HSI using patch-based learning without relying on computationally expensive 3D convolutions and enables hierarchical receptive field growth, allowing the network to capture both fine-grained textures and broader spatial contexts relevant to land-cover classification. The experiments are performed on well-known publicly available datasets and performance is evaluated on different evaluation metrics and it has been observed that the proposed method has obtained an OA of 99.96 ± 0.0003, 98.70 ± 0.0036 on the Pavia University and Indian Pines datasets, respectively.

© 2026 Published by Faculty of Engineering

1. INTRODUCTION

Over the past few decades, remote sensing technologies have gone through tremendous changes and have made advancements in capturing minute information about the objects. This is due to the growing number of high-resolution satellite and airborne sensors. Hyperspectral imagery (HSI) has become one of the best sources of data for detailed Earth observation and material analysis among these changes (He et al., 2018). Hyperspectral sensors gather information in hundreds of narrow, adjacent spectral bands that usually cover the electromagnetic spectrum. This information can be utilized for the identification of types of classes of land cover, material or object and can be used to gather the minute information about the land cover. With the help of HSI imagery we can detect the small variations in land cover which is not possible in multispectral (MS) imagery like Landsat (Mishra et al., 2024) and Sentinel-2 (Soni et al., 2021) which capture the data in a few broad spectral bands. The MS images are ideal for general land

[1] Corresponding author: Pramod Kumar Soni
Email: Pramod.soni@jaipur.manipal.edu

cover analysis, but their low spectral resolution can cause mixed pixels and misclassification, especially in areas that are complex or have a lot of different types of land. On the other hand, HSI facilitates the identification of objects having low spectral variation between them, for example different types of urban material. (like asphalt, concrete, and rooftops), soil types, or mineral compositions (Lu et al., 2019). Due to rich information availability the HSI data is widely used in applications like agriculture (Wang et al., 2023), environmental monitoring (Rajabi et al., 2024), document forgery (Singh et al., 2025), ecological studies (Sun et al., 2021), geology and mineral exploration (Peyghambari & Zhang, 2021), dense and security applications like target detection, camouflage discrimination, material identification, and surveillance (Shimoni et al., 2019). The classification of HSI is a very important and rapidly evolving field of research in remote sensing and has shown better performance than MS imagery.

Although the HSI data has numerous advantages but information extraction from it is a very challenging and tedious task, as it suffers from issues like high dimensionality, as it has hundreds of spectral bands, which require more computational capability and low spectral variability between these bands which makes it hard to discriminate even with high spectral resolution. Furthermore, during the acquisition of HSI data, the presence of sensor noise, atmospheric absorption effects, and calibration errors make HSI less accurate, which further poses challenges during classification. The preparation of ground truth data from HSI is a time-consuming and expensive task. For the classification of HSI, different types of techniques have been introduced by the researchers in the literature, which mainly include statistical and machine learning-based methods like k-nearest neighbours (KNN) (Tu et al., 2018), support vector machines (SVM) (Okwuashi & Ndehedehe, 2020), random forests (Wang et al., 2019), and Gaussian mixture models (GMM) (Li et al., 2013). These techniques along with dimensional reduction approach such as PCA (Kang et al., 2017), ICA (Jayaprakash et al., 2020) etc. have been employed successfully for classification and shown promising results. But they are primarily based on handcrafted features and are unable to handle the complex spectral-spatial relationships of HSI. Recent developments in computational capability introduced deep learning (DL) techniques have significantly transformed HSI classification, especially convolutional neural networks (CNNs) (Luo et al., 2018) have shown promising results to automatically extract hierarchical features from raw hyperspectral data. DL models can detect correlations in spectral, spatial, and spectral-spatial data all at once. This makes classification more accurate and reliable than older methods. Furthermore, the advancement in DL like transfer learning (He et al., 2019), attention mechanisms (Pan et al., 2023), and hybrid spectral-spatial networks are some of the methods that have made things even better, especially when in the presence of limited labelled data. But the DL based methods are computationally expensive and require more computing power. Although ML and DL techniques have attempted well to handle the challenges, the classification of HSI efficiently are making progress to solve the issues of HSI. To handle all the challenges involved in HSI imagery ML and DL methods have attempted well but they still suffer from the limited availability of ground-truth samples leads to overfitting, especially in DL architectures. Many The DL models, like 3D CNNs and attention-based networks, are hard to use in environments with limited resources or real-time needs because they take up a lot of memory and processing power. To handle these issues in this, work an Efficient Patch-Based 2D Convolutional Neural Network, ESSL that uses efficient spatial–spectral representation Learning, is designed to classify the HSI. The major contribution of the proposed CNN is as follows: -

- In this work a patch-based 2D convolutional neural network that effectively capture spatial–spectral relationship from HSI and fuses features through standard 2D convolutions, preserving spectral discriminability while significantly reducing model complexity and computational cost.
- The introduction of stacked convolutional layers with increasing filter depths enables hierarchical receptive field growth, allowing the network to capture both fine-grained textures to discriminate against the low spectral variations.
- The proposed method ESSL has shown strong generalization capabilities in computationally efficient way making it suitable for wide range of HSI datasets.

