COMPLEX-VALUED VS. REAL-VALUED CONVOLUTIONAL NEURAL NETWORK FOR POLSAR DATA CLASSIFICATION

Reza Mohammadi Asiyabi[1], Mihai Datcu[1,2], Holger Nies[3], Andrei Anghel[1]

[1] Research Center for Spatial Information (CEOSpaceTech), University POLITEHNICA of Bucharest (UPB), Bucharest, Romania
[2] Earth Observation Center (EOC), German Aerospace Center (DLR), Wessling, Germany
[3] Center for Sensor Systems (ZESS), University of Siegen, 57076 Siegen, Germany

IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-2792-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IGARSS46834.2022.9884081

ABSTRACT

Despite the state-of-the-art performance of the deep learning methods for Synthetic Aperture Radar (SAR) data classification, the Real-Valued (RV) networks neglect the phase component of the Complex-Valued (CV) SAR data and lose a lot of useful information. CV deep architectures have been developed in the recent years to exploit the amplitude and phase components of the CV data, in different fields. However, the superiority of CV models over RV models are proved to be different for each application, and more investigation into the advantages and disadvantages of implementing CV models for SAR data classification is necessary. In this study, the performance of the CV Convolutional Neural Network (CV-CNN) for Polarimetric SAR (PolSAR) data classification is compared with its RV equivalent network, in different contexts.

Index Terms— Complex-valued CNN, deep learning, Remote sensing, Classification, PolSAR

1. INTRODUCTION

The very large volume of the available data and the remarkable results of the deep learning methods have encouraged numerous studies to apply deep architectures for various applications. In remote sensing, several Earth Observation (EO) satellites are launched and acquire hundreds of terabytes of data every day, with diverse characteristics, e.g., Multi/hyper-spectral, Synthetic Aperture Radar (SAR), LIDAR, scatterometry, (very) high-resolution, wide-coverage, time series data, and many more varying properties [1]. A great number of studies have focused on utilizing deep learning methods for various remote sensing applications, and obtained state-of-the-art results [2]–[4].

Concurrently, unprecedented advancements of SAR sensors, which provided (very) high resolution SAR images in various imaging modes and polarization channels, and the unique capabilities of these data (e.g., day and night imagery and (almost) independent of atmospheric condition) opened up several opportunities for different applications. Many researchers conducted experiments to apply deep learning models to SAR data processing [5]–[7]. However, SAR data are in complex-domain by nature and processing them with real-valued operators will neglect the phase component of the data and utilize only the amplitude information [8]. Most of the previous attempts to develop deep architectures for SAR data processing have neglected the phase component of the data. In addition to the SAR community, there are several fields (e.g., communication, sonar, biomedicine, seismology, and physics) [9] that are using complex filtering operations and there is a huge interest in these communities for complex-valued operators. Consequently, a number of attempts for complex-valued deep network's development have been carried out in recent years [10]–[12].

A few studies proposed complex-valued deep architectures for SAR data processing, as well. Zhang et al. [8] proposed a Complex-Valued Convolutional Neural Network (CV-CNN) for Polarimetric SAR (PolSAR) data classification. In another study, Shang et al. [13] proposed a Complex-Valued Convolutional Autoencoder (CV-CAE) and reconstructed the PolSAR image patches. Later, the bottleneck features of the CV-CAE are extracted and used in a one layer fully connected classification network. Moreover, Zhang et al. [14] proposed a novel CV-CNN architecture, called "SAR4LCZ-Net", for Local Climate Zones (LCZ) classification in Gaofen-3 quad-pol SAR images. Furthermore, Sun et al. [15] proposed a complex-valued generative adversarial network (GAN) and utilized a semi-supervised classification procedure for PolSAR data classification.

One of the major problems for complex-valued neural network utilization, is the implementation of the network. Several conversions from real-valued mathematical operators into the complex domain have been proposed [8], [13], [16], but most of the well-known libraries and toolboxes for deep network's implementation, such as TensorFlow and PyTorch libraries in python environment, are defined for real-valued networks. Barrachina et al. [16] have developed a library for the Complex-Valued Neural Network (CVNN), based on the TensorFlow library in python environment.

