Automatic simulation of SAR images: comparing a deep-learning based method to a hybrid method

Nathan Letheule, Flora Weissgerber, Sylvain Lobry, Elise Colin

AUTOMATIC SIMULATION OF SAR IMAGES: COMPARING A DEEP-LEARNING BASED METHOD TO A HYBRID METHOD.

Nathan Letheule[1,2], *Flora Weissgerber*[1], *Sylvain Lobry*[2], *Elise Colin*[1]

[1]ONERA, DTIS-Lab, Université de Paris Saclay
[2]LIPADE, Université de Paris, Paris, France

ABSTRACT

This study compares two approaches for simulating synthetic aperture radar (SAR) images. The first approach uses a conditional Generative Adversarial Network (cGAN) to learn statistical image distributions from optical images. In a second approach, we generate SAR images using a electromagnetic simulator taking into input material maps obtained by segmenting optical images. We propose two metrics to evaluate the quality of the simulation. We evaluate the methods on existing Sentinel-1 SAR images of France using the DREAM database. The results suggest that the physical simulator with automatically created material maps is better suited for generating realistic SAR images compared to the cGAN approach, even if a lot of work remains to be done on the complexity of the description of the scene.

***Index Terms*—** Simulation, Radar, Deep Learning, Remote sensing, Semantic segmentation

1. INTRODUCTION

Radar images have diverse applications in civil, industrial, and defense sectors, such as land surveillance and mapping. Simulating radar images is crucial for generating large quantities of SAR images for deep learning, evaluating algorithmic performance, and meeting specific acquisition conditions.

Simulation tools face significant challenges in balancing computational complexity, fidelity to physics, and image size. Several tools exist for solving Maxwell's equations, such as RaySAR [1] which uses ray tracing to account for topography and interactions, CohRaS [2] which utilizes coherence simulation, and EMPRISE (www.emprise-em.fr) developed in Onera [3]. EMPRISE is capable of generating high-resolution images over large areas and in various operational scenarios. This capability is achieved through a combination of two modules. The first module accurately describes the Radar Cross Section (RCS) of objects by utilizing their detailed 3D mesh and material properties. The second module generates speckle zones by employing curves that describe the variation of σ_0 with respect to incidence angle and polarization. However, this simulation tool faces two significant challenges: ensuring the electromagnetic model is accurately represented in the resulting image, and properly describing the input data for the simulator. In the case of the EMPRISE simulator, the input is typically a manual annotation of the scene from an optical image. This study aims to replace this process with an automatic algorithm and evaluate the quality of the resulting simulation product.

We propose two approaches to automatically simulate SAR images, depicted in Figure 1. The first one uses a conditional generative adversarial network (cGAN) architecture to simulate a SAR image from its optical equivalent, without taking into account the physics. The second approach involves automatic segmentation of an optical image to generate a scene description, which can be used within our physical simulation tool EMPRISE.

In section 2, we describe the methods used for the simulation and the performance evaluation, including the implementation of the metrics. In section 3, we present the main results of our approaches before concluding.

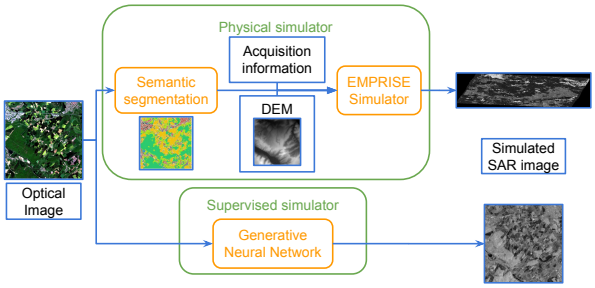

Fig. 1. Block diagram of the two SAR image simulation methods. Below: in GRD geometry with cGan; Up: in SLC geometry with EMPRISE

2. METHODS

2.1. Simulation Frameworks

2.1.1. Fully supervised simulation

To generate a SAR image from an optical image, we train a cGAN using a Pix2Pix architecture [4]. The cGAN incorporates an additional condition, such as a reference image or la-

bel, to guide the generation process. Specifically, we use the Pix2Pix architecture [4] to learn the mapping between optical and SAR domains, resulting in realistic SAR images from corresponding optical images.

