International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

Hyperspectral Image Classification by Using K-Nearest Neighbor Algorithm

Perepi. Rajarajeswari, R. Hemashri, S. Jayapriya and M.M. Ravikumar

Abstract--- *Recently, Deep learning has been acknowledged as one of the strong tool for feature-extraction to effectually address non linear problems and is employed in image processing tasks and has attained good performance. However, excessively increasing the depth of the network will lead to overfitting and gradient vanishing. To address these issues, a deep feature fusion network (DFFN) is introduced. With the application of k-nearest neighbor (KNN) algorithm in combination with residual learning, improves the classification accuracy by extracting much more discriminative features of HSI and also ease the training of deep network. It also extracts the hybrid features of the various classes in the image. The proposed model combines the results of various hierarchical layers that improve the classification accuracy.*

Index Terms--- *K-Nearest Neighbour (KNN), Hybrid Features, Residual Learning, Feature Fusion.*

I. INTRODUCTION

Hyperspectral imaging is the process of collecting and processing the information across the electromagnetic spectrum. It consists of hundreds of spectral bands ranging from visible spectrum to infrared spectrum. A high dimensional vector is used to represent each pixel in an HSI whose entries correlate to spectral reflectance in a specific wavelength. HSI gathers both special and spectral data that help in better extraction of features.

HSI has become a very important research topic in Remote Sensing domain. Fusion of spectral and spatial information is an efficient way of enhancing the accuracy of HSIs [1]. Various methods have been applied to HSIs like the Support Vector Machine (SVM)[2], Extended Morphological Profiles (EMP)[3], Large-scale Multi Label Learning (LMLL) method, Local Binary Pattern (LBP) method, etc., have been used to analyze the HSIs. The classification maps that were obtained from the pixel based classifiers were noisy as it did not consider the spatial contents [4], [5]. In later periods, the spectral-spatial feature based classification was introduced which constructed the spectral-spatial features by using multiple morphological operations. Recently deep learning techniques have been used for feature extraction to analyze the deep architecture and multi layered HSIs [6].

Deep models, though effectively extracts deep features, it couldn't utilize the spatial information of HSIs completely when the input transformation to 1-dimentional vector is done. Very recently, convolutional neural network (CNN) based HSI classification was proposed that directly processes the small cubes of HSIs[7], [8], extracting the deep pixel-pair features. But increasing depth creates issues like gradient vanishing, overfitting and accuracy degrading. To overcome this, a Deep Feature Fusion Network (DFFN) is introduced along with residual learning that gathers the discriminative features from deep layers [9]. A Convolutional neural network (CNN) is

Perepi. Rajarajeswari, Associate professor, Department of CSE, Kingston Engineering College,Vellore, Tamil Nadu, India.
R. Hemashri, UG Student, Department of CSE, Kingston Engineering College,Vellore, Tamil Nadu, India.
S. Jayapriya, UG Student, Department of CSE, Kingston Engineering College,Vellore, Tamil Nadu, India.
M.M. Ravikumar, Professor, Department of Mechanical Engineering, Kingston Engineering College, Vellore, Tamil Nadu, India.

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

used along with DFFN that uses multiple convolution layers to extract features and to finally combine the results of several layers to obtain desired output[9][12].

In this paper, we propose deep feature fusion network (DFFN) along with k-nearest neighbor algorithm (KNN), to extract the hybrid features of the HSIs and then produce the best output by choosing the nearest neighbour to the input, thereby, increasing the total accuracy and the individual accuracy of the inputted hyperspectral image.Related work is presented in section 2.Section 3 describes overall description of system.Proposed work and algorithms are given in section 3 and section4.Datasets description in section 6.Section 7 gives results and discussions conclusion is given Section 8.

II. RELATED WORK

Hyperspectral images are being used for various applications like Farming, Mining[14], Forestry, Environmental Management and Satellite Communication. Ahram et al. [10] classified images using convolutional neural network (CNN) to classify materials with similar spectral characteristics and attribute information.

Shutao Li et al. [4] proposed a classification technique by the probabilistic fusion of pixel-level and super pixel-level classifiers, by incorporating the spatial information from similar regions, which effectively eliminate the noisy appearance. Weiwei Song et al. [9] introduced deep feature fusion network (DFFN) for HSI classification in combination with residual learning technique to upgrade several convolutional layers.

III. SYSTEM OVERVIEW

In this paper, we present a deep feature fusion network in combination with the KNN algorithm for Hyperspectral image classification to solve the issues of increasing network depth. When an HSI image is taken, it consists of hundreds of hierarchical layers. Each layer is analyzed separately to extract the discriminant features and is processed.

