Land Cover Classification of Satellite Imagery using Deep Learning

Vismaya Prakasan
Pillai College of Engineering
University of Mumbai
Plot No. 10, Sector 16, Panvel, Maharashtra

Romita Pawar
Pillai College Of Engineering
University Of Mumbai
Plot No. 10, Sector 16, Panvel, Maharashtra

Aditee Pachpande
Pillai College Of Engineering
University Of Mumbai
Plot No. 10, Sector 16, Panvel, Maharashtra

ABSTRACT

In the assessment of remotely sensed imagery, hyper-spectral (HSI) image classifications are commonly employed. The hyper-spectral image includes various image bands. The Convolutional Neural Network (CNN) is an extensively useful deep learning algorithm for data visualization and processing. Both the Spatial along with Spectral information are important for HSI classes to be effective. Due to the higher computing complexity, only a few approaches have used 3D CNN. Hybrid Spectral Convolutional 2D-3D Network (HybridSN) is instituted for HSI classing in this paper. HybridSN involves a spatial and spectral 3D-CNN which is then trailed by a spatial 2D. A study of more abstract level spatial representation will continue with 2D-CNN over 3D-CNN. Furthermore, when compared to conventional CNNs, the employment of hybrid CNNs lessens the model's complexity. To see if this hybrid method works, a thorough HSI phase test was performed over Indian Pines, Salinas and Pavia University and results compared to ground truth. The presented HybridSN HSI classification model provides with best result.

General Terms

Convolutional Neural Network, Hybrid Spectral Network

Keywords

Hyper-spectral Image, Remote Sensing, Overall Accuracy, Hybrid Spectral Network, Principle Component Analysis, Confusion Matrix

1. INTRODUCTION

Mapping the quilt of land and land use (LULC) using the long distance sensory records has been a task for a long time because of diverse capabilities and mixtures, especially in all exclusive city areas.LULC facts are extensively used in Geographic facts device (GIS) applications. Accordingly, correct, high-quality maps and precise statistics are wanted anywhere and whenever. Despite the fact that images and remote sensing techniques,in addition with subject exploration, have been key resources for figuring out the land use characteristics, the inner dimensions of neighbourhood statistics based totally information series for one's capabilities have constantly been a hard step in terms of time, money, and so forth because of the reliability of the data.

Image classification is a rapid method and the most commonly used one in areas of computer vision incorporated with Hyper-spectral Image Classification (HSI) for remote sensing image survey. The Convolutional Neural Networks (CNN) are one of the most broadly utilized deep learning-based computing algorithms. HSI is made up of different image bands. Convolutional neural networks (CNNs) have achieved great success in the field of image classification over the past few years because of their fast and very precise feature extraction skills and continuously qualified network frameworks. CNN architecture comprises a series of hidden layers, with input and an output layer. The dissimilarity is that whereas the result of the convolutional computation is an attribute of the picture, the source of the CNN is an image (pixel matrix).

2. LITERATURE SURVEY

Yanhui Guo et al.(2019): To address the issue of spatial aspect extraction in spectral and spatial classification of the HSI, the researchers developed the guided filter's-based technique with Support Vector Machine. They began by extracting the guided filter's spatial information from the original image, which they did using a Principal Component Analysis technique. The second step was classifying spatial features using SVM. Finally, the classification was improved by using a guided filter once more. This technique was applied to the Indian Pine (IP), Salinas (SA), and Pavia Universilty (PU) data-sets. Calculation time along with implementation complexity was decreased by this technique. (1) Using a guided filter, spatial features are effectively extracted from HSI. (2) The guided filter's extracted feature is sufficient for HSI classification even in the absence of original data. (3) The SVM approach with two filtration provides a quick and accurate way to categorise HSI.[1]

Shambulinga M, G. Sadashivappa(2019):In this paper, Indian pine data-set is used. They put 10% of the training sample for each of the 16 classes. The classification process was done using support vector machine (SVM) with different kernels and filters. Using SVM, they could accomplish an overall accuracy (OA) close to 81.15%. Further to improve the OA, they used the SVM post-guide filter, thereby increasing their OA by 13.05% making their OA to about 89% which they got from RBF Kernel and to reduce computation time they used principal component analysis, calling the whole project 2 hours 30 minutes.[2]

