Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Hyperspectral Image Analysis using LSTM and 2D CNN and its Application in Remote Sensing

Amit Kumar Jha[1], Dr. Ram Krishna Maharjan[2,]

Dr. Nanda Bikram Adhikari[3*]

[1]Department of Electronics and Computer Engineering, Pulchowk Campus, Tribhuvan University, Lalitpur, Nepal.

[2]Professor, Department of Electronics and Computer Engineering, Pulchowk Campus, Tribhuvan University, Lalitpur, Nepal.

[3*]Associate Professor, Department of Electronics and Computer Engineering, Pulchowk Campus, Tribhuvan University, Lalitpur, Nepal.

E-mail: [1] 076msice002.amit@pcampus.edu.np, [2] rkmahajn@gmail.com, [3]adhikari@ioe.edu.np

Abstract

The land cover classification in urban areas is described in this research work. The use of hyperspectral image analysis is growing in popularity because it performs better than conventional machine learning techniques. Hypercubes, a type of three-dimensional dataset with two spatial dimensions and one spectral dimension, make up the Hyperspectral imaging (HSI). An overview of HSI's uses in remote sensing applications and the methods for classifying it are given in this research. In the field of HSI, numerous experiments are conducted with various deep learning methods for analysis and classification. The main components of this research is convolutional neural network (CNN)and long short-term memory (LSTM) that shows to be more effective than alternative models. In this case, spectral and spatial features are extracted using CNN and LSTM, respectively, and the results are then classified using support vector machines (SVM). The datasets utilized in this study were gathered using a ROSIS sensor/spectrometer at Pavia University and Indian Pines.

DOI: https://doi.org/10.36548/jiip.2023.4.002
Received: 07.10.2023, received in revised form: 04.11.2023, accepted: 20.11.2023, published: 01.12.2023

Keywords: LSTM, CNN, Gabor filter, SVM, Indian Pines, Principal Component Analysis (PCA).

1. Introduction

Target detection, material identification, land cover/use classification, archaeology, and art conservation are just a few of the many applications where the efficacy of HSI sensors has greatly increased. Remote sensing uses like surveillance, water resource management, and target detection have also profited from this advancement.

Hyperspectral images, which are primarily used in remote sensing applications, consist of several hundred spectral bands from specific areas of the earth's surface that were photographed with a spectrometer. Each pixel in an HSI is made up of a sequential spectrum that is analysed for reflectance or emissivity, and each pixel analyses the scene. One of the key elements of hyperspectral images is their spectral and spatial features, which need to be used effectively for improved classification [1].

Before the advent of deep learning, computer vision mainly used features that were manually used to develop the morphology, extract the texture, spectral information, and spatial properties from images [16]. These features were then combined with a conventional classifier, such as clustering, logistic regression, SVM, random forests, decision trees, or other well-established methods. These algorithms have been extensively studied for use in HSI data analysis for remote sensing. They are capable of multivariate, nonlinear, nonparametric regression or classification [2].

However, in order to be successfully retrieved, many machine learning techniques that rely on intricately created features that require a high level of domain expertise, and this is often challenging for complex and irregular domains. Conversely, deep learning algorithms are highly effective for these kinds of tasks. This is due, in part, to the fact that these models automatically extract from the input high-level, abstract, and discriminative characteristics. They automatically create a set of high-level, hierarchical features that are optimal for a given task [3].

In this work, Gabor filter a special class of band pass filter is used after PCA is used to decrease the image dimensionality. The Gabor filter makes sure that the image contains only particular band of frequency in certain directions surrounding the analysis region [4]. After PCA and whitening is applied to standardize each dimension's characteristics, spectral and

spatial features from the Gabor filter's output is extracted using the LSTM and CNN. This instance uses a SVM to classify the outcome.

1.1 Problem Definition

Analysis and manipulation of hyperspectral images have become more challenging since the introduction of HSI due to several challenges [5]. Spectroscopic technology was the first to hinder it, as a result of low-resolution hyperspectral sensors and inadequate data. Even though deep learning techniques have advanced and made things easier, there are still a few well-known non-dispersible problems. The following are some of them:

- Insufficient -resolution, noise-free earth observation images.
- Obstacles to feature extraction.
- Significant interclass similarity and spatial variability.
- Inadequate labelled data and a restriction on the number of training samples available.
- An uneven distribution of samples across classes.
- The dimension that is higher.

1.2 Objectives

The newest technology in biomedicine and remote sensing is called hyperspectral imaging, and it can efficiently provide spectral and spatial data about a ground object for a variety of analysis and research purposes.

The objectives of this research work are:

Detection: Differentiating target from similar background and locating and identification of small objects.

Identification: Identifying several unknown targets in the scene.

Differentiation: Differentiate among spectrally similar materials.

1.3 Scope of the Work

In this study, the model for classifying and identifying land cover in urban areas is trained using hyperspectral images from Indian Pines and Pavia University. The main components of the proposed work are CNN and LSTM, with SVM being utilized for classification. Speech recognition, natural language processing, and time series analysis, are

among the many applications for SVM because its classification accuracy is higher than that of conventional Deep learning models.

