Hyperspectral Image Classification using EfficientNet-B4 with Search and Rescue Operation Algorithm

S.Srinivasan[1*], Dr.K.Rajakumar[2]
[1]Corresponding Author: srinivasans.2016@vitstudent.ac.in
Research Scholar, School of Computer Science and Engineering, VIT University, Vellore, Tamilnadu, India
[2]Associate Professor, School of Computer Science and Engineering, VIT University, Vellore, Tamilnadu, India

Summary
In recent years, popularity of deep learning (DL) is increased due to its ability to extract features from Hyperspectral images. A lack of discrimination power in the features produced by traditional machine learning algorithms has resulted in poor classification results. It's also a study topic to find out how to get excellent classification results with limited samples without getting overfitting issues in hyperspectral images (HSIs). These issues can be addressed by utilising a new learning network structure developed in this study.EfficientNet-B4-Based Convolutional network (EN-B4), which is why it is critical to maintain a constant ratio between the dimensions of network resolution, width, and depth in order to achieve a balance. The weight of the proposed model is optimized by Search and Rescue Operations (SRO), which is inspired by the explorations carried out by humans during search and rescue processes. Tests were conducted on two datasets to verify the efficacy of EN-B4, with Indian Pines (IP) and the University of Pavia (UP) dataset. Experiments show that EN-B4 outperforms other state-of-the-art approaches in terms of classification accuracy.

Keywords:
Deep Learning; Features; EfficientNet-B4; Search and Rescue Operation; Learning Rate; Overfitting.

1. Introduction

The hyperspectral imaging spectrometer captures hyperspectral images (HSIs). There are several practical uses for the HSI, military target recognition, mineral examination, and agricultural manufacture [1-4]. Feature extraction and classification are the two most common processes in traditional machine learning approaches [5]. Traditional methods like feature mining and Markov random fields were used to classify HSI in the early stages of their development. In order to extract characteristics with high discrimination capacity, these strategies are ineffective [6]. A SVM algorithm was presented [7] to adapt to hyperspectral data's nonlinear structure, but this solution fails to effectively tackle the multi-classification problem. In addition, features are extracted using only spectral information from the images, the spatial information is left out [8].

Manuscript received December 5, 2023
Manuscript revised December 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.12.22

Since the introduction of deep learning (DL) technology, some DL-based algorithms are introduced to solve the above mentioned problems for HSI classification have been widely employed [9-10]. For diverse classification purposes, CNN-based algorithms show greater classification results than previous methods based on handcrafted features and can automatically extract high-level features from the HSI. On the other hand, deep CNN approaches have some drawbacks. The main problem is overfitting of large datasets, which must be addressed by the proposed model.An effective collection of feature vectors can be generated using the proposed EN-B4 method. The CNN is being pointed at an acyclic graph in this case. Non-linear functions can also be learned by this network, where the SRO approach is employed to maximise the learning rate.

The remaining paper is organized as follows: Section 2 presents the related works that are used for the classification of HIS. A brief explanation of the proposed EN-B4 network with optimized model is given in Section 3. The validation of proposed model with existing technique in terms of various metrics is provided in Section 4. Lastly, the conclusion of the research work is given in Section 5.

2. Related Works

Table 1: Study of existing techniques

Author with reference	Technique	Advantage
Roy et al. [11]	HybridSpectral Net (SN)	Extracts spatial and spectral data. High-precision classification is achieved using depth spatial features.
Zhu, L [12]	generative adversarial network (GAN)	
Paoletti et al. [13].	pyramid residual network (PyResNet)	
Zhu et al. [14]	Residual spectral-spatial attention network	Avoid the interference between the extracted spatial

Ma et al. [15]	Double-branch multi-attention (DBMA)	characteristics and spectral features, which improved the classification accuracy.
Li et al. [16]	Double-branch dual attention (DBDA)	The performance improves with the addition of modules for spatial attention and channel attention to the two branches
Cui et al. [17]	dual triple-attention network (DTAN)	Uses three branches to acquire cross-dimensional interactive information and to produce attention maps between different dimensions.
Roy et al. [18]	adaptive spectral–spatial kernel enhanced residual (A2S2K-ResNet) network	Aims to broaden the receptive field and extract more useful features.

