Performance Evaluation of Convolutional Neural Network at Hyperspectral and Multispectral Resolution for Classification

Subir Paul *[a], Poliyapram Vinayaraj[b], D. Nagesh Kumar[a,c], Ryosuke Nakamura[d]
[a]Department of Civil Engineering, Indian Institute of Science, Bengaluru-560012, India; [b]National Institute of Advanced Industrial Science and Technology (AIST)-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory (RWBC-OIL), Tokyo 152-8550, Japan; [c]Centre for Earth Sciences, Indian Institute of Science, Bengaluru-560012, India; [d]National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, Japan

ABSTRACT

Convolutional Neural Network (CNN) has established as an effective deep learning model for hyperspectral image classification by considering both spectral and spatial information. In this study, the performance of two-dimensional (2D) CNN architecture is evaluated at hyperspectral and multispectral resolution. Two types of multispectral data are analyzed viz., original and transformed multispectral data. Hyperspectral bands are transformed to spectral resolution of multispectral bands by averaging the reflectances of specific hyperspectral narrow bands which are falling within the spectral ranges of multispectral bands. The well-known Pavia University dataset and a new dataset of Pear orchard are investigated in this study. In case of Pear orchard dataset, classification is performed with both types of multispectral data. All the experiments are carried out with the same 2D CNN architecture. In case of Pavia University dataset, hyperspectral and transformed multispectral data achieve OA(%) of 94.29±1.28 and 94.27±2.01 respectively considering 20% samples as training. In case of Pear orchard dataset, hyperspectral, multispectral and transformed multispectral data achieve OA(%) of 91.59±0.89, 88.65±1.35, and 93.24±0.16 respectively considering 20% samples as training. It is evident that transformed multispectral data, which comprises of inherent hyperspectral information, provides similar or better performance compared to hyperspectral data. Further, with the use of 3D CNN architecture, classification performance improves in case of Pavia University dataset, whereas it remains statistically similar in case of Pear orchard dataset. The present promising results illustrates the performance of CNN even in small dataset which is comparable to several published state-of-the art results on the same dataset.

Keywords: Convolutional neural network, deep learning, hyperspectral data, multispectral data, spectral-spatial classification, 2D CNN, 3D CNN

1. INTRODUCTION

Remote sensing image acquisition and processing of those images for different Earth observation applications (e.g. agriculture or environmental monitoring, military surveillance etc.) have become very important in recent times [1]. The main differences between multispectral (MS) and hyperspectral (HS) data are the number of bands utilized for data acquisition and spectral resolution of each band. MS data generally consists of 3-10 bands with spectral resolution of 30-200 nm (approx.), whereas HS data have hundreds of narrower bands with spectral resolution of 1-10 nm (approx.). HS data provide abundant spectral information about any observed objects (e.g. agricultural fields, crop or soil types, land use land cover (LULC), minerals etc.), which are advantageous for detection or identification and classification of these objects [2, 3]. Regardless of the advantages, large number of contiguous narrow spectral bands are also basis of the issues like high dimensionality, Hughes phenomenon and data redundancy.

In recent years, deep learning techniques have been widely used for HS image classification [2-10] compared to support vector machine (SVM), random forest (RF) and other traditional (e.g. maximum likelihood, k-nearest neighbor, naïve Bayes etc.) classifiers. SVM is not efficient resourcefully in case of big data (e.g. HS data or spatial data over a large area)

---

* subir.paul.iisc@gmail.com

applications, whereas RF requires multiple features for efficient performance [11]. Different deep learning models are available, viz. Deep Belief Network (DBN), Recurrent Neural Network (RNN), Stacked Auto-Encoder (SAE), Convolutional Neural Network (CNN); among which CNN is originated to be most promising for diverse remote sensing applications (e.g. classification, segmentation, and object detection etc.) [12]. Deep learning technique, such as CNN has several applications in the spectral-spatial classification of HS images but very limited studies are available with MS images [12, 13]. Spectral-spatial classification techniques have advantages of considering additional spatial information in the algorithm which improves classification performances [2, 7, 12]. In case of analysis of multi-temporal MS images dimensionality of the dataset increases, which introduces the possibilities of using deep learning techniques for feature extraction or classification. Ji, et al. [14] introduced three-dimensional (3D) CNN for crop classification using multi-temporal MS remote sensing images and showed better performance compared to two-dimensional (2D) CNN and other conventional methods.

