Remote Sensing

Emeric AURIANT, Marien RENAUD

Abstract

Synthetic Apperture Radar (SAR) is a kind of remote sensing imaging technique that can captured regardless of weather or light conditions. However, it presents a significant noise, called speckle. This noise is a crucial problem for SAR images and a despeckeling step is compulsory before using the images. Neural network are often use to do this denoising task. Here, the neural network models are improved by using multi-temporal stacks of images. The SAR2SAR model was studied and used to introduce a new denoiser, the SARratio2SARratio model. This model is based on images called ratio-images. Finally the different despeckeling techniques are compared qualitatively and quantitatively.

Introduction

Synthetic Apperture Radar (SAR) is a popular remote sensing technique due to its numerous advantages compared to optical imaging technologies. An antenna positioned on a moving object emits pulses radio waves and record the received echos. Some signal processing of the echos allows to create a **Syntethic Antenna Aperture** whose resolution is higher than the one of the original antenna. Moreover, this type of imagery is possible in cloudy weather or even at night. The main problem of the SAR images is a strong noise called the **speckle phenomenon**.

The reduction of the speckle is a widely discussed topic and various methods have been tested. Lately, deep learning algorithm and particularly convolutionnal networks have been tested [1]. The **SAR2SAR model** [2] is self-supervised Convolutional Neural Network using a stack of images from the same location to learn to denoise them.

New kind of images can be defined with the image stack. The **super-image** is a mean of the images of the stack and allows to detect the variable and stationary components of the images. The **ratio-image** is simply a component-wise ratio between an image and the related super-image. These images can be used with the previous SAR2SAR model and improve the quality of the denoising [3].

To go further, in this document, the SAR2SAR model is trained directly with ratio-images and the denoised results are compared with the previously presented methods. The comparison is made with qualitative observations but also with a quantitative metric the **Peak signal-to-noise ratio (PSNR)**.

1 Preliminaries

Model for the intensity. An imaged is composed of pixels, in the considered model each pixel is independent. Furthermore, in a SAR context, the Goodman model [4] provides a model for the intensity I of each pixel. The intensity I is composed of the reflectivity R and of a multiplication noise, the speckle S. The same quantities can be defined in the logarithmic domain, the log-intensity is noted i, the log-reflectivity r and the log-speckle s. In the logarithmic space, the noise is additive. Formally:

$$I = S \times R \tag{1}$$
$$i = s + r \tag{2}$$

Moreover, the speckle S follows a Gamma distribution of parameters $(L, \frac{1}{L})$ with L the number of looks of the image and Γ the gamma function:

$$\forall S \geqslant 0, \quad \mathbb{P}(S) = \frac{L^L}{\Gamma(L)} S^{L-1} \exp(-LS) \tag{3}$$

In the logarithmic domain, the log-speckle s also follows a know distribution: the Fisher-Tippett distribution:

$$\forall s, \quad \mathbb{P}(s) = \frac{L^L}{\Gamma(L)} \exp(Ls) \exp(Le^s) \tag{4}$$

These distributions are well known and can be simulated easily to create artificial noisy images. All the notations are listed in Section 5.

Peak Signal-to-noise ratio. The PSNR metric is commonly used to measure the quality of reconstruction between two images or two signals. It is a logarithmic quantity expressed in decibels (dB), the typical values between a noisy image and the relative ground-truth image are of about 10 dB. The expression of the PSNR for an image of intensity $\mathbf{I}$) and a ground-truth intensity $\mathbf{I_{gt}}$ is:

$$PSNR = 10 \log_{10} \left(\frac{\max \mathbf{I_{gt}}^2}{MSE(\mathbf{I}, \mathbf{I_{gt}})} \right) \tag{5}$$

with MSE the mean squared error.

This quantity is useful to evaluate the quality of the reconstruction but only provides global information. The reconstructed images must also be analyzed locally.

Lee Filter. A common denoising filter for SAR images is the Lee Filter [5]. This denoising is performed by doing a weighted average between the original noisy intensity I and the intensity on the locally averaged image $\tilde{I}$. The weighting coefficient is local and depends on the coefficient of variation $\tilde{\gamma}_p$:

$$\tilde{\gamma}_p = \frac{\tilde{\sigma}_p}{\tilde{\mu}_p}$$

where, $\tilde{\sigma}_p$ is the standard deviation of the locally averaged image and $\tilde{\mu}_p$ the mean of the locally averaged image centered in the pixel p. The Lee filter is defined by:

$$\hat{R} = \tilde{I} + k(I - \tilde{I}) \tag{6}$$

where $k_p = 1 - \frac{1}{L\tilde{\gamma}_p}$ and $\hat{R}$ is an estimation of the reflectivity R. The Lee filter can be used as a simple speckle remover [5]. Despite its simplicity, this filter will be useful as a non-trivial reference of a denoiser algorithm.

2 SAR2SAR

The speckle introduced before makes SAR images difficult to analyse. Various methods have been developed to denoise such images. Here the focus will be on the **SAR2SAR** model, a self-supervised deep learning algorithm [2].