Here we gather the hybrid features i.e., considering multiple features of each class for the classification purpose (height, width, colour, shape, weight, etc.). Initially the ground tooth images of the Indian Pines image is gathered which shows the 16 different types of classes of pines available.

This image is then pre-processed to obtain the training values. The Principal Component Analysis (PCA) is applied to this dataset to derive the PC1 values and then PCA in combination with KNN is applied to obtain PC2. Feature extraction is done here and a classifier is obtained.

This set of data is analyzed using the k-nearest neighbor (KNN) algorithm, starting from n values to the last available data in the dataset. This process helps to find the value that is nearest to the testing sample from the trained dataset to obtain the best nearest matching value with the most accuracy. Application of the KNN algorithm improves the prediction rate of the system.

The proposed system is much more flexible when compared to the existing DFFN with CNN system and increases the processing speed and fetches the output in a lesser time.

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

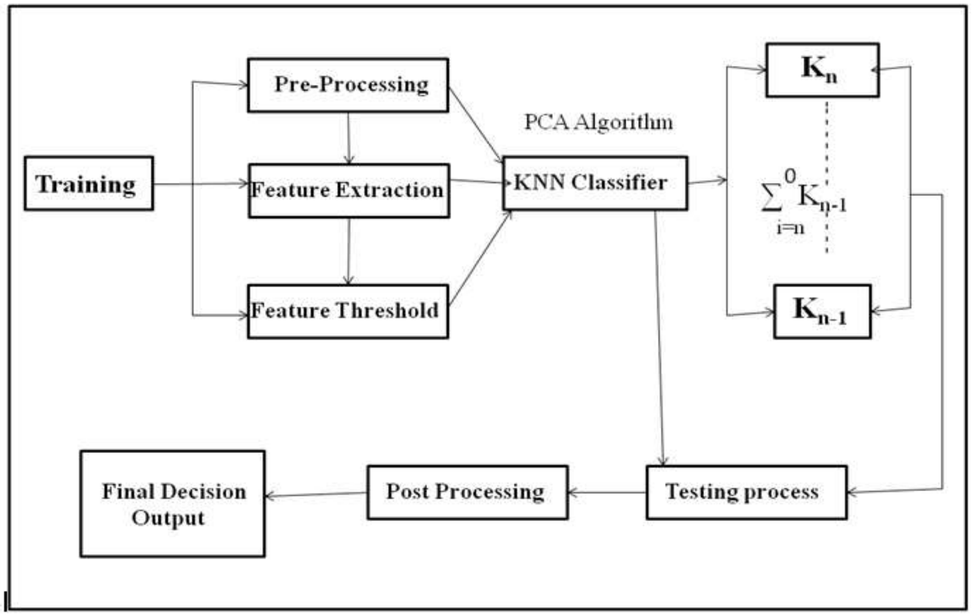

Fig. 1: Architecture Diagram

IV. PROPOSED WORK

1. Pre-Processing of Hyperspectral Image

Input the Ground tooth image obtained by applying the PCA algorithm to the original Indian pines hyperspectral image. Using dimensionality reduction techniques, high dimensional data is reduced to low dimensional data. To this apply the KNN Filters and perform resizing of the obtained image. From the resultant, obtain the data value for each image load it as .csv file.

2. Feature Extraction

The .csv file contains the training data. While performing feature extraction we consider the pre-trained data as an arbitrary feature extractor, allowing the input image to propogate forward, stopping at prespecified layer. Using the pines colour and number of classes as a main feature to set a threshold value based on which the most discriminant features are extracted.

3. Classification

1. Classification of hyperspectral images is more
2. Challenging task due to number of reasons like, availability
3. Of less number of training samples (Hughes phenomenon),
4. Presence of mixed pixels and the curse of dimensionality of
5. Data.

In this step, we apply the neural network algorithm i.e., the KNN algorithm to the obtained data values and define a decision rule to obtain the classification and mapping is done.