HS and MS data were compared for land cover classification [15], seagrass species mapping [16], urban land cover mapping [17], vegetation mapping [18] and better performances were reported with HS data. These comparison studies were investigated mostly using traditional classifiers. To the best of our knowledge, HS and MS data are not compared utilizing the deep learning technique, CNN. Hence, the primary objective of this study to compare the classification performances of HS and MS resolution datasets on employing 2D CNN. In order to accomplish the primary goal, HS bands are transformed to the spectral resolution of MS bands, since MS image was not available for multiclass classification problem. This step will address another research question that whether transforming of HS bands to MS resolution will affect the classification performance! In order to achieve these goals, two datasets are utilized. The first dataset is publicly available Pavia University dataset, which is used to compare the performances of HS bands and transformed MS bands. The second dataset is a new dataset covering a Pear orchard, where both HS and MS data are available. In case of Pear orchard dataset, classification performances are evaluated with HS bands, original MS bands and transformed MS bands.

2D CNN is mainly designed to perform spatial convolution operation, whereas 3D CNN performs convolution operation over an additional dimension (i.e. spectral or temporal) and extracts the features from the data [9, 14]. HS data comprises of contiguous narrow spectral bands, among which adjacent spectral bands are highly correlated and provide redundant information. In case of HS data, 2D CNN can only address the issue in spatial correlation, whereas 3D CNN deals with both spatial and spectral correlation for creation of feature maps [9]. Hence, 3D CNN should provide a better spectral-spatial representation of the HS datasets. Therefore, HS datasets are also classified using 3D CNN architecture to evaluate the improvements over the performances of 2D CNN.

The remainder of this paper is organized as follows. In Section 2, we have discussed the CNN architecture for classification in detail. Section 3 presents the details about the datasets. The results of all the experiments are reported and discussed in Section 4. To end, Section 5 summarizes the overall conclusions of this work.

2. METHODOLOGY

2.1 Convolutional Neural Network (CNN)

CNN is a deep learning algorithm, where deep (multi-layer) neural networks are used to learn the deep features. In case of CNN, neural networks utilize mathematical operation ‘convolution’ in one or more layers instead of simple matrix multiplication [19]. Convolution operation assists machine learning system to reduce the memory requirements and improve statistical efficiency with three important ideas viz., sparse interactions, parameter sharing, and equivariant representations [19]. CNN models comprise of convolution layer and pooling layer in the hidden layers and followed by a fully connected layer. Nonlinear activation functions are used for the creation of feature maps in the convolution layers. Pooling layers generally use max-pooling or average-pooling operation to provide a summary statistic of a certain location with reference to its surroundings. CNN models can be designed using one-dimensional (1D), 2D and 3D CNN. In case of HS images, 1D and 2D CNN perform convolution in the spectral and spatial domain respectively, whereas 3D CNN performs convolution in both spectral and spatial dimension.

2.2 CNN architecture

We have used both 2D CNN and 3D CNN architecture in this study. The graphical representation of 2D CNN architecture is shown in Figure 1. In the first step, small patches of (7×7) size are extracted corresponding to each pixel of the image by centering the pixel in the patch. This process of patches extraction is adopted from [9]. This patches extraction process provides a representative (7×7) patch with reference to each pixel, centered in that patch. Both CNN architectures are investigated with three convolutional layers, where 64, 32 and 16 filters are used respectively and rectified linear unit (ReLU) is considered as an activation function to create the feature maps. ReLU is the most popular and computationally efficient compared to the sigmoid function for the training of deep learning models, like CNN [11]. To avoid the reduction of spatial dimension in the feature maps, use of pooling layers is not considered in this work. Use of pooling layers on all the features can cause underfitting where precise spatial information preservation is necessary [19]. In case of 2D CNN, kernel size of (3×3) and strides of (1×1) have been used. To retain the spatial dimension in the feature maps, 'same' padding has been used in the convolution layers. Finally, the feature maps are flattened and connected to the next layer with 'Softmax' activation function to classify the Pavia University dataset into 9 LULC classes. In case of Pear orchard dataset, the last layer will have 2 classes (tree and non-tree). Training of the dataset is performed for 50 epochs with a batch size of 128. RMSprop has been used as optimizer function with a learning rate of 0.001.