2.1 Classical SAR2SAR approach

The classical SAR2SAR model is a Convolutional Neural Network (CNN) using a stack of log-intensities $\{i_1, ..., i_n\}$. As the distribution of the log-speckle is a Fisher-Tippett distribution, an efficient loss according to *E. Dalsasso et al.* [2] is the negative log-likelihood:

$$\mathcal{L}(x, y) = y - x + \exp(x - y)$$

where x and y are real values.

Most of the time, the ground truth is not available, the model has to be trained in a self-supervised setting. In this case, the ground truth was available and the model was trained by creating noisy data and then denoising them.

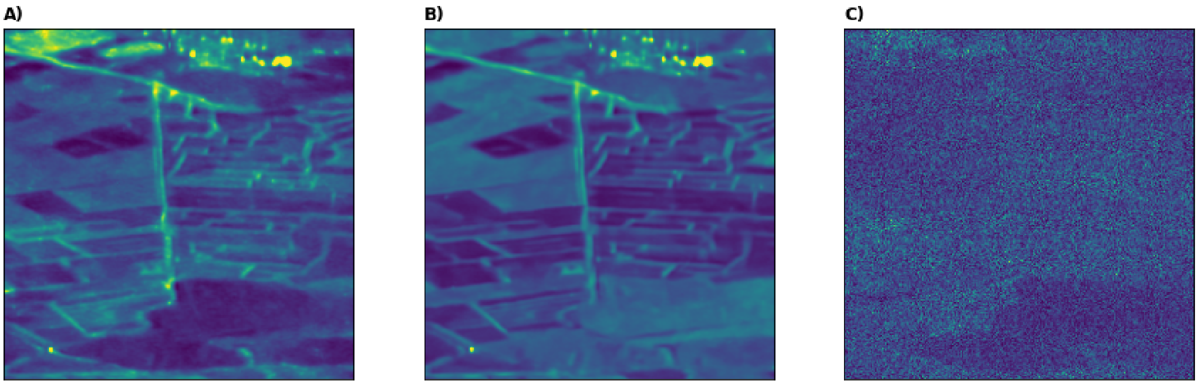

Figure 1: Various images for the *marais2* stack. The dimension is 256×256 pixels. **A)** Ground truth image. **B)** Super image. **C)** Ratio image.

2.2 Multi Temporal Ratio-based SAR2SAR

The speckle is independent between the images of the database $\{I_1, ..., I_n\}$. However, the reflectivity is the same for all the images. This leads to the idea of **super-images** which is simply a temporal averaging of the images of the database to reduce the speckle. The super-image intensity $\overline{I}$ and the log-super-image intensity $\overline{i}$ ($\overline{i} = \ln(\overline{I})$) can be defined as follow:

$$\overline{I} = \frac{1}{n} \sum_{k}^{n} I_k$$

By doing this averaging, the intensity of the speckle is reduced. A spatial filtering can be done to reduce even further the amplitude of the speckle. The figure 1 shows a ground truth image and the associated super image for the stack *marais2*. In the upper-left corner or the bottom-right corner the two images are different. This means that these areas changed over the various acquisitions. In contrast, the roads are very similar on both images because they undergo little variation over time. Overall, the speckle is reduced in the image.

This super-image can then be used to define a new category of image: the ratio images. For the intensity I_k, the ratio image intensity I_k^r is: $I_k^r = I_k/\overline{I}$. An example of ratio image is presented on the figure 1. The ratio image highlight different phenomenons. The speckle is clearly visible on the image. Some homogenous areas appears, for example in the upper-left or in the bottom-right corner of the image. These areas are highlighted because they reflectivity was not the same on all the image.

These ratio images can then be denoised, for example with the SAR2SAR algorithm [3]. Before using the SAR2SAR algorithm, the distribution of the pixels of the ratio images needs to approximately the same than the one of the images used for the training. To do so, each image is shifted (in the logarithmic domain) to have the first moments than the original image. The figure 2 shows the distribution for the entire stack. Even after the normalization, the distributions are not exactly the same. Normalizer according to the first two moments, a shift and a scale, was tested but gave slightly worse results. This difference between distributions could cause generalization errors during the denoising. Finally the denoised ratio-image intensity is multiplied by the super-image intensity to obtain the denoised image intensity.

