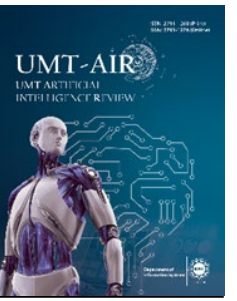
UMT Artificial Intelligence Review (UMT-AIR)
Volume 2 Issue 1, Summer 2022
ISSN$_{(P)}$: 2791-1276 ISSN$_{(E)}$: 2791-1268
Homepage: https://journals.umt.edu.pk/index.php/UMT-AIR

Article QR



Title:	**Deep Feature Learning And Classification Of Remote Sensing Images**
Author (s):	Zohaib Ahmad[1], Bushra Naz[2], Bhavani Shankar[1], Sara Ali[1], Zakir Shaikh[3]
Affiliation (s):	[1]Department IICT, Mehran University Of Engineering And Technology, Jamshoro. [2]Department of Computer Systems Engineering, Mehran University of Engineering and Technology, Jamshoro. [3]Department Electronics. Mehran University Of Engineering And Technology, Jamshoro
DOI:	https://doi.org/10.32350.umt-air.22.04
History:	Received: March 15, 2022, Revised: April 30, 2022, Accepted: June 4, 2022
Citation:	Z. Ahmad, B. Naz, S. Ali, and Z. Shaikh, "Deep feature learning and classification of remote sensing images," *UMT Artif. Intell. Rev.*, vol. 2, no. 1, pp. 00–00, 2022, doi: https://doi.org/10.32350.umt-air.21.04
Copyright:	
Licensing:	
Conflict of Interest:	Author(s) declared no conflict of interest

A publication of
Department of Information System, Dr. Hasan Murad School of Management
University of Management and Technology, Lahore, Pakistan

Deep Feature Learning and Classification of Remote Sensing Images

Zohaib Ahmad*, Bushra Naz, Bhavani Shankar, Zakir Shaikh, and Sara Ali

Mehran University of Engineering and Technology, Jamshoro, Pakistan.

***Abstract*- Hyperspectral imaging has been largely utilized in applications involving remote sensing to describe the composition of thousands of spectral bands in a single scene. Hyperspectral images (HSI) require an accurate training model for extracting the characteristics of scenes presented in an image. Image learning models involving spectral resolution present major challenges because of the complex nature of image frames. Several attempts have been made to address this complexity. Nevertheless, these models have failed to retain a deeper understanding of hyperspectral images. Since there are mixed pixels, limited training samples, and duplicate data, so the deep learning method solves the problem.In this method, spectral values (for every pixel) of the hyperspectral images are sequentially fed into spectral long-short-term memory (LSTM) through several routes to study the spectral features. Most of the existing state-of-the-art models are based on spectral-spatial frameworks. The added spatial features add more dimensions to hyperspectral images. However, these classification models do not take advantage of the sequential nature of these images. Due to the presence of mixed pixels, limited training samples, and redundant data, the utilization of deep learning techniques addresses the problems. This paper describes a method for the classification of hyperspectral images through spectral-spatial LSTM networks. For extracting the first principal constituent from such an image, principle component analysis (PCA) was used in spectral and spatial joint feature networks (SSJFN), as well as spectral and spatial individual extraction of the features via LSTM, to get the uniform end-to-end network. Furthermore, it was aimed to achieve the integration of all processes in a neural network by making a classifier to overcome the training error and backpropagation, which may lead to learning more features. During categorization, SoftMax classification considers the spatial and spectral characteristics of all the pixels independently to get two different outcomes. Afterwards, joint spectral-spatial results are gained by using the strategy of decision fusion. The classification accuracy improves by 2.69%, 1.53%, and 1.08% when compared to the rest of the state-of-art methods.**

* Corresponding Author: zohaibkk@hotmail.com

***Index Terms*- classification, deep learning, hyperspectral imaging, long-short-term memory (LSTM), spectral-spatial**

I. Introduction

Hyperspectral imaging is a three-dimensional data cube. It consists of 1-Dimensional spectral data about spectral bands, as well as 2-Dimensional spatial data about image features. In particular, spectral bands occupy a very small wavelength. At the same time, image features, such as the landcover feature and shape features show inconsistencies and relationships between adjoining pixels in different directions at the same wavelength.