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

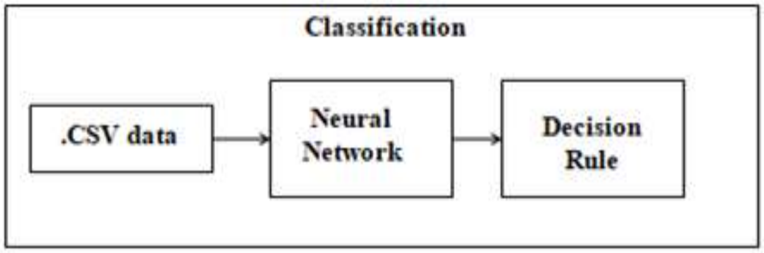

Fig. 2: Classification

4. Post Processing

In this step the classification accuracy is improved. Features extracted from the one test data are applied to synchronisation techniques and the obtained output is validated and analysed.

V. ALGORITHIMS

Principal Component Analysis (PCA)

Principal Component Analysis is used for dimensionality reduction of the input data that condenses the large set of data variables into a smaller set by applying some transformation techniques. Feature Extraction is done here in order to extract the most discriminant features from our input. The output of this step is in the form of Eigen vectors i.e., PC1, PC2, so on. Which are considered as the new axes.

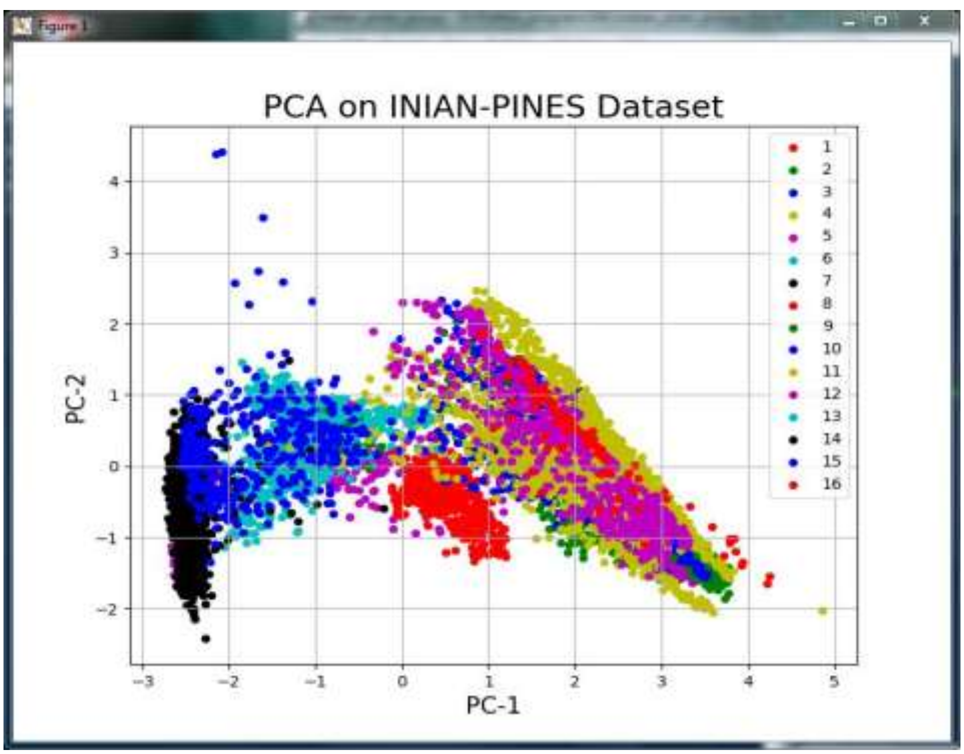

Fig. 3: PCA on Indian Pines Dataset

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

K-Nearest Neighbor Algorithm (KNN)

It is a Supervised Machine Learning algorithm used to solve classification and regression predictive problems. It is a Lazy Learning algorithm as no specialized training phase is present and it utilizes the entire available data for the training process, and also a Non-Parametric Learning algorithm since it does not consider anything about the underlying data.

Step 1: Input the training data and test data to the algorithm.

Step 2: Assign the value for K (any integer) based on the nearest data points.

Step 3: For every point in the test data:

- Calculate the distance between test data and each row of training data using the following techniques, namely Euclidean, Manhattan or Hamming distance. The most commonly used method is Euclidean.
- Based on the obtained distance value, arrange them in ascending order.
- Choose the top K rows from the sorted array.
- Assign a class to the test point based on most frequent class of the chosen rows.

Step 4: Stop.

VI. DATA SETS

Upon the application of multiple Hyperspectral image processing techniques and various algorithms repeatedly on the ground tooth image of the Indian Pines image, a large set of data values are obtained for each pixel in the image for the various classes present. This data is called the complete data for the specified HSI which is stored in a .csv format, for being used in the further process in our paper. This .csv file is the training data for our proposed system.

