Review On Hyperspectral Image Classification Using Deep Learning Method

[1] Dhrisya Krishna, [2] Keerthi Krishnan K
[1]M.Tech Student, [2]Assistant Professor
[1]Department of Electronics and Communication Engineering,
[1]NSS College of Engineering, APJ Abdul Kalam Technological University, Palakkad, India

Abstract : Deep learning is a booming technology for analysing remotely sensed imagery. There are different approaches for classifying the hyperspectral image based on supervised, unsupervised, semi-supervised classifications of deep learning methods. A hyperspectral image is emerging as an effective tool for extracting complex, invariant and important features due to its capability of covering a large surface of the earth and the characteristic of the hyperspectral image is also different and useful while compared to traditional imagery which is very much useful for various applications. Deep learning methods for classifying hyperspectral image are CNN, Deep Belief Network, Recurrent Neural Network, Restricted Boltzmann Machine, Generative Adversarial Neural Network, Auto encoders. Extraction of features in spectral, spatial, joint spectral and spatial domain using supervised, unsupervised and semi-supervised are compared in which unsupervised method work without labeled data, supervised technique processed with labeled data and semi-supervised method work with limited training samples. Comparison of each method had done in terms of its classification accuracy for different datasets.

***Index Terms* - CNN, DBN, RBM, RNN, GAN, Hyperspectral image, Auto encoder**

I. INTRODUCTION

A hyperspectral image is emerging as an effective tool to extract more robust, complex and invariant features for material mapping, target detection, material identification due to its capability of covering the large surface of the earth. The human eye is able to see the limited part of the electromagnetic spectrum but there are lots of information present in other ranges of the electromagnetic spectrum. HSI/ multispectral imagery able to capture the information ranging from visible to near-infrared region of the electromagnetic spectrum. Hyperspectral sensors measures bands at 10 to 20nm intervals [32], [46]. The main difference between HSI and the multispectral image is that HSI collects an image in which each pixel has contiguous bands of spectra whereas multispectral image contains discrete bands of information. Comparison of the hyperspectral image and the multispectral image is shown in Fig.1.

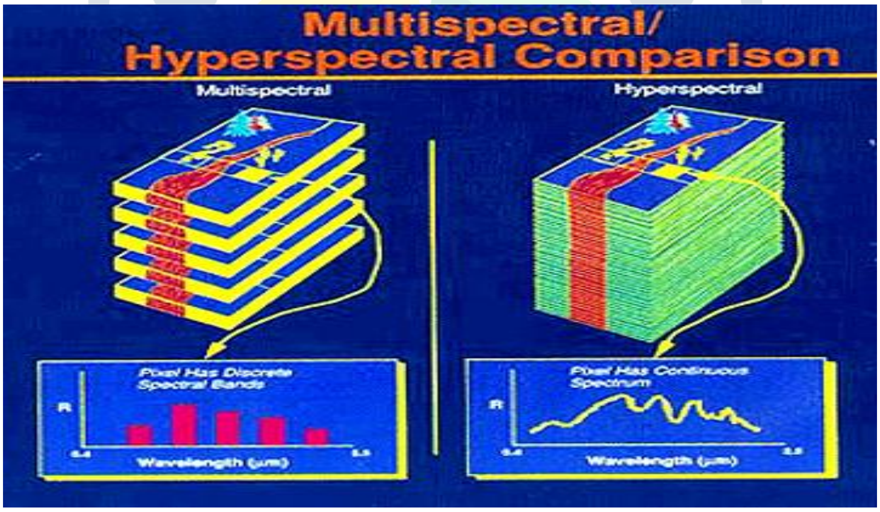


Fig.1 Comparison of hyperspectral image and multispectral image

The continuous and large number of spectral channel provides opportunity for detailed analysis of land cover materials, image classification, target detection etc. Traditional imagery (RGB sensors) lacks spectral range and precision to profile materials [15]. Difference between Hyperspectral and RGB image is shown in Fig.2.

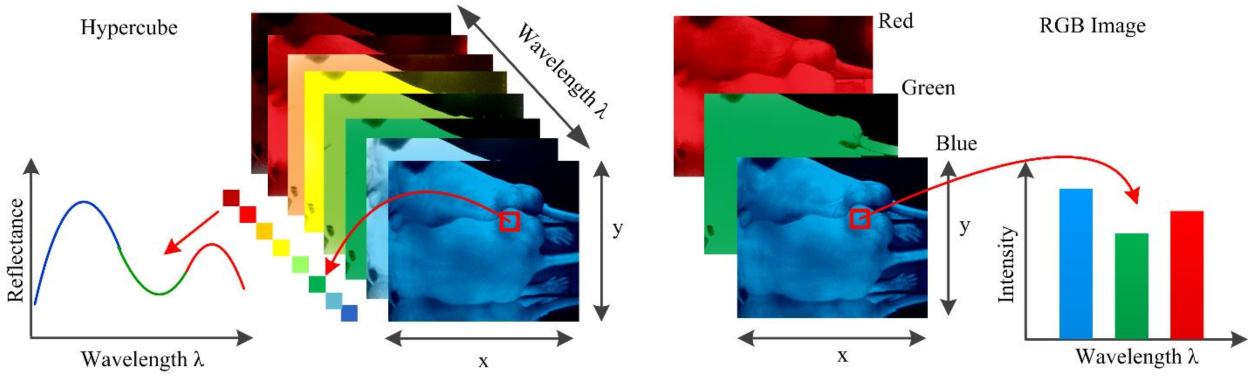


Fig.2 Difference between Hyperspectral and RGB image

Characteristics of HSI

A. <u>Spatial resolution</u>
It is the smallest discernable details in an image and it gives the clarity of image, not the number of pixels in an image. Spatial resolution is inversely proportional to the patch size. Smaller the patch size, higher the details can be interpreted from the observed scene.

B. <u>Spectral resolution</u>
It is the number of spectral bands and range of EM spectrum measured by the sensor. HSI sensors acquire extremely narrow spectral bands in mid infrared, near infrared and visible segments of EM spectrum. Spectrum of single pixel in HSI gives more information about the surface of materials than a normal image.

