Hyperspectral Imagery Classification Using Deep Learning

Indira Bidari
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
indira_bidari@kletech.ac.in

Satyadhyan Chickerur
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
chickerursr@kletech.ac.in

Harivijay Ranmale
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
harivijayranmale91@gmail.com

Sushmita Talawar
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
sushmitatalawar555@gmail.com

Harish Ramadurg
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
harishramadurg95@gmail.com

Rekha Talikoti
School of Computer Science and Engineering
KLE Technological University
Hubballi, India
rekhatalikoti2@gmail.com

***Abstract*—Hyperspectral Imagery (HSI) data analysis and processing is an emerging topic in the arena of remote sensing and earth observation technology. Recently land cover deep learning based classification algorithms have become an emerging research area and these techniques are used in majority of applications like agriculture, military surveillance, environmental analysis, urban investigation, mineral exploration. An end-to-end deep learning architecture is introduced in this paper which extracts band from spatial-spectral features and also performs classification with comparative classifier analysis and provides state-of-the-art efficiency.**

***Index Terms*—Convolutional Neural Networks, hyperspectral imagery, Deep Neural Networks, pattern classification, land cover classification.**

I. INTRODUCTION

Hyperspectral imaging (HSI) [1] investigates a wide spectrum of light to each pixel. The light which strikes each pixel is split into the different spectral bands. Spectral imaging uses multiple bands across the electromagnetic spectrum consisting of the infrared, the visible spectrum, the ultraviolet, X-Rays, etc. In HSI, an image each pixel's spectral information is added as 3-D of the values to the 2-D spatial images, generating the 3-D data cube referred to as hypercube data. It produces data volumes of 3-D (x, y), where x, y corresponds to the spatial dimensions and corresponds to the spectral dimension. HSI is over-determined spectrally, they provide sufficient spectral details for defining and distinguishing target objects. They have the ability to collect information more accurately and in greater details. Recently deep neural networks have been engaged for land cover classification in Hyperspectral imaging. Methods on deep learning for the HSI classification [2] focus entirely on spectral-spatial background modeling to tackle the issue of spectral signature temporal variation. The proposed solution is to develop a system to perform the comparative analysis of the traditional classifiers on Deep Neural Network i.e., Secular Neural Network (CNN) and the Multilayer Perceptron (MLP). This work is focused on Deep Learned Artificial Neural Networks [3] [4] development for vital land-cover classification of hyperspectral imagery.

The CNN algorithms [5] are self-learning and automatically the necessary features are extracted by the algorithm rather than the coder specifying the features to be extracted. A frequent combination of machine learning methods with CNN a feature extractor is used, in particular Support Vector Machines (SVM). CNN is an approach to dynamically learn features from the data. The methodology has two benefits from CNN side and as well as from classical machine learning side, as it exploits the capability to automatically retrieve a good feature set and the robustness to over fitting even on small datasets respectively. A CNN-SVM [6] model for Land cover classification on hyperspectral takes spectral information from its neighbors and a target pixel, arranges into a spectral-spatial multi-function cube without the need for spectral information an optional change to the CNN.

II. RELATED WORK

Despite the challenges and difficulties in Hyperspectral Imaging such as imbalance between high dimensionality, training samples incomplete accessibility and the huge spatial variability of the spectral signature. Hyperspectral Imaging remains a dynamic research in Remote Sensing (RS) and Artificial Intelligence (AI). The study on the HSI classification in the early stages exhibited that majority approaches have concentrated on discovering the role of the HSIs spectral signatures for the drive of classification. Many pixel-wise classification techniques have been proposed to classify HSIs, such as Support Vector Machines (SVM),

