ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2020.0304

SOFT COMPUTING APPROACHES FOR HYPERSPECTRAL IMAGE CLASSIFICATION

H.S. Prasantha, Moon Moon Chatterjee, Pasupuleti Sai Roshitha and V. Roshitha
Department of Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, India

Abstract

Hyperspectral image classification is one of the most emerging form of image classification. It is able to convey information about an image in a more detailed way as compared to RGB or multispectral data. When spectral measurement is performed using hundreds of narrow contiguous wavelength intervals, the resulting image is called a hyperspectral image. Spectral signature of thousands of materials have been measured in the laboratory and gathered into libraries. Library signatures are used as the basis for identification of materials in Hyperspectral Image (HSI) data. We analyze the spectral signature of the image to extract information. In HSI, each pixel is in fact a high dimensional vector typically containing reflectance measurement from hundreds of continuous narrow band spectral channels (FWHM between 2 and 20) and 400-2500 nm wavelength range. The range of spectrum in HSI data extends beyond the visible range. Hyperspectral data processing comes with many stages such as pre-processing, feature reduction, classification and followed by target detection. Various machine learning and deep learning algorithms have been used to classify HSI data where few of them are Support Vector Machine, Convolutional Neural Network, random forest, SSRN, etc. HSI is being used in variety of fields such as agriculture, mining, food quality, soil types, defense etc.

Keywords:

Image Classification, Convolutional Neural Network, Support Vector Machine, Hyperspectral

1. INTRODUCTION

Hyper Spectral Images (HSI) are captured by specialized remote sensors on the aircraft, and are picked up from the spectral data reflected by the ground objects in any desired area on earth. Here, contrary to recording the change of the spatial characteristics of image data, the change of the spectral characteristics of a single point called pixel along with the space in the region is mainly recorded. This enables the capture of two types of data, namely the position and distribution information of the surface object; and the spectrum, that is, the reflection intensity of each pixel in different wavelength bands. Hyperspectral images therefore contain a wealth of information, and the different substances belonging to the same species do not diminish their resolution as they are distinguished by the unique characteristics of their respective reflected spectral information. The data sets are considered as test samples and training samples.

There are many algorithms to classify hyperspectral images. In the proposed work, the author discusses the Convolutional Neural Network (CNN) and the Support Vector Machine (SVM) techniques for classifying the Hyperspectral images and evaluating the results. The following parts of this section briefly discuss HSI, and the classification of HSI using Deep Learning and Machine Learning.

1.1 HYPERSPECTRAL IMAGING

Hyperspectral imaging comes under spectral imaging, and deals with the imaging of narrow spectral bands over a spectral range which is continuous in nature, and produces the spectra of all pixels .Hyperspectral imaging not only measures each pixel in the image but also measures the reflection, emission and absorption of electromagnetic radiation. It also provides the unique spectral signature of every pixel, an attribute that can be employed in processing techniques to identify and discriminate materials.

1.2 DEEP LEARNING

Artificial intelligence has many divisions out of which one is Deep learning. Deep learning is inspired from the functioning of the human brain to process data and to make patterns. Deep learning is a subset of machine learning and is organized in a hierarchical level of artificial neural networks. It looks like a human brain where the nodes of the neuron are connected to each other. One of the benefits of deep leaning is its ability to extract features automatically from raw data. This is also known as feature learning. It is capable to perform with large number of unlabeled data as its efficiency increases with the large number of data. This study employs Convolutional Neural Networks (CNN), one of the neural networks in Deep Learning, for classifying Hyperspectral images.

1.3 MACHINE LEARNING

Machine learning is an application of artificial intelligence (AI) that will give systems the power to automatically learn and improve from experience without being explicitly programmed. The primary aim of machine learning is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly. Machine Learning is classified into Supervised and Unsupervised Machine Learning.

In Supervised Machine Learning algorithms can be applied to what has been learnt in the past to new data using labelled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to form predictions about the output values. The system is in a position to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly. In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labelled.

Unsupervised learning studies how systems can infer a function to explain a hidden structure from unlabeled data. The system does not figure out the right output, but it explores the data to draw inferences from datasets for describing hidden structures from unlabeled data. This paper uses Support Vector Machine (SVM) algorithm to classify Hyperspectral Images.

