

A Novel Approach of Hyperspectral Imaging Classification using Hybrid ConvNet

Soumyashree M Panchal¹

Assistant Professor

Department of Information Science & Engineering
HKBK College of Engineering, Bangalore, India

Shivaputra²

Assistant Professor

Electronics and Communication Engineering
Dr. AIT, Bangalore India

Abstract—In recent years, remote sensing applications have been booming, and with this hyperspectral imaging (HSI) has been used in many real-life applications. However, the classification of HSI is a significant problem due to the complex features of the captured hyperspectral scene. Moreover, the HSI is often inherently nonlinear and has very high-dimensional data. Recent years have seen a rise in deep learning applications for addressing nonlinear problems. However, deep learning tends to overfit when sparse or less training data is available. In this paper, the proposed work focuses on addressing the trade-off problem between classification performance and less training samples for classifying hyperspectral image data in a single training process. Thus, the study presents a hybrid multilayer learning system based on the joint approach of 2D and 3D convolutional kernels. The main reason is to utilize the spectral-spatial and spatial correlations in the learning process to achieve improved generalization of features in the training process for better HSI classification. The study outcome exhibits higher precision, recall rate, and F1-score performance. The overall accuracy is 99.9%, with a better convergence rate. The results prove that the proposed model is effective for HSI classification even with fewer training data samples.

Keywords—Hyperspectral image; convolution neural network; classification; spatial feature; spectral feature

I. INTRODUCTION

Hyperspectral imaging (HSI) consists of hundreds of narrow bands with rich spectral and spatial data in remote sensory applications. These spatial and spectral characteristics of hyperspectral remote sensing images can provide useful information for detecting and classifying objects [1]. Since the early 1990s, HSI has been widely applied in a variety of real-world contexts, including precision agriculture [2] and land management to healthcare and military target identification [3]. HSIs are high-dimensional data with a high correlation between adjacent spectral bands, making it more complex in time and space context and leading to the Hughes phenomenon [4]. Thus, reducing the amount of redundancy in HSI processing is a crucial concern. In the literature, most of the existing studies have focused on exploring the role of spectral features of HSI in classification. Indeed, HSIs also typically have spatial features where the adjacent pixels tend to be part of the same class. Because of these characteristics, two significant challenges encountered related to HSI processing viz. i) the high spatial inconsistency concerning spectral features and ii) constraints samples and the high dimensionality data. Several factors are usually responsible for the challenges

mentioned above, such as changes in lighting conditions, environment, surroundings impact, and temporal circumstances [5]. These challenges often result in problems for most traditional methods and reduce their classification accuracy [6]. To overcome these problems, the analysis of spatial features has been reported to be valuable in improving object identification and classification performance. According to recent literature, classification of the HSI object based on spectral-spatial information, incorporating spatial attributes into pixel-wise cataloging processes [7] using mathematical morphological operations, Artificial Neural Networks (ANN), and machine learning methods such as support vector machines (SVMs), Logistic Regression, and many others [8]. Moreover, existing researches have also tried to address the problem associated with feature engineering, using principal component analysis (PCA) and linear discriminant analysis (LDA) [9][10]. Despite this, the previous works heavily rely on shallow and manual feature descriptors and are usually created for specific purposes, limiting their effectiveness in real-time situations [11].

Several deep learning models, including ConvNet convolution neural networks (CNNs), recurrent neural networks (RNNs), and deep autoencoders, have recently made significant breakthroughs in computer vision tasks, such as image classification [12] object recognition [13], and language processing [14]. These applications have inspired HSI analysis, and deep learning has proved to be highly effective in detecting and classifying objects. In contrast to traditional manual methods, deep learning can extract valuable insights from input data samples through a sequence of hierarchical layers [15]. In the literature, deep learning-based research works for HSI classification have few flaws and rely on substantial labeled samples [16]. However, feature generalization can be fully automated, making deep learning more appropriate for various situations. Furthermore, the previous deep learning models adopt a very complex structure, lack the ability for the input data to be spatially invariant, and are prone to overfitting problems due to the high dimensionality and small sample size of HSI. Therefore, there is a need to develop an effective model that can perform precise feature analysis to classify HSI objects with data samples without posing an overfitting issue. Hence, the factor of motivation is to understand the fact the wider scope of usage of HSI could be more leveraged if these impending problems are addressed where the existing solution encounters problems associated with computational complexities and non-inclusion of various constraints. This

results in evolution of proposed solution towards addressing this point for classifying HSI objects. In this paper, the proposed work addresses the HSI classification issue by using a new hybrid deep learning mechanism to identify the object category of each pixel with a limited number of data samples. Specifically, the study emphasizes feature learning aspects of the proposed hybrid learning model, which uses both 2D ConvNet and 3D ConvNet to process hyperspectral cube structure. The proposed hybrid ConvNet assembles 2D and 3D convolution layers as complementary operations to attain rich contextual information in the learning process concerning both spatial feature and mixed feature (i.e., spatial-spectral) to achieve maximum possible accuracy. The significant contribution of this paper is highlighted as follows:

- The proposed research work enhances the function of deep learning techniques with stochastic data treatment and feature selection process for the optimal performance in HSI classification.
- Unlike previous schemes, the proposed work emphasizes balancing overall accuracy and computational efficiency.
- A hybridization is introduced in the learning model, providing less dependency on the training sample and quick convergence.
- The design of the proposed model for HSI classification is adaptive to different HSI data, thus meeting the requirement of the real-time deployment scenario.

The remaining sections of this paper are organized as follows: Section II presents a brief review of the existing works done in the context of HSI classification; Section III highlights the significant issue and the research gap explored based on the review analysis; Section IV discusses the proposed system design and methodology adopted; Section V focuses on the detailing the implementation procedures for processing hyperspectral cube and classifying the objects from the HSI; in Section VI results and discussion is carried out for the validation of the proposed work and finally Section VII concludes the real contribution of this paper.

