DUAL POLARIMETRIC SAR COVARIANCE MATRIX ESTIMATION USING DEEP LEARNING

Adugna G. Mullissa, Diego Marcos, Martin Herold and Johannes Reiche

Laboratory of Geo-Information Science and Remote Sensing, Wageningen University
Wageningen, the Netherlands
adugna.mullissa@wur.nl

IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9324333

ABSTRACT

A polarimetric Synthetic Aperture Radar (PolSAR) image is able to capture target backscattering properties in different polarimetric states, making it a rich source of information for target characterization. However, as with any SAR image, PolSAR images are affected by speckle. Therefore, to extract useful information about targets, the polarimetric covariance matrix has to be first estimated by reducing speckle. In this paper, we use a deep neural network to estimate the dual PolSAR covariance matrix. This application was compared against the state of the art PolSAR despeckling methods. Even if the method is agnostic on the structure of the covariance matrix, the deep learning based PolSAR covariance matrix estimation performed better than the state of the art PolSAR despeckling methods. These results showcase the potential of supervised deep learning for the improvement of PolSAR despeckling pipelines.

***Index Terms*—** Fully convolutional networks, deep learning, PolSAR, Sentinel-1, Speckle

1. INTRODUCTION

The advent of single and multiple polarization SAR images, such as Sentinel-1 dual polarimetric SAR images has been a game changer for all weather day/night global coverage land observation tasks. However, the exploitation is complicated by the presence of speckle. In polarimetric SAR images, due to the presence of speckle, the main interest is not in the scattering matrix itself, but the covariance matrix that specifies completely the probability density function (pdf) of the acquired SAR data vector. Therefore, for the effective characterization of targets in a distributed medium, we need to correctly estimate these covariance matrices [1], as they define the polarimetric properties of the scene under investigation.

Traditionally, the estimation of the polarimetric covariance matrix has been based on knowledge about the underlying data model. In a seminal work, the authors of [2] applied a blind low pass filter to estimate the covariance matrix. This filter was applied with the assumption of spatial homogeniety and ergodicity. This method was effective in preserving polarimetric information in homogenous regions at the expense of resolution. The authors in [1] overcame this issue by minimizing the mean square error of the trace of the covariance matrix in a series of edge aligned neighbourhood window to filter elements of the covariance matrix, whereas the authors in [3] applied target decomposition to determine the scattering mechanisms to use as a basis for similarity between pixels to average together so that the covariance matrix can be effectively estimated in a heterogeneous medium. The authors in [4] alternatively used a combination of additive-multiplicative speckle model for PolSAR data, which adapted to the local correlation of the PolSAR data to effectively estimate the PolSAR covariance matrix, whereas [5] used a non-local means approach to accurately estimate the covariance matrix in a heterogeneous medium without losing resolution. Recently, deep learning based single polarization SAR image despeckling techniques has gained attention. These methods operate by feeding a pair of noisy and clean images in the deep learning network so that the network learns a non-linear function to map the noisy input images to a noise free images [6]. However, the single polarization despeckling methods that are based on deep learning to our knowledge were never extended to the PolSAR domain.

In this paper, we apply a simple feed forward convolutional neural network to dual polarized Sentinel-1 SAR data to estimate the covariance matrix. We use a data driven method with supervision from temporally averaged PolSAR covariance matrix for the estimation of the covariance matrices of Sentinel-1 SAR data.

2. METHODOLOGY

A data vector in a dual polarimetric SAR sensor such as Sentinel-1 is given as:

$$k = \left[\begin{array}{cc} S_{VV} & 2S_{VH} \end{array} \right]^T. \quad (1)$$

where the complex scattering coefficient S_{XY} indexed as $X, Y = (V, H)$ represents the vertical (V) and horizontal (H) polarization states and ^{T} designates a matrix transpose. In a distributed medium k, which follows a zero mean multivariate complex circular Gaussian pdf given as,

$$p(k) = \frac{1}{\pi^3 |C|} \exp(-k^\dagger C^{-1} k), \quad (2)$$

is insufficient to describe the scattering process of the scene, so the second order statistics represented by the covariance matrix C that define the pdf of k is estimated to describe the scattering process. Here, C is given as [7] :

