Hyperspectral Image Segmentation using End-to-End CNN Architecture with built-in Feature Compressor for UAV Systems

Muhammad Bilal[1], Khalid Munawar[2], Muhammad Shafique Shaikh[3], Ubaid M. Al-Saggaf [4], Belkacem Kada[5]
Department of Electrical and Computer Engineering[1, 2, 3, 4]
Department of Aerospace Engineering[5]
Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia[1, 2, 3, 4, 5]

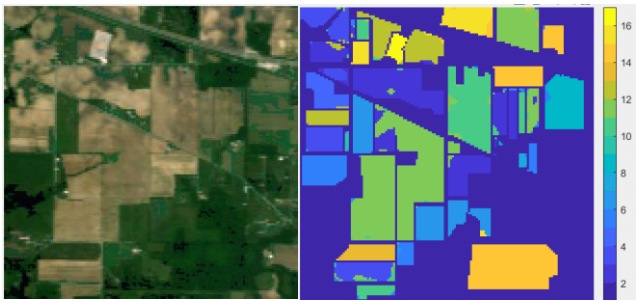
***Abstract*—Hyperspectral image segmentation is an important task for geographical surveying. Real-time processing of this operation is especially important for sensors mounted on-board Unmanned Aerial Vehicles in the context of visual servoing, landmarks recognition and data compression for efficient storage and transmission. To this end, this paper proposes a machine learning approach for segmentation using an efficient Convolutional Neural Network (CNN) which incorporates a feature compressor and a subsequent segmentation module based on 3D convolution operations. The experimental results demonstrate that the proposed approach gives segmentation accuracy at par with conventional approaches based on Principal Component Analysis (PCA) to reduce the feature dimensionality. Moreover, the proposed network is at least 35% faster than the conventional CNN-based approaches using 3D convolutions.**

***Keywords*—Hyperspectral images; CNN; dimensionality reduction; segmentation; PCA**

I. INTRODUCTION

Remote sensing is an important field which has helped urban and rural planning as well as environmental studies. Originally initiated as a satellite-based technology, it is being fast adopted to be used via sophisticated multispectral sensors that can be installed on Unmanned Aerial Vehicles (UAV) [1]. Other than being equipped with a standard RGB camera [2], several hyperspectral imaging (HSI) sensors [3] are now being deployed on-board these platforms to serve various applications in the field of security surveillance, town management, wild fires and agriculture etc. Another important application is to supplement the on-board navigation algorithms via visual servoing [4]. The spectral information provided by HSI is extremely rich and powerful. However, this raw information has to be processed extensively. Specifically, segmentation (classification) of each spatial data point in HSI has to be performed. Overall, the spatial and spectral data points constitute an information 'cube' [5]. Various researchers have put forward multiple segmentation methods of hyperspectral images to this end [6]. Various open-source datasets have also been made available to facilitate the research and development in this field [7]. Conventionally, HSI segmentation has been done through well-known machine learning methods such as Bayesian [8] distance, nearest-neighbor classification [9] and hand-coded features (e.g. Local Binary Pattern) [10] etc. Although, these methods have shown satisfactory performance on the standard datasets, recently CNN-based methods have shown exceptionally better results with near optimal segmentation [11, 12]. In this regard, 'Indian Pines' is a well-known HIS dataset [7] captured through Airborne Visible Infrared Imaging Spectrometer (AVIRIS) [13]. This scene (Fig. 1) consists of 145×145 pixels (spatial data points) and 224 spectral reflectance bands captured in the wavelength range [0.4 2.5] μ meters. This scene is marked by regions consisting of common agricultural land, a forest as well as wild natural vegetation. Other than this, highways, a railway line and some urban construction also exist. The ground truth (intended segmentation output) consists of sixteen classes (Table I).

Some earlier works have suggested reducing the number of spectral bands to *200* by removing bands corresponding to [104-108], [150-163], the so-called 'water absorption bands' to improve segmentation accuracy [14]. Moreover, it has been observed that there is a considerable spectral redundancy i.e. the observations in different spectral are highly correlated [15]. Thus, it has been found beneficial to consider dimensionality reduction through PCA before application of segmentation algorithms. Furthermore, it has been suggested to employ the spatial variations in spectral data, especially in the case of high-resolution data, to improve the segmentation accuracy. While, the former approach helps in increasing the inference time (lower computational load), the latter adds further computational complexity to the segmentation task. However, in order to improve segmentation performance, it has been proposed to use both spectral and spatial data while segmenting a particular data point [6]. Spatial information improves the segmentation accuracy because further discriminating information could be found in the texture and shape of geological structures. To illustrate this point, we have trained two classifiers on Indian Pines dataset. The results have been given in Table II.



a b

Fig. 1. Segmentation of a hyperspectral image (Indian pines). a) RGB image b) pseudocolor representation of segmented regions for 16 classes.