Many authors have contributed to the field of Hyperspectral Image classification using soft computing approaches. There are several attempts by researchers to classify the Hyperspectral images using various techniques or algorithms. The ensuing parts of this section strives to present a few techniques from such precedents.

Savardi et al. [1] used deep learning technologies to hyperspectral data from a multidisciplinary perspective. HSI-CNN and XG Boost methods are employed in [2] to prevent over fitting that occurs during analysis. In [3], the authors propose a novel framework that takes advantage of both CNNs and multiple feature learning to better predict the class labels for HSI pixels. Parallel morphological/neural classification algorithms are proposed to classify [4].

Vibhute and Kale [5] have used SVM classifier for the datasets received from an earth-observing-satellite, and have concluded that SVM classifier gives promising results in case of small sets of training samples. The analysis made in [6] shows that the spectral reluctance and wavelength are characteristics of the salinity of soil. Depending upon the spectral reluctance and wavelength, the soil can be classified as highly saline or less saline [6].

Shreshtha et al. [7] have studied and classified the features like absorption features, and soil clustering features. They concluded that the selection of end members is vital for adequate hyperspectral classification when applied to soil [7]. This paper intends to improve in such areas.

2. METHODOLOGY

The different steps involved in Hyperspectral image classification are discussed in detail. The preprocessing step involved, feature reduction and feature classification steps are discussed. All the steps using SVM and CNN are presented in detail.

2.1 PRE-PROCESSING OF HYPERSPECTRAL IMAGES

The first stage involves Preprocessing of Hyperspectral images to convert the given Hyperspectral images into suitable form. The Pre-processing required for CNN and SVM techniques as classifiers are presented.

2.1.1 Preprocessing Required for CNN Techniques:

The often highlighted benefit of HSI is its combination of detecting spectral and imaging dimensions techniques in the process of imaging the spatial feature, wherein many hundreds of narrow bands are scattered enabling the continuous spectral coverage of each spatial pixel. The data thus formed is called three-dimensional data blocks. There are three axes named as x, y and λ. Here, x and y are two dimensional coordinate axes representing the planar pixel information, while the third coordinate axis λ represents information about the wavelength. An individual pixel holding the spectral bands is termed as sample of category for training.

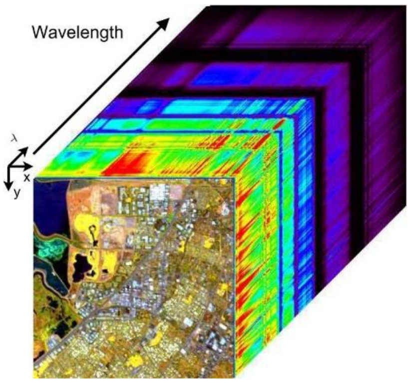

Fig.1. Data Cube of HSI

Following this method, the HSI will contain only the spectral information. There is high correlation between a pixel's neighborhood and itself, where the characteristics match and is similar.

Leng et al. [15] are the ones who directed and gave an idea to extract a spectral cube of varied spatial strategies, which individually means using single-pixel, 4-neighbor pixels and 8-neighbor pixels. Classification of the pixel located at the centre is needed. Results from experiment show that 8-neighbor pixels are the best, followed by 4-neighbor pixels and single pixel.

2.1.2 Preprocessing Required for SVM Techniques:

A hypercube has been created to observe the data and locate what categories are present. RGB false color is shown in the image in which we use few bands. Masking of image is done to classify the different classes in the dataset. Some of the steps that are used for pre-processing of data in SVM are:

- *Data Cleaning*: The data which have been collected might be incomplete sometimes. It is used to fill the missing values, resolve data inconsistencies, and smoothen data having noises and remove outliners.
- *Data Integration*: Combining data with various representation from many databases, removing redundant/similar data from many sources.
- *Data Transformation*: Before starting SVM classifier it is required to balance the training data. It can be done either by adding some new foreground segments or removing half of background segments. If we still have an outsized dataset, we will choose the latter choice to equalize the number of feature vectors for both clusters.
- *Data Reduction*: Obtaining reduced representation of data which produce similar analytical results. One well known technique is the Principal Component Analysis (PCA) method. It performs transformations on a dataset to define new vectors with the highest variance of our features. A newly created uncorrelated, orthogonal basis set for the data is generated from the vectors. As the vectors (principal components) are arranged from the highest variance to the lowest, most of the lower ones are rejected, and the resultant dimensions are thus reduced.

