Robustness to spatially-correlated speckle in Plug-and-Play PolSAR despeckling

Cristiano Ulondu-Mendes, Loïc Denis, Charles-Alban Deledalle, Florence Tupin

▶ **To cite this version:**

Cristiano Ulondu-Mendes, Loïc Denis, Charles-Alban Deledalle, Florence Tupin. Robustness to spatially-correlated speckle in Plug-and-Play PolSAR despeckling. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62, 10.1109/TGRS.2024.3432180 . hal-04159195v2

HAL Id: hal-04159195

https://telecom-paris.hal.science/hal-04159195v2

Submitted on 30 Aug 2024

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.

Robustness to spatially-correlated speckle in Plug-and-Play PolSAR despeckling

Cristiano Ulondu Mendes, Loïc Denis, Charles Deledalle, Florence Tupin

***Abstract*—Synthetic Aperture Radar (SAR) provides valuable information about the Earth's surface in all-weather and day-and-night conditions. Due to the inherent presence of speckle phenomenon, a filtering step is often required to improve the performance of downstream tasks. In this paper, we focus on dealing with the spatial correlations of speckle, which impacts negatively many of the existing speckle filters. Taking advantage of the flexibility of variational methods based on the Plug-and-Play strategy, we propose to use a Gaussian denoiser trained to restore SAR scenes corrupted by colored Gaussian noise with correlation structures typical of a range of radar sensors. Our approach improves the robustness of Plug-and-Play despeckling techniques. Experiments conducted on simulated and real polarimetric SAR images show that the proposed method removes speckle efficiently in the presence of spatial correlations without introducing artifacts, with a good level of detail preservation. Our method can be readily applied, without network re-training or fine-tuning, to filter SAR images from various sensors, acquisition modes (SAR, PolSAR, InSAR, PolInSAR) and spatial resolution. The code of the trained models is made freely available at https://gitlab.telecom-paris.fr/ring/mulog-drunet.**

***Index Terms*—SAR, polarimetry, image despeckling, deep learning, correlated speckle.**

I. Introduction

POLARIMETRIC synthetic aperture radar (SAR) images offer rich information of the back-scattering mechanisms occuring throughout the scene [1]. They are particularly useful to retrieve physical information for land, ice, ocean, forest, or urban applications [2].

Yet, due to the speckle phenomenon, these images, like all SAR images, are difficult to analyze. The speckle phenomenon occurs due to the coherent summation of many elementary echoes within a radar resolution cell. It appears in the images in the form of strong signal-dependent fluctuations. The development of efficient speckle filtering methods that preserve at best the content of the images is therefore crucial for many applications, in particular those requiring the full spatial resolution provided by the instrument. Many approaches have been proposed in the past three decades to address this problem. They can be divided into several groups:

Local filtering: These methods aim to attenuate the fluctuations by combining pixel values within a small neighbourhood. The simplest one, the boxcar filter, consists in a simple moving average with a fixed-size square window. In [3]–[5] adaptive windows construction techniques are developed to better preserve the scene's features, and the filtering is performed using averaging or minimum mean square error (MMSE) based approaches.

Patch-based techniques: The target pixel is still denoised by combining the values of pixels contained within a neighborhood (this time, a much larger neighborhood defined by the so-called search window is used). The importance of each pixel of the search window is derived based on patch similarity. By discarding dissimilar pixels, patch-based methods perform adaptive smoothing: speckle is strongly reduced in homogeneous areas while geometrical and textural structures are preserved [6]. This paradigm has been successfully used for PolSAR image filtering [7], [8] and has improved the performances of various approaches including TV-based methods [9] and linear MMSE-based methods [10]. Most of these methods rely on a statistical model of the speckle to define patch similarity but recently [11] pointed out the downsides of such practice and proposed instead a model-free similarity measure using the ratio and inverse ratio of patches.

Anisotropic diffusion (AD) based methods: This group of methods estimates the speckle-free image as the solution of a partial differential equation of the following form [12]:

$$\partial_t \mathbf{u} = \mathrm{div}(\mathrm{G}(\mathrm{E})\nabla \mathbf{u})$$

where the initial state is the noisy image, ∇ and div denote respectively the gradient and divergence operators, E represents an edge estimator and the diffusion coefficient $\mathrm{G}(\cdot)$ is a scalar positive value. The challenge of adapting such a method to PolSAR despeckling lies in defining a suitable edge estimator for covariance matrices. [13], [14] define such estimator using similarity measures between pixels based on the Wishart statistical model of PolSAR data. [15] proposed to detect edges using the magnitude of ratio gradients computed on covariance matrices in the horizontal and vertical directions.

Variational methods: They formulate the restoration problem as the minimization of a cost function involving a data fidelity term that accounts for the statistics of speckle and a regularization term that promotes noise-free images (for example: smooth images with sharp edges [9], [16], [17]). The Plug-and-Play (PnP) framework [18]

C. Ulondu Mendes and F. Tupin are with LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France, e-mail: forename.name@telecom-paris.fr.

L. Denis is with Université Jean Monnet Saint-Etienne, CNRS, Institut d Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France, and also with LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France, e-mail: loic.denis@univ-st-etienne.fr.

C. Deledalle is with Brain Corporation, 10182 Telesis Ct, San Diego, CA 92121, USA.

is a special case where a Gaussian denoiser is used as a regularizer instead of explicitly defining it.

Deep learning: Neural networks with a large number of parameters are trained in order to learn how speckle fluctuations can be separated from the content of SAR scenes [19], [20]. In the case of single-channel SAR images, many training strategies have been proposed [21]: (i) supervised learning [22], [23] where pairs of speckle-free and corrupted images are provided (first, a set of reference images is built considering optical images [24] or despeckled images produced by averaging a long time series of co-registered images [25], [26], or by degrading the resolution of a very high-resolution image, then, corrupted versions of these images can be generated with synthetic speckle when they are not already available); (ii) semi-supervised learning uses pairs of images of the same area to train the network to predict the second image with only the first image as input (if speckle is temporally decorrelated, only the underlying reflectivity is output by the network provided that adequate change compensation is performed [27], [28]); (iii) self-supervised learning masks partially the input image: the central pixel in blind-spot approaches [29] or either the real or the imaginary component in MERLIN [30]. Despeckling polarimetric SAR images with deep learning is much more challenging than the case of single-channel intensity images: polarimetric information is represented in the form of complex-valued covariance matrices. The networks have to adapt both to the statistics of speckle (involving scene-dependent correlations between polarimetric channels) and to the high variability of polarimetric SAR scenes. Several approaches have been considered to address these difficulties. Foucher *et al.* [31] follow an additive decomposition of polarimetric SAR covariance matrices into the speckle-free covariances and two additional signal-dependent corruption terms. They then derive an iterative despeckling scheme involving a shallow neural network to suppress these additive corruptions. Most other approaches use a matrix logarithm transform to turn the speckle into an almost signal-independent additive component. This component is then suppressed thanks to an iterative processing [32] or in a single-step with a network involving only real-valued operations [33], or complex-valued processing [34]. Another direction consists of using a neural network to compute the relative weights of pixels in the final estimation. This can be seen as an extension of pixel-selection methods introduced by the seminal works of Lee [3], [35] and later particularly developed by patch-based approaches. First introduced to despeckle intensity images [36], this idea has been recently extended to polarimetric SAR images [37].

Many of the proposed speckle removal methods perform poorly when applied to real SAR images because they assume that the speckle is spatially independent, while in practice it is spatially-correlated. Preprocessing steps that decorrelate the speckle either modify the appearance of the SAR scene (side-lobes of the bright targets) or degrades the spatial resolution (by subsampling) [38]. Rather than whitening speckle [39], [40], it is preferable to design restoration methods that are robust to the spatial correlations of speckle to preserve the spectral apodization and sampling of the original SAR images. Only few despeckling methods are robust to speckle correlations. In single-channel despeckling, training a neural network to process images with spatially-correlated speckle provides robustness to these correlations. For multi-channel SAR images such as in PolSAR, robustness to spatially-correlated speckle has not been addressed yet by deep-learning-based methods. To the best of our knowledge, the only type of methods that can be applied to PolSAR images without a preprocessing step to reduce speckle correlations is the non-local methods [7]. The goal of this paper is to propose a despeckling framework for monochannel and multi-channel SAR images that is robust to spatial correlation through the joint exploitation of statistical models and deep neural networks. Our method builds on MuLoG [18], [41], a variational despeckling technique based on the PnP framework that supports SAR images with an arbitrary number of channels. An overview of the proposed approach is given in Figure 1. To illustrate the necessity of taking into account the spatial correlation of speckle and give a preview of the ability of the proposed framework to handle it, Figure 2 shows the results of MuLoG and our approach on an unprocessed PolSAR image.

Our main contributions are summarized as follows:

1) We provide the first-ever despeckling method robust to spatial correlations of speckle observed for a range of sensors, without requiring a decorrelating preprocessing. The proposed method is both bandwidth-agnostic and resolution-agnostic. We show its successfull application to metric and submetric resolution images from 7 different sensors (airborne: E-SAR, F-SAR, SETHI, UAVSAR; satellite: RADARSAT-2, RCM, and TerraSAR-X).
2) The method can handle single-channel (intensity) or multi-channel (polarimetric and/or interferometric) SAR images from various sensors with a single neural network (no re-training necessary). The state-of-the-art capabilities of our approach are demonstrated on PolSAR and InSAR. Additional results on single-channel SAR and PolInSAR data can be found on supplementary material available at https://gitlab.telecom-paris.fr/ring/mulog-drunet.
3) we make the source code available for reproducible research.

II. MuLoG: a generic despeckling method for mono-channel and multi-channel SAR images

In this section, we recap the statistical model of speckle and outline the main steps involved in MuLoG.

A. Statistical model of multi-channel SAR images

An important step in the development of a variational filtering method is the selection of an adequate statistical model. A widely used model to describe speckle in SAR

[1]RADARSAT is an official brand of the Canadian Space Agency

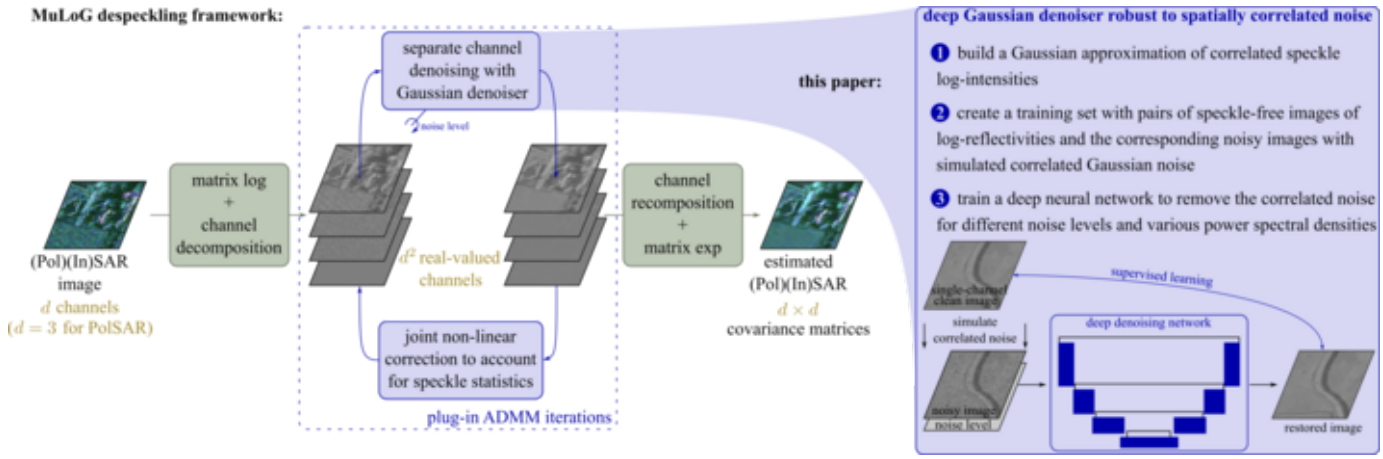
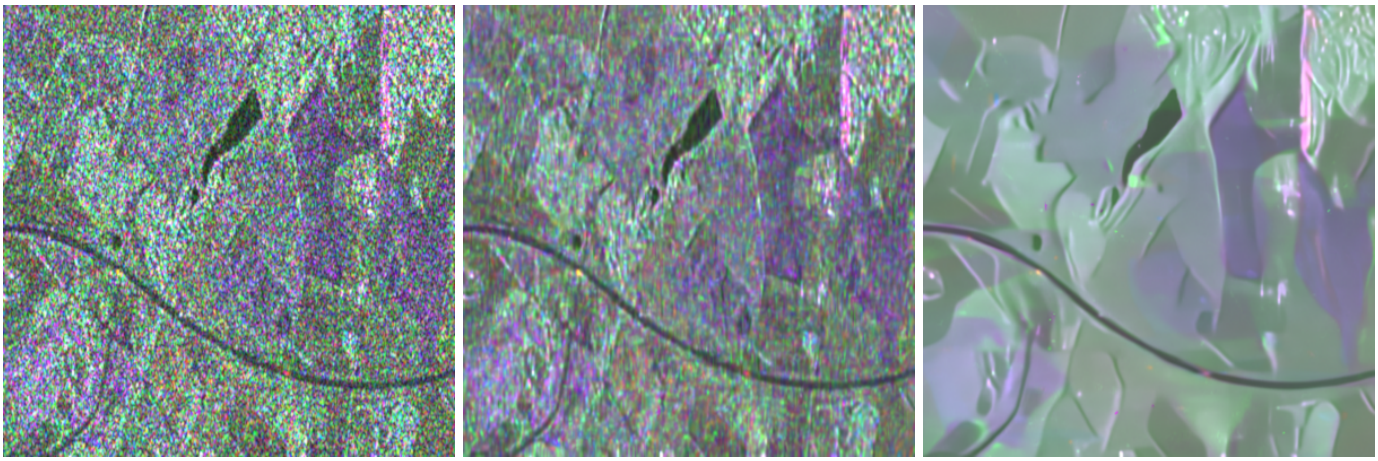


Figure 1: Proposed approach for addressing spatially-correlated speckle in multi-channel SAR image filtering.



