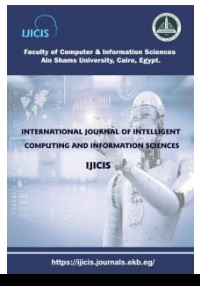
IJICIS, Vol.22, No.4, 146-158
DOI: 10.21608/ijicis.2022.147175.1198

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



HYPERSPECTRAL IMAGE ANALYSIS USING A CUSTOM SPECTRAL CONVOLUTION NEURAL NETWORK

Mayar A. Shafaey*

Faculty of Information Technology, National Egyptian E-Learning University, Cairo, Egypt
mayaraly22@gmail.com

Maryam N. Al-Berry

Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
maryam_nabil@cis.asu.edu.eg

Mohammed A.M. Salem

Faculty of Media Engineering and Technology, German University in Cairo, Egypt
salem@cis.asu.edu.eg

Hala M. Ebied

Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
halam@cis.asu.edu.eg

Elsayed A. El-Dahshan

Faculty of Information Technology, National Egyptian E-Learning University, Cairo, Egypt
seldahshan@eelu.edu.eg

Mohammed F. Tolba

Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
fahmytolba@gmail.com

Received 2022-06-26; Revised 2022-12-15; Accepted 2022-12-25

Abstract: *In recent time, the most applied classification method for hyperspectral images is based on the supervised deep learning approach. The hyperspectral images require special handling while it consists of hundreds of bands / channels. In this article, the experiments are conducted using one of the widespread deep learning models, Convolutional Neural Networks (CNNs), specifically, Csutom Spectral CNN architecture (CSCNN). The introduced network is based on the data reduction and data normalization. It firstly ommits the unnecessary data channels and retains the meaningful ones. Then, it passes the remaining data through the CNN layers (convolutional, rectified linear unit, fully connected, dropout,…etc) until reaches the classification layer. The experiments are applied on four benchmarcks [hyperspectral datasets], namely, Salinas-A, Kenndy Space Center (KSC), Indian Pines (IP), and Pavia University (Pavia-U). The proposed model achieved an overall accuracy more than 99.50 %. In last, a comparison versus the state of the art is also introduced.*

Keywords: *Pixel-based Image Classification, Principal Component Analysis (PCA), Deep Learning.*

1. Introduction

Today, the whole earth can be scanned during the day using the artificial satellites to produce a massive number of Remote Sensing (RS)/ satellite images for widespread applications [1]. Those captured images have different characteristics, i.e., they have high spatial, temporal, and spectral resolution. Firstly, spatial resolution represents how much the human eye can see the photographic image details which are detected

*Corresponding Author: Mayar A. Shafaey

Faculty of Information Technology, National Egyptian E-Learning University, Cairo, Egypt

Email address: mayaraly22@gmail.com

by a satellite sensor. However, temporal resolution refers to the time between images of the same geographical area. Finally, spectral resolution denotes the wavelengths of the electromagnetic spectrum detected by a satellite sensor. These characteristics and information make the process of analyzing and classifying the remote-sensing images the most exciting and fastest-growing research area recently [2-3].

The spectral image is an image that consists of multiple bands over the electromagnetic range. For a specific band or channel, the finer the spectral resolution, the narrower the wavelength ranges. In fact, the traditional cameras capture the images over the three wavelength bands (Red, Green, and Blue) which are known as RGB system. With the progress of time, several remote sensing systems started recording various scenes over multiple wavelength bands [more than three], which has been defined as multispectral sensors. Meanwhile, the innovative multispectral sensors are known as hyperspectral sensors. The acquired hyperspectral images are collected from hundreds of spectral bands for the particular scene.

Today, in the image analysis field, analyzing and classifying the large-scale data, i.e., hyperspectral images are introduced as hot topics. Moreover, to learn the machines for land use scene classification, the large number of remote-sensing datasets are required for training process. The specialists introduced several hyperspectral remote-sensing datasets to allow the researchers to conduct their training experiments [4].

The hyperspectral dataset has an important feature that aids in differentiating materials of interest. It has a detailed spectral information which raises the power of accuracy. Moreover, the recent operated sensors have a fine spatial resolution which allows the image analysis of small spatial structures. All the above-mentioned advances make the hyperspectral data a useful tool for a wide variety of image analysis applications, such as, earth monitoring and object detection, precision cultivation, thematic planning, land use mapping, or change detection [5].