TABLE I. GROUND TRUTH DATA FOR INDIAN PINES DATASET

#	Class	Number of Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

TABLE II. REFERENCE SEGMENTATION CLASSIFIER ON INDIAN PINES DATASET

Classifier	Dimensionality	Segmentation Accuracy
SVM	220	96.0%
SVM	30	96.1%
CNN	30	99.6%

The first classifier is based on a conventional machine learning classifier, i.e. Support Vector Machine (SVM) [16]. In order to verify the strong redundancy in the spectral bands, two different versions have been trained. The first version uses all 220 hyperspectral bands (including noisy bands) while the second reduces the dimensionality of these spectral bands through PCA. PCA works by projecting the *d* dimensional data to *m* dimensional data where $m < d$.

$$\boldsymbol{x}_i \in \mathbb{R}^d, i = 1,2,\cdots N$$

$$\boldsymbol{z}_i \in \mathbb{R}^m, i = 1,2,\cdots N$$

The projection works by exploiting the eigen values and corresponding eigen vectors to build the transform as follows.

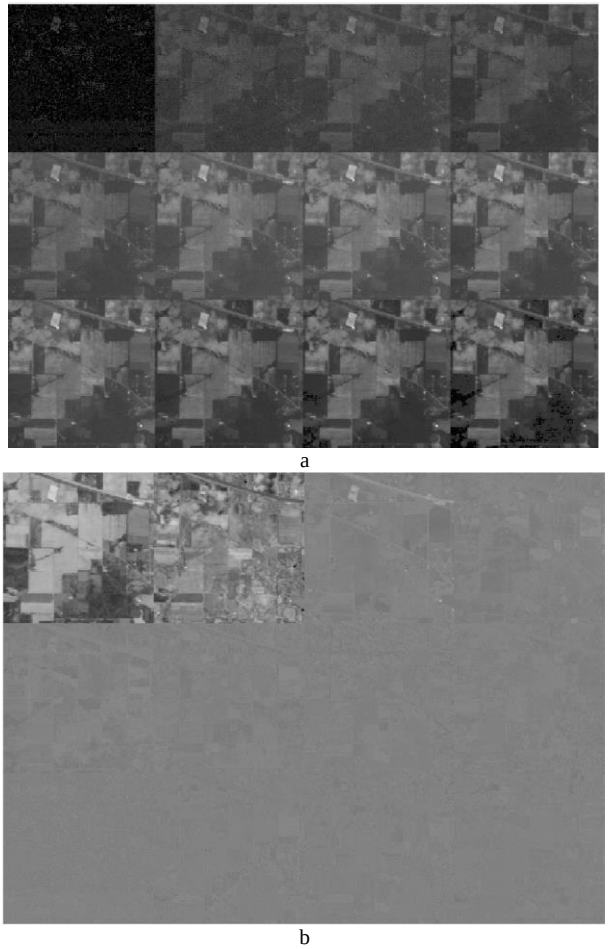
$$\boldsymbol{S}\boldsymbol{u}_{\boldsymbol{d}} = \lambda_d \boldsymbol{u}_{\boldsymbol{d}} \quad (1)$$

Where S is the normalized covariance matrix of the input features, λ_d are the eigen values while $\boldsymbol{u}_{\boldsymbol{d}}$ are the corresponding eigen vectors. The transformation is then given by.

$$z = \boldsymbol{u}_m^T \left(\frac{x_i - \mu}{\sigma}\right) \quad (2)$$

As can be noticed from the data given in Table II, the segmentation accuracy with and without dimensionality reduction is almost identical with the minor difference being without any statistical significance. The accuracy results have been obtained with the provided ground truth as benchmark.

Fig. 2 shows the original first twelve bands in the Indian Pines data as well as the twelve most significant bands after PCA has been applied. It can be clearly seen that the most information is contained within the first three bands only. Thus, the dimensionality reduction through PCA is merited to facilitate the further segmentation task. However, it should be noted that PCA does not ensure that the segmentation accuracy will remain intact in all the cases.



a b

Fig. 2. Application of PCA on HIS data from Indian Pines a) First twelve original bands b) Twelve transformed bands after PCA.

The second classifier considered in Table II is a CNN [6] classifier which also incorporates the spatial information while performing segmentation. In order to reduce the computational load, CNN has also been fed the same PCA-compressed spectral bands as for SVM classifier. It can be seen that this classifier almost achieves 100% accurate results. For both classifiers, 70% data points have been used for training while the remaining have been used for validation. Moreover, for CNN classifier, image patch size of 25×25 has been used to learn spatial information.