In this paper, the CVNN library, developed by [16], is utilized to define a Complex-Valued Convolutional Neural Network (CV-CNN) for PolSAR data classification, and the

classification is compared to the equivalent Real-Valued CNN (RV-CNN) in different contexts. Comprehensive comparisons between the two models are carried out in terms of the model convergence and trainset size, classification accuracy, and the computational efficiency. The rest of this paper is organized as follows: a brief mathematical background of the complex-valued functions is given in section 2. Section 3 consists of the dataset description, model architecture, and the experimental results. The comparisons between the complex- and real-valued networks is carried out in section 4, and the study in concluded in section 5.

2. MATHEMATICAL BACKGROUND

All of the weights and different functions in the CV-CNN should be in complex-domain, to be able to handle and preserve the structure of the complex-valued data. In order to train the CV-CNN, using the complex-valued backpropagation, the complex gradient has to be defined. According to the Wirtinger calculus [17], if z is a complex variable, $z = x + jy \in \mathbb{C}, (x, y) \in \mathbb{R}^2$, the partial derivatives of a complex function $f(z)$ with respect to z and $\bar{z}$ are as shown in equation 1:

$$\frac{\partial f}{\partial z} \triangleq \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right), \quad \frac{\partial f}{\partial \bar{z}} \triangleq \frac{1}{2}\left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right) \qquad (1).$$

Later, the complex gradient can be defined as equation (2):

$$\nabla_z f = 2\frac{\partial f}{\partial \bar{z}} \qquad (2).$$

Despite the complex-valued weights and functions, the loss function has to be real-valued to reduce the empirical risks during the implementation [16]. For the real-valued loss function $\mathcal{L}: \mathbb{C} \to \mathbb{R}$, the complex derivative of $\mathcal{L}$ with any complex function, $g: \mathbb{C} \to \mathbb{C}$ where $g(z) = r(z) + js(z)$ can be computed according to the chain rule as shown in equation (3):

$$\frac{\partial \mathcal{L} \circ g}{\partial \bar{z}} = \frac{\partial \mathcal{L}}{\partial r}\frac{\partial r}{\partial \bar{z}} + \frac{\partial \mathcal{L}}{\partial s}\frac{\partial s}{\partial \bar{z}} \qquad (3).$$

And the complex backpropagation can be implemented, using the equations (2) and (3).

3. EXPERIMENTAL RESULTS

The covariance matrix C contains the polarization information of the SAR measurements and is authenticated as a suitable representation of the PolSAR data in previous studies [8], [13], [18]. The diagonal elements of the covariance matrix are real-valued, however the other elements are complex-valued and conjugated at the symmetric position of the main diagonal. As a result, the three real-valued and three complex-valued elements of the upper triangle of the covariance matrix (i.e., $C11$, $C22$, $C33$, $C12$, $C13$, and $C23$, 6 channels) are used as the input features of the models. A PolSAR image, acquired by the NASA/JPL AirSAR system over an agricultural area in the Flevoland, Netherlands with the size of 750×1024 is used in this study. The Pauli RGB image and the corresponding Ground Truth (GT) map, with 15 semantic classes, are shown in Figure 1 (a) and (b), respectively.

The main purpose of this study is to compare the utilization of phase and amplitude components of the SAR data through complex-valued deep architectures with the conventional procedure, which only considers the amplitude components of the SAR data in real-valued networks. In order to preserve the equivalency between the networks, the same architecture is used for the CV-CNN and RV-CNN, illustrated in Figure 2. A 7×7 patch around each pixel is used as the representor of the pixel and the semantic label is assigned to the central pixel of the patch. By sliding the window through the whole image with one-pixel steps, a pixel-wise classification is achieved.

