A REVIEW OF DEEP-LEARNING TECHNIQUES FOR SAR IMAGE RESTORATION

Loïc Denis[†]*, Emanuele Dalsasso*[‡] *and Florence Tupin*[‡]

[†]Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d Optique Graduate School,
Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France

[‡]LTCI, Telecom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France

ABSTRACT

The speckle phenomenon remains a major hurdle for the analysis of SAR images. The development of speckle reduction methods closely follows methodological progress in the field of image restoration. The advent of deep neural networks has offered new ways to tackle this longstanding problem. Deep learning for speckle reduction is a very active research topic and already shows restoration performances that exceed that of the previous generations of methods based on the concepts of patches, sparsity, wavelet transform or total variation minimization.

The objective of this paper is to give an overview of the most recent works and point the main research directions and current challenges of deep learning for SAR image restoration.

Index Terms— SAR imaging, speckle, deep learning

1. INTRODUCTION

Speckle phenomenon arises due to the coherent summation of echoes produced by elementary scatterers that project into the same SAR pixel. Mitigating the strong fluctuations of speckle has been a major issue since the beginnings of SAR imaging.

Multilooking, i.e., averaging SAR intensities in a spatial window around the pixel of interest, reduces speckle fluctuations at the cost of a dramatic resolution loss. More subtle approaches have thus been proposed to prevent from blurring structures with very different reflectivities: *pixel-selection* methods restrict the average to intensities close to that of the current pixel, *window-based* methods adapt the shape of the window (by locally selecting a window among a set of oriented windows, or by region growing), *patch-based* methods compare patches to identify (possibly disconnected) pixels with similar neighborhoods, *transform-based* techniques apply a transform (such the wavelet transform) to separate noise from the useful signal, *regularization* or *variational* methods minimize a cost function that expresses a tradeoff between the proximity to the speckled observation and spatial smoothness properties. Deep learning is a much more recent approach to speckle reduction. The data-driven nature of this approach offers an improved flexibility and the ability to capture a wide variety of features observed in SAR images (point-like scatterers, lines, curves, textures). In the following we describe how deep learning methods are designed and describe the main challenges of this quickly evolving research topic.

The authors would like to thank the French space agency CNES for funding this work.

2. KEY INGREDIENTS OF A DEEP LEARNING APPROACH FOR SAR DESPECKLING

2.1. Building a training set

A first but crucial step to design a deep learning method for speckle reduction is the choice of a training strategy. The most conventional approach to train a network is *supervised training* (Fig.1, block ❶). This strategy requires the building of a training set with pairs of speckled / speckle-free images. Such pairs can be obtained by generating simulated speckle from a ground-truth image. It is however difficult to obtain such speckle-free images. The main approach consists in reducing speckle fluctuations by temporally averaging images from a long time series. A major limitation of numerically generated speckle, though, is that it generally neglects speckle correlations. The shift between the speckled images used during training and the real images used at test time produces strong artifacts unless adaptations are done, such as image downsampling [5], or training on regions of real images carefully selected to reject any area that changed during the time series [1]. To prevent these limitations, *self-supervised strategies* use only speckle-corrupted images in the training phase. Pairs of co-registered SAR images obtained at two different dates (chosen so that speckle is temporally decorrelated between the images) can be used to drive the network to predict an estimate from the first image that is as close as possible to the second image (Fig.1, block ❷). Single-image self-supervision introduces a form of masking: the network accesses only unmasked values and is asked to guess the masked values (Fig.1, block ❸). Given the random nature of speckle phenomenon, the best guess for the network is the underlying reflectivity (i.e., the noiseless value at the masked pixel).

