

From Patches to Deep Learning: Combining Self-Similarity and Neural Networks for Sar Image Despeckling

Loïc Denis, Charles-Alban Deledalle, Florence Tupin

▶ **To cite this version:**

Loïc Denis, Charles-Alban Deledalle, Florence Tupin. From Patches to Deep Learning: Combining Self-Similarity and Neural Networks for Sar Image Despeckling. 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019), Jul 2019, Yokohama, Japan. pp.5113-5116, 10.1109/IGARSS.2019.8898473 . ujm-03114605

HAL Id: ujm-03114605
https://ujm.hal.science/ujm-03114605

Submitted on 19 Jan 2021

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.

FROM PATCHES TO DEEP LEARNING: COMBINING SELF-SIMILARITY AND NEURAL NETWORKS FOR SAR IMAGE DESPECKLING

Loïc Denis[1], Charles-Alban Deledalle[2], Florence Tupin[3]

[1]Univ Lyon, UJM-Saint-Etienne, Institut d Optique Graduate School, Laboratoire Hubert Curien CNRS UMR 5516, Saint-Etienne, France

[2] IMB, CNRS, Univ. Bordeaux, Bordeaux INP, F-33405 Talence, France

[3]LTCI, Télécom ParisTech, Université Paris Saclay, Paris, France

ABSTRACT

Speckle reduction has benefited from the recent progress in image processing, in particular patch-based non-local filtering and deep learning techniques. These two families of methods offer complementary characteristics but have not yet been combined. We explore strategies to make the most of each approach.

***Index Terms*—** patches, non-local filtering, deep learning, speckle, SAR, PolSAR.

1. INTRODUCTION

The intensity in a SAR image fluctuates strongly because of constructive / destructive interferences, a phenomenon called *speckle*. These interferences occur between echoes received from multiple elementary scatterers located within each resolution cell.

The reduction of the speckle in SAR images has fueled a large amount of work in the image processing community since the 80s. This problem has been addressed through several methodological approaches: local filtering, neighbor filtering [1], Markov random fields [2], total variation minimization [3], wavelet transforms [4], and, more recently, patch-based approaches [5] and deep neural networks [6].

The introduction of patch-based filtering (a.k.a. non-local processing) [7] in image processing led to a significant improvement in the performance of image restoration methods. The key idea of these methods is to exploit the self-similarity of images: at the scale of small windows (the *patches*) of size about 10 by 10 pixels, many similar parts can generally be found within an extended neighborhood surrounding any given location. After identification of the matching patches, these patches can be combined in order to provide an estimate of a speckle-free value. Because only similar patches are combined, blurring of textures or sharp structures is avoided. Several extensions of patch-based processing have been proposed for SAR imagery, see the review [8] and the references therein.

The strength of patch-based methods lies in their capability to strongly smooth homogeneous areas while introducing very limited blur around sharp structures (preservation of the spatial resolution). The major weakness appears when not enough similar patches can be identified within the search window: a phenomenon often described as the *rare patch effect*. No substantial filtering can then be performed which leads to residual noise located around those isolated structures. This problem can be reduced by locally adjusting the sizes of patches and search windows, as in [5], but remains intrinsic to the method.

Following the numerous successes of deep learning in computer vision, convolutional neural networks (CNNs) have been adapted to the problem of speckle reduction. By training over a larger number of image patches, the networks learn an implicit model of the structures found in images and how to discriminate between those structures and the random fluctuations due to noise. The ability of deep architectures to capture a wide variety of geometrical structures and textures is at the core of the very good restoration capability of CNNs. A weakness of deep learning methods is the apparition of some phantom structures, in some cases. These artifacts are difficult to identify visually and to discard because they are very similar to structures found in actual images.

Given the complementarity of patch-based methods and deep learning techniques, it is natural to try to combine them. In section 2, we review several approaches that perform this combination in the recent literature. We then describe in section 3 a method that is well-adapted to SAR imaging.