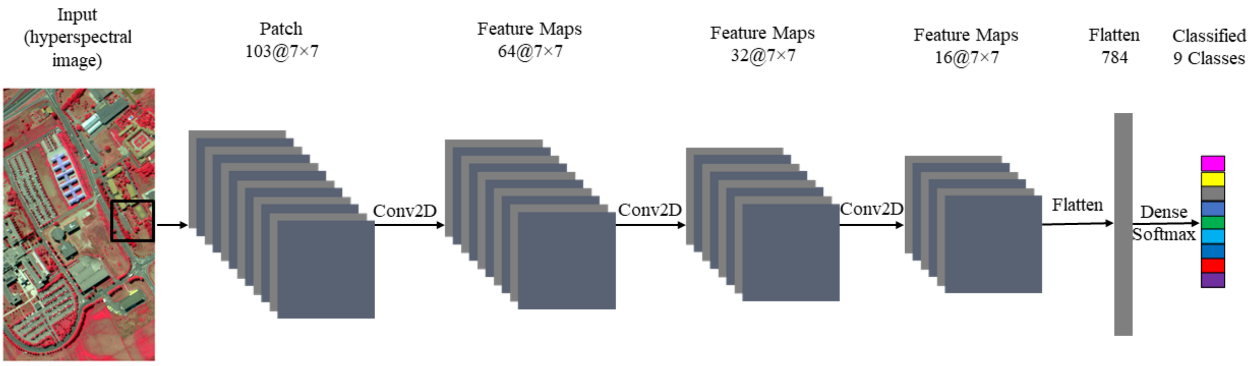

Figure 1. 2D CNN architecture used in this study.

In case of 3D CNN architecture 'Conv2D' (in Figure 1) is replaced with 'Conv3D'. The 3D CNN architecture is experimented with kernel size of (3×3×3), strides of (1×1×3), and rest of the configurations are same as 2D CNN.

2.3 HS to MS transformation

The HS narrow bands of Pavia University dataset are transformed to MS broad-bands for the comparative study. The wavelength ranges of Landsat-8 bands are used as a baseline for transformation to MS resolution. Five MS bands are created by averaging the reflectances of HS narrow bands which are falling under the wavelength ranges of specific Landsat-8 bands. Table 1 shows the wavelength ranges of MS bands and the corresponding HS bands which are averaged for the transformation.

Though, Pear orchard data comprises of both HS and MS data, HS bands are transformed to spectral resolution of MS bands to compare the performances of original MS data and transformed MS data. In this case, wavelength ranges of 4 MS bands are used as a reference for HS to MS transformation and the details of the bands are reported in Table 1.

Table 1. Details of HS bands and wavelengths utilized for MS bands creation.

HS band numbers	**Wavelength ranges (µm)**	**MS band numbers**
Pavia University dataset		
1 – 5	0.430 – 0.447	1
6 – 20	0.451 – 0.510	2
25 – 39	0.530 – 0.590	3
51 – 58	0.640 – 0.670	4

101 – 103	0.850 – 0.860	5
Pear orchard dataset		
72 – 93	0.530 – 0.570	1
132 – 153	0.640 – 0.680	2
180 – 185	0.730 – 0.740	3
202 – 223	0.770 – 0.810	4

2.4 Experiment setup

The datasets are divided into three parts (i.e. training, validation and testing) for all the experiments. Initially, all the experiments are carried out considering 20, 20, and 60% of the labelled samples as training, validation and testing dataset respectively. Later, performance of 2D CNN is evaluated with Paiva University dataset by increasing the training sample size to 60% of the labelled samples, where rest of the labelled samples are divided equally into two parts to be used as validation and testing data. All the data partitioning is done through random sampling and such random samplings are performed 5 times for each experiment. The mean and standard deviation (SD) of overall accuracy (OA) and kappa statistic (k) are reported from the testing results of 5 random trials, performed for each experiment. Precision and recall of each class are evaluated for all the experiments. Precision represents a fraction of predicted pixels identified correctly, whereas recall is the fraction of ground-truth pixels which are identified correctly. Precision and recall are equivalent to user and producer accuracy respectively. CNN architectures are modelled and run using deep learning library 'Keras' [20] in Python.