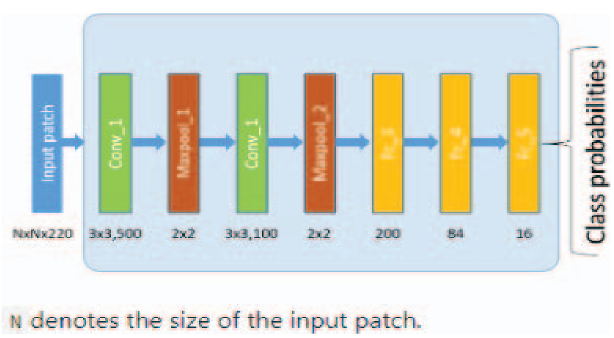
978-1-7281-6823-4/20/$31.00 ©2020 IEEE

Multinomial Logistic Regression and the Convolutional Neural Network. Some additional classification approaches, such as Individual Component Analysis (ICA), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA), have focused on the design of dimension reduction approaches. Nevertheless, the classification maps achieved by these pixel-wise classifiers are inadequate as the spatial contexts are not considered. Recently, the spatial characteristics have been expressed as useful in illumination, the representation of the hyperspectral data and the classification accuracy. Silvia Valero and Jocelyn Chanussot proposed data preprocessing methods like hyperspectral data representation, redundancy reduction and limited annotated training samples. In hyperspectral data representation, the two fundamental representations are spatial and spectral information of Hyperspectral Images (HSI). In Redundancy Reduction, band selection is accomplished using maximum class separability criterion or methods based on information theory. Pixel pair functions have been used to reduce the small training samples issues. The methodologies for object detection used in this work are grouped into two types "Two Step Approach" and "One Step Approach". Traditional methods of extraction traits such as linear extraction, Independent Component Analysis (ICA), Principal Component Analysis (PCA), and Minimum Noise Function (MNF) are used in a two-step approach. In the second step of this approach conventional classifiers such as K-Nearest Neighbor (KNN) and SVM (Support Vector Machine) [7] [8] are used for classification. Methods based on deep learning are used in one step approach for automatic feature extraction from the neural network's internal layer, and then these features are provided as traditional classifier inputs. Many deep learning frameworks and methods have been proposed for classification of scene based indoor location recognition [10] [12]. And several classification approaches using transfer learning on cultural heritage sites [14].

III. METHODOLOGY AND IMPLEMENTATION

CNN or ConvNet [9] is an artificial neural network intended for processing data with a grid structure. This architecture is grounded on parameter sharing and sparse interactions. Spatial invariances in the images can be learnt more effectively thought this methodology.

A. Convolutional Neural Network



A typical ConvNet architecture includes 4 types of layers:

1) Convolutional (conv): Convolutional Neural Network contains deep models, the functionality of this to apply training filters on the input data which results as an increasing complex features. It belongs to the class of deep neural network. It consisting of neurons having learnable weights and the biases. The network still expresses a function of a single differentiable ranking.

2) Pooling (pool): In the ConvNet architecture, the purpose of pooling layer is to reduce the spatial size in order to reduce the number of parameters and computation, and henceforth also to manage over fit. The pooling layer works on each input depth slice independently and it resizes spatially using the MAX operation.

3) Fully connected (affine): A completely connected layer that forms the network's last few layers, the data from previous pooling or (flattened) convolution layers is input to this. In this layer, the one-layer inputs have complete connections to all next-layer activation modules. After a bias offset, the activations are calculated with a matrix multiplication.

4) Rectifying Linear Unit (ReLU): It is a linear function by component, if it is positive then the input will then be outputted directly, otherwise the output will be zero. Within a neural network, the activation function is responsible for transmuting the pooled weighted input from the node into the output for that node input or activate.