Limitation:-

Despite the use of many outstanding classification algorithms, the classification of HSI remains a huge issue in extracting features with strong discrimination capacity due to overfitting issues. In addition, for high-HSI classifications, these aforementioned methods are effective but not sufficient due to complex structure of DL. To solve this overfitting issues, optimized proposed model is introduced and achieving high-precision image classification, which is explained in the upcoming section.

3. Proposed Methodology

In this section present the proposed methodologyas EfficientNet-B4-Based Convolutional network (EN-B4), which is why it is critical to maintain a constant ratio between the dimensions of network resolution, width, and depth in order to achieve a balance. The weight of the proposed model is optimized by SRO, which is inspired by the explorations carried out by humans during search and rescue processes. Tests were conducted on two datasets to verify the efficacy of EN-B4, with Indian Pines (IP) and the UP dataset

3.1. Data Pre-processing

Iterative linear clustering [19] and principal component analysis (PCA) [20] are applied to the HSI data to reduce computing complexity. In the HSI data, the features of all superpixels are denoted by $X = \{x_1, x_2, \ldots . x_n\} \in \mathbb{R}^{n \times s}$, where $n\ and\ s$ the sum of superpixels and spectral are bands, respectively. The feature x is determined by the average of its all pixels. The i[th] spectral band's superpixels are shown here as $Z_{S_{a1}} = \{x_{1i}, x_{2i}, \ldots x_{ni}\}$ for $i = 1,2, \ldots . , s,$ super pixel is denoted as $Z_{Se_j} = \{x_{j1}, x_{j2}, \ldots x_{js}\}$ for $j = 1, 2, \ldots \ldots \ldots, n.$ The pre-processed data is fed as an input to the classifier, which is described as follows,

3.2. EfficientNet-B4-Based Classification

EN-B4 is used to build a usable usual of feature vectors from the pre-processed HSI images. Acyclic graphs are shown on the CNN. Furthermore, this network is capable of learning very nonlinear functions as well. Inside a CNN, neurons are the most fundamental unit. The neurons that make up a CNN's numerous layers are called "neurons." It's possible to connect these neurons together by using the following equation: the output of neuron l becomes neuron l + 1:

$$a^{(l+1)} = f(w^{(l)} a^{(l)} + b^{(l)}) \quad (1)$$

where $w^{(l)}$ signifies the matrix of layers l, $b^{(l)}$ denotes bias term, and f indicates the activation function. The activation for layer l is characterized as $a(l)$. In train a CNN, keep the cost functions to a minimum, it's critical to learn$W\ and\ b$ for each layer. The weight W and the bias b should be distinct to minimise the cost, i.e. the discrepancies between the chosen output y and the actual output $f(W, b)$ of the training set $\{(x^{(1)}, y^{(1)}), \ldots, (x^{(m)}, y^{(m)})\}$ with (m) . Separate training examples' cost functions are determined as shadows:

$$J(W, b; x, y) = \frac{1}{2} ||h_{w,b}(x) - y||^2 \quad (2)$$

The activation of the final layer is represented by $h(x)$. The gradient descent method is used to perform the minimization procedure iteratively. In this method, weight is taken into account while computing cost functions' partial derivatives, and the weight is then updated by using SRO algorithm.

3.2.1. Parameter Optimization using SRO Algorithm

The SRO algorithm is exploited to fine tune the weight parameters involved in the EN-B4 model. SRO algorithm which is simulated as the research shown by humans at the searching and rescue methods. It is kept in the memory matrix (O) and the position matrix (P) is where the human location is recorded when a member of the group searches (W). Size N* D is used to build the clue matrix B, which is the left clue and human location.

$$B = \begin{bmatrix} W \\ 0 \end{bmatrix} = \begin{bmatrix} W_{11} & \cdots & W_{1D} \\ \vdots & \ddots & \vdots \\ W_{N1} & \cdots & W_{ND} \\ O_{11} & \cdots & O_{1D} \\ \vdots & \ddots & \vdots \\ O_{N1} & \cdots & O_{ND} \end{bmatrix} \quad (3)$$

The 2 step of human searching are modelling in the subsequent. 1) Social stage: the search way is signified as $SD_i = (W_i - B_k)$ in which $k \neq i$. The new solution is produced as:

$$W'_{ij} = \begin{cases} B_{ij} + r_1(W_{ij} - B_{ij}), if\ f(B_i) > f(W_i) \\ W_{ij} + r_1(W_{ij} - B_{ij}), \quad otherwishe \text{ if } r_2 > \\ W_{ij}, otherwise \\ SE \quad (4) \end{cases}$$