Simultaneously obtaining images with high spatial and spectral resolutions has been easier and more useful after the development of hyperspectral sensors. From being used to monitor the surface of the Earth to being applied in agriculture, chemical imaging, environmental sciences, and physics-related fields, hyperspectral data has become a significant tool that requires identifying the label of every pixel through image classification [1, 2]. Several strategies for hyperspectral image (HSI) classification have been proposed. Conventional ways such as k-nearest neighbors (KNN) frequently utilize knowledge about the spectral characteristics, resulting in a "dimensionality curse" [3]. Resolving this issue requires methods for dimensionality reduction, such as principle component analysis (PCA) [4, 5] coupled with linear discriminant analysis (LDA) [6, 7]. According to [8], the support vector machine (SVM) method has been used for HSI classification. Although, it has limited sensitivity to input along with the dimensions of highness and shortness in the size of the sample. As compared to other methods, [9] SVM-based classifiers can usually perform more effectively. Still, SVM remains a superficial architecture. Shallow architectures are useful in a variety of simple or constrained issues. Whereas, in complex cases, their limited modeling and representing capacity remains inadequate, as described in [10]. Great success has been achieved in a range of machine-learning tasks using deep-learning techniques in recent years as per [3], thanks to the advancements in computing power and the availability of wide-ranging datasets. CNN, due to its limited connectivity with the properties that share weightage [10, 11], is acknowledged as being a state-of-the-art characteristic in the method of extraction, keeping in view a

variety of tasks which include computer vision [2]. Moreover, recurrent neural networks (RNNs) [12] along with their varied points have been extensively utilized with the data of sequences related to model applications, such as the recognition of spoken discourse and translation [1, 13].

Deep learning has gained importance in recent years in the remote sensing community, particularly in the classification of HSI [14, 15], for instance, when a stacked autoencoder model was put forward for unsupervised extraction of high-level features [1]. Tao et al. offered an enhanced version of the autoencoder model [16], which included a regularization concept regarding the role of energy.

As stated by [5], a deep belief network (DBN) was used for the extraction of features. It was followed by the classification of results using logistic regression classifier to which the inputs of the models were high-dimensional vectors. An alternative method for learning the spatial feature from an HSI is to flatten an image patched in the nature of locality in a vector. Even so, the technique might result in the loss of spatial information by destroying the two-dimensional structure of images. The study [17] presents a two-dimensional CNN model to resolve this concern, which may lose spectral information for using the first element of an HSI as the value input. Three-dimensional CNN [18] uses local cubes as inputs to learn both spectral and spatial features, simultaneously. Presumably, hyperspectral data correlates different spectral bands as they are compactly sampled from the entire spectrum. Firstly, it is easy to notice that adjacent spectral bands in any material are likely to have extremely similar values, implying that neighboring spectral bands largely rely upon one another. Additionally, some materials exhibit long-term interdependence amid neighboring spectral bands [13]. As described in this study, every single hyperspectral pixel is approached as a sequence of data and it represents the spectral domain dependency using long-short-term memory (LSTM) [19]. Like spectral channels, image pixels in the spatial domain are interdependent. Hence, LSTM may be utilized to extract spatial information as well. SoftMax classifiers are fed into spectral and spatial data extracted for every pixel. Combining the categorization yields the combined spectral-spatial product.

The rest of the article is laid out in this manner. The authors go through the fundamentals of LSTM in Section 2. Section 3 covers the

suggested method in depth. Section 4 discusses the trials and Section 5 offres the conclusion.

Long Short-Term Memory (LSTM): Recurrent neural networks deal with the learning of problems in sequence by involving the edges of a recurrent node for joining the neuron to itself during different times [20]. Consider that the sequence of input $\{x_1, x_2, \ldots, x_T\}$ states of hidden layers are $\{h_1, h_2, \ldots, h_T\}$. Therefore, an input x_t is received at time t by a recurrent edge node, along with the value of its previous output h_{t-1} at time $t-1$. The weighted sum of the output can be as follows:

$$h_t = \sigma\left(W_{hx}x_t + W_{hh}h_{t-1} + b\right) \quad (1)$$

- W_{hx} = weight between the recurrent hidden node to the input node
- W_{hh} = weight between itself from the previous time step and recurrent hidden node
- b = bias
- σ = nonlinear activation function

However, in training RNN models it was found that there exists an issue with the model.

The equation above assists the recurrent node of hidden h_m at time m to itself, while h_n at time n may shift towards infinity or zero as n-m increases, whether $|W_{hh}| < 1$ or $|W_{hh}| > 1$.

Due to the backpropogation error the gradient starts to either explode or vanish which makes it difficult for RNN to deal with long-term dependency problems. Therefore, a memory cell replaced recurrent hidden nodes as proposed by LSTM to solve these problems. The memory cell is shown in the figure which $\otimes$ shows dot product, whereas $\oplus$ shows matrix addition [21]. A node along with a self-connected edge of recurrent node is present in a memory cell (that has a fixed weight). This makes sure that the gradient takes transversely numerous steps unaccompanied by vanishing or detonating gradient. Based on LSTM units (consisting of input gate, for-gate, output gate plus candidate cell value), the memory cell as well as the output are given below.