C. <u>Temporal resolution</u>
In hyperspectral remote sensing, the temporal resolution depends on the orbital characteristics of imaging sensors. It is the time needed by the sensor platform to revisit and obtain data from the exact same location. It is represented in days.

Information provided by HSI

Hyperspectral remote sensing mainly benefits for distinguishing between spectrally similar materials as shown in Fig.3

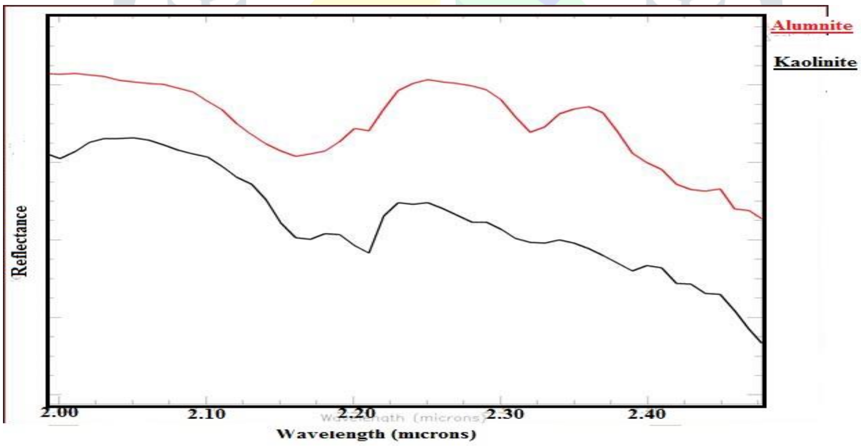


Fig.3 Pixel spectra from an AVIRIS hyperspectral image. The red spectrum is from a pixel filled with the mineral Aluminite, and the white spectrum is from a pixel filled with the mineral Kaolinite.

Fig.3 shows the plots of spectral of kaolinite and aluminite as measured by NASA's hyperspectral AVIRIS sensor. Both mineral have absorption feature of 2.2um. But from the figure it clearly shows that Kaolinite displays a double dip in the absorption feature while aluminite shows only a single dip whereas multispectral Landsat ETM sensor measures this entire spectral region with one channel and doesn't provide sufficient details to differentiate between kaolinite and aluminite using double dip or single dip. Several methods have been developed for classification of hyperspectral images from those that use only spatial or spectral information to those that combine both kinds of data [4][7].This includes unsupervised techniques[4][5][6] but supervised classifiers are preferred due to their capacity to prove high classification accuracies [1][2][3]. The problem caused due to limitations of training samples in hyperspectral remote sensing is overcome by introducing semi supervised methods which only requires limited training samples compared to other techniques [7]. In this paper, different approaches for classifying hyperspectral image using various deep learning method is mentioned and comparison of each method is discussed through their classification accuracy. In section II and section III, applications and issues while working with hyperspectral remote sensing is discussed. In section IV, deep learning methods like unsupervised, supervised and semi supervised techniques are discussed.

II. APPLICATIONS OF HYPERSPECTRAL REMOTE SENSING

Hyperspectral remote sensing is used in wide variety of applications which usually have one of the following objectives:

- Target detection
- Material mapping
- Material identification
- Mapping details of surface properties
- Agriculture

For particular crops and in particular climates, hyperspectral remote sensing use is increasing for monitoring the development and health of crops. Traditionally, crop monitoring for disease, water stress, nutrients and insect attack was carried out by manual visual inspection from the ground. But this method, sometimes failed to detect the disease at the early stages [32] [34]. High water level stresses leads to changes in photosynthetic pigments [35] [36]. These changes results to yellowish tint in crops due to increasing reflectance of red wavelength. HSI sensors can detect these changes at earlier stages whereas human eyes can't.

1. Water Resource and Flood Management

Hyperspectral remote sensing technology has found enormous applications in water resource management. Accurate estimates of water resource parameters are possible by analysing spatial, spectral and temporal variations in water bodies. Hyperspectral remote sensing provides conclusive and certain information about water quality parameters which contain biochemical, hydro-physical and biological attributes. It enable us to measure chlorophyll, turbidity and chemical oxygen demands and phosphorous in water resource. It also helped in detailed understanding of vegetation characteristics of ecosystem.

2. Defence and Homeland security

HSI can detect strategic deployments in unpopulated areas such as forests, deserts and mountains where targets such as military vehicles and mines are distinct from the background even if camouflaged.

3. Mineralogy

Geographical specimens (drill centres) can be quickly mapped for almost all minerals of business enthusiasm with hyperspectral imaging. Combination of SWIR and LWIR unearthly imaging is standard for the recognition of minerals in the feldspar, silica, calcite, garnet, as these minerals have their most unique and strongest ghastly signature in the LWIR districts.

III. HSI CLASSIFICATION ISSUES

There are some problems faced in the classifications of hyperspectral data:

1. Curse of dimensionality due to its high number of spectral channels;
2. A limited number of labelled training samples;
3. Large spatial variability of spectral signature

For classification of HSI, the large number of bands could aid the classification of features, but adding a large number of bands to a classifier will deteriorate the accuracy of classification [15]. To increase the accuracy of classification, the number of training samples is required. With insufficient training sets, the estimation of the statistical parameter decreases. But a large number of training samples leads to an increase of dimensionality called the curse of dimensionality.

The high dimension has the following properties:

- The volume of hypercube concentrate in corner and volume of hyper sphere concentrate in outer shell.

This implies that with limited training data, much of hyperspectral data space is empty. Two solutions exist:

1. Providing large set of training data.
2. Reduce the dimensionality of data by extracting relevant features from HSI.

   Former solution results to increasing the size of training data exponentially while increasing dimension, thus it is impractical because HSI contains 100's of spectral bands. Later solution must be considerable because it reduces the data space dimension by extracting relevant features only that allows for separation of classes without losing the original information.