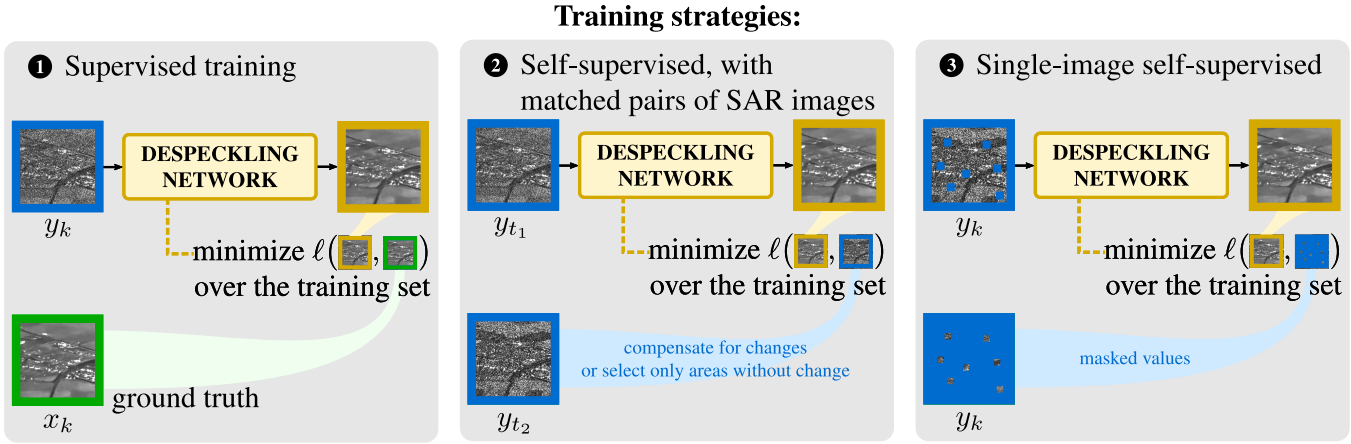


Fig. 1. Three training strategies have been explored in the literature: ❶ supervised training, using ground-truth images that match the speckled images provided as input to the network; ❷ self-supervised training, using co-registered pairs of SAR images captured at different dates; ❸ self-supervised training, using single images and a masking strategy: the network is trained to correctly infer the masked pixels of the input image.

2.2. Choosing a network architecture

There is a wide variety of network architectures available for image denoising. Two kinds of networks are generally used for SAR despeckling: (i) the convolutional structure of DnCNN [6] (obtained by stacking 15 to 20 layers formed by convolutions, possibly with dilation [7], batch normalization and a ReLU activation function), trained in a residual fashion, and (ii) the U-Net [8] (originally used for image segmentation, that takes the form of a particular auto-encoder with skip-connections).

2.3. Handling the high dynamic range of SAR images

Due to the physics of SAR imaging, the dynamic range between echoes produced by weakly scattering surfaces and the very strong returns generated by trihedral structures typically spans several orders of magnitude. Normalization and compression of the range of SAR image intensities is a crucial step: it strongly reduces the risk of falling outside the domain covered during the training phase of the network. Many works apply a logarithm transform to the SAR intensities before the deep neural network. This has two beneficial effects: it compresses the range of input values (so that it is much less likely to find strongly out-of-range values at test time) and it stabilizes the variance of speckle fluctuations (which may simplify despeckling). When SAR images are processed by the network in the original domain (i.e., without log-transform), the largest values are typically clipped to reduce the dynamic range, see for example [4].

2.4. Selecting a loss function

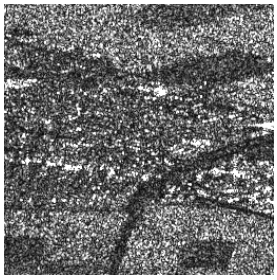
The most widely used loss function for regression is the squared ℓ_2 norm. To reduce the impact of the training samples that are poorly modeled, an ℓ_1 norm can be preferred. Total variation is sometimes considered as an additional term to penalize oscillations and thus limit the apparition of artifacts when applied to images that differ from the distribution of images considered during training (e.g., when speckle is spatially correlated at test time) [5]. Loss terms that enforce a good fit with the theoretical distribution of speckle have also been recently considered [9]. Perceptual losses can be used in supervised training strategies to give more weight to artifacts that may be interpreted as visual clues of meaningful content in the image. Generative Adversarial Networks (GANs) can be used to train a discriminator whose aim is to recognize restored images based on some artifacts of the restoration technique. Training the restoration network to fool the discriminator is then a way to obtain more plausible restoration results, at the cost of increasing the risk of also fooling the human by adding fake content that looks realistic [10].