(a) Real noisy PolSAR image (b) MuLoG + BM3D (c) Our approach

Figure 2: Comparing the original version of MuLoG and our proposed framework on a PolSAR image acquired with RCM [1]©Government of Canada(RADARSAT Constellation Mission) sensor. In its original version, MuLoG is based on the hypothesis that the speckle is spatially-uncorrelated causing severe artifacts when applied directly to a real SAR image 2b. Our approach, taking into account the spatial correlation, provides a high-quality filtered image 2c without pre-processing of the original image.

images is the Goodman model [42] in which the diffusion vector $\boldsymbol{k} \in \mathbb{C}^d$ at a given pixel of a d-channel SAR image (InSAR: $d = 2$, PolSAR: $d = 3$, PolInSAR: $d = 6 \ldots$) follows a d-variate complex circular Gaussian distribution $\mathcal{N}_c(\boldsymbol{\Sigma})$ defined by the probability density function (PDF):

$$p(\boldsymbol{k}|\boldsymbol{\Sigma}) = \frac{1}{\pi^d |\boldsymbol{\Sigma}|} \exp\left(-\boldsymbol{k}^* \boldsymbol{\Sigma}^{-1} \boldsymbol{k}\right) \tag{1}$$

where * refers to the Hermitian transpose and $|\boldsymbol{\Sigma}|$ denotes the determinant of matrix $\boldsymbol{\Sigma}$. This PDF is fully characterized by the covariance matrix $\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{k}\boldsymbol{k}^*] \in \mathbb{C}^{d \times d}$, a Hermitian positive definite matrix that contains all the polarimetric and/or interferometric information of interest. It is common practice to pre-estimate $\boldsymbol{\Sigma}$ by computing the empirical covariance $\mathbf{C}$ over a relatively small neighborhood ω (e.g., a square window centered on the pixel of interest). In the absence of spatial correlations of speckle, the number of pixels L in ω defines the number of looks (the number of independent looks is strictly smaller in the case of correlations). The empirical covariance $\mathbf{C}$ follows a Wishart distribution [43] parameterized by $\boldsymbol{\Sigma}$ and L: $\mathbf{C} \sim \mathcal{W}(\boldsymbol{\Sigma}; L)$, and its PDF is well-defined when $L \geq d$:

$$p_{\mathbf{C}}(\mathbf{C}|\boldsymbol{\Sigma}) = \frac{L^{dL} |\mathbf{C}|^{L-d}}{\Gamma_d(L) |\boldsymbol{\Sigma}|^L} \exp(-L \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})) \tag{2}$$

where $\Gamma_d(L) = \pi^{d(d-1)/2} \prod_{l=1}^{d} \Gamma(L-l+1)$ and Γ is the gamma function.

B. Brief recall of the MuLoG framework

The MuLoG framework corresponds to the extension to the multi-channel case of the variational approach MIDAL (Multiplicative Image Denoising by Augmented Lagrangian) [44] and the application of the Plug-and-Play (PnP) approach. MIDAL was introduced to estimate, in the logarithmic domain, the underlying reflectivity of a scene from the L-look intensity image. The authors of [44] highlight three benefits of this

domain change. Firstly, taking the logarithm of the intensity transforms the multiplicative nature of the speckle into an additive fluctuation with constant variance; secondly, it facilitates the resolution of the optimization problem derived from the MAP formulation by removing the positivity constraint under the multiplicative model; and finally, it makes the data fidelity term convex. MuLoG builds on these advantages to denoise PolSAR images which are much more challenging due to the complex relationships binding the different terms of the covariance matrices that must be preserved during the filtering process. Rather than defining an explicit regularization term, it adopts a PnP strategy: it includes a denoising step performed by an off-the-shelf denoiser suitable to remove additive white Gaussian noise (AWGN) in images. It can benefit readily from the continuing progress of research on Gaussian denoisers.

Let $\mathbf{C}$ be the speckled version of an n-pixels image of covariance matrices $\mathbf{\Sigma} \in \mathbb{C}^{n\times d\times d}$, where d is the number of channels ($d = 3$ in monostatic polarimetry). If the input data are available in the form of SLC diffusion vectors $\boldsymbol{k}$, full-rank covariance matrices $\mathbf{C}$ with a controlled condition number can be formed by modifying the eigenvalues of the rank-one matrices $\boldsymbol{k}\boldsymbol{k}^*$ (see Algorithm 5 and Sec. 4.2 in [41]), without altering the spatial resolution.

The main steps to filter the image $\mathbf{C}$ with the MuLoG filtering framework are the following [18], [41]:

1) at each pixel i, the matrix logarithm of the covariance matrix $\mathbf{C}_i$ is computed, leading to an image $\tilde{\mathbf{C}} \in \mathbb{C}^{n\times d\times d}$.
2) the d^2 real-valued degrees of freedom of these log-transformed covariance matrices are extracted and decorrelated using principal component analysis, leading to a d^2-independent-channel image $\mathbf{y} \in \mathbb{R}^{n\times d^2}$.
3) the maximum *a posteriori* estimator associated with the filtering of $\mathbf{y}$ is given by:

$$\hat{\mathbf{x}} \in \underset{\mathbf{x}\in\mathbb{R}^{n\times d^2}}{\operatorname{argmin}} \; -\log(p_{\mathbf{y}}(\mathbf{y}|\mathbf{x})) \underbrace{-\log(p_{\mathbf{x}}(\mathbf{x}))}_{\mathcal{R}(\mathbf{x})} \tag{3}$$

Solving this problem with the alternating directions method of multipliers (ADMM) consists of iterating the following steps:

$$\hat{\mathbf{z}}^i \leftarrow \underset{\mathbf{z}^i\in\mathbb{R}^n}{\operatorname{argmin}} \frac{\beta}{2}\|\mathbf{z}^i - \hat{\mathbf{x}}^i + \hat{\mathbf{m}}^i\|^2 + \mathcal{R}_i(\mathbf{z}^i)$$

for each real-valued channel i from 1 to d^2, (4)

$$\hat{\mathbf{m}} \leftarrow \hat{\mathbf{m}} + \hat{\mathbf{z}} - \hat{\mathbf{x}} \tag{5}$$

(Lagrange multipliers update)

$$\hat{\mathbf{x}} \leftarrow \underset{\mathbf{x}\in\mathbb{R}^{n\times d^2}}{\operatorname{argmin}} \frac{\beta}{2}\|\mathbf{x} - \hat{\mathbf{z}} - \hat{\mathbf{m}}\|^2 - \log(p_{\mathbf{y}}(\mathbf{y}|\mathbf{x}))\,. \tag{6}$$

The PnP strategy consists of replacing Equation (4) by the application of a Gaussian denoiser $\mathcal{D}_\sigma$, corresponding to an implicit regularization:

$$\forall i \in [\![1, d^2]\!], \quad \hat{\mathbf{z}}^i = \mathcal{D}_\sigma(\hat{\mathbf{x}}^i - \hat{\mathbf{m}}^i) \tag{7}$$

where $\sigma = \frac{1}{\sqrt{\beta}}$ controls the filtering strength.

4) At the end of the ADMM algorithm, the d^2 real-valued channels of $\hat{\mathbf{x}}$ are recombined into an image of complex covariance matrices, and a matrix exponential is applied to obtain the final result $\hat{\mathbf{\Sigma}}$.

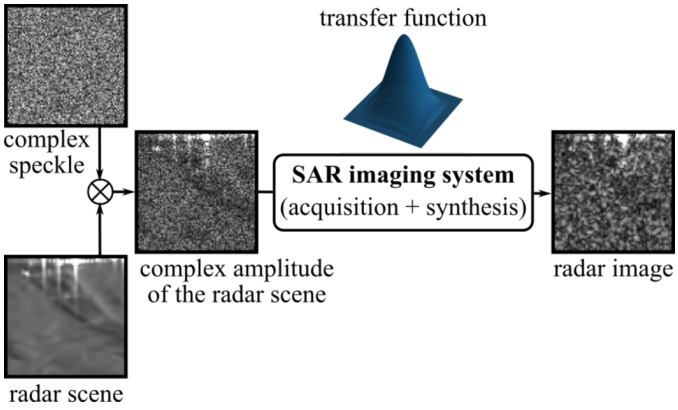


Figure 3: Generative model of SAR images: as the transfer function depends on specific choices (range and azimuth pixel size, spectral apodization) made during the radar image synthesis, the spatial correlation of the speckle noise may vary from one sensor to another.

Like MIDAL, MuLoG assumes the independence of speckle from one pixel to the next and artifacts appear if speckle is spatially-correlated. To suppress these artifacts, we propose in section III-B to improve the robustness of the Gaussian denoiser to spatially-correlated noise.

III. Improvement of MuLoG's robustness to spatially-correlated speckle

In this section, we introduce our robustification strategy for a specific sensor and its extension to a range of different sensors.

A. Addressing Speckle Spatial Correlations via Neural Networks

Most radar image providers apply spectral apodization and oversampling to reduce the sidelobes of strong scatterers and adjust the pixel size (see Figure 3), resulting in images corrupted by spatially-correlated speckle noise. Spatial correlation of the speckle leads to severe artifacts with most despeckling algorithms, including the original version of MuLoG (see figure 2b), unless a proper preprocessing step is applied [38].

Exploiting the learning capabilities of neural networks in this context, we propose to account for the spatial correlation of the speckle by considering spatially-correlated Gaussian noise with appropriately adapted statistics during the training process of the Gaussian denoiser destined to be used within MuLoG. The following paragraph describes how to simulate such Gaussian noise.

B. Generation of spatially-correlated Gaussian noise

The formation model of mono-channel SAR images illustrated in Figure 3 can be used to simulate spatially-correlated speckle fields in accordance with the characteristics of a given radar sensor. Speckle is a multiplicative noise: in a

fully developed speckle area, the complex amplitude of the scene can be written as the product of the scene's amplitude $\sqrt{\boldsymbol{r}} \in \mathbb{R}_+^n$ with the complex speckle field $\boldsymbol{s} \in \mathbb{C}^n$: $\sqrt{\boldsymbol{r}} \odot \boldsymbol{s}$, with $\odot$ the element-wise product. The transfer function of the SAR imaging system introduces a spectral apodization that can be modeled as a 2D convolution product with the impulse response $\boldsymbol{h}$, see for example [45]. The intensity $|(\sqrt{\boldsymbol{r}} \odot \boldsymbol{s}) * \boldsymbol{h}|^2$ of the resulting complex image is the quantity designated by the term "radar image" in the right part of Figure 3. As mentioned in Section II-B, the logarithm (or matrix logarithm, in the case of polarimetric SAR images) is applied at the first step of the MuLoG algorithm. It has the effect of making the speckle approximatively additive: $\boldsymbol{y} = \log|(\sqrt{\boldsymbol{r}} \odot \boldsymbol{s}) * \boldsymbol{h}|^2 \approx \boldsymbol{x} + \boldsymbol{\eta}$ where $\boldsymbol{x} = \log|\sqrt{\boldsymbol{r}} * \boldsymbol{h}|^2$ is the component describing the radar scene and $\boldsymbol{\eta} = \log|\boldsymbol{s} * \boldsymbol{h}|^2$ is the spatially-correlated speckle component after the log transformation (this approximate decomposition is using the fact that $\boldsymbol{h}$ is very close to a Dirac to distribute the convolution operator over the element-wise multiplication; it is exact for an ideal SAR response, i.e., when $\boldsymbol{h}$ corresponds to a Dirac distribution).

To adapt the Gaussian denoiser, we propose to train it to remove an additive noise component $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\boldsymbol{\xi}})$, spatially-correlated according to a covariance matrix $\mathbf{C}_{\boldsymbol{\xi}} \in \mathbb{R}^{n \times n}$ that matches the covariance matrix of the log-transformed speckle $\boldsymbol{\eta}$, instead of a white Gaussian noise $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{Id})$.