IV. DEEP LEARNING APPROACH

Feature extraction using deep learning is a new approach that has achieved remarkable results. The high dimensionality of HSI and the limited number of labelled training samples makes deep learning an appealing approach for analysing hyperspectral data. Different deep learning method for hyperspectral image classification is discussed in the Table 1.

Table 1. Comparison of different deep learning approach for hyperspectral image classification

Deep Learning Method		Author	Features	Layers	Dataset	Overall Accuracy (%)
Supervised Learning (discriminative deep network)	1. Deep Stacking Network (DSN)	[1] Mingyi He, Xiaohui Li, Yifan Zhang, Weigang Wang	• Simplicity • Batch mode learning • Nonlinear activation function is employed	Number of Hidden layers = 120	KSC	96.59
	2. Convolutional Neural Network (CNN)	[2] Konstantinos Makantasis, Konstantinos Karantzalos, Anastasios Doulamis, Nikolaos Doulamis	• Encode pixel's spectral and spatial information. • Use Multilayer Perceptron for classification task • Hierarchically constructs high level features. • For training, standard back propagation algorithm is employed.	1. First layer = Convolutional layer of 3x3. 2. Second layer= convolutional layer of 3 x principal components. 3. Hidden units in MLP= 6xprincipal components.	Pavia Centre	99.91
					Salinas	99.53
					Indian Pines	98.88
	3. Recurrent Neural Network (RNN)	[3] Lichao Mou, Pedram Ghamisi, and Xiao Xiang Zhu	• Analyse hyperspectral pixel as sequential data • Employed parametric rectified tanh (PRetanh) as activation function.	1. Single recurrent layer with GRUs of size 64 with sigmoid activation function. 2. PRetanh activation functions for hidden. Representation. 3. Output layer uses softmax activation.	Indian Pines	88.63
Unsupervised (Generative learning)	1.Stacked Auto encoder	[4] Yushi Chen, Zhouhan Lin, Xing Zhao,Gang Wang and Yanfeng Gu	• Applied auto encoder • Extracting spatial dominated information for classification • Proposed joint spectral-spatial information.	• Single layer AE with 100 hidden units	KSC	96.73
					Pavia	95.14
	2. Deep Belief Network	[5] Yushi Chen, Xing Zhao and Xiuping Jia	• Combines spectral-spatial feature extraction • Hybrid of PCA, hierarchical learning-based FE and logistic regression	• Single layer RBMs with different number of hidden units (10, 50, 100, 150, 200). . • Spectral classification: number of hidden unit for each hidden	Indian Pines	
					1. Spectral feature	90.81
					2.Spatial feature	93.20
					3.Spectral-Spatial feature	95.95

				layers is 30. • Spatial classification: 50 hidden units • Spectral-spatial classification: 60 and 50 hidden units for Indian pines and Pavia dataset.	Pavia 1.Spectral feature 2.Spatial feature 3.Spectral-spatial feature	 96.42 98.62 99.05
	3. Restricted Boltzmann Machine	[6] Midhun E M , Sarath R Nair ,Nidhin Prabhakar T V , Sachin Kumar S	• Capture distribution among pixel at hidden level • Band-by-band nonlinear diffusion introduced.	*3 layer of RBM* • First layer: vector (pixel) of size 200. • Hidden layer 1 of size 60units is the input to second layer of RBM. • Output layer: 200 units.	Indian Pines	79.34
Hybrid deep networks (Semi-supervised methods)	Generative Adversarial Networks	[7] Lin Zhu, Yushi Chen , Pedram Ghamisi and Jón Atli Benediktsson [8] Ying Zhan , Dan Hu, Yuntao Wang, and Xianchuan Yu	• A CNN designed to discriminate the inputs and another CNN used to generate fake inputs. • Improves generalization capability.	• Prinipal components= 3 for spectral and spatial features in 3D GAN and 10 principal components for spectral features in 1D GAN. • Convolutional layer=5. • Sigmoid and softmax classifier used.	*Spectral - Spatial features* Salinas KSC Indian Pines	*3D GAN* 93.02 96.89 89.09

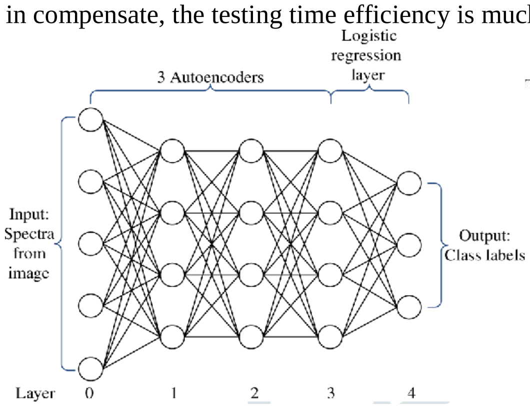
ANN (Deep Neural Network) techniques have been known since the 1980s and Auto Encoder is ANN trained to reconstruct its input as its outputs. CNN (Restricted Boltzmann Machines) is another key technique in DL. A CNN is an ANN which is composed of multiple pair of layers – a convolutional layer and a pooling layer. The main benefit of this learning is that feature maps used in the classifications are learnt from data without any handcrafted feature extraction.

1. UNSUPERVISED LEARNING

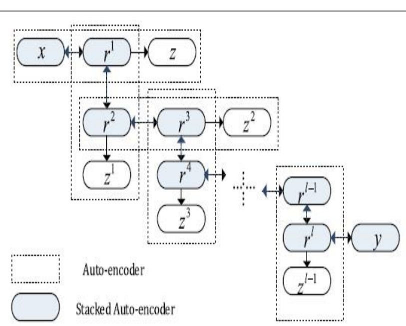
Unsupervised deep networks work without labelled classes, looking for patterns between pixels through capturing high order correlation of data [4].