B. Implementation

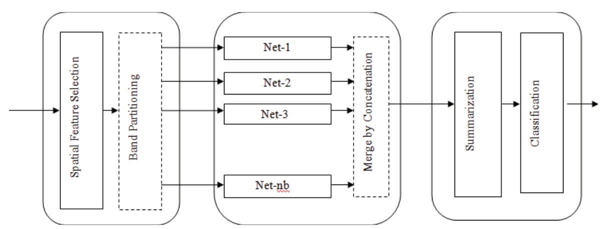


Fig. 2. Block diagram of the system

1) Data Handling: The Hyperspectral Image (HSI) data is usually accessible in the .mat format; this can be retrieved using programming language like python. Hyperspectral data have a different spatial–spectral point of view. Substantial traditional deep learning approaches exploit data pixel-wise (1-dimensional approaches), concentrating in a spectral direction. Pixels extraction of the HSI data is one of the significant preprocessing tasks. This step helps to handle HSI data and to implement the classification algorithms.

2) Pixel Extraction: Pixels extraction of the HSI is one of the imperative preprocessing mechanisms. It assists to handle the data and to implement the machine learning algorithms such as clustering, classification, etc. The individual elements in the Hyperspectral Image (HSI) are pixels which are vectors

3) Dimensionality reduction: To tackle redundancy in the spectral information, we use this approach. Principal Component Analysis (PCA) [9] is a classic way to proceed for different dimensionality reduction technique. A 2-D process can be a practical approach. Spatial processing is exploited by extracting spatial features on 2-D patches or from the whole bands. PCA [11] is a technique to handle data dimensionality [13] and in numerous deep learning pipelines to pre-process data as well.

4) K-Nearest Neighbor (KNN): A version of the KNN classifier algorithm, by finding the nearest neighbor class it predicts the target label. Using the distance measures like Euclidean distance the closest class will be identified. The result is calculated as the class with KNN for classification with the maximum frequency from the K-most similar instances. Class achieves the votes from each instance. The prediction is considered by taking the class which has the maximum votes. For the new data instance, class probabilities is measured by considering each class in the set of K most similar instances as the normalized frequency of samples.

5) Support Vector Machine (SVM): To distinct data points into many classes the SVM makes use of hyperplanes. Support vectors make sure that the width of the margin is maximized. A single supplementary point of data can especially affect the locality of the hyperplane. In its unique origin, the technique is offered with a set of the labeled data.

6) Multi-Layer Perceptron: It is an ANN with one or many hidden layers. A perceptron is a precursor to the larger neural networks consisting of single neural model. It models high non-linear functions which forms the basis of deep learning neural network model.

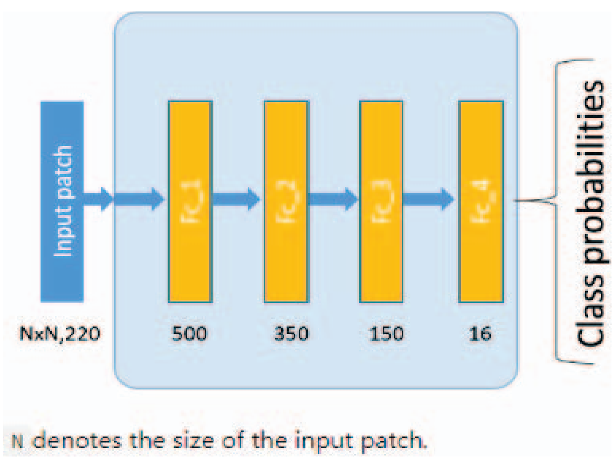


Fig. 3. Full architecture of the Multi-Layer Perceptron

7) CNN as a Feature Extractor: A frequent combination of machine learning methods with CNN a feature extractor is used, in particular Support Vector Machines (SVM). CNN is a methodology to dynamically learn features from the data. The methodology has two benefits from CNN side and as well as from classical machine learning side, as it exploits the capability to automatically retrieve a good feature set and the robustness to over fitting even on small datasets respectively. For the hyperspectral land-cover classification, a hybrid CNN-SVM [8] [9] the spectral information of its neighbors and the target pixel are arranged into the spectral–spatial multi-feature cube without the additional change of the CNN.

8) Specification of the learning algorithm: The learning algorithms details are: Parameter update algorithm being used with batch size of 200, learning rate of 0.01and number of steps is until the best validation performance.