The rest of the paper is structured as follows: in Section 2 related work for the HSI classification is discussed in detail which is followed by the description of dataset and methodology in Section 3. The experimental results are discussed in Section 4, the comparison of the performance of the proposed method is presented in Section 5 and finally the conclusion and future scope are presented in Section 6

2. RELATED WORK

Due to the importance of HSI in different domains various techniques have been designed for the classification of HSI imagery. The initial HSI classification techniques are based on statistical decision theory and probabilistic modelling to describe how different types of land cover behave in terms of their spectra and includes like MLC (Chang, 2019), Bayesian (Borges et al., 2011), Linear and Quadratic Discriminant Analysis (Imani & Ghassemian, 2015). The main advantage of these methods is their ease of use and computational efficiency; however, their performance is constrained by the assumptions made about the data and their inability to manage noise and low spectral variations. Machine learning based supervised

techniques such as SVM (Mercier & Lennon, 2003), RF(Lawrence et al., 2006), ANN (Du et al., 2008) are also widely used for classification and have shown better results as compared to statistical based methods. But these techniques use some dimensionality reduction methods like PCA (Mounika et al., 2021), ICA (Li et al., 2014) etc., along with them to make the model computationally efficient. A variant of RF, cascaded RF(CRF) is designed by using neighbourhood rough sets with a hierarchical random subspace to enhance accuracy and diversity (Zhang et al., 2018). From the detailed experiment it has been observed that CRF has shown competitive and reliable classification performance compared to existing RF and SVM techniques. Among all ML techniques SVM with RBF kernel has shown promising results in handling HSI data.

In recent year DL based techniques have gained popularity in discriminate HSI classes and ML based methods are also combined with these methods. An integrated framework of SVM and DL is designed (Okwuashi & Ndehedehe, 2020). This makes the framework computationally efficient and significantly improved the classification accuracy which confirms the robustness and generalization capability in handling complex HSI data. Furthermore, the effect of different kernels function for the initialization of weights is also observed and from the experiment it has been observed that polynomial and radial basis variants yield the highest accuracies. The results confirm the robustness, generalization capability of the method for complex HSI classification tasks. Different variants of CNN like 1D, 2D and 3D are employed for the classification of HSI data. For the identification of textile fiber from HSI data 1D CNN based on spectral features is employed and it has outperformed the traditional classifiers by a significant margin (Huang et al., 2022). Similarly, 1D CNN is also used for the identification of pesticides on Hami melon using HSI (Hu et al., 2023). A hybrid framework that uses 1D and 2D CNN for capturing higher level features and 3D CNN for spectral-spatial features is used to classify HSI data efficiently and accurately (Liu et al., 2023). Similarly, a hybrid CNN framework that use both 2D and 3D CNN is designed, and the effect of different dimensional reduction techniques is also evaluated on diverse HSI datasets (Fırat et al., 2022). It has been observed that the performance of reduction method depends heavily on the complexity of datasets. A Multi-Feature Fusion framework (MFFCG) is designed by integrating a 3D-CNN and graph-based learning modules (Bhatti et al., 2023). MFFCG has shown better performance even with very few training samples in time constrained environment. The performance of different DL based methods like 1-D, 2D, 3D- CNN, LSTM and vision transformers are evaluated on wide range of datasets, and the use case of these architectures have been discussed in detail (Viel et al., 2023).

In recent years more computationally efficient DL models based on transfer learning that use pretrained networks are widely used in this domain as they require fewer parameters and efficient ones without affecting the overall accuracy. A lightweight 3-D separable ResNet (3-D-SRNet) is designed using transfer learning (Jiang et al., 2019). The complexity of the model is reduced by replacing 3-D convolutions with spatial and spectral separable convolutions. In the first stage the model is trained on source HSI dataset with sufficient samples and then fine-tuned on a target dataset of different sensors. The framework has shown obtain better accuracy and has low computational complexity as compared to traditional 3-D model. Similarly, an integrated spectral–spatial feature extraction with meta-transfer learning is designed and the uses of dense connections and three-dimensional convolutional residual structures to fully exploit spectral and spatial information (Zhou et al., 2024). Experimental results show that the method significantly outperforms existing small-sample and state-of-the-art classifiers, confirming their robustness and suitability for HSI classification. Although researchers have attempted well to address the issues of HSI classification, the issue of higher dimensions and low spectral variability between bands still posing a challenge.

3. DATASET AND METHODOLOGY

In this section, the dataset and methodology adopted for the classification of HSI data is discussed.

3.1 Dataset

In this to assess the performance of the proposed method three different publicly available dataset namely Pavia University, Indian pines and Salinas are used. The Pavia University dataset has a spatial resolution of 1.3 m and size of $610 \times 340 \times 115$, whereas the Indian pines have resolution of 20 m and dimension of $145 \times 145 \times 224$. The Salinas dataset has dimension of $517 \times 217 \times 224$ and spatial resolution of 3.7 m. These datasets have different spatial resolutions, classes and are from different landscapes which ensures the testing of the generalization and robustness of the proposed method. All the datasets are divided into 70% for training, 20% for validation and 10% for testing. The details about the number of classes, number of training testing and validation samples are present in Table 1, 2 and 3.