$$C = E\{kk^\dagger\} \quad (3)$$

where $E\{\}$ is the mathematical expectation, $|C|$ is the determinant of C and † is the matrix conjugate transpose. Here, C is an unknown deterministic quantity whereas, the random component of the observed covariance matrix $\hat{C}$ follows a multiplicative noise model, i.e. $\hat{C} = C^{1/2} N C^{1/2}$, where N is speckle. The multiplicative noise model can be converted to a signal independent additive noise by taking the matrix logarithm of the rank 1 covariance matrix $\hat{C}$. The estimated output can be converted back to the multiplicative model by taking the matrix exponent. In this paper, we convert the rank 1 $\hat{C}$ to the additive noise model and apply a deep learning based additive Gaussian denoiser to estimate the covariance matrix. Therefore the deep learning objective can be formulated as:

$$\arg\min_{\theta} \sum_{i=1} L(f_\theta(\hat{C}), \tilde{C}). \quad (4)$$

Here, $\hat{C}$ is the observed rank 1 covariance matrix, $\tilde{C}$ is the reference covariance matrix synthesized by taking the temporal avergae of rank 1 covariance matrices by assuming polarimetric stationaity in time and f_θ is the deep neural network parametrized by the learned parameters θ. We adapt a feed forward fully convolutional network (FCN) architecture [6] to estimate the denoised rank 2 covariance matrix (C) (Figure 1). The Loss function (L) in (4) is a mean square error (MSE) loss function that takes a temporally averaged $\tilde{C}$ as a reference and the network output from the input rank 1 covariance matrix $f_\theta(\hat{C})$, given as:

$$L(\mathbf{f}_\theta(\hat{\mathbf{C}}), \tilde{\mathbf{C}}) = \frac{1}{n} \sum_{i=1}^{n} (f_\theta(\hat{C})_i - \tilde{C}_i)^2. \quad (5)$$

The used network consists of four main building blocks which are convolution, batch normalization, non-linear activation and loss function. The architecture does not use any pooling layers to increase the receptive field of the network.

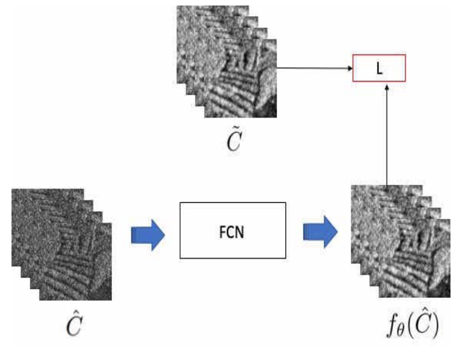


Fig. 1. The proposed architecture.

We have not used pooling layers to avoid the loss of spatial information. Instead we have opted to maintain the sizes of feature maps in the intermediate layers and increasing the depth of the network in order to attain a sufficiently large receptive field.

The network consists of 17 convolutional layers that are interleaved by batch normalization and non-linear activation layers. The stride of the filter is set at 1 to avoid downsampling and a zero padding is applied to maintain the same dimension in the feature maps created from the input image. The parameters of the designed FCN are shown in Table 1. The networks used in this paper are implemented using the MatConvNet package in a Matlab 2018a environment run on a Linux operating system with Intel® Xeon(R) E-2176M CPU and Quadro P2000 GPU.

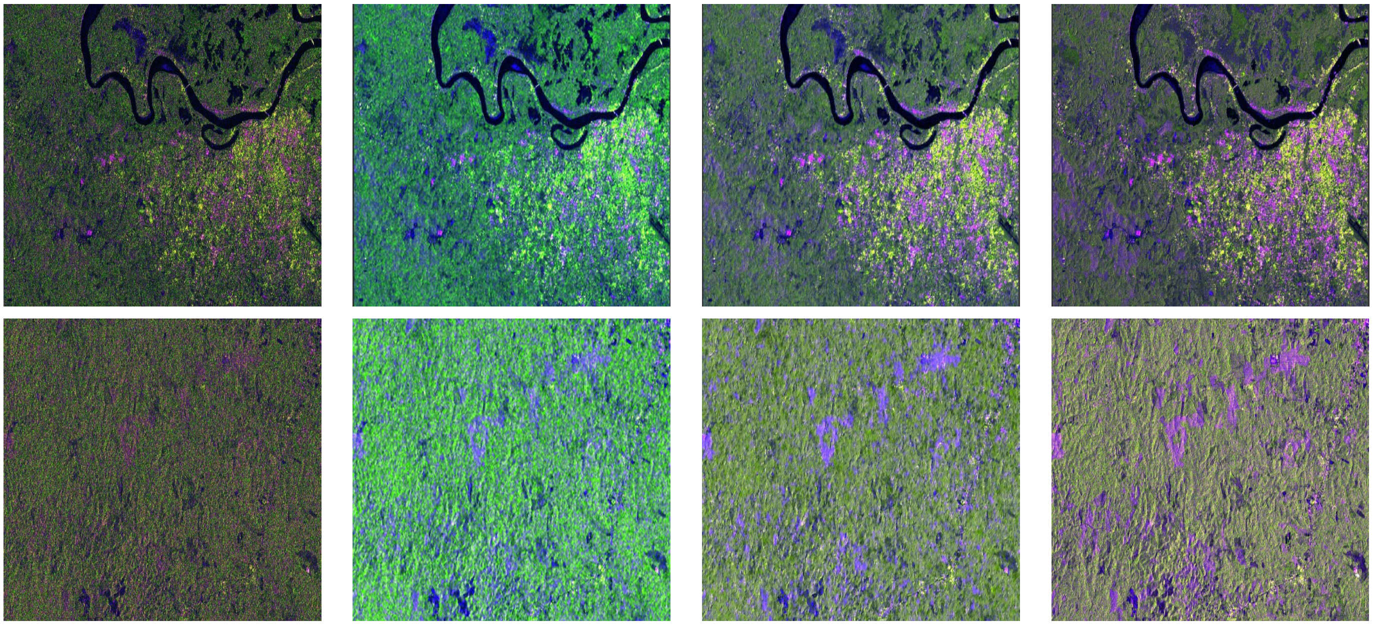
Table 1: Architecture of the FCN blocks for the proposed method.

