HAL
open science

SAR image despeckling through convolutional neural networks

Giovanni Chierchia, Davide Cozzolino, Giovanni Poggi, Luisa Verdoliva

To cite this version:

Giovanni Chierchia, Davide Cozzolino, Giovanni Poggi, Luisa Verdoliva. SAR image despeckling through convolutional neural networks. IEEE International Geoscience and Remote Sensing Symposium, Jul 2017, Fort Worth, Texas, United States. hal-01710036

HAL Id: hal-01710036

https://hal.science/hal-01710036

Submitted on 15 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SAR IMAGE DESPECKLING THROUGH CONVOLUTIONAL NEURAL NETWORKS

G. Chierchia

Université Paris Est
LIGM UMR 8049, CNRS, ESIEE Paris
F-93162 Noisy-le-Grand (France)

D. Cozzolino, G. Poggi, L. Verdoliva

University Federico II of Naples
DIETI
Via Claudio 21, 80125 Naples (Italy)

arXiv:1704.00275v2 [cs.CV] 3 May 2017

ABSTRACT

In this paper we investigate the use of discriminative model learning through Convolutional Neural Networks (CNNs) for SAR image despeckling. The network uses a residual learning strategy, hence it does not recover the filtered image, but the speckle component, which is then subtracted from the noisy one. Training is carried out by considering a large multitemporal SAR image and its multilook version, in order to approximate a *clean* image. Experimental results, both on synthetic and real SAR data, show the method to achieve better performance with respect to state-of-the-art techniques.

***Index Terms*—** SAR, speckle, multiplicative noise, convolutional neural networks.

1. INTRODUCTION

SAR images are affected by a strong multiplicative noise, the speckle, which may severely impair the performance of automatic operations, like classification and segmentation, aimed at extracting valuable information for the end user. As more and more images are acquired every day, automatic analysis is mandatory, making of image despeckling a central issue. A number of approaches have been proposed in the last few years to suppress speckle while preserving relevant image features [1]. Wavelet shrinkage [2], sparse representations [3], and especially nonlocal filtering [4, 5, 6, 7], represent arguably the current state-of-the-art.

Most of these approaches rely on detailed statistical models of signal and speckle, either in the original or in a transform domain [1]. However, depending on the sensor, the acquisition modality, the possible use of multilooking, and a number of other factors, including of course the land cover, statistics may vary significantly from case to case (see Fig. 1). A well-known example concerns high-resolution data such as those acquired by TerraSAR-X, COSMO-SkyMed, and RADARSAT-2 systems.

In this work, we propose to avoid altogether the modeling problem by resorting to the machine learning approach, implemented through a convolutional neural network (CNN). Given a suitable set of images, the network is trained to learn an implicit model of the data which allows the effective despeckling of all new data of the same type. In the last few years, several authors have proposed CNN-based methods for AWGN image denoising [8, 9]. Here, we follow the paradigm proposed in [10], which resorts to residual learning to guarantee a faster convergence in the presence of limited training data. Adaptation to SAR is obtained by handling multiplicative noise and by using an ad hoc procedure to build a suitable training set. To the best of our knowledge, this is the first paper investigating CNNs for SAR image despeckling.

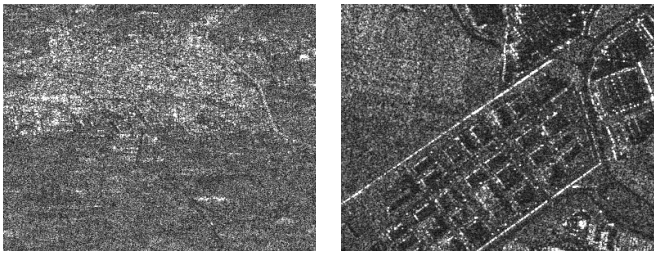

Fig. 1. Examples of low-resolution (left) and high-resolution (right) SAR images. Statistical differences appear even by visual inspection.

In the following sections, we describe the proposed method, we present experimental results on both synthetic and real data, and finally we draw conclusions.