Table 1. Pavia University dataset

Class ID	Training	Validation	Testing
CL1	4,642	1,326	663
CL2	13,054	3,730	1,865
CL3	1,469	420	210
CL4	2,145	613	306
CL5	942	269	134
CL6	3,520	1,006	503
CL7	931	266	133
CL8	2,577	736	369
CL9	663	189	95

Table 2. Description of Indian Pines dataset

Class ID	Training	Validation	Testing
CL1	32	9	5
CL2	1,000	286	142
CL3	581	166	83
CL4	166	47	24
CL5	338	97	48
CL6	511	146	73
CL7	20	6	2
CL8	335	96	47
CL9	14	4	2
CL10	680	194	98
CL11	1,719	491	245
CL12	415	119	59
CL13	144	41	20
CL14	886	253	126
CL15	270	77	39
CL16	65	19	9

Table 3. Description of Salinas dataset

Class ID	Training	Validation	Testing
CL1	1,406	402	201
CL2	2,608	745	373
CL3	1,383	395	198
CL4	976	279	139
CL5	1,875	536	267
CL6	2,771	792	396
CL7	2,505	716	358
CL8	7,890	2,254	1,127
CL9	4,342	1241	620
CL10	2,294	656	328
CL11	748	214	106
CL12	1,349	385	193
CL13	641	183	92
CL14	749	214	107
CL15	5,088	1,454	726
CL16	1,265	361	180

3.2 Methodology

The detailed flow chart of the proposed patch based 2D-CNN is presented in the Figure 1 which is divided into three stages the data obtained by eliminating noisy bands and spectral normalization, the dataset is transformed into patches for capturing the localized information which is followed by the training and validation of the network and finally the performance of the proposed method is evaluated on different evaluation metrics.

The first step of the ESSL method is the transformation of the raw HSI into a form suitable for CNN based classification. For this the HSI data is divided into fixed-size spatial–spectral patches cantered on labelled pixels which ensuring that both local spatial structure and spectral information are preserved by excluding background pixels. Each patch represents a localized neighbourhood that enhances class separability compared to using single-pixel spectra. Similarly, the labelled data converted to zero-based indexing to maintain compatibility with CNN the proposed CNN framework.

The HSI data cube can be represented as

$$\mathbf{X} \in \mathbb{R}^{H \times W \times B} \qquad (1)$$

where H and W denote the spatial dimensions and B denotes the number of spectral bands. For each labeled pixel at spatial location (i, j), a spatial–spectral patch $\mathbf{P}_{i,j} \in \mathbb{R}^{k \times k \times B}$ is extracted using a sliding window approach, where k denotes the patch size.

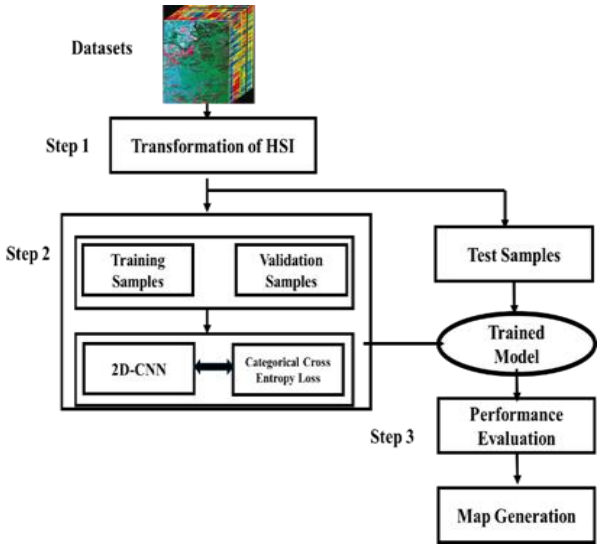


Figure 1. Flow chart of the proposed ESSL method

The architecture of the proposed ESSL consists of a series of 2D convolutional, Batch Normalization and ReLU layers, with a kernel size of 3×3 as shown in Figure 2. These layers have proven effective in extracting up fine-grained local spatial patterns while still being efficient in terms of computation. The kernel size of 3 x 3 makes the receptive field grow deeper, which facilitates the network to capture both local textures and larger spatial structures.

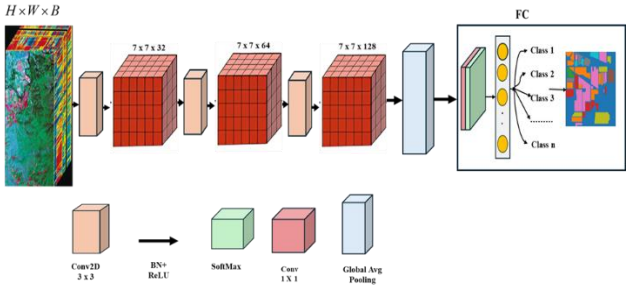


Figure 2. Architecture of the ESSL

The number of convolutional filters grows from layer to layer (32, 64, and 128), which helps the network to acquire a hierarchy of spatial features. Lower layers with fewer filters focus on basic spatial primitives, while deeper layers with more filters encode complex spatial configurations that are related to different types of land cover classes. Even though the convolution operators are two-dimensional, the spectral dimension is thought of as

channel depth. Every convolutional kernel covers all spectral bands and performs a weighted aggregation across the spectral dimension and perform spectral fusion This makes the computations easier and perform efficient discrimination between classes.

For the i^{th} convolutional layer, the feature extraction process is mathematically expressed as

$$\mathbf{F}^{(l)} = \sigma(\mathrm{BN}(\sum_{c=1}^{C_{l-1}} \mathbf{W}_c^{(l)} * \mathbf{F}_c^{(l-1)} + \mathbf{b}^{(l)})) \quad (2)$$

where $*$ denotes the 2D convolution operator, $\mathbf{W}^{(l)}$and $\mathbf{b}^{(l)}$are learnable parameters, $\mathrm{BN}(\cdot)$represents batch normalization, and $\sigma(\cdot)$is the ReLU activation function. By stacking multiple convolutional layers with increasing filter sizes, the network learns hierarchical spatial representations while implicitly fusing spectral information across channels.