2.2 FEATURE REDUCTION

The feature reduction required for CNN and SVM techniques as classifiers is presented in this section.

2.2.1 Feature Reduction Required for CNN Techniques:

Reducing the dimension is an important step in pre-processing stage of Hyperspectral Image Classification. Hence, pooling is used to reduce the spatial dimension to be fed into the convolutional neural network. There are many types of pooling such as max pooling, mean pooling, and average pooling that have their individual benefits. Pooling is applied to the hyper data cube containing spectral dimension. To reduce the network parameters, CNN comprises of a special architecture with local connection and having shared weights.

Local connection can reduce the number of parameters present in a network but it is not be enough. If all the neurons in every feature map consists of same weights and biases named as shared weights, the parameters present in a network will be further reduced. CNN is composed of many layers named as input layer, convolutional layer, pooling layers and fully connected layer. Convolutional input data contains different feature maps with various convolutional kernels, and a feature is represented by each feature map.

The feature map's resolution can be reduced by subsampling thus offering invariance and one of the common method of subsampling is pooling such as max pooling, mean pooling etc. The Fig.2 shows the common method for pooling such as Max-pooling and mean pooling. With many layers of convolution and pooling a 1D vector can be obtained from the feature map and it gets connected fully to the front layer. In the last layer a classifier is selected which can be softmax or any other classifier. Back propagation algorithm is used to train the whole CNN network

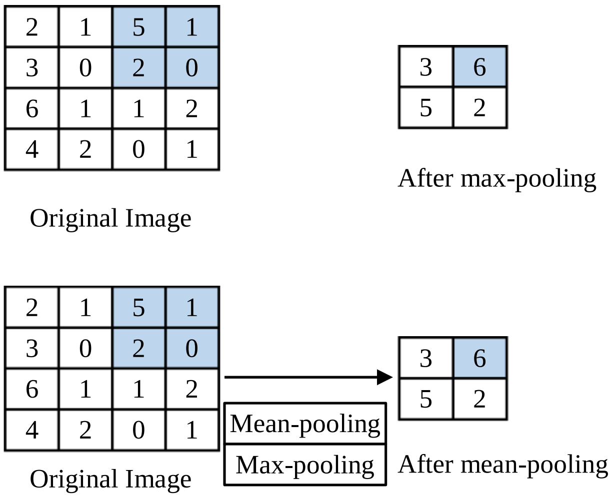

Fig.2. Max-pooling and mean pooling

The feature map's resolution can be reduced by subsampling thus offering invariance, One of the common methods of subsampling is pooling such as max pooling, mean pooling, etc. With many layers of convolution and pooling, a 1D vector can be obtained from the feature map and it gets connected fully to the front layer. In the last layer a classifier is selected which can be softmax or any other classifier. Back propagation algorithm is used to train the whole CNN network.

2.2.2 Feature Reduction Required for SVM Techniques:

High dimensionality, which is known as Hughes Phenomenon is a hindrance for classifying hyperspectral images. In Hughes phenomenon, as the dimensionality of the data increases with fixed number of training samples and labels, the accuracy of the supervised classification decreases. This can be overcome by feature reduction, which comprises of feature selection and extraction. Feature extraction includes Principle Component Analysis (PCA), Linear Discriminate Analysis (LDA), etc. Among all the feature extraction methods PCA will store most of the spectral information of a hyperspectral data in small principal components. The highest class separability is achieved by finding the best spectral bands in hyperspectral images which is done by feature selection techniques.

PCA is used to reduce the dimension of hyperspectral images, which have correlated bands that represent similar information. The bands are observed and changed to reduce the correlation. Eigenvalue decomposition of the covariance matrix is the basic principal used in PCA. PCA images have K bands at first and these bands are known as pixel vector. The first band of PCA images has the highest contrast or variance and the last band has the lowest contrast or variance. Therefore, the first few bands of the hyperspectral images in PCA contain majority of information.

After using PCA, we use Extended Morphological Profiles (EMP). It is known as opening and closing and is based on mathematical morphological operators. Bright and dark structures of an image are isolated using EMP. Small objects from foreground are removed by opening and placing them in the background and small holes in the foreground are removed by closing and converting small islands from the background into foreground. The objects can be thus restored completely by the opening and the closing processes using reconstruction method. In this fashion the dark and bright structures in each morphological profile can be reduced through the opening and closing procedure using reconstruction to obtain a more homogenous version of the original image. At each iteration, the Structuring Element (SE) is increased to capture further spatial information.