Alou Diakite et.al (2021): In this article, they worked on four data-sets and for training purposes, 20% of the sample data was taken. The four records were Indian Pines, KSC, UP, Salinas. The optimizer they used was RMSprop and they trained their model for 50 epochs. Several window sizes were chosen for each of the four data sets. However, due to hardware limitations, they could not work with a 15x15 window size. The best OA was obtained with HybridSN. By using HybridSN, they were able to increase accuracy by 6% on 3 of the four data-sets compared to traditional methods. They also came to an important conclusion that as the training sample increased, the accuracy increased, but when they increased the window size to a threshold of 11x11, the accuracy for IP and Salinas decreased. The paper also notes that their model could have been much more accurate had more training data been included.[3]

Junru Yin et. al(2021): A framework was initiated for a unified network by using a Bi-LSTM 3D CNN network. Bi-LSTM network served them as a spectral feature extractor for HSI Classification. They basically used three data-sets, wherein the Indian Pines data-set using 10% and was chosen arbitrarily for labelled sections of individual classes and that too in ratio 1:9, for the respective training part and the testing part. Here the overall accuracy which they gained was 95.41% on 10×10 window and the average accuracy was illustrated as 98.50 whereas, in the Pavia University and Salina Valley data-sets, they received the model having less training data i.e., 5%. They gained 96.35% again an overall accuracy on 10×10 window and the average accuracy was obtained as 99.22% whereas an OA i.e., overall accuracy gain was 95.0% and A.O. average accuracy gain was 99.83% for respective data-set. By comparing with the model, the researchers got to know that the Bi-LSTM CNN method could obtain superior classifying outcomes than the 3D-CNN.[4]

Douglas Omwenga Nyabuga(2021): The early layers of the proposed Hybrid network and transfer learning model uses 3D convolution for extracting information (Spatial-Spectral) the 2D convolutional layers are added on top to primarily handle semantic abstraction. In this Hybrid-SN model, the ResNeXt-50 block was used in the model to make the network for image classification simple and extremely modular and 3D convolutions to extract spatial-spectral information. Prior to feeding the network with data from HSI's, principal component analysis (PCA) was enforced to find optimal orthogonal vectors for transferring the data. This enhanced the separability inside the classes and the balance of the inter-class and also the intra-class criteria. The experimental finding demonstrates that the model can effectively enhance the classification of HSI while also bringing forth an immediate representation of the spectral-spatial data.[5]

2.1 Inference of Literature Review

- After going through a number of papers from 2019 to 2021 on topic of Land Cover Classification it was inferred that most of the researchers have used a lot of data-sets and methods.
- The Deep Learning techniques mostly used till date are SVM and Simple CNN but, the latest papers used 1Dimensional CNN, 2Dimensional CNN and 3Dimensional CNN.
- Few papers have encountered a high accuracy but with the old traditional methods i.e., SVM and CNN.

Table 1. Summary of Literature Survey

Sr. No.	Year of Publication	Research Paper Details	Techniques Used	Performance Analysis
1.	2019	Hyperspectral image classification with SVM and guided filter[1]	SVM	IP-81.02%, PU-93.36%, SA-92.21%
2.	2019	Hyperspectral image classification using 3D 2D CNN[2]	3D-2D CNN	IP-95.78%, SA-99.71%
3.	2021	A 3D-2D Convolutional Neural Network and Transfer Learning for Hyperspectral Image Classification[3]	3D-2D CNN	IP-98.78%, PU-98.80%, SA-99.00%
4.	2021	Hyperspectral Image Classification using Convolutional Neural Networks[4]	SVM, CNN	91.28%, CNN-98.28%
5.	2021	Spatial-Spectral Network for Hyperspectral Image Classification: A 3-D CNN and Bi-LSTM Framework [5]	3D CNN	IP-95.41%, PU-96.35%, SA-95.01%

3. DATA-SET DESCRIPTION

Description along with their true labelled colour coded classes for the three data-sets used are discussed in this chapter.

3.1 Indian Pines

Indian Pine data-set is obtained through AVIRIS sensor in the region of Indiana. It has spatial resolution of about 20 meters, and it consists of 145×145 spatial dimensions with two hundred and twenty-four spectral bands.Twenty-four spectral bands having area under water absorption are removed. Hence, it forms three dimensional data of size $145 \times 145 \times 200$. There are 16 different types of agricultural crops that are planted in the area.