A. *Deep Auto-Encoders*

An auto encoder (AE) is one of the deep architecture based models, to learn deep features of hyperspectral data in an unsupervised manner. In paper [4], a single layer AE and multi-layer stacked AE (SAE) to learn shallow and deep features of hyperspectral data. A stacked auto encoder (SAE) is a deep network model consisting of multiple layers of auto encoders (AEs) in which the output of one layer is wired to the input of the successive layer as shown in Fig.4 (b). An AE is consists of two parts: an encoder and a decoder. The encoder takes an input vector x and maps it to a hidden representation and get the code r by the parameter W and bias b, while the decoder reconstruct the input from the hidden representation. No class labels are required for the training since AE only map the input onto itself (via hidden representation) and independent of any class labels. Multilayer AEs, SAEs, are constructed using a greedy layer – wise strategy. In the cited paper, the author integrated the feature extraction stage with the classification stage. The author had configured the multiple layers of stacked auto-encoders with final logistic regression layer as shown in Fig.4(a). It is shown that AE-extracted features are useful for classification, and it helps to increase the accuracy of SVM and logistic regression while obtaining the highest accuracy when compared with other feature extraction techniques like PCA, KPCA and NMF. The cited paper had introduced the spatial dominated feature based classifications; both SAE-LR and SVM yield higher accuracy than traditional spectral information based methods. The first AE maps inputs in the zeroth layer feature in the first layer. After training the first layer AE as shown in Fig.4 (b), subsequent layers of AEs are trained through the output of the previous layer. The classification scheme is shown in Fig. 5(a) has five layers: one input layer, three hidden layers of AEs and its output layer of logistic regression. The layers shown in Fig. 5(a) are for classifying HSI with spectral features. To extract the spatial features as shown in Fig. 5(b), PCA is introduced in the first layer to reduce the data dimension. In the second layer, the neighbourhood region of the pixel is extracted from the condensed data which has only principle components in its spectral dimensions. Then flattening of data is applied in the third layer which is in 1-D vector and feeds it into a SAE. The drawback of SAE-LR is its training time, but in compensate, the testing time efficiency is much faster than other methods like SVM or KNN.

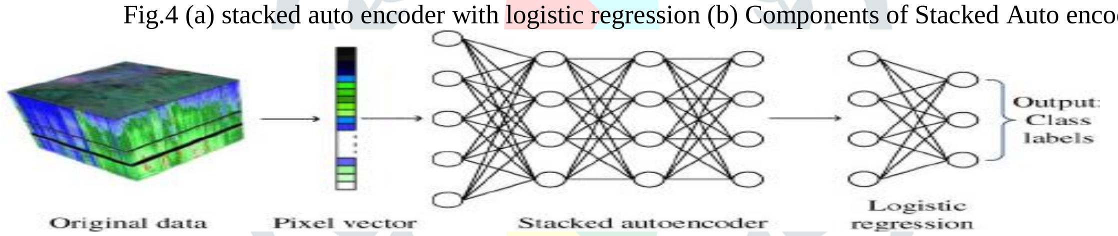


(a)

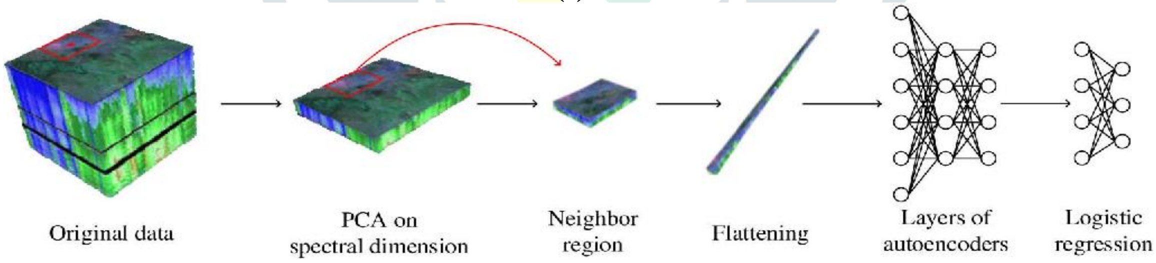


(b)

Fig.4 (a) stacked auto encoder with logistic regression (b) Components of Stacked Auto encoder



(a)



(b)

Fig. 5(a) Spectral dominated classification, (b) Spatial dominated classification.

B. *Deep Belief Network*

DBNs are the effective learning algorithms available. In paper [5], proposed a new feature extraction and image classification framework for hyperspectral data analysis based on deep belief network. In the cited paper, the author focused on single layer Restricted Boltzmann machine (RBM) – based model to learn the shallow and deep features of hyperspectral data. The learnt features are then used in LR to address the classification problem of hyperspectral data. The spectral-spatial classification strategy based on DBN was proposed. It can be assessed that DBN is an effective FE method, which reduces the dimension of features. RBM is commonly used as a layer-wise training model in the construction of a DBN. It is a two layer network as shown in Fig. 7. Features extraction in spectral and spatial domain is same as in case of stack auto encoder with the difference of replacing auto encoder with RBM. Integration of spectral and spatial domain together provides a discriminating power for the pixel-wise classification shown in Fig. 6. For each pixel, the 1-D vector processing using RBM is added to the end of spectral vector. After forming a hybrid set of spectral-spatial features, it is feed into DBN-LR without any pre-processing of FE and selection. Following the pre-training and fine-tuning steps discussed in the cited paper, assigned a class label to each pixel. It is conceded that the training complexity of DBN is a disadvantage, but they are superfast on the testing time. The superfast classification stage is a great advantage when large hyperspectral images are processed.