Layer	Module type	Dimension
Conv1	Conv ReLU	$3 \times 3 \times 4 \times 64$
Conv2 ($\times$15)	Conv BN ReLU	$3 \times 3 \times 64 \times 64$
Prediction	Conv	$3 \times 3 \times 64 \times 4$

3. EXPERIMENTAL SETUP

3.1. Dataset

The proposed FCN is tested using a rank 1 covariance matrix synthesized from a dual polarization Sentinel-1 single look complex (SLC) images acquired near the city of Jambi, Sumatra, Indonesia. The images have a nominal ground resolution of 14.05×14.05 meters. The Sentinel-1 sensor acquires data in C band for both the single and dual polarimetric images. The test areas contained $1,528 \times 1,016$ and $1,366 \times 852$ pix-



(a) (b) (c) (d)

Fig. 2. (a) A false color composite of the two test images synthesized using a) A single look $\hat{C}$ of the test areas b) Lee sigma filter c) The applied deep learning based method d) the reference temporal average covariance matrix $\tilde{C}$. The red color in the image represents C_{11} (the first row and first column element of C), green is represented by C_{22} and blue is represented by C_{11}/C_{22}.

els and cover $21.3\ \mathrm{km} \times 14.05$ and $19.1\ \mathrm{km} \times 11.9\ \mathrm{km}$ in range and azimuth directions, respectively. The top half of the image was used for training and validation, and the bottom half for testing and obtaining the reported metrics. The first Sentinel-1 test image scene covers a mixed urban and natural scene whereas the second scene covers a natural environment (Figure 2a, test area).

3.2. Network training

The input data to the network was prepared by vectorizing the 2×2 covariance matrix into a 4 channel image representing the two real valued diagonal elements and the real and imaginary parts of the complex off diagonal element. Since, $\hat{C}$ is a Hermitian positive semi-definite matrix we only focus on the upper triangular elements of $\hat{C}$. We also used a temporal average of 18 covariance matrices of the same region to prepare the reference $\tilde{C}$. We took the matrix logarithm to convert the data to an additive signal independent Gaussian noise model to be fed to the network.

The networks were trained using Adam optimization method. Batch normalization was used for every convolution layer except the prediction layer. The networks were trained for 30 epochs with a learning rate of 10^{-4} and an additional 20 epochs with a rate of 10^{-5}. To apply this we used a training set of 40800 randomly selected labeled patches of size $40 \times 40 \times 4$. A mini-batches of 128 samples and a weight decay factor of 5×10^{-4} was used. The trained network was finally applied to the test tile.

The result from the proposed method was compared with the state of the art PolSAR coherency matrix estimation methods: improved Sigma filter [8] and Refined Lee filter [9].