To simulate this correlated noise, we consider the case of a homogeneous radar scene with a unitary reflectivity ($\boldsymbol{r} = \mathbf{1}$), the log-intensity image then corresponds to a log-transformed speckle: $\boldsymbol{y} = \log|\boldsymbol{s} * \boldsymbol{h}|^2 = \boldsymbol{\eta} \approx \boldsymbol{\xi}$ (the approximation stems from the discrepancy between the Fisher-Tippett distribution that is followed by the log intensities in the absence of spatial correlations and the Gaussian distribution). Let's suppose for now that the impulse response $\boldsymbol{h}$ is known, then in order to simulate Gaussian noises that are statistically as close as possible to $\boldsymbol{y}$, we set $\mathbf{C}_{\boldsymbol{\xi}} = \mathbf{C}_{\boldsymbol{\eta}} := \mathrm{Cov}[\boldsymbol{y}]$ which we estimate from draws of $\boldsymbol{y}$ obtained from simulated white complex speckle fields $\boldsymbol{s}$ colored with $\boldsymbol{h}$. Since the speckle is spatially stationary (the radar imaging system is characterized by its transfer function), the covariance matrix $\mathbf{C}_{\boldsymbol{\eta}}$ is itself stationary and can be characterized in the frequency domain by the power spectral density (PSD) $\boldsymbol{\Psi}_\eta$ (i.e., $\mathbf{C}_{\boldsymbol{\eta}}$ is diagonalizable by the Fourier transform). We apply the periodogram method, consisting of averaging the power spectra of several random draws of $\log|\boldsymbol{s} * \boldsymbol{h}|^2$ in order to estimate a spatial filter $\boldsymbol{g}$ capable of producing the expected spatial correlations:

$$\boldsymbol{g} = \mathcal{F}^{-1}\big[\widehat{\boldsymbol{\Psi}}_\eta^{1/2}\big], \text{ with } \widehat{\boldsymbol{\Psi}}_\eta = \frac{1}{K}\sum_{k=1}^{K}\big|\mathcal{F}\{\log|\boldsymbol{s}_k * \boldsymbol{h}|^2\}\big|^2 \tag{8}$$

and where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the direct and inverse Fourier transforms. Thus, we can draw white noises $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and turn them into correlated noises $\boldsymbol{\xi} = \boldsymbol{g} * \boldsymbol{\epsilon}$ such that $\mathrm{Cov}[\boldsymbol{\xi}] \approx \mathrm{Cov}[\boldsymbol{y}]$. In practice, the closed-form of the impulse response $\boldsymbol{h}$ is often unknown and has to be estimated from the data as detailed in Algorithm 1. This estimation requires only a single mono-channel complex-valued image $\boldsymbol{u}$ from the sensor under consideration. The accuracy of this estimation can be assessed by comparing the Fourier spectrum of $\boldsymbol{u}$ with that of the simulated spatially-correlated speckle noise, as shown in Figure 4. In summary, we have shown that it is possible, in practice, to simulate spatially-correlated Gaussian noise $\boldsymbol{\xi}$ with a covariance matrix $\mathrm{Cov}[\boldsymbol{\xi}]$ that closely matches $\mathrm{Cov}[\boldsymbol{y}]$.

Algorithm 1 Simulation of a spatially correlated Gaussian noise.

Input: a complex-valued mono-channel SAR image $\boldsymbol{u} = (\sqrt{\boldsymbol{r}} \odot \boldsymbol{s}) * \boldsymbol{h} \in \mathbb{C}^{n_1 \times n_2}$

Output: a real-valued spatially-correlated Gaussian noise $\boldsymbol{\xi} \in \mathbb{R}^{n_1 \times n_2}$ such that $\mathrm{Cov}[\boldsymbol{\xi}] = \mathrm{Cov}[\log|\boldsymbol{s} * \boldsymbol{h}|^2]$

1: estimation of the transfer function $\boldsymbol{H} = \mathcal{F}(\boldsymbol{h})$
   1) estimate $\Omega = \Omega_1 \times \Omega_2$ the rectangular support of $\boldsymbol{H}$
   2) $\hat{\boldsymbol{H}}(i,j) = \frac{\mathbb{1}_{\hat{\Omega}}(i,j)}{\mathrm{card}(\hat{\Omega})} \sum_{j' \in \hat{\Omega}_2} |\mathcal{F}(\boldsymbol{u})(i,j')| \sum_{i' \in \hat{\Omega}_1} |\mathcal{F}(\boldsymbol{u})(i',j)|$
2: estimation of the coloring spatial filter $\boldsymbol{g}$
   1) simulate an arbitrary number K of white speckle fields $\boldsymbol{s}_k$
   2) $\boldsymbol{g} = \mathcal{F}^{-1}\big[\widehat{\boldsymbol{\Psi}}_\eta^{1/2}\big]$, with $\widehat{\boldsymbol{\Psi}}_\eta = \frac{1}{K}\sum_{k=1}^{K}\big|\mathcal{F}\{\log|\boldsymbol{s}_k * \mathcal{F}^{-1}(\hat{\boldsymbol{H}})|^2\}\big|^2$
3: generate spatially correlated Gaussian noise $\boldsymbol{\xi}$ as

$$\boldsymbol{\xi} = \boldsymbol{g} * \boldsymbol{\epsilon}, \text{ with } \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

C. Toward a "generic" Gaussian denoiser robust to a variety of spatial correlations

The transfer function involved in the processing of SAR data (see Figure 3) may vary depending on the data provider. Consequently, different noise spatial correlations can be expected for images from different sensors. These variations are illustrated using Canadian sensors RADARSAT-2 and RCM in Figure 5. The figure shows 2D and 3D representations of the auto-correlation function estimated from homogeneous areas of the RADARSAT-2 and RCM images, respectively. In both cases, the region of non-negligible correlation coefficients is constrained to a small neighborhood that we will refer to as the correlation structure of the speckle noise. However, as the correlation structure seems symmetrical for RADARSAT-2 5a, it is more extended in the vertical direction for RCM 5b. A Gaussian denoiser trained for one specific type of noise spatial correlation might not be efficient on data affected by noise with a different spatial correlation. This can be shown by considering the interesting case of the RCM sensor 5b. Because of the non-symmetry of the noise correlation structure in this case, an image and its rotated version by $\pm 90°$ have different noise spatial correlations. Therefore, a Gaussian denoiser trained with Gaussian noise simulated using a transfer function estimated from one orientation may produce artifacts for the other orientation as shown by Figure 6.

To obtain a network that works well on images captured by various sensors, we propose to train a "generic" Gaussian denoiser by considering various spatial correlations during the

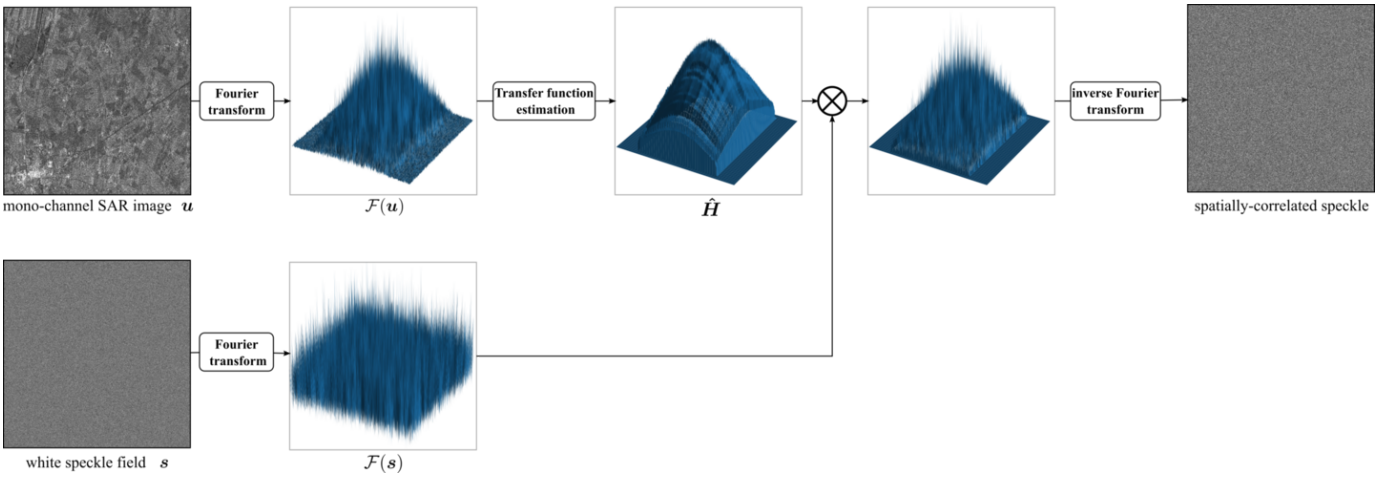
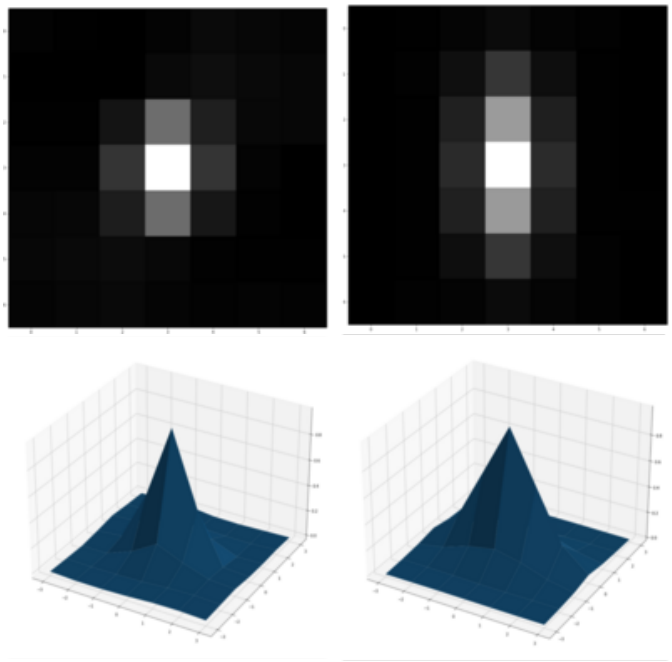


Figure 4: Accuracy of the estimation of the transfer function: the Fourier spectrum of the simulated spatially-correlated speckle noise matches that of the mono-channel SAR image showing a good estimation of the transfer function.



(a) RADARSAT-2 (b) RCM

Figure 5: Auto-correlation function estimates from homogeneous areas of RADARSAT-2 and RCM images: differences in shape of these 2D and 3D representations indicate distinct spatial correlations of the speckle noise.

training step. We replace $\mathbf{\Psi}_\eta$ in Equation 8 with a random convex combination of PSDs each corresponding to a different sensor or acquisition mode.

IV. Network architecture and training of the Gaussian denoiser

The proposed framework to handle spatial correlation is independent of the particular neural network architecture of the Gaussian denoiser. Nevertheless, the more efficient the denoiser, the better the results. In this section, we present the choice made for our experiments and the training details.

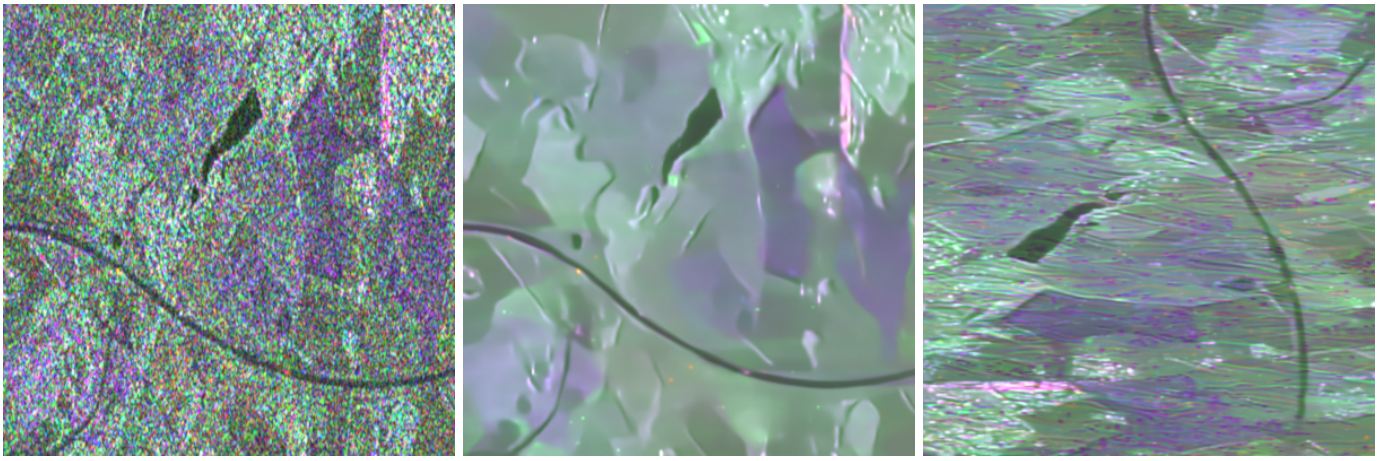
A. DRUNet: A versatile architecture for Plug-and-Play strategies

It is established that the noise level considered in the denoising step of PnP strategies can vary from one iteration to the next. Therefore, we have considered the DRUNet architecture, recently introduced in [46], for its capacity to handle a wide range of noise levels within a single model. For the sake of completeness, we briefly describe its architecture. The DRUNet architecture, shown in Figure 7, integrates elements from three different network architectures. Its ability to handle various noise levels is achieved through an additional input channel that is constant and equal to the noise standard deviation, as introduced by FFDNet [47]. The network's main structure leverages two efficient architectures: UNet [48] and ResNet [49]. UNet is known for its frugal training-set requirements, while ResNet introduces residual blocks for improved performance. Like the UNet, DRUNet consists of a contraction (downscaling) followed by a symmetric expansion (upscaling) but where the convolutional layers (conv+ReLU) are replaced with residual blocks (Id + [conv + ReLU + conv]). The first layer increases the number of channels to 64, and is then doubled after each downsampling operation.

In addition, DRUNet may achieve better results than the architectures it is based on thanks to the additional use of stride convolutions for the contraction phase instead of max-pooling (followed by a non-strided convolution) [50] (see Figure 7). Stride convolutions apply the kernel to every element of the input, whereas max-pooling applies it to a subset of the input. This means stride convolutions preserve more spatial information, increasing the receptive field of neurons in deeper layers of the network. Consequently, the network can learn more complex features.

B. DRUNet training details

In this section, we detail the training process of DRUNet. The network is trained in a supervised manner to remove



(a) noisy RCM PolSAR image (b) MuLoG + NN_S (c) Rotation + MuLoG + NN_S

Figure 6: Limits of using, within MuLoG, a Gaussian denoiser NN_S trained for one specific noise spatial correlation: NN_S is trained to handle specifically the correlation structure of the speckle of the orientation corresponding to 6a, artifacts appear when filtering a 90°-rotated version 6c due to the asymmetry of the speckle correlation structure.