II. REVIEW OF LITERATURE

This section briefly reviews the existing solutions in this context and highlights the significant problem explored based on the review analysis. Although the HSI classification has been extensively studied in the past recent years. The existence of noise seriously affects the classification accuracy of the model. The work carried out by Lu et al. [17] suggested a different technique named penalized linear discriminant (PLD) with principal components to address the issue of noise in HSI data. PLD analysis is implemented to determine the optimal covariance matrix of noisy data, and then it is eliminated using a principal component transformation scheme. The study outcome shows that this method removed noise significantly without losing spectral fidelity. Hou et al. [18] presented a supervised dimensionality reduction scheme based on the kernel-based possibilistic clustering mechanism in the same line of work. The fundamental principle of this kernel-based possibilistic clustering scheme is the construction of the

weights to generalize effective transformation directions for executing HSI classification. However, deciding a suitable kernel is quite tricky, and similar performance may not be achieved on the other HSI dataset. In this direction, Hang et al. [19] reported the suitability of applying local graph discriminant (LGD) embedding. However, this lacks consideration of the spatial features of the HSI data. The authors have developed a regularization scheme that considers the spatial information in LGD embedding, thereby boosting classification performance. The study has also shown that implementing this method can improve the performance of the kernel-based methods. Jia et al. [20] emphasized addressing the problem of labeling data samples in the classification task. This study suggests an unsupervised model based on the combined approach of Gabor filters and LAD to extract the most revealing and refined features for classification. However, LDA is quite popular, but it ignores the local structure of the data, which limits the applicability of LDA in real-time HSI classification. To address this problem, Wang et al. [21] presented a locality adaptive LDA method to generalize an illustrative subspace of data sample and determine the points closely associated with spectral and spatial domains. LDA heavily relies on certain assumptions, limiting its scope to a specific context. In this regard, another most popular method is principal component analysis (PCA), which is an unsupervised dimensionality reduction technique. The application of PCA is used in the study of Tu et al. [22] for the HSI dimensionality reduction. The authors have extracted sub-cubes in the further steps, which is then decomposed into texture and background layers. The obtained texture layer is introduced to the pixel-wise classifier. The result shows the effectiveness of the presented approach under fewer training samples. In the work of Chen et al. [23], PCA is integrated with a feature engineering process based on the local binary pattern that produces multifeatured vectors. Further, a kernel extreme learning mechanism is employed for the classification task, and its parameter is optimized using the gray wolf optimization algorithm. However, due to a complex implementation strategy, the method may pose a huge computational complexity issue while executing the model training process. Despite many works, the kernel-based methods suffer convex problems and adequate selection of an adequate kernel. Recently, multilayer learning models have been recognized as advanced classification methods for HSI classification. For example, deep ConvNet via hashing semantic attribute is presented by Yu et al. [24]. In this study, a series of hash functions are produced to improve the generalization of classes and discriminative learning mechanisms into the input HSI. A large CNN is configured to perform HSI classification task. However, the presented CNN model is complex and lacks a trade-off between precise feature generalization and network complexity. The work in the context of effective feature learning is carried out by Zhang et al. [25], where the authors have presented an unsupervised learning-based feature extraction mechanism. The presented mechanism is devised using a recursive autoencoder that considers both spatial and spectral information to construct a high-level features vector for the learning model. The authors in the study of Liu et al. [26] have tried to enhance the performance of extreme learning machines by introducing the concept of transfer learning for

HSI classification. Transfer learning introduces weights and concealed biases by using instances in the source domain. The application of the feedback attention technique in CNN is presented in the work of Yu et al. [27] for HSI classification. The feedback attention is integrated to improve the feature extraction process with the semantic information from the top dense layer. This model considers spatial-spectral information for the feature analysis. Also, the computational complexity is controlled band attention technique is incorporated in the learning model. Dong et al. [28] focused on addressing the problem caused by the small training sample size by designing pixel cluster CovNet with a spatial-spectral synthesis mechanism. A co-occurrence matrix is created to store spatial attributes, and band superposition is then applied to fuse the spatial attributes with spectral features. The authors have devised a certain policy to increase the training sample size, which is then subjected to the Covnet model for the classification. However, increasing sample size may introduce non-linearity and redundancy in the training data sample, which may impact the real performance of the learning model. Different from the other works, Zhang et al. [29] presented graph convolution networks that produce operative local spectral-spatial attributes for effective HSI classification.

Hence, various research works have been done to date for the HSI classification using different approaches. However, there is still a substantial problem concerning the model complexity, overfitting issue, learning, and classification performance that needs further effort by the researchers with evolved solutions. The next section highlights significant research problems explored based on the above discussion.

III. RESEARCH PROBLEM

The prominent research gap explored in existing solution is associated with narrowed classification performance for HSI which demands more lightweight feature extraction technique using machine learning. Further, the significant research issues explored based on the review of existing literature.

- HSI data is high-dimensional, which makes supervised classification techniques very difficult to implement. In this case, the complexity of the HSI data and the limited number of training samples are the major challenges.
- Space and spectral information are essential for applying effective classification mechanisms to HSI data but are not considered in most existing studies.
- Most machine learning methods implement the recursive nature of algorithms that do not consider the characteristics and quality of the data.
- The existing deep learning-based solution is often subjected to the overfitting issue due to the lack of proper labeled and large data samples. Such a model also requires multiple attempts of training and tuning to meet the targeted requirements of the performance.
- HSSI requires higher computing resources and longer execution times which are not as prominently emphasized in the previous works. Such existing solution is not much suitable for time-sensitive and mission-critical applications.

Hence, there is a requirement to evolve up with good solution that leads towards an effective processing and classifying HSI.

