CPS580 - Convolutional Neural Network for Hyperspectral Image

Ashish Pushkar

Central Michigan University, Computer Science, Mt. Pleasant, 48858, USA

pushk1a@cmich.edu

ABSTRACT

Hyperspectral imaging, like other spectral imaging, collects and processes information from across the electromagnetic spectrum. The goal of hyperspectral imaging is to obtain the spectrum for each pixel in the image of a scene, with the purpose of finding objects, identifying materials, or detecting processes. Convolutional Neural Network (CNN) is one of the most utilized deep learning architecture for visual data processing. Most of these methods are on 2-D CNN. But the HSI arrangement is subject to both spatial and spectral information. Since, 3-D CNN increases the computational complexity because of high dimensionality and limited amount of labelled data from Indian Pines, Pavia Centre and University and Salinas, not many methods have utilized it. To address these challenges, this paper proposes a deep learning architecture using Hybrid Spectral Convolutional Neural Networks which has a spectral-spatial 3D-CNN followed by spatial 2-D-CNN. The 2D-CNN will learn spatial representation in more conceptual way after 3D-CNN utilizes the stack of spectral bands to give joint spatial-spectral feature. The utilization of mixed 3D and 2D CNN decreases the complex nature of the model in contrast with only 3D-CNN model. The performance of the model is compared with end-to-end deep learning models and the performance is exceptionally well using Hybrid Spectral CNN.

Introduction

Hyperspectral imaging (HSI) is a technique that analyzes a wide spectrum of light instead of just assigning primary colors (red, green, blue) to each pixel. The light striking each pixel is broken down into many different spectral bands in order to provide more information on what is imaged. The algorithms and the image processing methodologies associated with HSI are a product of military research, and were primarily used to identify targets and other objects against background clutter. In the past, HSI has seen civil applications, and has particularly been useful in satellite technology. It might become an inexpensive, promising, and quick tool for the assessment of tissue conditions at diagnosis and during surgery. The medical applications include forensics, detection of colorectal and gastric cancer or ulcers.

The examination of Hyperspectral images (HSIs) are significant as a result of its applications, all things considered, [1]. Hyperspectral imaging brings about different groups of pictures and creates a major volume of information that make it difficult to examine. The otherworldly and the spatial connection between's various groups pass on helpful data about the location of interest. As of late, Camps-Valls et al. [2] have reviewed the advances in HSI grouping.

Convolutional neural networks(CNN) have indicated potential in hyperspectral imagery arrangement as they utilize broad quantities of parameters for feature learning [4]. Notwithstanding the probability, high changeable signature properties of HSIs muddles the relating CNN plans; time devouring and costly manual naming of HSIs has restricted the quantity of training samples. These issues have discouraged and decreased the prescient intensity of CNN models.

Existing CNN models for HSI arrangement are generally founded on one dimensional or two dimensional CNN structures. The more seasoned arrangement of models is utilized for spectral feature learning [5], while the more current arrangements of models search for nearby spatial feature learning [6]. These models don't perform very well in feature extraction on various measurements. Three-dimensional CNN is explored by specialists [7] and it is all the more computationally intricate and this by itself appears to perform surprisingly more dreadful. Customary machine learning strategies including Logistic Regression and Kernel-based SVM [8] [9] are proposed by experts yet it appears to be that arrangement, precision are substandard and less ideal.

In this paper, I propose a Hybrid Spectral CNN architecture that handles the limitations of the past models. This model exploits the programmed highlight learning capacity of both 2-D and 3-D CNNs. Basically, the proposed model plays out a mix of 3D-CNN and 2D-CNN followed by a fully associated layer. We have diminished the spectral redundancy from the data. At that point, the data is sent to 3 layers of 3D-CNN, and afterward, the data is sent to 1 layer of 2D-CNN. At last, the output is sent to fully connected layers with dropout to forestall over-fitting. My architecture is demonstrated to be vigorous and stable where the model is more exact contrasted with BASS Net and 2D-CNN [10].

