MULTIFREQUENCY POLSAR IMAGE CLASSIFICATION USING DUAL-BAND 1D CONVOLUTIONAL NEURAL NETWORKS

Mete Ahishali[1], Serkan Kiranyaz[2], Turker Ince[3], and Moncef Gabbouj[1]

[1]Department of Computing Sciences, Tampere University, Tampere, Finland
[2]Department of Electrical Engineering, Qatar University, Doha, Qatar
[3]Department of Electrical and Electronics Engineering, Izmir University of Economics, Izmir, Turkey

ABSTRACT

In this work, we propose a novel classification approach based on dual-band one-dimensional Convolutional Neural Networks (1D-CNNs) for classification of multifrequency polarimetric SAR (PolSAR) data. The proposed approach can jointly learn from C- and L-band data and improve the single band classification accuracy. To the best of our knowledge, this is the first study that introduces 1D-CNNs to land use/land cover classification domain using PolSAR data. The proposed approach aims to achieve maximum classification accuracy by one-time training over multiple frequency bands with limited labelled data. Moreover, the proposed dual-band 1D-CNN approach yields a superior computational efficiency compared to the deep 2D-CNN based approaches. The performed experiments using AIRSAR PolSAR image over San Diego region at C- and L-bands have shown that the proposed approach is able to simultaneously learn from the C- and L-band SAR data and achieves an elegant classification performance with minimal complexity.

***Index Terms*—** Polarimetric Synthetic Aperture Radar (PolSAR), 1D Convolutional Neural Networks, land use/land cover classification, multifrequency classification

1. INTRODUCTION

Synthetic Aperture Radar (SAR) has been extensively used in many applications for earth surface observation since it is an active remote sensing method, and hence unlike optical sensors it can operate under any weather circumstances and day/night conditions. There are a considerable amount of studies [1]–[5] about accurate land use/land cover classification of Polarimetric SAR (PolSAR) images, which makes the terrain classification a crucial task. Recent developments in PolSAR imaging enables to use of many target decomposition (TD) theorems that can be utilized for the classification task. In general, the existing conventional classification approaches [2], [5], [6] can be grouped in terms of their way to describe multi-dimensional SAR data. For instance, some SAR classification methods only use low-level SAR EM features, while the majority proposes to use numerous TD theorems to extract various polarimetric features. These conventional approaches have two separate stages: feature extraction and feature classification, and they [2], [5], [6] are based on to use as many features as possible to increase the classification accuracy.

Convolutional Neural Networks (CNNs) [7] have become the *de-facto* standard in many visual recognition tasks including image segmentation, recognition, and object tracking. However, deep CNNs usually require a massive amount of data for training, and this drawback is also valid for recently proposed SAR classification methods [8], [9] with deep 2D CNNs that used 75% or even higher of the SAR data for training. However, in practice, such proportion of annotated SAR data is unfeasible because of the effort and cost of the labelling process in remote sensing. Nevertheless, the study in [1] proposes to use compact CNN configurations to cope with the limited number of training data. Instead of conventional deep 2D-CNNs, several methods propose to use 1D-CNNs: real-time ECG classification [10], motor-fault detection [11] and many other signal processing applications [12]. Overall, these studies focus on the computational efficiency of 1D convolutional operations. Although 1D-CNNs have been used in remote sensing for hyperspectral image classification [13], we believe that 1D-CNNs are not fully studied for SAR terrain classification and the proposed approach is a pioneer study in this domain.

Using multifrequency polarimetric data for SAR classification has been studied in [3], [4]. These studies use P-, L-, and C-band PolSAR data together and evaluate the performance improvements over Flevoland and San Francisco AIRSAR images. Different frequency bands have different sensitivities to various radar targets due to different penetration power and their scattering characteristics can be better distinguished through various combinations of data from available frequency bands. In this work, we propose a novel approach based on compact 1D-CNNs. The proposed classification framework for PolSAR images has two branches which can take multifrequency band input; the first branch takes C-band and the second takes L-band PolSAR data. The performance comparisons are made with implemented single branch 1D-CNN classifier over San Diego region AIRSAR data, and it is shown that the classification accuracy is significantly improved compared to the single-band accuracies.