2. PROPOSED METHOD

The architecture of the proposed CNN, inspired by [10], is shown in Fig. 2. The network comprises 17 full convolutional layers, with no pooling. Each layer extracts 64 feature maps, using filters of size $3\times3\times64$, except the first and last layers which have single-band input and output, respectively. Rather than the clean image, the network recovers the speckle component, which is then subtracted from the noisy image.

A fundamental difference of our approach with respect to [10] lies in the criterion to be optimized during training. Indeed, outside of the AWGN realm, the Euclidean distance is not optimal anymore. To deal with multiplicative noise, we use the homomorphic approach with coupled log and exp

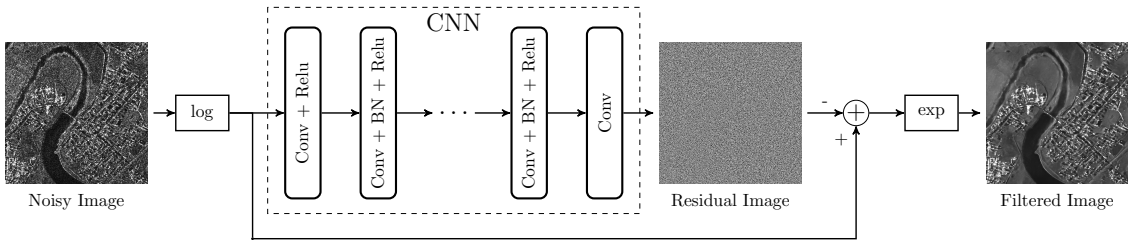

Fig. 2. Proposed CNN architecture for SAR image despeckling.

transforms, in synergy with the similarity measure for speckle noise distribution [11], leading to the loss function[1]

$$\ell(\mathbf{\Theta}) = \sum_{i=1}^{N} \mathbf{1}^{\top} \log \left(\cosh \left(\mathcal{R}_{\mathbf{\Theta}}(\log \mathbf{y}_i) + c - \log \frac{\mathbf{y}_i}{\mathbf{x}_i} \right) \right), \tag{1}$$

where $\mathcal{R}_{\mathbf{\Theta}}$ is the network output, Θ denotes the trainable parameters, $(\mathbf{x}_i, \mathbf{y}_i)$ is a pair of clean-noisy training patches in amplitude format, and c is the nonzero mean of log-speckle.

This strategy, called residual learning [12], is key to speed up the training process and helps improving the performance. In fact, it has been observed experimentally [12] that training a CNN may be quite slow when the desired output is very similar to the input. This is the case of many restoration tasks, such as denoising or super-resolution. By setting the dual goal of reproducing the noise (hence, removing the clean signal) training becomes much more effective. This is extremely important for SAR applications, given the inherent scarcity of training data. In fact, while for the "conventional" case of fully developed speckle one can generate a large dataset by simulation, this is not possible in other cases, such as high-resolution data, due to the lack of satisfactory models.

Therefore, we propose an ad hoc procedure for dataset generation, described graphically in Fig.3. We assume that a relatively large multitemporal SAR image is available. The clean image is obtained by averaging the temporal components (multilooking) and keeping only the regions with no significant temporal changes. Of course, the more temporal instances are available, the more reliable the clean reference is. Eventually, a number of noisy patches are extracted with their clean version and used to train the network. The use of residual learning, together with batch normalization and a suitable optimization algorithm [13], allows us to obtain satisfactory training. The trained network can now deal with any SAR images acquired in the same modality.

3. EXPERIMENTAL RESULTS

In SAR image despeckling, the performance assessment is quite challenging, due to the lack of original noiseless signal. Therefore, we split the numerical validation in two parts.

[1] log and cosh are meant element-wise, whereas $\mathbf{1}^{\top} z = \sum_{k=1}^{K} z_k$.

Table 1. CPU time for despeckling a 512×512 clip.

PPB	NL-SAR	SAR-BM3D	SAR-CNN
49.7 s	10.1 s	87.4 s	4.6 s

First, we present experiments carried out on synthetic SAR images corrupted by simulated speckle, and make comparison using the usual performance indexes, such as the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM). Afterwards, we experiment with real-world SAR images, focusing on the challenging high-resolution case.