HSIs contain spectrum data for every pixel in the image of a scene, where each spatial pixel is a spectral vector made out of many adjacent restricted electromagnetic groups reflected by the recognizing materials. HSI classification includes caterorical class name to each unlabelled pixel dependent on the relating ghastly and additionally spatial component. With the coming of new hyperspectral far off detecting instruments and their expanded resolutions, the measure of high dimensional hyperspectral datra is expanding. This outcome in viable and hypothetical issues because of the high dimensionality where conventional algorithms created for multi-spectral images may not, at this point be reasonable.

Methods

Data

The data utilized in this experiment are freely accessible HSI data sets, named Indian Pines (IP), University of Pavia (UP), and Salinas Scene (SA), assembled from [http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes]. The Indian Pines data set has images with 145 X 145 spatial measurements and 224 spectral bands in the frequency scope of 400 to 2500 nm, out of which 24 spectral bands covering the district of water retention have been disposed of. The ground truth accessible is assigned into 16 classes of vegetation. The University of Pavia data set comprises of 610 X 340 spatial pixels with 103 spectral bands in the frequency scope of 430–860 nm. The ground truth is partitioned into nine metropolitan land-cover classes. The Salinas data set contains the images with 512 X 217 spatial measurements and 224 spectral bands in the frequency scope of 360–2500 nm. The 20 water-engrossing spectral bands have been disposed of. Altogether, 16 classes are available in this data set.

Each HSI pixel in Original Input contains various spectral bands measures and structures a one-hot encoded vector. The hyperspectral pixels show the blended land-cover classes, presenting the high intraclass fluctuation and interclass comparability into the Original Input. It is a test for any model to handle this issue. To eliminate the spectral redundancy, the conventional principal component analysis (PCA) is applied over the first HSI data alongwith spectral bands. The PCA decreases the quantity of spectral bands, maintaining spatial dimensions. Just spectral bands are diminished with the end goal that it safeguards the spatial data which is significant for perceiving an data.

For utilizing image classification strategies, the HSI data is partitioned into 3D patches, trust labels of that are chosen by the label of the focused pixel.

Development Environment

The experiments are conducted using Python development environment, CoLab that runs in the browser using Google Cloud.The environment supports Keras stack. The experiment is done using a GPU runtime. To make a fair comparision, spatialdimension was extracted in 3D patched of input volume, such as 25X25X30 for Indian Pines and 25X25X15 for University ofPavia and Salinas Scene.

Training and Evaluation Protocol

The input data are convolved with 2-D kernels in 2D-CNN. The convolution occurs by calculating the sum of the dot product between input data and kernel. The kernel is stridden over the data to cover the spatial dimension. The convolved features are gone through the activation function to present the nonlinearity in the model.

The 3-D convolution [11] is finished by convolving a 3-D kernel with the 3-D data. In the proposed model for HSI data, the component guides of convolution layer are produced utilizing the 3-D kernel over numerous bands in the input layer, this catches the spectral data.

The boundaries of CNN, for example, the bias and the kernel weight, are normally prepared utilizing directed methodologies [12] with the assistance of gradient descent. In 2-D CNNs, the convolutions are applied over the spatial measurements just, covering all the feature maps of the past layer, to process the 2-D feature maps. Again, for the HSI classification, it is better to catch the spectral information encoded in various groups alongside the spatial data. The 2-D-CNNs can't deal with the spectral information. Then again, the 3-D-CNN kernal can separate the spectral and spatial feature all the while from HSI data however at the expense of expanded computational intricacy. So as to exploit the feature learning ability of both 2-D and 3-D CNNs, I propose a mixed learning framework called Hybrid Spectral CNN for HSI. The flow outline of the proposed Hybrid Spectral CNN network is appeared in Fig. 1. It contains three 3-D convolutions, one 2-D convolution, and three fully connected layers.