$$f_t = \sigma(W_{hf}.h_{t-1} + W_{xf}.x_t + b_f)$$

$$i_t = \sigma(W_{hi}.h_{t-1} + W_{xi}.x_t + b_i)$$

$$\tilde{C}_t = tanh\left(W_{hc} \cdot h_{t-1} + W_{xC} \cdot x_t + b_C\right)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t$$

$$o_t = \sigma\left(W_{ho} \cdot h_{t-1} + W_{xo} \cdot x_t + b_o\right)$$

$$h_t = o_t \circ tanh\left(C_t\right)$$

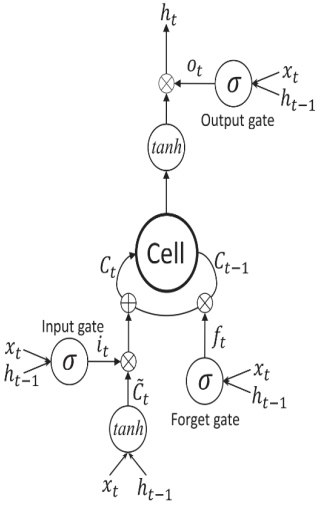


Fig. 1. Memory cell

II. Methodolgy

Figure 2 depicts the flowchart for the suggested spectral and spatial joint feature network (SSJFN). The diagram shows that SSJFNs have two main constituents: Spectral-LSTM and Spatial-LSTM. The spectral values of each pixel in a particular HSI were fed to the Spectral-LSTM, in terms of learning the spectral feature and also to subsequently obtain a classification outcome. Likewise, every pixel's local patch was fed into a spatial LSTM for extracting the spatial feature in order to get a classification result. Eventually, the resuls of the two categorizations were merged in a weighted sum approach to combine the spectral-spatial results. Each one of these steps is discussed in-depth in the subsections that follow.

A. Spatial-LSTM

The spatial feature of a pixel was extracted keeping in view its neighborhood region. Due to the presence of hundreds of channels of spectral dimenions, they have thousands of dimensions at all times. A big neighboring area would lead to a classifier with an excessively large input dimension and redundancy [1]. PCA was utilized initially to extract the first principal constituent, as inspired by the work in [1, 17]. Then, a neighborhood Xi ∈ RS × S centered on a particular pixel xi was used. The rows in this neighborhood were then transformed into an S-length sequence $\left\{X_i^1, \ldots, X_i^l, \ldots, X_i^s\right\}$, wherein X_i^l denotes the lth row X_i. Lastly, the spatial feature X_i was extracted from the sequence using LSTM. The study ended up using LSTM's final output like a new input in the layer of SoftMax in the same way that spectral features-based classifications were used to obtain the probability value $P_{spa}\left(y = j \mid x_i\right), j \in \{1, 2, \ldots, C\}$. The configurations of loss function and optimization algorithm in Spatial-LSTM are identical to those

in Spectral-LSTM. Figure 3 depicts the hypothesized spatial features-based categorization using flowchart technique.

B. Spectral-LSTM

HSI contain spectral bands in hundreds that give various spectral features of the object in the same region. Spectra exhibit multiple variations due to the complex circumstances of lighting, sensor rotations, and varied atmospheric scattering circumstances. As a result, robust and invariant characteristics must be extracted for categorization. Deep architectures seemingly have the potential of leading towards increasingly greater abstract features at higher layers, with most abstract features being resistant to the most input variations. The current research employed LSTM to extract spectral features for HSI classification using the spectral values of distinct channels as a sequence of inputs. The overview of the suggested spectral characteristics categorization technique is shown in Figure 2. To begin with, the authors selected the pixel vector $x_i \in R^{1 \times K}$, where K denotes the number of spectral bands in a specific HSI. Next, the vector was converted to a K-length series F of $\left\{ X_i^1 , \ldots, X_i^k , \ldots, X_i^k \right\}$ in which $x_i^k \in R^{1\times1}$ represents the pixel value of the k-th spectral band. This series was then fed one by one into LSTM, with the final output being fed to the Softmax classifier. Cross entropy $CE = - \sum Y \log \tilde{Y}$ was used as the loss function, with $\tilde{Y}$ and Y representing the actual and estimated labels of a pixel, accordingly. Adam algorithm [22] was used to modify this loss function. Lastly, the authors calculated $P_{spa}\left(y = j \mid x_i\right), j \in \{1, 2, \ldots, C\}$. , where C denotes the number of classes.