The proposed system can be applied in the various fields such:

- Agriculture
- Environmental Monitoring,
- Target Detection,
- Identification,
- Land cover/use Classification

2. Related Work

Shenming, Xiang and Zhihua 2022 [1] proposed a method for classifying HSI data that incorporates spatial and spectral features based on the Gabor filter and random patch convolution. This method increases computing efficiency by first analyzing the image using PCA dimensionality reduction, which provides valuable information. The images spatial and spectral features are then extracted employing the Gabor filter and random patches convolution, which effectively uses the spectral and the spatial information. Lastly, the spatial and spectral information is fused using the feature stacking method, and the image is classified using support vector machines.

Shahraki, et al. 2020 [4] performed analysis on two real world HSI analysis task which are tissue histology and remote sensing and evaluated the performance of 1D, 2D and 3D CNN and RNN and compared their performance also concluded that RNN worked better with remote sensing images when compared to tissue histology. The datasets used in this research were publicly available.

Giri and Panta 2022 [9] performed analysis on a 3D CNN approach for hyperspectral image classification. The factor analysis is applied for dimensionality reduction and 3D CNN is engaged for spatial and spectral feature extraction. Evaluation was done using datasets from Indian pines and Salinas.

Yin, et al. 2021 [6] has suggested a unified network using Bi-LSTM and 3D-CNN for hyperspectral image classification. The performance of the model proposed was evaluated using three datasets the results demonstrate better performance than other methods.

“Chen, Zhao and Jia 2015” [13] utilized deep belief network to perform hyperspectral image classification. The major components in this research are PCA, feature extraction based

on hierarchical learning, and logistic regression. Results with hyperspectral data indicate that given model provided competitive solution and the deep belief network has huge potential for HSI classification.

Jiang, Ma and Liu 2020 [11**]** performed analysis considering the effect of noise that is produced by human and other environmental factors as it is unavoidable and degrades the SSG (spatial spectral graph) model performance. To overcome this issue, the label noise cleansing method was proposed in their work.

Zhou, et al. 2019 [8] proposed a hyperspectral image classification using spatial-spectral LSTM in which for each pixel, the spectral value is fed one by one to spectral LSTM in order to learn the spectral features. Thereafter, PCA is applied, and row vector of extracted patch is feed to spatial LSTM to learn spatial features. For classification, SoftMax classifier is used.

Zhou, et al. 2021 [14] developed a new collaborative representation for HSI classification using spatial peak-aware collaborative representation, which displays spatial-spectral information among super-pixel clusters. The Pavia University and the Indian pine datasets were utilized to demonstrate how well the suggested method worked with a small number of samples without the need for any processing.

Xu, et al. 2018 [3] performed hyperspectral image classification based on random patch network (RPNet), where the random patches are extracted from the input HSI without training. One advantage of RPNet is its multi-scale capability, which allows for improved HSI adaptation because various objects have varying scales. The research proved that the RPNet can yield better performance than existing methods.

Qiao, et al. 2018 [15] carried out HSI analysis as part of their study project. Combined bilateral filtering and sparse representation based on spectral similarity. By utilizing joint bilateral filters and sparse representation classification (SRC), it is efficient for feature extraction as well as hyperspectral imaging data classification. The proposed method was able to improve both the kappa coefficient and classification accuracy.

3. Proposed Work

The procedure and the workflow of the suggested model are outlined in this section. Here, the model was trained using datasets of Indian pines and conducted analysis on the classification of land cover in urban areas. The major components of the model are 2D CNN and LSTM.

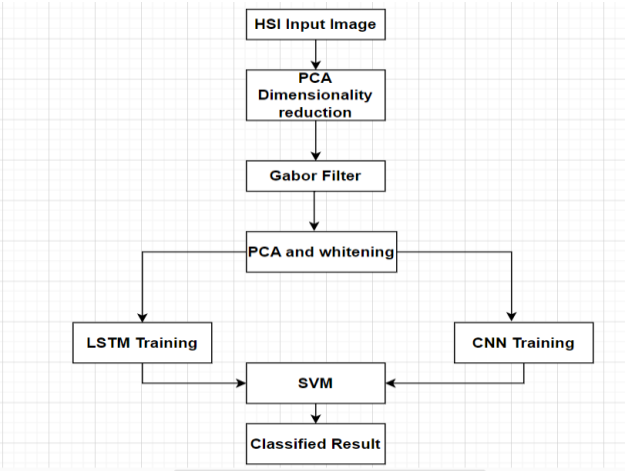

Figure 1. System Architecture

The significance of spatial components in HSI classification has increased with the advancement of image processing technologies. Because Gabor filters can provide a variety of feature information, they have attracted a lot of attention among the current methods for extracting spatial information. An efficient technique for unsupervised feature extraction is the Gabor filter.

It is less dependent on training examples, more adept at describing texture and spatial features, and capable of extracting HSI texture and spatial structure information. Furthermore, it has shown in their research that employing Gabor filters for HSI classification can result in improved results. This study suggests a hybrid spectral-spatial feature approach for HSI classification using CNN and LSTM.