3. DATASETS

Experiments are conducted using two datasets viz., Pavia University dataset[†] (publicly available HS data), and Pear orchard data from Belair Hesbania 2017 dataset [21]. In this study, Pavia University dataset is preferred among all the publicly available HS datasets since this dataset has large numbers of samples corresponding to each class and CNN requires massive training datasets [22]. These two datasets have their own significance for the comparison study. Pavia University dataset is a multiclass classification problem, whereas Pear orchard data is a binary classification problem and also comprises of original MS data.

Pavia University dataset was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor, having 103 spectral bands in the wavelength range of 0.43–0.86 µm, during a flight campaign over Pavia, northern Italy. This dataset covers an urban area of 610×340 pixels with 1.3 m spatial resolution, out of which 42776 pixels have class label information (9 types of land-cover classes). The details of land-cover classes are shown in Table 2; a color composite image and a map of the land-cover classes are presented in Figure 2(a) and 2(b) respectively.

Belair Hesbania 2017 dataset comprises of pear and apple orchard data acquired in Belgium. We have used the Pear orchard dataset (Figure 3) for tree and non-tree classification. This dataset includes both HS and MS airborne sensor acquisitions and hence apt for our comparative study with CNN. The HS and MS images, which are used in this work, were acquired using the Headwall Micro-Hyperspec and Parrot Sequoia sensor respectively on 14th June, 2017. The HS image contains spectral information of 326 bands (with ~2 nm bandwidth and covering the spectral range of 400 to 1000 nm) at 5 cm spatial resolution. The spectral bands around 400 nm and 900-1000 nm were found to be noisy [21], hence 264 bands (405-895 nm) are utilized in this study. The MS image contains spectral information of 4 bands viz., green (530-570 nm), red (640-680 nm), red-edge (730-740 nm) and near-infrared (770 to 810 nm) region at 8 cm spatial resolution. The HS image is co-registered with the MS image and upscaled to 8 cm spatial resolution for the comparative study. The tree canopy areas of 178 Pear trees are provided in the Belair Hesbania 2017 dataset, which are used as a reference for tree and non-tree classification.

[†] http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University

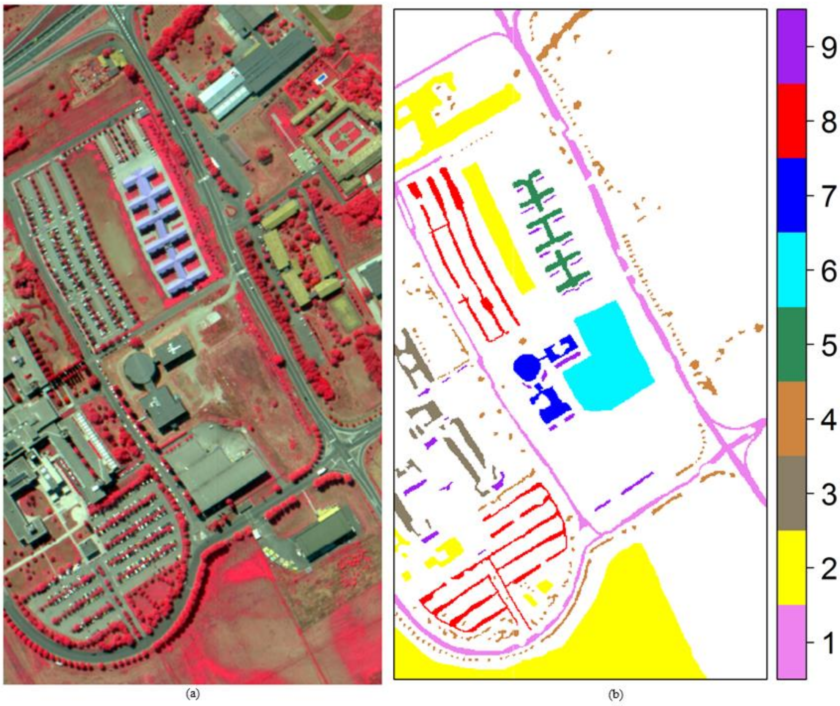

Figure 2. (a) Color composite image [R: 804, G: 657, B: 552 nm]; and (b) Ground-truth or class label maps of the Pavia University dataset.