At this point, $f(B_i)$& $f(W_i)$ signifies the fitness function value for B_i &W_i, r_1&r_2 reperesents the arbitary value from the interavel of $[-1,1]$ and $[0,1]$, SE has been model variable range in zero and on. Based on the human places define its new place and i^{th} human is signified S:

$$w'_i = W_i + r_3(B_k - B_m), i \neq k \neq m \quad (5)$$

Every solution can defined from the solution space once the novel place is outer the solution space as:

$$\hat{W_{ij}} = \begin{cases} \frac{W_{i,j}+W_j^{max}}{2} & if\ \hat{W_{ij}} > W_j^{max} \\ \frac{W_{i,j}+W_j^{min}}{2} & if\ \hat{W_{ij}} < W_j^{max} \end{cases} \quad (6)$$

In which W_j^{max}&W_j^{min} represents the higher and lower thresholds. The efficent of detectiong the global optimal solution is enhanced as

$$ME_n = \begin{cases} W_i & if\ f(W'_i) > f(W_i) \\ ME_n & otherwise \end{cases} \quad (7)$$

$$W_i = \begin{cases} W'_i & if\ f(W'_i) > f(W_i) \\ W_i & otherwise \end{cases} \quad (8)$$

Where as ME_n signifies the n_{th} saved clues located from the memory matrix and n signifies the arbitary integer sum from 1 and N. In the clues search function, if an optimal clue could not initiate neighbouring the existing place then aproximately count of searches, humans carry to the new place.

$$USN_i = \begin{cases} USN_i + 1 & if\ (W'_1) > f(W_i) \\ 0 & otherwise \end{cases} \quad (9)$$

But the unsuccessful search number (USN) value has maximum if related to the higher in effective search sum, the human become an arbitary located from the search space by. Eq. (10), and the value of USN_i is set to 0) for that human.

$$W_{EN-B4} = W_{ij} = W_j^{min} + r_4(W_j^{max} - W_j^{min}), \ i = 1, \ldots . D \quad (10)$$

The variables W_{EN-B4} and b are updated in the following way by a single iteration of gradient descent:

$$W^{(I)} = W^{(I)} - \alpha \frac{\partial}{\partial W^{(I)}} J(W_{EN-B4}, b),$$

$$b^{(I)} = b^{(I)} - \alpha \frac{\partial}{\partial b^{(I)}} J(W_{EN-B4}, b) \quad (11)$$

It is necessary to use the BP model in order to derive an incomplete cost function partial derivative. 104 x N connections in the input layer are required for a 380×380×3 pixel image, where N is the sum of neurons in the initial layer. In order to build a sparse connection, the convolution layer assigns parameters to neurons. There are fewer parameters in the convolution layer than there are in the FC layer. As a result, it is simple to train. Fig.1 shows the construction of the projectedideal.

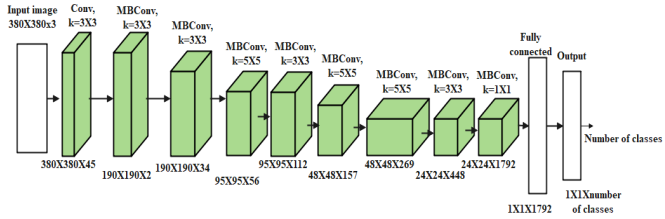

Fig. 1 Construction of the EN-B4 model

Expanding the network width, increasing the network depth, and improving input image solutions are the most commonly used approaches to increase the model's precision in DL training procedure models. It's significant to balance respectively dimension of network

resolution, width, or depth so that it could be $Y_i = F_i(X_i)$, where F_i represents an operator (op), while Y_i represents an output-tensor tensor and X_i represents an input-tensor shape tensor shape $< H_i, W_i, C_i >$ where C_i, H_i, and W_i signify the number of channels, height and width of input image. A sequence of layers could be used to identify a CNN: In other words, $Net = F_k \odot \ldots \odot F_2 \odot F_1(X_1) = \odot_{,j=1\ldots k} F_j(X_1)$. In practise, CNN layers are typically used in multiple stages of the application process, with a consistent network foundation throughout each stage. As a result, it can be concluded as:

