International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

# CLASSIFICATION OF HYPERSPECTRAL IMAGES USING CNN

**G Nithin[1], G Prabhas[2], G Punith[3], G Deepak[4]**

[1234] *Artificial Intelligence & Machine Learning, Malla Reddy University*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** The resilience, precision, and efficiency of spectral-spatial information-based algorithms have recently attracted increased attention. This work proposes an CNN-based classification algorithm that extracts features while taking into account both spectral and spatial information. The suggested method makes use of CNN to encode pixel's spectral and spatial data and also employed for classification tasks. Investigated is a clear contrast between the relative gain obtained with the inclusion of spatial features and its spectral counterpart. Three benchmark datasets—Indian Pines, Pavia University, and Salinas—were used in the investigation.

The proposed approach outperforms the classification methods K closest neighbours, linear discriminant analysis, Naive Bayes, and decision tree, according to experiments.

The classification of hyperspectral images (HSIs) has grown in popularity in the remote sensing community. broadly speaking.

Because of their reliability, precision, and efficiency, spectral-spatial information based algorithms are recently receiving greater attention. An CNN-based classification approach that extracts features while taking into account both spectral and spatial information has been developed in this study. The suggested technique uses CNN to encrypt the spectral and spatial information contained in each pixel and is also used to perform a classification task. It is also looked into how well the relative gain obtained by including spatial features compares to its spectral counterpart. The experiment was run using the Indian Pines, Pavia University, and Salinas benchmark datasets. It has been demonstrated through experiments that the suggested approach outperforms the classification techniques K closest neighbours, linear discriminant analysis, Naive Bayes, and decision tree.

In the area of remote sensing, categorization of hyperspectral images (HSI) has gained popularity.

## 1. INTRODUCTION

The problem of an unbalanced class distribution in hyperspectral data is one research gap. Classes in hyperspectral pictures frequently have different sample sizes, which can lead to inaccurate classification outcomes.

To solve this problem, oversampling or undersampling strategies must be developed, or the CNN algorithm must be changed to take imbalanced data into consideration. • The incorporation of spatial data into the classification process represents another area for future investigation. The accuracy of classification can be increased by using spatial information since hyperspectral images frequently contain spatially associated pixels. To do this, hybrid approaches that integrate CNN with other spatial classification methods, including Markov random fields or object-based classification, can be developed. An important job in classifying hyperspectral photographs is to detect and label the various types of materials or land cover that are present in the scene. Due to its prowess in handling high-dimensional data and nonlinear relationships, Convolutional Neural Network (CNN), a sophisticated machine learning method, have been extensively employed for the categorization of hyperspectral images. CNN divides the data into various groups while maximising the margin between them in order to function. • In this project, we want to create a hyperspectral image classification method that is accurate and effective. In this project, the data is preprocessed, its dimensions are reduced, and an CNN classifier is trained using labelled training data. Using test data, we will assess the classifier's performance and compare it to that of other cutting-edge systems.

## 2. RELATED WORK

In pattern recognition and picture classification, maximum likelihood is a very common parametric classifier. Compared to other conventional classification algorithms, it typically achieves higher classification accuracy. It is assumed that each band has a normal distribution and that the classes in the training samples have been carefully established. Finding the effective training pixels (exhaustively defined) for the identification of land cover classes in hyperspectral data with tens of hundreds of spectral bands is a difficult task. In contrast, the maximum likelihood classifier's classification accuracy is reliant on the carefully chosen training examples. Therefore, it is advisable to use a different multiclass classifier than the traditional one for hyperspectral data with weakly represented labelled training examples.

Compared to other conventional classification algorithms, it typically achieves higher classification accuracy. It is assumed that each band has a normal distribution and that the classes in the training samples have been carefully established. Finding the effective training pixels (exhaustively defined) for the identification of land cover classes in hyperspectral data with tens of hundreds of spectral bands is a difficult task. In contrast, the ML classifier's classification accuracy is reliant on the carefully chosen training examples. Therefore, it is advisable to use a different multiclass classifier than the traditional one for hyperspectral data with weakly represented labelled training examples.