To reduce dimensionality and enhance generalization, a global average pooling (GAP) layer is employed, which aggregates each feature map as;

$$g_m = \frac{1}{k^2}\sum_{i=1}^{k} \sum_{j=1}^{k} F_m\,(i,j) \quad (3)$$

The pooled feature vector is subsequently transformed by a fully connected layer with ReLU activation and dropout regularization to mitigate overfitting. The final classification is performed using a SoftMax layer, which estimates the posterior probability of each class $c \in \{1, \dots, C\}$ as

$$P(y = c \mid \mathbf{P}_{i,j}) = \frac{\exp(\mathbf{w}_c^{\top}\mathbf{z})}{\sum_{k=1}^{C} exp(\mathbf{w}_k^{\top}\mathbf{z})'} \quad (4)$$

where $\mathbf{z}$ denotes the learned feature representation and the network is optimized by categorical cross entropy loss function which can be defined as;

$$\mathcal{L}_{CCE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic} log(\hat{y}_{ic})$$

Where y_i and $\hat{y}$ are represents labelled and predicted samples from the N samples from C classes respectively.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this Section, the experimental results of the ESSL for the classification of HSI along with the experimental settings, evaluation metrics are presented and discussed.

4.1 Experimental Settings

The experiments are conducted on a GPU-equipped system featuring 32 GB RAM, and an NVIDIA RTX graphics card. For extracting patches from the HSI data a sliding window of 7 is used and all datasets are divided into training, testing, validation subsets in a 70:20:10 ratio to avoid overlapping between these sets. To prevent overfitting, a dropout rate of 0.2 is incorporated as a regularization technique, while a learning rate of 0.001 is used to ensure stable and efficient convergence. The model used convolutional filters of progressively increasing sizes (32, 64 and 128). A batch size of 64 was used, and the CCE loss was effectively selected for the estimation of the performance. The Adam optimizer, with an adaptive learning rate, is used over 100 epoch to fine-tune the model parameters, resulting in robust training and precise outcomes. To assess the performance of the proposed method the evaluation metrics such as overall accuracy (OAC), Kappa coefficients (κ) and average accuracy (AA) are used which can be computed as follows:

$$OAC = \frac{\sum_{j=1}^{C} P_{jj}}{P} \quad (4)$$

$$AA = \frac{1}{C}\sum_{j=1}^{C} \frac{n_{jj}}{N_i} \quad (5)$$

where P_{jj} is the number of correctly classified samples of class j

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (6)$$

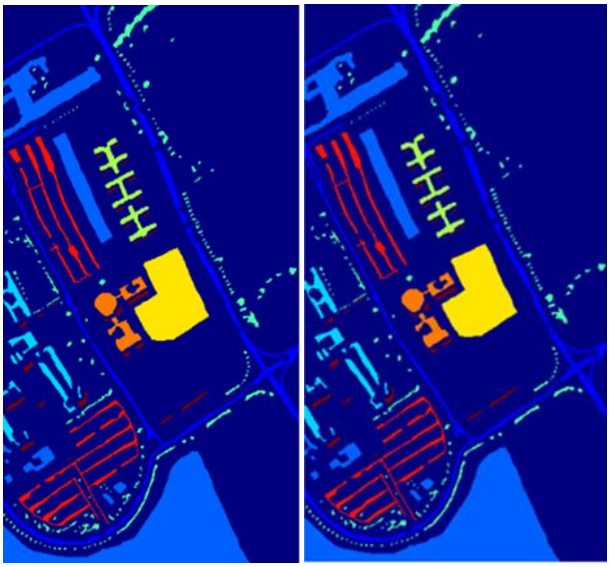
where P_o and P_e is the observed and expected accuracy.

4.2 Experimental Results and Discussion

The experimental results of the proposed method on Paiva University are visualized in Figure 3 and presented in Table 4, and it has been observed that ESSL has demonstrate the excellent performance in class-wise accuracies, with eight out of nine classes attaining 100% accuracy, indicating strong discriminative capability in extracting spectral–spatial features. A slight reduction is observed for CL4, likely due to spectral similarity in complex urban regions; however, the low standard deviation confirms stable performance. An OAC of 99.96 ± 0.0003 shows that most of the pixels are classified correctly, similarly, on AA of 99.93 ± 0.0005 indicates uniform performance across all classes without bias toward dominant categories. Furthermore, κ = 0.9994 ± 0.0004 signifies almost perfect agreement between predicted labels and ground truth.

Table 4. Performance of the proposed method on Pavia University Dataset

Class_ID	Accuracy
CL1	100 ± 0.0000
CL2	100 ± 0.0000
CL3	100 ± 0.0000
CL4	99.4 ± 0.0049
CL5	100 ± 0.0000
CL6	100 ± 0.0000
CL7	100 ± 0.0000
CL8	100 ± 0.0000
CL9	100 ± 0.0000
OAC	99.96 ± 0.0003
AA	99.93 ± 0.0005
κ	0.9994 ± 0.0004



(a) (b)

Figure 3. Experimental results on Pavia University dataset (a) Ground Truth (b) classified Map

The classification map of the proposed ESSL on the Indian Pines dataset is shown in Figure 4 and its performance is on different metrics is presented in Table 5. The effect of fusion of spectral and spatial features extracted by the ESSL is clearly seen as it has shown strong overall performance of 98.70 ± 0.0036, an AA of 97.38 ± 0.0274, and a high κ of 0.9852 ± 0.0041, indicating excellent agreement between predicted labels and ground truth. While in terms of class wise accuracy it has obtained 100% in most of the classes, highlighting the effectiveness of patch-based spectral–spatial feature learning. High accuracies are also observed for most remaining classes, such as CL10, CL11, and CL14, exceeding 98%, which confirms the robustness of the model. However, comparatively lower accuracies for CL4 of 91.25 ± 0.0599 and CL9 of 82.50 ± 0.3800 suggest challenges arising from spectral similarity, limited training samples, and mixed pixels, which are well-known issues in the Indian Pines dataset.