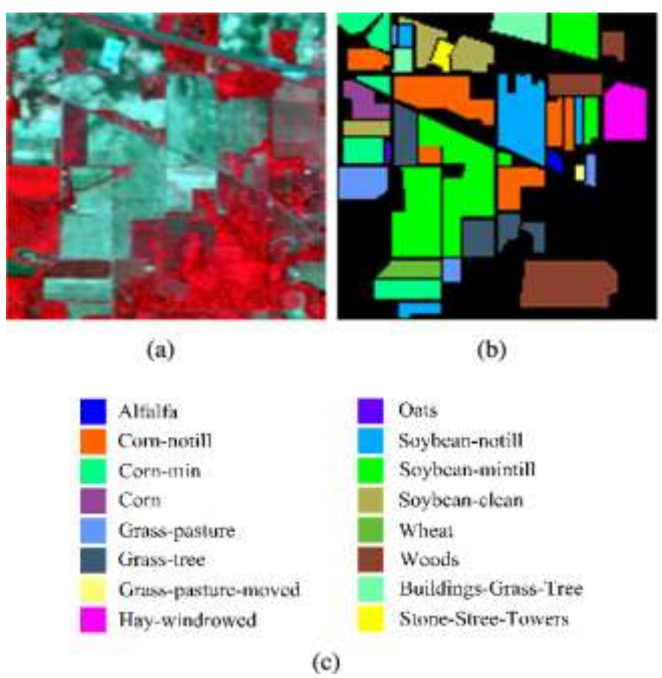

Fig 4: Indian Pines Image (a) Three Band False Colour Composite (b) Ground Tooth Image (c) Colour Code for 16 Classes

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

VII. RESULTS AND DISCUSSION

In this paper, the main aim is to enhance the accuracy of the HSI classification, by applying the neural network algorithm (KNN), thereby increasing the overall accuracies of the classification system[15]. Estimates of accuracy are:

1. OVERALL ACCURACY
2. AVERAGE ACCURACY
3. KAPPA COEFFICIENT

From all the observations it is certain that the accuracy of the classification is enhanced when DFFN with KNN is applied to the HSI classification compared to the remaining methods that were previously employed: i.e., we achieve OA=99.02%, AA=99.10%, KA=99.17% accuracy.

VIII. CONCLUSION

- The proposed KNN based algorithm has succeeded in taking advantage of spectral-spatial information.
- Introduction of deep feature fusion network (DFFN) and residual learning has upgraded several convolutional layers which eases the training of a deep network and benefit from increasing depth.
- The experimental result demonstrates the superiority of the proposed technique in terms of the classification map and quantitative metrics over other classical and neural networks. In future works, a further more effective feature extraction technique can be used to achieve greater accuracies.

REFERENCES

[1] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4186–4201, Aug. 2015.

[2] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[3] X. Kang, S. Li, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.

[4] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[5] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.

[6] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[7] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

[8] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.

[9] Weiwei Song, Shutao Li , Leyuan Fang, Ting Lu, "Hyperspectral Image Classification With Deep Feature Fusion Network," *IEEE Trans. on Geosci. and Remote Sens.*, Volume: 56 , Issue: 6, June 2018.

[10] Song, Ahram, Kim, Yongil , "Deep Learning-based Hyperspectral Image Classification with Application to Environmental Geographic Information Systems," *Korean Journal of Remote Sensing*, Volume: 33, Issue 6_2, Dec 2017.

DOI: 10.37200/IJPR/V24I5/PR2020214
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

[11] S. Chaib, H. Liu, Y. Gu, and H. Yao, "Deep feature fusion for VHR remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4775–4784, Aug. 2017.

[12] S. Li, T. Lu, L. Fang, X. Jia, and J. A. Benediktsson, "Probabilistic fusion of pixel-level and superpixel-level hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7416–7430,Dec. 2016.

[13] Y. Chen, X. Zhao, and X. Jia, "Spectral–spatial classification of hyper-spectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, Jun. 2015.

[14] P.Rajarajeswari etal "Impressive order Invention in Pattern Evolution for Text Mining ", *International journal of Latest trends in Engineering and Technology* ,vol 2,issue 4,July 2013 ISSN-2278-621X ,impact factor 0.685.

[15] P.Rajarajeswari etal "Consistency Evolution of Process models based on Structural Analysis and Behavioral Profiles ",is published in *International Journal of Modern Engineering Research (IJMER)*Vol.2, Issue.4, July-Aug. 2012 pp-2568- ISSN: 2249-6645,impactfactor 2.72