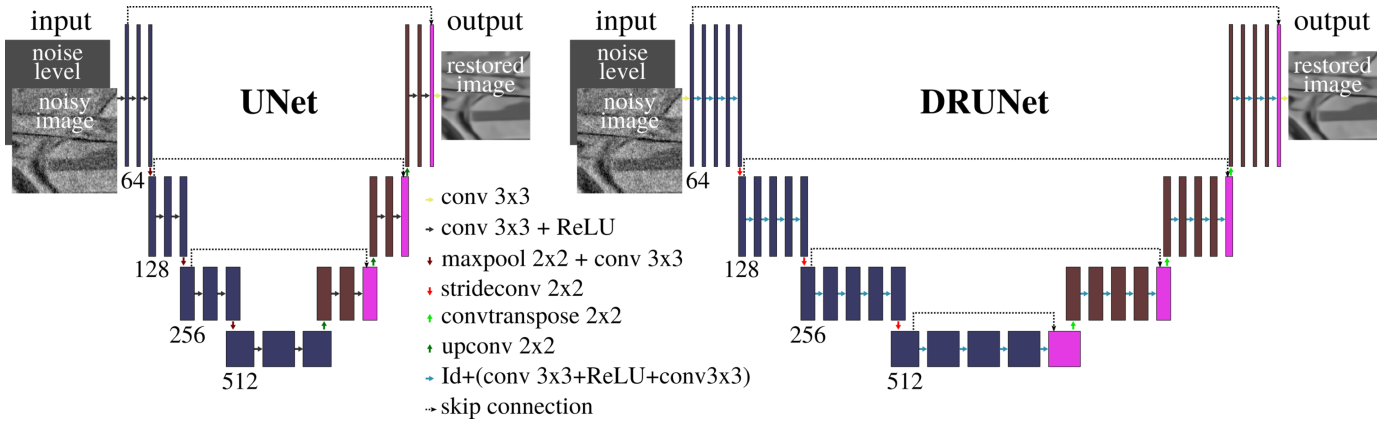


Figure 7: Comparison of UNet and DRUNet architectures.

additive spatially-correlated Gaussian noise with pairs of clean and corrupted images. The clean reference images are made by first averaging a very large time series of Sentinel 1 mono-channel SAR images of the same scene to almost suppress the speckle noise. Then the log-intensity is computed and normalized to $[0, 1]$. The corrupted versions are then built by adding spatially-correlated Gaussian noise with a noise standard deviation chosen randomly in $[0, \frac{50}{255}]$ during training, as explained in [46]. The network input corresponds to a corrupted image and a constant map equal to the noise standard deviation. Training involves minimizing the L_1 loss between the reference image and the log-intensity image estimated by the network from the noisy image, using the ADAM optimizer with mini-batches of 16 patches of size 128×128. We divided our initial dataset into training and validation datasets, and we monitored the loss on the validation dataset and performed an early stopping to avoid overfitting. The process takes about 30 minutes on Pytorch with an Nvidia Tesla A100 GPU.

V. Experiments on polarimetric and interferometric SAR images

Our method is multimodal, but as we focus on SAR polarimetry in this paper, our approach will be mainly validated on that modality. We will refer to the network trained on a unique spatial correlation and the one trained for a family of sensors as DRUNet_S and DRUNet_G, respectively, as explained in section III.C. Section IV.A present the metrics that will be used in the following sections. Section V-B and V-C present the comparison of our approach to existing despeckling methods on simulated and real SLC PolSAR images respectively. The unavailability of network weights for deep-learning-based methods makes it difficult to compare with these methods. Therefore, the following methods are considered: the 5×5 boxcar filter, the 7×7 refined Lee filter [3], and the patch-based filter NL-SAR [7]. Finally, our method is evaluated on real SLC InSAR data in Section V-D and compared with two deep-learning-based methods: Phi-Net [51] and InSAR-MONet [52]. The source code of NL-SAR and, the network architectures and weights of Phi-Net and InSAR-MONet have been made publicly available by the

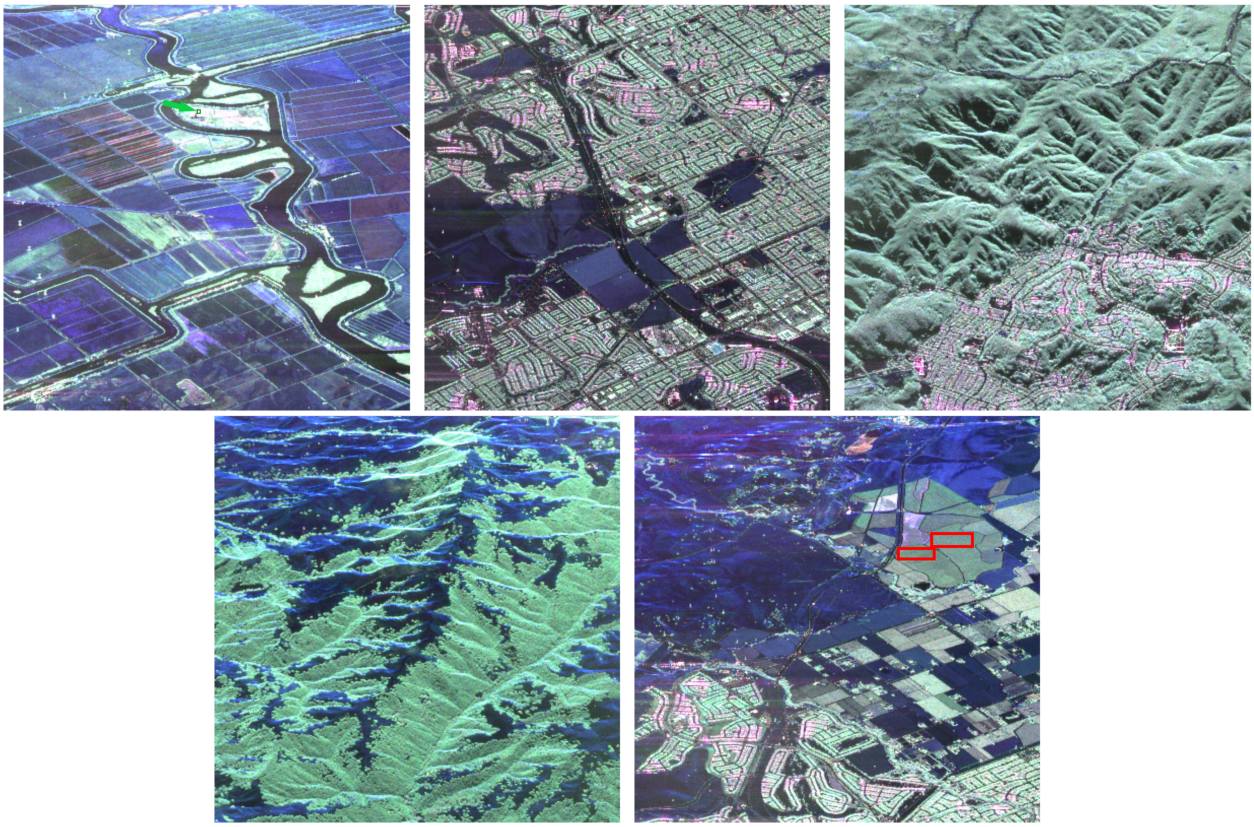


Figure 8: Ground truths computed from UAVSAR data ©NASA/JPL-Caltech. The colored boxes mark the regions where some of the metrics are computed: the green box indicated by the green arrow in the first image for co-pol and cross-pol, and the red boxes in the last image for ENL.

Metric	Boxcar	Refined Lee	NL-SAR	proposed	
				MuLoG+DRUNet$_S$	MuLoG+DRUNet$_G$
MSSIM $\uparrow^1$	0.814±0.0003	0.810±0.01	0.857±0.0005	**0.860**±0.0007	0.857±0.0009
GSIM $\downarrow_0$	0.169±0.0002	0.174±0.009	0.165±0.0004	**0.122**±0.0002	**0.123**±0.0003
ENL $\uparrow$	12.0	15.13	25.7	146.1	**291.8**

Table I: Quantitative evaluation of despeckling methods on images with simulated speckle (best performances are indicated in bold).

authors of the corresponding papers. We used PolSARpro v6.0 implementation of refined Lee filter.

A. Performance assessment metrics

It has been stated that a good speckle filter method should have the following characteristics [53]:

1) strong speckle reduction in homogeneous areas;
2) scene feature preservation (texture, edges,...);
3) radiometric preservation;
4) absence of artifacts.

To evaluate our method on PolSAR data, we use the following criteria:

- Equivalent Number of Looks (ENL): it measures the degree of speckle removal in homogeneous areas and is defined as:

$$\text{ENL} = \frac{\mu^2_{\hat{I}_\mathcal{H}}}{\sigma^2_{\hat{I}_\mathcal{H}}}, \quad \hat{I}_\mathcal{H} = \{\hat{I}_\rho, \rho \in \mathcal{H}\} \tag{9}$$

where $\mathcal{H}$ is the set of pixel indices defining a homogeneous area in image $\hat{I}$, $\mu^2_{\hat{I}_\mathcal{H}}$ and $\sigma^2_{\hat{I}_\mathcal{H}}$ are the empirical mean and the empirical variance of $\hat{I}_\mathcal{Z}$, respectively. A higher value of the ENL indicates a stronger suppression of speckle.

- Mean Structural Similarity Index Measure(MSSIM) [54]: this parameter is given by:

$$\text{MSSIM}(I, \hat{I}) = \frac{1}{n}\sum_{k=1}^{n} \text{SSIM}(I_k, \hat{I}_k) \tag{10}$$

with,

$$\text{SSIM}(I_k, \hat{I}_k) = \frac{(2\mu_{I_k}\mu_{\hat{I}_k} + \kappa_1)(2\sigma_{I_k\hat{I}_k} + \kappa_2)}{(\mu^2_{I_k} + \mu^2_{\hat{I}_k} + \kappa_1)(\sigma^2_{I_k} + \sigma^2_{\hat{I}_k} + \kappa_2)} \tag{11}$$

where n is the number of pixels of I, κ_1 and κ_2 are constants and μ_{I_k}, $\sigma_{I_k\hat{I}_k}$, and $\sigma^2_{I_k}$ are the empirical

expectation, covariance, and variance, respectively, computed locally around pixel k. This criterion takes values between -1 and 1 which are the worst and best values, respectively.

- Global SIMilarity (GSIM) [55]: the global similarity of two n-pixels multi-channel SAR images $\mathbf{\Sigma}, \hat{\mathbf{\Sigma}} \in \mathbb{C}^{n\times d\times d}$ is defined as

$$\text{GSIM}(\Sigma, \hat{\Sigma}) = \frac{1}{nd^2}\sum_{k=1}^{n} \| \log(\mathbf{\Sigma}_k) - \log(\widehat{\mathbf{\Sigma}}_k)\|_F \quad (12)$$

where $\| \cdots \|_F$ corresponds to the Frobenius norm and $\log$ is the matrix logarithm. A lower value of this criterion indicates a higher degree of similarity.

- The Cloude-Pottier polarimetric decomposition parameters (entropy H, anisotropy A, and angle α) [56], [57]: they have been shown to be well-adapted for covariance matrix classification and thus can be used to indicate if covariance matrices within the same class are less scattered in the H-A-α space thanks to the speckle filtering.
- The polarimetric signatures (co-pol and cross-pol) [16], [57], [58]: the polarimetric signature $\text{P} \in \mathbb{R}^+$ of a target represented by the covariance matrix $\mathbf{C}_k \in \mathbb{C}^{d\times d}$ at a pixel k allows to visualize the response of the target under different settings (polarization of the waves, relative position of the target to the illumination source, ...). It is defined as follows:

$$\begin{aligned}\text{P} = &(1 \quad \cos(2\phi_i)\cos(2\psi_i) \quad \cos(2\phi_i)\sin(2\psi_i) \quad \sin(2\phi_i)) \\ &\times \mathbf{K} \times \\ &(1 \quad \cos(2\phi_j)\cos(2\psi_j) \quad \cos(2\phi_j)\sin(2\psi_j) \quad \sin(2\phi_j))^t\end{aligned} \quad (13)$$

where $\mathbf{K}$ is the Kennaugh matrix [59] associated to $\mathbf{C}_k$, the pairs (ϕ_i, ψ_i), and (ϕ_j, ψ_j) of orientation angle $\psi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and ellipticity angle $\phi \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ correspond respectively to the parameters describing the polarization state of the incident and backscattered waves. There is an infinity of possible signatures but the most used in despeckling quality assessment are the co-pol which corresponds to the case where the polarization is invariant in emission and reception i.e $\phi_j = \phi_i$ and $\psi_j = \psi_i$, and the cross-pol which corresponds to the case where the polarization in emission and reception are orthogonal i.e $\phi_j = -\phi_i$ and $\psi_j = \psi_i + \frac{\pi}{2}$.

The first two of our list are defined for intensity images, we extend them to PolSAR images by averaging their values over the three intensity images corresponding to the diagonal values of the covariance matrices. MSSIM, GSIM, and the polarimetric signatures are meant for simulated experiments since the computation of MSSIM and GSIM require a ground truth, and the polarimetric signatures computed from estimated images need to be visually compared to ground truth signatures.