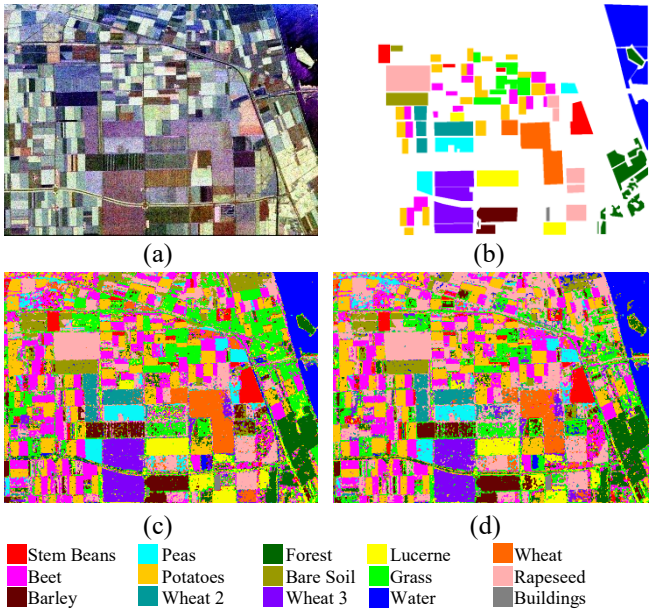

Figure 1. (a) Pauli RGB composite and (b) the Ground Truth map of the Flevoland PolSAR data. White color represents unlabeled pixels. And the classified maps with the (c) CV-CNN and (d) RV-CNN models.

In the architecture of the models, as illustrated in Figure 2, a convolutional layer with the kernel size 3×3, stride 1, and ReLU activation function is applied to the input (i.e., the 7×7×6 patch), which resulted in twelve 7×7 feature maps. An average pooling with the kernel size 2×2 is applied to the feature maps and reduced the size to 3×3 patches. Another similar convolutional layer is applied to achieve eighteen 3×3 feature maps. Because of the small size of the input patch, zero padding is applied in each convolution layer. Later the feature maps are flattened and a vector with 162 features is achieved. Two fully connected layers with ReLU activation functions are used to reduce the size of the vector to 80 and 40, respectively. Finally, a fully connected layer with absolute value activation function is used to obtain the output vector with the size of 15 (i.e., number of the semantic classes in the dataset). Figure 1 (c) and (d) illustrate the classified maps with the CV-CNN and RV-CNN models, respectively.

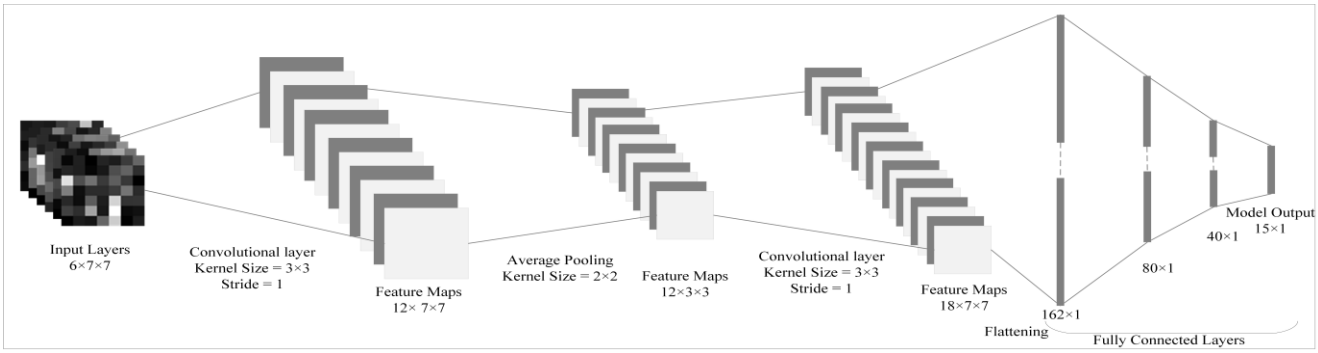

Figure 2. Architecture of the models. In the CV-CNN, all of the components of the model including the convolutional, pooling and fully connected layers are in complex-domain.