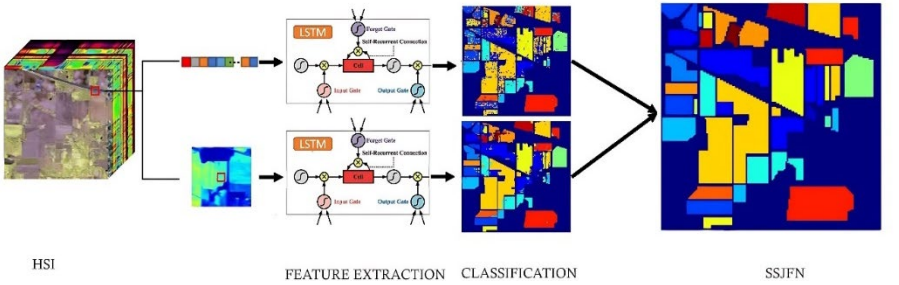


Fig. 2. Flow chart for the proposed SSJFN

C. Joint Spatial-Spectral Classification

The spectral and spatial features-based categorization techniques were introduced in the two subsections above. HSI with very high spatial resolution may now be obtained with current sensors, thanks to the development of image spectroscopy technologies [23]. As a result, the pixels in a limited spatial neighborhood are likely to belong to the same class. While, the pixels in a large homogenous region may have varied spectral sensitivities. The pixels grouped into distinct subregions in case one utilize the characteristics of the specter solely. On the other hand, if spatial information is utilized solely to classify numerous surrounding regions, they would all be categorized as the same. Hence, accurate classification needs simultaneous consideration of spatial and spectral information [7]. Using the maximum posterior probabilities $P_{spe}\left(y = j \mid x_i\right)$ and $P_{spa}\left(y = j \mid x_i\right)$, an obvious way for merging the spatial and spectral features is to combine the outcomes from the above equations including their sum manner weight that may be expressed as $P\left(y = j \mid x_i\right) = w_{spe} P_{spe}\left(y = j \mid x_i\right) + w_{spa} P_{spa}\left(y = j \mid x_i\right)$, where, w_{spe} and w_{spa} are fusion weights that satisfy $w_{spe} + w_{spa} = 1$. The current research utilized constant weights as the approach for convenience, namely $w_{spe} = w_{spa} = \frac{1}{2}$.

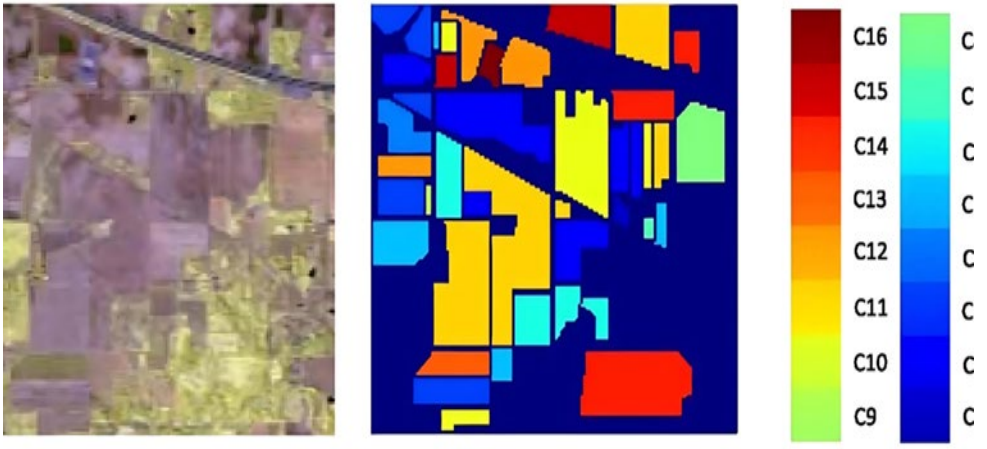


Fig. 3. IP Dataset: False-color composite with ground-truth image

Table I
Number of Pixels for Training/Testing and The Total Number of Pixels for Each Class in IP Ground Truth Map

Class	Class name	Training	Test
1	Alfalfa	30	16
2	Corn-notill	150	1278
3	Corn-mintill	150	680
4	Corn	100	137
5	Grass-pasture	150	333
6	Grass-tree	150	580
7	Grass-pasture-mowed	20	8
8	Hay-windrowed	150	328
9	Oats	15	5
10	Soybean-notill	150	822
11	Soybean-mintill	150	2305
12	Soybean-clean	150	443
13	Wheat	150	55
14	Woods	150	1115
15	Buildings-Grass-Trees-Drives	50	336
16	Stone-Steel-Towers	50	43
	Total	1765	8484