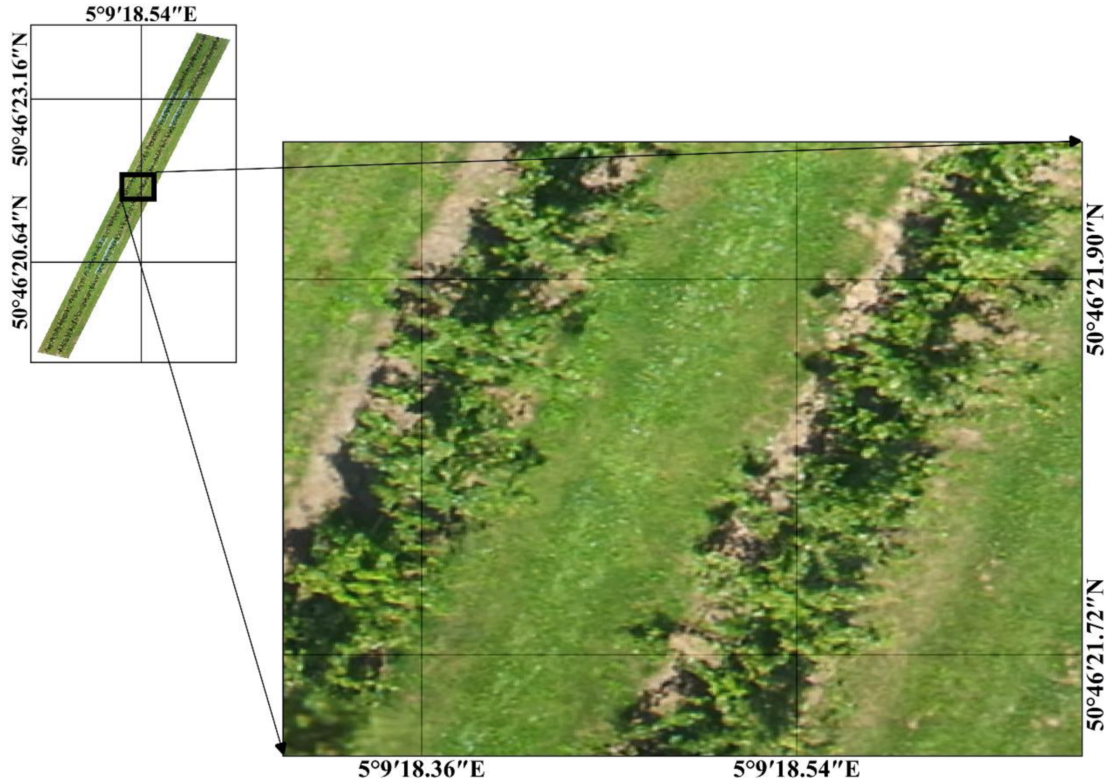

Figure 3. True color composite (R: Red, G: Green, B: Blue band) image of the location of Pear orchard (Belair Hesbania 2017 dataset).

Table 2. Details of the labelled samples of the Pavia University dataset.

Class Sl. No.	Class Name	Total Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947
Total Samples		**42776**

4. RESULTS AND DISCUSSION

4.1 Performances of Pavia University dataset with 2D CNN

The Pavia University data has experimented with 2D CNN architecture using the original HS and transformed MS bands. The average performances of classification metrics are presented in Table 3. The transformed MS bands require less computational time to train the 2D CNN network when compared with the HS bands. It has been observed that both HS and transformed MS data are providing almost similar average OA and k. The SD of OA and k, estimated from five random trials, are higher in case of transformed MS dataset, which implies that the performance of transformed MS data is not as consistent as HS data. A confusion matrix (Table 4) is also created from the results of testing samples of a single trial, where 2D CNN is trained with HS bands. It has been observed from the confusion matrix that the misclassification of Class 2 (i.e. Meadows) and 6 (i.e. Bare soil) are mostly occurred to each other. Analyzing the precision and recall of each class, it has been observed that all the classes are providing similar performances. Class 3 (i.e. Gravel) and Class 6 (i.e. Bare soil) are providing significantly lower recall values compared the precision and recall of other classes. These comparable performances of HS and transformed MS datasets can be caused for two reasons. Firstly, the MS data is transformed from the HS data, hence MS data also comprises of some inherent HS information. Second possible reason can be the advantage of using 2D CNN network, which learns the salient deep features even from the MS data to classify the land cover classes. Therefore, in order to make more conclusive remarks, we have investigated this comparison with another dataset (Pear orchard), where original MS data is available.