4. COMPARISON AND DISCUSSION

In this section, the performance of the CV-CNN and RV-CNN are compared in terms of the model convergence with different trainset sizes, classification accuracy, and the computational efficiency of the models.

4.1. Model convergence and trainset size

In order to compare the convergence rate of the models, the size of the trainset is changed between 5% and 50% of the GT and the models are trained with 100 epochs. Figure 3 illustrates the average testing accuracy of the models for each 10 epochs. For better visualization, only the accuracies above 80% are shown. As it is obvious in Figure 3, the test accuracy of the CV model with only 5% trainset size reaches about 92% and 94% after 50 and 100 epochs, respectively. However, the RV model needed at least 30% trainset size for the similar performance. With 10% trainset size, CV model achieves more than 95% testing accuracy, while the RV model did not reach that performance, even with 50% trainset size. CV-CNN has a remarkable testing accuracy of more than 98% with 50% trainset size.

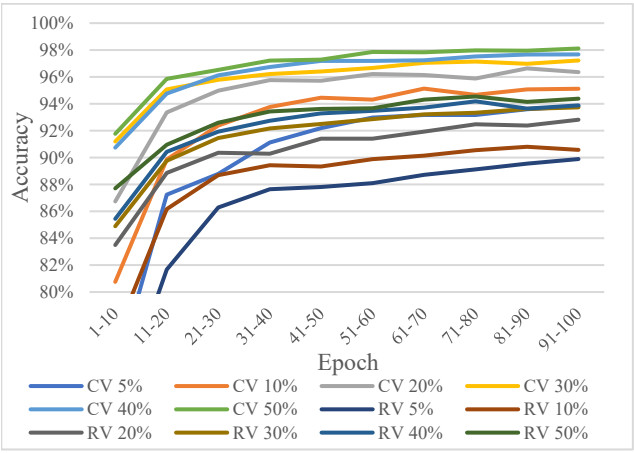

Figure 3. Average test accuracies for every 10 epochs with different training set sizes for both CV-CNN and RV-CNN models.

4.2. Classification Accuracy

With 10% trainset size, the CV-CNN and RV-CNN models achieved more than 95% and 90% Overall Accuracies (OA) for the test set, respectively. The classification accuracies for the 15 semantic classes are shown in Table 1. Only in two classes, Forest and Bare soil, the RV model achieved higher classification accuracies, however, the accuracy of the CV-CNN is remarkably higher in other 13 semantic classes. The highest classification accuracy for one semantic class is achieved by the CV model for water class, with more than 99% accuracy, about 9% higher than RV-CNN. The lowest classification accuracy is observed in the Wheat 2 semantic class for the RV model with less than 80%, while the CV-CNN increased the accuracy of this class with about 15 %.

Table 1. Classification accuracies

Class	CV-CNN	RV-CNN
Stem Beans	**97.12%**	94.21%
Peas	**95.83%**	94.47%
Forest	95.53%	**97.54%**
Lucerne	**96.61%**	91.22%
Wheat	**92.22%**	91.01%
Beet	**94.08%**	89.42%
Potatoes	**95.57%**	87.95%
Bare Soil	95.53%	**95.88%**
Grass	**88.81%**	81.79%
Rapeseed	**95.15%**	89.45%
Barley	**99.29%**	94.28%
Wheat 2	**94.05%**	79.23%
Wheat 3	**96.35%**	90.78%
Water	**99.39%**	90.54%
Buildings	**96.14%**	94.22%
OA	**95.60%**	**90.36%**

4.3. Computational efficiency

Despite the superior performance, CV-CNN has two times more trainable parameters than RV-CNN, 39,034 and 19,517 parameters, respectively, and as a result the training time is much higher for the CV model. However, CV model reaches higher OA with less training epochs. Consequently. One can argue that the CV model can reach a target OA faster, despite

the higher trainable parameters. For instance, for the target OA of more than 90%, CV model needs 17 epochs and 173.32 seconds training time, while the RV model needs 39 epochs and 192.26 seconds training time.