The problem which remains is that the automation degree for remote-sensing images analysis is low despite decades of research, whether in the accuracy percentage value or in the processing time. Indeed, the common paradigm to analyze the hyperspectral data is a pixel-based approach, in which the individual image pixels are classified by the spectral information that they contain. This is the traditional approach of the classification since the pixel is a fundamental unit of the satellite image [6].

2. Related Work

The traditional models for pixel-based classification are based on measuring the distance of the class's mean value. The means and variances are computed based on the statistics of the class. There are several schemes apply this approach, i.e., Minimum Mahalanobis distance, Minimum distance to mean, and Maximum likelihood.

Recently, the new concept of deep Convolutional Neural Networks (CNNs) has an impressive effect on several applications regarding to image analysis and pixels classification, object detection, and image segmentation, regardless the image type. CNNs [7] are very similar to the traditional Neural Networks. It has the same parameters provided for regular Neural Networks i.e. learnable weights and biases. The main difference is that the input of CNN models is in image format, which allows specific features to be encoded into the model. Those features then vastly reduce the number of parameters used in the network. That is bringing the deeper concept with fewer parameters [8].

The deep learning approach showed a remarkable capability of the feature representation. The following table shows some of authors who applied the deep learning approach in their experiments for instance.

Table 1 Review on deep learning methods for remote-sensing images classification

Authors in	Deep learning approach	Datasets	Performance
[9, 10, and 21]	CNNs	UCMerced dataset [11]	93.48%, 98.6%, and 92.4%, respectively
[12]	Deep networks	Malaysian dataset	96.88%
[13]	Region-based CNN	UAV images	97.8%
[14]	CaffeNet, and GoogLeNet [15]	two remote sensing datasets	97 %, 95.48 %
[16]	Deep CNN	Remote sensing images	93.48%
[17]	CNN	Remote sensing scenes	92.43 %

Therefore, the target of this research is to replace the old-fashion multi-layers networks with the deep learning models for pixel-based hyperspectral data analysis.

3. Proposed Network Architecture

The proposed CNN model is consisting of several consecutive layers, the input layer, followed by several layers of convolutional, rectified linear unit, fully connected, dropout, and finally we obtain an output layer, as shown in Figure. 1. The output from the final layer [classification layer] can be a class or a probability of classes that best describes the image.

In this work, the hyperspectral input image is firstly preprocessed by two consecutive steps: data reduction and data normalization. The hyperspectral data include dozens of channels. Hence, there are enormous features in every single input. Thus, some of those data must be omitted to keep the most unique signatures. The principal component analysis (PCA) approach is applied here to reduce the number of spectral bands [18].