IV. SYSTEM DESIGN

This section illustrates the design of the proposed system for the HSI classification using a unique and hybrid multilayer learning model. Therefore, the proposed study's ultimate goal is to extract the comprehensive and precise feature concerning 2D spatial information and 1D spectral information with neighbor pixels in the center that needed to be classified. However, it is well known that the constraint of training sample heavily impacts the learning model performance with increase in feature dimension. To this end the modelling of the proposed system aims to address Hughes phenomenon problem, pixel mixing and achieve a good trade-off between the number of limited or unbalanced training samples, and model performance, and control overfitting. The design of the proposed system for HSI classification is demonstrated in Fig. 1.

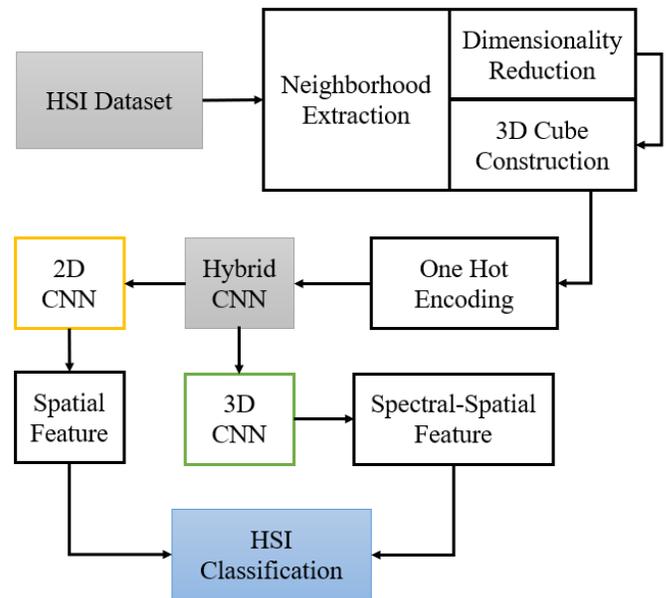


Fig. 1. Schematic Architecture of the Proposed System.

The HSI classification dataset is considered the Indian pine data collected through “airborne visible/infrared imaging spectrometer sensor” in north-western Indiana. Further, the system modelling process emphasizes on the neighborhood extraction process that incorporates two distinct operations. The first operation is subjected to dimensionality reduction as the HSI data exhibit the mixing pixel property, introducing the high intra-class variability and inter-class similarity. To this end, the study implements principal component analysis (PCA), which leads in reduction of redundant spectral information without losing spatial information for the object identification and classification. On the other hand, the second operation is executed to construct 3D cube using frequency and spatial domain information. This operation leads to generation of HSI into several images of 5x5 pixel (5 neighboring pixel) visualization with respect to spectral analysis, spatial analysis and spectral-spatial analysis in transform domain. The obtained

3D cubes are vectorized using one-hot encoding mechanism, to make it suitable for the proposed deep learning model. The study proposes hybrid deep learning model which is a combined implementation of 2D ConvNet and 3D ConvNet that synchronizes both spectral both spatio-spectral feature concerning 2D and 3D convolution operation for processing 3D HSI cube to extract precise attributes closely related to objects of the HSI. In this way, the proposed system is computationally efficient, and can achieve better HSI classification performance without posing overfitting problem.

A. 2D ConvNet

The design of Convnet is inspired by the visual system that does not need human involvement in the feature extraction process. The 2D ConvNet utilizes 2D kernels to extract spatial features map followed by convolution operation to map input observation to the output prediction class. Essentially, the convolution is an algebraic operation executed based on the summation of the scaler product between input HSI and filter (kernel) employed to extract features from the 3D HSI cube. This kernel is a matrix that moves or strides over the input HSI data, executes scaler product with the sub-region of spatial dimension. Further, a non-linearity is introduced to the model by passing the obtained feature map through the activation function given as follows:

$$v_{i,j}^{x,y} = f \left(b_{i,j} + \sum_{m=1}^{d_{l-1}} \sum_{\rho=0}^{\gamma_{i-1}} \sum_{\sigma=0}^{\delta_{i-1}} w_{i,j,m}^{\sigma,\rho} \times v_{i-1,m}^{x+\rho,y+\sigma} \right) \quad (1)$$

Equation (1) exhibits activation $v_{i,j}^{x,y}$ of convolution operation in 2D ConvNet, with spatial positions (x,y) in j^{th} feature map at every i^{th} layer under consideration, where $f(\cdot)$ denotes activation function, $b_{i,j}$ denotes bias term at i^{th} layer and for j^{th} feature map, d_{l-1} denotes the number of feature map in $(l-1)^{\text{th}}$ layer and the kernel depth $w_{i,j}^{m,\sigma,\rho}$ at position (ρ,σ) for the j^{th} feature map of i^{th} layer, γ_i represents the width of kernel, δ_i indicates height of kernel and $w_{i,j}$ represents weight term for j^{th} feature map of i^{th} layer. In the proposed model 2D ConvNet, Relu is considered as the activation function numerically expressed as follows:

$$f(x) = f_{max}(0, x) \quad (2)$$

In the proposed 2D ConvNet only 1 convolution layer is configured and pooling layers are not taken under consideration to keep significant attributes of each pixel. The 2D ConvNet is implemented before the flattening layer and 3 fully connected layers. The reason is that with 2D ConvNet, the spatial attributes within the different spectral bands can be captured powerfully without losing vital spectral features, which is a crucial for the effective classification of HSI data.

B. 3D ConvNet

The configuration of the 3D Covnet is quite similar to the 2DCovnet. The significant difference is that it has additional layer of reordering where convolution operation is carried out using 3D kernel with multiple contagious spectral bands in the input layer using 3D information. It preserves their correlations under a spectral context by sequentially ordering images of similar bands. The operation of the 3D Covnet can be expressed as follows:

$$v_{i,j}^{x,y,z} = f \left(b_{i,j} + \sum_{m=1}^{d_{l-1}} \sum_{\rho=0}^{\gamma_{i-1}} \sum_{\sigma=0}^{\delta_{i-1}} \sum_{n=0}^{n_{i-1}} w_{i,j,m}^{\sigma,\rho,n} \times v_{i-1,m}^{x+\rho,y+\sigma,z+n} \right) \quad (3)$$

Where, n_i denotes the size of the 3-D kernel along the spectral dimension and remaining parameters are similar as expressed in equation (1). In the proposed 2D ConvNet 3 convolution layer is configured and Relu is used as activation function.