This method combines CNN and LSTM network features with all spatial Gabor features, providing a high degree of discrimination [6]. By utilizing principal component analysis (PCA) to preprocess images to reduce dimensionality as well as projecting the image data in a low-dimensional feature space, this procedure prevents the "dimensionality catastrophe". Next, information about the texture as well as the internal spatial organization of the image is extracted using the Gabor filter.

PCA and whitening are done simultaneously to normalize the characteristics of each dimension in the input from the Gabor filter. After the whitened image has been processed, the spectral and spatial information is retrieved and passed to the CNN and LSTM networks. Following the stacking of spatial and spectral data for classification, during feature activation, the feature density for the returned spectral features is sparsely enhanced using a modified linear unit activation function.

3.1 PCA (Principal Component Analysis)

PCA reduces the image's dimensions while maintaining the most information possible to analyze high-dimensional data with improved interpretability. In order to achieve this, the data are linearly transformed into new coordinates that allow for a more dimensional description of the data [7].

3.2 Gabor Filter

The Gabor filter is a type of linear filter that is specifically designed for textural analysis. Its primary function is to determine whether a particular frequency is present in an image in particular directions surrounding the analysis point [1].

$$G(x, y, \lambda, \theta, \psi, \sigma, y) = \exp((-x'^2 + r^2 y'^2)/2\sigma^2) \exp[I((2\pi x')/\theta + \psi)] \quad (1)$$

- Where $X' = x\cos\theta + y\sin\theta$ and $y' = -\sin\theta + y\cos\theta$
- λ = wavelength, θ is the angel of Gabor kernel direction,
- ψ = phase offset,
- y = spatial aspect ratio,
- σ = standard deviation of the gaussian envelope,
- λ = spatial frequency bandwidth.

A 2D Gabor filter is applied to each image separately to begin the process of extracting Gabor features. Gabor's uncertainty principle, which states that the product of frequency resolution and time must be greater than a constant, serves as the process's compass. The idea facilitates a more intelligent choice of frequency and orientation.

3.3 PCA and Whitening

It describes the procedure that follows PCA dimensionality reduction for standardizing each dimension's features. Due to the similar variance and decreased correlation between wavebands, the task of classifying images can be completed more easily.

3.4 LSTM (Long Short-Term Memory)

LSTMs (RNNs) are intended to solve the vanishing gradient problem that traditional RNNs have. Because it provides RNNs with a short-term memory that lasts for thousands of time steps, it is known as LSTM. It can be applied to time-series-based data classification and prediction. [8].

A typical LSTM unit consists of an input gate, an output gate, a forget gate, and a cell. A gate controls the information flow, and a cell stores a value for any duration of time. By assigning a value between 0 and 1 to the previous state in relation to the current input, the forget gate discards information from the previous state. One value is utilized for information storage, and zero is used for discarding.

With the addition of a recurrent edge to connect the neuron to itself over time, RNNs are an effective tool for solving sequence learning problems. Suppose we have an input sequence of (x_1, x_2…..x_t) and a sequence of hidden states ($h1, h2 \ldots \ldots h_t$). At time t, the node with the recurrent edge receives input $x_t$and its previous value (h_t-1) at time instant t-1 and gives the weighted sum of them, which is formulated below.

$$H_t = \sigma (w_{hx} x_t + w_{hh} ht - 1 + b) \quad (2)$$

where w_{hx} is weight between input and recurrent hidden node, w_{hh} is weight between recurrent hidden node and itself, and the bias, whereas σ is nonlinear activation function.

In the same way that an output gate chooses which portion of a message in the current state to output while taking into account both the previous and current states, an input gate determines which portion of a new message to place in the current state. The LSTM obtains long-term dependencies for prediction, both in the present and the future, by selectively outputting messages from the current state.

3.5 CNN (Convolution Neural Network)

It is difficult to classify and recognize images in the modern world of computer vision and image processing. CNN came into existence as a result of the image's high dimensionality and complexity. When it comes to image recognition and generating precise and scalable output, CNN performs exceptionally well. The Convolution layer, fully connected layer, and pooling layer are the three layers that make up it [9].

The main distinction between CNN and other neural network models is that the former is made up of a number of convolution layers that operate on input data through convolution.

3.5.1 Convolution Layer

The convolution layer is made up of several filters or kernels that are used to identify particular patterns or features. These filters work their way across the image, producing a map that indicates the locations of these features. The result of applying a filter to an input image is combined to create a feature map, which is the output of the convolution layer.

3.5.2 Pooling Layer

Reducing the spatial dimension of the input makes processing simpler and requires less memory. It is the second layer after the convolution layer. Because this layer decreases the parameters and weights, it helps prevent overfitting and expedites the model's training.

3.5.3 Fully Connected Layer

One of the key layers in CNN, this layer implies that every neuron in the FC layer is connected to every other neuron in the layer above. This layer is added at the end and uses the information gleaned from the max pooling and convolution layers to classify the input. High dimensional data from earlier layers is received by the fully connected layer, which flattens it into a single dimension and maps it to the intended output.

3.6 Support Vector Machine

Regression and classification are performed using the Support Vector Machine a supervised machine learning technique. Nevertheless, the ideal application for regression issues is classification. The objective of the support vector algorithm is to generate an N-dimensional space hyperplane that correctly classifies the input points. The number of features determines the hyperplane's dimension. [1].