3 Ratio Variation

The SAR2SAR algorithm is trained on SAR images, to go further and avoid generalization errors, an idea would be to train the SAR2SAR model directly on ratio images. To do so, the noise of the ratio images needs to be modeled. The ratio-image of an image I_k can be rewritten as:

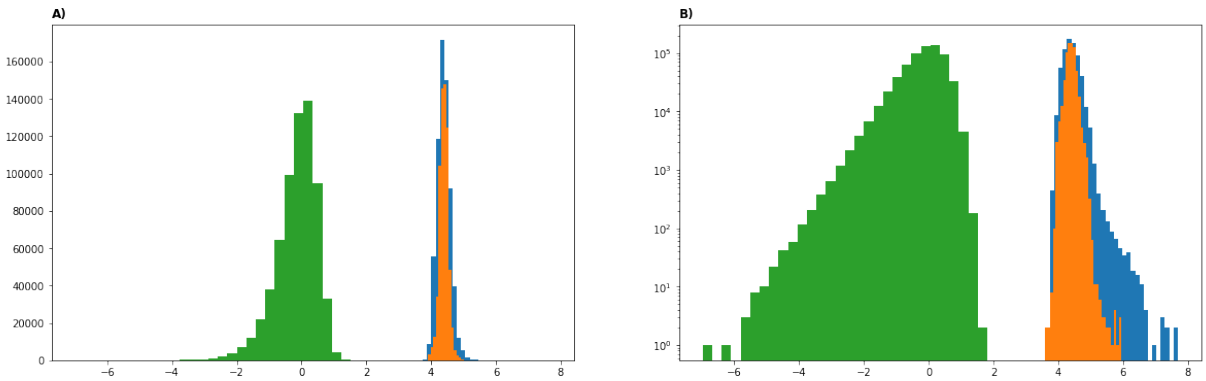

Figure 2: Distributions of the values for each pixels in the natural domain **A)** and in the logarithmic domain **B)**. **Blue.** Ground Truth. **Green.** Ratio images. **Orange.** Normalized Ratio images.

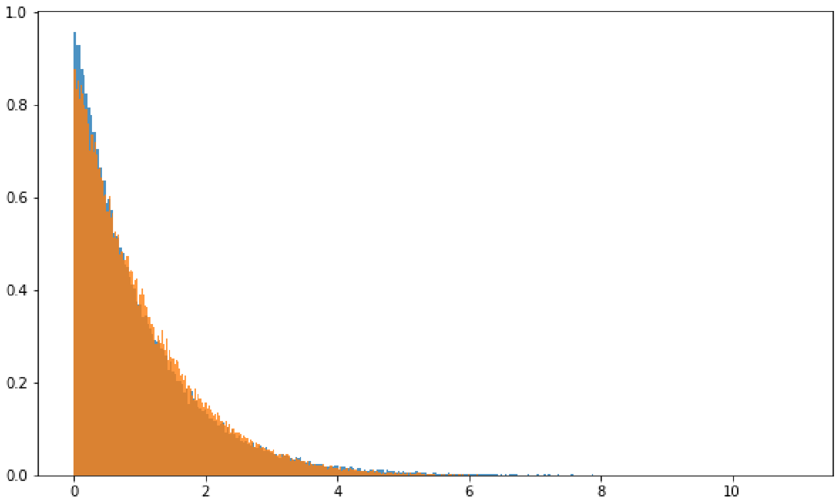

Figure 3: Distribution of the noise with 10000 samples. **Blue.** Distribution of the speckle ($\Gamma(1,1)$). **Orange.** Distribution of the ratio speckle.

$$I_k^r = \frac{I_k}{\bar{I}} = \frac{S_k R_k}{\frac{1}{n}\sum_l^n S_l R_l} = \underbrace{\frac{S_k \frac{1}{n}\sum_l^n R_l}{\frac{1}{n}\sum_l^n S_l R_l}}_{\text{random component}} \times \underbrace{\frac{R_k}{\frac{1}{n}\sum_l^n R_l}}_{\text{deterministic component}} \tag{7}$$

Two terms can be identified in the previous equation. The first one is a multiplicative noise that does not follow a Gamma distribution. This noise is difficult to estimate as it depends on the sum of the reflectivities R_k which is unknown. The second one is a deterministic ratio of the reflectivities of the images composing the stack. To generate this noise, a simplifying hypothesis is made. In the case where the reflectivity is constant ($R_k = R$):

$$I_k^r = \frac{I_k}{\frac{1}{n}\sum_k^n I_k} = \frac{S_k}{\frac{1}{n}\sum_l^n S_l} \tag{8}$$

As the S_k are independents, the denominator is a sum of independent Gamma variables. This phenomenon can thus easily be generated as a division of two non-independent random variables. This hypothesis is tested by simulating the noise and comparing it with a Gamma distribution 3. As the distributions are close, this hypothesis is credible. The noise of the ratio image can be generate as a classical speckle noise without much errors.

The distribution of the noise in the original images and in the ratio image are similar, the SAR2SAR model can be trained with ratio images. This new model is called SARratio2SARratio.