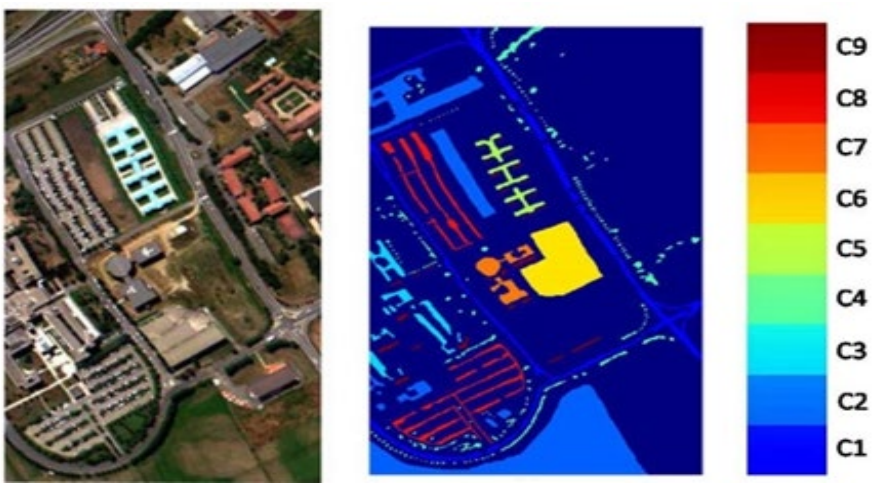


Fig. 4. PU dataset: False-colour Composite with Ground-truth Image

Table II
Number of Pixels for Training/Testing and The Total Number Of Pixels for Each Class in PUS Ground Truth Map

Class	Class name	Training	Test
1	Asphalt	548	6083
2	Meadows	540	18109
3	Gravel	392	1707
4	Trees	542	2522
5	Metal sheets	256	1089
6	Bare soil	532	4497
7	Bitumen	375	955
8	Bricks	514	3168
9	Shadows	231	716
	Total	3930	38846

III. Experimental Results

A. Datasets

We put the suggested method to test on three well-known HSI datasets commonly used to assess classification methods.

The first dataset, that is, Indian Pines (IP) which covers 224 bands of spectral energy, was captured via the sensor of AVIRIS on June 12, 1992 on the Indian Pine test site in northwestern Indiana, USA. Two-hundred bands were used after deleting four bands having zero value as well as twenty bands influenced with water absorptive way. The image consists of a spatial resolution of 20 m and a spatial size of 145 × 145 pixels. Figure 5 shows the ground-truth map and the false-color hybrid image. Table I shows the total number of samples provided, which ranges from 20 to 2455 in each class.

Pavia University (PUS): The ROSIS sensor collected the second dataset on July 8, 2002 during a flight campaign over Pavia, northern Italy. The original image has 115 spectral channels with wavelengths of the range 0.43-0.86 m. Following the removal of noisy bands, 103 bands were utilized. The image has a resolution of 1.3 m and a size of 610 × 340 pixels. A three-band false-color composite image, as well as the ground-truth map, is

shown in Figure 3. There are nine classifications of land coverings in the ground-truth map, each of which goes by the extension of 1000 labeled pixels, as displayed in Table II.

B. Experimental Setups

The performance of Spectral-LSTM, Spatial-LSTM, and SSJFNs was evaluated, both quantitatively and qualitatively, to illustrate the usefulness of the suggested LSTM-based categorization approach. They were also compared against several state-of-the-art approaches, including PCA, LDA, CNN, ELN^2 -RegMLR, RNN-LSTM, and RNN-GRU-PRetanh [10, 16]. The researchers also utilized original pixels as a benchmark, directly. To solve a singular problem in LDA, the within-class scatter matrix S_W is substituted with $S_W + \varepsilon I$, where ε = 10−3. [2, 24] is used to select the best deducted dimensions for PCA, LDA, NWFE, and RLDE. The best window size for MDA is chosen from a set of three, five, seven, nine, and eleven [25]. The number of layers and filter sizes remain the same for CNN as they are for networking [26]. The researcher only utilized a confined layer in LSTM. Moreover, a numerical range of optimal confined nodes was also chosen among the set of {16, 32, 64, 128, 256}.

Furthermore, 10% of pixels per class were chosen at random for training sets, while the remaining pixels were for the datasets of IP selected as testing sets. The authors chose 3921 pixels at random as the training set, with the remaining pixels were selected as the testing set for the PUS dataset [27]. Tables 1–3 show the precise training numbers and testing samples. Every algorithm was run five times to limit the impact of random selection and the average results are provided. Overall accuracy (OA), average accuracy (AA), per-class accuracy, and Kappa coefficient were used to assess classification performance. OA is the percentage of agreement adjusted by the number of agreements that would be expected simply by chance. While, AA is accuracy average of all the classes.

C. Parameter Selections

The size of neighboring areas and the number of hidden nodes are two crucial criteria in the proposed classification system. To begin with, hidden node numbers were specified and the best area size was chosen from a range of {8 × 8, 16 × 16, 32 × 32, 64 × 64}. Table IV shows the OAs of the SSLSTMs technique. Regarding the PUS dataset, it can be seen that as the

region size grows larger, OA increases at first then it decreases, nominating 32 × 32 as the best size. OA grows as the size of IP and KSC datasets becomes larger. Larger sizes, on the other hand, dramatically increase processing time. Thus, for both IP and KSC datasets, the optimal size was determined to be 64 × 64.