This approach of reducing the spectral information before applying the classifier has been considered by various early researchers [17]. The other approach of inferring spatial information before incorporating spectral data has also been considered in some works [18]. Adaptive Markov Random Fields [19] and graph cut [20] algorithms have also been considered to regularize the segmentation results in the spatial domain. Joint processing of spatial and spectral features has been considered by Zhon et al. [21] and Guijarro et al. [22]. These efforts to combine spatial and hyperspectral data to extract joint features is especially important since latent redundancy could be exploited. As noted earlier, HIS data cube hold valuable information in the local neighborhoods both with respect to spatial and spectral dimensions. However, this also presents a challenge in terms of computation complexity. This is especially relevant for on-board processing of data on battery-powered UAVs.

In contrast to the conventional machine learning models, deep learning or CNN-based approaches have been shown to be exceptionally well in exploring the data dependencies and discriminating features. Thus, several approaches have been proposed by researchers in the recent years related to this technique [23-25]. In this regard, U-Net [26-28] is of special interest because it has been successfully applied to image segmentation tasks other than HSI. However, since like other CNN-based segmentation frameworks, the computational load is a consideration for these methods, PCA is generally applied to reduce the HSI data dimensionality as shown above. However, since PCA and the subsequent CNN deep feature extraction modules are separate, the joint spatial-spectral redundant information [27, 29] may not be captured for optimal segmentation. To this end, Ying et al. [6] proposed using 3D CNN operations for HSI segmentation. 3D convolution kernels can learn the patterns found in both the spatial and the spectral dimensions of the HSI cube in local neighborhoods. This information has been shown to be of crucial value for segmentation. This structure can take the full advantage of all structural similarities within the HIS cube. 3D CNNs have shown promise in a variety of computer vision applications where the input is simple RGB images. Moreover, they have been applied to video frames to make the best use of temporal information. While using 3D CNNs for HSI segmentation, not only the pre-processing (PCA) overhead is reduced, but also the spectral and spatial features are learnt jointly. However, this presents a computational overhead since 3D convolutions are even more complicated than their 2D counterpart. In this paper, we have proposed an approach to circumvent this problem by adding an '*Encoder'* stage before 3D CNNs. The '*Encoder'* stage is inspired by the autoencoders [30] which have been shown to infer even non-linear statistical dependencies. In contrast PCA can only infer linear correlations. The proposed 'Encoder' stage compresses the spectral features, just like PCA, but in tandem with the subsequent segmentation stage based on 3D CNN. In this regard, Khan et al. [31, 32] have proposed using a traditional Autoencoder for the segmentation of HIS but they do not explicitly target the improvement in the inference speed. On the other hand, the experimental results demonstrate that the proposed architecture works at par with the state of the art while being up to 35% faster.

The remaining paper has been structured as follows. Section II describes the proposed architecture in detail with respect to the inclusion of encoder section before a 2D CNN. Section III has been devoted to the provision of results of the proposed architecture on standard datasets followed by a discussion. Section IV concludes the discussion by summarizing the results and contributions of the current work.

II. PROPOSED ARCHITECTURE

Autoencoders have been extensively used in computer vision applications to learn the underlying data decencies and have been shown to work better than PCA in the presence of non-linear structures. Autoencoders are made up of two parts i.e. encoder and decoder. The encoder finds a compressed representation of the input while the decoder replicates the input using the compressed (low dimensional) encoding. Inspired by their architecture, we propose adding only the 1D encoding stage before a 3D CNN stage for HIS cube processing. The idea is to find non-linear data decencies in the spectral bands data which is then subsequently processed by the 3D CNN architecture to exploit spatial information as well. However, since the two modules are part of one learnable architecture, both spatial and spectral information are learnt jointly. Moreover, it is faster since the early part of the network (closer to the input) only processes the spectral dimension. In contrast the 3D CNN architecture proposed in [6] processes the whole cube starting from the input which leads to a high computational load.

The proposed architecture depicted in Fig. 3 consists of two stages as mentioned earlier. The 'Encoder' stage reduces the spectral dimension in two stages. The first stage uses 3×3 convolutions with a stride of two to reduce the dimensions from 220 to 110. The second stage uses 3×3 convolutions with a stride of fourj to reduce the dimensions from 110 to 27. The spatial dimensions (25×25) remain the same. Thus, at the end of the 'Encoder' stage, a 25×25×27 cube is presented the 3D CNN stage for further segmentation. This network consists of four 3D convolutional layers i.e. 3×3×7, 3×3×5, 3×3×3 and 3×3×1. There are 8, 16, 32 and 8 filter kernels for these layers respectively. A 256 long vector is generated through a fully connected layer followed by another fully connected layer generating a 128 long vector. The final vector (16 output classes) is generated through another fully connected layer. Thus, the 25×25×220 HSI input cube is gradually processed through convolutions to generate the final segmentations. The hallmark of this architecture, as mentioned earlier, is that the 'Encoder' stage is light-weight since it only processes the spectral data. However, despite only processing spectral data, it

is connected with the subsequent spatial processing layers as well through the forward propagation. Thus, the whole network learns spectral and spatial information jointly without being computationally complex. The proposed architecture has been implemented using the Matlab Deep Learning Toolbox and trained in the same environment using NVIDIA GeForce 940 MX GPU on a intel i7 powered CPU (3.48 GHz) with 12 GB RAM. For training and testing of the proposed architecture, Indian Pine, Salinas, SalinasA, Pavia and PaviaU datasets [7] have been used. A sample training session has been shown in Fig. 4. It can be clearly seen that the training process quickly converges for both training and validation sets. Thus, not only is the proposed architecture suitable for the task at hand but also does not overfit showing excellent harmony between spatial and spectral data points.