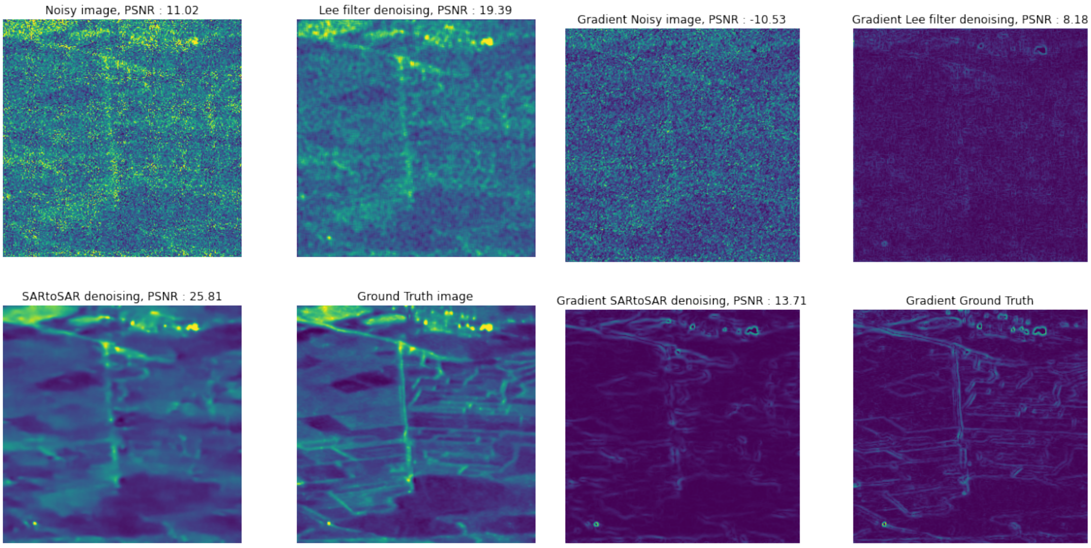

Figure 4: The denoising of one image.

4 Results

Experiment The neural network architecture used for this work is the same as SAR2SAR [2]. The database is constituted of 3 stacks of 10 images. The neural network was trained with 30 epochs, images of size 256×256, a stride of 128 and 500 batches per epoch. All other parameters are the same as in [2]. The model, SARratio2SARratio has the same architecture and the same parameters than the SAR2SAR model. The SARratio2SARratio model was trained with a database composed of ratio-images. These ratio-images were created using the same stacks of images. After training the networks, the experiments were realised on Google Colab. The source code is disponible on GitHub [1].

Denoising of a single image As shown on Figure 4, two despeckling techniques were tested. The neural networks method gives a better reconstruction than a simple Lee Filter. Also, the PSNR of the SAR2SAR model is high, the quality of the denoising is good. But the norm of the gradient of the images, see Figure 4, clearly show that the reconstruction is not perfect with the SAR2SAR neural network, especially to reconstruct the edges, which are essentials for detection task [6] or object recognition [7]. This consideration motivates the implementation of techniques using ratio-images.

4.1 Quantitative aspects

The newly introduced ratio techniques (see Sections 2.2 and 3) now have to be compared in a quantitative way. Three patches of size 256×256 are selected from the images of database (see Figure 5) in order to understand the behavior of the methods on various landscapes such as: fields, city or water. Statistics were then made on these 30 images (3 patches for each of the 10 SAR images). These results are shown on Table 1. The PSNR between the norm of the gradient is used to evaluate the reconstruction of the outline. A method performs well if it is able to correctly reconstruct the image (high PSNR) and if it is able to correctly reconstruct the contours of that image (high PSNR on the norm of the gradient image).

The results are that the SAR2SAR method has good performances to reconstruct the image but cannot correctly reconstruct the edges. The SAR2SAR method used with ratio-images or the SARratio2SARratio techniques are equivalent with this PSNR metric. These methods give very similar quantitative results, it is not possible to discriminate between the two after this analysis.

[1] `https://github.com/Marien-RENAUD/Speckle_reduction_in_SAR_time_series_using_an_updating_strategy`

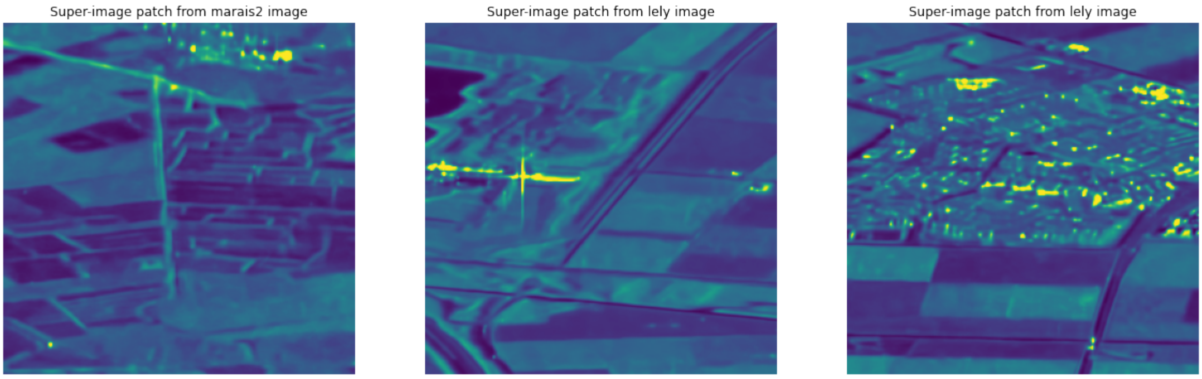

Figure 5: The super-image from the three selected patches.