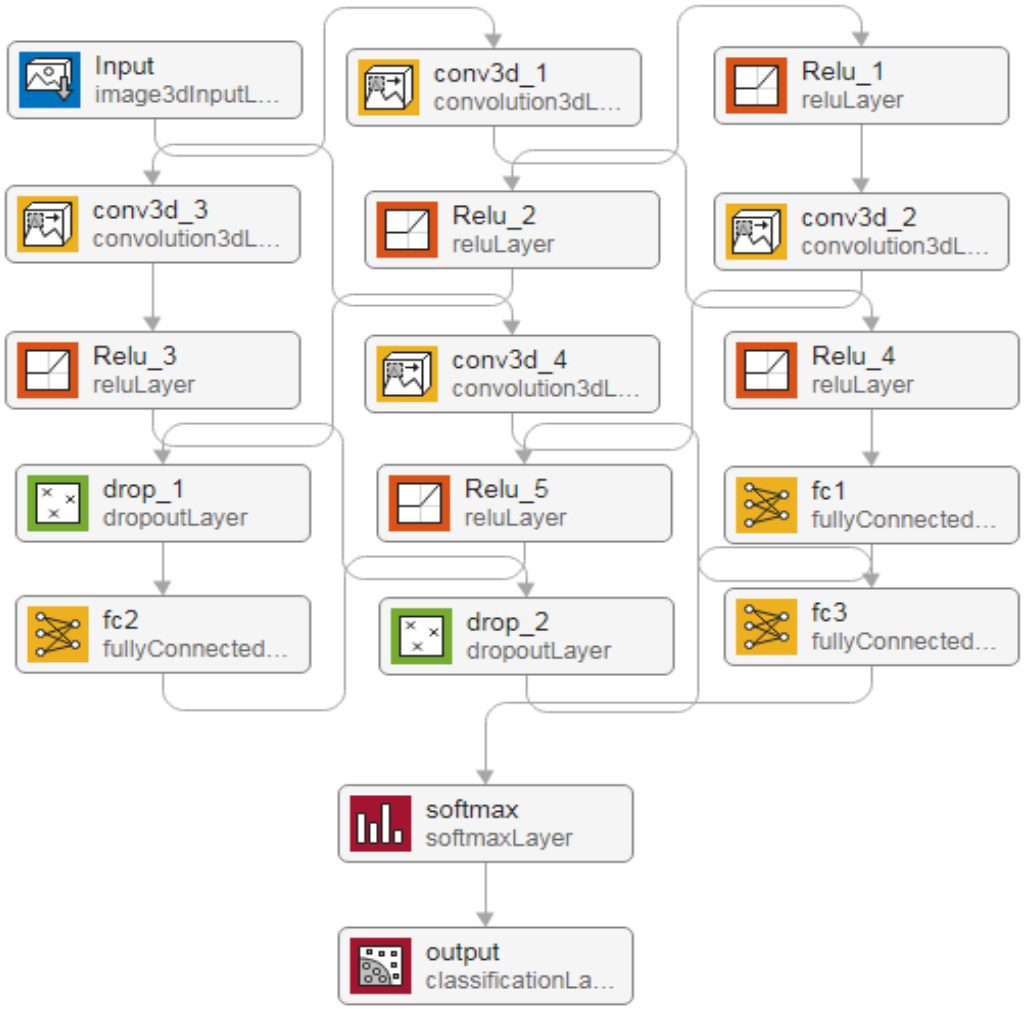


Figure. 1: The CNN layers, starting from the input image, going through several convolutional, dropout, and fully connected layers, then finally ending by the output layer

In principle, PCA maps the image data into a new, uncorrelated vector space for data compression and features extraction. The second preprocessing step is normalizing the input image. The normalization step is carried out by computing the standard deviation of the image [19]. Finally, and for training simplicity, the hyperspectral image is split up into patches. Those patches are then randomly divided into training and testing sets.

Actually, the convolution operation is the most important step in CNNs cycle. It firstly convolves the input by the filter banks to extract the main features then passes it to the subsequent layers. CNNs may include a nonlinear gating function i.e. Rectified Linear Unit (ReLU) to keep only the positive values [20].

It may also include the dropout layers to prevent the training data overfitting. Finally, the fully connected layers are applying its job similarly to the traditional NN. It connects each neuron in certain layer to every neuron in the successive layer.

The convolutional and ReLU layers are responsible for extracting the features of the input and generating the global feature map, while the fully connected layers are responsible for classifying that map to obtain the closer class that matches the input.

4. Experimental Work

4.1. Datasets

In the experiments, we assess the proposed CSCNN network on four common hyperspectral scenes named Salinas-A [22], Kenndy Space Center (KSC) [23], PaviaU [24], and Indian Pines [25]. Firstly, Salinas-A is a sub-scene which acquired from Salinas Valley, USA. It is captured by the AVIRIS sensor [~ 3.7 m/pixel spatial resolution].

It covers the area which includes 86 lines by 83 samples. This scene includes bare soils, vegetables, and vineyard grounds. It includes only six classes and background, as shown in Figure. 2. Table 2 indicates the ground truth classes for the Salinas-A scene and their respective samples number.

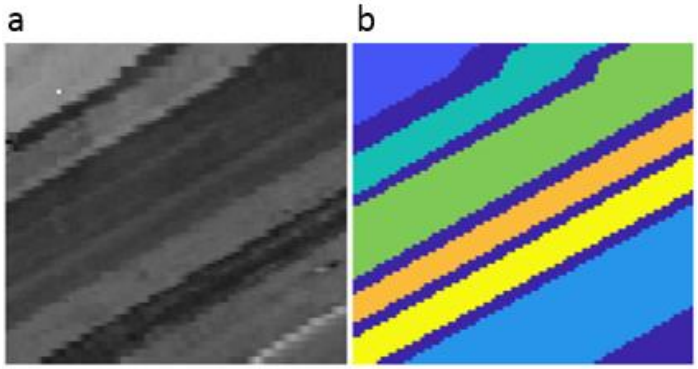


Figure. 2: Salinas-A [original] (a) and Salinas-A [ground-truth] (b)

Table 2 The salinas-a classes and their respective ground truth sample number

No.	Field Name	Pixels Counts
1	Lettuce_romaine_7W	799
2	Lettuce_romaine_6W	674
3	Lettuce_romaine_5W	1525
4	Lettuce_romaine_4W	616
5	Corn_scenced_green_weeds	1343
6	Brocoli_green_weeds	391

Secondly, the KSC data set was obtained by 18-m spectrometer over Florida, USA [23]. That scene has a size of 512 × 614 × 176, and contains 13 classes as ground truth, see Figure. 3 (b) and Table 3.