V. SYSTEM IMPLEMENTATION

Mathematically, the HSI data considered in the study expressed as $I = (I_1, I_2, I_3 \dots I_D)^T \in R^{D \times (M \times N)}$, where D indicates number of spectral bands containing $(M \times N)$ images per band subjected to the output class $Y = (Y_1, Y_2, Y_3 \dots Y_C) \in R^{1 \times 1 \times C}$, where C indicates object classes. The major operation in HSI classification is assign an exclusive label to each pixel according to the both spatial and spectral features. Therefore, the classification of HSI can be considered as domain mapping problem, where mapping function $f(x)$ takes input data I and after applying some transformation operation, the function should provide matching class Y given as follows:

$$Y = f(I, \theta)$$

Where θ denotes learning adjustable parameter required in the feature learning and mapping process. The following are the steps carried out for implementing proposed hybrid Covnet for processing 3D HSI data.

Step: 1 Load the HSI dataset $\rightarrow I \in R^{D \times (M \times N)}$. The input data contains a dimension of $145 \times 145 \times 200$, where (145×145) is the dimension of the image and 200 is the number of spectral bands. Fig. 2 presents a sample visualization of image at random bands under range of 200.

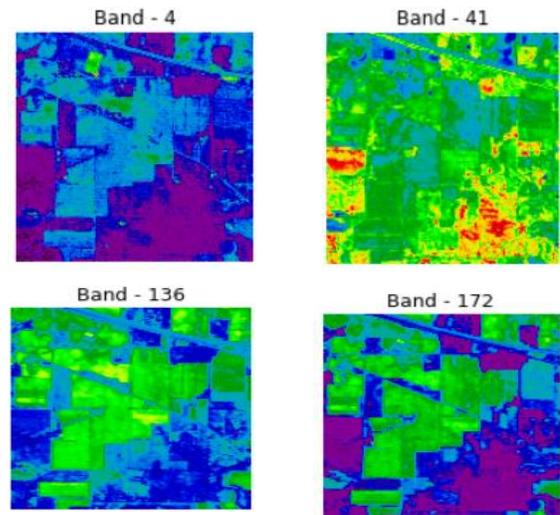


Fig. 2. Sample Visualization of HSI Bands.

Step: 2 Since, the HSI is of high dimension and often contains mixed pixel posing huge inter-class similarity and intra-class variability. Therefore, the study applies a PCA technique to overcome these issues to an extent by reducing redundant spectral information. As a result, reduced spectral band is attained while preserving spatial information. Mathematically, this operation can be given as follows:

$$I \in \mathbb{R}^{B \times (M \times N)} \leftarrow f_{PCA}(I \in \mathbb{R}^{D \times (M \times N)})$$

The processing of input data (I) using PCA returns a reduced number of spectral bands (B) such that $D \rightarrow B$, while keeping spatial information ($M \times N$) same for carrying out object classification.

Step: 3 The next operation is to perform neighborhood extraction which is subjected to construction of set of 3D cubes each representing 3D images, where object class is decided by its centering pixel. Numerically, this operation given as follows

$$I \in \mathbb{R}^{B \times (M \times N)} \rightarrow I' \in \mathbb{R}^{B \times (w \times w)}$$

Where, ($w \times w$) denotes window size covering all 'B' at centering pixel at spatial location (p,q). In the proposed system the size of window ($w \times w$) is considered equal to (5×5). The constructed 3D cubes from input data $I \in \mathbb{R}^{B \times (M \times N)}$ is the $(M - w + 1) \times (N - w + 1)$. Therefore, the 3D data cubes (D) with spatial location (p,q) denoted as $D_{(p,q)}$ covers width $p - (w - 1)/2$ to $p + (w - 1)/2$ and height from $q - (w - 1)/2$ to $q + (w - 1)/2$ and all 'B' of the dimensionality reduced data i.e., $\in \mathbb{R}^{B \times (M \times N)}$.

Step: 4 Split the obtained set of 3D data cubes into training and testing set. Further, apply one hot encoding operation to vectorize the 3D data cubes in the training samples.

Step: 5 Develop a hybrid Covnet model using 3D Covnet and 2D Covnet. Since, it has been discussed that 2D Covnet does not process spectral information and 3D Covnet is able to process both spectral and spatial data simultaneously.

In order to attain comprehensive and precise feature learning, the proposed study performs hybridization of the 3D CovNet and 2D CovNet to leverage capability of both model in the HSI classification task. The configuration details of the proposed hybrid CovNet is mentioned in Table I. The flow procedure of the proposed system using hybrid CovNet is illustrated in Fig. 3. The proposed system comprises of many operational steps. In the first step the HSI data is loaded, and further it is subjected to the PCA for the dimensionality reduction. As a result, the original HSI data of dimension ($145 \times 145 \times 200$) is reduced to the ($145 \times 145 \times 30$). In the next process, neighborhood windowing is carried out with window size (5×5) followed by zero padding operation. This operation leads to generation of 10249 number of data cubes having dimension of ($25 \times 25 \times 30$) where 30 is the number of spectral bands and 25×25 is the spatial resolution. The next vital operation is executed to perform modelling of hybrid ConvNet which comprises of total 3 layer of 3D ConvNet and 1 layer of 2D ConvNet. The first layer of 3D ConvNet comprises kernel size of ($3 \times 3 \times 7$) with filter size 8 which after convolving provides feature map of size ($23 \times 23 \times 24 \times 8$) where (23×23) is the resolution size of 3D data cube (two spatial information) with (24) spectral bands (one spectral information) and filter size 8 subjected for the further convolving operation at next layer of 3D ConvNet. The convolution operation in Covnet is most critical process. For example, at the first layer of 3D ConvNet the input data cube