5. CONCLUSION

A comparison between the CV-CNN and RV-CNN for PolSAR data classification is carried out in this study. The obtained results, demonstrated the superiority of the CV model, in terms of the better classified map and higher classification accuracy. However, the CV model has two times more trainable parameters and requires more training time. But it can achieve the target OA with smaller trainset size and less training epochs, which makes it more efficient than the RV equivalent network, in many case studies.

In conclusion, using the phase component of the PolSAR data in CV deep architectures can boost the classification results, and will require smaller trainset and less training epochs and time. In the future studies, the effects of a CV and RV autoencoder as the latent feature extractor before the CNN models should be investigated.

6. AKNOWLEDGMENT

We thank the authors of [16] for the interesting contribution to the field and making the CVNN library publicly available. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860370.

11. REFERENCES

[1] R. M. Asiyabi and M. Datcu, "Earth Observation Semantic Data Mining: Latent Dirichlet Allocation-based Approach," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2022.

[2] X. Yuan, J. Shi, and L. Gu, "A review of deep learning methods for semantic segmentation of remote sensing imagery," *Expert Syst. Appl.*, vol. 169, p. 114417, 2021.

[3] S. T. Seydi, M. Hasanlou, M. Amani, and W. Huang, "Oil Spill Detection Based on Multiscale Multidimensional Residual CNN for Optical Remote Sensing Imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 10941–10952, 2021.

[4] S. Tong, X. Liu, Q. Chen, Z. Zhang, and G. Xie, "Multi-feature based ocean oil spill detection for polarimetric SAR data using random forest and the self-similarity parameter," *Remote Sens.*, vol. 11, no. 4, p. 451, 2019.

[5] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, "A SAR dataset of ship detection for deep learning under complex backgrounds," *Remote Sens.*, vol. 11, no. 7, p. 765, 2019.

[6] J. Zhao, Z. Zhang, W. Yao, M. Datcu, H. Xiong, and W. Yu, "OpenSARUrban: A Sentinel-1 SAR Image Dataset for Urban Interpretation," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 187–203, 2020.

[7] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "Deep SAR-Net: Learning objects from signals," *ISPRS J. Photogramm. Remote Sens.*, vol. 161, pp. 179–193, 2020.

[8] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, 2017.

[9] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge university press, 2010.

[10] A. Hirose, *Complex-valued neural networks*, vol. 400. Springer Science & Business Media, 2012.

[11] B. Zhou and Q. Song, "Boundedness and complete stability of complex-valued neural networks with time delay," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 24, no. 8, pp. 1227–1238, 2013.

[12] A. Hirose and S. Yoshida, "Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 23, no. 4, pp. 541–551, 2012.

[13] R. Shang, G. Wang, M. A Okoth, and L. Jiao, "Complex-valued convolutional autoencoder and spatial pixel-squares refinement for polarimetric SAR image classification," *Remote Sens.*, vol. 11, no. 5, p. 522, 2019.

[14] R. Zhang, Y. Wang, J. Hu, W. Yang, J. Chen, and X. X. Zhu, "SAR4LCZ-Net: A Complex-valued Convolutional Neural Network for Local Climate Zones Classification Using Gaofen-3 Quad-pol SAR Data," *IEEE Trans. Geosci. Remote Sens.*, 2021.

[15] Q. Sun, X. Li, L. Li, X. Liu, F. Liu, and L. Jiao, "Semi-supervised complex-valued GAN for polarimetric SAR image classification," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 3245–3248.

[16] J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, and J.-P. Ovarlez, "Complex-Valued vs. Real-Valued Neural Networks for Classification Perspectives: An Example on Non-Circular Data," *arXiv Prepr. arXiv2009.08340*, 2020.

[17] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Math. Ann.*, vol. 97, no. 1, pp. 357–375, 1927.

[18] R. M. Asiyabi, M. R. Sahebi, and A. Ghorbanian, "Segment-based bag of visual words model for urban land cover mapping using polarimetric SAR data," *Adv. Sp. Res.*, 2021.