Secondly, region size is fixed and uses four alternative combinations of {16, 32}, {32, 64}, {64, 128}, and {128, 256} to find the the optimal number of hidden nodes for Spectral-LSTM and Spatial-LSTM. The SSJFN approach obtains maximum OA on the IP dataset when the number of hidden nodes for Spectral-LSTM and Spatial-LSTM is set to 64 and 128 accordingly, as shown in Table V. Whenever, the number of conceived nodes for Spectral-LSTM and Spatial-LSTM are organized up to 128 and 256, SSJFNs correspondingly achieve the highest OA on the PUS datasets.

D. Performance Comparison

Table III shows the OA and AA for multiple classification methods applied on the Indian Pines dataset. OA, as well as AA, are the lowest for PCA among all the available methods. This is because PCA does not take into account spatial features while classifying spectral features. Moreover, LDA shows better accuracy because it uses labeled data for training. Furthermore, MDA and RIDE are comparatively better than previous LDA-based methods because of their ability to use spectral and spatial features, simultaneously. CNN, for the same reason, outperforms all of them as it uses spatial and spectral features for making its predictions. It is noteworthy that ELN^2-RegMLR showed the second best results with an overall accuracy of 97.93%, followed by RNN-GRU-PRetanh, and RNN-LSTM. This makes it evident that spatial features are equally important while classifying spectral objects. The proposed algorithm, that is, SSJFN outperforms all its predecessors, including regularized machine learning algorithms, such as ELN^2-RegMLR and deep neural networks, such as RNN-LSTM. The reason is that it incorporates all the features and also has the ability to capture the non-linear distribution of hyperspectral data.

Table III
Classification Accuracy (%) for the Indian Pines Image Using Training and Testing Samples

Class	Orignal	PCA	LDA	CNN	$ELN^{2} - \text{Re}\, gMLR$	RNN-LSTM	RNN-GRU-PRetanh	SSJFN
1	56.96	59.57	63.04	73.17	97.92	46.03	70.59	99.87
2	79.75	68.75	72.04	93.48	96.67	61.73	70.28	98.16
3	66.60	53.95	57.54	84.02	97.07	86.96	81.52	99.24
4	59.24	55.19	46.58	83.57	100.0	87.02	90.16	99.81
5	90.31	83.85	91.76	96.69	98.21	86.66	91.97	98.97
6	95.78	91.23	94.41	99.15	99.40	97.49	96.13	98.51
7	80.00	82.86	72.14	93.60	13.04	59.69	84.75	98.08
8	97.41	93.97	98.74	99.91	100.0	64.89	59.64	99.91
9	35.00	34.00	26.00	63.33	22.27	60.46	86.17	99.44
10	66.32	64.18	60.91	82.15	99.68	98.77	99.38	100.0
11	70.77	74.96	76.45	92.76	98.55	75.32	84.75	100.0
12	64.42	41.72	67.45	91.35	98.42	71.82	77.58	99.76
13	95.41	93.46	96.00	99.13	98.97	91.11	95.56	97.98
14	92.66	89.45	93.79	98.22	100.0	79.49	84.62	99.80
15	60.88	47.77	65.54	87.84	96.49	90.91	90.91	99.59
16	87.53	88.17	83.66	94.29	95.29	100.0	100.0	99.86
OA	77.44	72.58	76.67	80.97	97.93	80.52	88.63	98.37
AA	74.94	70.19	72.88	80.94	88.41	78.65	85.26	94.63
K	74.32	68.58	73.27	78.25	97.64	63.72	73.66	98.10

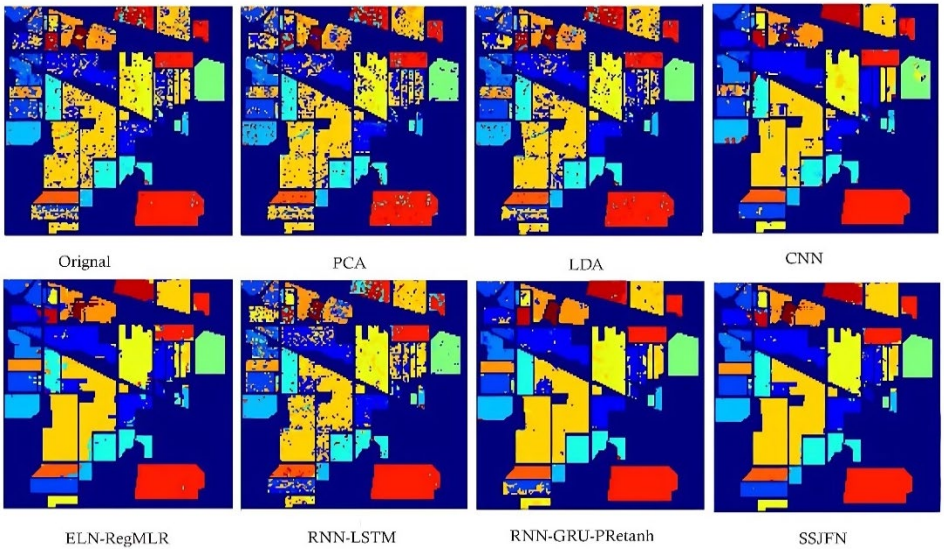


Fig. 5. *Classification maps on the IP dataset*

Table IV
Classification Accuracy (%) for the University of Pavia Image Using Training Samples and Testing Samples