Table 3. Classification performances of testing dataset (60% of labelled samples) of Pavia University dataset with 2D CNN.

Class	**HS**		**transformed MS**	
#	**Precision**	**Recall**	**Precision**	**Recall**
1	0.9620±0.0164	0.9580±0.0110	0.9800±0.0100	0.9600±0.0265
2	0.9500±0.0346	0.9800±0.0122	0.9467±0.0351	0.9733±0.0208
3	0.9120±0.0661	0.7780±0.1453	0.9133±0.0416	0.7867±0.2223
4	0.9960±0.0055	0.9900±0	0.9967±0.0058	0.9867±0.0058
5	1±0	1±0	1±0	1±0
6	0.9200±0.0354	0.8080±0.1303	0.9033±0.0643	0.8100±0.1253
7	0.9360±0.0472	0.9200±0.0534	0.9667±0.0404	0.9133±0.0058
8	0.8800±0.0447	0.9540±0.0230	0.8800±0.1217	0.9667±0.0306
9	0.9980±0.0045	0.9940±0.0055	1±0	0.9967±0.0058
OA (%)	94.29±1.28		94.27±2.01	
k	0.9247±0.0178		0.9244±0.0274	

Table 4. Confusion matrix created from the testing performance of a single trial using HS bands of Pavia University dataset in the 2D CNN.

		Predicted (output) class								
	Class #	1	2	3	4	5	6	7	8	9
True (target) class	1	3741	0	40	0	0	0	52	91	0
	2	1	10159	0	5	0	257	1	5	0
	3	20	0	1016	0	0	0	0	203	0
	4	6	6	0	1753	0	2	0	0	0
	5	0	0	0	0	802	0	0	0	0
	6	2	436	33	2	0	2545	0	21	0
	7	37	0	0	0	0	1	767	10	0
	8	52	0	40	1	0	0	0	2157	0
	9	1	0	0	0	0	0	0	0	559

4.2 Performances of Pear orchard dataset with 2D CNN

Pear orchard dataset comprises of both HS and MS datasets. Hence, this dataset is more preferable for the comparison study. But this data is a binary classification problem, having tree and non-tree classes. The same 2D CNN network, used with Pavia University dataset, is trained with HS and MS data of Pear orchard to classify the tree and non-tree classes and performance metrics are reported in Table 5. Further, the HS bands are transformed into MS bands using the same process as explained in Section 2.3 to evaluate the performance of transformed MS data. The transformed MS data are used to train the 2D CNN architecture for tree and non-tree classification and the results are presented in Table 5. It has been observed that non-tree class is consistently achieving better precision and recall compared to tree class with all the three experiments. The non-tree class has almost three times samples than that of the tree class, hence the non-tree class samples are trained more precisely and providing better testing accuracy. It is evident from the results that transformed MS data is providing the best performance, where the average OA and k are improved by 1.8% and 5.96% respectively compared to HS data. Precision of non-tree class and recall of tree class have been improved significantly with the transformed MS data. The computational time required to train the 2D CNN network with transformed MS bands is comparatively reduced than use of HS bands, since very few bands are utilized during training. The better performance with the transformed MS data could be caused because of the inherent HS information and using the spectral bands from some specific wavelength regions instead of using all the bands, which could have some redundant information. However, HS data is achieving better accuracy compared to the original MS data with an increase of 3.32% and 10.85% in average OA and k. A confusion matrix (Table 6) is created from the results of testing samples of a single trial, where 2D CNN is trained with the transformed MS bands. It has been observed that almost similar numbers of samples from both the classes are misclassified, but non-tree class has comparatively large numbers of samples as well as more numbers of correctly classified samples. Hence, the precision and recall of non-tree class are significantly larger than tree class.

Table 5. Classification performances of testing dataset (60% of labelled samples) of Pear orchard dataset with 2D CNN.