2. A REVIEW OF METHODS FOR COMBINING NON-LOCAL APPROACHES AND DEEP LEARNING

The integration of the non-local processing ideas (i.e., use of image parts extracted within an extended search window, on the basis of patch similarity) within deep network architectures has been considered at various levels in the recent literature.

Davy *et al.* [9] propose to augment the region of the

This work has been funded by ANR (the French National Research Agency) and DGA (Direction Générale de l'Armement) under ALYS project ANR-15-ASTR-0002.

image that is fed to the neural network with so-called non-local features which are values collected from the most similar patches. A patch-based processing step is thus applied in order to add information from farther regions of the image. After that step, a conventional CNN architecture is used. Plötz and Roth [10] consider a similar but more integrated approach. They design a non-local block that learns features that are then used to compute distances, a soft version of the k-nearest neighbors is followed in order to produce the non-local features that are appended to the input image. In contrast to [9], both the similarity between patches and the selection / combination of patches based on the similarities are learned. This non-local block can be embedded with a CNN architectures (and used at several depth of the network). Wang *et al.* [11] also design a block that incorporates a non-local processing. They learn both a similarity between patches and a combination of the patches. However, no increase of the dimension is performed: only a non-local filtering is applied. Like the approach of [10], this non-local block can be included at various levels of a deep network architecture. Cruz *et al.* [12] follow a very different approach with a clear separation between the CNN step and the non-local processing. Moreover, the non-local step is not used in order to enrich the input of the network, but rather as a post-processing step to reduce the artifacts produced by the network.

Given the difficulty to train neural networks specifically to the processing of SAR images (in particular, for multi-channel images like in SAR polarimetry and/or SAR interferometry), we base our approach on the non-integrated framework of Cruz *et al.* [12]. That way, we can use pre-trained networks and design a method that can be applied both to single channel images and to multi-channel images.

3. PROPOSED APPROACH

As a reference CNN-based speckle reduction method, we consider MuLoG framework [13] with the Gaussian denoiser DnCNN [14]. It can be applied both to single-channel intensity images and to multi-channel images such as polarimetric [15] or interferometric [16] images.

3.1. Single-channel SAR images

In order to improve the restoration result and reduce artifacts produced by the CNN-based method, we propose to add a non-local filtering step similar to that proposed in [12]. We identify the 32 most similar patches (of size 10×10) within the search window, on the noisy image. We then extract the *filtered* patches from the restored image obtained after the CNN step at the locations previously identified. These collections of patches form 3D blocks that are then filtered by applying a 1D Haar transform in the third dimension, followed by a shrinkage according to the Wiener filtering described in [12]. Restored patches are then aggregated to recover the image. We perform the operations of patch matching and collaborative filtering in the log domain.

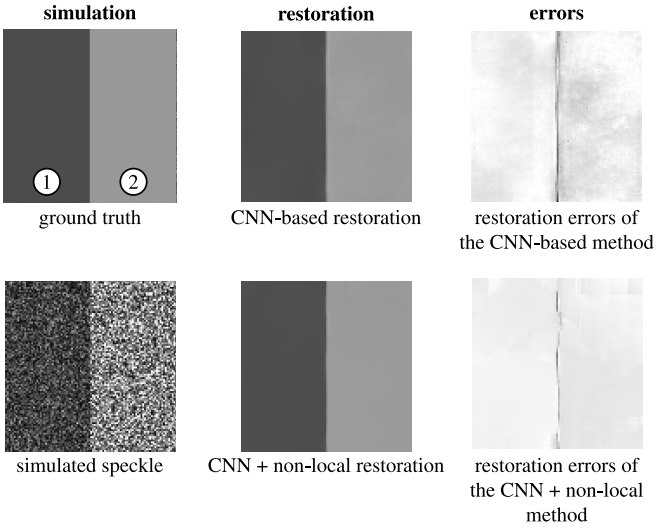


Fig. 1: Illustration of the two-steps restoration of an intensity image: the first step (top row) is performed using a CNN, the second step is a non-local filtering based on patch similarity (bottom row).

region	mean	variance	error
ground truth (1)	0.2	0	0
ground truth (2)	0.8	0	
CNN (1)	0.211	6.8×10^{-5}	0.033
CNN (2)	0.828	1.2×10^{-3}	
CNN + non-local (1)	0.211	4.8×10^{-5}	0.025
CNN + non-local (2)	0.829	2.4×10^{-4}	

Table 1: Quantitative evaluation of the restorations depicted in Fig. 1. The two regions (1) and (2) indicated on the ground truth are analyzed in terms of mean value, and variance. The root mean square error is given over the whole image.