It is assumed that each band has a normal distribution and that the classes in the training samples have been carefully established. Finding the effective training pixels (exhaustively defined) for the identification of land cover classes in hyperspectral data with tens of hundreds of spectral bands is a difficult task. In contrast, the ML classifier's classification accuracy is reliant on the carefully chosen training examples. Therefore, it is advisable to use a different multiclass classifier than the traditional one for hyperspectral data with weakly represented labelled training examples. One method is to use the Spectral Angle Mapper (SAM) methodology, which instead of comparing each pixel to the training pixels compares it to each endmember for each class

## 3. METHODOLOGY

Two steps make up the approach used in this investigation. Finding an accurate classifier should come after first extracting the best features suited for classifying different types of land cover. In this study, three classification techniques—Maximum Likelihood, Spectral Angle Mapper, and Convolutional Neural Network —have been proposed for comparison analysis.

### 3.1 Data pre-processing

Minimum Noise Fraction (MNF) is the solution used in this study to overcome the problem of dimensionality by choosing the suitable features for the classification of HYDICE sensor data. The HYDICE sensor data initially includes 191 hyperspectral bands. In order to further analyse the HYDICE sensor data for the classification of the land cover, MNF was utilised in order to reduce the dimensionality and computational needs.

### 3.2 Classification Techniques

The land cover classes from the HYDICE sensor data are classified in this study using three different classification approaches.

### 3.3 Maximum Likelihood

The parametric classifier Maximum Likelihood (ML) is based on the presumptions of normally distributed data for each class and an exhaustively chosen set of classes. Numerous research employed the ML classifier for classifying land cover classes as a benchmark to compare its classification accuracy to that of the other recently created classifiers [8]. It is regarded as a common method for thematic mapping using imagery from remote sensing systems. Rarely is the distribution's nature known in practical applications. Use non-parametric classifiers that don't make any assumptions whenever possible. ML classification is used to compare classification accuracies and confirm the applicability of the non-parametric classifier for categorising the land cover classes in this study.

### 3.4 Spectral Angle Mapper

A supervised classification system called Spectral Angle Mapper (SAM) uses spectral angular information to categorise hyperspectral image data. By comparing the picture spectrums to reference reflectance spectra and estimating their spectral similarity, it enables quick classification . Either field measurements or a picture obtained directly can be used to obtain the refrerance spectra.The reference spectra used in this investigation were directly extracted from images. By calculating the angle between the image and referance spectra and considering them as vectors in the n-dimensional feature space, SAM determines the degree of spectral similarity. High resemblance is indicated by smaller angles between the two spectra, and vice versa.This classifier is unaffected by factors related to sun illumination. Additionally, it has a strong classifier because to the influence of shading.

### 3.5 Convolutional Neural Network

Convolutional Neural Networks, or CNNs, are a particular kind of deep learning model made for processing and analysing structured grid-like input, such images or sequential data. In domains including image classification, object recognition, and image segmentation, it has achieved outstanding performance and is widely utilised in computer vision tasks.

The convolution technique, which involves sliding a tiny filter (sometimes referred to as a kernel) across the input data and performing element-wise multiplication and summing to build a feature map, is the fundamental idea behind CNNs. The network may extract pertinent features from the input data at various spatial locations using this process. The network can learn increasingly complicated features and hierarchies of representations by stacking many convolutional layers together.

### 3.6 Accuracy Assesment

Using the provided ground truth pixels, the confusion matrix was utilised to evaluate the accuracy metrics for each of the three classification methods. Additionally, the producer's accuracy (PA), user's accuracy (UA), and the Kappa coefficient (K) were closely scrutinised. The user's accuracy and producer's accuracy give details on the commission and omission mistakes related to the various classes, whereas the overall accuracy is the percentage of all validation pixels correctly identified. Kappa considers the potential of agreements arising by chance in a random categorization, in contrast to the total accuracy.

## 4. RESULTS

### 4.1 Training and Validation datasets

The ML classification considerably overestimated the Roof class and underestimated the Shadow class, as can be seen by visually comparing the classification results in Figure 2 to the original image in Figure 1. The SAM classification, on the other hand, vastly overestimated every class—aside from the Water and Grass divisions. This is because the reference spectra were extracted straight from the image. Due to the heterogeneous nature of the image and the fact that reference spectra did not account for sub-pixel information, the majority of the classes are incorrectly assigned. However, the CNN classifier managed to classify all of the land cover classes with a fair amount of accuracy.