For binary classification the equation of linear hyperplane can be written as:

$$w^{\mathrm{T}}x + b = 0 \tag{3}$$

Where w represents the normal vector to the hyperplane and b represents the distance from the origin along the normal vector w of the hyperplane.

The distance between data point x_i and decision boundary is calculated as:

$$d_{\mathrm{i}} = (w^{\mathrm{T}}x_{\mathrm{i}} + b)/||w|| \tag{4}$$

Where ||w|| is the Euclidian norm of the weight vector w.

3.7Data Set Used

3.7.1 Indian Pines

Indian Pines is an HSI image segmentation dataset made up of hyperspectral band images with 145*145 pixels that were taken over the Indian Pines test site in northwest Indiana using an aerial visible/infrared imaging sensor or spectrometer. It has 200 spectral bands with a 20-m spatial resolution per pixel (24 water absorption bands are eliminated). The range of

wavelengths is 0.4 to 2.5 µm. The samples that make up the ground truth are sporadically distributed across 16 classes [10].

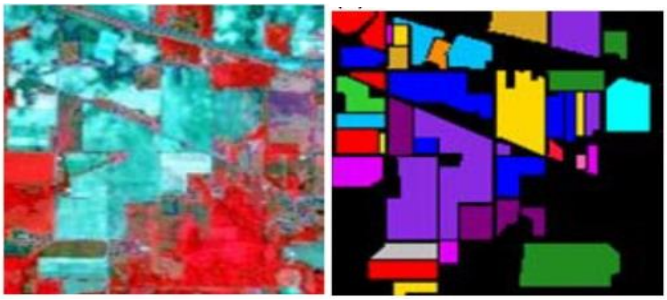

Figure 2. (A) Input Image, (B) Ground Truth

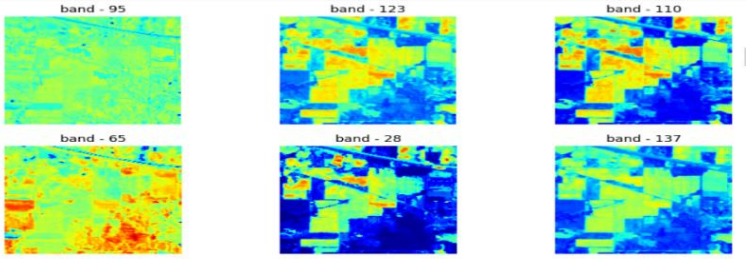

Figure 3. Sample Extracted Bands

3.7.2 Pavia University

The reflective optics system imaging spectrometer (ROSIS-3) sensor produced the hyperspectral images that make up the Pavia University dataset. These images were taken over the city of Pavia, Italy. The image has 610 x 340 pixels with a spatial resolution of 1.3 meters and 115 spectral bands. In all, 42,776 labeled samples from 9 classes are asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, brick, and the shadow present in the picture [10].

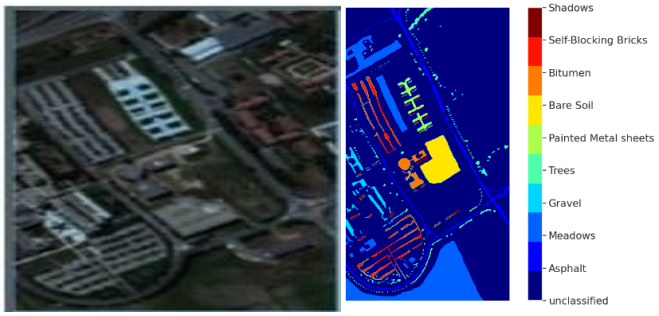

Figure 4. (A) Input Image, (B) Ground Truth

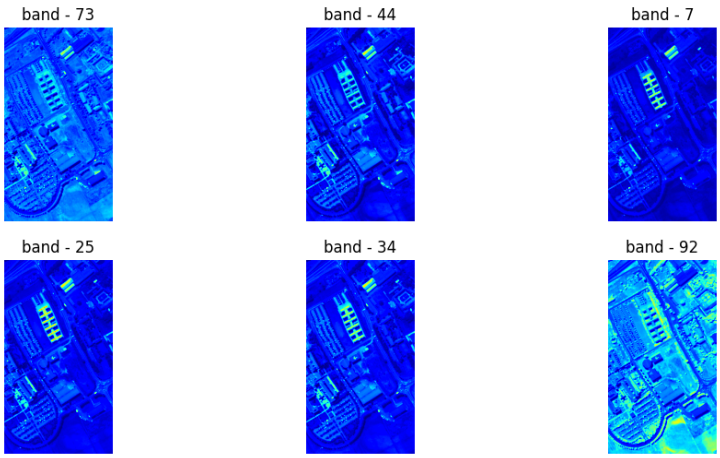

Figure 5. Sample Extracted Bands

3.8 Tools and Libraries Used

3.8.1 Python

Python is currently the most widely used language for building and training neural networks. TensorFlow, Keras, and PyTorch are just a few of the major deep learning frameworks that support Python. For machine learning and deep learning, additional Python frameworks and libraries are available, such as scikit-learn and NumPy. Because there is a large community supporting the language, finding solutions to problems might be easier.