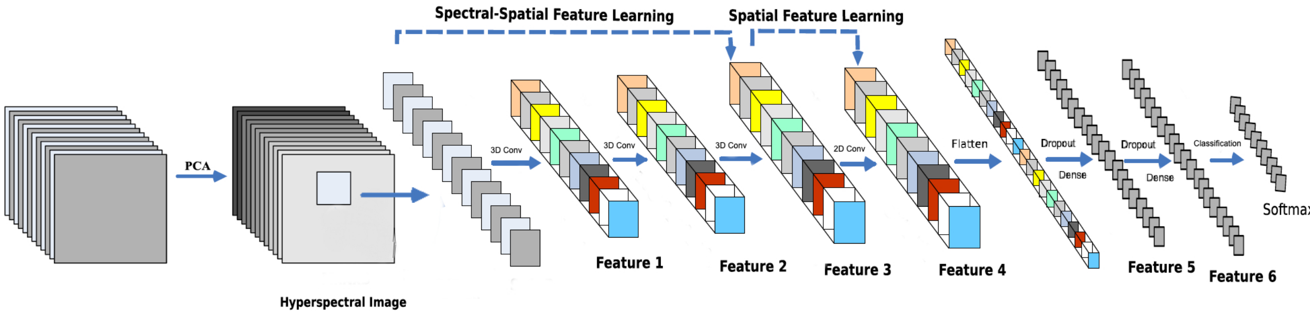

Fig. 1. Proposed Hybrid Spectral CNN model that integrates 3-D and 2-D convolutions for HSI classification.

In the Hybrid Spectral CNN framework, the components of 3-D convolutional kernels are 8X3X3X7X1, 16X3X3X5X8, and 32X3X3X3X16 in the resulting first, second, and third convolution layers, separately, where 16X3X3X5X8 methods 16 3-D-kernels of measurement 3X3X5 for every one of the eight 3-D input feature maps. Then again, the element of the 2-D convolution part is 64X3X3X576 where 64 is the number of 2-D-kernels, 3X3 tells the spatial component of 2-D-kernel, and 576 is the quantity of 2-D feature maps. To expand the quantity of spectral spatial feature maps, 3-D convolutions are applied threefold and can protect the spectral information from HSI data in the yield volume. The 2-D convolution is applied once before the flatten layer by remembering that it firmly segregates the spatial data inside various otherworldly groups without significant loss of spectral data, which is significant for HSI data. A summary of the proposed model as far as the layer types and yield map measurements are given in Table 1. The quantity of nodes in the last dense layer is 16, which is equivalent to the number of classes in the Indian Pines. In this manner, the absolute number of boundaries in the proposed model relies upon the number of classes in a data set. The absolute number of teachable weight boundaries in Hybrid Spectral CNN is 5, 122, 176 for the Indian Pines. All weights are arbitrarily introduced and prepared to utilize a back-propagation algorithm with the Adam optimizer by utilizing the softmax loss. We utilize small batches of size 256 and train the organization for 100 epochs.

Table 1

SUMMARY OF THE PROPOSED ARCHITECTURE

Layer	Output Shape
Input Layer	(25, 25, 30, 1)
Conv3D	(23, 23, 24, 8)
Conv3D	(21, 21, 20, 16)
Conv3D	(19, 19, 18, 32)
Reshape	(19, 19, 576)
Conv2D	(17, 17, 64)
Flatten	(18496)
Dense	(256)
Dropout	(256)
Dense	(128)
Dropout	(128)
Dense	(16)

In this, we have utilized the overall accuracy and average accuracy to pass judgment on the HSI classification. The consequences of the proposed Hybrid Spectral CNN are contrasted and the most broadly utilized supervised methods, for example, 2-D-CNN[13], 3-D-CNN[14], and multi-scale 3-D deep convolutional neural network (M3D)- CNN[15]. The 30% and 70% of the data are partitioned into training and testing data.The code is taken from "https://github.com/eecn/Hyperspectral-Classification" to process the outcomes.

The outcomes regarding the Overall Accuracy and Average Accuracy for techniques referenced appear in Table 2. From Table 2, we can see that the Hybrid Spectral CNN performs outstandingly well when contrasted with different strategies on each data set.