convoluted with learnable filters such as filters and 3D kernel, characterized by the weighting and bias parameter resulting in generation of the feature map. Already activation of this operation is shown in equation (3). Similarly, the second layer of the 3D ConvNet takes feature map of first layer and after convolution it provides a feature map ($21 \times 21 \times 20 \times 16$) with two spatial information (21×21) and one spectral information (20) with filter size 16. This layer comprises a kernel size of ($3 \times 3 \times 5$). On the other hand, the third layer of 3D ConvNet comprises kernel size of ($3 \times 3 \times 3$) with filter size equal to 32 which after convolving provides a feature map of size ($19 \times 19 \times 18 \times 32$) where (19×19) is the resolution size of 3D data cube (two spatial information) with (18) spectral bands (one spatial information) and filter size 32. On the sub-sequent layer 2D ConvNet is implemented with single layer which comprises kernel size of (3×3) with filter size 64 which after convolving provides feature map of size ($17 \times 17 \times 64$) where (17×17) is the spatial resolution of 3D data cube and filter size 64. As it has been already discussed that 2D ConvNet are not able to process spectral information, whereas 3D ConvNet efficiently processes spatial-spectral information. Furthermore, the reason behind implementing 3D ConvNet at three layers is due to the fact that it increases spectral-spatial feature maps for better feature generalization process. Also, 2D ConvNet is implemented at single layer is due to the fact that it efficiently recognizes spatial attributes from different spectral information without compromising the spectral information. Further, flattening layer is introduced after 2D ConvNet to flatten the multi-dimensional feature map to a single dimension vector for further processing at fully connected layers used to extract more precise features by reshaping feature maps into an n-dimension vector. The last layer of proposed Hybrid ConvNet is the classification layer which uses SoftMax classifier for the classification of HSI objects. The training of model is carried out using back-propagation algorithm with Adam optimizer.

TABLE I. CONFIGURATION DETAILS OF PROPOSED HYBRID COVNET

Layers	Shape
InputLayer	25,25,30,1
Conv3D_L1	23,23,24,8
Conv3D_L2	21,21,20,16
Conv3D_L3	19,19,18,32
Reshape	19,19,576
Conv2D_L1	17,17,64
Flatten	18496
Dense_L1	256
Dropout	256
Dense_L1	128
Dropout	128
Dense_L1	16

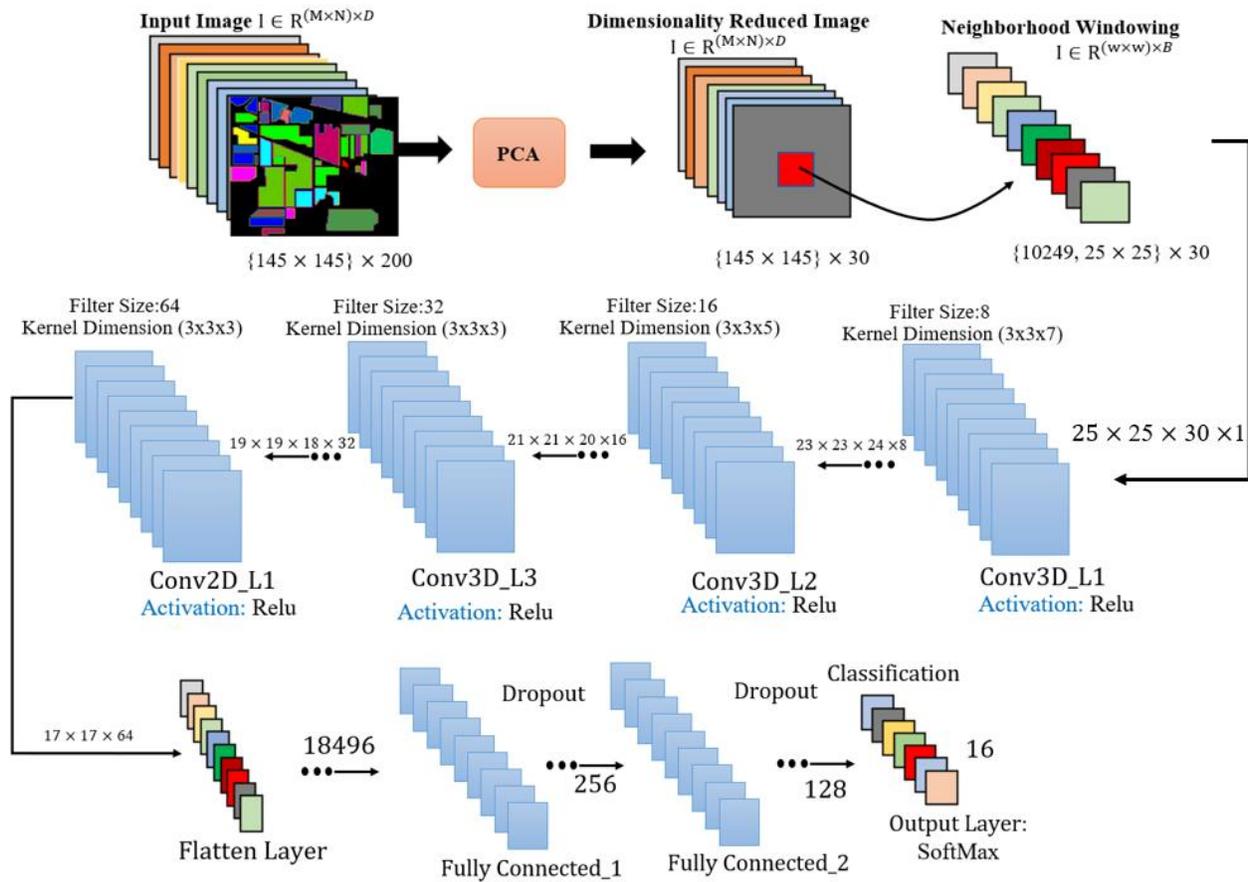


Fig. 3. Process Flow of the Proposed Hybrid CovNet.