Figure 1 illustrates on a numerical simulation how the restoration produced by the CNN can be further improved by the non-local filtering step: errors are reduced in the homogeneous regions without blurring the sharp transition between the regions. In the simulation, the speckle is single-look and uncorrelated. A quantitative analysis of this numerical simulation is provided in table 1. Compared to a single CNN step, adding the non-local processing reduces the errors (mostly by reducing the variance in the homogeneous regions).

3.2. Multi-channel SAR images

The extension of this two-steps strategy to muti-channel images requires to handle the statistics of polarimetric and/or interferometric channels. The first step is readily performed by MuLoG and the same CNN denoiser as for the single-channel case. MuLoG uses a matrix logarithm transform and a chan-

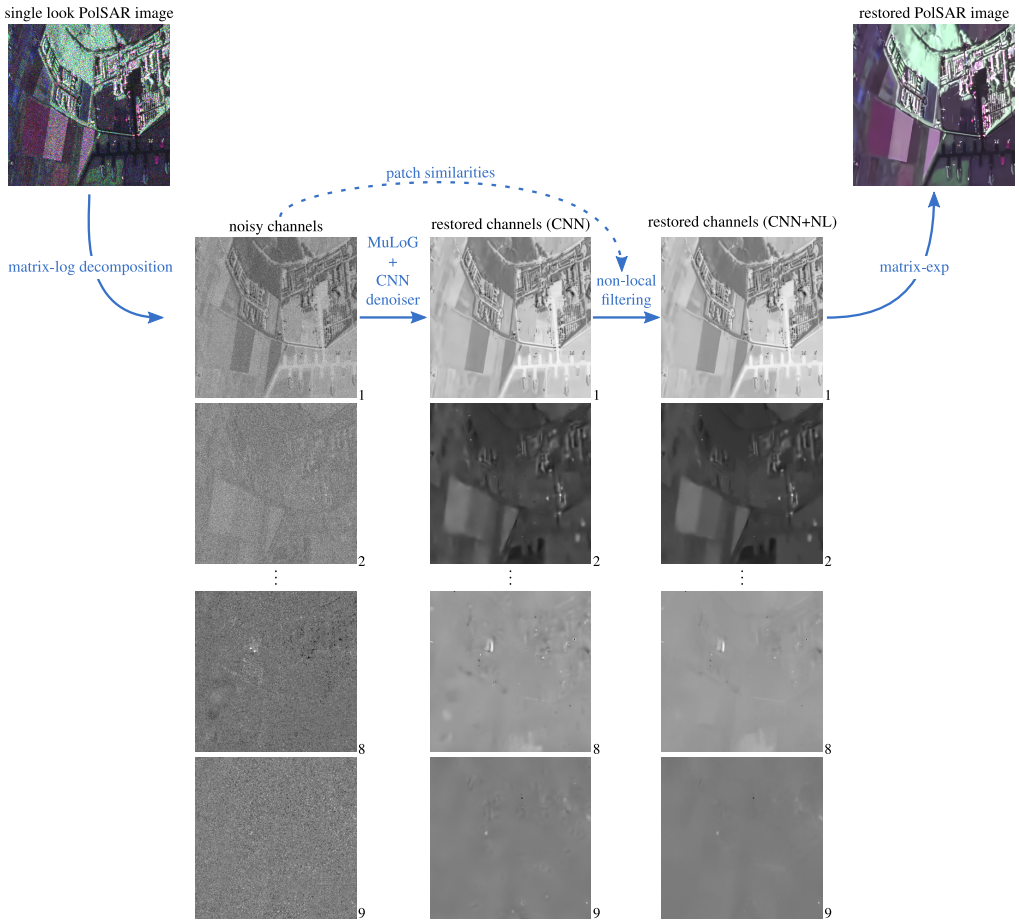


Fig. 2: Principle of the proposed two-step restoration method that combines a CNN-based denoiser and a non-local filtering step.