Class	HS		MS		transformed MS	
	Precision	Recall	Precision	Recall	Precision	Recall
Tree	0.8540±0.0195	0.8240±0.0594	0.7980±0.0377	0.7780±0.1154	0.8620±0.0084	0.8880±0.0110
Non-tree	0.9380±0.0179	0.9480±0.0110	0.9220±0.0335	0.9260±0.0261	0.9600±0.0000	0.9480±0.0045
OA (%)	91.59±0.89		88.65±1.35		93.24±0.16	
k	0.7815±0.0281		0.7050±0.0531		0.8281±0.0038	

Table 6. Confusion matrix created from the testing performance of a single trial using transformed MS bands of Pear orchard dataset in the 2D CNN.

		Predicted (output class)	
	Class	Non-tree	Tree
True (target) class	Non-tree	85403	4260
	Tree	4002	28540

4.3 Impact of training sample sizes

In order to evaluate the impact of training sample size, experiments are performed with 2D CNN using the Pavia University dataset. The training sample size is increased from 20% to 60% of the labelled samples. The changes in classification performances with the increase in training sample size are reported in Table 7. The average OAs have been improved by 2.61% and 4.8% with the use of increased training samples (60%) from HS and transformed MS data respectively. Similarly, average of k is also improved by 3.24% and 6.48% in case of HS and transformed MS data respectively. It has been observed that transformed MS dataset is providing better classification performances compared to HS data with the increase in training sample size.

Table 7. Classification performances of testing samples of Pavia University dataset for different training sample size with 2D CNN.

Training sample size		HS	transformed MS
20%	**OA (%)**	94.29±1.28	94.27±2.01
	k	0.9247±0.0178	0.9244±0.0274
60%	**OA (%)**	96.75±1.52	98.80±0.11
	k	0.9574±0.0202	0.9843±0.0015

4.4 Performances with 3D CNN

MS data has very few (i.e. ~ 4-6) spectral bands, which are not contiguous and also inter-band correlation is not as significant as HS bands. Hence, 3D CNN is investigated with HS narrow bands only. The Pavia University dataset with 3D CNN provides OA (%) and k of 96.64±0.63 and 0.9559±0.0083 respectively. Use of 3D CNN improve the average OA by 2.5% compared to 2D CNN. The advantage of 3D CNN architecture is it considers the context of spectral correlation in the feature extraction process; hence the improvement is observed compared to 2D CNN. Though, this improvement with 3D CNN is accomplished at the cost of computational time, since it requires more computational time to train the 3D CNN network as compared to 2D CNN.

The HS data of Pear orchard is also trained with 3D CNN network and the performance metrics are evaluated with the testing dataset. The average OA (%) of 92.62±0.21 and k of 0.8038±0.0039 have been estimated. The performance (in terms of OA) of 3D CNN is increased by 1.12% compared to 2D CNN and reduced by 0.66% compared to the performance of transformed MS with 2D CNN. Theoretically, 3D CNN should provide better performance compared to 2D CNN network, which is evident with the original HS data. However, the transformed MS data with 2D CNN is producing statistically equivalent performances as compared to the performance of HS data with 3D CNN. Hence, the advantage of 3D CNN network is not noticeable with the Pear orchard dataset. The possible reason to be highlighted against such results is that the dataset is a binary classification problem which may not have sufficient additional information (about the variability between samples of two classes) to be captured by the 3D CNN network.

5. CONCLUSIONS

The comparison of classification performances of 2D CNN network has been investigated at HS and MS resolution. Two datasets are used for this comparison study, which can be distinguished as binary (i.e. Pear orchard dataset) and multi-class (i.e. Pavia University dataset) classification problem. The HS bands are utilized to create the transformed MS bands by averaging the spectral responses of specific wavelength regions. These transformed MS bands are proven to be providing either equivalent or better performance compared to the HS bands. Hence, HS narrow bands can be transformed to MS bands prior to train the 2D CNN network, which will be advantageous in reducing the computational time without affecting the classification performance. In case of Pear orchard dataset, performances of 2D CNN network with original HS and MS data are compared and better performance is observed with HS data. The excellent performances of transformed MS data are evident because of inherent presence of HS information and use of spectral bands from some specific wavelength regions. The performance of 3D CNN network is experimented with the HS data and better accuracy is observed compared to 2D CNN in the cost of computational time. The impact of training sample size is also investigated with 2D CNN network and it can be concluded that with the increase in training sample sizes testing performance improves. This study lacks training of 2D CNN network with original MS data for multi-class classification problem, which should be investigated for better comparison. The imbalance in the number of samples of each class can affect the classification problem, which requires further investigation.