Self-supervised training strategies require adapted loss functions. In the case of self-supervision with matched pairs of SAR images, it is important to compensate for changes that occurred between the two dates [3]. Single-image self-supervision requires to limit the computation of the loss to the masked pixels, or the use of a specific network architecture that prevents the receptive field to contain the central pixel [4].

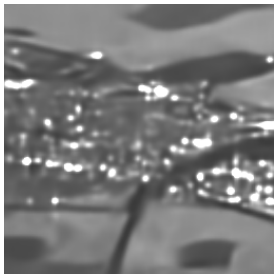
3. CURRENT CHALLENGES AND TRENDS

3.1. Self-supervision

In remote sensing, huge amounts of images are available but ground truths are scarce and costly to produce. Numerical simulations only imperfectly reproduce the complexity of actual systems. The development of learning strategies that rely



(a) Sentinel-1 (SLC)







(b) restoration with SARCNN (c) restoration with SAR2SAR (d) restoration with speckle2void.

Fig. 2. Restoration of the single-look Sentinel-1 image shown in (a) with deep-learning methods illustrative of the 3 training strategies shown in Fig.1: (b) SAR2SAR [1] uses a **supervised training** strategy (here, the training is performed on synthetic speckle and the Sentinel-1 image is downsampled by a factor 2 to limit speckle correlation, see [2]); (c) restoration with SAR2SAR [3], a network trained with a **self-supervised approach with pairs** of Sentinel-1 images of the same area captured at different dates; (d) restoration with the **single-image self-supervised** method speckle2void [4].

solely on actual observations is thus very appealing. Specific challenges face these strategies, however, such as the compensation of temporal changes (when co-registered image pairs acquired at different dates are considered) or the correlation of speckle (in particular for masking approaches).

3.2. Extensions to polarimetric and/or interferometric SAR

Most deep learning approaches for speckle reduction focused on the case of intensity images. Multi-channel complex-valued SAR images, as in SAR polarimetry or in SAR interferometry, raise other challenges. Polarimetric and interferometric information are encoded in complex-valued covariance matrices. Restricting the estimated matrices to the cone of positive definite covariance matrices requires an adequate design of the learning strategy and/or of the network. Due to the increase of the dimensionality of the data and of the unknowns, the learning task becomes more complex and it is expected that many more training samples are required to capture all spatial and polarimetric/interferometric configurations during the learning phase.

A notable approach to address these issues consists in applying a plug-in ADMM strategy to account for the statistics of speckle in polarimetric and interferometric SAR imaging [11]. By decomposing the SAR images into almost independent channels, deep neural networks can be readily applied, see Fig.3 and [12].

3.3. Extension to time series

Satellite constellations such as ESA's Sentinel-1 provide very long time series. The frequent revisit time and the temporal decorrelation of speckle offer the potential of very effective speckle suppression by (spatio)-temporal filtering. Versatile networks able to process temporal stacks of various size would be of great value to analyze these images.

3.4. Understanding and characterizing the restoration results

A limitation of deep learning methods is their lack of explainablity: due to the highly non-linear nature of the networks and their numerous parameters, it is very hard to grasp how a network produced a given result and to characterize the different artifacts that may be produced at test time. An approach to improve the explainability of deep learning methods is to combine them with more traditionnal processing techniques such as patch-based methods [13, 14].