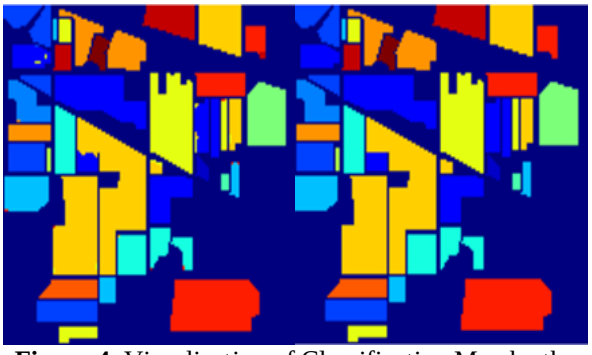


Figure 4. Visualization of Classification Map by the proposed method on Indian pines dataset (a) Ground Truth (b) Classified Map

To test the robustness and generalizability of the proposed framework a more complex Salinas dataset is used and it has been observed that the methods has shown an exceptionally well both visually as presented in Figure 5 and quantitatively in Table 6. In most of the classes it has achieve perfect classification accuracy of 100% demonstrating highly stable and consistent predictions. Slight performance variations are observed only for CL8 of 99.82 ± 0.0015 and CL15 of 99.71 ± 0.0015, which may be attributed to spectral similarity among vegetation types and boundary-level mixing effects inherent in the Salinas scene. The OAC of 99.92 ± 0.0004 and AA of 99.97 ± 0.0002 confirms that nearly all pixels are correctly classified class wise also. Furthermore, the κ of 0.9991 ± 0.0005 signifies almost perfect agreement with the ground truth. Overall, these results on diverse landscape of varying resolutions validate the effectiveness and reliability of the proposed approach for high-resolution agricultural hyperspectral image classification.

Table 5. Performance of the proposed method on Indian Pines Dataset

Class	Class Wise accuracy
CL1	100 ± 0.0000
CL2	96.55 ± 0.0150
CL3	100 ± 0.0000
CL4	91.25 ± 0.0599
CL5	95.65 ± 0.0320
CL6	100.00 ± 0.0000
CL7	95 ± 0.2179
CL8	100 ± 0.0000
CL9	82.50 ± 0.3800
CL10	98.91 ± 0.0106
CL11	99.14 ± 0.0061
CL12	100 ± 0.0000
CL13	100 ± 0.0000
CL14	99.09 ± 0.0090
CL15	100 ± 0.0000
CL16	100 ± 0.0000
OAC	98.71 ± 0.0036
AA	97.38 ± 0.0274
κ	0.9853 ± 0.0041

Table 6. Performance of the proposed method on Salinas dataset

Class	Class Wise accuracy
CL1	100 ± 0.0000
CL2	100 ± 0.0000
CL3	100 ± 0.0000
CL4	100 ± 0.0000
CL5	100 ± 0.0000
CL6	100 ± 0.0000
CL7	100 ± 0.0000
CL8	99.82 ± 0.0015
CL9	100 ± 0.0000
CL10	100 ± 0.0000
CL11	100 ± 0.0000
CL12	100 ± 0.0000
CL13	100 ± 0.0000
CL14	100 ± 0.0000
CL15	99.71 ± 0.0015
CL16	100 ± 0.0000
OAC	99.92 ± 0.0004
AA	99.97 ± 0.0002
κ	0.9991 ± 0.0005

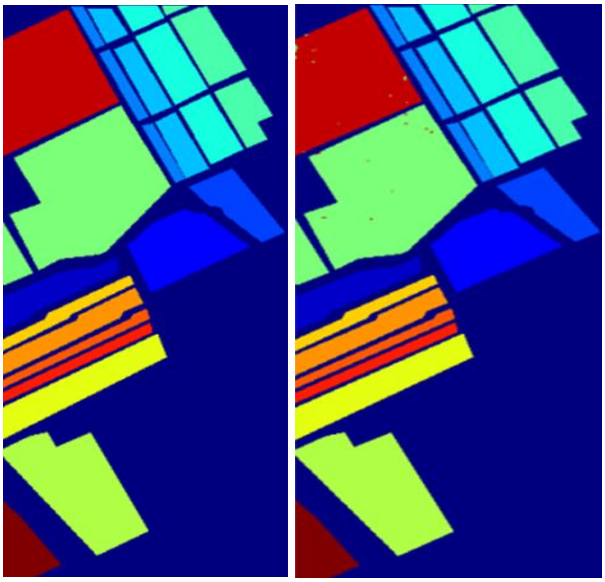


Figure 5. Visualization of Classification Map by proposed method on Salinas dataset (a) Ground Truth (b) Classified Map