We compare results with three despeckling algorithms, PPB [4], SAR-BM3D [5], and NL-SAR [7], chosen for their competitive performance and the availability of software code. For all these algorithms parameters are set as suggested in the reference papers. Turning to the proposed method, called SAR-CNN from now on, a training set of 2000×128 patches (40×40 pixels) is used, with the ADAM gradient-based optimization method [13], minibatches of 128 patches, and the batch normalization strategy of [14]. Training proceeds for 30 epochs with learning rate 0.001 and, only for synthetic data, further 20 epochs with learning rate 0.0001. All experiments were carried out in Matlab R2016b with the *MatConvNet* toolbox [15], with an Intel Xeon CPU at 2.10GHz and an Nvidia P100 GPU. Training took about 8 hours. Interestingly, once training is over, SAR-CNN exhibits the lowest run-time complexity as shown in Tab.1.

3.1. Results on simulated SAR images

We generate a number of SAR-like images by injecting single-look speckle in amplitude format on optical images. Training patches are extracted from 400 different images. Tab. 2 reports PSNR results for some out-of-training images often used for testing. In all but one case SAR-CNN provides the best performance, with an average gain over the reference techniques of about 1 dB, 2 dB, and 2.5 dB, respectively. Similar considerations apply for the SSIM index (Tab. 3). Such good results are confirmed by visual inspection, see Fig. 4, with an impressive improvement in detail preservation.

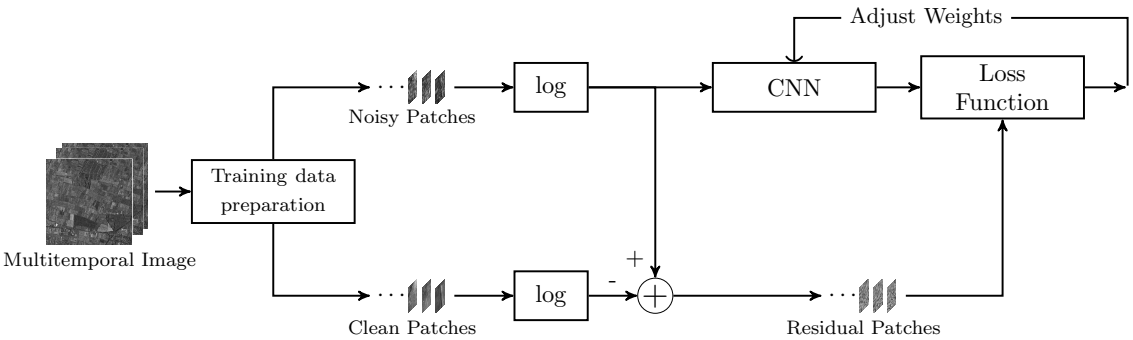

Fig. 3. Training procedure.