2.3 CLASSIFICATION TECHNIQUES USING CNN

CNN algorithm is used to classify the Hyperspectral image.

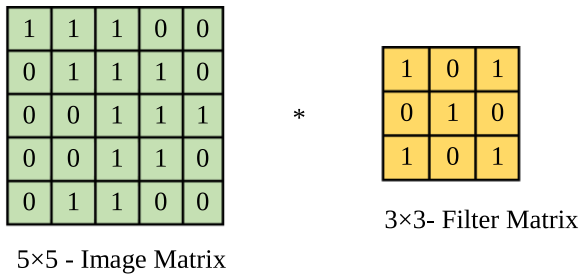

Fig.3. Convolving a Matrix

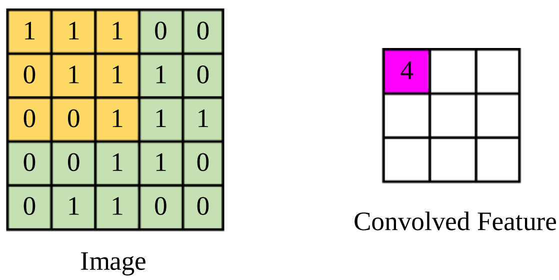

Fig.4. Feature Map

In CNN models, to test and train the data, the input images go through a series of steps which include convolution with appropriate filters, pooling, fully connected layers, and applying softmax functions. Input image and the classification is done based on the values given.

Convolution Layer is the first layer that extracts features from an input image. This performs the mathematical operation that takes on two inputs which are image matrix and a filter or kernel. Convolution of an image is used to perform edge detection, and apply different filters to blur and sharpen. The output of image matrix and filter matrix is called as Feature map. Sometimes, the filter does not perfectly fit the input image. In such cases, we have two options:

1. Pad the picture with zeros (zero-padding) so that it fits, and

2. Drop the part of the image where the filter did not fit. This is called valid padding which keeps only a valid part of the image.

- *Non Linearity (ReLU)*: ReLU stands for Rectified Linear Unit for a non-linear operation. The output is $f(x) = \max(0, x)$. Why ReLU is important: ReLU's purpose is to introduce non-linearity in our ConvNet.

Pooling the layers section would scale back the number of parameters when the images are overlarge. Spatial pooling, also called sub sampling or down sampling, reduces the dimensionality of each map but retains important information. Spatial Pooling is of different types, such as Max Pooling, Mean Pooling, and Sum Pooling. Max pooling takes the most important element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements within the feature map is called as sum pooling.

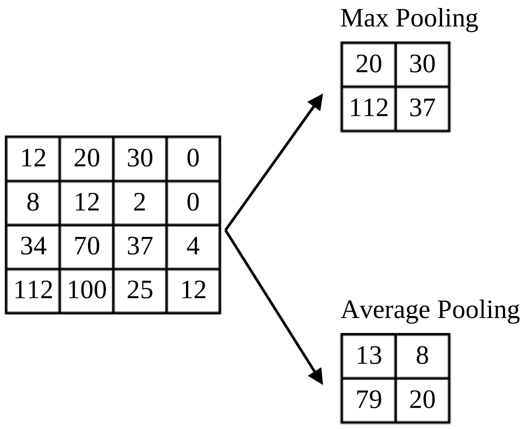

Fig.5. Pooling Fully Connected Layer

The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like a neural network.

The feature map matrix is converted as vector (x_1, x_2, x_3,…). With the fully connected layers, we combined these features together to create a model. Finally, we have an activation function such as softmax or sigmoid to classify the hyperspectral images.

CNN classifier parameter space has been introduced. A random value between -0.05 and 0.05 has been initialized to all the trainable parameters of CNN. The training process is followed by two steps: Forward Propagation and Back Propagation. The current parameters are computed with the actual classification result in forward propagation. The back propagation is used to update the trainable parameters so as to make the discrepancy between the actual classification output and the desired classification output as small as possible.