nel decorrelation step, relying on principal components analysis, in order to iteratively denoise each channel. We propose to use the same channels in order to perform the non-local filtering step. As in the single-channel case, the identification of similar patches is performed on the noisy image (similarities are evaluated based on the L2 distance between patches extracted from the first channel of the principal components analysis, i.e., the channel with the highest signal-to-noise ratio, in the log-domain). The denoised patches obtained by the CNN step are filtered in the second step in order to smooth some artifacts. Figure 2 illustrates the process.

To better reduce the artifacts, we filter more strongly the channels with the smallest weight (i.e., the last vectors of the principal components analysis). Figure 3 displays restoration results on an airborne full-polar SAR image obtained with the E-SAR sensor of the DLR (image over Oberpfaffenhofen provided with PolSARpro). The restoration obtained after the CNN-based step is shown on the left part of the figure, the restoration after the second step (non-local filtering) is displayed on the right part. Differences between the two results are indicated on a difference image obtained by computing, at each pixel, the dissimilarity between the estimated covariance matrices (see [5]):

$$\log\det \widehat{\boldsymbol{C}}^{(\mathrm{CNN})} + \log\det \widehat{\boldsymbol{C}}^{(\mathrm{CNN+NL})} - 2\log\det\left[\frac{1}{2}\widehat{\boldsymbol{C}}^{(\mathrm{CNN})} + \frac{1}{2}\widehat{\boldsymbol{C}}^{(\mathrm{CNN+NL})}\right] \quad (1)$$

where $\widehat{\boldsymbol{C}}^{(\mathrm{CNN})}$ is the estimated polarimetric covariance matrix after the first step and $\widehat{\boldsymbol{C}}^{(\mathrm{CNN+NL})}$ is the estimated polarimetric covariance matrix after the second step.

4. CONCLUSION

Combining ideas from patch-based non-local processing and deep learning techniques is a promising direction. We described a simple approach that improves the multi-channel speckle reduction framework MuLoG. Our method does not require to train a network specifically for SAR or PolSAR images. Learning dedicated networks is more challenging but could potentially lead to even better restoration results.