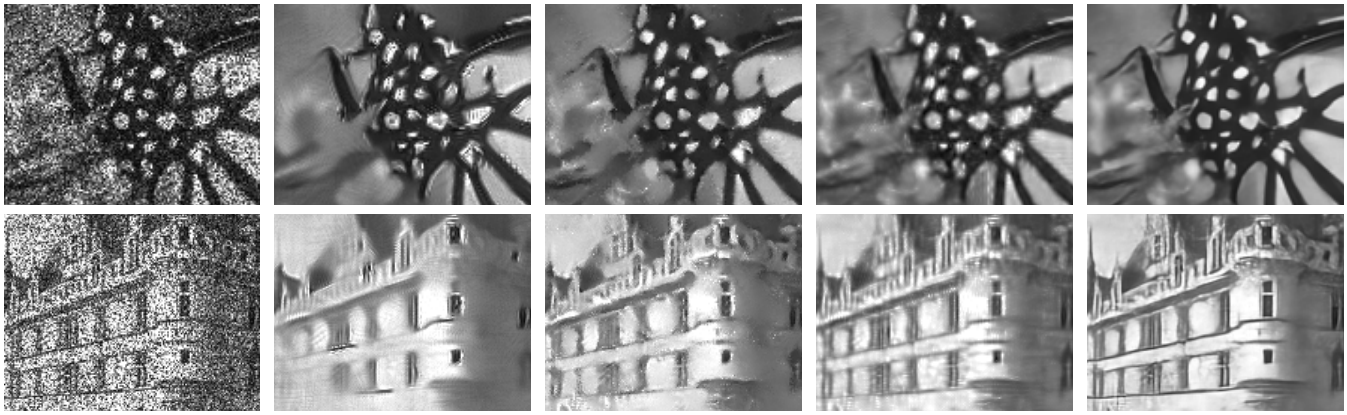

Fig. 4. Sample results on simulated SAR images. Left to right: original single-look, PPB, NL-SAR, SAR-BM3D, SAR-CNN.

Table 2. PSNR over synthetic SAR images.

Image	PPB	NL-SAR	SAR-BM3D	SAR-CNN
Cameraman	23.02	24.37	24.76	26.15
House	25.51	25.75	27.55	28.60
Peppers	23.85	23.62	24.92	26.02
Starfish	21.13	21.84	22.71	23.37
Butterfly	22.76	23.82	24.48	26.05
Airplane	21.22	21.83	22.71	23.93
Parrot	21.88	24.13	24.17	25.92
Lena	26.64	26.80	27.85	28.70
Barbara	24.08	23.13	25.37	24.70
Boat	24.22	24.55	25.43	26.05
Average	23.43	23.98	24.99	**25.95**

Table 3. SSIM over synthetic SAR images.

Image	PPB	NL-SAR	SAR-BM3D	SAR-CNN
Cameraman	0.661	0.716	0.750	0.792
House	0.651	0.686	0.751	0.791
Peppers	0.680	0.716	0.747	0.793
Starfish	0.563	0.609	0.664	0.702
Butterfly	0.714	0.752	0.792	0.841
Airplane	0.533	0.620	0.672	0.724
Parrot	0.685	0.732	0.771	0.805
Lena	0.680	0.714	0.763	0.800
Barbara	0.652	0.631	0.729	0.718
Boat	0.573	0.602	0.650	0.675
Average	0.639	0.677	0.729	**0.764**

3.2. Results on high-resolution SAR images

In the second experiment we consider a single-look COSMO-SkyMed image of size 16000×16000 with 25 co-registered temporal components. Training patches are extracted from one half of the image, whereas numerical evaluation is carried out on several 512×512 clips from the other half. Fig. 5 shows results for some of these clips. Lacking a clean reference, visual inspection is the main tool for quality evaluation. In our assessment, SAR-CNN looks extremely promising, showing the same speckle suppression ability of NL-SAR, but with a better detail preservation, comparable to that of SAR-BM3D. These observations are supported also by results of Tab. 4, reporting two no-reference metrics, the equivalent number of looks (ENL), evaluated on homogeneous blocks, and the $\alpha\beta$ index [16]. The best scores are achieved by SAR-CNN and NL-SAR, indicating a better speckle suppression and detail preservation. The improvement w.r.t. competitors is not as striking as in the previous case. However, note that the network did not see clean patches during training but only well despeckled ones, with a sure impact on performance.