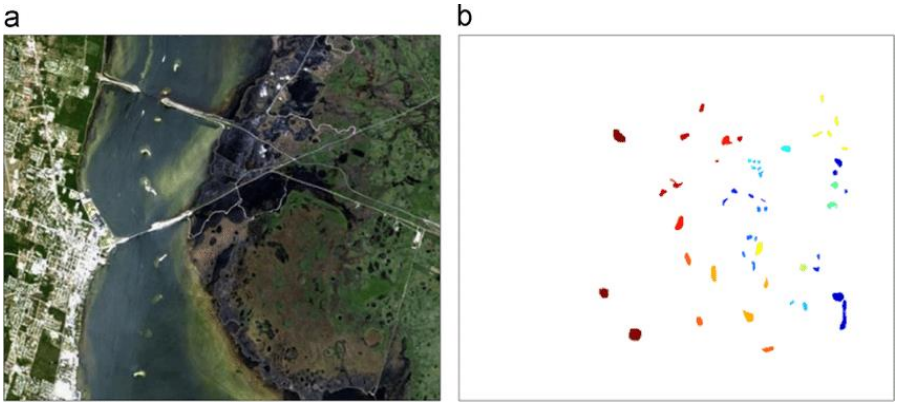


Figure. 3: KSC [original] (a) and KSC [ground-truth] (b)

Table 3 The ksc classes and their respective ground truth sample number

No.	Field Name	Pixels Counts
1	Spartina march	520
2	Cattail march	404
3	CP hammock	256
4	Slash Pine	252
5	Oak	161
6	Hardwood	229
7	Mud flats	503
8	water	927
9	Swamp	105
10	Graminoid march	390
11	Scrub	347
12	Willow swamp	243
13	Salt march	419

Thirdly, The University of Pavia scene [24] was captured in 2003 from a flight over Northern Italy (Pavia) by the ROSIS sensor (~1.3-m/ pixel). Its dimensions are 610 × 340 × 103, and it has 9 ground truth classes which cover the ground as shown in Figure. 4 (b) and Table 4.

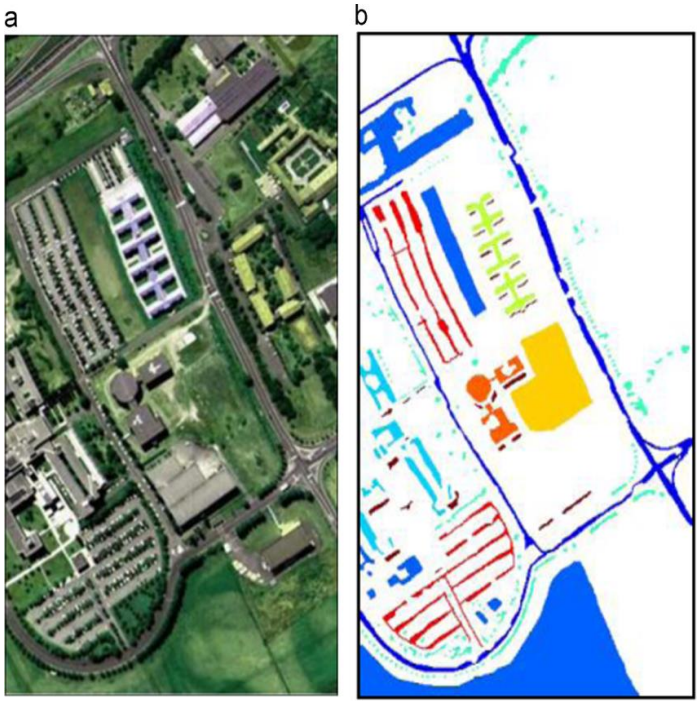


Table 4 The paviaU classes and their respective ground truth sample number

No	Field Name	Pixels Counts
1	Bare soil	5029
2	Bitumen	1330
3	Trees	3064
4	Shadows	947
5	Painted metal sheets	1345
6	Asphalt	6631
7	Meadows	18649
8	Self-blocking bricks	3682
9	Gravel	2099

Figure. 4: Pavia-U [original] (a) and Pavia-U [ground-truth] (b)

Finally, the scene in the test site 'Indian pines' was collected over North-western Indiana. It is captured by the 224-band AVIRIS sensor, by the 0.4–2.5 10^{-6} m wavelength range. It composes of 145 x 145 pixels. It covers 16 classes of agricultural, forest, and road areas, as shown in Figure. 5. Table 5 indicates the classes of the ground truth for the Indian pines scene and its corresponding samples number.

