Self-supervised SAR anomaly detection guided with RX detector

Max Muzeau, Chengfang Ren, Sébastien Angelliaume, Mihai Datcu, Jean-Philippe Ovarlez

SELF-SUPERVISED SAR ANOMALY DETECTION GUIDED WITH RX DETECTOR

M. Muzeau[1,2] *C. Ren*[1] *S. Angelliaume*[2] *M. Datcu*[3] *J.-P. Ovarlez*[1,2]

[1] SONDRA, CentraleSupélec, Université Paris-Saclay, 91192 Gif-sur-Yvette, France
[2] DEMR, ONERA, Université Paris-Saclay, 91120 Palaiseau, France
[3] University Politehnica of Bucharest (UPB), Romania and German Aerospace Center (DLR), Germany

ABSTRACT

Anomaly detection in Synthetic Aperture Radar (SAR) images is an important topic. However, the task is challenging due to the scarcity of anomalous samples and the lack of annotated data, which has led most algorithms in this field to be unsupervised. To address the issue, this article proposes a new loss that adds prior information. One of the main functions of an autoencoder is to reconstruct the input data as accurately as possible after encoding them in a latent vector. The proposed loss function guides the network using the Reed-Xiaoli (RX) detector and replaces any pixels in the input data deemed too abnormal with normal surrounding values. This approach incorporates a priori information in addition to the assumption that anomalies are largely under-represented compared to the rest of the image. An ablation study demonstrates that the proposed loss function improves detection performance.

***Index Terms*—** anomaly detection, deep learning, self-supervised learning, adversarial autoencoder, SAR

1. INTRODUCTION

Remote sensing is a field of study that is of growing interest due to the increasing number of satellites in orbit, which provides a vast amount of data, particularly in the case of synthetic aperture radar (SAR) images. However, a major issue in this field is the lack of annotations. To train practical deep learning algorithms, it is typically necessary to have access to a well-labeled dataset. Without such annotations, we must use self-supervised algorithms that can learn independently. This presents a challenge because finding a suitable self-supervision task is not straightforward.

The case of anomaly detection, which is the detection of patterns that differ from their local background, is detailed in the article for polarimetric SAR images. This research area has been extensively studied in hyperspectral images [1], medical images [2], and industrial vision [3]. Some research has been made in SAR [4] but not as much as in the previously mentioned fields, although it has a lot of applications, such as oil slick detection, turbulent ship wake, levee anomaly or archaeology. This is mainly due to the problem of annotations and also to the fact that until recently, no denoising algorithms did not degrade the image's resolution. One of the most used algorithms for anomaly detection is the autoencoder or his modified version, the adversarial autoencoder (AAE) [5]. Because SAR data is corrupted by a strong noise called speckle [6], the use of these algorithms is inefficient and a despeckling pre-processing step has to be applied [7, 8]. Furthermore, the goal of an autoencoder is to reconstruct an image as best as possible, which is not the goal of anomaly detection.

In this paper, we propose a solution for the issue of limited labeled data in SAR by introducing a new reconstruction loss designed for anomaly detection with encoding-decoding networks. This is achieved by incorporating prior information given by any conventional detectors. In our case, An anomaly probability score for each pixel is generated by the RX detector [9] before the training. Then, the AAE training loss is guided by this score, a threshold is fixed to define what should be considered an anomaly in the reconstruction loss. Each pixel considered abnormal will then be replaced by surrounding "normal" values in the input image to force a lousy reconstruction of these anomalous areas.

2. PROPOSED METHOD

A deep-learning self-supervised algorithm developed explicitly for SAR anomaly detection [8] is used to test the proposed RX self-supervision method. The standard L_1 loss function is replaced with the newly defined loss. A pixel-per-pixel error and a covariance matrix difference are computed to determine the anomaly score. Let denote $\mathbf{X}^c \in \mathbb{C}^{h\times w\times p}$ the original complex-valued polarimetric SAR image of size $h \times w$ and depth p and $\mathbf{X} \in \mathbb{R}^{h\times w\times p}$ the same image in log-intensity after the despeckling process explained in [10]. The notation $\mathbf{x}_{i,j}$ characterizes the vector of size p, extracted from the matrix $\mathbf{X}$ at position (i, j). An overview of the architecture is displayed in Fig. 1. Each block will be explained in the following subsections.