2.1.2. EMPRISE-based simulation

Our second approach, based on hybrid simulation, utilizes the EMPRISE simulator which requires a material map as input for scene simulation. To generate the material map, we propose assigning a label to each pixel in an image through semantic segmentation. In recent years, deep learning, particularly convolutional neural networks (CNNs), has emerged as the dominant paradigm for image semantic segmentation [5]. U-Net [6], a CNN specifically designed for biomedical image segmentation, has become a popular choice for this task. We train a U-Net-based neural network on both the optical image and OpenStreetMap (OSM) data rasterized in the DREAM database and used as classification labels. Once the segmentation map **S** is obtained, it is used as input for automatically simulating SAR images with the EMPRISE simulator. Each label of **S** is assigned a backscatter curve for each class type, which is parameterized by the local incidence angle and acquisition frequency. This allows us to generate simulated SAR images with accurate backscatter properties for each labeled class in the material map.

2.2. Evaluation Frameworks

To evaluate the quality of simulated results, we propose dedicated metrics considering the specific statistical properties of SAR. Previous literature commonly uses metrics like the structural similarity index measure (SSIM) [4] and mean square error (MSE) for image assessment. However, it is important to note that speckle can cause MSE variations even when comparing different real images of the same area. Similarly, the SSIM may not accurately distinguish between texture and statistical differences in images. In this study, we propose two new metrics designed for SAR images to address limitations in previous approaches. These metrics offer a comprehensive evaluation of SAR simulations.

2.2.1. Speckle statistics

The first metric focuses on analyzing the statistical properties of speckle. When speckle satisfies Goodman's model assumptions [7], the second and third order log-cumulants become independent of the first moment of the speckle distribution. To compare the micro-texture properties of two speckle realizations, regardless of their mean values, we propose to compute the Bhattacharyya distance between the bivariate distribution of the log-cumulants κ_2 and κ_3 [8] and we call this metrics **BDLC**. The distribution is estimated through the joint histogram of these log-cumulants. We recommend estimating the log-cumulant of the intensity distributions over 5x5 pixel estimation windows. To ensure that the estimated Bhattacharyya distance is as independent as possible from content of the image, we propose to estimate the histogram on the support of $[0, 5\sigma_{\kappa_2,L}]$ and $[\mu_{\kappa_3,L}-9\sigma_{\kappa_3,L}, \mu_{\kappa_3,L}+9\sigma_{\kappa_3,L}]$ where $\mu_{\kappa_i,L}$ and σ_{κ_i} are respectively the theoretical mean and theoretical standard deviation for $\kappa_i (i = 2, 3)$ of a gamma distribution. The number of bins required for histogram discretization is chosen to be equal to the square root of the number of pixels in the considered area. We found that an adequate image area size for calculating this criterion is approximately 90,000 pixels. With these parameters, the Bhattacharyya distance established on these two statistical parameters can effectively distinguish different distributions such as gamma distributions with different equivalent number of looks L or distinguish gamma distribution with log-normal and inverse gamma.

2.2.2. Bright scatterer detection and matching

We introduce our second evaluation process, which aims to identify bright scatterers in the scene. These scatterers are characterized by their high reflectivity, typically resulting from the material or geometric structure of objects. Our approach involves comparing the presence of these bright scatterers in both real and simulated images. To achieve this, we propose a multiscale detection method inspired by the work of Armand and Lopes [9]. This method entails computing intensity averages over central and peripheral neighborhoods. By employing various neighborhood sizes, we are able to detect bright scatterers at different scales. For each detection involving a set of related pixels, we retain only the pixel with the highest value, resulting in a refined final detection.

We introduce three metrics for comparing the sets of points extracted from both the real and simulated images. These metrics include the Chamfer distance [10], the recall comparing at a fixed matching distance of N pixels with the true image as the reference, called **RBMP** for *Recall-Based Matching Performance*, and the minimum matching distance required to attain a recall score of 0.9 with the simulated image as the reference called **90RD**.

The Chamfer distance can be interpreted as the average distance between two point clouds. When the points in the simulated image closely match those in the real image, the Chamfer distance will be small. Conversely, if no bright points are detected in the simulated image, the distance will be infinite, indicating a lack of correspondence between the two point clouds.

The recall measures the ability of our simulator to accurately locate and match real points in the simulated image. It calculates the percentage of real points that have a corresponding point in the simulated image within a certain distance threshold, fixed here to N=5 pixels. A high recall indicates accurate localization and matching of the simulated image. The distance to a fixed recall score, assesses the abil-

ity of our simulator to reproduce the points present in the real image. It returns the maximum distance threshold required to match 90% of the bright scatterers in the simulated image to the points in the real image. A smaller distance indicates a closer resemblance between the simulated points and the real image. If the density of bright scatterers in the real image deviates significantly from that of the simulated images, the first and third metrics will reflect this discrepancy by increasing their respective values.