5. COMPARTIVE ANALYSIS

To test the efficacy of the ESSL on Pavia and Indian pines dataset is matched with existing state of art techniques like 3D-ResNet (Jiang et al., 2019) that used 3D convolutional operations, IGroupSS-Mamba (He et al., 2024) which capture multi global and multiscale information, ENL (Shen et al., 2020) based on non-local information and cross attention based CASST method (Peng et al., 2022). On Pavia University dataset the detailed comparison of proposed ESSL is presented in Table 7 and it has been observed that it has achieved the best performance on as compared to counterpart methods with an OAC of 99.96, AA of 99.93, and κ of 0.9994, indicating near-perfect classification. The IGroupSS-Mamba has also shown competitive results and follows the ESSL very closely with an OAC of 99.75. whereas ENL shows moderate performance with an OAC of 99.08 and a κ of 0.9878 and 3DResNet attains and OAC of 98.88, demonstrating effective learning but limited overall consistency and CASST records the lowest OAC of 98.53 among the compared methods. Overall, ESSL has significantly outperformed conventional deep architectures on this dataset.

Table 7. Comparative Evaluation of ESSL on Pavia University Dataset

Method	OAC	AA	κ
3DResNet	98.88	99.30	0.9984
CASST	98.53	99.01	0.9937
IGroupSS-Mamba	99.75	99.46	0.9946
ENL	99.08	98.20	0.9878
ESSL	99.96	99.93	0.9994

The classification of the Indian pines dataset is challenging as it has overlapping pixels and uneven class distribution. Compared to computationally intensive architectures such as 3DResNet, ESSL delivers comparable accuracy of 98.71 as presented in Table 8 with improved efficiency, making it more suitable for real-world HSI. With an AA of 97.38, ESSL shows consistent performance across different land-cover classes and outperformed CASST, ENL and remain competitive as compared to other methods like IGroupSS-Mamba, 3D-ResNet and ENL. Whereas the κ value of 0.9853 by ESSL further indicates a high level of agreement between predicted labels and reference data, validating the robustness of the proposed method. While ENL achieves a marginally higher κ,.

Table 8. Comparative Analysis of ESSL on Indian Pines Dataset

Method	OAC	AA	κ
3DResNet	98.71	99.37	0.9850
CASST	97.71	94.98	0.9738
IGroupSS-Mamba	98.71	98.33	0.9853
ENL	98.85	97.35	0.9869
ESSL	98.71	97.38	0.9853

Similarly, the performance of the proposed ESSL on Salinas valley dataset which predominantly contains vegetative regions and suffers from the issue of fine grained class boundaries, is matched with the current DL based methods like HSI-Mamaba (Liao & Wang, 2025), Transfer learning-based CNN (TL-CNN) (Giri et al., 2024), hierarchical feature based CNN (H-CNN)(Feng et al., 2023) and traditional SVM. In Salinas dataset the detailed comparative analysis of the proposed method is presented in Table 9, and it has been observed that ESSL consistently outperforms all competing methods across OA, AA, and κ metrics by obtaining the best values of OAC of 99.92, AA of 99.97 and k of 0.9991. Although, the TL-CNN has shown competitive performance as compared to other methods in handling the challenges of complex heterogenous landcover. The superior performance of the ESSL demonstrate the robustness of the proposed framework in handling class imbalance, low spectral variation and in extracting discriminative spatial–spectral features and underline its novel contribution toward achieving highly accurate and reliable HSI classification.

Table 9. Comparative Analysis of ESSL on Salinas Dataset

Method	OAC	AA	κ
TL-CNN	99.12	99.40	0.9901
H-CNN	91.71	88.45	0.9076
SVM	87.54	92.84	0.8689
HSI-Mamba	96.40	98.44	0.9599
ESSL	99.92	99.97	0.9991

In terms of computational efficiency, the proposed ESSL framework has 170608 trainable parameters with 14.83 MFLOPs. In terms of memory the ESSL require only 666 KB of memory making it suitable for efficient HSI classification with a very low computational complexity.

6. CONCLUSION

In this work, a computationally efficient ESSL CNN framework is proposed for the classification of HSI data. The proposed method uses a patch-based 2D CNN architecture with hierarchical receptive field expansion, enabling multi-level feature abstraction while avoiding the high computational cost and memory overhead associated with 3D convolutional operations. The performance of the ESSL is tested on three widely used benchmark datasets and evaluation on different evaluation metrics. The experimental results reveal that ESSL consistently achieves superior or highly competitive results when compared with state-of-the-art DL framework such as 3D-ResNet, CASST, ENL, IGroupSS-Mamba, TL-CNN, and HIS-Mamba. By efficiently integrating spatial context and spectral information through hierarchical feature learning, ESSL demonstrates strong generalization and robustness across diverse HSI scenes. Overall, the proposed method provides a scalable, accurate, and computationally viable solution for the classification of HSI and establishes a promising foundation for future research in large-scale remote sensing analysis, precision agriculture, and land-cover mapping.