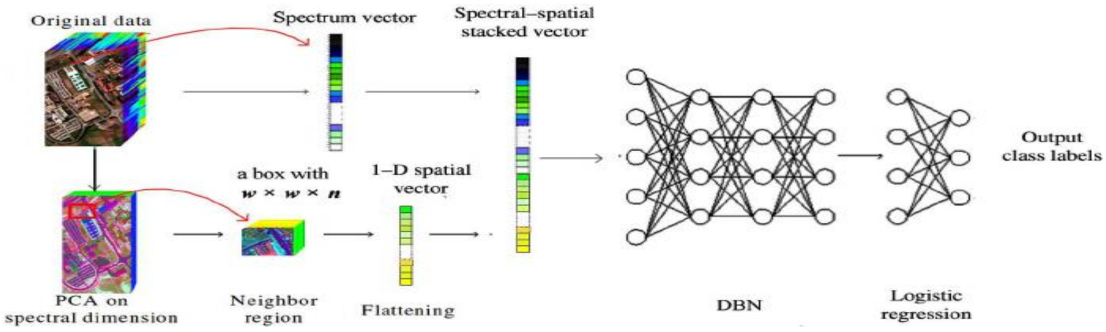


Fig.6 Spectral-spatial classification DBN-LR

C. Restricted Boltzmann Machine

The key aspects of deep learning are that it attempts to model high-level abstraction of data by using architectures composed of multiple non-linear transformations. Feature representations are the backbone of image classification. The ability of undirected graphical models like RBM, to capture distribution among pixels at the hidden level is utilized in the paper [6], to extract features for each bands in the hyperspectral image. The generative feature vectors are feed as input to different classifiers for the classification. An RBM consists of two types of units [6]. One is visible units and other is hidden units. The visible units constitute the first layer and correspond to the components of observation (one visible unit for each pixel of the digital input image). The hidden unit model relate to dependencies between the pixels in the image. Feature extraction using RBM provides better accuracy than other methods. The layout of layers used in RBM is shown in Fig. 7.

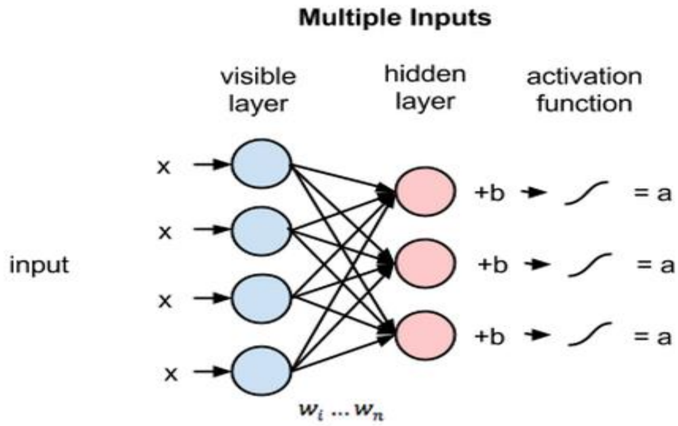


Fig.7 RBM layers

2. SUPERVISED LEARNING

Supervised deep network works with labelled information and categorize the input data in these labels.

A. *Deep Stacking Network*

DSN owns an advantage over other deep models for its simplicity when processing in batch mode learning. In paper [1], the feature extraction is gradually obtained by employing a nonlinear activation function on the hidden layer nodes of each module. The cited paper shows, this method achieves improved classification performance compared to SVM and NN methods. A DSN includes a variable number of Layered modules, wherein each module is a specialized neural network consisting of a single hidden layer and two trainable sets of weights. Due to the nonlinear characteristics of hyperspectral data, the nonlinear activation function is used to calculate the hidden layer node output of each module. The hidden layer of the lowest module of a DSN comprises a set of non-linear units that are mapped to the input units by way of a first, lower-layer weight matrix W. The non-linear units in each module of the DSN may be mapped to a set of the linear output units by way of a second, upper-layer weight matrix, which we denote by U. The classification of hyperspectral data is realized by combining DSN with logistic regression (DSN-LR). General structures of layers of stacking network are shown in Fig.8.

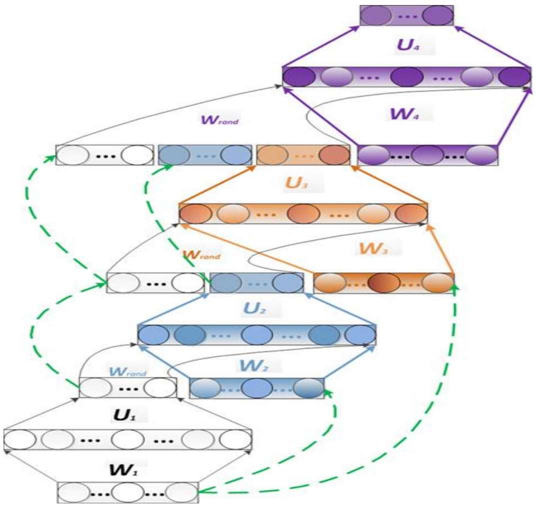


Fig. 8 General structure of layers of stacking network

B. Convolutional Neural Network

In a paper [2], deep convolutional neural networks are employed to classify hyperspectral images directly in spectral domain. The architecture of proposed classifier in this paper contains five layers with weights which consist of input layer, the convolutional layer, the max pooling layer, the full connection layer, and the output layer. These five layers are realized on each spectral signature to discriminate against each others. This method achieves better classification performance than some traditional methods such as SVM and the conventional deep learning based methods.

CNN also have been incorporate spatial information. In the cited paper, it is incorporated spatial information by decomposing the hyperspectral image cube into patches, each of which contains spectral and spatial information for a specific pixel. This paper exploits CNN to encode pixel's spectral and spatial information and a Multi-Layer Perceptron to conduct the classification task as shown in Fig. 9.