### 4.2 Model Summary

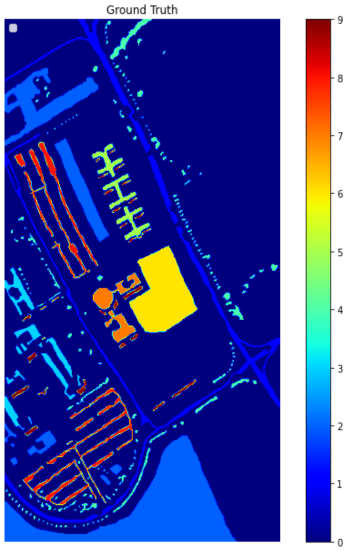
The remaining pixels in the hyperspectral image would then be classified using the trained CNN model, creating a thematic map of the land use and cover in the area of interest. The classification map could be subjected to post-processing methods like majority filtering or morphological procedures to increase accuracy and reduce noise. Metrics like overall accuracy, kappa coefficient, and confusion matrix would be used to assess the CNN classifier's performance. To evaluate the classification's accuracy and dependability, the results could also be contrasted with real-world data or other independent datasets.

Because it serves as a standard for evaluating CNN classifier performance, the accuracy of the ML classifier was calculated. The shadow class is the most important class for the analysis . In comparison to the SAM (41.67%) and SVM (37.50%), the classification result of ML for the shadow class reveals an omission error that is 93.75% greater; as a result, ML is underestimating the shadow class by ignoring the shadow pixels. Figure 3 helps you better visualise this. Table 2 also shows that the commission error for the roof class is larger, at 58.62%, than it is for the SAM (42.51%) and SVM (45.03), indicating that the roof class is significantly over estimated. Consequently, ML classifiers are used.

Finding a good classifier for the classification of land cover is very important for the proper monitoring of environmental changes. As a result, the confusion matrix for the SAM classifier is created for the comparison analysis. With the exception of the classes for grass and water, which have user accuracy of 86.79% and 97.18%, respectively.

**Figure -3**: Sample Outcome

Figure -3 refers to a few examples of hyper spectrical images found by our CNN-based model.




## 5. CONCLUSIONS

This work uses Maximum Likelihood (ML), Spectral Angle Mapper (SAM), and Convolutional Neural Network (CNN) to tackle the issue of classifying hyperspectral remote sensing data. The main goal taken into consideration was to find an efficient classifier by extracting the best features using MNF, appropriate for the land cover classification, in order to evaluate the efficacy of the land cover classification approaches. According to the results from the HYDICE sensor dataset, CNN is significantly more efficient than other traditional classifiers (such as the ML and the SAM classifier) in terms of classification accuracy, computational efficiency, and parameter setting stability.

## ACKNOWLEDGEMENT

We would like to thank everyone who offered their assistance and recommendations in developing this application. Our sincere gratitude goes out to all of the people whose assistance, motivating ideas, and support assisted us throughout the project's growth. a

## REFERENCES

[1] (2016). Zhu, X., Tuia, D. using support vector machines and texture features to categorise hyperspectral pictures. 119, 166–177, ISPRS Journal of Photogrammetry and Remote Sensing.

[2] (2017) Kumar, V., and Garg, P. K. categorization of hyperspectral images using support vector machines and geographical information. 55(4), 2195–2204 IEEE Transactions on Geoscience and Remote Sensing.

[3] Li, J., Zhang, L., & Wang, J. (2017). Hyperspectral image categorization employing combination spatial and spectral features on an SVM basis. Applied Remote Sensing Journal, 11(1), 015009.

[4] Huang, J., Li, H., Y., and Li (2018). Hyperspectral image categorization using SVM and spatial spectral data. 583. Remote Sensing, 10(4).

[5] Land cover mapping principles: a return to interpretative basics, King R B 2002. Global Journal of Remote Sensing 23 (2): 3546

[6] Fuentes (2016). Zhang, L., Zhang, J., & Li, J. Hyperspectral image categorization using an SVM and a composite kernel