3.8.2 Google Collaboratory

The robust tool known as Google Collaboratory allows users to write and run code directly within their browser. It offers free usage of Google's processing power, including the CPU, GPU, and TPU. Moreover, a ton of pre-installed libraries are there. It works best when a large data set is needed to train the model.

3.8.3 Overleaf

It is cloud-based latex editor which is used for writing, editing and publishing scientific research documents.

3.8.4 Pytorch

PyTorch is an open-source machine learning framework based on the Torch library, largely developed by Facebook's AI Research division. It is utilized for applications such as computer vision and natural language processing (NLP)

4. Results and Discussion

On a widely used public data set of Indian pines, several experiments are carried out to verify the efficacy of the model proposed in this paper. The data sets are hyperspectral images (HSI) from water mist that have had the absorption band removed. On a local PC, the entire experiment was run. Windows 10 Enterprise Edition was the operating system (OS), and the hardware included an Intel Core i7-4570S @ 2.90 GHz, an NVIDIA GeForce, and 8 GB of RAM.

This method is compared with other deep learning models to confirm its superiority in classifying HSI images. Here, dimensionality reduction is achieved by PCA on the input HSI image, which is subsequently passed to the Gabor filter. The features are extracted using a Gabor filter, and the data is then sent to PCA and whitening to standardize the features. CNN and LSTM networks are used to process the image and extract the spatial and spectral information. Given that various objects have varying wavelengths, the SVM is utilized to categorize the data.

The LSTM and CNN networks used in this study have strong feature extraction capabilities. Furthermore, this research's methods outperform RPNet and Gabor Based in terms of classification accuracy. The benefit of combining CNN and LSTM to extract spatial information, such as edges and textures, is also demonstrated. Alfalfa, soybean, and grass pastures in the Indian pine dataset are just a few of the features for which this method can be used to increase the classification accuracy. It also illustrates the benefits of applying different feature information at different scales to determine classification accuracy using LSTM networks.

There are very few samples available because hyperspectral datasets are expensive, hard to come by, and of poor quality. Different numbers of training samples were chosen from Indian Pine datasets and compared with LSTM for comparison experiments in order to further highlight the model's efficacy. As the number of training samples rises, we can observe that overall accuracy rises as well, with higher classification accuracy over a range of training sample counts.

4.1 Accuracy and Loss Curve

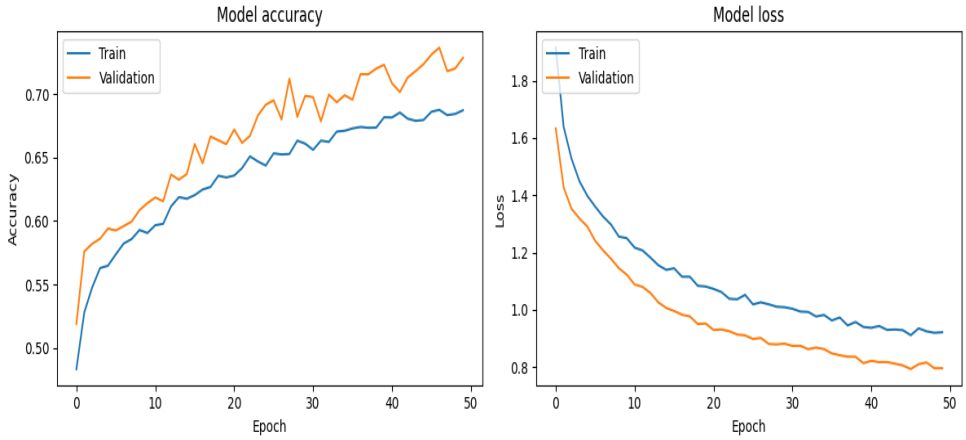

Figure 6. Accuracy and Loss Curve of Indian Pines Dataset.

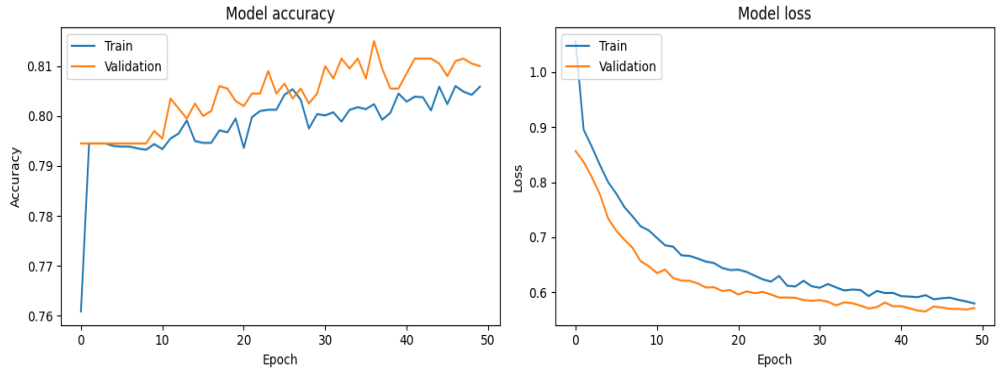

Figure 7. Accuracy and Loss Curve of Pavia University Dataset

As shown in the above Figure 7. we can see that as the number of epochs increases the loss decreases and accuracy increases. If the number of epochs is set too low the model may not get enough time to learn and can cause underfitting whereas if set too high it may cause overfitting.