3.2 Salinas

The Salinas scene data-set is obtained through the AVIRIS sensors having 224 bands in the Salinas Valley, having pixels of about 3.7 meters. The covered land is 512 x 217 in acres in size. 20 absorption bands were destroyed. This image is only available in the form of At-sensor radiation data. It has veggies, barren ground with vines. Salina has about 16 classes.

3.3 Pavia University

The data-set of Pavia (PU) is a HSI data-set gained from (ROSIS-03) Reflective Optics System Imaging Spectrometer in Pavia region of Italy. The data-set comprises of a pixel size of 610 x 340 with 115 bands. The images are divided into 9 classes with a total of 42,776 labelled patterns such as asphalt, meadow, gravel, trees, tin, bare earth, bitumen, bricks and shades.

4. EXPERIMENT AND DISCUSSION

Information about training and classification results for each of the three used data-sets along with the implemented steps and results for the confusion matrix is showcased in this chapter.

4.1 Implemented Steps:

(1) **Data Collected and loaded the data set:** Data-set Indian Pines i.e. IP, Salinas which is also known as SA and Pavia University i.e. PU were initially downloaded; later loaded into jupyter.

(2) **Data Pre-processing:** In this step, a technique called PCA, i.e., Principal component analysis was used. This is an approach whereby the dimensionality of the data is reduced by conserving most of the feasible information involved in the ground truth as much as possible. To do this, the given data are convolved onto a subspace having low dimension which keeps the maximum value between the data points.

(3) **Padded with zero:** Zero padding is an approach in which the original input size is maintained. This is specified for each of the convolution layers. Zero padding is done to introduce a zero valued pixel on the input image's perimeter. It adds/pads zeros to the border of the image, hence known as zero padding.

(4) **Divided the data into Training Testing part:** The data were divided into training and testing part; so, the data-set was divided in the ratio of 70:30.

(5) **Created patches:** The CNN kernel / filter processes only one patch at a time and not the entire image. This is because the filter is designed to process small pieces of the image to the detect features (e.g., edges). It also has good regularisation properties due to the small number of parameters it estimates. These parameters should always be "good" in many areas of each image, and in many areas of all other training images. Patches are input to the kernel.

(6) **Created the model:** Then CNN / convolutional neural network model is created using Keras in TensorFlow.

(7) **Tested our model:** The first parameter is of course, the data-set (training set) that trains the model. This corresponds exactly to the second parameter, which specifies the validation data (test set) here. This is a set that evaluates CNNs. Here the test set is the ground truth model.

(8) **Evaluated Accuracy:** Lastly, the epochs parameter, which is nothing but the number of epochs, was chosen as 10 epochs to converge the accuracy not only on the testing set but also on training set.

(9) **Visualized Result:** Thus,after all these steps we were able to visualize the results.

4.2 Confusion Matrix

Lastly, the epochs parameter, which is nothing but the number of epochs, here it was chosen as 10 epochs to converge the accuracy not only on the testing set but also on training set.

By looking at a confusion matrix (CM), one can easily observe the predicted labels present on abscissa and the true/original labels on the ordinate. Here is a confusion matrix chart for the three data-sets that we have performed on. The very first confusion matrix i.e., fig 1, depicts the chart of Indian pine data-set wherein it shows the prediction for the 16 classes ranging from Alfalfa to the Woods. The second confusion matrix chart, i.e., fig.2, indicates the chart of Pavia University data-set wherein; it shows us the prediction for the 16 specific classes same as the previous data-set which ranges from Asphalt to Shadows. Whereas, the last confusion matrix reveals the prediction for 9 classes starting from Broccoli to Vineyard of Salinas Data-set as shown in fig 3.