The quantitative analysis on InSAR data is carried out with the number of residues (NoR) [60] commonly used to assess the performances of interferometric phase filters. As its name implies, it totals up the residues in the phase $\Phi \in \mathbb{R}^{m\times p}$. The lower this value, the more effective the filter. The presence of a residue in a pixel (i, j) is indicated by the sum of the phase differences clockwise around the 4-neighborhood of (i, j):

$$\begin{aligned}R = \frac{-1}{2\pi}\Big(&\Delta[\Phi_{i,j+1}, \Phi_{i,j}] + \Delta[\Phi_{i+1,j+1}, \Phi_{i,j+1}] \\ &+ \Delta[\Phi_{i+1,j}, \Phi_{i+1,j+1}] + \Delta[\Phi_{i,j}, \Phi_{i+1,j}]\Big)\end{aligned} \quad (14)$$

with

$$\Delta[\alpha, \beta] = \begin{cases}\alpha - \beta & \text{if } |\alpha - \beta| \leq \pi \\ \alpha - \beta - 2\pi & \text{if } \alpha - \beta > \pi \\ \alpha - \beta + 2\pi & \text{if } \alpha - \beta < -\pi\end{cases} \quad (15)$$

for two adjacent pixels p and q. R can only be equal to 0 or ± 1 indicating the absence of residue, a positive or a negative residue, respectively.

B. Results on simulated PolSAR data

The objective of this section is to demonstrate the effectiveness of our approach in achieving satisfactory results in the presence of spatial correlations. This is accomplished by comparing our method with state-of-the-art despeckling techniques. The experiments have been conducted considering the spatial correlation associated to RADARSAT-2 high-resolution images (12×12 meters). The ground truths were obtained by applying a 7×15 multilooking followed by downsampling to a UAVSAR noisy SLC PolSAR image with 1.8×0.8 meters pixels (respectively in the range and azimuth directions). This multilooking produced almost speckle-free PolSAR images (an equivalent number of looks equal to 22 has been estimated on these images) with a spatial resolution close to RADARSAT-2 images, as can be seen in Figure 8 representing five 400×400 pixel images with a variety of areas including water, fields, mountains, forests, and human constructions.

For each of the five images represented in Figure 8, we simulated 10 independent speckled SLC versions with the following procedure. Starting from the ground truth PolSAR image of covariance matrices $\mathbf{\Sigma} \in \mathbb{C}^{n\times 3\times 3}$, we perform the pixel-wise Cholesky decomposition: at each pixel i, we find the unique lower triangular matrix with real positive diagonal elements matrix $\mathbf{A}_i$ such that $\mathbf{\Sigma}_i = \mathbf{A}_i\mathbf{A}_i{}^*$. Diffusion vectors $\boldsymbol{k}_i \sim \mathcal{N}_c(\mathbf{\Sigma}_i)$ distributed according to Equation (1) can then be obtained by drawing standard circular Gaussian random vectors $\boldsymbol{\epsilon} \sim \mathcal{N}_c(\mathbf{I})$ and multiplying them with matrix $\mathbf{A}_i$: $\boldsymbol{k}_i = \mathbf{A}_i\boldsymbol{\epsilon}$. The SAR transfer function can then be applied to each channel of the simulated complex-valued SLC images $\boldsymbol{k}$, according to the model discussed in Section III-B.

Table I shows the results of the quantitative comparison of the proposed method with the Boxcar filter, the Refined Lee filter, and NL-SAR on the simulated data using the first five metrics of the above list. It represents the global mean of the obtained results and the mean of the corresponding standard deviations computed separately for each ground truth. As it can be seen our method outperforms the others (the results used for this analysis can be viewed in a pdf file shared with the code). The co-pol and cross-pol signatures shown in Figure 9 also indicate that the proposed method restores polarimetric

	Refined Lee			NL-SAR			proposed MuLoG+DRUNet$_G$		
	①	②	③	①	②	③	①	②	③
ENL	7.04	7.71	8	9.19	26.53	14	**22.90**	**62.57**	**84**

Table II: Equivalent number of looks (ENL) reached by reference methods and the proposed polarimetric restoration technique on 3 high-resolution airborne PolSAR images: ① E-SAR image from the DLR, ② F-SAR image from the DLR, ③ SETHI image from the ONERA.

properties close to the reference signatures. Values of the ENL are large ($>$100) with our method, which indicates a strong smoothing in the output images. Only the order of magnitude of the ENL should be compared since the sole maximization of the ENL is not recommended as it captures both positive effects (speckle suppression) and negative effects (over-smoothing and loss of texture/details).

C. Results on real PolSAR data

In this section, we illustrate the efficiency of our method on PolSAR data from different sensors using a unique generic Gaussian denoiser trained by combining the PSD of RADARSAT-2, RCM and UAVSAR. Figure 10 shows that such a generic denoiser leads to results that are comparable to those obtained with a denoiser trained specifically for the PSD of the sensor.

Experiments are also conducted on high-resolution images captured by E-SAR (DLR), F-SAR (DLR), and SETHI (ONERA). The pixel dimensions of the images are 512×512, 1014×1014, and 1024×1024 pixels, respectively. The despeckled images presented in Figure 11 show that speckle fluctuations are reduced by all filtering techniques and that the proposed method performs a stronger smoothing of homogeneous areas while preserving geometrical details. Table II reports the ENL values for each despeckling method. This indicator confirms that our method more strongly smooths homogeneous areas.

Finally, Figure 12 presents 3D scattergrams of the Cloude-Pottier polarimetric decomposition parameters for three distinct homogenous areas of the E-SAR image. we can observe that, as expected from a despeckling method, the distribution of the parameters for the despeckled images 12(b-d) is much more clustered than for the speckled image 12(a). The clusters are even more separable with our method, indicating a better performance compared to NL-SAR or the refined Lee filter.

These results show that our method generalizes well to a large range of sensors and outperforms state-of-the-art methods.

D. Results on real InSAR data

In this section, we validate our approach on InSAR data, illustrating its capacity to handle various multi-channel SAR data. Two SLC SAR images taken in the interferometric configuration with TerraSAR-X sensor are used. An area of 1024×1024 pixels is selected. By computing the covariance matrix in each pixel our method produces jointly an estimation of the two reflectivities, the coherence and the interferometric phase, while Phi-Net estimates coherence and phase, and InSAR-MONet estimates only the interferometric phase. Figure 13 shows the results obtained by our method and the reference methods on this InSAR image. We can notice that the intensity (first row) is well restored by our method without any noticeable artifacts. The proposed method also obtains the best results in terms of phase restoration: the reference methods produce a good estimation of the phase but they seem to be less efficient in low-coherence areas. Our method is the only one that manages to connect the fringes in those areas, as can be seen in the third row of the figure. This observation is confirmed by the number of residues and the residue maps (fifth row): our method produces the lower number of residues. The residues of the other methods are concentrated in the low-coherence areas where their performance is worse.

VI. Discussion and ways of improvement

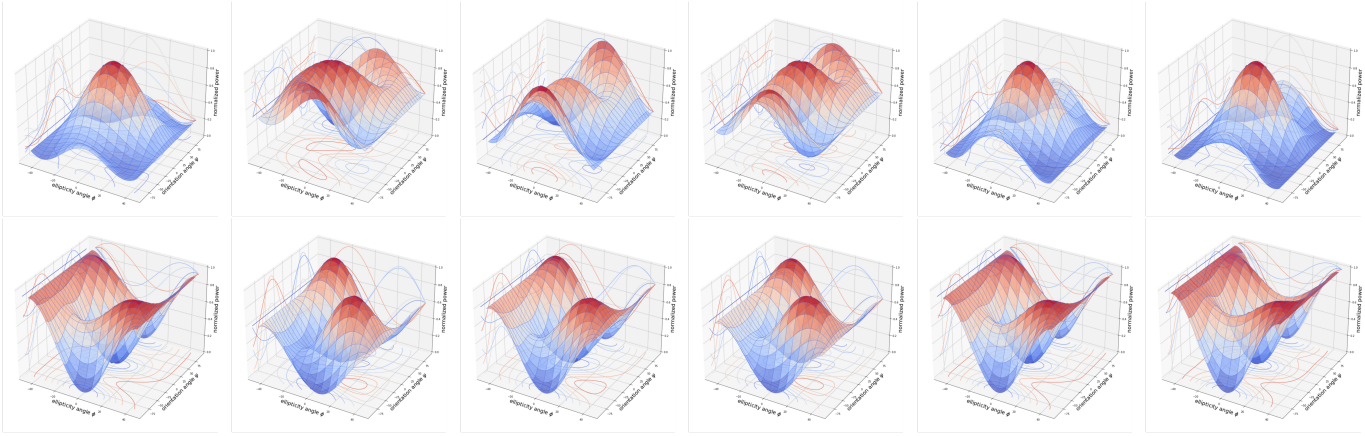
Our exhaustive experiments on simulated and real SAR images corrupted with spatially-correlated noise indicate a good performance of the proposed approach. Although providing the best metrics, the results seem visually over-smoothed in some areas. In this section, we discussed possible areas of improvement of our results.

A. Choice of the network architecture

The PnP methods can benefit from the advancements of Gaussian noise filters. Our approach is a great example of their capacity to evolve. We make use of the learning capacity of CNNs to obtain a Gaussian denoiser robust to spatial correlations encountered in real SAR images allowing to be more effective at filtering those images when using MuLoG. Here, we show that the choice of the architecture has an impact on the final results of MuLoG. We compare the DRUNet architecture with the UNet architecture. For that purpose, we train the UNet following the DRUNet training process described in Section IV-B. The results shown in Figure 14 correspond to the case of RADARSAT-2 images. The spatial correlation is well handled with both architectures but, we can notice that DRUNet tends to produce sharper images than UNet. We can imagine that with an architecture more efficient than DRUNet, the current results can be further improved.

B. Influence of the reference images of the training dataset

The choice of the training dataset is of great importance when developing a deep-learning-based method. The difficulty in accessing noise-free SAR images leads some to use optical images as a substitute to train a neural network for SAR image parameter estimation. In our approach, we train a Gaussian denoiser to be applied individually on the real-valued channels from the log-channel decomposition which corresponds to steps 1 and 2 in the description of MuLoG given in Section II-B. In the case of monochannel SAR images, the log-channel decomposition boils down to the logarithm of the intensity. Figure 15 provides a visualization of the log-channels decomposition of a PolSAR image from RCM: the useful information



(a) Original (b) Boxcar (c) Refined Lee (d) NLSAR (e) MuLoG+DRUNet$_S$ (f) MuLoG+DRUNet$_G$

Figure 9: Co-pol (first row) and cross-pol (second) computed from the average covariance matrix of the 3×3 area of the first image in Figure 8 mark by a green box.

seems to be concentrated in the first channel which is similar to a log-intensity (in terms of content) while the others contain mostly noise. From this observation, valid most of the time, we decided to train our network using noise-free log-intensities as reference images.

The reference images used in the experiments presented in Section V are obtained from a Sentinel 1 image time series as described in Section III-B. A sample of the training dataset is shown in the first two rows of Figure 16. While representing a large variety of scenes, the images of our dataset lack fine structures and texture. This can lead to a model not being able to differentiate fine texture from noise or to restore small structures. We believe that it might be the cause of the over-smoothing observed when using our DRUNet model within MuLoG. To remove any doubt on this matter, we propose to enrich the training dataset with high-resolution images. The additional speckle-free images are obtained by spatial multilooking of submetric resolution images from Capella and Umbra sensors (see the two last rows of Figure 16). A total of 2025 patches of size 128×128 is added to the initial dataset. The DRUNet architecture is trained considering the single correlation case for RADARSAT2 and RCM images.

Figure 17 compares the results of MuLoG with the models from our previous experiments DRUNet$_S$ and with the models trained on the extended dataset DRUNet$_{S+}$. We can notice some improvements in terms of detail preservation with DRUNet$_{S+}$ like in the areas delimited by the red boxes. We also have a better separation of different terrains with this new model (more visible in the second row of the figure). These results are very encouraging as for the capacity of our method to handle noise spatial correlation with a minimal loss of information and we can expect even better results with the addition of high-quality images to the current training dataset.

VII. Conclusion

Due to the difficulty of handling complex-valued covariance matrices, the literature on polarimetric SAR image despeckling is much less extended than for intensity SAR images. By performing a matrix-logarithm and decomposition into real-valued channels that can be denoised independently, MuLoG simplifies this problem. Yet, the statistical model of speckle used as starting point to derive the MuLoG algorithm ignores spatial correlations of speckle that are present in actual SAR images. In this paper, we proposed a solution to this problem that exploits the learning capabilities of deep-learning models in the PnP approach adopted in MuLoG. We proposed to train a deep-learning-based Gaussian denoiser to be robust to spatial correlations. We have shown that it is possible to improve the robustness of our despeckling method to the spatial correlations of a range of sensors with a single neural network. Our approach leads to a versatile method for multi-channel SAR image despeckling. From a practical point of view, it simplifies the training problem of deep neural networks for PolSAR imagery which is reduced to mono-channel filtering (simplifying the constitution of the training set). The downside of this simplified single-channel processing step is the requirement to perform several iterations to produce a despeckled image ($T \times d^2$ single-channel denoising operations for T steps of the ADMM loop on d-channel SAR images, see table III). Despeckling techniques based on a single pass in a feed-forward network can be typically T times faster, but are harder to train and can not readily handle SAR images with different number of channels, e.g. InSAR or PolInSAR. We hope that, by providing the code and network weights of our one-network-fits-all MuLoG approach, deep-learning-based PolInSAR despeckling will be more routinely applied in a wide range of applications of SAR imaging.

Acknowledgments

This project has been partially funded by ASTRAL project (ANR-21-ASTR-0011) and by the grant n°R-S19/OT-0003-086 of the French space agency (CNES).