References:

Bhatti, U. A., Huang, M., Neira-Molina, H., Marjan, S., Baryalai, M., Tang, H., … Bazai, S. U. (2023). MFFCG – Multi feature fusion for hyperspectral image classification using graph attention network. *Expert Systems with Applications*, *229*, 120496. https://doi.org/10.1016/J.ESWA.2023.120496

Borges, J. S., Bioucas-Dias, J. M., & Marcal, A. R. S. (2011). Bayesian Hyperspectral Image Segmentation With Discriminative Class Learning. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(6), 2151–2164. https://doi.org/10.1109/TGRS.2010.2097268

Chang, C.-I. (2019). Statistical Detection Theory Approach to Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(4), 2057–2074. https://doi.org/10.1109/TGRS.2018.2870980

Du, P., Tan, K., Zhang, W., & Yan, Z. (2008). ANN classification of OMIS hyperspectral remotely sensed imagery: Experiments and analysis. *2008 Congress on Image and Signal Processing*, *4*, 692–696. https://doi.org/10.1109/CISP.2008.656

Feng, Z., Liu, X., Yang, S., Zhang, K., & Jiao, L. (2023). Hierarchical feature fusion and selection for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, *20*, 1–5. https://doi.org/10.1109/LGRS.2023.3236672

Fırat, H., Asker, M. E., & Hanbay, D. (2022). Classification of hyperspectral remote sensing images using different dimension reduction methods with 3D/2D CNN. *Remote Sensing Applications: Society and Environment*, *25*, 100694. https://doi.org/10.1016/J.RSASE.2022.100694

Giri, R. N., Janghel, R. R., Pandey, S. K., Govil, H., & Sinha, A. (2024). Enhanced hyperspectral image classification through pretrained CNN model for robust spatial feature extraction. *Journal of Optics*, *53*(3), 2287–2300. https://doi.org/10.1007/s12596-023-01473-7

He, L., Li, J., Liu, C., & Li, S. (2018). Recent Advances on Spectral–Spatial Hyperspectral Image Classification: An Overview and New Guidelines. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(3), 1579–1597. https://doi.org/10.1109/TGRS.2017.2765364

He, X., Chen, Y., & Ghamisi, P. (2019). Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(5), 3246–3263. https://doi.org/10.1109/TGRS.2019.2951445

He, Y., Tu, B., Jiang, P., Liu, B., Li, J., & Plaza, A. (2024). IGroupSS-Mamba: Interval group spatial-spectral mamba for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2024.3502055

Hu, Y., Ma, B., Wang, H., Zhang, Y., Li, Y., & Yu, G. (2023). Detecting different pesticide residues on Hami melon surface using hyperspectral imaging combined with 1D-CNN and information fusion. *Frontiers in Plant Science*, *14*, 1105601. https://doi.org/10.3389/fpls.2023.1105601

Huang, J., He, H., Lv, R., Zhang, G., Zhou, Z., & Wang, X. (2022). Non-destructive detection and classification of textile fibres based on hyperspectral imaging and 1D-CNN. *Analytica Chimica Acta*, *1224*, 340238. https://doi.org/10.1016/J.ACA.2022.340238

Imani, M., & Ghassemian, H. (2015). Two dimensional linear discriminant analyses for hyperspectral data. *Photogrammetric Engineering & Remote Sensing*, *81*(10), 777–786. https://doi.org/10.14358/PERS.81.10.777

Jayaprakash, C., Damodaran, B. B., Viswanathan, S., & Soman, K. P. (2020). Randomized independent component analysis and linear discriminant analysis dimensionality reduction methods for hyperspectral image classification. *Journal of Applied Remote Sensing*, *14*(3), 36507. https://doi.org/10.1117/1.JRS.14.036507

Jiang, Y., Li, Y., & Zhang, H. (2019). Hyperspectral image classification based on 3-D separable ResNet and transfer learning. *IEEE Geoscience and Remote Sensing Letters*, *16*(12), 1949–1953. https://doi.org/10.1109/LGRS.2019.2913011

Kang, X., Xiang, X., Li, S., & Benediktsson, J. A. (2017). PCA-based edge-preserving features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 7140–7151. https://doi.org/10.1109/TGRS.2017.2743102

Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment*, *100*(3), 356–362. https://doi.org/10.1016/J.RSE.2005.10.014

Li, C., Yin, J., & Zhao, J. (2014). Using improved ICA method for hyperspectral data classification. *Arabian Journal for Science and Engineering*, *39*(1), 181–189. https://doi.org/10.1007/s13369-013-0831-7

Li, W., Prasad, S., & Fowler, J. E. (2013). Hyperspectral image classification using Gaussian mixture models and Markov random fields. *IEEE Geoscience and Remote Sensing Letters*, *11*(1), 153–157. https://doi.org/10.1109/LGRS.2013.2250905

Liao, J., & Wang, L. (2025). HyperspectralMamba: a novel state space model architecture for hyperspectral image classification. *Remote Sensing*, *17*(15), 2577. https://doi.org/10.3390/rs17152577

Liu, J., Wang, T., Skidmore, A., Sun, Y., Jia, P., & Zhang, K. (2023). Integrated 1D, 2D, and 3D CNNs enable robust and efficient land cover classification from hyperspectral imagery. *Remote Sensing*, *15*(19), 4797. https://doi.org/10.3390/rs15194797

Lu, B., He, Y., & Dao, P. D. (2019). Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(6), 1784–1797. https://doi.org/10.1109/JSTARS.2019.2910558

Luo, Y., Zou, J., Yao, C., Zhao, X., Li, T., & Bai, G. (2018). HSI-CNN: A novel convolution neural network for hyperspectral image. *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, 464–469. https://doi.org/10.1109/ICALIP.2018.8455251

Mercier, G., & Lennon, M. (2003). Support vector machines for hyperspectral image classification with spectral-based kernels. *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, *1*, 288–290. https://doi.org/10.1109/IGARSS.2003.1293752