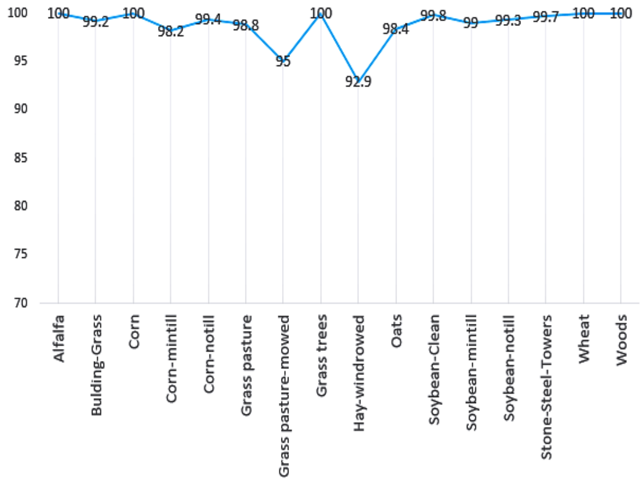

Fig. 1. Confusion Matrix of Indian Pine

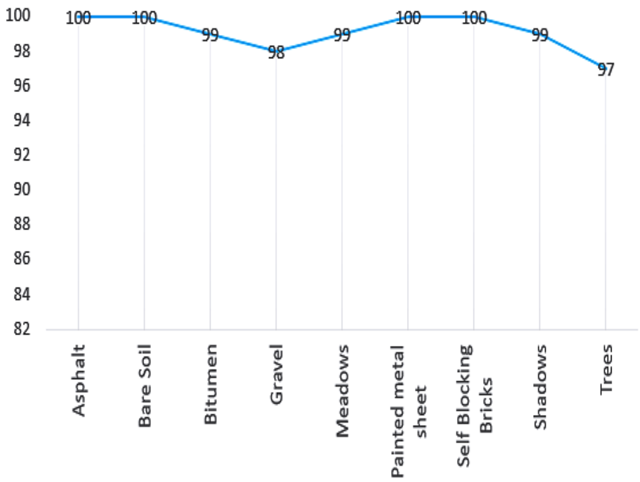

Fig. 2. Confusion Matrix of Pavia(PU)

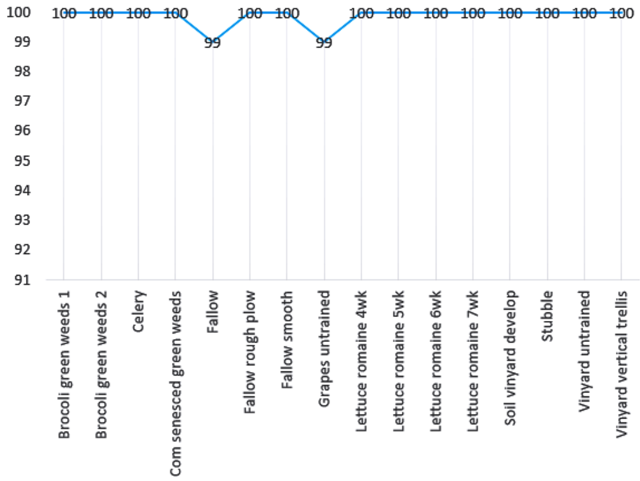

Fig. 3. Confusion Matrix for Salinas

4.3 Ground Truth vs Predicted Image

The Indian Pines Data-set's predicted image is respectively depicted below in fig.4 and its colour code in fig.5. The predicted image of Indian Pine Data-set was derived with an accuracy of 98.37%.

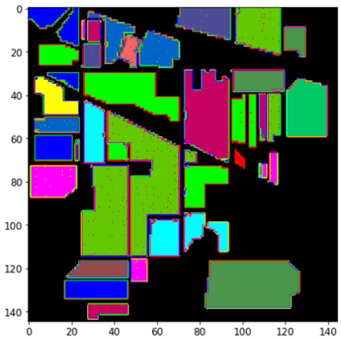

Fig. 4. Indian Pines Predicted Image

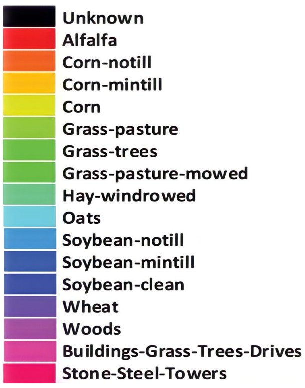

Fig. 5. Colour Code for Indian Pines

The Salinas Data-set's predicted image is shown in fig.6 and its colour code in fig.7. The predicted image of Salinas Data-set was derived with an accuracy of 99.37%.

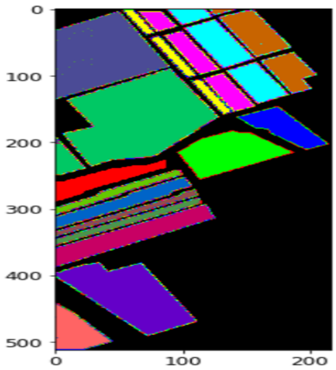

Fig. 6. Salinas Predicted Image

Fig. 7. Salinas Colour Code

Pavia University's predicted image is shown in fig 8 and its colour code in fig 9. The predicted image of Pavia University was derived with an accuracy of 99.08%.