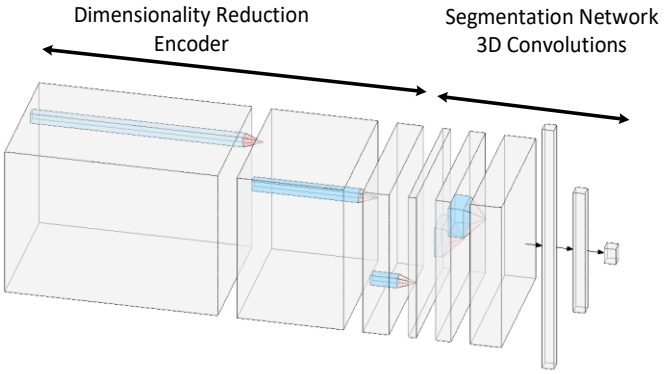


Fig. 3. Proposed HSI cube segmentation through Encoder + 3D Convolution stages.

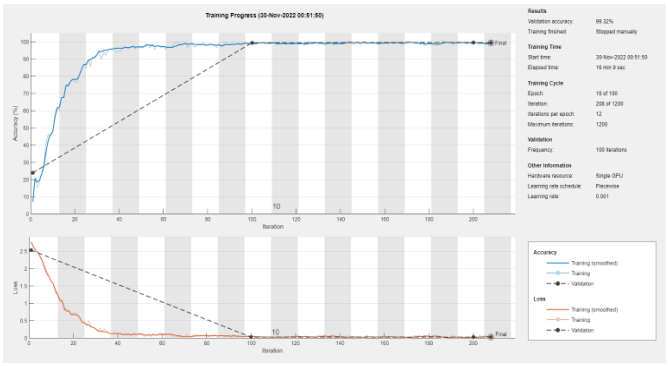


Fig. 4. Training of the proposed network in Matlab environment depicting fast convergence without overfitting.

III. Results and Discussion

The proposed architecture has been trained and tested on Indian Pine, Salinas, SalinasA, Pavia and PaviaU datasets. The former three datasets have been captured by AVIRIS sensor with 220 bands. The latter two datasets are captured by ROSIS sensor and have 102 and 103 bands respectively. To cater for the fewer number of bands on ROSIS datasets, the first stage of the 'Encoder' uses 3×3 convolutions with a stride of two to reduce the dimensions from 102/103 to 50. The second stage uses 3×3 convolutions with a stride of two to reduce the dimensions from 50 to 25. The rest of the architecture remains same.

The test results have been reproduced in Table III. It can be seen that the proposed network either performs better or at par with the 3D CNN network. Moreover, it is worth mentioning that this particular network has been tuned for each dataset separately by the authors [6] while we have used the same universal architecture for all cases. Thus, the proposed architecture is universally applicable to all scenarios.

It is worth mentioning that the proposed architecture also operates faster than the competing 3D CNN architecture due to its simpler 'Encoder' stage. On average, the proposed network processes an HSI cube of size 145×145×220 in *7.14 s* while a 3D CNN without encoder stage processes the same in *9.64 s*. Thus, the proposed approach is at least *35%* faster while being equally accurate.

Fig. 5 shows the reconstruction of spectral data for four example data points in Indian Pine dataset. This data has been generated by a reference Autoencoder with both 'Encoder' and'Decoder' parts for visualization. It can be seen that the spectral information can be fairly accurately recovered even after being compressed to 27 channels from original 220 (12%). However, it must be noted that the purpose in the application at hand is not reconstruction but use in subsequent segmentation where upto 100% accuracy has been achieved on SalinasA dataset even with this much compression. Visual results using psuedocolors for segmentation results on the mentioned datasets have been depicted in Fig. 6, Fig. 7, Fig. 8 and Fig. 9 respectively.