3. RESULTS

3.1. Our global Approach

We evaluate the performance of our two simulation approaches on Sentinel-1 images, using a Sentinel-2 optical image as input. However, comparing the results is not straightforward due to differences in simulation geometries. Specifically, for the cGAN approach, we simulate images in the terrain-corrected Ground Range Detected (GRD) geometry, which allows for a direct transfer from the optical images. On the other hand, EMPRISE generates images in the antenna reference frame, resulting in Single Look Complex (SLC) products. Despite these differences, we analyze the simulation results using our proposed metrics, as shown in Figure 3.

In both of our approaches, we rely on the DREAM database [11]. DREAM is a multimodal database containing optical images from Sentinel-2, material map composed of 7 different labels extracted from OSM, and SAR images from Sentinel-1 acquired within less than 10 days apart from the optical images. To validate the hybrid simulations, we also download Single Look Complex (SLC) images corresponding to the same products than the initial GRD images.

Figure 2 consists of a real Sentinel-1 image on the left, a cGAN-simulated image on the right, and an EMPRISE-simulated image in the center after automated segmentation and rescaled to the same squared grid for consistency. In a qualitative assessment, the overall results appear to be more realistic with cGAN. However, our metrics will provide a more accurate and detailed evaluation of these simulations.

3.2. Accuracy of statistics

Taking into account the entire image footprint, both the cGAN and EMPRISE approaches closely approximate the speckle distributions observed in real images. However, at the local level, this is no longer the case for cGAN.

Specifically, in the case of EMPRISE, the log-cumulant diagrams exhibit remarkable similarity between the real and simulated data. Conversely, cGAN fails to adequately capture the microtexture. In Figure 3, the example diagram clearly illustrates that the points in the log-cumulants plane for cGAN are considerably less dispersed around the theoretical value, indicating a notable deviation from the desired representation.

Simu \ Metrics	BDLC	Chamfer	90RD	RBMP
EMPRISE	0.013	16.86	43.57	24.27%
cGAN	1.81	inf	inf	0%

Table 1. Evaluation metrics for simulation results

Furthermore, a complementary parametric analysis of EMPRISE indicates that the fidelity of the simulation depends much more on the number of materials used, and especially on the attribution of their electromagnetic properties, than on the performance of the U-Net segmentation.

3.3. Accuracy of bright points

Both types of simulators struggle to retrieve bright scatterers. cGAN fails to generate any bright point. For this simulation scenario, the Chamfer distance thus is infinite. This issue may be attributed to the convolutional filters used in conjunction with the loss functions, which tend to smooth the results at finer scales, hindering the representation of locally higher intensities. The EMPRISE approach performs slightly better with a Chamfer distance of 21 pixels. It produces an excess of bright points, due to the selection of a material for urban areas with inadapted directivity. The attribution of electromagnetic behaviors to labels plays a critical role in achieving an adequate number of bright points. More accurate modeling would require either the use of more materials or the use of object with RCS computed from canonical mechanisms.

The table 1 gathers all the criteria computed on the relative test dataset for the U-net and cGAN learning methods. It should be noted that it is not possible to compare the **BDLC** criterion in this situation, as it depends on the equivalent number of looks, which differs in our two cases (L=4.4 for the GRD data of cGAN, and L=1 for the SLC data of EMPRISE).

4. CONCLUSION

In conclusion, two approaches are proposed and tested to simulate SAR images from optical data. The first approach uses a cGAN while the second approach uses a physical simulator based on electromagnetic simulation of segmented areas and the 3D geometry of the terrain. We propose metrics to evaluate the simulation results, showing that EMPRISE is better at reproducing local statistics, while both approaches struggle to retrieve bright scatterers. Future progress should focus on assigning materials to segmented areas through supervision and adding specific 3D object modeling, rather than improving segmentation accuracy. Diffusion networks could be used to guide physical simulation by distributing these specific objects, such as vehicles or urban details, with realistic distributions.