SETHI images were acquired by ONERA and processed in the framework of ALYS project (ANR-15-ASTR-0002).

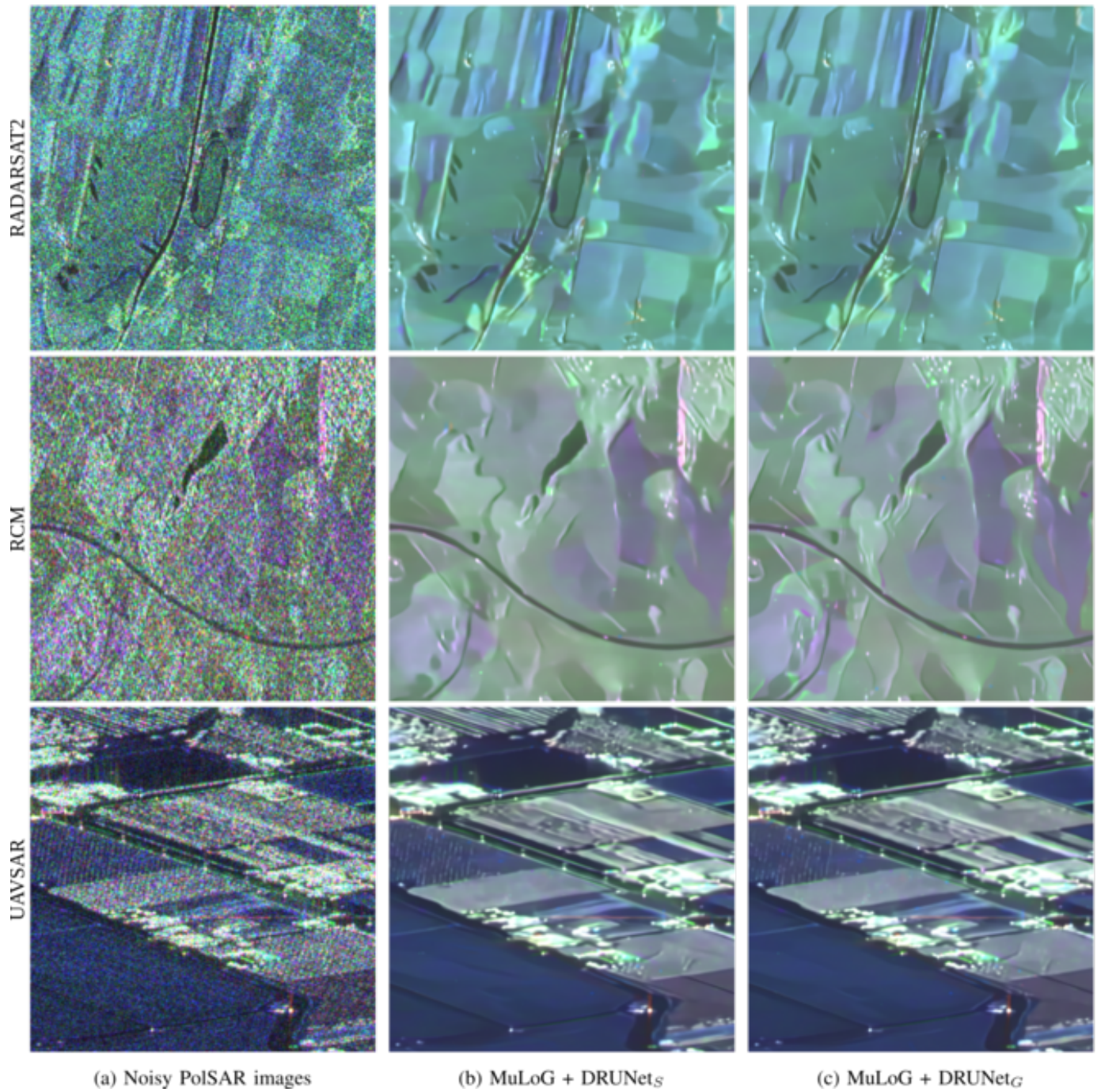


Figure 10: Visual comparison of MuLoG filtering results using DRUNet$_S$ trained on a unique spatial correlation (fixed sensor) and DRUNet$_G$ trained as a generic denoiser for a family of sensors: from top to bottom are shown images acquired with RADARSAT-2, RCM ©Government of Canada (RADARSAT Constellation Mission) and UAVSAR ©NASA/JPL-Caltech sensors.

RADARSAT Constellation Mission images have been provided by the Canadian Space agency, and E-SAR and F-SAR aerial images have been provided by the German Space Agency (DLR).

REFERENCES

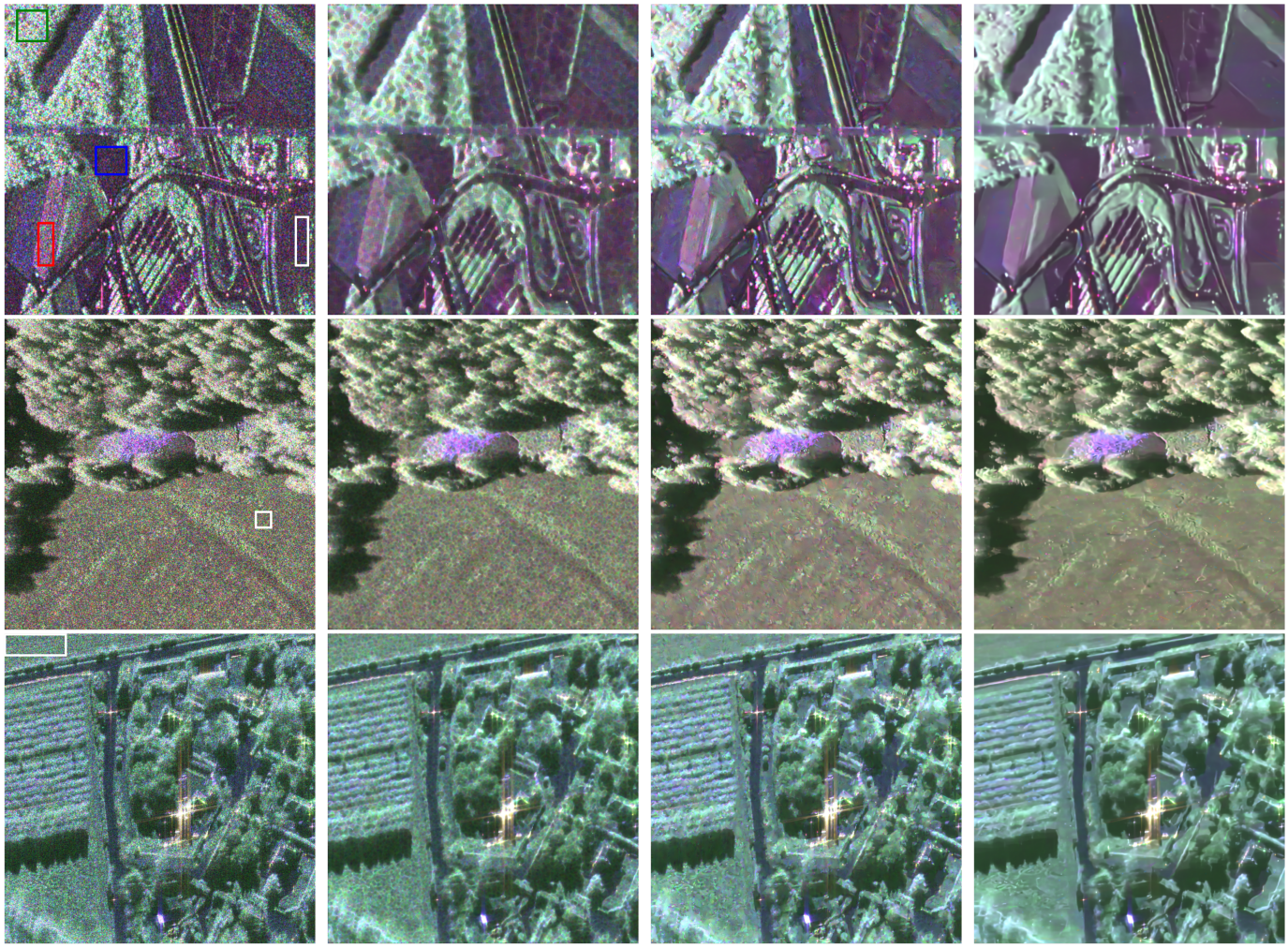
[1] J.-S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*. CRC press, 2017.

[2] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[3] J.-S. Lee, M. Grunes, and G. 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.

[4] G. Vasile, E. Trouvé, J.-S. Lee, and V. Buzuloiu, "Intensity-Driven-Adaptive-Neighborhood Technique for Polarimetric and Interferometric SAR Parameters Estimation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, pp. 1609 – 1621, 07 2006.

[5] B. S. Sahel Mahdavi and B. Brisco, "Speckle filtering of synthetic



(a) Initial images (b) Refined Lee (c) NLSAR (d) MuLoG+DRUNet$_G$

Figure 11: Filtering results of high-resolution PolSAR images acquired by E-SAR ©DLR, F-SAR ©DLR and SETHI ©ONERA sensors. The colored boxes in the left column images show the boundaries of the regions where the metrics are computed: white for ENL and the others for the Cloude-Pottier polarimetric decomposition.

aperture radar images using filters with object-size-adapted windows," *International Journal of Digital Earth*, vol. 11, no. 7, pp. 703–729, 2018.

[6] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: The nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69–78, 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 Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2021–2038, 2015.

[8] J. A. Agersborg, S. N. Anfinsen, and J. U. Jepsen, "Guided Nonlocal Means Estimation of Polarimetric Covariance for Canopy State Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.

[9] X. Nie, H. Qiao, B. Zhang, and X. Huang, "A Nonlocal TV-Based Variational Method for PolSAR Data Speckle Reduction," *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2620–2634, 2016.

[10] Y. Shen, X. Ma, S. Zhu, and J. Xu, "Polarimetric SAR Speckle Filtering Using a Nonlocal Weighted LMMSE Filter," *Sensors*, no. 21, 2021.

[11] H. Aghababaei, G. Ferraioli, S. Vitale, R. Zamani, G. Schirinzi, and V. Pascazio, "Nonlocal Model-Free Denoising Algorithm for Single- and Multichannel SAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[12] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[13] X. Ma, H. Shen, L. Zhang, J. Yang, and H. Zhang, "Adaptive Anisotropic Diffusion Method for Polarimetric SAR Speckle Filtering," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 3, pp. 1041–1050, 2015.

[14] D. Santana-Cedrés, L. Gomez, L. Alvarez, and A. C. Frery, "Despeckling PolSAR Images With a Structure Tensor Filter," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 357–361, 2020.

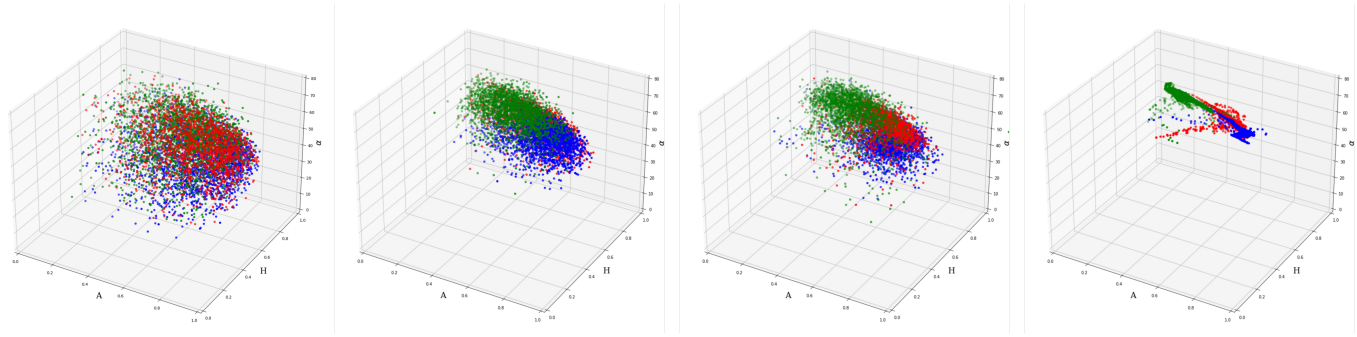
[15] H. Wang, L. Zeng, T. Zhang, J. Yin, and J. Yang, "A PolSAR Despeckling Method based on Wishart Gradient and Anisotropic Diffusion," *Electronics Letters*, vol. PP, pp. 1–3, 01 2021.

[16] X. Nie, H. Qiao, and B. Zhang, "A Variational Model for PolSAR Data Speckle Reduction Based on the Wishart Distribution," *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1209–1222, 2015.

[17] D. Tucker and L. C. Potter, "Speckle Suppression in Multi-Channel Coherent Imaging: A Tractable Bayesian Approach," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1429–1439, 2020.

[18] 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.

[19] G. Fracastoro, E. Magli, G. Poggi, G. Scarpa, D. Valsesia, and L. Verdoliva, "Deep learning methods for synthetic aperture radar image despeckling: An overview of trends and perspectives," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 2, pp. 29–51, 2021.



(a) Boxcar (b) Refined Lee (c) NL-SAR (d) MuLoG+DRUNet$_G$

Figure 12: $3D$ scattergram of the Cloude-Pottier polarimetric decomposition parameters for the three homogeneous regions in the E-SAR image indicated by the red, green and blue boxes.