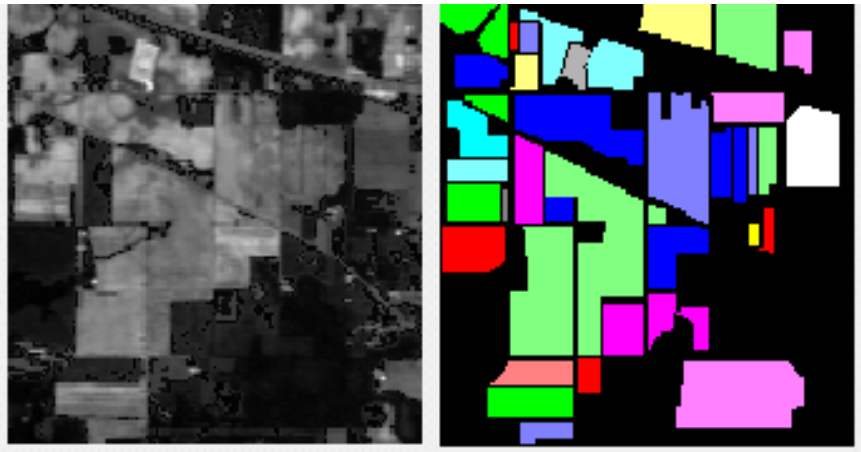


Figure. 5: Indian Pines [original] (a) and Indian Pines [ground-truth] (b)

Table 5 The indian pines classes and their respective ground truth sample number

No.	Field Name	Pixels Counts
1	Oats	20
2	Stone-Steel-Towers	93
3	Wheat	205
4	Soybean-clean	593
5	Buildings-Grass-Trees-Drives	386
6	Soybean-mintill	2455
7	Hay-windrowed	478
8	Woods	1265
9	Alfalfa	46
10	Soybean-notill	972
11	Grass-trees	730
12	Corn-notill	1428
13	Grass-pasture	483
14	Grass-pasture-mowed	28
15	Corn	237
16	Corn-mintill	830

4.2 Experimental procedure

The practical implementation is done using Matlab® software on a hardware machine supplied with NVIDIA GTX 1050 4G Processor by Computing Capability: 6.1, Intel® Core™ i7-7700HQ @ 2.20GHz GPU, and 16 GB RAM. In this article, the results of our experiments are represented using 17 layers custom spectral convolution neural network.

Table 6 shows the architecture of this network including the parameters of each convolutional layer in addition to the dropouts and classification layers. The proposed network is trained by an initial learning rate of 0.001 for 100 epochs, a 256 batch size, momentum 0.9, learning rate factor 0.01, and *Adam* optimization.

Table 6 The parameters of each layer in the CSCNN

Layer #	Layer Name	Layer Parameters
1	Input	The input data are divided randomly into 30: 70 training: testing ratio
2	3D Convolutional	# filter banks: 8 Filter size: [3 × 3 × 7]
3	ReLU	
4	3D Convolutional	# filter banks: 16 Filter size: [3 × 3 × 5]
5	ReLU	
6	3D Convolutional	# filter banks: 32 Filter size: [3 × 3 × 3]
7	ReLU	
8	3D Convolutional	# filter banks: 8 Filter size: [3 × 3 × 1]
9	ReLU	
10	Fully Connected	Number of neurons: 256
11	ReLU	
12	Dropout	Retention Probability: 0.4
13	Fully Connected	Number of neurons: 128
14	Dropout	Retention Probability: 0.4
15	Fully Connected	
16	Softmax	
17	Output	Classification layer

By using the afore-mentioned datasets, there are numerous experiments were carried out to assess the CSCNN performance. In our experiments, we started by splitting the images into 10% training set and repeated the experiment eight times until splitting the training set into 90% of the images.

The best partitioning credits to 30:70 training: testing ratio. The large training set has enhanced the classification accuracy by a small fractional number, but, on the other hand, it consumed a lot of the execution time. Hence, the authors divide the input data randomly into 30:70 training: testing ratio. After that, the training data are sent to the convolutional, ReLU, and dropout layers reaching to the output layer.