$$Net = \bigodot_{i=1,\ldots,s} F_i^{L_i}(X_{<H_i,W_i,C_i>}) \quad (12)$$

where $F_i^{L_i}$represents the layer $F_i$which is continued$L_i$time in a phase i and $< H_i, W_i, C_i >$ represents the width, Identifying an optimal layer frameworkF_i is an important part of the typical CNN design. F_i baseline network architecture states that model scaling mainly spreads the network's resolution H_i, length and/or width (L and/or C). Meanwhile, by setting F_i , model scaling solves the implementation issues caused by a new resource limitation. As a sample design space, they could also explore L_i, H_i, W_i, C_i separately for each layer. Scaling each layer by a constant ratio will reduce the design space, according to EN. As an optimization problem, the goal is to meaningfullyprogress the accuracy of the models in the given resource constrained context:

$$\underset{d,w,r}{Max} Accuracy\big(Net(d,w,r)\big),$$

$$S.t. Net(d,w,r) = \bigodot_{j=1\ldots k} \hat{F}_i^{d.\hat{L}_i}\big(X_{<r.\hat{H}_i, r.\hat{W}_i, \omega.\hat{C}_i>}\big) \quad (13)$$

$$Memory\ (Net) \leq target_{memory},$$

$$FLOPS(Net) \leq target_{flops}$$

Where $w, d, and\ r$ represent the network scaling coefficients; $\hat{F}_i, \hat{L}_i, \hat{H}_i, \hat{W}_i$ and $\hat{C}_i$ represent the baseline network's preset parameters. An innovative compound scaling technique is then used to ensure that the network's depth and width are maintained at a constant rate.

$$Depth: d = a^{\phi},$$

$$Width: w = \beta^{\phi},$$

$$Resolution: r = \gamma^{\phi},$$

$$S.t. \alpha.\beta^2.\gamma^2 = 2, \quad (14)$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$

One of the most important aspects of the model expansion process is the allocation of additional resources to resolution, width, and depth of the network. Squeeze and excitation optimization, as well as the mobile-inverted bottleneck MBConv, are the core components. Because the EN-B4 uses the W (EN-B4) to extract spectral features, overfitting is avoided. The proposed model's additional trainable parameters are also changed during training, resulting in more refined spatial and spectral characteristics for HSI categorization.

4. Experimental Results and Analysis

Using two different sets of data, this part tests EN-ability B4's to correctly classify objects. A 32 GB RAM server, an Intel (R) i9-9900k CPU, and an NVIDIA Geforce RTX 2080TI GPU are used in all of the experiments.

4.1. Datasets Description

Two data sets were selected such as IP and UP [15], where including Infrared imaging spectrometer (AVIRIS) sensors is used to acquire the IP. Data from the ROSIS-3 sensor and the tiny airborne spectrum imager is used to create the dataset for the UP dataset.Specifically, IP has 16 feature categories with a space size of 145 × 145, and 200 spectral bands can be used for experiments. Compared with IP, UP has fewer feature categories, only nine, and the image size is 610 × 340. In addition to 13 noise bands, 103 bands are used in the experiment. Table 2 shows the description of IP and UP with its number.

Table 2: Experimental IP dataset information

IP Dataset Information		UP Dataset Information	
Number	**Class**	**Number**	**Class**
386	Bldg-Grass-Tree-Drivers	520	Spartina-marsh
93	Stone-Steel-Towers	105	Swap
46	Alfalfa	431	Graminoid-marsh
1428	Cornnotill	243	Willow-swamp
830	Cornmintill	761	Scrub
237	Corn	161	Oak/Broadleaf
483	Grass/pasture	229	Hardwood
730	Grass/trees	252	Slash-pine
28	Grass/pasture-mowed	256	CP-hammock
478	Hay-	-	-

	windrowed		
20	Oats	-	-
972	Soybeannotill	-	-
2455	Soybeanmintill	-	-
593	Soybean-clean	-	-
205	Wheat	-	-
1265	Woods	-	-

4.2. Evaluation Index

Three evaluation indicators were adopted in the experiments, the measurement unitsof these evaluation indicators are all dimensionless. The confusion matrix $H = (a_{i,j})_{n\times n}$ is constructed with the real category information of the original pixel and the predictedcategory information, where $n$is the number of categories, and $a_{i,j}$ is the sum of samplesclassified as category i by category j.