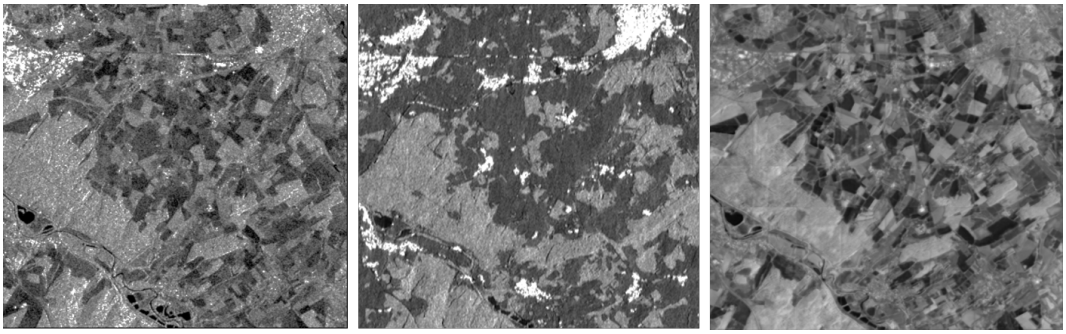

Fig. 2. Left: initial real Sentinel-1 image. Middle: EMPRISE-simulation approach. Right: cGAN approach

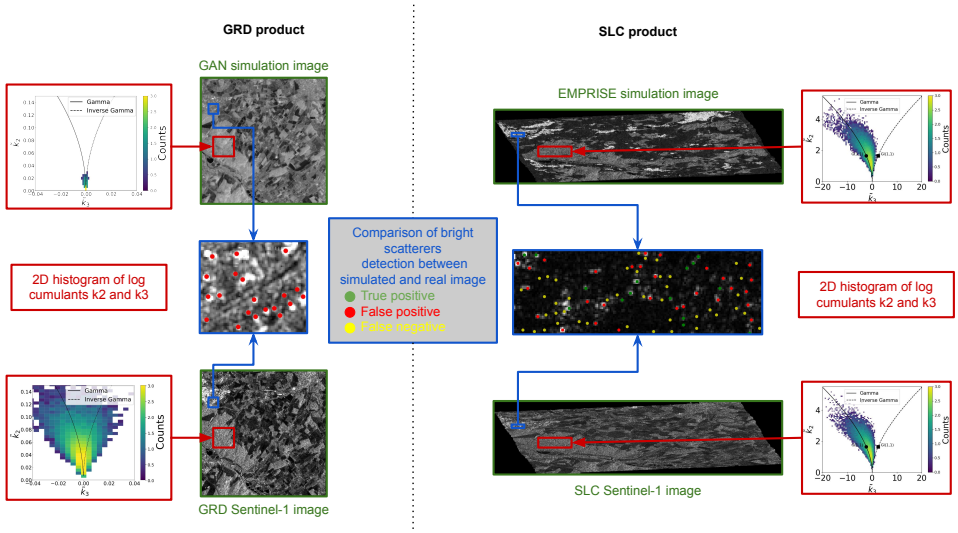

Fig. 3. Metric results on simulated (on the top) and real (on the bottom) SLC (on the right) and GRD (on the left) images. Display of 2D histograms of κ_2 and κ_3 cumulants of a forest area (red) on the four images as well as bright scatterers detection on a city area (blue).

5. REFERENCES

[1] S. Auer and al., "RaySAR - 3D SAR simulator: Now open source," in *IGARSS*, 2016.

[2] H. Hammer and al., "Coherent simulation of SAR images," in *Image and Signal Processing for Remote Sensing XV*. 2009, SPIE.

[3] C. Lebarbu and al., "Complete radar simulation chain: Application to maritime patrol surveillance using SAR/ISAR modes," in *EUSAR*, 2021.

[4] R. Dietrich-Sussner and al., "Synthetic glacier SAR image generation from arbitrary masks using pix2pix algorithm," *IGARSS*, 2021.

[5] I. Kotaridis and al., "Remote sensing image segmentation advances: A meta-analysis," *ISPRS*, 2021.

[6] O. Ronneberger and al., "U-net: Convolutional networks for biomedical image segmentation," in *ICMICCAI*. Springer, 2015.

[7] J. W Goodman, "Some fundamental properties of speckle," *JOSA*, vol. 66, no. 11, 1976.

[8] J-M. Nicolas and al., "Statistical models for sar amplitude data: a unified vision through mellin transform and meijer functions," in *EUSIPCO*. IEEE, 2016.

[9] A. Lopes and al., "Optimal target detection using one channel sar complex imagery: application to ship detection," in *IGARSS*, 1998, vol. 2.

[10] T. Wu and al., "Density-aware chamfer distance as a comprehensive metric for point cloud completion," in *NeurIPS*, 2021.

[11] E. Colin and al., "The DREAM database: a multimode database including optics, radar, DSM (SRTM) and OSM labels for deep machine learning purposes," in *IGARSS 2021*, Bruxelles, Belgium, Dec. 2021.