2. POLSAR DATA PROCESSING

PolSAR systems measure the complex 2×2 backscattering matrix [S], in fully polarization mode:

$$E^r = \begin{bmatrix} |S_{hh}|e^{j\phi_{hh}} & |S_{hv}|e^{j\phi_{hv}} \\ |S_{vh}|e^{j\phi_{vh}} & |S_{vv}|e^{j\phi_{vv}} \end{bmatrix} \begin{bmatrix} E_h^t \\ E_v^t \end{bmatrix} \quad (1)$$

In monostatic system configuration $S_{hv} = S_{vh}$ holds [14]. The target is now represented by three absolutes $|S_{hh}|, |S_{vv}|$ and $|S_{hv}|/|S_{vh}|$ and two relative phases $\phi_{vv-hh}, \phi_{hv-hh}$. PolSAR classification methods generally utilize two types of polarimetric information: 1) the information directly extracted from PolSAR data, backscattering matrix, and its second-order descriptors, 2) the TDs consisting of coherent and incoherent decompositions.

Coherent TDs try to express [S] matrix as the sum of coherent scattering responses; for example, Pauli decomposition [14], in which vector *k* is computed with by the summation of the diagonal elements after the defined transformation:

$$k = V([S]) = Tr([S]\varphi_P) \ where \ \varphi_P = \left\{ \begin{matrix} [S]_a = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, [S]_b = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \\ [S]_c = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{matrix} \right\} \quad (2)$$

Since SAR sensors suffer from speckle noise, the data is often multi-looked processed by averaging *n*-looks. Using Pauli based scattering vector, the multi-look coherency matrix can be written for pixel *i* as the following: $\langle [T] \rangle = \frac{1}{n}\sum_{i=1}^{n} k_i k_i^{*T}$, where $k_i = [S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv}]^T / \sqrt{2}$ from Pauli decomposition in (2). The lexicographic basis matrices [14], φ_L, can also be used to compute the covariance matrix as $\langle [C] \rangle = \langle \varphi_L . \varphi_L{}^T \rangle$.

Incoherent TDs [14] analyze distributed scatters by decomposing incoherently averaged second-order target descriptors (the $\langle [T] \rangle$ and $\langle [C] \rangle$ matrices). For instance, the Cloude-Pottier decomposition [14] performs eigenanalysis over $\langle [T] \rangle$, using the following orthonormal eigenvectors:

$$\langle [T] \rangle = [U_3]\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}[U_3]^{*T}, \ where \ [U_3] = \begin{bmatrix} cos\alpha_1 & \cos\alpha_2 & cos\alpha_3 \\ sin\alpha_1 cos\beta_1 e^{i\delta_1} & sin\alpha_2 cos\beta_2 e^{i\delta_2} & sin\alpha_3 cos\beta_3 e^{i\delta_3} \\ sin\alpha_1 sin\beta_1 e^{i\gamma_1} & sin\alpha_2 sin\beta_2 e^{i\gamma_2} & sin\alpha_3 sin\beta_3 e^{i\gamma_3} \end{bmatrix} \quad (3)$$

Several polarimetric parameters are computed by Cloude-Pottier's eigenanalysis to exploit scattering characteristics of the medium: the entropy *H*, anisotropy *A*, and the average set of four angles $\underline{f} = [\bar{\alpha} \ \bar{\beta} \ \bar{\delta} \ \bar{\gamma}]$ for the real eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$:

$$H = -\sum_{i=1}^{3} p_i \log_3 p_i, \quad where \ p_i = \frac{\lambda_i}{\sum_{i=1}^{3}\lambda_i},$$
$$\bar{\alpha} = \sum_{i=1}^{3} p_i \alpha_i, \quad \bar{\beta} = \sum_{i=1}^{3} p_i \beta_i, \quad (4)$$
$$\bar{\delta} = \sum_{i=1}^{3} p_i \delta_i, \bar{\gamma} = \sum_{i=1}^{3} p_i \gamma_i, \quad A = \frac{(p_2 - p_3)}{(p_2 + p_3)}$$

In addition to explained above features, the three complex correlation coefficients, $(\rho_{12}, \rho_{13}, \rho_{23})$, calculated over the backscattering matrix, and the total scattering power *Span* are used in this work for the proposed classification approach. *Span* can be computed as: $Span = \sum_{i=1}^{3}\lambda_i = |\alpha|^2 + |\beta|^2 + |\gamma|^2$.