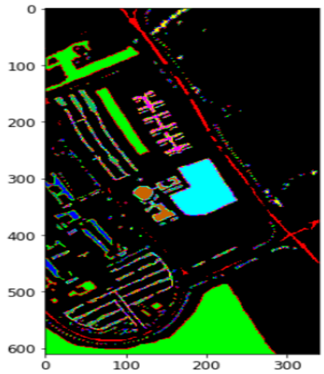

Fig. 8. Pavia Predicted Image

Fig. 9. Pavia University Colour Code

4.4 epoch vs accuracy for 0.01 learning rate

Figure 10 describes the graph of the Indian Pines Data-set for epoch Vs accuracy by keeping learning rate constant at 0.001 with a decay of weight 1e-06 respectively. Here, Adam optimizer was employed. Most of the time Adam optimizer is used. When there are less training samples, over-fitting is triggered. Each experiment was run for three epochs that are 5, 10, 20.

It can be seen that accuracy is affected after changing the number of epoch. The accuracy at 5 epoch is 35.91% and at 20 it is 97%. Moreover, the highest percentage of accuracy obtained is approximately 97.55% at 10 epochs.

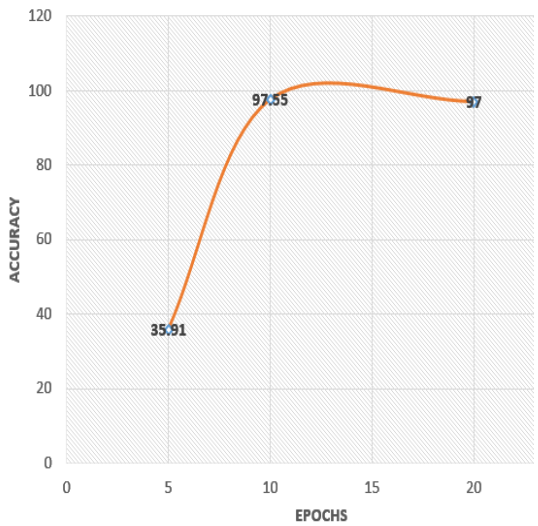

Fig. 10. epoch vs accuracy for learning rate 0.01. (Without GPU)

The figure 11 describes the graph for epoch vs accuracy now by alternating the learning rate at 0.002 for 5, 10, 20 number of epochs. A gradual rise can be seen in the percentage of accuracy as the number of epochs are increased. The highest percentage is obtained at 20 epochs having accuracy of 99.21% while the least is obtained at 5 epoch that is 76.22%.

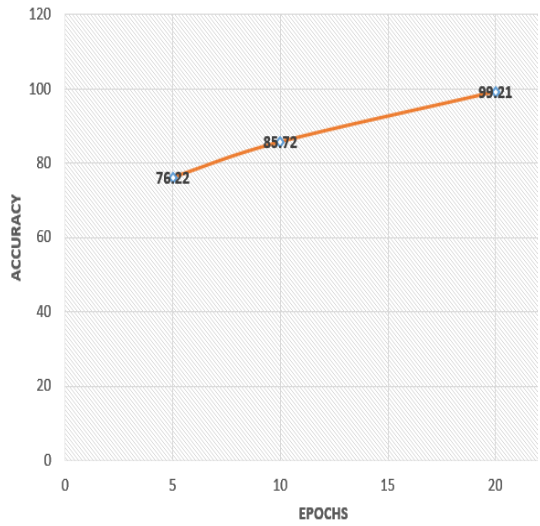

Fig. 11. epoch vs accuracy for learning rate 0.02. (Without GPU)

In conclusion, a smaller learning rate and more training epochs gives the highest accuracy. The graphs up top compare the amount of time it took to finish the experiment across all three data-sets.

Without a GPU, the Time vs. Data-set graph is shown in fig. 12. Where the hardware requires 120 minutes to finish the experiment. Due to large data size of the Salinas and Pavia University data-sets, the training was not completed. All three data-sets successfully finished training after using the GPU. The time taken by Indian Pines was 52 minutes, Salinas was 60 minutes, while the University of Pavia needed 40 minutes.

Processor AMD Ryzen Threadripper 2970w x 24 core x 48, Memory 125.7GB, Graphics NVIDIA, and Disk Capacity 2 were all present in the GPU that was used.