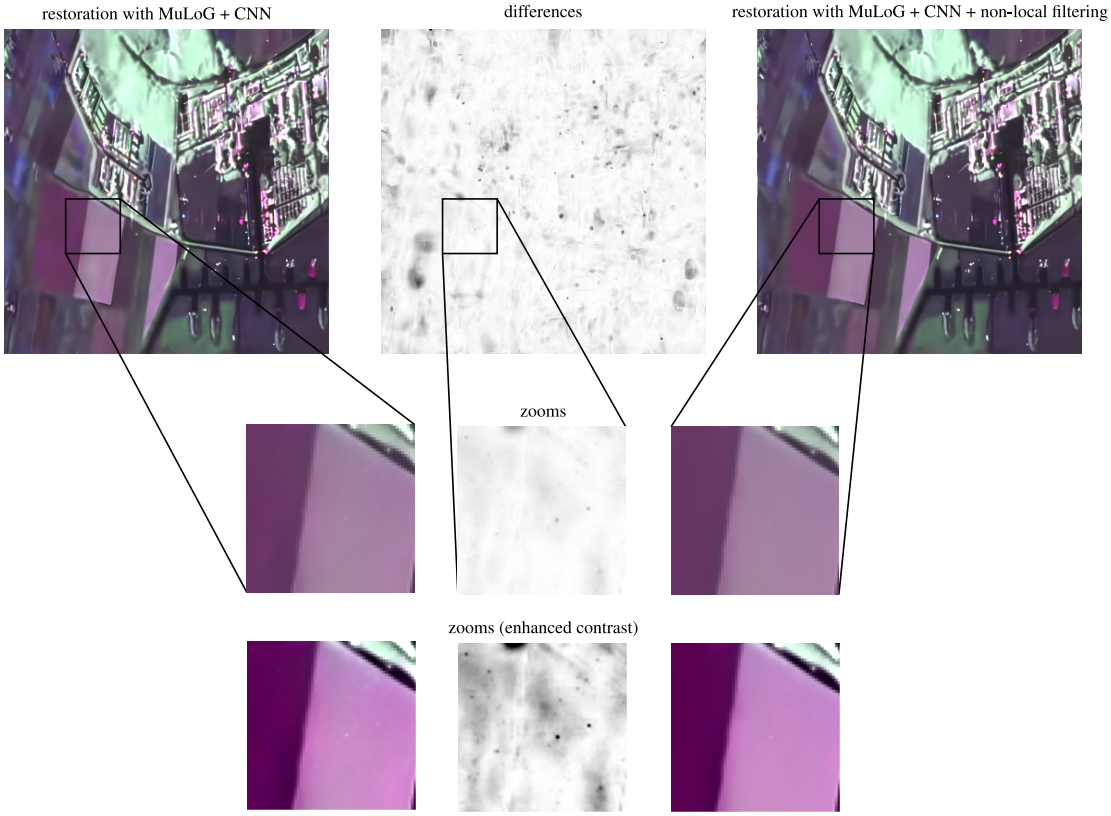


Fig. 3: Comparison of the restoration obtained after the CNN step and after the non-local filtering. Some point-like artifacts in homogeneous areas are strongly reduced by the non-local filtering.

5. REFERENCES

[1] J-S Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Comp. graph. image proc.*, vol. 17, no. 1, 1981.

[2] H Xie, L E Pierce, and F T Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE trans. Geosc. Remote Sens.*, vol. 40, no. 10, 2002.

[3] J-F Aujol, G Aubert, L Blanc-Féraud, and A Chambolle, "Image decomposition application to SAR images," in *Int. Conf. Scale-Space Th. in Comp. Vision*. Springer, 2003.

[4] F Argenti and L Alparone, "Speckle removal from SAR images in the undecimated wavelet domain," *IEEE Trans. Geosc. and Remote Sens.*, vol. 40, no. 11, 2002.

[5] C 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, 2015.

[6] G Chierchia, D Cozzolino, G Poggi, and L Verdoliva, "SAR image despeckling through convolutional neural networks," in *IEEE IGARSS*, 2017, pp. 5438–5441.

[7] A Buades, B Coll, and J-M Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[8] C Deledalle, L Denis, G Poggi, F Tupin, and L Verdoliva, "Exploiting patch similarity for SAR image processing: the nonlocal paradigm," *IEEE Signal Process. Mag.*, 2014.

[9] A Davy, T Ehret, G Facciolo, J-M Morel, and P Arias, "Non-Local Video Denoising by CNN," *arXiv preprint arXiv:1811.12758*, 2018.

[10] T Plötz and S Roth, "Neural nearest neighbors networks," in *NIPS*, 2018, pp. 1095–1106.

[11] X Wang, R Girshick, A Gupta, and K He, "Non-local neural networks," in *IEEE CVPR*, 2018, vol. 1.

[12] C Cruz, A Foi, V Katkovnik, and K Egiazarian, "Nonlocality-reinforced convolutional neural networks for image denoising," *arXiv preprint arXiv:1803.02112*, 2018.

[13] C. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or How to Apply Gaussian Denoisers to Multi-Channel SAR Speckle Reduction?," *IEEE TIP*, vol. 26, no. 9, Sept 2017.

[14] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE TIP*, vol. 26, no. 7, 2017.

[15] C-A Deledalle, L Denis, F Tupin, and S Lobry, "Speckle reduction in PolSAR by multi-channel variance stabilization and Gaussian denoising: MuLoG," in *EUSAR*, 2018, pp. 1–5.

[16] C-A Deledalle, L Denis, and F Tupin, "MuLoG: A Generic Variance-Stabilization Approach for Speckle Reduction in SAR Interferometry and SAR Polarimetry," in *IGARSS*, 2018.