- *Forward Propagation*: The (L+1) layer of CNN network consists of n_1 input units in input layer, output layer has n_5 output units, and layers C_2, M_3, and F_4 consists of several hidden units. The proposed CNN classifier is a multiclass classifier and as a result the output of layer F_3 is fed to n_5 softmax function which produces a distribution over the n_5 class labels. The final probability of all the classes in the current iteration is given by the output vector $y = x_L+1$ in the output layer.
- *Back Propagation*: The gradient descent method is used to update the trainable parameters in back propagation stage. It is done by minimizing a cost function and computing the partial derivative of the cost function with respect to each trainable parameter. The actual output is closer to the desired output is indicated by the smaller value of the return of cost function because the number of training iteration increases. Finally the trained CNN is ready for HSI classification.

2.3.1 Classification Techniques using SVM Techniques:

Support vector networks are supervised learning techniques associated with learning algorithms that examine data used for classification and regression analysis. The goal of SVM algorithm is to detect a hyper plane in N dimensional space.

A Method is used for the HSI classification with SVM and guided filter. First, the spatial features of HSI are extracted by the guided filter that are obtained by principal component analysis (PCA) method from the original HSI. Following spatial features are classified by SVM. Then the guided filter are employed at final stage to optimize the classification.

- *Extraction of Spatial Features by Guided Filter*: A guidance image is obtained by PCA and first three principal components are taken as color guidance image. Given a dataset $D = \{d_1, d_2, \ldots, d_s\}$, PCA is adopted to get the result. So, the guidance image is $G = [g_1, g_2, g_3]$. Based on the input image d_1 and guidance image G, we can get the output u_1 by filtering. In the same way, we can yield all the u_i which constructs a new hyper spectral image $U = \{u_1, u_2, \ldots, u_S\}$.
- *Classifying HSI by SVM*: After obtaining the image $U = \{u_1, u_2, \ldots, u_S\}$ by the guided filter, we can rewrite it as $V = \{v_1, v_2, \cdots, v_N\}$, where $v_n = \{v_{n,1}, v_{n,2}, ..., v_{n,S}\}$ is the spectral feature vector.
- *Optimizing Classification Map*: First, we convert the classification map C into a probability map $P = \{p_1, p_2, ..., p_n\}$, where $p_{i,n}$ is the initial probability with a value of 0 or 1. And n denotes the number of categories to classify. If a pixel i belongs to the n^{th} class, p_i, n is set to 1. Other else, p_i, n is set to 0.

3. IMPLEMENTATION DETAILS

For the proposed work, the input is considered with the data sets available in the website. The experimentation uses two publicly available hyperspectral image datasets, namely Indian Pines and University of Pavia. The experimentation are conducted using python programming.

3.1 THE INDIAN PINES (IP) DATASET

The Indian Pines dataset had been gathered by AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor over the Indian Pines test situated in North-western Indiana in 1992. IP has images with 145×145 spatial dimension and 224 spectral bands in the wavelength ranging from 400 to 2500 nm, out of which 24 spectral bands from the area absorbed by water have been discarded. The available ground truth is split into 16 classes of vegetation. The IP scene contains two-thirds agriculture, and one-third forest or other natural perennial vegetation. There are two major dual lane highways, a rail line, some rarity housing, other built structures, and smaller roads. The ground truth available is designated into sixteen classes of which not all could be considered mutually exclusive.

3.2 UNIVERSITY OF PAVIA (UP) DATASET

The University of Pavia dataset had been acquired by the ROSIS (Reflective Optics System Imaging Spectrometer) sensor during a flight campaign over Pavia situated in Northern Italy in 2001. The University of Pavia (UP) dataset consists of 610 × 340 spatial dimension pixels with 103 spectral bands and the ground truth is split into 9 urban land-cover classes.

Table.1. Details about the Dataset

	Dataset 1	Dataset 2
Dataset	Indian Pines	University of Pavia
Source	AVIRIS sensor	ROSIS sensor
Spatial Dimension	145*145	610*340
Spectral bands	224	103
Class	16 classes of vegetation	9 urban land cover classes

Table.2. Number of training and testing samples in Indian Pines

Number	Class	Training	Test
1	Corn-notill	200	1228
2	Corn-mintill	200	630
3	Grass-pasture	200	283
4	Hy-windrowed	200	278
5	Soybean-notill	200	772
6	Soybean-mintill	200	2255
7	Soybean-clean	200	393
8	Woods	200	1065
Total		**1600**	**6904**

Table.3. Number of training and testing samples in university of Pavia

Number	Class	Training	Test
1	Asphalt	200	6431
2	Meadows	200	18449
3	Gravel	200	1899
4	Trees	200	2864
5	Sheets	200	1145
6	Bare soil	200	4829
7	Bitumen	200	1130
8	Bricks	200	3482
9	Shadows	200	747
Total		**1800**	**40976**

3.2.1 Ground Truth Images:

Ground truth refers to information collected on location. Ground truth allows image data to be associated with real features and materials on the ground. Ground truth also helps with atmospheric correction and is shown in Fig.6.