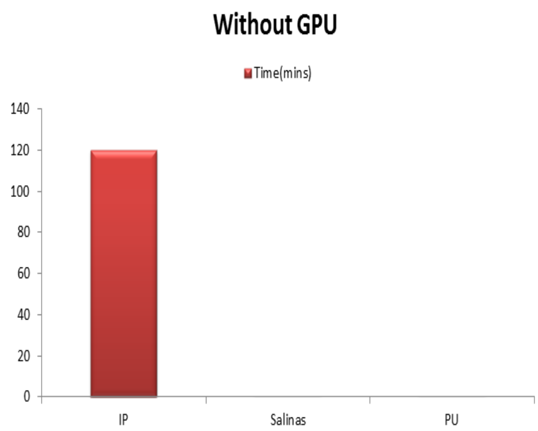

Fig. 12. Time(training) VS Data-set

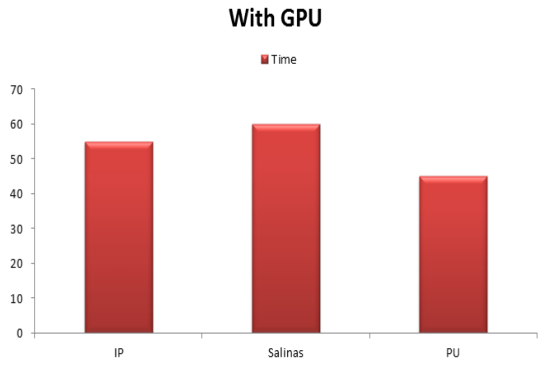

Fig. 13. Time VS Data-set

4.5 Layer wise Summary

The below table 2 and 3 illustrates layer wise summarizing of the model proposed which uses 2D-3D convolutional neural network having window size of 25 x 25 for all the three used data-sets.

Table 2. Layer wise Summary of Indian Pines Data-set & Salina Data-set

Layer/Type	Shape of Output	Parameters
input_1 (Input Layer)	[(0, 25, 25, 30, 1)]	0
convo_3d(Convo3D)	(0, 23, 23, 24, 8)	512
convo_3d_1(Convo3D)	(0, 21, 21, 20, 16)	5776
convo_3d_2(Convo3D)	(0, 19, 19, 18, 32)	13856
re-shape(Reshape)	(0, 19, 19, 576)	0
convo_2d(Convo2D)	(0, 17, 17, 64)	331840
flaten (Flaten)	(0, 18496)	0
dense (Dense)	(0, 256)	4735232
drop-out (Dropout)	(0, 256)	0
dense_1 (Dense)	(0, 128)	32896
drop-out_1 (Dropout)	(0, 128)	0
dense_2 (Dense)	(0, 17)	2065
Total Trainable Parameters: 5,122,177		

Table 3. Layer wise Summary of Pavia Data-set

Layer/Type	Shape of Output	Parameters
input_1 (Input Layer)	[(0, 25, 25, 30, 1)]	0
convo_3d(Convo3D)	(0, 23, 23, 24, 8)	512
convo_3d_1(Convo3D)	(0, 21, 21, 20, 16)	5776
convo_3d_2(Convo3D)	(0, 19, 19, 18, 32)	13856
re-shape(Reshape)	(0, 19, 19, 576)	0
convo_2d(Convo2D)	(0, 17, 17, 64)	331840
flaten (Flaten)	(0, 18496)	0
dense (Dense)	(0, 256)	4735232
drop-out (Dropout)	(0, 256)	0
dense_1 (Dense)	(0, 128)	32896
drop-out_1 (Dropout)	(0, 128)	0
dense_2 (Dense)	(0, 9)	2064
Total Trainable Parameters: 5,121,273		

5. CONCLUSION

A 2D-3D hybridization process was used for the hyperspectral image classification. Thus, by using three different datasets, the method was found to be more accurate and faster than traditional methods. This hybridization technique works well on both small and large datasets. For dimensionality reduction, PCA was used. Additionally, Adam optimizer was used in this case to reduce the drop.

The larger the training sample, the longer the training time and the shorter the testing time, so 30 percent training was chosen for the project. The test results were compared by changing the learning rate. With the help of the GPU, it was able to significantly reduce the training time when working with a large dataset, which also helped to get a good accuracy score.

6. REFERENCES

[1] Yanhui Guo, Xijie Yin, Xuechen Zhao, Dongxin Yang, and Yu Bai. Hyperspectral image classification with svm and guided filter. *EURASIP Journal on Wireless Communications and Networking*, 2019(1):1–9, 2019.