VI. RESULT AND ANALYSIS

The implementation of the proposed system is carried out using Python programming language. The study considers Indian pine HSI dataset for the model execution. The proposed hybrid ConvNet model is trained using assignment of random weights at initial process, mini-batch size is kept equal to 256 after empirical analysis, the model is trained for 100 epochs. This section discusses the outcome obtained and performance analysis of the proposed system to justify the proposed contribution in the field of HSI processing.

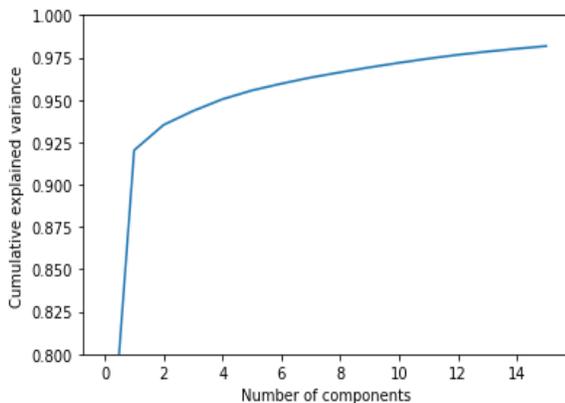


Fig. 4. Analysis of Cumulative Variance (%).

Fig. 4 shows the performance of the PCA applied for dimensionality reduction. The graph trend exhibits slightly a linear trend in the percentage variance explained by each component.

Fig. 5 shows performance analysis of proposed model regarding its loss curve in training. The graph trends lower value of loss is maintained for 80% epochs. However, at initial the loss rate is high but after 10 epochs the loss gets lower and stabilized from 20 epochs to 100 epochs.

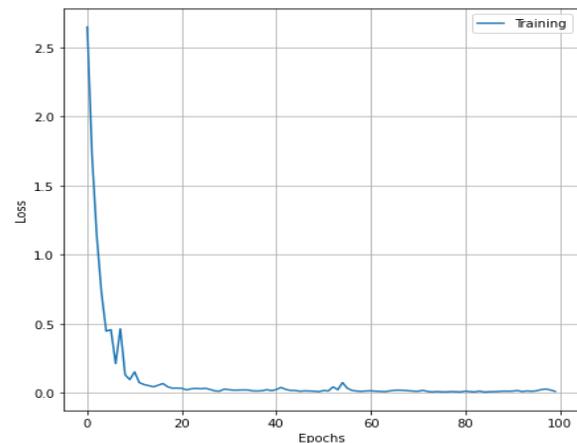


Fig. 5. Analysis of Training Loss.

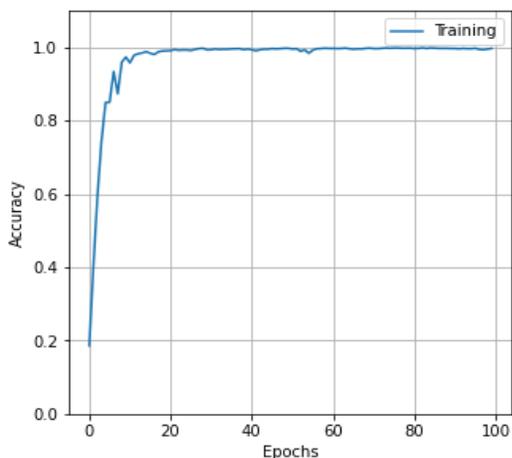


Fig. 6. Analysis of Training Accuracy.

Similar, observation can be found in Fig. 6 for the training accuracy analysis. The graph trend shows constantly higher training accuracy from 20th epoch to 100 epochs.

It can be seen from Table II that the proposed system has achieved good performance for the classification of the HSI objects. The results show a 100% precision score obtained for each class of HSI dataset. Similarly, the outcome exhibits 100% recall rate for each class except one class namely ‘Buildings-Grass-Trees-Drive’ that exhibits 99% recall rate. On the other hand, the result shows 100% F1-score for 14 classes of HSI data, and 99% for two classes namely ‘Grass-trees’ and ‘Buildings-Grass-Trees-Drive’. The performance in terms of overall accuracy is 99.9%. Therefore, the proposed model proved to be efficient and effective for the processing of the HSI data without compromising the classification performance which also evident through the confusion plot shown in Fig. 7. The performance analysis from the human visual system perspective, the ground truth of input HSI data is shown in Fig. 8 and the visualization of predictive classified outcome is presented in Fig. 9. The comparison of both figure shows that the predictive outcome is almost similar to the ground truth data, thereby exhibiting the effectiveness of the proposed system using a hybrid learning model. A closer look into the existing system showcase that proposed system is capable enough to be processed on varied number of HSI with better classification performance with respect to accuracy. Further, the learning method involved the proposed study is of hybrid form that can be used for identifying and localizing multiple form of standard land area in the HSI image.

The findings of this study based on simulation analysis also show that the proposed model has a better convergence rate. The reason behind this is that the features extracted by the proposed Hybrid ConvNet consist of fine and precise contextual attributes of HSI images. The implementation of multilayer 3D ConvNet effectively exploited both spatial and spectral information and the single-layer 2D ConvNet, exploiting rich spatial context analysis without losing spectral information. Finally, it has been found that the proposed multilayer hybrid deep learning model effectively synchronizes correlation between spatial and spectral features and provides better classification results with less training data samples.