Methods	**PSNR**	**PSNR between gradient norm**
Noisy	11.25 ± 1.75	-4.80 ± 5.38
SAR2SAR	22.81 ± 2.62	13.14 ± 3.88
SAR2SAR (ratio-images, noisy super-image)	19.58 ± 1.47	4.41 ± 5.23
SARratio2SARratio (noisy super-image)	19.61 ± 1.46	4.44 ± 5.23
SAR2SAR (ratio-image, Lee filter)	23.39 ± 1.77	12.69 ± 2.40
SARratio2SARratio (Lee filter)	23.24 ± 1.83	12.59 ± 2.36
SAR2SAR (ratio-image, BM3D)	26.17 ± 1.46	16.38 ± 2.66
SARratio2SARratio (BM3D)	26.17 ± 1.46	16.35 ± 2.58
SAR2SAR (ratio-image, GT)	28.08 ± 1.48	19.29 ± 2.72
SARratio2SARratio (GT)	28.18 ± 1.57	19.28 ± 2.76

Table 1: Comparison of PSNRs for various denoising methods.

It is clear, by looking at the Table 1, that the quality of the super-image is a key point regarding the performance of the denoising method. When the super-image is untouched (SAR2SAR with ratio-images made with the noisy super-image or SARratio2SARratio with the noisy super-image), the performance is worse than with the SAR2SAR model. In this case, the residual speckle of the super-image is not removed and will be present in the final image.

As soon as the super-image is denoised, for example with a Lee filter (SAR2SAR with ratio-images and Lee filter or SARratio2SARratio with Lee filter) performances are improved. With a better denoising method, such as BM3D in the logarithmic domain (SAR2SAR with ratio-images and BM3D or SARratio2SARratio with BM3D), the performances starts to be comparable to the images obtained with the ground truth super-image, both on the image itself and on its edges. BM3D is a gaussian denoiser which take the noise level at an input. This algorithm was choosen because the log-speckle is approximated to be gaussian [8]. The noise level is approximated by an image processing technique based on convolutions.

The last used method consists in using the ground truth super-image (SAR2SAR with ration-image made with the ground truth super-image and SARratio2SARratio with the ground truth super-image). This denoising gives an upper bound to the ratio-images methods. In practice, this ideal super-image is not accessible but can be estimated. The problem of denoising the super-image is easier than the initial problem because the noise-level is reduced thanks to the averaging.

4.2 Qualitative aspects

The PSNR metric gives a global idea of the quality of the denoising. To go further, qualitative analyses are crucial to understand the algorithm behaviour and detect local issues in the reconstruction.

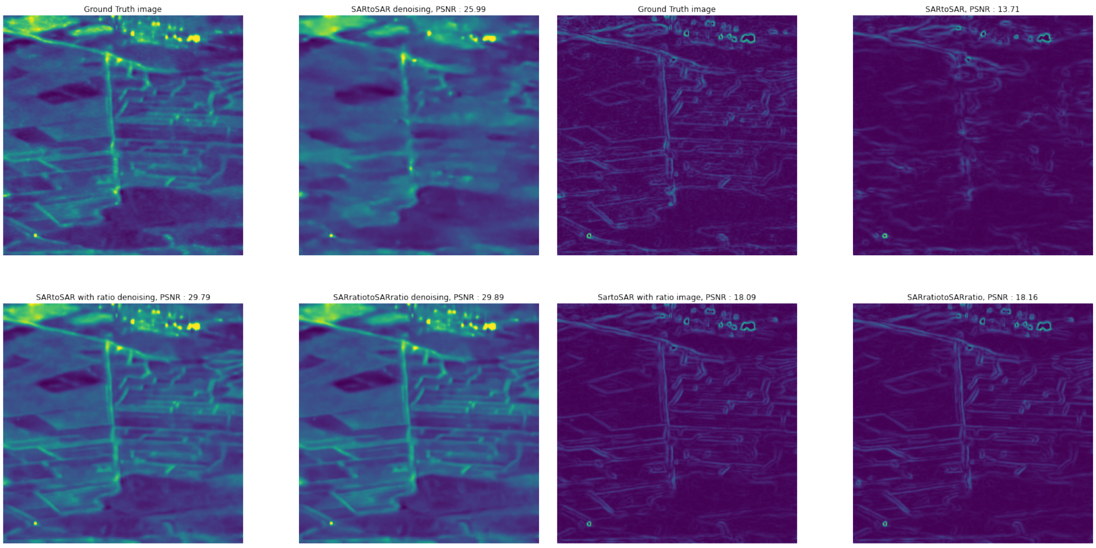

Figure 6: Comparison of methods with a ground-truth super-image.