4.2 Distribution of Samples in Each Band

Table 1. Distribution of Samples in Each Band of Indian Pines Dataset

	Band 1	**Band2**	**Band 3**	**Band 4**	**……..**	**Band 200**
Count	21025	21025	21025	21025	…….	21025
Mean	2957.36	4091.32	4277.50	4169.95	…….	1008.51

Std	354.91	230.39	257.82	280.761	……	7.050
Min	2560	2709	3649	2810		981.00
25%	2602	3889	4066	3954		1004.00
50%	2780	4106	4237	4126		1009.00
75%	3179	4247	4479	4350		1014.00
Max	4536	5744	6361	6362		1036.00

As seen in Table.1, there are 200 bands ranging from 400nm to 2500nm with a nominal spectral resolution 10nm. The bands between 104-108,150-163 and 220 do not contain useful information and are hence removed to enhance the classification accuracy. The data set contains two third agriculture and one third vegetation or forest.

Table 2. Distribution of Samples in Each Band of Pavia University Dataset

	Band 1	**Band2**	**Band 3**	**Band 4**	**……..**	**Band 103**
Count	207400	207400	207400	207400	…….	207400
Mean	925.32	850.23	800.36	809.88	…….	2253.45
Std	444.42	468.46	493.59	527.23	……	844.16
Min	0.00	0.00	0.00	0.00	……	0.00
25%	649.00	555.00	483.00	473.00	……	1692.00
50%	857.00	762.00	694.00	690.00	……	2261.00
75%	1111.00	1046.00	1015.00	1036.00	……	2715.00

Max	8000	8000	8000	8000	…….	8000

As shown in the above Table 2, there are 103 bands out of 115 bands, as 12 bands were removed due to noise. The size is 610 * 340 pixels per band and the spatial resolution is 1.3m per pixel. There are nine classes of interest, as shown in the ground truth.

The table above shows the distribution of samples in each band. 25 percentile means that 25% of the data lies below the given value and 75% of it lies above. In the same way, 50% and 75% are evaluated. In total, there are 21025 points of interest in each band of Indian pines and 207400 in the Pavia University, dataset and the mean and standard deviation are also evaluated.

4.3 Confusion Matrix

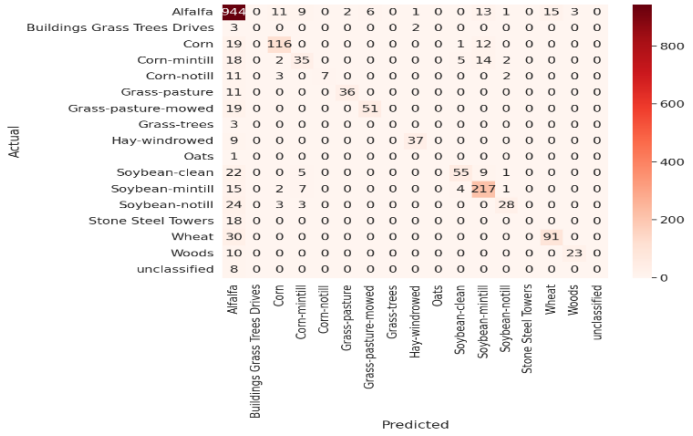

Figure 8. Confusion Matrix of Indian Pines Dataset

As seen in the above Figure 8. The diagonal represents the correctly predicted samples. Correctly predicted Alfalfa is 944 and there are some wrongly predicted too in the same way the correctly predicted corn is 116 whereas 12 soyabean-mintill is also predicted as corn.

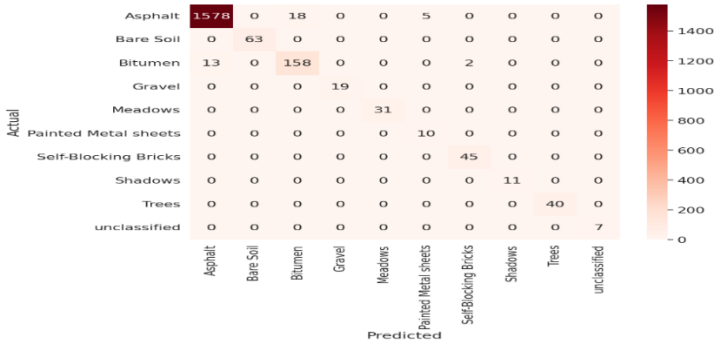

Figure 9. Pavia University Dataset Confusion Matrix

The above Figure 9, shows the confusion matrix resulting from the proposed model using Pavia university dataset. Same as above we can see here 1578 asphalt has been correctly classified whereas 18 bitumen are wrongly classified as Asphalt. This wrong classification leads to reduced accuracy hence few performance metrics are described below to determine the classification accuracy of the model.