$$\boldsymbol{OA} = \frac{\sum_{i=1}^{n} \boldsymbol{a_{i,i}}}{\boldsymbol{M}} \times \mathbf{100}\% \quad (15)$$

Where $\boldsymbol{a_{i,i}}$ is the correctly classified element in the confusion matrix. Similarly, AA is the average value of classified accuracy for each class,

$$\boldsymbol{AA} = \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \frac{\boldsymbol{a_{i,j}}}{\sum_{\boldsymbol{j=a}}^{\boldsymbol{n}} \boldsymbol{a_{i,j}}} \times \mathbf{100} \quad (16)$$

The Kappa matrix is another performance evaluation index. The specific calculation isas follows:

$$Kappa = \frac{\sum_{i=1}^{n} a_{i,i} \frac{\sum_{i=a}^{n}(a_i, a_i)}{M}}{M - \frac{\sum_{i=1}^{n}(a_i, a_i)}{M}} \quad (17)$$

where a_i and a_i represent all column elements in row $i$and all row elements in column Iof confusion matrix H, respectively.

4.3. Performance Analysis of proposed model

Initially, the performance of projected model is associated with existing techniques from [11] and [17] for all classes in the IP dataset in terms of accuracy, which is shown in Fig 2.

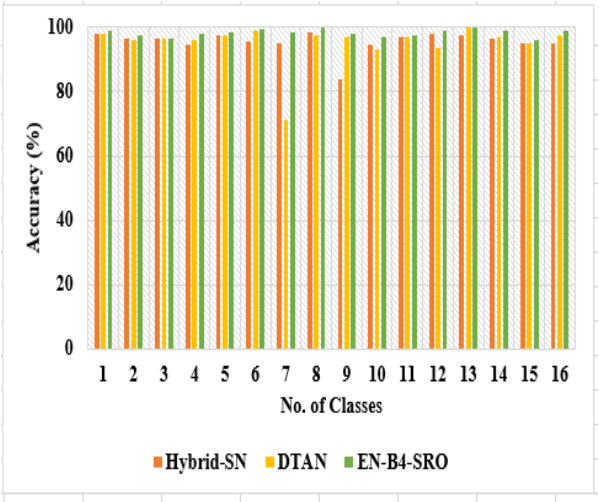

Fig. 2 Graphical Representation of various techniques on IP dataset for 16 classes.

In deep neural network (DNN) replicas such as CNN are prone to overfitting because of the large sum of parameters and the small amount of training data used to train the model. Consequently, the model is able to perform well on the training samples, but its results on the validation or testing samples are poor. Therefore, the model cannot be generalized to new data samples. This issue of overfitting can be worse in the field of remote sensing due to the complexity of data such as hyperspectral images which are self-possessed of thousands of spectral channels. To reduce the impact of such an issue, the proposed model EN-B4 is used and the weight is optimized by SRO technique.In network resoultion, each dimension or width or depth is balanced by scaling all the dimensions at a continuous ratio. Fig 3 shows the experimental analysis of proposed model on UP dataset in terms of accuracy.

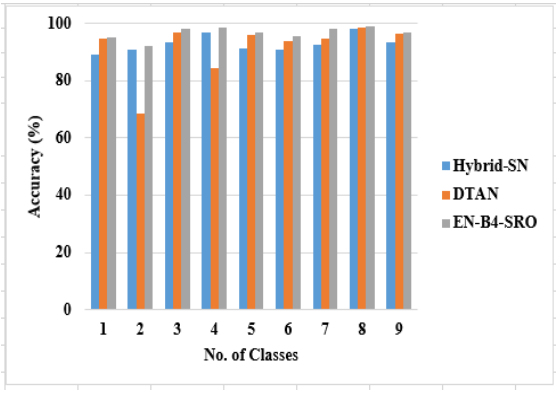

Fig. 3 Graphical Representation of various techniques on UP dataset for 9 classes.