Ideal Case In the ideal case, the super-image is perfectly denoised. The artefacts are only induced by the method. This is the case with the *SAR2SAR model with ratio-images with ground truth super-image* and the *SARratio2SARratio model with ground truth super-image*. In these cases, the artefacts can be analysed and compared with the ground truth image. The figure 6 shows that the methods with ratio-images give a very good reconstruction of the original image with a high PSNR. Furthermore, the norm of the gradient of these images are well reconstructed. These methods correctly reconstruct the edges.

After looking at the image as a whole, a closer look at the artifacts generated by these methods on patches allows a precise comparison of the methods. The Figure 7 shows a zoom on a patch of size 50×50. The colored squares on the side of the Figure correspond to zoomed images of the respective square on the images. The structures in the bright part at the top-left of the image are lost in all the reconstructions, the light points in this area are no longer discernible. The SAR2SAR reconstruction also caused a loss of connexity of the road and artificially created a shadow in the center of the patch. This shadow is less accentuated but present in the SAR2SAR with ratio-image reconstruction. Finally, the SARratio2SARratio method does not have this artifact. The bright area at the bottom-left of the patch is not reconstructed by any of the methods. In general, the reconstructions look blurrier than the initial image, which is expected in this denoising context.

On Figure8, the patch is correctly reconstructed by the ratio-images techniques, not by the classical one. In the top-left of the patch, the road is only reconstructed by the ratio-images methods. In addition, both ratio-images methods have two vertically aligned light points at the bottom-center of the patch. Unfortunately, one of these two points does not exist in the initial image. This artifact highlights an issue, a detail was created during the denoising. This is very problematic and could create false alarms for detection application.

The qualitative analysis of these patches completes the demonstration of the clear superiority of the ratio-images methods over the SAR2SAR method. The two ratio-images methods provide similar results, but the SARratio2SARratio method seems to more accurately portray the initial image. However, both methods can add details, this may be due to the use of neural networks.

Real Case After this qualitative study of ratio-images methods in the ideal case, the focus is moved to the real case, the super-image is noisy. In the quantitative analysis, the super-image reconstruction method was highlighted as an essential part of the denoising process. With a well reconstructed super-image, the real case is comparable to the ideal case. This part aims to analyse the various techniques to reconstruct the super-image.

The super-image encodes all the stationary structures. On the other hand, the ratio-image highlight the time-varying component. Thus, the super-image is very structured while the ratio-image is not. This can be seen on Figure 1. It is very important to denoise the ratio-image