Table 2						
CLASSIFICATION ACCURACY (IN PERCENTAGES) ON VALIDATION DATA SETS USING DIFFERENT METHODS						
——	Indian Pines Dataset		University of Pavia Dataset		Salinas Scene Dataset	
Methods	Overall Accuracy	Average Accuracy	Overall Accuracy	Average Accuracy	Overall Accuracy	Average Accuracy
2D-CNN	89.48	86.14	97.86	96.55	97.38	98.84
3D-CNN	91.10	91.58	96.53	97.57	93.96	97.01
M3D-CNN	95.32	96.41	95.76	95.08	94.79	96.25
Hybrid Spectral CNN	99.75	99.63	99.98	99.97	100	100

The accuracy and loss for 100 epochs of training and validation sets are appeared in Fig. 2 for Hybrid Spectral CNN. It very well may be seen that the convergence is achieved in around 50 epochs which shows that the strategy is quick. The effect of spatial dimension over the performance of the model is appeared in Table 3.

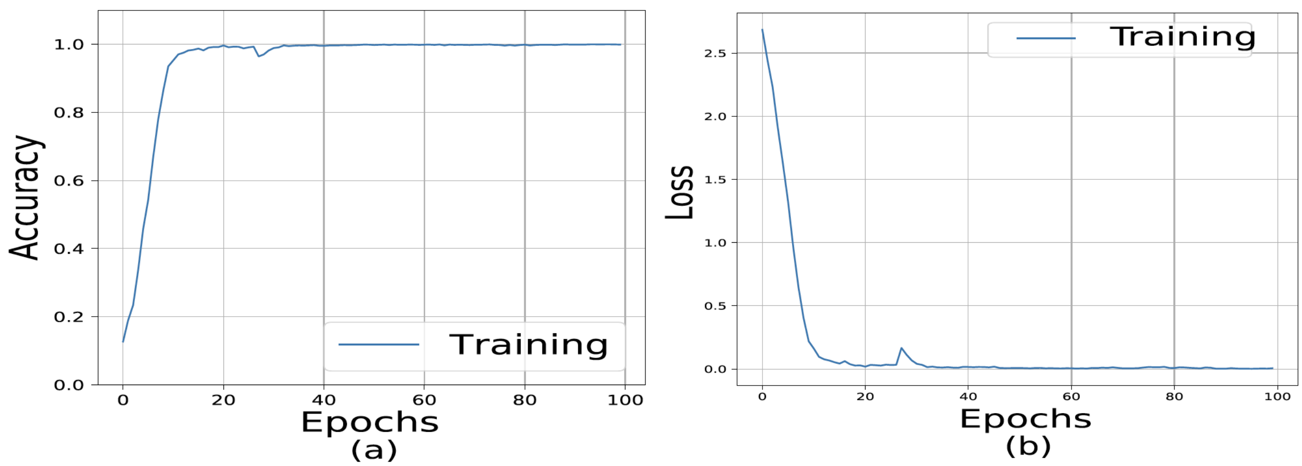

Fig. 2. Accuracy and loss versus epochs over the Indian Pines dataset.

Table 3			
IMPACT OF THE SPATIAL WINDOW SIZE ON THE PERFORMANCE OF HYBRID Spectral CNN			
Window	Indian Pines(%)	University of Pavia(%)	Salinas Scene(%)
19X19	99.74	99.98	99.99
21X21	99.73	99.90	99.69
23X23	99.31	99.96	99.71
25X25	99.75	99.98	100

Discussion

The Hyperspectral Images are very huge in number, however, it has very limited labeled data sets, also it includes varying bands of images. Thus, it makes it difficult for a 2D-CNN to give a good performance because the performance will depend on both spatial and spectral information. This problem can be tacked by a 3D-CNN, however, because of the high computational complexity, it makes it non-efficient. In both the 2D and 3D models there are some shortcomings like missing channel relationship information or very complex model.

My approach to handling this problem is to combine both 3D and 2D CNN to make it efficient and perform better taking both spatial and spectral information into consideration. Here, the 3-D-CNN and 2-D-CNN layers are assembled so that they use both the spectral as well as spatial feature maps to their full extent to accomplish maximum possible accuracy.