TABLE III. Segmentation Accuracy on Test Datasets

Architecture	Indian Pine	Salinas	SalinasA	Pavia	PaviaU
3D CNN [6]	99.07%	-	-	-	99.4%
Proposed	99.6%	94.3%	100%	98.0%	98.3%

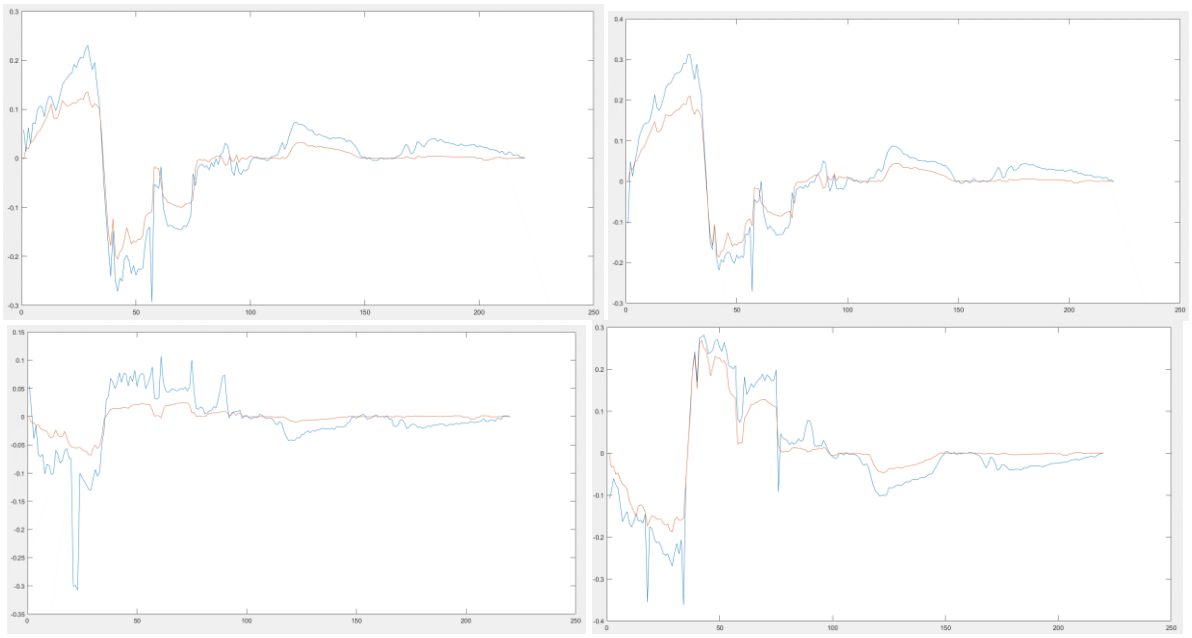
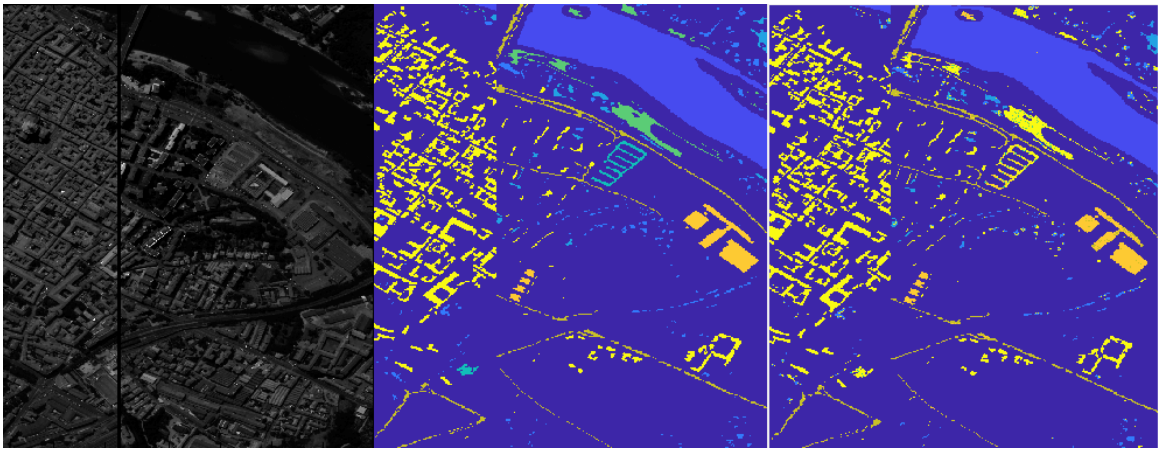
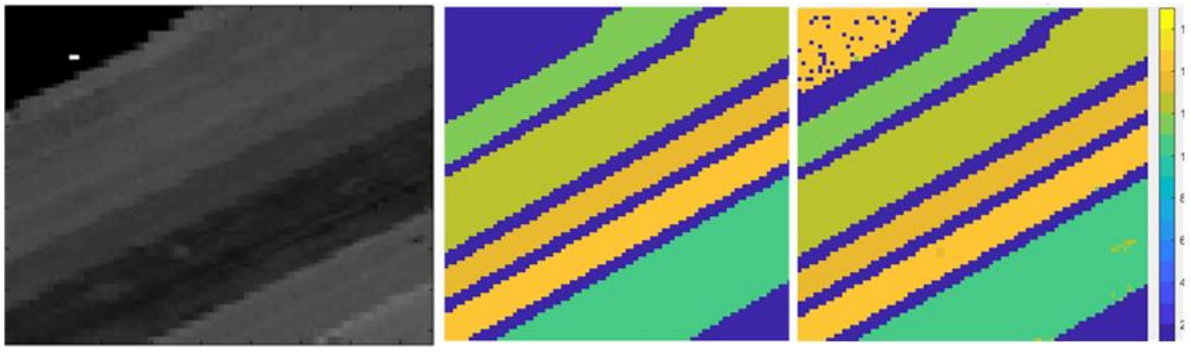