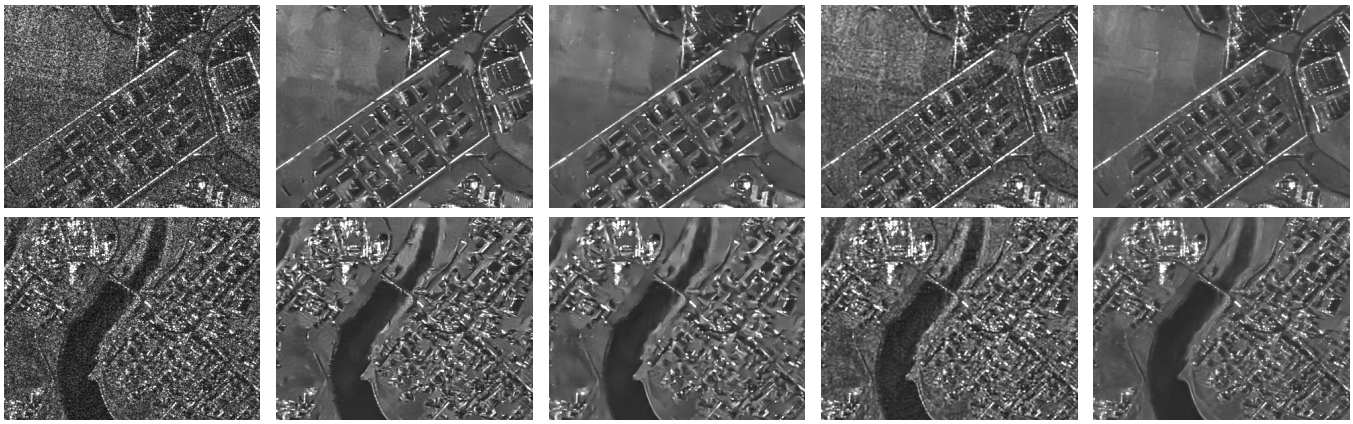

Fig. 5. Results on a COSMO-SkyMed image. Left to right: original single-look clip, PPB, NL-SAR, SAR-BM3D, SAR-CNN.

Table 4. ENL and $\alpha\beta$ over the clips in Fig. 5.

# clip	Index	PPB	NL-SAR	SAR-BM3D	SAR-CNN
1	ENL	47.61	154.10	4.87	129.10
2	ENL	25.28	52.12	4.71	56.32
1	$\alpha\beta$	0.162	0.076	0.530	0.187
2	$\alpha\beta$	0.171	0.065	0.511	0.182

4. CONCLUSION

In this paper we investigated the use of Convolutional Neural Networks for SAR image despeckling. A residual learning strategy is applied together with a suitable training phase that is carried out by using multitemporal SAR data of the same scene (both the original data and their mutilook version). Results on synthetic and real SAR data show promising results both considering objective and visual assessment.

5. REFERENCES

[1] F. Argenti, A. Lapini, L. Alparone, and T. Bianchi, "A tutorial on speckle reduction in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, pp. 6–35, 2013.

[2] T. Bianchi, F. Argenti, and L. Alparone, "Segmentation-Based MAP Despeckling of SAR Images in the Undecimated Wavelet Domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2728–2742, 2008.

[3] S. Foucher, "SAR image filtering via learned dictionaries and sparse representations," in *IEEE IGARSS*, 2008, pp. 229–232.

[4] C.-A. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Trans. on Image Process.*, vol. 18, no. 12, pp. 2661–2672, 2009.

[5] S. Parrilli, M. Poderico, C.V. Angelino, and L. Verdoliva, "A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 606–616, Feb. 2012.

[6] D. Cozzolino, S. Parrilli, G. Scarpa, G. Poggi, and L. Verdoliva, "Fast adaptive nonlocal SAR despeckling," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 2, pp. 524–528, 2014.

[7] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In)SAR denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2021–2038, 2015.

[8] V. Jain and S. Seung, "Natural image denoising with convolutional networks," in *Advances in Neural Information Processing Systems*, 2009, pp. 769–776.

[9] H.C. Burger, C.J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?," in *IEEE CVPR*, 2012, pp. 2392–2399.

[10] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *arXiv:1608.03981v1*, 2016.

[11] C.A. Deledalle, L. Denis, and F. Tupin, "How to compare noisy patches? patch similarity beyond gaussian noise," *International Journal of Computer Vision*, vol. 99, pp. 86–102, 2012.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *IEEE CVPR*, 2016, pp. 770–778.

[13] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conf. on Learning Representations (ICLR)*, 2015.

[14] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167v3*, 2015.

[15] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for Matlab," in *ACM Conf. on Multimedia Conference*, 2015, pp. 689–692.

[16] L. Gomez, M.E. Buemi, J.C. Jacobo-Berlles, and M.E. Mejail, "A new image quality index for objectively evaluating despeckling filtering in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1297–1307, 2016.