It was interesting to use hybrid model which does not get affected with smaller data-set. The Hybrid Spectral CNN outperformed all the traditional models as shown in Table 2. Spatial information when using patches of size 19X19, it seems to be sufficient to reach a remarkably good accuracy as shown in Table 3. The use of spatial information is highly effective, the conjunction of spatial and spectral information achieves the best classification results. The combination of 2D and 3D CNN are able to achieve an Overall Accuracy near 100% with small amount of data. It is observed that the spectral information is able to reduce the uncertainty of the classifier when few training data is available. The 2D kernels allows to reduce the over-fitting in comparison to only 3D model.

Also, the efficiency of the model is way better than the 3D-CNN as shown in Table 4 because the complexity has been removed by using 2D-CNN. The model is more complex than the 2D-CNN and taking more time to train as well, however, we are taking the spectral features as well into consideration which is also increasing the size of the data and it is increasing the performance of the model.

Table 4

TRAINING TIME IN MINUTES (m) AND TEST TIME IN SECONDS (s)

——	2D-CNN		3D-CNN		Hybrid Spectral CNN	
Data	Train(m)	Test(s)	Train(m)	Test(s)	Train(m)	Test(s)
Indian Pines	1.9	1.1	15.2	4.3	14.1	4.8
University of Pavia	1.8	1.3	58.0	10.6	20.3	6.6
Salinas Scene	2.2	2.0	74	15.2	25.5	9

The performance has increased with the use of Hybrid Spectral CNN, however, it can reach 100% with this method for India Pines and University of Pavia dataset as well, if more of the labeled data are available. Or the model can perform better for the 2 data set if K-fold or data augmentation is applied to them.

Conclusion

This paper has presented a mixture of 3-D CNN and 2-D CNN models for HSI classification. The proposed Hybrid Spectral CNN model fundamentally consolidates the integral data of spatial-spectral and spectral as 3-D and 2-D convolutions, individually. The model created on three data sets is compared with the latest methods and the outcome of the Hybrid Spectral CNN is way better than the others. It is also showing better performance on smaller training data sets.

References

1. C.-I. Chang, Hyperspectral Imaging: Techniques for Spectral Detection and Classification, vol. 1. Springer Science & Business Media, 2003.

2.G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, “Advances in hyperspectral image classification: Earth monitoring with statistical learning methods,” IEEE Signal Process. Mag., vol. 31, no. 1, pp. 45–54, Jan. 2014.

3. S. P. Sabale and C. R. Jadhav, “Hyperspectral Image Classification Methods in Remote Sensing - A Review,” in 2015 International Conference on Computing Communication Control and Automation, 2015, pp. 679–683.

4. Y. Luo et al., “HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image,” in 2018 International Conference on Audio, Language and Image Processing (ICALIP), 2018, pp. 464–469.

5. A. Santara et al., “BASS Net: Band-Adaptive Spectral-Spatial Feature Learning Neural Network for Hyperspectral Image Classification,” IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 5293–5301, Sep. 2017.

6. Y. Chen et al., “Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks,” IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232–6251, Oct 2016.

7. Z. Tian et al., “Hyperspectral Classification Via Spatial Context Exploration with Multi-Scale CNN,” in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2018, pp. 2563–2566.

8. J. M. Haut et al., “Cloud Implementation of Logistic Regression for Hyperspectral Image Classification,” in 17th International Conference on Computational and Mathematical Methods in Science and Engineering (CMMSE), 2017, pp. 1030–1041.

9. J. Leng et al., “Cube-CNN-SVM: A Novel Hyperspectral Image Classification Method,” 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1027–1034, 2016.

10. A. Santara, K. Mani, P. Hatwar et al., “BASS net: band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification,” IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 5293–5301, 2017.

11. S. Ji, W. Xu, M. Yang, and K. Yu, “3D convolutional neural networks for human action recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 221–231, Jan. 2013.

12. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc.

Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.
13. K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, “Deep supervised learning for hyperspectral data classification through convolutional neural networks,” in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Jul. 2015, pp. 4959–4962.
14. A. B. Hamida, A. Benoit, P. Lambert, and C. B. Amar, “3-D deep learning approach for remote sensing image classification,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4420–4434, Aug. 2018.
15. M. He, B. Li, and H. Chen, “Multi-scale 3D deep convolutional neural network for hyperspectral image classification,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 3904–3908.