Confusion matrix is a performance evaluation tool in machine learning that shows how accurate the model is. A confusion matrix is a N*N matrix compares the actual data with the data predicted by model. There are four important terms in confusion matrix true positive, true negative, false positive and false negative which are illustrated below.

4.4 Performance Evaluation

Table 3. Performance Evaluation of Indian Pines Dataset

	Precision	Recall	F1-score	Support
Alfalfa	0.81	0.94	0.87	1005
Corn-notill	0.00	0.00	0.00	5
Corn-mintill	0.85	0.78	0.81	148
Corn	0.59	0.46	0.52	76
Grass-Pasture	1.00	0.30	0.47	23
Grass-Trees	0.95	0.77	0.85	47
Grass-pasture-mowed	0.89	0.73	0.80	70
Hay-windrowed	0.00	0.00	0.00	3
Oats	0.93	0.80	0.86	46
Soyabean-notill	0.00	0.00	0.00	1
Soyabean-mintill	0.85	0.60	0.70	92
Soyabean-clean	0.82	0.88	0.85	246
Wheat	0.80	0.48	0.60	58

Woods	0.00	0.00	0.00	18
Trees- Grass- Drives- Building	0.86	0.75	0.80	121
Towers made of stone steel	0.88	0.70	0.78	33
unclassified	0.00	0.00	0.00	8
Accuracy			0.82	2000
Macro avg	0.60	0.48	0.52	2000
Weighted avg	0.81	0.82	0.81	2000

Table 4. Performance Evaluation of Pavia University Dataset

	Precision	**Recall**	**F1-score**	**Support**
Unclassified	0.82	0.96	0.88	1589
Asphalt	0.75	0.05	0.09	62
Meadows	0.55	0.31	0.39	176
Gravel	0.00	0.00	0.00	17
Trees	0.20	0.10	0.13	30
Painted sheets (in metal)	0.33	0.15	0.21	13
Soil (bare)	0.00	0.00	0.00	51
Bitumen	0.00	0.00	0.00	12
Self- Blocking Bricks	0.00	0.00	0.00	41
Shadows	0.00	0.00	0.00	9
Accuracy			0.80	2000
Macro avg	0.27	0.16	0.17	2000
Weighted avg	0.72	0.80	0.74	2000

The Table.3-4 above shows the performance evaluation the model on the basis of precision, recall and F1 score. As no machine learning system is 100% accurate, to determine the accuracy of the model, we have used these metrics. The precision value of 1 indicates that the system has classified 100% accurately which cannot be achieved.

- The precision is the ratio $[\boldsymbol{True\ positive}/(\boldsymbol{True\ positive} + \boldsymbol{False\ positive})]$
- The recall is the ratio $\boldsymbol{Truepositive}/(\boldsymbol{True\ positive} + \boldsymbol{False\ negative})$.
- The F1 score is the ratio $\boldsymbol{2} * [(\boldsymbol{precision} * \boldsymbol{Recall})/(\boldsymbol{precision} + \boldsymbol{Recall}]$
- The support is the number of occurrences of each class.

4.5 Ground Truth Predicted by Model

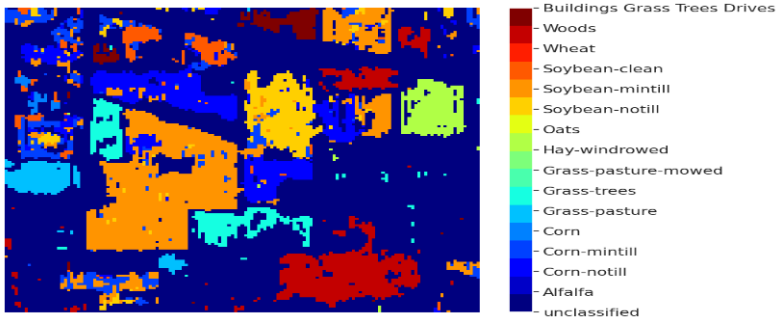

Figure 10. Predicted Ground Truth of Indian Pines Dataset

As seen in Figure 10**,** the color images show 16 different classes in 16 different colors extracted from the HSI. These spectral reflectance signatures show that classes representing various material qualities frequently show a distinctive spectral reflectance response. Therefore, deep learning techniques would produce better classification performance as they used spectral information in addition to spatial information.

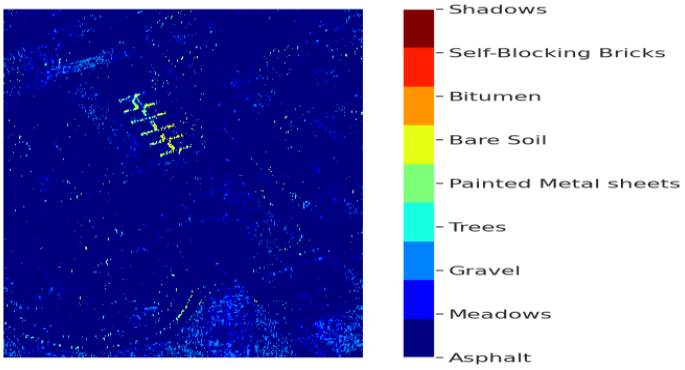

Figure 11. Predicted Ground Truth of Pavia University Dataset

Figure 11 reflects the ground truth of Pavia university dataset,showing nine different classes by different colors in the right. As different objects have differing spectral reflectance, it can easily categorize the objects by sensing its spectral signature.