Image size	d^2	**Gaussian denoiser**	**Time for GD**	**Time for FTP**
512 × 512	1×1	BM3D DRUNet	10.02s 1.32s	0.42s
1024 × 1024	1×1	BM3D DRUNet	41.13s 4.62s	3.03s
512× 512	3×3	BM3D DRUNet	1min 33.88s 8.67s	25.19s
1024× 1024	3×3	BM3D DRUNet	6min 15s 35.58s	2min 1.81s

Table III: Computation time on an 'Intel(R) Core(TM) i5-9300H CPU @ 2.40GHz': Time for GD and Time for FTP correspond to the execution time of the Gaussian denoiser on the $\mathbf{d}^2$ channels and the estimation of the fidelity-term proximal operator, respectively; the second task is performed in parallel on different cores using the separability of the corresponding Equation [18].

[20] B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image restoration for remote sensing: Overview and toolbox," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 2, pp. 201–230, 2021.

[21] E. Dalsasso, L. Denis, M. Muzeau, and F. Tupin, "Self-supervised training strategies for SAR image despeckling with deep neural networks," in *EUSAR 2022; 14th European Conference on Synthetic Aperture Radar*. VDE, 2022, pp. 1–6.

[22] 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.

[23] 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, p. 2636, 2020.

[24] 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, 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/2/196

[25] S. Vitale, G. Ferraioli, A. C. Frery, V. Pascazio, D.-X. Yue, and F. Xu, "SAR Despeckling Using Multiobjective Neural Network Trained With Generic Statistical Samples," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–12, 2023.

[26] A. Mazza, G. Scarpa, L. Verdoliva, and G. Poggi, "Impact of training set design in cnn-based sar image despeckling," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 415–418.

[27] E. Dalsasso, L. Denis, and F. Tupin, "SAR2SAR: A Semi-Supervised Despeckling Algorithm for SAR Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4321–4329, 2020.

[28] X. Ma, C. Wang, Z. Yin, and P. Wu, "SAR Image Despeckling by Noisy Reference-Based Deep Learning Method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8807–8818, 2020.

[29] A. B. Molini, D. Valsesia, G. Fracastoro, and E. Magli, "Speckle2Void: Deep Self-Supervised SAR Despeckling With Blind-Spot Convolutional Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2020.

[30] E. Dalsasso, L. Denis, and F. Tupin, "As if by magic: self-supervised training of deep despeckling networks with MERLIN," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2021.

[31] S. Foucher, M. Beaulieu, F. Cavayas, and M. Dahmane, "Revisiting an Iterative Speckle Filtering Technique," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 5213–5216.

[32] X. Yang, L. Denis, F. Tupin, and W. Yang, "SAR image despeckling using pre-trained convolutional neural network models," in *2019 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 2019, pp. 1–4.

[33] D. Tucker and L. C. Potter, "Polarimetric SAR Despeckling With Convolutional Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[34] A. G. Mullissa, C. Persello, and J. Reiche, "Despeckling Polarimetric SAR Data Using a Multistream Complex-Valued Fully Convolutional Network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[35] J.-S. Lee, "Digital image smoothing and the sigma filter," *Computer vision, graphics, and image processing*, vol. 24, no. 2, pp. 255–269, 1983.

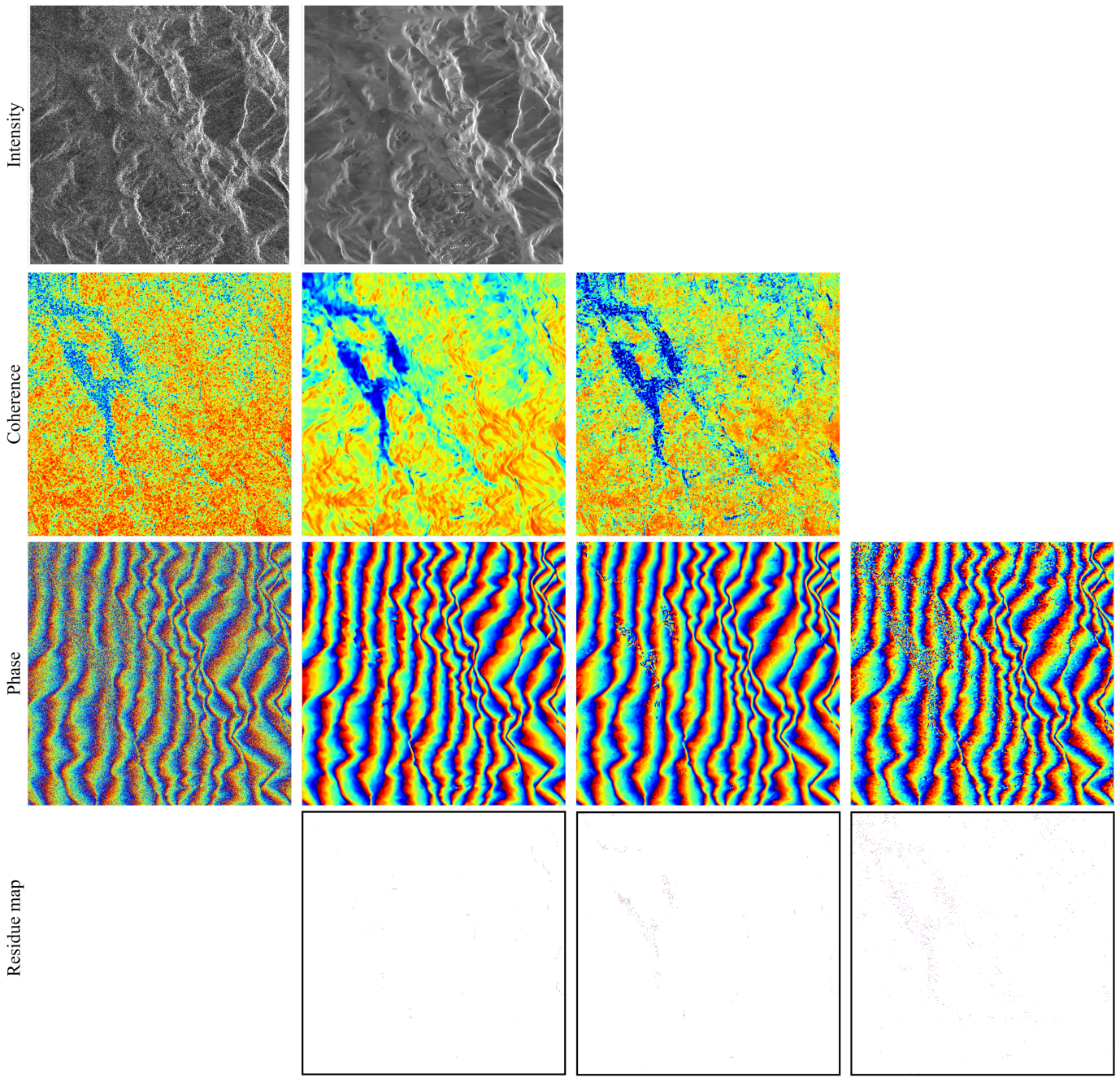
[36] D. Cozzolino, L. Verdoliva, G. Scarpa, and G. Poggi, "Nonlocal CNN SAR image despeckling," *Remote Sensing*, vol. 12, no. 6, p. 1006, 2020.

[37] H. Lin, K. Jin, J. Yin, J. Yang, T. Zhang, F. Xu, and Y.-Q. Jin, "Residual In Residual Scaling Networks for Polarimetric SAR Image Despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–17, 2023.

[38] E. Dalsasso, L. Denis, and F. Tupin, "How to handle spatial correlations in SAR despeckling? Resampling strategies and deep learning approaches," in *EUSAR 2021; 13th European Conference on Synthetic Aperture Radar*. VDE, 2021, pp. 1–6.

[39] A. Lapini, T. Bianchi, F. Argenti, and L. Alparone, "Blind Speckle Decorrelation for SAR Image Despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1044–1058, 2014.

[40] A. Arienzo, F. Argenti, L. Alparone, and M. Gherardelli, "Accurate Despeckling and Estimation of Polarimetric Features by Means of a Spatial Decorrelation of the Noise in Complex PolSAR



(a) Noisy images (b) MuLoG + DRUNet$_G$ (c) Phi-Net (d) InSAR-MONet

Figure 13: Filtering results of a real SLC TerraSAR-X InSAR image. The 'noisy' coherence is computed after a spatial multilooking of the initial image with a 5×5 window. The phase and coherence are shown in false color in $[-\pi, \pi]$ and $[0, 1]$ respectively. The red and blue dots in the residue map correspond to negative and positive residues, respectively.

Data," *Remote Sensing*, vol. 12, no. 2, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/2/331

[41] C.-A. Deledalle, L. Denis, and F. Tupin, "Speckle Reduction in Matrix-Log Domain for Synthetic Aperture Radar Imaging," *Journal of Mathematical Imaging and Vision*, vol. 64, no. 3, p. 298–320, 2022.

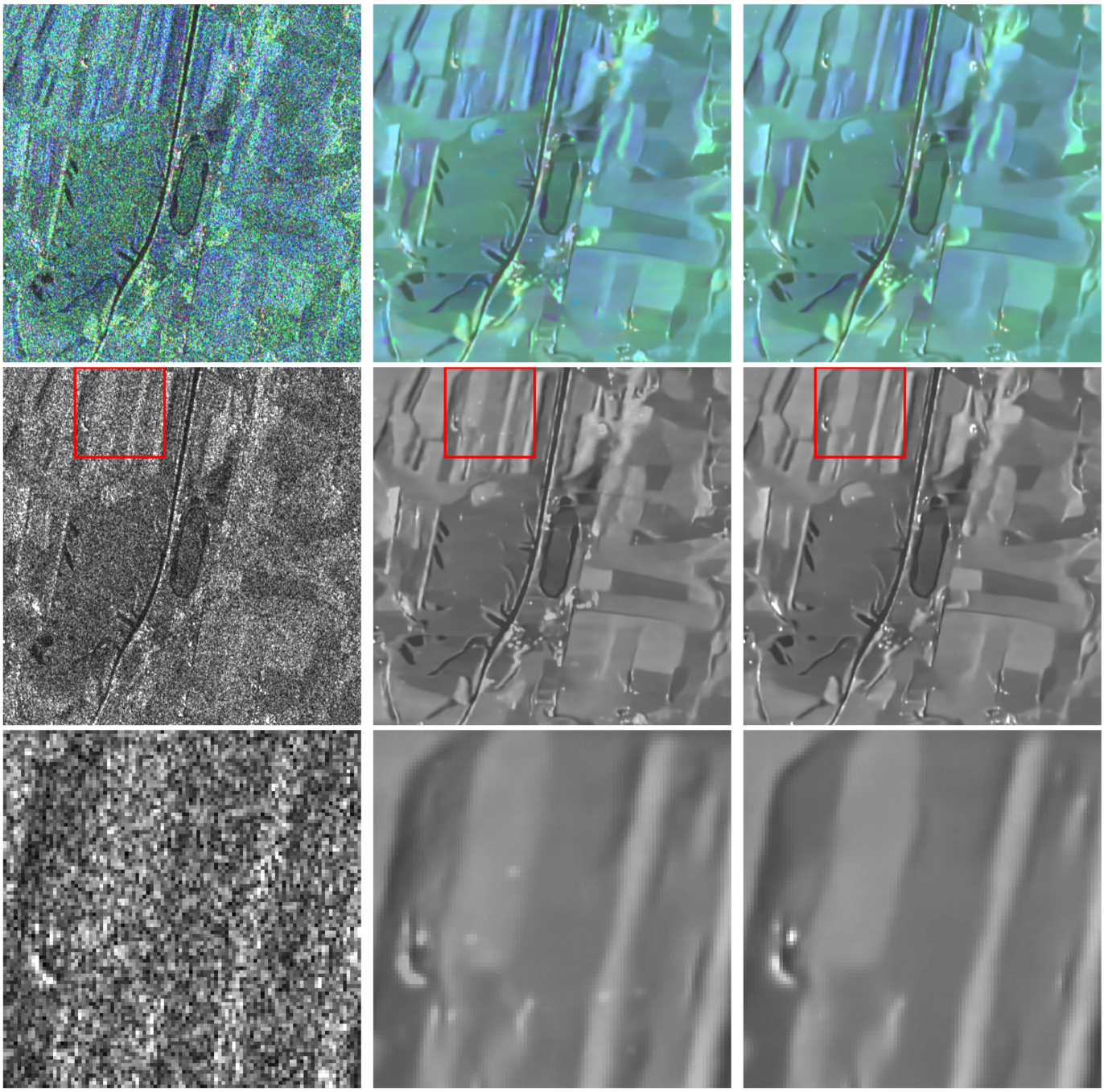
[42] J. W. Goodman, *Laser Speckle and Related Phenomena*. Springer Berlin, Heidelberg, 1984, ch. Statistical properties of laser speckle patterns, pp. 9–75.

[43] N. 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. [Online]. Available: https://doi.org/10.1214/aoms/1177704250

[44] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Multiplicative Noise Removal Using Variable Splitting and Constrained Optimization," *IEEE Transactions on Image Processing*, vol. 19, pp. 1720–1730, 2009.

[45] I. Meraoumia, E. Dalsasso, L. Denis, R. Abergel, and F. Tupin, "Multitemporal Speckle Reduction With Self-Supervised Deep Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.

[46] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Gool, and R. Timofte, "Plug-and-



(a) Initial RADARSAT-2 PolSAR image (b) MuLoG + UNet (c) MuLoG + DRUNet

Figure 14: Comparison of UNet and DRUNet architectures: filtering results of a single-look RADARSAT-2 polarimetric image with each architecture. First row: Pauli RGB representation of covariance matrix images, second row: intensity of the HH channel, last row: close up views of the area delineated by the red squares.

Play Image Restoration With Deep Denoiser Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, pp. 1–1, 06 2021.