Class	Original	PCA	LDA	CNN	$ELN^{'}$ – Re $gMLR$	RNN-LSTM	RNN-GRU-PRetanh	SSJFN
1	87.25	87.07	82.91	96.72	96.09	77.45	84.45	99.89
2	89.10	88.38	80.68	96.31	89.32	61.83	85.24	93.01
3	81.99	81.96	69.21	97.15	99.65	64.60	54.31	99.66
4	95.65	95.14	95.99	96.16	94.69	97.98	95.17	99.19
5	99.76	99.76	99.90	99.81	99.32	99.18	99.93	98.79
6	88.78	88.06	89.53	94.87	99.96	91.19	80.99	98.59
7	85.92	85.32	81.11	97.44	100.0	90.90	88.35	99.80
8	86.14	86.06	85.81	98.23	98.96	92.29	88.62	99.93
9	99.92	99.92	99.92	98.04	99.02	97.47	99.89	99.14
OA	89.12	88.63	76.67	96.55	97.93	77.99	88.85	98.91
AA	90.50	90.18	72.88	97.19	88.41	85.88	86.33	98.63
K	85.81	85.18	73.27	95.30	97.64	70.28	80.48	97.91

Similarly, Table IV shows the OA and AA for the PUS dataset for multiple training models. In this case, LDA shows the lowest accuracy with 76.67%, followed closely by RNN-LSTM which gives the OA of 77.99%. However, the AA of RNN-LSTM is much better at 85.88%. Akin to the IP dataset, CNN, ELN^2-RegMLR, and SSJFN outperform all others showing an accuracy of 96.55%, 97.93%, and 98.91%, respectively. Moreover, AA, as well as κ, follow the same trend.

IV. Conclusion

An HSI classification strategy that has its basis in an LSTM network was discussed in this study. The spectral and spatial characteristics extracting complications were perceived as problems involving sequence learning, while LSTM was used to tackle those challenges of gradients, long-term dependencies, information extraction and classification performance. To understand spectral properties regarding a specific pixel in the respective HSI, its spectral values in multiple channels were fed into LSTM, individually. A common patch of images concerned with the center of the pixel was first extracted from the first basic element of the HSI for spatial feature extraction. Later on, the lines of the row of the image patch remained as input to LSTM The researchers compared the methods to state-of-the-art-methods such as CNN by carrying out experimental reviews on the three HSI obtained by various equipment (AVIRIS and ROSIS). In comparison to merely employing spectral-informed values, the experimental results showed that combining spectral and spatial information enhances performance in categorization and provides results in greater homogeneous regional values in the classified maps.

References

[1] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural network for hyperspectral image classification," *IEEE Transac. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639-3655, 2007, doi: https://doi.org/10.1109/TGRS.2016.2636241

[2] S. Singh and S. S. Kasana, "Spectral-Spatial hyperspectral image classification using deep learning," presented at the *2019 Amity Int. Conf. Artif. Intell.*, Dubai, United Arab Emirates, Feb. 04-06, 2019, doi: https://doi.org/10.1109/AICAI.2019.8701243

[3] R. Hang et al., "Robust matrix discriminative analysis for feature extraction from hyperspectral images," *IEEE J. Trans.Geosci. Remote Sens.*, vol. 10, no. 5, pp. 202-211, May. 2017, doi: https://doi.org/10.1109/JSTARS.2017.2658948

[4] H. Tulapurkar, B. Banerjee, and B. K. Mohan, "Effective and efficient dimensionality reduction of hyperspectral image using CNN and LSTM network," presented at *2020 IEEE India Geosci. Remote Sens. Symp.*, Dec. 01-04, 2020, pp. 213-216, doi: https://doi.org/10.1109/InGARSS48198.2020.9358957

[5] R. Hang , Q. Liu , H. Song , Y. Sun , F. Zhu , and H. Pei, "Graph regularized nonlinear ridge regression for remote sensing data analysis," *IEEE J. Select. Top. Appl. Earth Obser. Remote Sens.*, vol. 10, no. 1, pp. 277–285, Jan. 2017, doi: https://doi.org/10.1109/JSTARS.2016.2574802

[6] M. Ghassemi, H. Ghassemian, and M. Imani, "Deep belief network for feature fusion in Hyperspectral image Classification," presented at the *IEEE Int. Conf. Aerospace Elect. Remote Sens. Technol.*, Sep. 20-21, 2018, doi: https://doi.org/10.1109/ICARES.2018.8547136