3.3. Results

To evaluate the performance of the proposed FCN we qualitatively and quantitatively compared the results from the improved sigma filter and the refined Lee filter. The quantitative comparisons were based on the peak signal to noise ratio (PSNR) and structural similarity index (SSIM) of the diagonal elements of the covariance matrix. PSNR estimates the quality of the reconstructed noise free diagonal element of the covariance matrix resemblance to the reference data, in this case the temporally averaged diagonal elements of the covariance matrix as:

$$PSNR = 20 \log_{10}(\frac{\hat{C}_{max}}{\sqrt{MSE}}). \quad (6)$$

Here, $\hat{C}_{max}$ is the maximum power of the diagonal elements of $\hat{C}$ given as 255, if we're using 8bit data. The MSE is computed between the diagonal elements of the label (C)

Table 2: PSNR/SSIM values computed for the diagonal elements of the covariance matrix.

Area	Parameter	Input	Refined Lee	Improved Sigma	Proposed
Test area1	C11	15.27/0.36	18.00/0.38	18.99/0.46	**21.13/0.61**
	C22	25.18/0.57	27.08/0.78	27.81/0.81	**30.97/0.87**
Test area2	C11	13.88/0.0.005	23.28/0.36	23.79/0.40	**24.16/0.48**
	C22	26.47/0.05	35.25/0.85	35.67/0.87	**37.33/0.89**

and the reconstructed covariance matrix ($\hat{C}$). SSIM estimates the structural similarity between the label (C) and the reconstructed ($\hat{C}$) as:

$$SSIM(\hat{C}, C) = \frac{(2\mu_{\hat{C}}\mu_C + c_1)(2\sigma_{\hat{C}C} + c_2)}{(\mu_{\hat{C}}^2 + \mu_C^2 + c_1)(\sigma_{\hat{C}}^2 + \sigma_C^2 + c_2)}, \quad (7)$$

Where μ_C is the mean of C and σ_C is its standard deviation. In addition to this, we derived the mean absolute error of α (physical scattering mechanisms) and H (Entropy) values derived between the baseline methods and the ones derived from the temporal average reference covariance matrix. These quality metrics uses the full covariance matrix as described in [10]. In the qualitative comparison, the deep learning based approach achieved superior quality estimating the covariance matrix as exemplified in the displayed intensity image composites. This was also confirmed when comparing the PSNR and SSIM values based on the diagonal elements of the covariance matrix and when comparing the absolute error from α and H (Figure 2 and Table 2, 3).

Table 3: Mean absolute error of the α and H values for the different baseline methods.

Area	Parameter	Refined Lee	Improved Sigma	Proposed
Test area1	α	5.2	4.61	**3.86**
	H	0.11	0.10	**0.08**
Test area2	α	4.70	4.11	**3.78**
	H	0.096	0.087	**0.072**

4. CONCLUSIONS

In this paper, we explore the use of a conventional deep learning method, based on fully convolutional neural networks, for the estimation of the covariance matrices of dual polarized Sentinal-1 data. We used a model trained with temporally averaged covariance matrix as a reference and obtain better results than traditional, hand-crafted approaches.

In the literature we have seen how the injection of knowledge about the signal and the noise properties has led to substantial improvements in the performance of traditional methods. We expect that injecting this knowledge into deep learning-based models will, once more, result in improvements over the baseline presented in this paper.

5. REFERENCES

[1] Jong-Sen Lee, Mitchell R Grunes, and Gianfranco De Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Transactions on Geoscience and remote sensing*, vol. 37, no. 5, pp. 2363–2373, 1999.

[2] Nathaniel R Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 152–177, 1963.

[3] Jong-Sen Lee, Mitchell R Grunes, Dale L Schuler, Eric Pottier, and Laurent Ferro-Famil, "Scattering-model-based speckle filtering of polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 176–187, 2005.

[4] Carlos Lopez-Martinez and Xavier Fabregas, "Model-based polarimetric SAR speckle filter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3894–3907, 2008.

[5] Charles-Alban Deledalle, Loïc Denis, Florence Tupin, Andreas Reigber, and Marc Jäger, "Nl-sar: A unified nonlocal framework for resolution-preserving (pol)(in) SAR denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2021–2038, 2014.

[6] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang, "Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[7] Adugna G Mullissa, Daniele Perissin, Valentyn A Tolpekin, and Alfred Stein, "Polarimetry-based distributed scatterer processing method for PSI applications," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3371–3382, 2018.

[8] Jong-Sen Lee, Jen-Hung Wen, Thomas L Ainsworth, Kun-Shan Chen, and Abel J Chen, "Improved sigma filter for speckle filtering of SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 202–213, 2008.

[9] Jong-Sen Lee and Eric Pottier, *Polarimetric radar imaging: from basics to applications*, CRC press, 2009.

[10] Jong-Sen Lee, Mitchell R Grunes, Eric Pottier, and Laurent Ferro-Famil, "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 4, pp. 722–731, 2004.