ACKNOWLEDGMENTS

Authors would like to thank the HYPERTEMP Project group and BELAIR HESBANIA 2017 dataset provider for sharing the data. First author wishes to thank National Institute of Advanced Industrial Science and Technology (AIST), Japan for providing the opportunity of Technical Training and financial support to carry out this research work with the Geoinformation Science Research Team (GSRT), Artificial Intelligence Research Center (AIRC), AIST.

REFERENCES

[1] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "A new deep convolutional neural network for fast hyperspectral image classification," *ISPRS journal of photogrammetry and remote sensing*, vol. 145, pp. 120-147, 2018.
[2] S. Paul and D. N. Kumar, "Spectral-spatial classification of hyperspectral data with mutual information based segmented stacked autoencoder approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 138, pp. 265-280, 2018.
[3] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing*, vol. 219, pp. 88-98, 2017.
[4] N. Audebert, B. Le Saux, and S. Lefèvre, "Deep Learning for Classification of Hyperspectral Data: A Comparative Review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, pp. 159-173, 2019.
[5] 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 Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 6232-6251, 2016.
[6] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE Journal of Selected topics in applied earth observations and remote sensing*, vol. 7, pp. 2094-2107, 2014.
[7] Y. Chen, X. Zhao, and X. Jia, "Spectral–spatial classification of hyperspectral data based on deep belief network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, pp. 2381-2392, 2015.
[8] Y. Li, H. Zhang, and Q. Shen, "Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sensing*, vol. 9, p. 67, 2017.
[9] X. Yang, Y. Ye, X. Li, R. Y. Lau, X. Zhang, and X. Huang, "Hyperspectral image classification with deep learning models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, pp. 5408-5423, 2018.
[10] H. Zhang, Y. Li, Y. Jiang, P. Wang, Q. Shen, and C. Shen, "Hyperspectral Classification Based on Lightweight 3-D-CNN With Transfer Learning," *IEEE Transactions on Geoscience and Remote Sensing*, 2019.
[11] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 778-782, 2017.
[12] M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang, "Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery," *Remote Sensing*, vol. 10, p. 1119, 2018.
[13] B. Huang, B. Zhao, and Y. Song, "Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery," *Remote Sensing of Environment*, vol. 214, pp. 73-86, 2018.
[14] S. Ji, C. Zhang, A. Xu, Y. Shi, and Y. Duan, "3D convolutional neural networks for crop classification with multi-temporal remote sensing images," *Remote Sensing*, vol. 10, p. 75, 2018.
[15] L.-J. Ferrato and K. W. Forsythe, "Comparing hyperspectral and multispectral imagery for land classification of the Lower Don River, Toronto," *Journal of Geography and Geology*, vol. 5, pp. 92-107, 2013.
[16] S. Phinn, C. Roelfsema, A. Dekker, V. Brando, and J. Anstee, "Mapping seagrass species, cover and biomass in shallow waters: An assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in Moreton Bay (Australia)," *Remote sensing of Environment*, vol. 112, pp. 3413-3425, 2008.
[17] F. Dell'Acqua, P. Gamba, and G. Lisini, "Urban land cover mapping using hyperspectral and multispectral VHR sensors: Spatial versus spectral resolution," in *Proc. URBAN, IAPRS*, 2005.
[18] M. Govender, K. Chetty, V. Naiken, and H. Bulcock, "A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation," *Water SA*, vol. 34, pp. 147-154, 2008.

[19]I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning. Book in preparation for MIT Press," *URL¡ http://www. deeplearningbook. org,* 2016.
[20]F. Chollet, "Keras," ed, 2015.
[21]Y. Vanbrabant, L. Tits, S. Delalieux, K. Pauly, W. Verjans, and B. Somers, "Multitemporal Chlorophyll Mapping in Pome Fruit Orchards from Remotely Piloted Aircraft Systems," *Remote Sensing,* vol. 11, p. 1468, 2019.
[22]T. Liu and A. Abd-Elrahman, "Deep convolutional neural network training enrichment using multi-view object-based analysis of Unmanned Aerial systems imagery for wetlands classification," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 139, pp. 154-170, 2018.