3. THE PROPOSED METHOD

The proposed approach is based on using dual-band 1D-CNN classifier for the multifrequency classification of PolSAR data as presented in Fig. 1. In the classification framework, after the data processing step, each labelled pixel is used to train the classifier. After training the 1D-CNN, to obtain final terrain segmentation mask, each individual pixel is fed to the classifier to predict the terrain labels. As shown in Fig. 1, the upper branch of the 1D-CNN takes C- band, and the secondary branch takes L-band PolSAR images. The dual-band classifier is then trained simultaneously using the samples from both bands. Once the training is completed, the same classifier is used over both band images. This brings two advantages over single band classification: joint learning from two bands significantly improves the overall classification accuracy, and there is no need to retrain the classifier for the other band. In the proposed approach, there is only one hyper-parameter needed to be adjusted besides the network configuration: the number of EM features, which is equal to the size of the input signal, is varied from 12 to 28 in the experiments. The effect of varied input size on classification accuracy will be discussed in Section 4. Note that methods based on 2D-CNNs, i.e., [1], [8], [9] have to set the size (*N*) of neighborhood as well, and it is generally kept high since deep learning methods consist of many layers. As a consequence, using a large neighborhood deteriorates the classification resolution and especially the finer details in the segmentation mask.

In this study, we propose the following compact 1D-CNN configuration to make the learning possible with a minimal amount of labelled data. 1) Two input layers, 2) two convolutional layers per branch having 80 neurons each with 1×3 weight kernels and *tanh* activation functions, 3) Concatenation of the output of last convolution layers, 4) a dense layer with 40 hidden neurons and an output layer with SoftMax activation function. Note that the number of neurons at the output is set according to the number of classes, and sub-sampling factor of the last convolution is adjusted based on the varied size of the input signal to produce scalar values for the dense layer. On the other hand, the single band classification results are obtained using the same 1D-CNN network configuration but with a single branch.

Moreover, unlike 2D-CNN based methods [1], 1D convolutions improve computational efficiency. Considering inference complexity, the proposed dual-band 1D-CNN network requires only ~450K operations for convolutional part (436K multiplication + 14K addition) which is two times less than the compact network in [1].

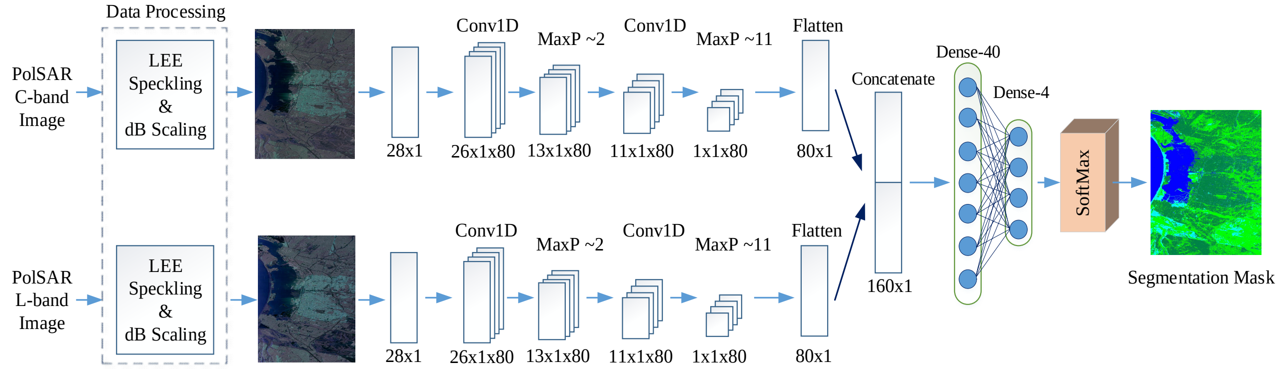
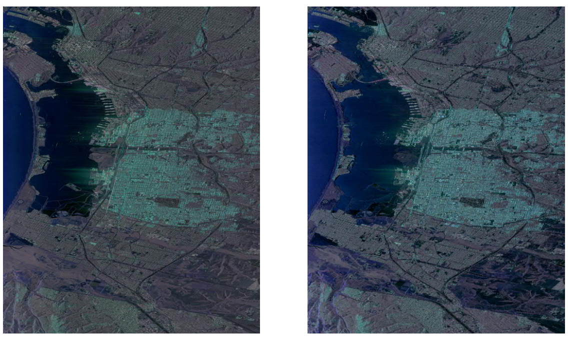


Fig. 1. The proposed classification approach for multifrequency PolSAR images using dual-band 1D-CNN classifier with 1×3 filter size. The dual-band 1D-CNN takes two inputs and jointly learns from C- and L-band PolSAR data.