It is found that the projected method outdoes other methods such as hybrid-SN and DTAN on all 9 classes. For instance, the hybrid-SN achieved 91.23% of accuracy, DTAN achieved 95% of accuracy and proposed model acheived 96.78% of accuracy in the class 5. In most of the various classes, the proposed model acheived 98% to 99% of accuracy, where the existing techniques achieved only 96% to 98% of accuracy only. EN-54, on the other hand, is able to learn the spatial and spectral correlations of HSI data successfully. In addition, SRO finds the appropriate weight values for exploring supplementary information collected from the extracted spatial and spectral features and suppressing noise interference. Detailed ablation experiments are carried out to evaluate how they contribute to the classification accuracy to demonstrate the usefulness of the proposed model compared to existing methodologies.Table 3shows for IP dataset and Table 4 shows for UP dataset.

Table 3: Ablation study results on IP dataset.

Methods	DTAN	Hybrid-SN	EN-B4-SRO
AA	93.37 ± 1.77	95.26 ± 1.52	97.13 ± 0.53
OA	92.36 ± 2.38	96.57 ± 0.92	97.81 ± 0.82
Kappa	95.79 ± 2.96	96.13 ± 0.57	97.92 ± 0.76

In the analysis of OA for IP dataset, hybrid-SN achieved nearly 96.57%, DTAN achieved nearly 92.36% and proposed model achieved 97.81%. The reason for better performance is that the weight is optimized using SRO technique. Not only OA, the AA is also increased, while implementing the proposed model on IP dataset. The existing techniques achieved only 93% to 95% of AA, where the proposed model achieved 97% of AA on IP dataset.

Table 4: Ablation study results on UP dataset

Methods	Hybrid-SN	EN-B4-SRO	DTAN
OA	96.53 ± 0.61	98.33 ± 0.28	96.74 ± 2.96
AA	95.29 ± 0.72	98.15 ± 0.41	97.13 ± 3.74
Kappa	95.74 ± 0.93	98.07 ± 0.62	94.92 ± 2.62

In the analysis of Kappa, the DTAN and Hybrid-SN achieved nearly 94% to 95%, where the proposed model achieved 98.07% on UP dataset. The DTAN achieved 96% of OA and 97% of AA, hybrid-SN achieved 96% of OA and 95% of AA, but the proposed model acheived 98% of AA and OA on UP dataset. HSI classification can benefit from learning spatial spectrum information in conjunction with EN-B4-SRO, as demonstrated in these experiments. The suggested model, however, outperformed the previous strategies, but by a smaller margin of error.

5. Conclusion

It was concluded that the EN-B4 model was superior to the prior state-of-the-art approaches in terms of HSI accuracy and speed. The SRO approach is used to reduce the weight of the proposed design. For IP and PU data, the experimental findings showed that using deep learning to minimise the model complexity and effectively handle the overfitting problem proved to be successful. Additionally, HSI spectral bands' computational complexity was reduced via iterative linear grouping and PCA.The results shows that the proposed model achieved 97.81% of OA on IP dataset and 98.33% of OA on UP dataset, where the Hybrid-SN model achieved 96.57% of OA on IP dataset and 96.53% of OA on UP dataset. The overfitting issue is effectively resolved and it is verified by the experimental results. However, the vanishing gradient is another issue faced by deep CNN models. To avoid this problem, an appropriate learning rate should be used, making the gradient to perform appropriate steps until convergence. However, finding the optimal learning rate is another issue, which can be resolved in the future work.

References

[1] Ghosh P. Deep Learning to Diagnose Diseases and Security in 5G Healthcare Informatics. InMachine Learning and Deep Learning Techniques for Medical Science 2022 (pp. 279-331). CRC Press.

[2]. Kwon HE, Kim JN, Kwon MJ, Lee JR, Kim SC, Nam JH, Kim BJ. The traditional medicine bojungikki-tang increases intestinal motility. Pharmacognosy Magazine. 2021 Jan 1;17(5):1.

[3] Manoharan, J. S., Sakthivel, T. G., Sangeetha, M., Sivamani. & Md. Saad Patel. (2022). A hybrid fuzzy based cross neighbor filtering (HF-CNF) for image enhancement of fine and coarse powder scanned electron microscopy (SEM) images. Journal of Intelligent and Fuzzy Systems, 42(6), 6159-6169. doi:10.3233/JIFS-212561.