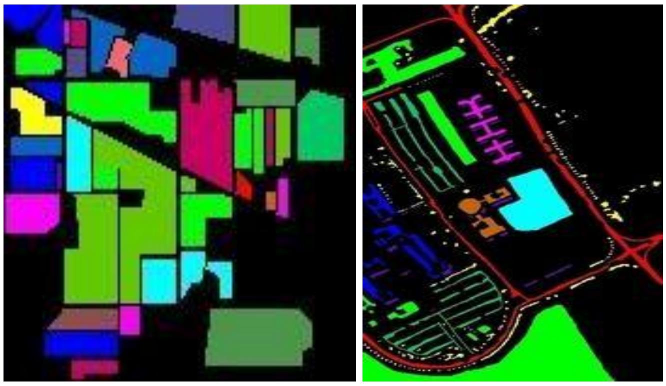

Fig.6. Ground Truth Image of Indian Pines and University of Pavia

3.2.2 False Color Images:

The wavelengths that are not visible through human eye can be seen through false color composite. The interpretability of the data can be enhanced by using bands such as near infrared which increases spectral separation and is shown in Fig.7.

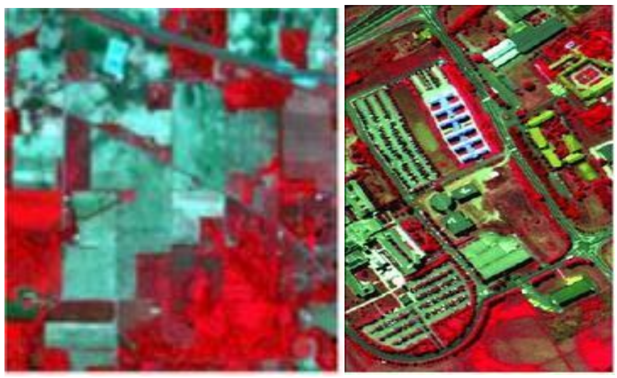

Fig.7. False color image of Indian Pines and University of Pavia

4. RESULTS AND DISCUSSIONS

This section presents the results obtained from different sections with detailed discussion.

4.1 PRE-PROCESSING AND FEATURE REDUCTION

In CNN, the pre-processing and feature reduction is done using the three layers, namely convolutional layer, non-linearity layer, and pooling layer. A deep CNN can hierarchically extract the features of inputs, which tend to be invariant and robust. The hyperspectral pixels constitute mixed-land cover classes producing interclass similarity and high interclass variability. As a result PCA is applied to reduce spectral redundancy. The PCA reduces the spectral bands while maintaining spatial dimensions. The training and testing datasets have been divided into 30% and 70% respectively for classification. The network is trained with 100 epochs with no batch normalization and data augmentation.

In SVM, PCA and EMP are used. The amount of bands in the hypercube are reduced by applying the Principal Components Analysis (PCA). The use of Principal Components (PC) gives us at least 96% of the total explained variance in the data to use as our main set of features.

It can be seen that main visual information present in the whole hypercube could be captured by using only 4 components. Hence, these 4 images are considered as main set of features to apply the EMP method. Bright and dark structures of an image are isolated using EMP. A more homogenous version of the original image can be obtained by reducing dark and bright structures by each morphological profiles through opening and closing using reconstruction. At each iteration, the Structuring Element (SE) is increased to capture further spatial information. A structuring element with disk shape of 4 pixels size (diameter), an increment step of 2 pixels and 4 openings and closings has been used. It will result in 36 features in EMP.