5. Results and Discussion

The results are presented as an Overall Accuracy (OA). The OA is computed as the average of Producer's Accuracy and User's Accuracy, see the following equations [26]:

$$\text{Producer's Accuracy} = \frac{\text{correctly identified pixels}}{\text{total number of pixels/class}} \quad (1),$$

$$\text{User's Accuracy} = \frac{\text{correctly identified pixels}}{\text{correctly identified pixels+incorrectly identified pixels}} \quad (2)$$

The following figures 6, 7, 8, and 9 display the original ground-truth image and the predicted classification scenes. The classification accuracy for the test data set reached up 100%, 99.53%, 99.35%, and 99.56% for Salinas-A, KSC, Pavia-U, and Indian Pines respectively. Some of simple images has achieved 100% classification accuracy due to its small dimensions and small number of ground-truth classes.

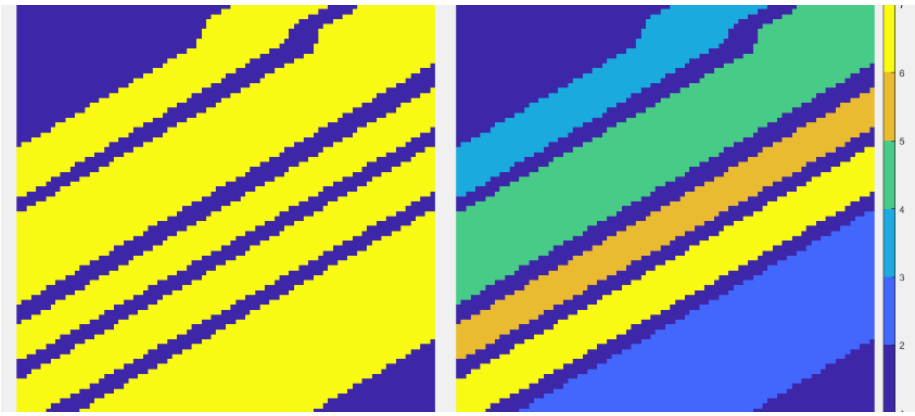


Figure. 6: ground-truth of Salinas-A (left) and the classification result (right)

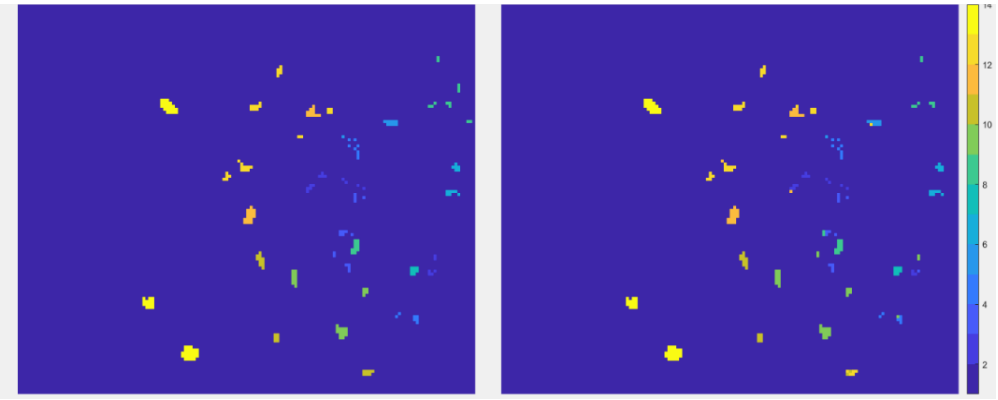


Figure. 7: ground-truth of KSC (left) and the classification result (right)

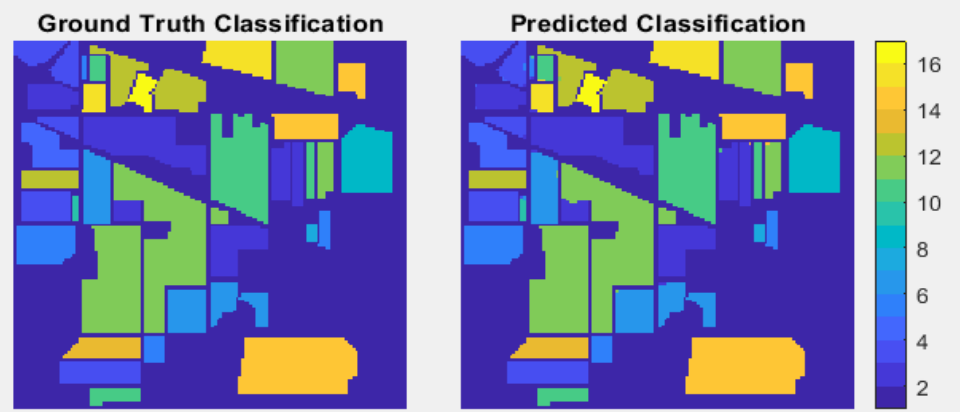


Figure. 8: ground-truth of Indian Pines (left) and the classification result (right)