4. EXPERIMENTAL RESULTS

In this work, San Diego PolSAR images obtained by AIRSAR of NASA Jet Propulsion Laboratory are used at C- and L-bands which is presented in Fig. 2 for the performance evaluations. Correspondingly, a sub-region is selected for the experiments with 2437×3109 dimensions. Both C- and L-band images are filtered with Lee's speckle filter [14]. The ground-truth data (GTD) are constructed by visual inspection of Pauli RGB coded images. Accordingly, the five main classes, including natural and human-made terrain types, are *water*, *urban*, *flat zone*, and *forest*. The number of samples selected for training is 9200, and whole GTD size is 216440. Notice only ~4.5% of GTD is used in the training which corresponds to 0.12% of entire PolSAR data. Using such compact 1D-CNN configuration allows training with such limited annotated data.

In this section, the performance gain of combining multiple bands with the proposed approach is presented both quantitatively and qualitatively by presenting improved classification accuracy compared to single-band results with 1D-CNN classifier and visual analysis of final segmentation masks, respectively.



(a) (b)

Fig. 2. San Diego PolSAR image at C-band (a) and L-band (b), diagonal elements of $\langle[\boldsymbol{T}]\rangle$ are assigned to R, G, B channels for the illustration purpose.

The experiments in this work are carried out on a workstation with Intel ® Xeon(R) CPU E5-2637 v4 at 3.50GHz, four Nvidia® TITAN-X GPU cards and 128 GB memory. Tensorflow [15] library is used with Python. We utilize ADAM optimizer for the training process using the default parameters (learning rate=0.001, β_1=0.9, and β_2=0.999) with only 20 Back-Propagation iterations.

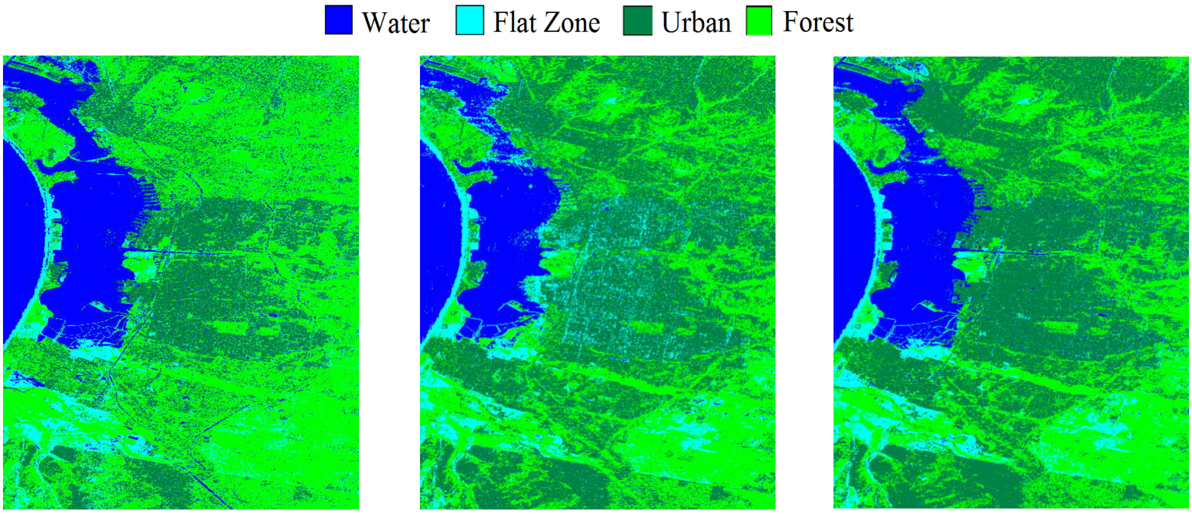
Features used to form the input signal for 1D-CNN classifier are grouped into four sets as shown in Table 1. In the experiments, these features are used to demonstrate that combining C- and L-bands with dual branch 1D-CNNs increases the classification performance. Initially, we have started from a 12-D setup: diagonal elements of $\langle[T]\rangle$ and complete $\langle[C]\rangle$, and then increase the dimension size to 28 by iteratively concatenating feature vectors. Thus we aim to demonstrate that the proposed compact CNN configuration is able to work with such a small number of features as observed in Table 2. When two bands are combined by the proposed approach, the classification accuracies are improved significantly for each feature set compared to single-band results. The maximum accuracy is achieved as 97% with 22-D features, whereas L-band classification can achieve only 94% with 28-D features. In addition to quantitative comparisons, the final segmentation masks of single-band and dual-band classifications are given in Fig. 3. The clear differences in the masks: L-band classification failed mostly with *water* class in Fig. 3b, while the proposed approach, in Fig. 3c, significantly improves the segmentation of *urban* class which was confused by *forest* in both single-band segmentation masks.