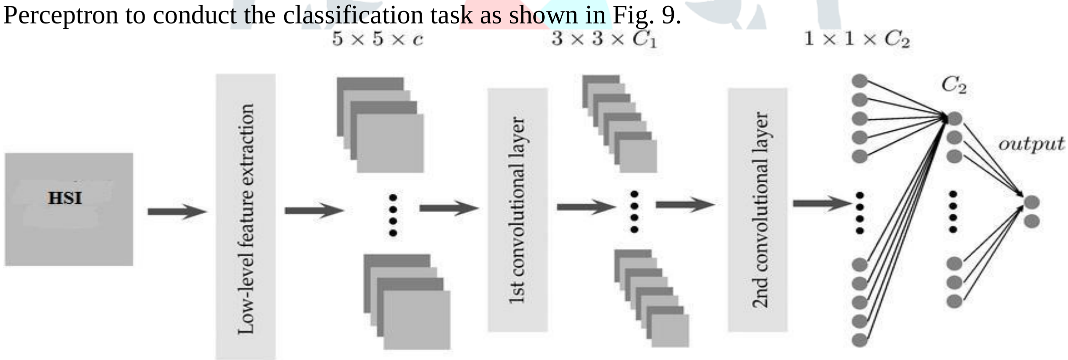


Fig. 9 Overall architecture of CNN

C. Recurrent Neural Network

An RNN is an important branch of the deep learning family which is mainly designed to handle sequential data. All the supervised models are a vector- based methodologies. These vector based approaches can lead to data loss when representing hyperspectral pixels. In paper [3], RNN model can effectively analyse hyperspectral pixels as sequential data instead of vector based approach and make use of an activation function called parametric rectified tanh (PRetanh) for hyperspectral sequential data analysis instead of the rectified linear unit. The activation function used in this paper makes it possible to use fairly high learning rates without the risk of divergence during the training procedure. An RNN is a network that uses recurrent connections between neural activations at consecutive time steps; such a network uses hidden layers or memory cells to learn the states that model the underlying dynamics of the input sequence for sequential data over time. In RNN, first, the value of existing spectral band xk is fed into the input layer as shown in Fig.10. Then, the recurrent layer receives x^k and calculates the hidden state information for the current band; it also restores that information in the meantime. Then, the next band x^{k+1} is input to the recurrent layer simultaneously with the state information of x^k, and the activation at spectral band k+1 is calculated by linear interpolation between suggested activation and the activation of previous band k. Then, the RNN can predict the label of the input hyperspectral pixel by looping through the entire hyperspectral pixel sequence.

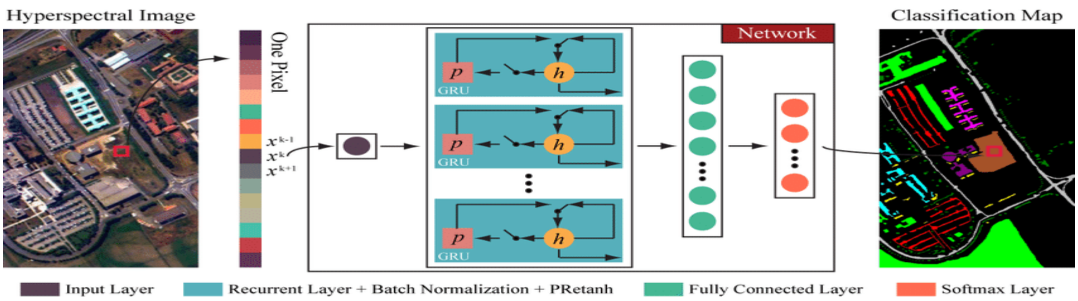


Fig.10 Overview of architecture of RNN

3. HYBRID DEEP NETWORKS (SEMI-SUPERVISED METHODS)

These methods make use of both generative and discriminative model components (they work with and without labelled data). Semi supervised learning is very useful in hyperspectral image classification to deal with the limited training samples problem.

Generative Adversarial Networks (GANs)

A GAN contains a generative network and a discriminative network. In paper [7], the usefulness and effectiveness of GAN for classification of hyperspectral images are explored for the first time. In this paper, two schemes were proposed: 1) a well-designed 1D-GAN as a spectral classifier and 2) a robust 3D-GAN as a spectral-spatial classifier. A GAN regarded as a regularization method which can mitigate the overfitting phenomenon which is the serious problem faced by deep learning methods due to limited training samples. With the help of GANs, the deep CNN achieves better performance in terms of classification accuracy compared with that of traditional CNN. The overfitting problem raised by CNN is mitigated. In the cited paper, PCA is used to reduce the high dimensionality of inputs, which is really important to stabilize the training procedure. GAN consist of both a generator and a discriminator. The generator produces fake samples which are considered as augmented data that increase the number of training samples. Both fake samples and the true samples are fed into the networks in order to optimize the nets. In 3D GAN, the generator accepts noise as input and transforms its shape to the same size as real data with principle components in the spectral domain. Then the discriminator acquires the real data or the generated fake samples as input data, and it uses the sigmoid classifier to give the real and fake classification results and softmax classifier to give the classification map. The overall architecture of 3D GAN is shown in Fig.11.

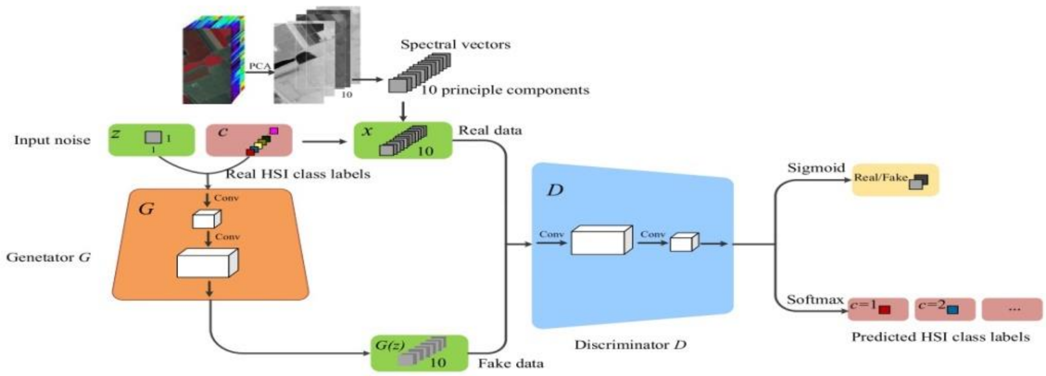


Fig. 11 Architecture of 3D GAN for hyperspectral image classification

V. CONCLUSIONS

Hyperspectral imagery is used to study the details of surface properties that are undetectable using other types of imagery. Classification of these images using traditional classifier results to reduce in accuracy. Thus deep learning is an emerging technology for handling complex features of the hyperspectral image which gives better accuracy which is discussed in this paper. Various deep learning classification methods are compared which includes supervised, unsupervised and semi-supervised techniques. Each method contains its own benefits over other methods which give improved classification accuracy for different datasets that were reviewed in this paper. From an overall comparison of each classification techniques for hyperspectral image, supervised deep network is most preferred due to its high classification accuracies.