Fig. 5. Reconstruction of the spectral data for 4 sample spatial points. Original-Blue, Reconstructed-Red.



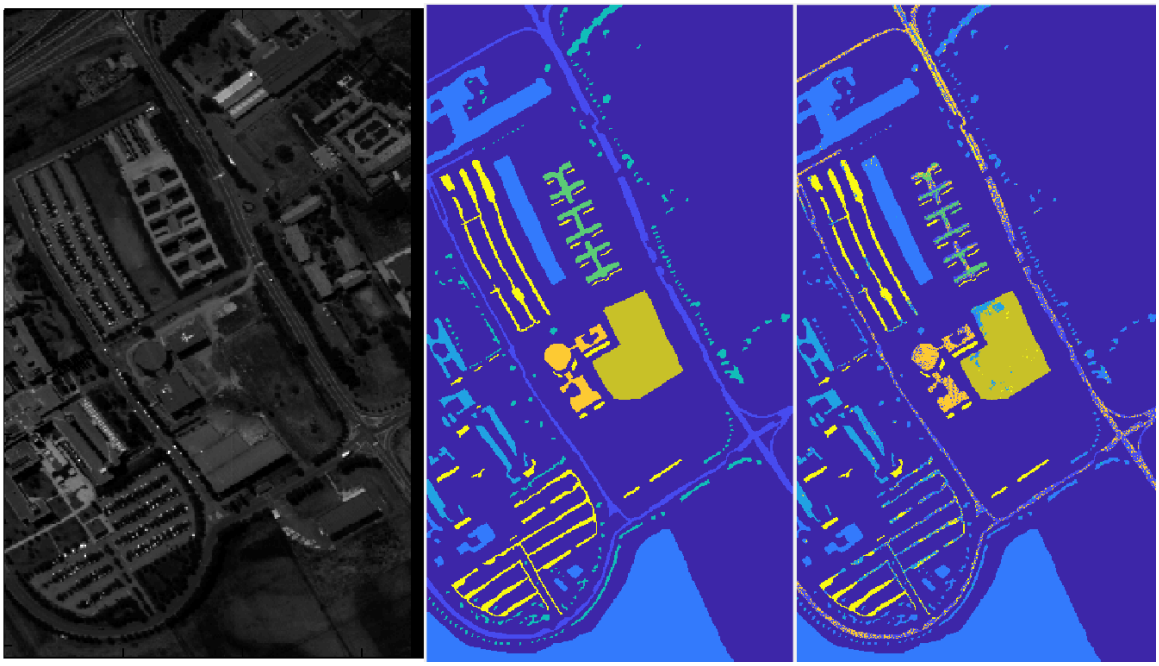
Input Grayscale Ground Truth Proposed Approach

Fig. 6. Pavia centre segmentation.