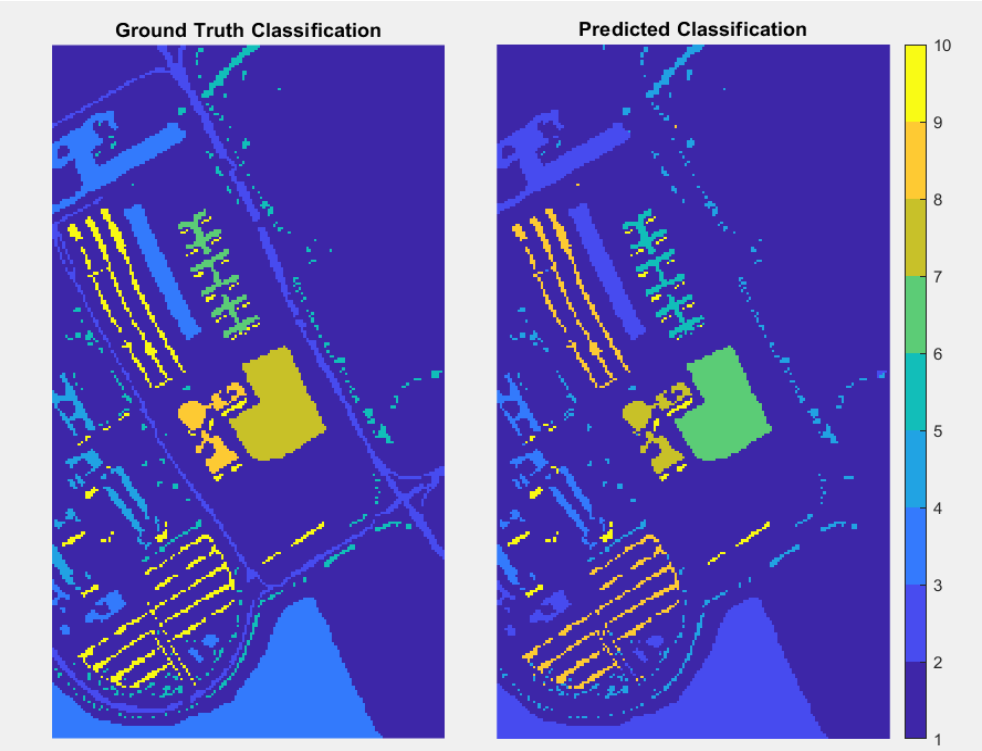


Figure. 9: ground-truth of Pavia-U (left) and the classification result (right)

The following table, Table 7, compares between the proposed CSCNN and the state-of-the-art for Salinas-A, KSC, Pavia-U, and Indian Pines datasets. Most of the authors divided randomly the labeled data into 70:30 training: testing ratio.

By a quick look and comparing to the published results of hyperspectral image classification, the new concept of custom spectral through CNN model gave an impressive result for remote-sensing images classification. The CSCNN filters the image from the unwanted channels. It preserves the significant features which improves the classification accuracy and accelerate the classification process.

Table 7 The CSCNN vs. the state of-the-art

Deep Network	Accuracy of			
	Salinas-A	KSC	Pavia-U	IP
Res-Net [27, 28]	98.09%	93.23%	98.40%	99.03%
Alex-Net [29, 30]	97.22%	96.30%	92.62%	88.26%
RBF-SVM [30]	91.66%	97.25%	90.52%	87.60%
1D- CNN [31]	95.47%	91.80%	93.50%	83.40%
2D- CNN [30]	98.90%	97.81%	98.78%	96.37%
3D- CNN [28]	-	97.87%	99.68%	98.87%
CSCNN	**100%**	**99.53%**	**99.35%**	**99.56%**

6. Conclusions

In this article, the deep learning behavior through CSCNN model was applied for hyperspectral scenes classification. Firstly, the authors explained the description of CNN model and reviewed the deep learning methods of remote-sensing images classification. Then, the experiments were carried out using the pretrained CSCNN on four well-known hyperspectral images, namely, *Salinas-A*, *KSC*, *Pavia-U*, and *Indian Pines*. The proposed network provided excellent results by comparing the classification accuracy and timing over the state-of-the-art.