REFERENCES

[1] Mingyi He, Xiaohui Li, Yifan Zhang, Jing Zhang, Weigang Wang,' Hyperspectral image classification based on deep stacking network', IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3286-3289, 2016.

[2] Konstantinos Makantasis , Konstantinos Karantzalos ,Anastasios Doulamis , Nikolaos Doulamis,' Deep supervised learning for hyperspectral data classification through convolutional neural networks', IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 4959-4962, 2015.

[3] Lichao Mou , Pedram Ghamisi , Xiao Xiang Zhu ,'Deep Recurrent Neural Networks for Hyperspectral Image Classification', IEEE Transactions on Geoscience and Remote Sensing, Volume: 55 , Issue: 7 , pp. 3639 – 3655,July 2017.

[4] Yushi Chen , Zhouhan Lin , Xing Zhao , Gang Wang , Yanfeng Gu,' Deep Learning-Based Classification of Hyperspectral Data', IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing , Volume: 7 , Issue: 6 , pp. 2094 – 2107,June 2014.

[5] Yushi Chen , Xing Zhao , Xiuping Jia,' Spectral–Spatial Classification of Hyperspectral Data Based on Deep Belief Network', IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Volume: 8 , Issue: 6 , pp. 2381 – 2392, June 2015.

[6] Midhun E M , Sarath R Nair , Nidhin Prabhakar T V , Sachin Kumar S,, 'Deep Model for Classification of Hyperspectral image using Restricted Boltzmann Machine', ICONIAAC '14 Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing, 2014.

[7] Lin Zhu ; Yushi Chen ; Pedram Ghamisi ; Jón Atli Benediktsson,' Generative Adversarial Networks for Hyperspectral Image Classification', IEEE Transactions on Geoscience and Remote Sensing, Volume: 56 , Issue: 9 , pp. 5046 – 5063, Sept. 2018.

[8] Ying Zhan ; Dan Hu ; Yuntao Wang ; Xianchuan Yu,' Semi supervised Hyperspectral Image Classification Based on Generative Adversarial Networks', IEEE Geoscience and Remote Sensing Letters , Volume: 15 , Issue: 2 , pp. 212 – 216, Feb. 2018.

[9] M.E.Paoletti,J.M.Haut, J.Plaza, A.Plaza,' A new deep convolutional neural network for fast hyperspectral image classification', ISPRS Journal of Photogrammetry and Remote Sensing, Volume 145, Part A ,pp. 120-147, November 2018.

[10] Wei Hu, Yangyu Huang, Li Wei, Fan Zhang, and Hengchao Li, 'Deep Convolutional Neural Networks for Hyperspectral Image Classification' , Journal of Sensors, pp.1-12,Volume 2015.

[11] Li deng, 'A tutorial survey of architectures, algorithms, and applications for deep learning', APSIPA Transactions on Signal and Information Processing Volume 3, 2014.

[12] Henrik Petersson , David Gustafsson , David Bergstrom, 'Hyperspectral image analysis using deep learning — A review', 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA) , 2016.

[13] Savita P. Sabale, Chhaya R. Jadhav, 'Hyperspectral Image Classification Methods in Remote Sensing - A Review', International Conference on Computing Communication Control and Automation, 2015.

[14] Xiaorui Ma, Jie Geng, Hongyu Wang, "Hyperspectral image classification via contextual deep learning", EURASIP Journal on Image and Video Processing,2015.

[15] Muhammad Jaleed Khan ; Hamid Saeed Khan ; Adeel Yousaf ; Khurram Khurshid ; Asad Abbas,'Modern Trends in Hyperspectral Image Analysis: A Review', IEEE Access (Volume: 6) , pp. 14118 – 14129, 2018.

[16] Jun Li ; Mahdi Khodadadzadeh ; Antonio Plaza ; Xiuping Jia ; José M. Bioucas-Dias," A Discontinuity Preserving Relaxation Scheme for Spectral–Spatial Hyperspectral Image Classification", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing ,Volume: 9 , Issue: 2 , pp. 625 – 639, Feb. 2016.

[17] CeZhang, XinPa,HuapengLi, Andy Gardiner, IsabelSargent, JonathonHare , Peter M.Atkinson" A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification", ISPRS Journal of Photogrammetry and Remote Sensing, Volume 140, Pages 133- 144, June 2018.

[18] Jiaojiao Li ; Xi Zhao ; Yunsong Li ; Qian Du ; Bobo Xi ; Jing Hu ," Classification of Hyperspectral Imagery Using a New Fully Convolutional Neural Network", IEEE Geoscience and Remote Sensing Letters ,Volume: 15 , Issue: 2 , pp. 292 – 296, Feb. 2018.

[19] ShiqiYu, SenJia, ChunyanXu, "Convolutional neural networks for hyperspectral image classification", Neurocomputing, Volume 219, pp. 88-98, 5 January 2017.

[20] Yushi Chen ; Zhouhan Lin ; Xing Zhao ; Gang Wang ; Yanfeng Gu," Deep Learning-Based Classification of Hyperspectral Data", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing , Volume: 7 , Issue: 6 , pp. 2094 – 2107, June 2014.

[21] Xudong Kang; Shutao Li; Jon Atli Benediktsson," Hyperspectral image classification: A benchmark", 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2017.

[22] Xuefeng Liu ; Qiaoqiao Sun ; Bin Liu ; Biao Huang ; Min Fu," Hyperspectral image classification based on convolutional neural network and dimension reduction", Chinese Automation Congress (CAC), IEEE, 2018.