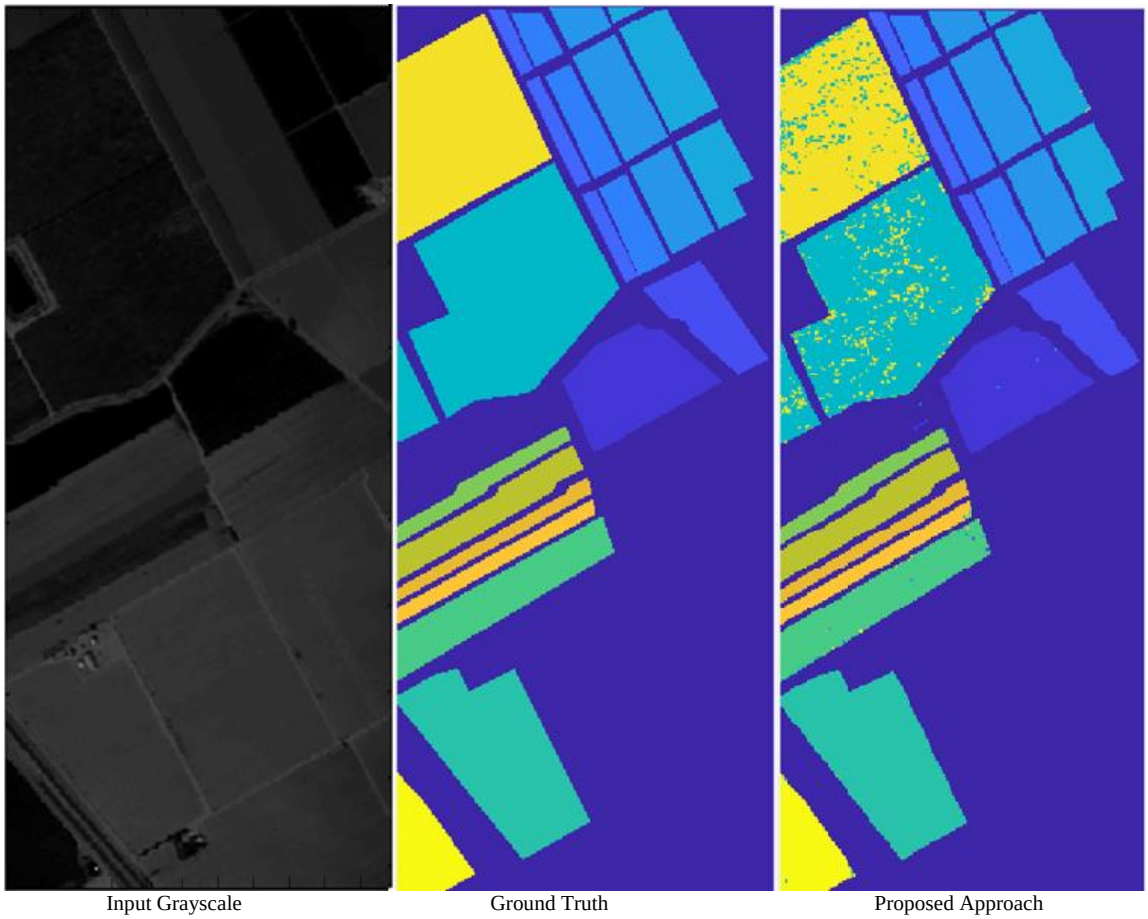
Input Grayscale Ground Truth Proposed Approach

Fig. 7. SalinasA Segmentation.



Input Grayscale Ground Truth Proposed Approach

Fig. 8. Pavia University segmentation.



Input Grayscale Ground Truth Proposed Approach

Fig. 9. Salinas segmentation.

IV. CONCLUSION

This paper has proposed a 3D convolution-based segmentation architecture for hyperspectral images. The proposed architecture achieves detection accuracy at-par with the state of-the-art with the added benefit that it achieves up to 35% higher speed due to a feature compressor module at the input. Moreover, this module works almost as good as conventional PCA in compressing the spectral information while being trained in tandem with the spatial data. Thus, the network has the ability to jointly learn spatial and spectral information. In conclusion, the proposed architecture provides a better alternative than both PCA-CNN and 3D-CNN architectures, combining the best of both approaches in terms of computational speed, learning flexibility and segmentation accuracy. For future work, the proposed technique could be applied to HS images from more satellite imagers to test the generalization of the approach.

ACKNOWLEDGMENT

This project was funded by the National Plan for Science, Technology and Innovation (MAARIFAH) – King Abdulaziz City for Science and Technology - the Kingdom of Saudi Arabia – award number (11-SPA2059-03). The authors also, acknowledge with thanks Science and Technology Unit, King Abdulaziz University for technical support.

REFERENCES

[1] K. A. M. F. Asmala Ahmad, Mohd Mawardy Abdullah, Suliadi Firdaus Sufahani, Mohd Yazid Abu Sari and Abd Rahman Mat Amin, "Noise and Restoration of UAV Remote Sensing Images," International Journal of Advanced Computer Science and Applications, vol. 11, no. 12, 2020.

[2] G. Zhang et al., "UAV Low-Altitude Aerial Image Stitching Based on Semantic Segmentation and ORB Algorithm for Urban Traffic," Remote Sensing, vol. 14, no. 23. doi: 10.3390/rs14236013.

[3] J. M. Haut and M. E. Paoletti, "Cloud Implementation of Multinomial Logistic Regression for UAV Hyperspectral Images," IEEE Journal on Miniaturization for Air and Space Systems, vol. 1, no. 3, pp. 163-171, 2020.

[4] J. Pomares, "Visual Servoing in Robotics," Electronics, vol. 8, no. 11. doi: 10.3390/electronics8111298.

[5] G. Martín, J. M. Bioucas-Dias, and A. Plaza, "HYCA: A New Technique for Hyperspectral Compressive Sensing," IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 5, pp. 2819-2831, 2015.

[6] Y. Li, H. Zhang, and Q. Shen, "Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network," Remote Sensing, vol. 9, no. 1. doi: 10.3390/rs9010067.