IV. DATASET

The Indian Pines (IP) Dataset is collected in Northwestern Indian and consists of 145×145 pixels and 224 spectral reflectance bands of 0.4–2.5 10^{-6} meters wavelength range. This scene is one of a wider sub-set. The IP scene contains two-thirds of forestry, and one-third of perennial forest or other natural vegetation. The accessible ground truth is divided into 16 grades, and is not all separated from each other.

Sensor	AVIRIS
Place	North Western Indiana
Frequency Band	0.4-0.45 µm
Spatial Resolution	20m
No. Of Channels	220
No. Of Classes	16

Fig. 4. Pine dataset specification

V. RESULTS

The experiments are performed on convolutional neural network varying input size and filter size. Test results on IP for various patch-size input selections are shown in Figure 5. Better classification accuracy is accomplished by increasing the patch-size and providing more spatial meaning as well.

Figure 6. shows different input patch sizes generated from decoding.

Evaluation metrics of the proposed architecture:

1) Rank-specific accuracy for a Class Ci is measured as the class portion of the samples that are categorized correctly.
2) Total precision: total number of samples listed correctly / total number of samples in all groups.

Figure 7. shows results of a study of approaches proposed based on in-depth learning with conventional classifiers. The suggested structure takes priority over all other approaches. The test accuracy of Convolutional Neural Network outperforms all other methods with overall accuracy of 87.97%. The test accuracy of CNN on Indian Pine dataset exceeds SVM and K-NN by 11.67% and 15.32% respectively.