[4] Wang Y, Mao H, Xu G, Zhang X, Zhang Y. A Rapid Detection Method for Fungal Spores from Greenhouse Crops Based on CMOS Image Sensors and Diffraction Fingerprint Feature Processing. Journal of Fungi. 2022 Apr 6;8(4):374.

[5] Li L, Li W, Du Q, Tao R. Low-rank and sparse decomposition with mixture of Gaussian for hyperspectral

anomaly detection. IEEE Trans Cybern 2021;51(9): 4363e72.

[6] Xu X, Zhao M, Yang J, Xiong Y, Pang F, Tan Z, Luo M. 3D laser scanning strategy based on cascaded deep neural network. Def Techno 2021.

[7]. Chen, Y.; Zhu, K.; Zhu, L.; He, X.; Ghamisi, P.; Benediktsson, J.A. Automatic design of convolutional neural network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2019, 57, 7048–7066.

[8]. Huang, H.; Duan, Y.; He, H.; Shi, G. Local linear spatial–spectral probabilistic distribution for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2020, 58, 1259–1272.

[9] Li Y, Zhang H, Shen Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. Rem Sens 2017;9(1):67.

[10] Zhu L, Chen Y, Ghamisi P, Benediktsson J. Generative adversarial networks for hyperspectral image classification. IEEE Trans Geosci Rem Sens 2018;56(9): 5046e63.

[11] Roy, S.K.; Krishna, G.; Dubey, S.R.; Chaudhuri, B.B. HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 2020, 17, 277–281.

[12] Zhu, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Generative adversarial networks for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2018, 56, 5046–5063.

[13] Paoletti, M.E.; Haut, J.M.; Fernandez-Beltran, R.; Plaza, J.; Plaza, A.J.; Pla, F. Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2019, 57, 740–754.

[14] Zhu, M.; Jiao, L.; Liu, F.; Yang, S.; Wang, J. Residual spectral-spatial attention network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2021, 59, 449–462.

[15] Ma, W.; Yang, Q.; Wu, Y.; Zhao, W.; Zhang, X. Double-branch multiattention mechanism network for hyperspectral image classification. Remote Sens. 2019, 11, 1307.

[16] Li, R.; Zheng, S.; Duan, C.; Yang, Y.; Wang, X. Classification of Hyperspectral Image Based on Double-Branch Dual-Attention Mechanism Network. Remote Sens. 2020, 12, 582.

[17] Cui, Y.; Yu, Z.; Han, J.; Gao, S.; Wang, L. Dual-Triple Attention Network for Hyperspectral Image Classification Using Limited Training Samples. IEEE Geosci. Remote Sens. Lett. 2021, 19, 1–5.

[18] Roy, S.K.; Manna, S.; Song, T.; Bruzzone, L. Attention-Based Adaptive Spectral Spatial Kernel ResNet for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. 2020, 59, 7831–7843.

[19] P. Zhong B. Du L. Zhang S. Wan, C. Gong and J. Yang. Multiscale dynamic graph convolutional network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens., 58(5):3162–3177, 2019.

[20] Sun, W., Yang, G., Peng, J. and Du, Q., 2019. Lateral-slice sparse tensor robust principal component analysis for hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters, 17(1), pp.107-111.

S Srinivasan received his Bachelor of Engineering in Computer science and Engineering from Anna University, Tamilnadu in 2008. He received his Master of Technology in Computer science and Engineering from PRIST University, Tamilnadu in 2014. He is currently pursuing his Ph.D degree with focus of Deep learning techniques in VIT University, Vellore, Tamilnadu.
His research area includes Image processing, Machine Learning with Deep learning techniques.
Email:srinivasans.2016@vitstudent.ac.in

Dr K Rajakumar received his Bachelor of Engineering in Electronics and communication engineering from Madras University, Tamilnadu. He received his Master of Engineering in Applied electronics from College of Engineering, Guindy, Anna University, Tamilnadu. He received his PhD in Image Processing from College of engineering, Guindy, Anna University, Tamilnadu.
He is currently working as an Associate Professor in School of computer science and engineering at VIT University, Vellore, Tamilnadu. He also acts as a recognized supervisor at Anna University, Chennai, Tamilnadu. He has published more than 50 papers in SCI, Scopus and IEEE Xplore.
His research area includes Image processing, Machine Learning, Big data, Deep learning and Artificial Intelligence.
Email:rajakumar.krishnan@vit.ac.in