4. CONCLUSION

SAR image restoration with deep neural networks is an extremely active research area, with very convincing results and several open research directions. The limited space of this paper was unsufficient to adequately cite the quickly growing literature on the subject. We focused on providing a broad

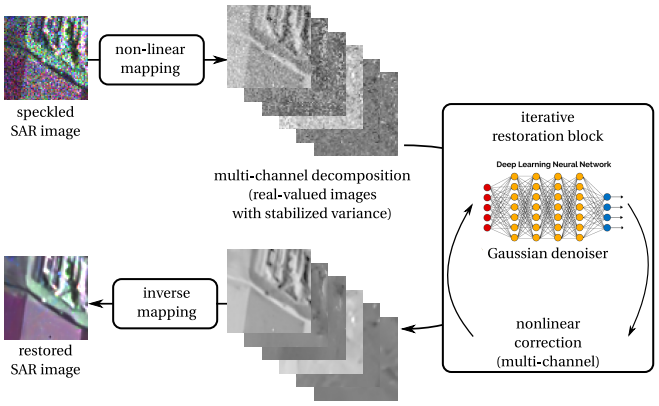


Fig. 3. MuLog [11] is one of the first approaches to apply deep neural networks to speckle reduction in polarimetric and interferometric SAR restoration. It works in a transformed domain in which complex-valued polarimetric and/or interferometric matrices are decomposed into real-valued channels with an approximately stabilized variance. In this domain, a deep neural network is applied iteratively until the channels are restored. Extending deep learning methods to polarimetric and/or interferometric SAR data is a hot topic.

view on the key elements of deep learning techniques for speckle reduction and invite the interested reader to refer to much more extensive reviews such as [15] and [16].

5. REFERENCES

[1] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "SAR image despeckling through convolutional neural networks," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017, pp. 5438–5441.

[2] E. Dalsasso, X. Yang, L. Denis, F. Tupin, and W. Yang, "SAR Image Despeckling by Deep Neural Networks: from a pre-trained model to an end-to-end training strategy," *Remote Sensing*, vol. 12, no. 16, pp. 2636, 2020.

[3] E. Dalsasso, L. Denis, and F. Tupin, "SAR2SAR: a self-supervised despeckling algorithm for SAR images," *arXiv preprint arXiv:2006.15037*, 2020.

[4] A. B. Molini, D. Valsesia, G. Fracastoro, and E. Magli, "Speckle2Void: Deep Self-Supervised SAR Despeckling with Blind-Spot Convolutional Neural Networks," *arXiv preprint arXiv:2007.02075*, 2020.

[5] E. Dalsasso, L. Denis, and F. Tupin, "How to handle spatial correlations in SAR despeckling? Resampling strategies and deep learning approaches," *hal-02538046*, 2020.

[6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. 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] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a dilated residual network for SAR image despeckling," *Remote Sensing*, vol. 10, no. 2, pp. 196, 2018.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[9] S. Vitale, G. Ferraioli, and V. Pascazio, "Multi-Objective CNN-Based Algorithm for SAR Despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[10] P. Wang, H. Zhang, and V. M. Patel, "Generative adversarial network-based restoration of speckled SAR images," in *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. IEEE, 2017, pp. 1–5.

[11] C.-A. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or how to apply Gaussian denoisers to multi-channel SAR speckle reduction?," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4389–4403, 2017.

[12] Charles-Alban Deledalle, Loïc Denis, and Florence Tupin, "MuLoG: a generic variance-stabilization approach for speckle reduction in SAR interferometry and SAR polarimetry," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 5816–5819.

[13] L. Denis, C.-A. Deledalle, and F. Tupin, "From patches to deep learning: combining self-similarity and neural networks for SAR image despeckling," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019, pp. 5113–5116.

[14] D. Cozzolino, L. Verdoliva, G. Scarpa, and G. Poggi, "Nonlocal CNN SAR Image Despeckling," *Remote Sensing*, vol. 12, no. 6, pp. 1006, 2020.

[15] X. X. Zhu, S. Montazeri, M. Ali, Y. Hua, Y. Wang, L. Mou, Y. Shi, F. Xu, and R. Bamler, "Deep learning meets SAR," *arXiv preprint arXiv:2006.10027*, 2020.

[16] G. Fracastoro, E. Magli, G. Poggi, G. Scarpa, D. Valsesia, and L. Verdoliva, "Deep learning methods for SAR image despeckling: trends and perspectives," *arXiv preprint arXiv:2012.05508*, 2020.