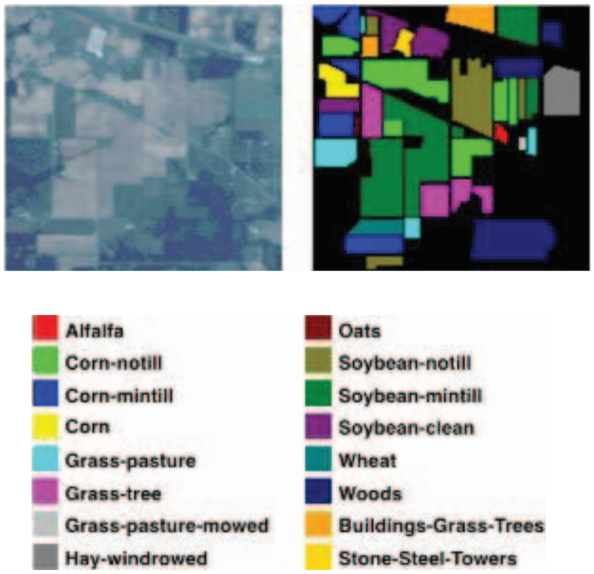


Fig. 5. Input and target output images

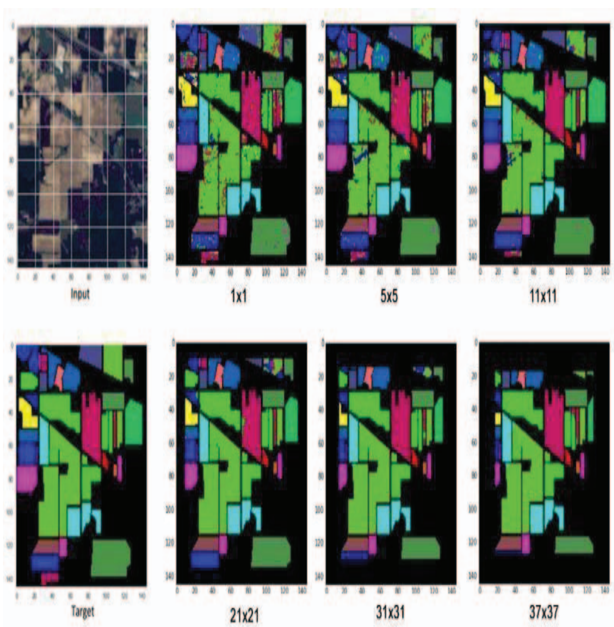


Fig. 6. Decoding generated for different input patch sizes

Class	K-NN	SVM	MLP	CNN
1	61.83	76.8	77.74	86.06
2	71	71.3	79.05	88.19
3	75.56	73.68	94.70	96.07
4	69	79.6	98.11	99.73
5	89	69.9	99.64	100
6	88	80.02	83.68	90.02
7	87.45	67	79.60	71.00
8	68.89	64.2	89.31	95.62
9	81.82	63.3	98.12	98.66
10	70.1	62.1	76.08	85.32
11	72.36	63.34	87.62	91.50
12	75.89	68.45	74.60	78.64
13	75.7	72.34	93.78	94.72
14	68.4	72.12	82.63	91.33
15	64.6	72.8	71.85	90.6
16	69.8	73.45	98..65	89.54
OA	72.65	76.3	86.51	87.97

Fig. 7. Class specific accuracy (%) and Overall Accuracy (%)

VI. CONCLUSION

The article proposes an end-to - end architecture of profound learning of the neural network to conduct band-specific spectral-spatial feature learning for superior modeling. Curse of dimensionality are tackled by Principal Component Analysis (PCA) reduction technique. Experiments on Hyperspectral image classification data shows superior classification performance with overall accuracy of 87.97%.

VII. ACKNOWLEDGEMENTS

We take this opportunity to express our special thanks to KLE Technological University for their support and encouragement.

REFERENCES

[1] D. Landgrebe, “Hyperspectral image data analysis,” IEEE Signal Process. Mag., vol. 19, no. 1, pp. 17–28, 2002.
[2] J. Richards, Remote Sensing Digital Image Analysis: An Introduction. Springer, New York, NY, USA, 2013.
[3] L. Zhang, L. Zhang, and B. Du, “Deep learning for remote sensing data: A technical tutorial on the state of the art,” IEEE Geosci. Remote Sens. Mag., vol. 4, no. 2, pp. 22–40, Jun. 2016.
[4] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, no. 3, pp. 436–444.
[5] I. Goodfellow, Y. Bengio, and A. Courville, “Deep learning.” Book in preparation for MIT Press, 2016.
[6] Y.LeCun,L.Bottou,Y.Bengio,andP.Haffner,“Gradient-basedlearning applied to document recognition,” Proceedings of the IEEE, pp. 2278–2324, 1998.

[7] F.Melgani and B.Lorenzo, “Classification of hyperspectral remote sensing images with support vector machines,” IEEE Trans. Geosci. Remote Sens., vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
[8] L. G. et al., “Subspace-based support vector machines for hyperspectral image classification,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 2, pp. 349–353, Feb. 2015.
[9] S. Y. et al., “Convolutional neural networks for hyperspectral image classification,” Neurocomputing, 2016.
[10] Hanni, Akkamahadevi, Satyadhyan Chickerur, and Indira Bidari. "Deep learning framework for scene based indoor location recognition." In 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy), pp. 1-8. IEEE, 2017.
[11] L. G. et al., “Subspace-based support vector machines for hyperspectral image classification,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 2, pp. 349–353, Feb. 2015.
[12] Pujar, Karthik, Satyadhyan Chickerur, and Mahesh S. Patil. "Combining RGB and Depth Images for Indoor Scene Classification Using Deep Learning." In 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-8. IEEE, 2017.
[13] B. Waske, S. van der Linden, J. A. Benediktsson, A. Rabe, and P. Hostert, “Sensitivity of support vector machines to random feature selection in classification of hyperspectral data,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 7, pp. 2880–2889, Jul. 2010.
[14] Kulkarni, Uday, S. M. Meena, Sunil V. Gurlahosur, and Uma Mudengudi. ”Classification of Cultural Heritage Sites Using Transfer Learning.” In 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), pp. 391-397. IEEE, 2019.