[7] B. N. Soomro, L. Xiao, S. H. Soomro, and M. Molaei, "Bilayer elastic net regression model for supervised spectral-spatial hyperspectral image classification," *IEEE J. Select. Topics Appl. Earth Obser. Remote Sens.*, vol. 9, no. 9, pp. 4102-4116, Sept. 2016, doi: https://doi.org/10.1109/JSTARS.2016.2559524

[8] W. Zhao and S. Du, "Spectral-Spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Transac. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544-4554, Aug. 2016, doi: https://doi.org/10.1109/TGRS.2016.2543748

[9] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017a, doi: https://doi.org/10.1109/TGRS.2016.2616355

[10] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning,"

Nature, vol. 521, pp. 436-444, May 2015, doi: https://doi.org/10.1038/nature14539

[11] M. Xu, Y. Wu, P. Lv, H. Jiang, M. Luo, and Y. Ye, "miSFM: on combination of mu- tual information and social force model towards simulating crowd evacuation," *Neurocomputing*, vol. 168, pp. 529-537, Nov. 2015, doi: https://doi.org/10.1016/j.neucom.2015.05.074

[12] W. Li, G. Wu, and Q. Du, "Transferred deep learning for anomaly detection in hyper-spectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 597-601, May 2017b, doi: https://doi.org/10.1109/LGRS.2017.2657818

[13] Y. Chen, H. Jiang, C. Li, and X. Jia, "Deep feature extraction and classification of hy- per spectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp.1-20, Oct. 2016, doi: https://doi.org/10.1109/TGRS.2016.2584107

[14] R. Girshick, "Fast r-cnn," *Proc. IEEE Int. Conf. Comput. Vision*, 2015, pp. 1440-1448.

[15] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Trans. Geosci. Re- mote Sens.*, vol. 56, no. 2, pp. 937-949, Feb. 2018, doi: https://doi.org/10.1109/TGRS.2017.2756851

[16] Q. Liu, R. Hang, H. Song, and Z. Li , "Learning multiscale deep features for high-reso-lution satellite image scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 117-126, Jan. 2018, doi: https://doi.org/10.1109/TGRS.2017.2743243

[17] Y. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," *IEEE J. Select. Top. Appl. Earth Obser. Remote Sens.*, vol. 8, no. 6, pp. 2381-2392, June 2015, doi: https://doi.org/10.1109/JSTARS.2015.2388577

[18] C. Tao, H. Pan, Y. Li, and Z. Zou, "Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no.

12, pp. 2438-2442, Dec. 2015, doi: https://doi.org/10.1109/LGRS.2015.2482520

[19] W. Zhao and S. Du, "Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544-4554, Aug. 2016, doi: https://doi.org/10.1109/TGRS.2016.2543748

[20] Q. Liu, F. Zhou, R. Hang, and X. Yuan, "Bidirectional-convolutional lstm based spectral-spatial feature learning for hyperspectral image classification," *Remote Sens.*, vol. 9, no. 12, pp. e1330, Dec. 2017, doi: https://doi.org/10.3390/rs9121330

[21] H. Wu and S. Prasad, "Convolutional recurrent neural networks for hyperspectral data classification," *Remote Sens.*, vol. 9, no. 3, pp. e298, Mar. 2017, doi: https://doi.org/10.3390/rs9030298

[22] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural net- works for sequence learning," *arVix*, 2015, doi: https://doi.org/10.48550/arXiv.1506.00019

[23] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv*, 2015, doi; https://doi.org/10.48550/arXiv.1412.6980

[24] Y. Zhou, J. Peng, and C. L. P. Chen, "Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classifica- tion," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1082-1095, Feb. 2015, doi: https://doi.org/10.1109/TGRS.2014.2333539

[25] R. Hang, Q. Liu, H. Song, and Y. Sun, "Matrix-based discriminant subspace ensemble for hyperspectral image spatial-spectral feature fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 783-794, Feb. 2016, doi: https://doi.org/10.1109/TGRS.2015.2465899

[26] B. N. Soomro, N. A. jaffar, S. bhatti, and L. A. thebo, "A unique spectral spatial bayesian framework via elastic net regression for the classification of hyperspectral images",

SindhUniv. Res. Jour. (Sci. Ser.), vol. 51, no. 3, pp. 555-564, 2019, doi: http://doi.org/10.26692/sujo/2019.03.87

[27] M. Xu, H. Fang, P. Lv, L. Cui, S. Zhang, and B. Zhou, "D-stc: deep learning with spa- tio-temporal constraints for train drivers detection from videos," *Pattern Recog. Letters.*, vol. 119, pp. 222-228, Mar. 2019, doi: https://doi.org/10.1016/j.patrec.2017.09.040