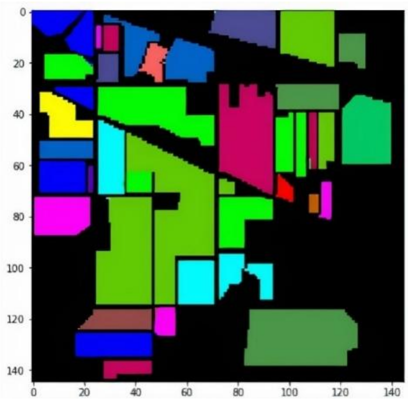

Fig.8. Classified image of Indian Pines dataset using CNN

4.2 CONFUSION MATRIX

The synopsis of prediction results on a classified hassle is given through confusion matrix. The count values offer the summarized result of the wide variety of correct and wrong predictions that are split into their respective class. The confusion matrix suggests various ways through which classification version is pressured, while also making predictions. It provides insights regarding the errors by means of a classifier, and, more significantly, regarding the forms of errors that are being created.

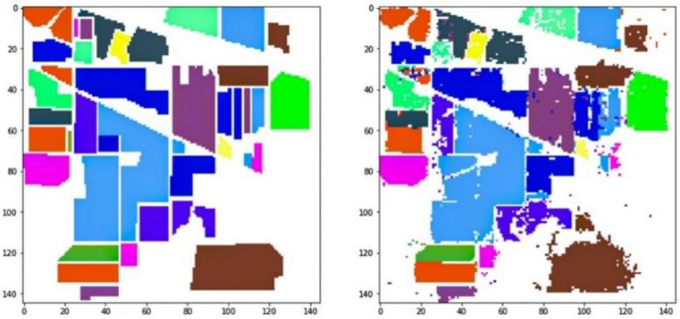

Fig.9. Classified image of Indian Pines dataset using SVM

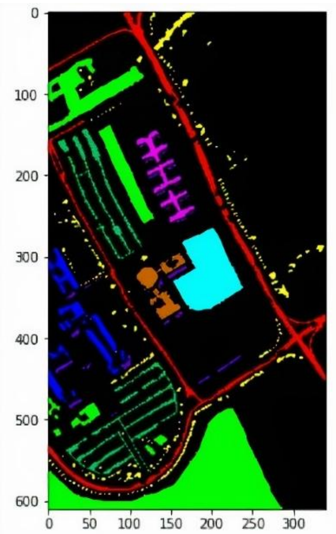

Fig.10. Classified image of University of Pavia using CNN

Fig.11. Colors marks to different classes in an image

In CNN the training and testing datasets have been divided into 30% and 70% respectively for classification. In SVM the training and testing datasets have been divided into 25% and 75% respectively for classification.

Table.4. Accuracy results of two classifiers with respective datasets

Dataset	Accuracy for CNN (%)			Accuracy for SVM (%)		
	OA	Kappa	AA	OA	Kappa	AA
Indian pines	99.75	99.71	99.63	88.7	88.7	88
University of Pavia	99.89	99.8	99.9			

HSI classification accuracy has to be analyzed with the help of overall accuracy (OA), Kappa coefficient and Average Accuracy (AA). Overall accuracy takes the total samples and tell us about the number of samples correctly classified. Average accuracy takes the average of the accuracies of all the classes present. The kappa is used to control only those instances that may have been correctly classified by chance. We have used 100 for training and validation. This can be calculated using both the observed (total) accuracy and the random accuracy. Kappa can be calculated as:

Kappa = (*total accuracy* – *random accuracy*)/(1-*random accuracy*)

5. CONCLUSIONS

In this proposed work, our focus is to investigate the classification of Hyperspectral image using CNN and SVM models. A hybrid CNN model is introduced that basically combines spatio-spectral and spectral complementary information respectively. The above model is more efficient than the 3D CNN model. On the other hand, SVM algorithm is proposed to classify various types of groups in the dataset. This combines the SVM and guided filter, and also adopts two spectral and spatial fusion methods. The model discussed above gives accuracy as 93% for Indian pines, and 92% for Pavia University in CNN, while the accuracy is given as 88% for Indian Pines in SVM.

From the work done so far, we can conclude that spatial information can be extracted efficiently in HSI by using guided filter and SVM algorithm with double filtrations in an easy and effective way for the classification of hyperspectral image.

6. FUTURE SCOPE FOR IMPROVEMENT

In the future, a network architecture called Siamese Network could be used, as it has proved to be robust in situations with small number of training samples per category.