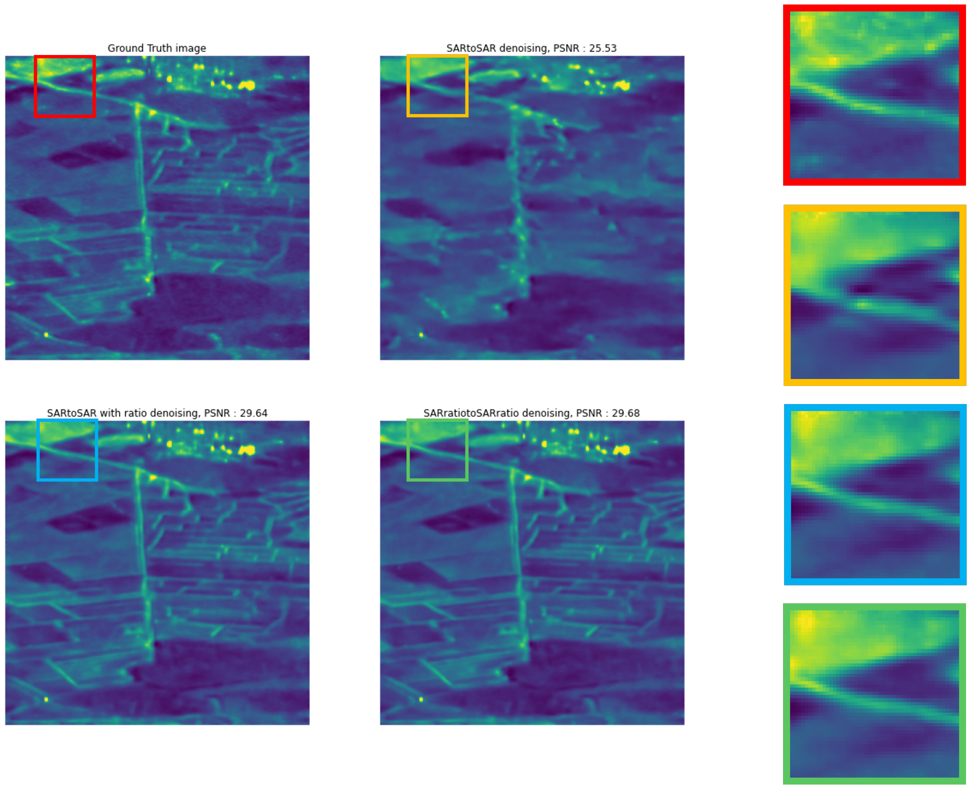

Figure 7: Artefacts deformation of the algorithm on the *marais* image

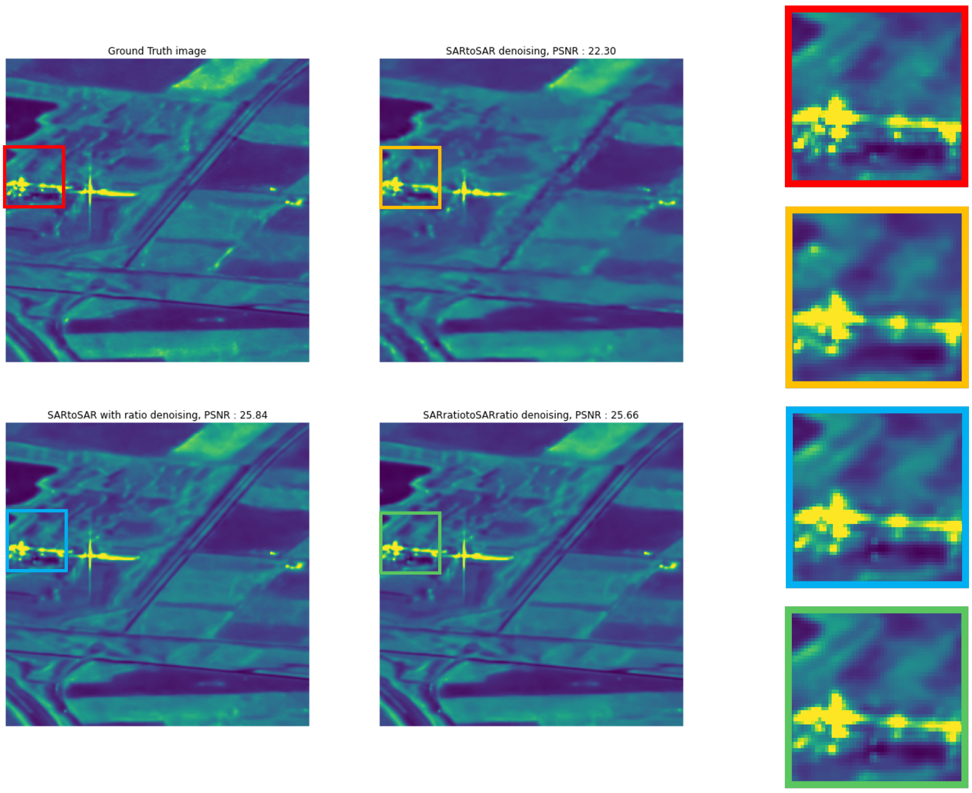

Figure 8: Artefacts deformation of the algorithm on lely image

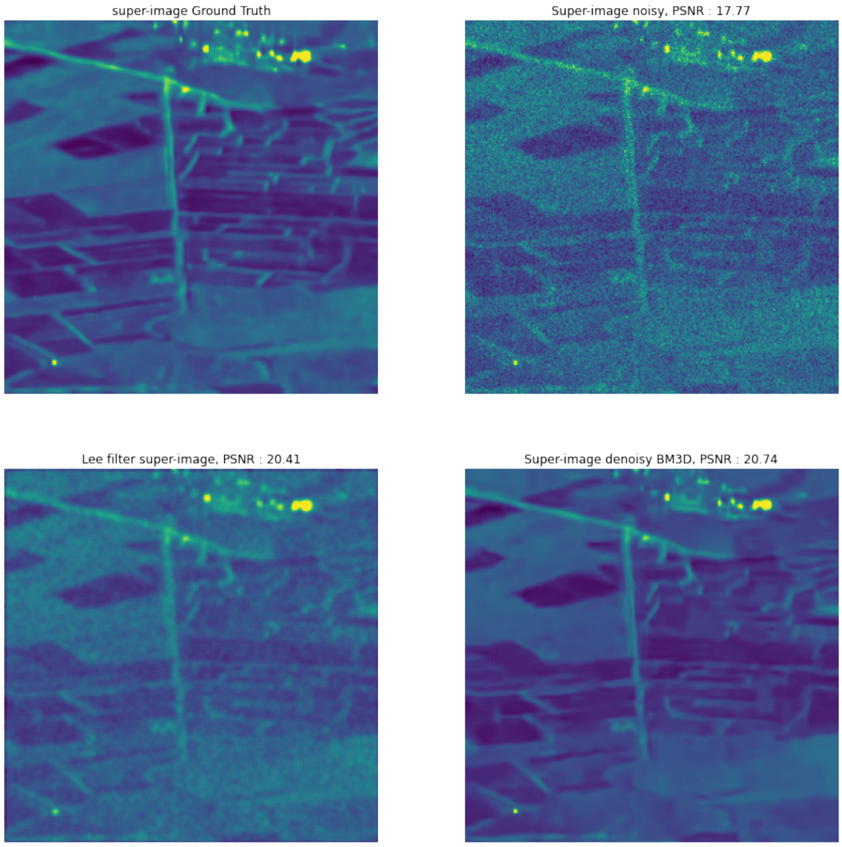

Figure 9: Super-image reconstruction with the three developed techniques

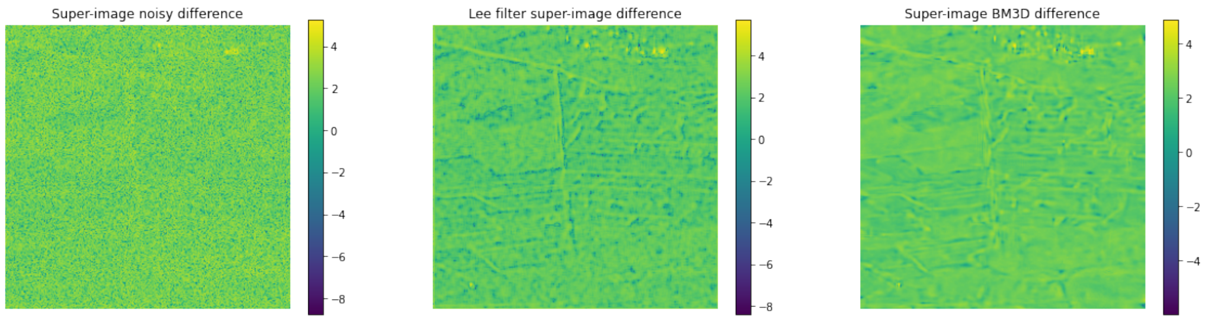

Figure 10: Differencies between the ground truth super-image and reconstructed in log-scale

without destroying these structures (*ie.* the edges).

By observing Figure 9, the Lee filter clearly does not totally eliminate the speckle. Indeed, the Lee filter uses local windows to average and blur the image on a spatial level. An average window of size 4×4 was used here and is not sufficient to eliminate the speckle. However, with a larger window, structures are destroyed and the ratio-image reconstruction is worse.

On Figure 9, the BM3D denoiser allows to reconstruct finer structures and gives an overall better denoising. The differences in logarithmic scale are presented on Figure 10. The speckle is suppressed with the BM3D algorithm but not with the Lee filter. In these two cases, information on the structure have been loss and the edges are not reconstructed perfectly. This denoising of the super-image is the main axis of improvement of our work. The reference article [3] uses the mu-log algorithm, powerful but difficult to implement in this project.

5 Conclusion

Taking into account multi-temporal information allows a better despeckling of SAR images. The SAR2SAR model and its variation the SAR2SAR model trained with ratio-images and the SARratio2SARratio model provide a neural network based denoising. The SARratio2SARratio with the BM3D denoising of the super-image is the more efficient technique developed in this work, both in terms of reconstruction of homogeneous areas and of edges.