5. Conclusion

This research suggests a method of classifying hyperspectral images based on CNN and LSTM that integrates information on both the spectral and spatial features in order to address the issues of low classification accuracy and fewer training samples. Owing to the large dimensionality of the available input, HSI PCA is used to reduce dimensionality. After that, we apply a Gabor filter, PCA, and whitening to standardize the input image's lost features.

After that, it is sent to CNN and LSTM for the extraction of spectral and spatial data, and SVM is used for classification. The suggested strategy outperforms other deep learning techniques in terms of classification accuracy, as has been determined. The model scales well and can train more quickly, even with the limited number of datasets available.

References

[1] Shenming, Qu, Li Xiang, and Gan Zhihua. "A new hyperspectral image classification method based on spatial-spectral features." Scientific Reports 12, no. 1 (2022): 1541.

[2] Hu, Wei, Yangyu Huang, Li Wei, Fan Zhang, and Hengchao Li. "Deep convolutional neural networks for hyperspectral image classification." Journal of Sensors 2015 (2015): 1-12.

[3] Xu, Yonghao, Bo Du, Fan Zhang, and Liangpei Zhang. "Hyperspectral image classification via a random patches network." ISPRS journal of photogrammetry and remote sensing 142 (2018): 344-357.

[4] Shahraki, Farideh Foroozandeh, Leila Saadatifard, Sebastian Berisha, Mahsa Lotfollahi, David Mayerich, and Saurabh Prasad. "Deep learning for hyperspectral image analysis, part ii: Applications to remote sensing and biomedicine." Hyperspectral Image Analysis: Advances in Machine Learning and Signal Processing (2020): 69-115.

[5] Datta, Debaleena, Pradeep Kumar Mallick, Akash Kumar Bhoi, Muhammad Fazal Ijaz, Jana Shafi, and Jaeyoung Choi. "Hyperspectral image classification: Potentials, challenges, and future directions." Computational Intelligence and Neuroscience 2022 (2022).

[6] Yin, Junru, Changsheng Qi, Qiqiang Chen, and Jiantao Qu. "Spatial-spectral network for hyperspectral image classification: A 3-D CNN and Bi-LSTM framework." Remote Sensing 13, no. 12 (2021): 2353.

[7] Yue, Jun, Wenzhi Zhao, Shanjun Mao, and Hui Liu. "Spectral–spatial classification of hyperspectral images using deep convolutional neural networks." Remote Sensing Letters 6, no. 6 (2015): 468-477.

[8] Zhou, Feng, Renlong Hang, Qingshan Liu, and Xiaotong Yuan. "Hyperspectral image classification using spectral-spatial LSTMs." Neurocomputing 328 (2019): 39-47.

[9] Giri, Ravi, and Dibakar Raj Pant. "A 3D CNN Approach for Hyperspectral Image Classification." (2022)..

[10] S. kakarla, "Towards data science," 20 August 2020. [Online]. Available: https://towardsdatascience.com/hyperspectral-image-analysis-classification-c41f69ac447f.

[11] Jiang, Junjun, Jiayi Ma, and Xianming Liu. "Multilayer spectral–spatial graphs for label noisy robust hyperspectral image classification." IEEE Transactions on Neural Networks and Learning Systems 33, no. 2 (2020): 839-852.

[12] Tu, Bing, Chengle Zhou, Xiaolong Liao, Guoyun Zhang, and Yishu Peng. "Spectral–spatial hyperspectral classification via structural-kernel collaborative representation." IEEE Geoscience and Remote Sensing Letters 18, no. 5 (2020): 861-865.

[13] Chen, Yushi, Xing Zhao, and Xiuping Jia. "Spectral–spatial classification of hyperspectral data based on deep belief network." IEEE journal of selected topics in applied earth observations and remote sensing 8, no. 6 (2015): 2381-2392.

[14] Zhou, Chengle, Bing Tu, Qi Ren, and Siyuan Chen. "Spatial peak-aware collaborative representation for hyperspectral imagery classification." IEEE Geoscience and Remote Sensing Letters 19 (2021): 1-5.

[15] Qiao, Tong, Zhijing Yang, Jinchang Ren, Peter Yuen, Huimin Zhao, Genyun Sun, Stephen Marshall, and Jon Atli Benediktsson. "Joint bilateral filtering and spectral similarity-based sparse representation: A generic framework for effective feature extraction and data classification in hyperspectral imaging." Pattern Recognition 77 (2018): 316-328.

[16] Anand, R., S. Veni, and J. Aravinth. "Robust classification technique for hyperspectral images based on 3D-discrete wavelet transform." Remote Sensing 13, no. 7 (2021): 1255.

[17] Huang, Wei, Yao Huang, Zebin Wu, Junru Yin, and Qiqiang Chen. "A multi-kernel mode using a local binary pattern and random patch convolution for hyperspectral image classification." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021): 4607-4620.