[7] M. Graña, M. Veganzons, and B. Ayerdi. (2022). Hyperspectral Remote Sensing Scenes. Available: https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.

[8] A. Ahmad, H. Sakidin, M. Y. A. Sari, A. R. M. Amin, S. F. Sufahani, and A. W. Rasib, "Naïve Bayes Classification of High-Resolution Aerial Imagery," International Journal of Advanced Computer Science and Applications, vol. 12, no. 11, 2021.

[9] B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou, and L. Guo, "KNN-Based Representation of Superpixels for Hyperspectral Image Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 11, pp. 4032-4047, 2018.

[10] W. Li, C. Chen, H. Su, and Q. Du, "Local Binary Patterns and Extreme Learning Machine for Hyperspectral Imagery Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 7, pp. 3681-3693, 2015.

[11] Y. Kong, X. Wang, and Y. Cheng, "Spectral–Spatial Feature Extraction for HSI Classification Based on Supervised Hypergraph and Sample Expanded CNN," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 11, pp. 4128-4140, 2018.

[12] X. Xue, H. Zhang, B. Fang, Z. Bai, and Y. Li, "Grafting Transformer on Automatically Designed Convolutional Neural Network for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022.

[13] C. J. Bruegge et al., "Vicarious Calibration of eMAS, AirMSPI, and AVIRIS Sensors During FIREX-AQ," IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 12, pp. 10286-10297, 2021.

[14] J. Wang, J. Zhou, and W. Huang, "Attend in Bands: Hyperspectral Band Weighting and Selection for Image Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 12, pp. 4712-4727, 2019.

[15] X. Shang et al., "Target-Constrained Interference-Minimized Band Selection for Hyperspectral Target Detection," IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 7, pp. 6044-6064, 2021.

[16] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-Based Method for Accurate Classification of Hyperspectral Images," IEEE Geoscience and Remote Sensing Letters, vol. 7, no. 4, pp. 736-740, 2010.

[17] M. D. Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of Hyperspectral Images by Using Extended Morphological Attribute Profiles and Independent Component Analysis," IEEE Geoscience and Remote Sensing Letters, vol. 8, no. 3, pp. 542-546, 2011.

[18] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 480-491, 2005.

[19] B. Zhang, S. Li, X. Jia, L. Gao, and M. Peng, "Adaptive Markov Random Field Approach for Classification of Hyperspectral Imagery," IEEE Geoscience and Remote Sensing Letters, vol. 8, no. 5, pp. 973-977, 2011.

[20] Y. Tarabalka and A. Rana, "Graph-cut-based model for spectral-spatial classification of hyperspectral images," in 2014 IEEE Geoscience and Remote Sensing Symposium, 2014, pp. 3418-3421.

[21] Y. Zhong, A. Ma, and L. Zhang, "An Adaptive Memetic Fuzzy Clustering Algorithm With Spatial Information for Remote Sensing Imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 4, pp. 1235-1248, 2014.

[22] M. Guijarro, G. Pajares, and P. J. Herrera, "Image-Based Airborne Sensors: A Combined Approach for Spectral Signatures Classification through Deterministic Simulated Annealing," Sensors, vol. 9, no. 9, pp. 7132-7149. doi: 10.3390/s90907132.

[23] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 152, pp. 166-177, 2019/06/01/ 2019.

[24] B. Xue et al., "A Subpixel Target Detection Approach to Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 5093-5114, 2017.

[25] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "Feedback Attention-Based Dense CNN for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022.

[26] F. Guo, Z. Li, Z. Xin, X. Zhu, L. Wang, and J. Zhang, "Dual Graph U-Nets for Hyperspectral Image Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 8160-8170, 2021.

[27] Q. Liu, L. Xiao, J. Yang, and Z. Wei, "Multilevel Superpixel Structured Graph U-Nets for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-15, 2022.

[28] S. Hao, W. Wang, and M. Salzmann, "Geometry-Aware Deep Recurrent Neural Networks for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 3, pp. 2448-2460, 2021.

[29] W. Zhao and S. Du, "Spectral–Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep

Learning Approach," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4544-4554, 2016.

[30] X. Ma, H. Wang, and J. Geng, "Spectral–Spatial Classification of Hyperspectral Image Based on Deep Auto-Encoder," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 9, pp. 4073-4085, 2016.

[31] S. D. Khan, L. Alarabi, and S. Basalamah, "An Encoder–Decoder Deep Learning Framework for Building Footprints Extraction from Aerial Imagery," Arabian Journal for Science and Engineering, 2022/03/29 2022.

[32] S. D. Khan, L. Alarabi, and S. Basalamah, "Deep Hybrid Network for Land Cover Semantic Segmentation in High-Spatial Resolution Satellite Images," Information, vol. 12, no. 6. doi: 10.3390/info12060230.