Unsupervised learning can be employed to train CNNs to significantly reduce the requirement of labeled samples. Deep learning, especially deep CNNs, should have great potentiality for HSI classification in the future.

Hyper spectral imaging (HSI) system could be developed for various food processes including cooking, drying, chilling, freezing & storage, and salt curing.

The characterization of plant disease symptoms by hyperspectral imaging is often limited as it cannot investigate early and still invisible symptoms. To overcome this drawback, automatic timely tracing of the symptom position on the leaf back could be a promising approach.

REFERENCES

[1] Javier Plaza, Antonio Plaza, Roza Perez and Pablo Martinez, “Parallel Classification of Hyperspetral Images using Neural Networks”, *Computational Intelligence for Remote Sensing*, Vol. 133, pp. 193-216, 2018.

[2] Alberto Signoroni, Mattia Sarvadia, Annalisa Baronio and Serigo Bennini, “Deep Learning Meets Hyperspectral Image Analysis: A Multidisciplinary Review”, *Journal of Imaging*, Vol. 5, No. 5, pp. 1-32, 2019.

[3] Amol D. Vibhute, Karbhari V. Kale, Rajesh K.Dhumal, Ajay D. Nagne and Suresh C. Mehrotra, “Identification, Classification and Mapping of Surface Soil Types using Hyperspetral Remote Sensing Datasets”, *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, Vol. 3, No. 1, pp. 921-932, 2018.

[4] M.K. Arora and S. Mathur, “Multi-Source Classification using Artificial Neural Network in Rugged Terrain”, *Geocarto International*, Vol. 16, No. 3, pp. 37-44, 2001.

[5] A. Goetz, G. Vane, J.E. Solomon and B. Rock, “Imaging Spectrometry for Earth Remote Sensing”, *Science*, Vol. 228, pp. 1147-1153, 1985.

[6] P. NilaRekha, R Gangadharan, S. M. Pillai, G. Ramnathan and A. Panigrahi, “Hyperspectral Image Processing to Detect the Soil Salinity”, *Proceedings of 4th IEEE International Conference on Advanced Computing*, pp. 1-5, 2012.

[7] S. Mei, J. Ji, J. Hou, X. Li and Q. Du, “Learning Sensor-Specific Spatial-Spectral Features of Hyperspectral Images via Convolutional Neural Networks”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 55, No. 8, pp. 4520-4533, 2017.

[8] M.T. Eismann, “*Hyperspectral Remote Sensing*”, SPIE Press, 2012.

[9] Yanan Luo, Jie Zou, Chengfei Yao, Tao Li and Ganga Bai, “HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image”, *Proceedings of 4th IEEE International Conference on Pattern Recognition*, pp. 1-7, 2018.

[10] Mehmoodul Hassan, Saleem Ullah, Muhmmad Jaleed Khan and Khurram Khurshid, “Comparative Analysis of SVM, ANN and CNN for Classifying Vegetation Species using Hyperspectral Thermal Infrared Data”, *Proceedings of International Workshop on ISPRS Geospatial Week*, pp. 1-7, 2019.

[11] D.P. Shrestha, D.E. Margate, F. Van Der Meer and H.V. Anh, “Analysis and Classification of Hyperspectral Data for Land Degradation: An Application in Southern Spain”, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 7, No. 2, pp. 85-96, 2005.

[12] Amol D. Vibhute and K.V. Kale, “Soil Type Classification and Mapping using Remote Sensing Data”, *Proceedings of International Conference on Man and Machine Interfacing*, pp. 17-21, 2015.

[13] Sushma Suresh, H.S. Prashantha H.S and S Sandya, “Satellite Image Classification using Clustering Algorithms with Edge Detection Operators”, *International Journal on Recent and Innovation Trends in Computing and Communication*, Vol. 3, No. 10, pp. 1-12, 2015.

[14] K.S. Gunasheela and H.S. Prasantha, “Satellite Image Compression-Detailed Survey of the Algorithms”, *Proceedings of International Conference on Cognition and Recognition*, pp. 187-198, 2017.

[15] Jiabing Leng, Tao Li and Gang Bai, “Cube-CNN-SVM: A Novel Hyperspectral Image Classification Method”, *Proceedings of IEEE International Conference on Tools with Artificial Intelligence*, pp. 1027-1034, 2016.