Mishra, V. K., Verma, K. K., Pant, T., Upadhyay, G. M., Singh, P. S., & Soni, P. K. (2024). Analysis of Urbanization Impact on Land Surface Temperature Variability by Using Landsat Imagery. *SN Computer Science*, *5*(7), 863. https://doi.org/10.1007/s42979-024-03226-0

Mounika, K., Aravind, K., Yamini, M., Navyasri, P., Dash, S., & Suryanarayana, V. (2021). Hyperspectral image classification using SVM with PCA. *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, 470–475. https://doi.org/10.1109/ISPCC53510.2021.9609461

Okwuashi, O., & Ndehedehe, C. E. (2020). Deep support vector machine for hyperspectral image classification. *Pattern Recognition*, *103*, 107298. https://doi.org/10.1016/j.patcog.2020.107298

Pan, H., Zhao, X., Ge, H., Liu, M., & Shi, C. (2023). Hyperspectral image classification based on multiscale hybrid networks and attention mechanisms. *Remote Sensing*, *15*(11), 2720. https://doi.org/10.3390/rs15112720

Peng, Y., Zhang, Y., Tu, B., Li, Q., & Li, W. (2022). Spatial–spectral transformer with cross-attention for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–15. https://doi.org/10.1109/TGRS.2022.3203476

Peyghambari, S., & Zhang, Y. (2021). Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review. *Journal of Applied Remote Sensing*, *15*(3), 31501. https://doi.org/10.1117/1.JRS.15.031501

Rajabi, R., Zehtabian, A., Singh, K. D., Tabatabaeenejad, A., Ghamisi, P., & Homayouni, S. (2024). Editorial: Hyperspectral imaging in environmental monitoring and analysis. *Frontiers in Environmental Science, Volume 11-2023*. https://doi.org/10.3389/fenvs.2023.1353447

Shen, Y., Zhu, S., Chen, C., Du, Q., Xiao, L., Chen, J., & Pan, D. (2020). Efficient deep learning of nonlocal features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(7), 6029–6043. https://doi.org/10.1109/TGRS.2020.3014286

Shimoni, M., Haelterman, R., & Perneel, C. (2019). Hypersectral imaging for military and security applications: Combining myriad processing and sensing techniques. *IEEE Geoscience and Remote Sensing Magazine*, *7*(2), 101–117. https://doi.org/10.1109/MGRS.2019.2902525

Singh, P. S., Karthikeyan, S., Upadhayay, G. M., Sadhukhan, S., Soni, P. K., & Malche, T. (2025). Unsupervised deep autoencoder-based reconstruction for ink mismatch detection in hyperspectral document images. *Discover Computing*, *28*(1), 333. https://doi.org/10.1007/s10791-025-09837-2

Soni, P. K., Rajpal, N., Mehta, R., & Mishra, V. K. (2021). Urban land cover and land use classification using multispectral sentinal-2 imagery. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-021-10991-0

Sun, G., Jiao, Z., Zhang, A., Li, F., Fu, H., & Li, Z. (2021). Hyperspectral image-based vegetation index (HSVI): A new vegetation index for urban ecological research. *International Journal of Applied Earth Observation and Geoinformation*, *103*, 102529. https://doi.org/10.1016/j.jag.2021.102529

Tu, B., Wang, J., Kang, X., Zhang, G., Ou, X., & Guo, L. (2018). KNN-based representation of superpixels for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(11), 4032–4047. https://doi.org/10.1109/JSTARS.2018.2872969

Viel, F., Maciel, R. C., Seman, L. O., Zeferino, C. A., Bezerra, E. A., & Leithardt, V. R. Q. (2023). Hyperspectral image classification: An analysis employing CNN, LSTM, transformer, and attention mechanism. *IEEE Access*, *11*, 24835–24850. https://doi.org/10.1109/ACCESS.2023.3255164

Wang, A., Wang, Y., & Chen, Y. (2019). Hyperspectral image classification based on convolutional neural network and random forest. *Remote Sensing Letters*, *10*(11), 1086–1094. https://doi.org/10.1080/2150704X.2019.1649736

Wang, B., Sun, J., Xia, L., Liu, J., Wang, Z., Li, P., … Sun, X. (2023). The applications of hyperspectral imaging technology for agricultural products quality analysis: a review. *Food Reviews International*, *39*(2), 1043–1062. https://doi.org/10.1080/87559129.2021.1929297

Zhang, Y., Cao, G., Li, X., & Wang, B. (2018). Cascaded Random Forest for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(4), 1082–1094. https://doi.org/10.1109/JSTARS.2018.2809781

Zhou, H., Wang, X., Xia, K., Ma, Y., & Yuan, G. (2024). Transfer learning-based hyperspectral image classification using residual dense connection networks. *Sensors*, *24*(9), 2664. https://doi.org/10.3390/s24092664

Abhishek Kumar Gupta
School of engineering and technology
K.R.Mangalam University,
Gurugram,
Haryana
India
abhishek130685@gmail.com
ORCID 0009-0000-9701-246X

Pankaj Agarwal
School of engineering and technology
K.R.Mangalam University,
Gurugram,
Haryana
India
pankaj7877@gmail.com
ORCID 0000-0002-1027-2419

Pramod Kumar Soni
Department of Computer Applications,
Manipal University Jaipur,
Jaipur,
Rajasthan
India
Pramod.soni@jaipur.manipal.edu
ORCID 0000-0003-0632-5044