TABLE II. ANALYSIS OF CLASSIFICATION OUTCOME

HSI classes	Precision	Recall	F1-Score
Alfalfa	1.00	1.00	1.00
Corn-notill	1.00	1.00	1.00
Corn-mintill	1.00	1.00	1.00
Corn	1.00	1.00	1.00
Grass-pasture	1.00	1.00	1.00
Grass-trees	1.00	1.00	0.99
Grass-pasture-mowed	1.00	1.00	1.00
Hay-windrowed	1.00	1.00	1.00
Oats	1.00	1.00	1.00
Soybean-notill	1.00	1.00	1.00
Soybean-mintill	1.00	1.00	1.00
Soybean-clean	1.00	1.00	1.00
Wheat	1.00	1.00	1.00
Woods	1.00	1.00	1.00
Buildings-Grass-Trees-Drive	1.00	0.99	0.99
Stone-Steel-Towers	1.00	1.00	1.00
Over all Accuracy (%)			99.9 %

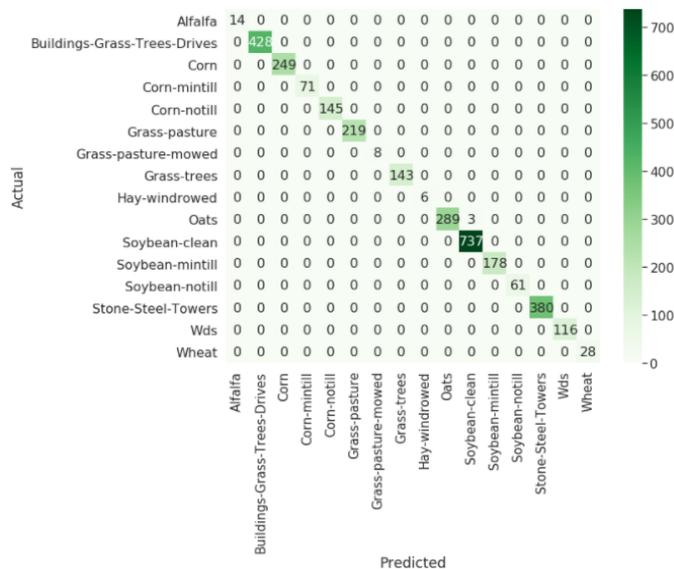


Fig. 7. Analysis of the Confusion Plot.

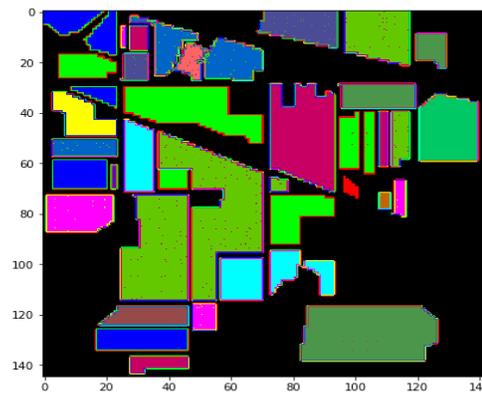


Fig. 8. Visualization of the Ground Truth of Input HSI Data.

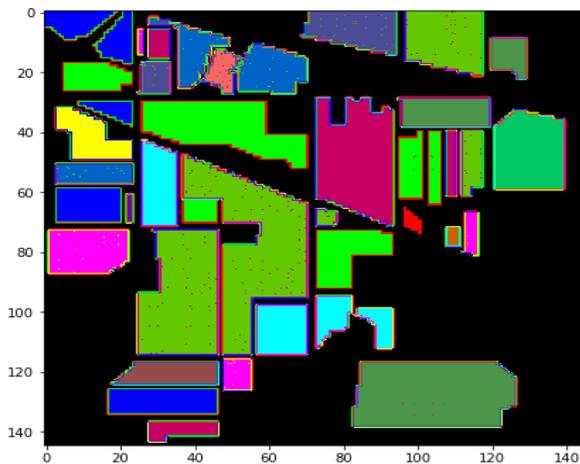


Fig. 9. Visualization of the Predicted Classified Objects.

VII. CONCLUSION

In this paper, the proposed study has explored the effectiveness of deep learning techniques for addressing issues associated with HSI classification. The proposed study has suggested modeling of hybrid learning mechanism emphasizing the trade-off between the classification performance and model overfitting problem due to the limited training data samples. The hybridization is carried out considering the application of 3D ConvNet and 2D ConvNet, which are good at exploring the spatial-spectral and spatial features. The study outcome exhibits superiority of the proposed system regarding classification performance and convergence rate. The proposed hybrid model is computationally inexpensive compared to the conventional or standalone complex 3D ConvNet. Despite the effectiveness of the proposed system, it has been realized that more optimization is required in the proposed deep learning mechanism to make it more adaptive and flexible to meet the requirement of real-time implementation. The proposed model can be introduced with other data modeling process like different preprocessing and data reduction mechanism that suits most of HSI dataset. Accordingly, it will be also very interesting to explore the application of transfer learning in future research work. The limitation of the study is associated with more extensive analysis of the outcome, which at present is restricted to Indian pine HSI dataset.