[47] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, sep 2018. [Online]. Available: https://doi.org/10.1109/tip.2018.2839891

[48] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2015.

[50] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. A. Riedmiller, "Striving for Simplicity: The All Convolutional Net," *in ICLR (workshop)*, vol. abs/1412.6806, 2014.

[51] F. Sica, G. Gobbi, P. Rizzoli, and L. Bruzzone, "Φ-Net: Deep Residual Learning for InSAR Parameters Estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 3917–3941, 2021.

[52] S. Vitale, G. Ferraioli, V. Pascazio, and G. Schirinzi, "InSAR-MONet: Interferometric SAR Phase Denoising Using a Multiobjective Neural

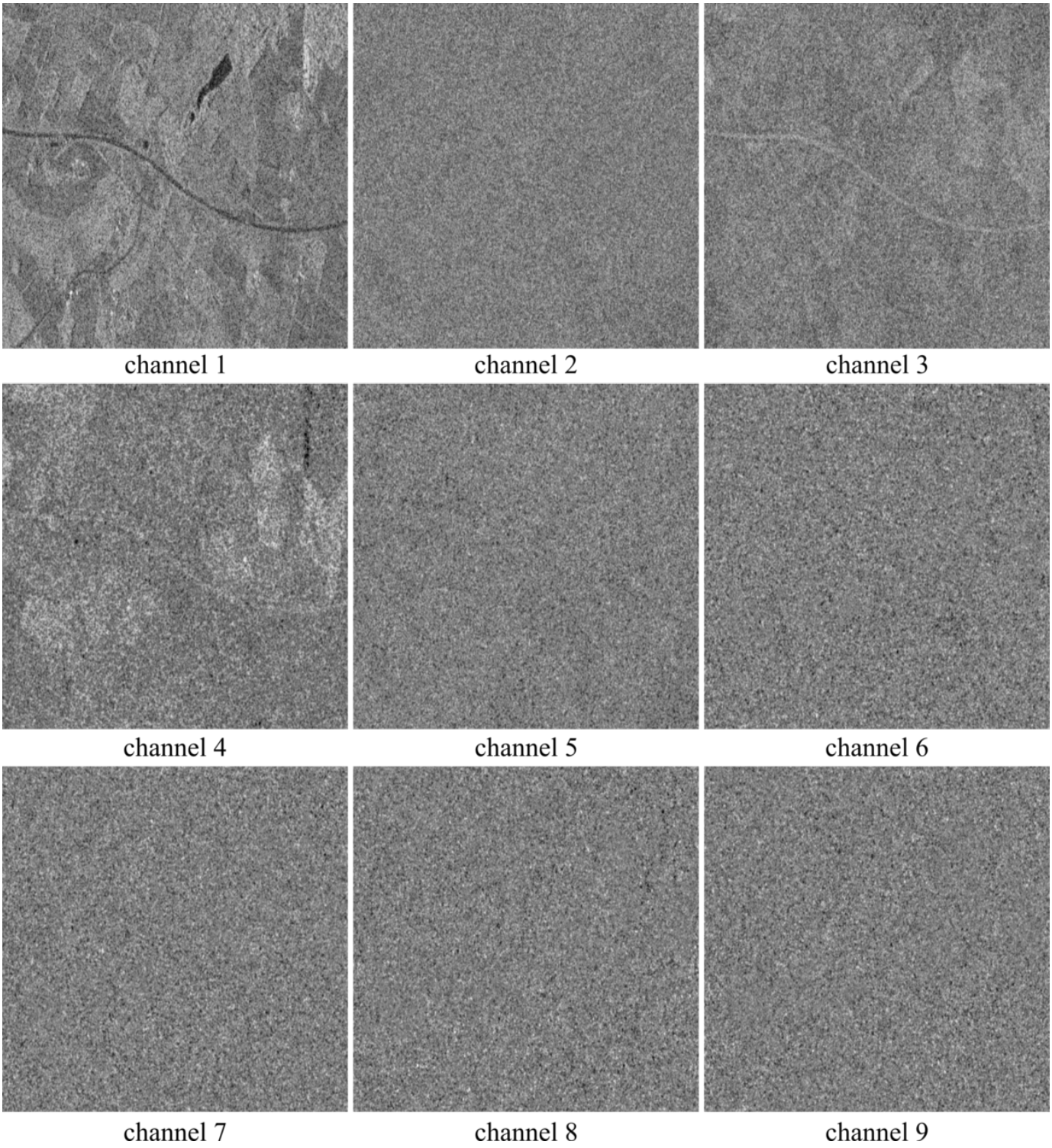


Figure 15: Log-channels decomposition: input of the PnP-ADMM algorithm in MuLoG.

Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[53] G. Di Martino, M. Poderico, G. Poggi, D. Riccio, and L. Verdoliva, "Benchmarking Framework for SAR Despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1596–1615, 2014.

[54] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[55] J. Amao-Oliva, D. Torres-Román, I. Yanez-Vargas, A. Reigber, and M. Jäger, "The Beltrami SAR Framework for Multichannel Despeckling," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2989–3003, 2019.

[56] S. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 68–78, 1997.

[57] X. Ma, P. Wu, Y. Wu, and H. Shen, "A Review on Recent Developments in Fully Polarimetric SAR Image Despeckling," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 743–758, 2018.

[58] D. Tucker and L. C. Potter, "Polarimetric SAR Despeckling With

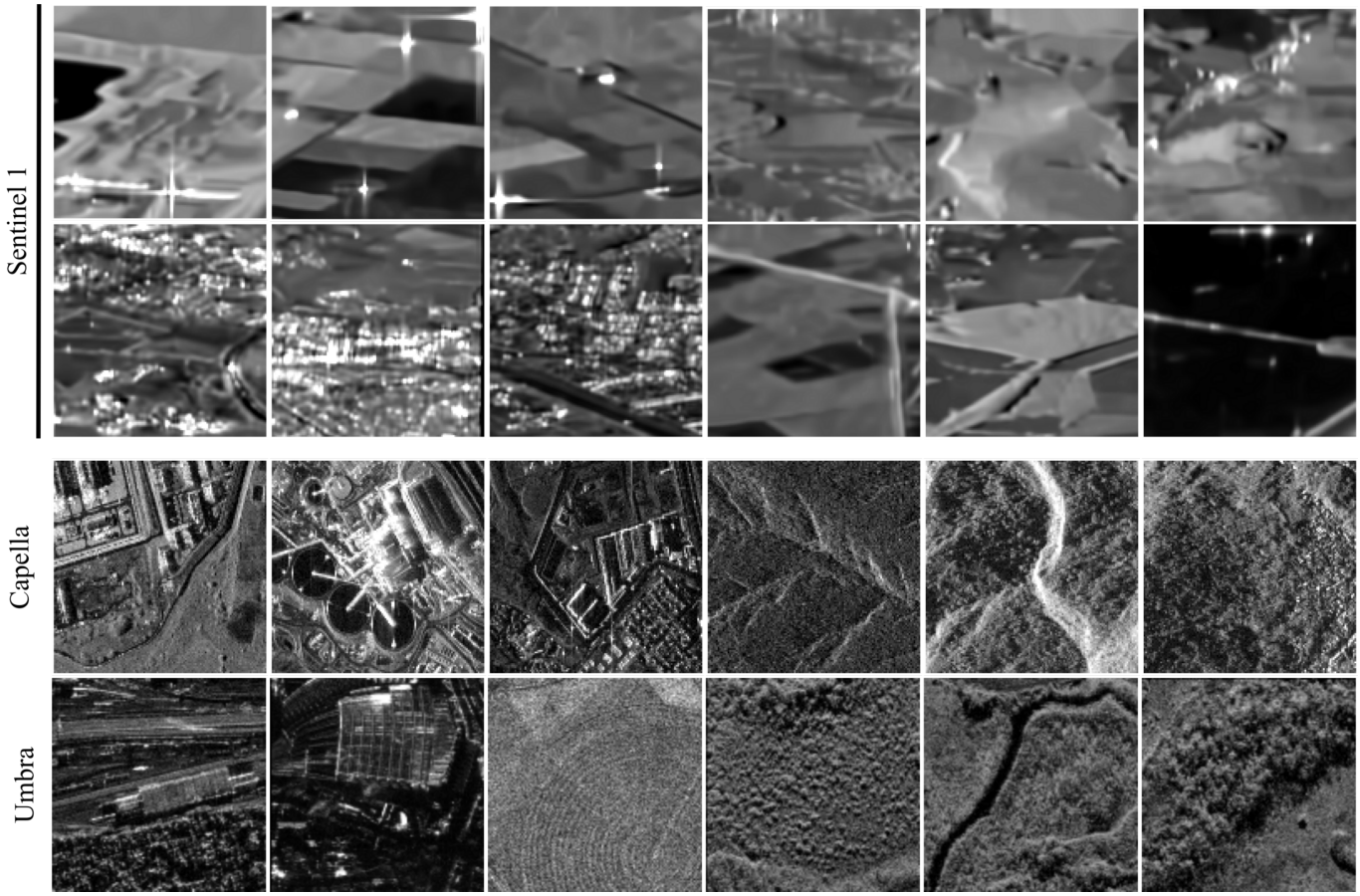
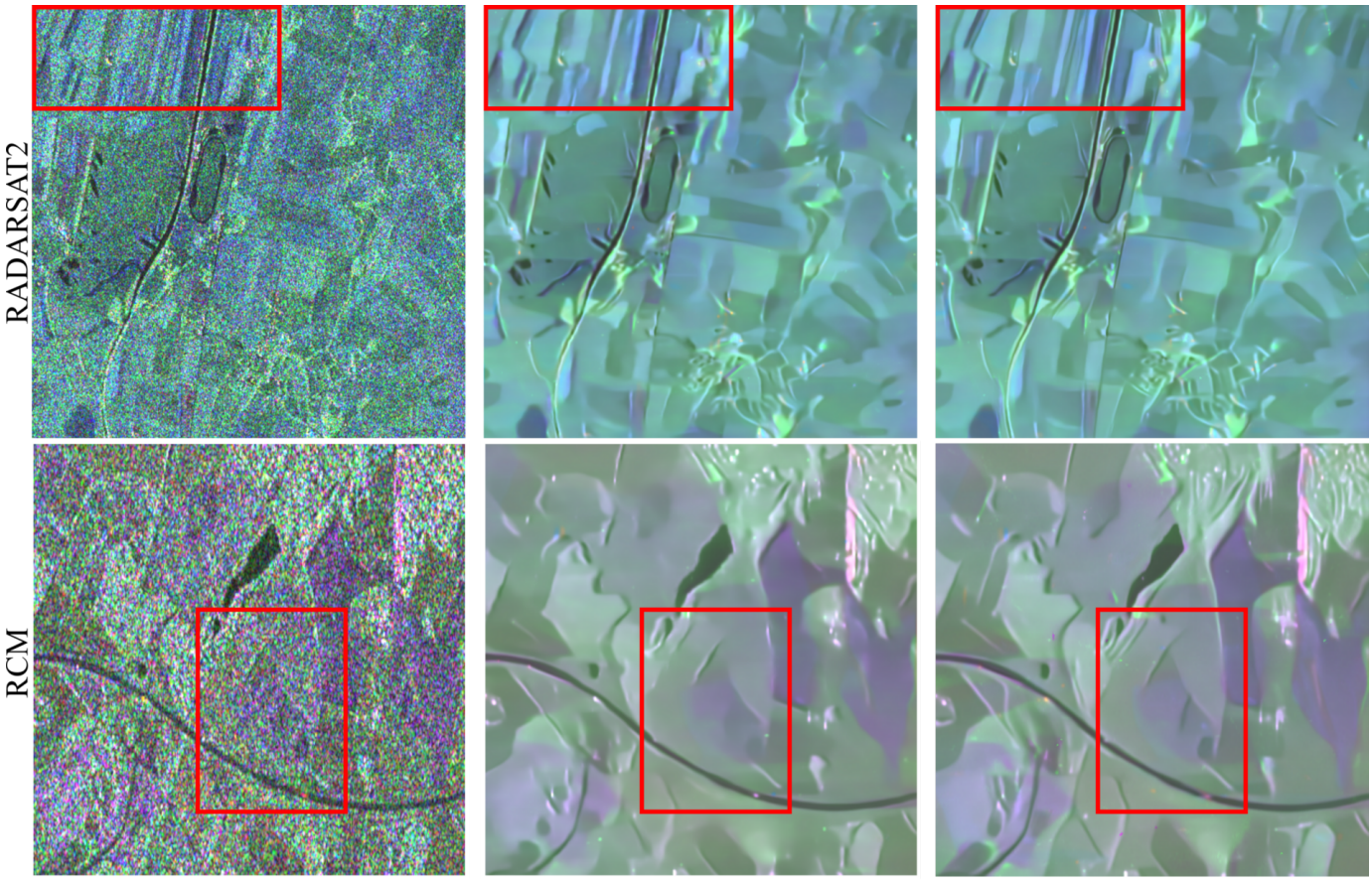


Figure 16: Sample reference images of our training dataset: the images are of size 128×128 pixels.

Convolutional Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[59] B. You, J. Yang, J. Yin, and B. Xu, "Decomposition of the Kennaugh Matrix Based on a New Norm," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, 04 2014.

[60] R. Goldstein, H. Zebker, and C. Werner, "Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping," *Radio Science*, vol. 23, 09 1988.



(a) Noisy PolSAR images (b) MuLoG + DRUNet$_S$ (c) MuLoG + DRUNet$_S$+

Figure 17: Impact of the training dataset of the Gaussian denoiser on the final results of MuLoG: the red boxes highlight, for each image, one area with significant improvements.