[2] M Shambulinga and G Sadashivappa. Hyperspectral image classification using support vector machine with guided image filter. *IJACSA*, 10:271–276, 2019.

[3] Alou Diakite, Gui Jiangsheng, and Fu Xiaping. Hyperspectral image classification using 3d 2d cnn. *IET Image Processing*, 15(5):1083–1092, 2021.

[4] Junru Yin, Changsheng Qi, Qiqiang Chen, and Jiantao Qu. Spatial-spectral network for hyperspectral image classification: A 3-d cnn and bi-lstm framework. *Remote Sensing*, 13(12):2353, 2021.

[5] Douglas Omwenga Nyabuga, Jinling Song, Guohua Liu, and Michael Adjeisah. A 3d-2d convolutional neural network and transfer learning for hyperspectral image classification. *Computational Intelligence and Neuroscience*, 2021, 2021.

[6] Xudong Kang, Shutao Li, and Jon Atli Benediktsson. Spectral–spatial hyperspectral image classification with edge-preserving filtering. *IEEE transactions on geoscience and remote sensing*, 52(5):2666–2677, 2013.

[7] Xinyu Lei, Hongguang Pan, and Xiangdong Huang. A dilated cnn model for image classification. *IEEE Access*, 7:124087–124095, 2019.

[8] D Jeevalakshmi, S Narayana Reddy, and B Manikiam. Land cover classification based on ndvi using landsat8 time series: a case study tirupati region. In *2016 International Conference on Communication and Signal Processing (ICCSP)*, pages 1332–1335. IEEE, 2016.

[9] Lei Ma, Tengyu Fu, Thomas Blaschke, Manchun Li, Dirk Tiede, Zhenjin Zhou, Xiaoxue Ma, and Deliang Chen. Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers. *ISPRS International Journal of Geo-Information*, 6(2):51, 2017.

[10] Bing Liu, Anzhu Yu, Xibing Zuo, Zhixiang Xue, Kuiliang Gao, and Wenyue Guo. Spatial-spectral feature classification of hyperspectral image using a pretrained deep convolutional neural network. *European Journal of Remote Sensing*, 54(1):385–397, 2021.

[11] Miae Kim, Junghee Lee, Daehyeon Han, Minso Shin, Jungho Im, Junghye Lee, Lindi J Quackenbush, and Zhu

Gu. Convolutional neural network-based land cover classification using 2-d spectral reflectance curve graphs with multitemporal satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(12):4604–4617, 2018.

[12] Fuding Xie, Fangfei Li, Cunkuan Lei, and Lina Ke. Representative band selection for hyperspectral image classification. *ISPRS International Journal of Geo-Information*, 7(9):338, 2018.

[13] Hayder Hasan, Helmi ZM Shafri, and Mohammed Habshi. A comparison between support vector machine (svm) and convolutional neural network (cnn) models for hyperspectral image classification. In *IOP Conference Series: Earth and Environmental Science*, volume 357, page 012035. IOP Publishing, 2019.

[14] Alexander FH Goetz. Three decades of hyperspectral remote sensing of the earth: A personal view. *Remote Sensing of Environment*, 113:S5–S16, 2009.

[15] SAROJ KUMAR. *" Jnana Sangama" Belgaum-590018*. PhD thesis, Visvesvaraya Technological University, 2012.

[16] Bing Lu, Phuong D Dao, Jiangui Liu, Yuhong He, and Jiali Shang. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, 12(16):2659, 2020.

[17] Yushi Chen, Hanlu Jiang, Chunyang Li, Xiuping Jia, and Pedram Ghamisi. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10):6232–6251, 2016.

[18] Ava Vali, Sara Comai, and Matteo Matteucci. Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*, 12(15):2495, 2020.

[19] Yanan Luo, Jie Zou, Chengfei Yao, Xiaosong Zhao, Tao Li, and Gang Bai. Hsi-cnn: a novel convolution neural network for hyperspectral image. In *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, pages 464–469. IEEE, 2018.

[20] Gavneet Singh Chadha, Jan Niclas Reimann, and Andreas Schwung. Generalized dilation structures in convolutional neural networks. In *ICPRAM*, pages 79–88, 2021.

[21] Ying Li, Haokui Zhang, Xizhe Xue, Yenan Jiang, and Qiang Shen. Deep learning for remote sensing image classification: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(6):e1264, 2018.