As a perspective, the denoising of the super-image is the main axe of improvement. To go further, a finer choose of hyper-parameter for the neural network models may also improve the reconstruction quality and the supervised approach could be replaced by a non-supervised setting, where ground-truth images are not necessary. Finally, to solve the issue of creation of details, a probabilistic approach could be developed. For a example, a model could return an *a posteriori law* of the image knowing the observation. This would allow to give a score to each of the reconstructed pixels.

Notations

- I : Intensity image
- $i = \log(I)$: Log-intensity image
- S : Speckle
- $s = \log(S)$: Log-speckle
- R : Reflectivity
- $r = \log(R)$: Log-reflectivity
- $\overline{I}$: Super-image
- $I^r = I/\overline{I}$: Ratio image

References

[1] Puyang Wang, He Zhang, and Vishal M. Patel. Sar image despeckling using a convolutional neural network. *IEEE Signal Processing Letters*, 24(12):1763–1767, 2017.

[2] Emanuele Dalsasso, Loic Denis, and Florence Tupin. SAR2sar: A semi-supervised despeckling algorithm for SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:4321–4329, 2021.

[3] Emanuele Dalsasso, Ines Meraoumia, Loic Denis, and Florence Tupin. Exploiting multitemporal information for improved speckle reduction of sentinel-1 SAR images by deep learning. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, jul 2021.

[4] J. W. Goodman. Some fundamental properties of speckle∗. *J. Opt. Soc. Am.*, 66(11):1145–1150, Nov 1976.

[5] Aiyeola Sikiru Yommy, Rongke Liu, Wu, and Shuang. Sar image despeckling using refined lee filter. In *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, volume 2, pages 260–265, 2015.

[6] P Trivero, B Fiscella, F Gomez, and P Pavese. Sar detection and characterization of sea surface slicks. *International Journal of Remote Sensing*, 19(3):543–548, 1998.

[7] Fei Gao, Yue Yang, Jun Wang, Jinping Sun, Erfu Yang, and Huiyu Zhou. A deep convolutional generative adversarial networks (dcgans)-based semi-supervised method for object recognition in synthetic aperture radar (sar) images. *Remote Sensing*, 10(6):846, 2018.

[8] Hua Xie, Leland E Pierce, and Fawwaz T Ulaby. Statistical properties of logarithmically transformed speckle. *IEEE transactions on geoscience and remote sensing*, 40(3):721–727, 2002.