[23] YonghaoXu, BoDu, FanZhang,LiangpeiZhang," Hyperspectral image classification via a random patches network", ISPRS Journal of Photogrammetry and Remote Sensing, Volume 142, pp. 344-357, August 2018.

[24] Xiaofei Yang ; Yunming Ye ; Xutao Li ; Raymond Y. K. Lau ; Xiaofeng Zhang ; Xiaohui Huang ,"Hyperspectral Image Classification With Deep Learning Models", IEEE Transactions on Geoscience and Remote Sensing ,Volume: 56 , Issue: 9 , pp. 5408 – 5423, Sept. 2018.

[25] José M. Bioucas-Dias ; Antonio Plaza ; Nicolas Dobigeon ; Mario Parente ; Qian Du ; Paul Gader ; Jocelyn, "Hyperspectral Unmixing Overview: Geometrical,Statistical, and Sparse Regression-Based Approaches", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing , Volume: 5 , Issue: 2 , pp. 354 – 379, April 2012.

[26] Salah Bourennane ; Caroline Fossati ; Alexis Cailly," Improvement of Classification for Hyperspectral Images Based on Tensor Modeling" IEEE Geoscience and Remote Sensing Letters Volume: 7 , Issue: 4 , pp. 801 – 805, Oct. 2010.

[27] Abhishek Agarwal ; Tarek El-Ghazawi ; Hesham El-Askary ; Jacquline Le-Moigne," Efficient Hierarchical-PCA Dimension Reduction for Hyperspectral Imagery", IEEE International Symposium on Signal Processing and Information Technology, **ISSN:** 2162-7843, Egypt, 2008.

[28] Yuemei Ren ; Liang Liao ; Stephen John Maybank ; Yanning Zhang ; Xin Liu," Hyperspectral Image Spectral-Spatial Feature Extraction via Tensor Principal Component Analysis", IEEE Geoscience and Remote Sensing Letters ,Volume: 14 , Issue: 9 , pp. 1431 – 1435, Sept. 2017.

[29] Fenghua Huang, YingYu, Tinghao Feng," Hyperspectral remote sensing image change detection based on tensor and deep learning", Journal of Visual Communication and Image Representation, Volume 58, pp. 233-244,January 2019.

[30] P. M. Atkinson &A. R. L. Tatnall," Introduction Neural networks in remote sensing", International Journal of Remote Sensing , Volume 18, - Issue 4, Pages 699-709 , 2010.

[31] J. A. Richards, 'Remote Sensing Digital Image Analysis: An Introduction'. New York, NY, USA: Springer, 2013.

[32] D. Landgrebe, "Hyperspectral image data analysis," IEEE Signal Process. Mag., vol. 19, no. 1, pp. 17–28, Jan. 2002.

[33] F. V. D. Meer, "Analysis of spectral absorption features in hyperspectral imagery," Int. J. Appl. Earth Observ. Geoinf. vol. 5, no. 1, pp. 55–68, Jan. 2004.

[34] P. W. Yuen and M. Richardson, "An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition," Imaging Sci. J., vol. 58, no. 5, pp. 241–253, May 2010.

[35] E. K. Hege et al., "Hyperspectral imaging for astronomy and space surveillance," in Proc. SPIE's 48th Annu. Meet. Opt. Sci. Technol., San Diego, CA, USA, 2004, pp. 380–391.

[36] J. Bioucas-Dias et al., "Hyperspectral remote sensing data analysis and future challenges," Geosci. Remote Sens. Mag., vol. 1, no. 2, pp. 6–36, Feb. 2013.

[37] S. Rajan, J. Ghosh, and M. M. Crawford, "An active learning approach to hyperspectral data classification," IEEE Trans. Geosci. Remote Sens., vol. 46, no. 4, pp. 1231–1242, Apr. 2008.

[38] S. B. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectral remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 39, no. 7, pp. 1360–1367, Jul. 2001.

[39] A. R. Mohamed, T. N. Sainath, and G. Dahl, "Deep belief networks using discriminative features for phone recognition," in Proc. Acoust. Speech Signal Process. (ICASSP), Prague, Czech Republic, 2011, pp. 5060–5063.

[40] Y Zhou, Y Wei. "Learning Hierarchical Spectral-Spatial Features for Hyperspectral Image Classification". *IEEE Trans. On Cybern.* 2015.

[41] C. Chang, Hyperspectral Imaging: Techniques for Spectral Detection and Classification. Dordrecht, the Netherlands: Kluwer, 2003.

[42] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," Proc. IEEE, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[43] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, "Advanced spectral classifiers for hyperspectral images: A review," IEEE Geosci. Remote Sens. Mag., vol. 5, no. 1, pp. 8–32, Mar. 2017.

[44] A. Radford, L. Metz, and S. Chintala. (2016). "Unsupervised representation learning with deep convolutional generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1511.06434.

[45] Y. Chen, L. Zhu, P. Ghamisi, X. Jia, G. Li, and L. Tang, "Hyperspectral images classification with Gabor filtering and convolutional neural network," IEEE Geosci. Remote Sens. Lett. vol. 14, no. 12, pp. 2355–2359, Dec. 2017.

[46] J Bioucas-Dias. Hyperspectral remote sensing data analysis and future challenges. Geosci. Remote Sens, 1(2):6–36, 2013.

[47] Xueyun Chen, Shiming Xiang, Cheng-Lin Liu, and Chun-Hong Pan. Vehicle detection in satellite images by hybrid deep convolutional neural networks. Geoscience and Remote Sensing Letters, IEEE, 11(10):1797–1801, October 2014.

[48] Li Deng and Dong Yu. Deep learning: Methods and applications. Foundations and Trends in Signal Processing, 7(3-4):197–387, 2013.

[49] Michael Theodore Eismann. Hyperspectral Remote sensing. SPIE Press, 2012.

[50] Zhouhan Lin, Yushi Chen, Xing Zhao, and Gang Wang. 'Spectral-spatial classification of hyperspectral image using auto encoders'. In Information, Communications and Signal Processing (ICICS) 2013 9th International Conference on, pages 1–5, Dec 2013.