Table 1. Collected features for the classification.

	Features	Dimension
FV1	$[T]$ and $[C]$ Matrices	12
FV2	Span, $H/A/Alpha$	7
FV3	Eigenanalysis - Eigenvalues	3
FV4	Correlation Coefficients	6

Table 2. Classification accuracies with different feature sets (FS) over San Diego AIRSAR data at C- and L-bands using 1D-CNN, and improved accuracy by combining C- & L-bands with the proposed dual-band 1D-CNN.

	Features	C-band	L-band	C & L-bands
FS1	FV1	0.8561	0.8279	**0.915**
FS2	FV2 + FV1	0.9221	0.8884	**0.9456**
FS3	FV3 + (FS2)	0.9208	0.8867	**0.9723**
FS4	FV4 + (FS3)	0.9376	0.9399	**0.9535**



(a) (b) (c)

Fig. 3. The segmentation masks are obtained using single-band 1D-CNN approach in (a) and (b). The mask in (a) is obtained using C-band data with FS3, and the mask in (b) using L-band data with FS4. The proposed dual-band 1D-CNN approach is used to learn from C- and L-bands simultaneously during the training and produced mask given in (c)**.**

5. CONCLUSIONS

We have proposed a novel approach for multifrequency classification of PolSAR images. The proposed approach consists of dual-band 1D-CNNs, which take inputs from both C- and L-bands and perform learning simultaneously from both image samples. The experimental evaluations show that with minimal network complexity and training data, the proposed approach can achieve about 97% classification accuracy. Moreover, the classification accuracy is significantly improved by the proposed approach with multifrequency data compared to single-band accuracies with 1D-CNN classifier. The proposed approach inherits the same characteristics from conventional CNNs; therefore, learn-to-extract feature-capability allows that the proposed approach can perform classification with a small number of features. In the future work, three bands (C, L, and P) will be studied for achieving further performance improvements.

6. REFERENCES

[1] M. Ahishali, S. Kiranyaz, T. Ince, and M. Gabbouj, "Dual and single polarized sar image classification using compact convolutional neural networks," *Remote Sens.*, vol. 11, no. 11, p. 1340, 2019.

[2] T. Ince, M. Ahishali, and S. Kiranyaz, "Comparison of polarimetric SAR features for terrain classification using incremental training," in *Proc. of PIERS*, 2017, pp. 3258–3262.

[3] F. Yang, W. Gao, B. Xu, and J. Yang, "Multi-frequency polarimetric SAR classification based on Riemannian manifold and simultaneous sparse representation," *Remote Sens.*, vol. 7, no. 7, pp. 8469–8488, 2015.

[4] H. T. Teng, H. T. Ewe, and S. L. Tan, "Multifractal dimension and its geometrical terrain properties for classification of multi-band multi-polarized SAR image," *Prog. Electromagn. Res.*, vol. 104, pp. 221–237, 2010.

[5] S. Kiranyaz, T. Ince, S. Uhlmann, and M. Gabbouj, "Collective Network of Binary Classifier Framework for Polarimetric SAR Image Classification: An Evolutionary Approach," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 42, no. 4, pp. 1169–1186, 2012.

[6] S. Uhlmann and S. Kiranyaz, "Classification of dual- and single polarized SAR images by incorporating visual features," *ISPRS J. Photogramm. Remote Sens.*, vol. 90, pp. 10–22, 2014.

[7] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proc. of NIPS 25*, 2012, pp. 1097–1105.

[8] Y. Zhou, H. Wang, F. Xu, and Y.-Q. Jin, "Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1935–1939, 2016.

[9] F. Gao, T. Huang, J. Wang, J. Sun, A. Hussain, and E. Yang, "Dual-Branch Deep Convolution Neural Network for Polarimetric SAR Image Classification," *Appl. Sci.*, vol. 7, no. 5, p. 447, 2017.

[10] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, 2016.

[11] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, 2016.

[12] S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci, and M. Gabbouj, "1-D Convolutional Neural Networks for Signal Processing Applications," in *Proc. of ICASSP*, 2019.

[13] 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 Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, 2016.

[14] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. CRC press, 2009.

[15] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *CoRR*, vol. abs/1603.0, 2016.