References

1. L. Zhang, et al., Deep Learning for Remote Sensing Image Understanding, Sensors. (2016).
2. R. Shaheera, and D. Nicolas, A Split-and-Merge Approach for Hyperspectral Band Selection, IEEE Geoscience and Remote Sensing Letters. (2017).
3. M. Shafaey, et al. (2018). Comparison of CNNs for Remote Sensing Scene Classification. 13th Inter. Conf. on Comp. Eng. & Sys.
4. M. Shafaey, et al., (2019). Deep Learning for Satellite Image Classification. AISC.
5. C. W. Adam, G. A. Vincent, and A. H. Everett, Unmanned Aircraft Systems in Remote Sensing and Scientific Research: Classification and Considerations of Use, Remote Sensing. (2012).
6. M. Shafaey, et al., (2020). Hyperspectral Image Classification Using Deep Learning Technique. Proc. of the Intern. Conf. on AICV.
7. L. Hansen, et al., Neural Network Ensembles, IEEE Trans. on Pattern Analysis & Mach. Intell. (1990).
8. T. Guo, et al. (2017). Simple convolutional neural network on image classification. IEEE 2nd Inter. Conf. on Big Data Analysis.
9. F. Luus, B. Salmon, F. Van Den Bergh, and B. Maharaj, Multiview deep learning for land-use classification, IEEE Geosci. Remote Sens. Lett. (2015).

10. I. Ševo and A. Avramović, Convolutional Neural Network Based Automatic Object Detection on Aerial Images, IEEE Geosci. Remote Sens. Lett. (2016).
11. UC-Merceed: http://weegee.vision.ucmerced.edu/datasets/landuse.html, last access date: 2020/3/4.
12. R. Dong et al., Oil palm plantation mapping from high-resolution remote sensing images using deep learning, International Journal of Remote Sensing. (2019).
13. X. Liu et al., Automatic Detection of Oil Palm Tree from UAV Images Based on the Deep Learning Method, Applied Artificial Intelligence. (2020).
14. M. Castelluccio, et al. (2017). Training convolutional neural networks for semantic classification of remote sensing imagery. Joint Urban Remote Sens. Event.
15. C. Szegedy, et al, Going Deeper with Convolutions, CVF. (2015).
16. B. Zhao, Y. Zhong, and L. Zhang, A spectral–structural bag-of-features scene classifier for very high spatial resolution remote sensing imagery, ISPRS J. Photogramm. Remote Sens. (2016).
17. K. Simonyan, and A. Zisserman. (2015). Very deep convolutional networks for large-scale image recognition. in Proc. Int. Conf. Learn. Represent.
18. T. Ian, and J. Cadima, Principal component analysis: a review and recent developments, Phil. Trans. R. Soc. (2016).
19. D. Lee, J. In, and S. Lee, Standard deviation and standard error of the mean, Korean Journal of Anesthesiology. (2015).
20. N. Laban, et al, Enhanced pixel based urban area classification of satellite images using convolutional neural network, Intelligent Computing and Inf. Sciences. (2021).
21. D. Marmanis, Datcu, M., Esch, T., and Stilla, U, Deep Learning Earth Observation Classification Using ImageNet Pretrained Networks, IEEE Geosci. Remote Sens. Lett. (2015).
22. Salinas: https://rslab.ut.ac.ir/data, last access date: 2021/12/5.
23. Kenndy Space Center: https://rslab.ut.ac.ir/data, last access date: 2021/12/5.
24. Pavia University: https://rslab.ut.ac.ir/data, last access date: 2021/12/5.
25. Indian Pines: https://rslab.ut.ac.ir/data, last access date: 2021/12/5.
26. F. Habeeb, et al, Reducing error rate of deep learning using auto encoder and genetic algorithms, International J. of Intelligent Computing and Inf. Sciences. (2016).
27. Z. Wang, et al., Hyperspectral Image Classification Based on Spectral and Spatial Information Using Multi-Scale ResNet, Appl. Sci. (2019).
28. H. Zhang, et al., Hyperspectral classification based on light weight 3D-CNN with transfer learning, IEEE Trans. on Geo. and RS. (2020).
29. C. Li, et al., Hyperspectral Remote Sensing Image Classification Based on Maximum Overlap Pooling Convolutional Neural Network, J. of Sensors. (2018).
30. X. Yang et al., Hyperspectral Image Classification with Deep Learning Models, IEEE Trans. on Geo. and RS. (2018).
31. T. Hsieh, & J. Kiang, Comparison of CNN Algorithms on Hyperspectral Image Classification in Agricultural Lands, Sensors. (2020).