REFERENCES

- [1] Wang, C.L.; Ren, J.; Wang, H.W.; Zhang, Y.; Wen, J. Spectral-spatial classification of hyperspectral data using spectral-domain local binary patterns. *Multimed. Tools Appl.* 2018, 77, 29889–29903.
- [2] C. M. Gevaert, J. Suomalainen, J. Tang, and L. Kooistra, "Generation of spectral-temporal response surfaces by combining multispectral satellite and hyperspectral uav imagery for precision agriculture applications," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3140–3146, June 2015.
- [3] Tiwari, K.C.; Arora, M.K.; Singh, D. An assessment of independent component analysis for detection of military targets from hyperspectral images. *Int. J. Appl. Earth Obs. Geoinf.* 2011, 13, 730–740.
- [4] Hidalgo, D.R., Cortés, B.B. and Bravo, E.C., 2021. Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps. *Information Processing in Agriculture*, 8(2), pp.310-327.
- [5] Stuart, Mary B., Andrew JS McGonigle, and Jon R. Willmott. "Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems." *Sensors* 19, no. 14 (2019): 3071.
- [6] Zhang, T., Fu, Y., Wang, L. and Huang, H., 2019. Hyperspectral image reconstruction using deep external and internal learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 8559-8568).
- [7] Borzov, S.M. and Potaturkin, O.I., 2017. Efficiency of the spectral-spatial classification of hyperspectral imaging data. *Optoelectronics, instrumentation and data processing*, 53(1), pp.26-34.
- [8] Gewali, U.B., Monteiro, S.T. and Saber, E., 2018. Machine learning based hyperspectral image analysis: a survey. *arXiv preprint arXiv:1802.08701*.
- [9] Koonsanit K., Jaruskulchai C., Eiumnoh A. (2012) Band Selection for Hyperspectral Imagery with PCA-MIG. In: Bao Z. et al. (eds) *Web-Age Information Management. WAIM 2012. Lecture Notes in Computer Science*, vol 7419. Springer, Berlin, Heidelberg.
- [10] D. M. S. Arsa, H. R. Sanabila, M. F. Rachmadi, A. Gamal and W. Jatmiko, "Improving Principal Component Analysis Performance for Reducing Spectral Dimension in Hyperspectral Image Classification," *2018 International Workshop on Big Data and Information Security (IW BIS)*, 2018, pp. 123-128, doi: 10.1109/IWBIS.2018.8471705.
- [11] Q. Hou, Y. Wang, L. Jing and H. Chen, "Linear Discriminant Analysis Based on Kernel-Based Possibilistic C-Means for Hyperspectral Images," in *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1259-1263, Aug. 2019, doi: 10.1109/LGRS.2019.2894470.
- [12] Guo, Y., Liu, Y., Bakker, E.M., Guo, Y. and Lew, M.S., 2018. CNN-RNN: a large-scale hierarchical image classification framework. *Multimedia tools and applications*.
- [13] Zhao, Z.Q., Zheng, P., Xu, S.T. and Wu, X., 2019. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11), pp.3212-3232.
- [14] Attardi, G., 2015, June. Deepnl: a deep learning nlp pipeline. In *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing* (pp. 109-115).
- [15] Ahmad, M., Shabbir, S., Roy, S.K., Hong, D., Wu, X., Yao, J., Khan, A.M., Mazzara, M., Distefano, S. and Chanussot, J., 2021. Hyperspectral Image Classification-Traditional to Deep Models: A Survey for Future Prospects. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- [16] Li, Y., Zhang, H., Xue, X., Jiang, Y. and Shen, Q., 2018. Deep learning for remote sensing image classification: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(6), p.e1264.
- [17] M. Lu et al., "Penalized Linear Discriminant Analysis of Hyperspectral Imagery for Noise Removal," in *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 3, pp. 359-363, March 2017, doi: 10.1109/LGRS.2016.2643001.
- [18] Q. Hou, Y. Wang, L. Jing and H. Chen, "Linear Discriminant Analysis Based on Kernel-Based Possibilistic C-Means for Hyperspectral Images," in *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1259-1263, Aug. 2019, doi: 10.1109/LGRS.2019.2894470.
- [19] R. Hang and Q. Liu, "Dimensionality Reduction of Hyperspectral Image Using Spatial Regularized Local Graph Discriminant Embedding," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3262-3271, Sept. 2018, doi: 10.1109/JSTARS.2018.2847042.
- [20] S. Jia et al., "Flexible Gabor-Based Superpixel-Level Unsupervised LDA for Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 12, pp. 10394-10409, Dec. 2021, doi: 10.1109/TGRS.2020.3048994.
- [21] X. Wang and F. Liu, "Weighted Low-Rank Representation-Based Dimension Reduction for Hyperspectral Image Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 1938-1942, Nov. 2017, doi: 10.1109/LGRS.2017.2743018.
- [22] B. Tu, J. Wang, G. Zhang, X. Zhang and W. He, "Texture Pattern Separation for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*,

- vol. 12, no. 9, pp. 3602-3614, Sept. 2019, doi: 10.1109/JSTARS.2019.2924930.
- [23] H. Chen, F. Miao, Y. Chen, Y. Xiong and T. Chen, "A Hyperspectral Image Classification Method Using Multifeature Vectors and Optimized KELM," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 2781-2795, 2021, doi: 10.1109/JSTARS.2021.3059451.
- [24] C. Yu et al., "Hyperspectral Image Classification Method Based on CNN Architecture Embedding With Hashing Semantic Feature," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 6, pp. 1866-1881, June 2019, doi: 10.1109/JSTARS.2019.2911987.
- [25] X. Zhang, Y. Liang, C. Li, N. Huan, L. Jiao and H. Zhou, "Recursive Autoencoders-Based Unsupervised Feature Learning for Hyperspectral Image Classification," in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 11, pp. 1928-1932, Nov. 2017, doi: 10.1109/LGRS.2017.2737823.
- [26] X. Liu, Q. Hu, Y. Cai and Z. Cai, "Extreme Learning Machine-Based Ensemble Transfer Learning for Hyperspectral Image Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 3892-3902, 2020, doi: 10.1109/JSTARS.2020.3006879.
- [27] C. Yu, R. Han, M. Song, C. Liu and C. -I. Chang, "Feedback Attention-Based Dense CNN for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022, Art no. 5501916, doi: 10.1109/TGRS.2021.3058549.
- [28] S. Dong, Y. Quan, W. Feng, G. Dauphin, L. Gao and M. Xing, "A Pixel Cluster CNN and Spectral-Spatial Fusion Algorithm for Hyperspectral Image Classification With Small-Size Training Samples," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4101-4114, 2021, doi: 10.1109/JSTARS.2021.3068864.
- [29] X. Zhang, S. Chen, P. Zhu, X. Tang, J. Feng and L. Jiao, "Spatial Pooling Graph Convolutional Network for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, doi: 10.1109/TGRS.2022.3140353.