2.1. Speckle-free reconstruction network

The deep learning anomaly detection algorithm is composed of two distinct steps. First, there is a speckle filtering preprocessing, and second, there is an AAE used to make an anomaly-free estimation of its input.

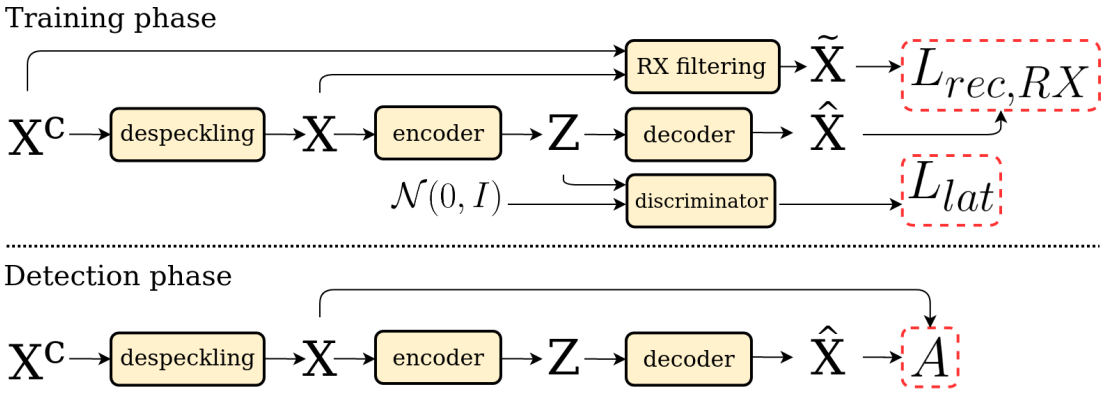


Fig. 1. Architecture of the AAE with the RX-guided reconstruction loss for the training phase and the detection phase.

SAR images contain a strong speckle noise. It corrupts the observations, and when it comes to anomaly detection, it creates a lot of false alarms. This explains why we need to remove this noise beforehand. It is essential to notice that SAR pixels are complex-valued, but with the despeckling algorithm, the phase is lost. We will then work with the intensity values, but the information we gain from removing the noise is more important than the information we lose by discarding the phase. This is done with the network [10] that is pre-trained to denoise ONERA X-band data in each polarization.

Once the speckle is removed, an AAE is used. The principle is to compress an image patch into a small vector containing the most information. The vector is then decoded into an estimation of the input patch. Because of the assumption that an anomaly rarely occurs, the network will not learn how to estimate them accurately. Detecting abnormal patterns in the input patch will be done by comparing the input patch and its estimation.

To train an AAE, there are two losses: a reconstruction loss corresponding to the L_1 norm between a patch and its estimation and a latent loss computed by a discriminator in the latent space to constrain it to follow a Gaussian distribution.

In practice, even if anomalies rarely occur in the training dataset, the network will, in some cases, be able to reconstruct them approximately. The statistically informed loss will help to address this issue.

2.2. Statisically informed reconstruction loss

For the purpose of anomaly detection, an efficient autoencoder is not defined by the quality of its reconstruction. The challenging part is to force a wrong recovery for the anomalies while preserving a good reconstruction in normal areas. This is why the standard reconstruction loss $L_{rec} = \|\mathbf{x} - \hat{\mathbf{x}}\|_1$ is not adapted for such a task. To add a priori information on what should be restored or not, L_{rec} between $\mathbf{X}$ and $\hat{\mathbf{X}}$ is replaced by the new loss $L_{rec,RX}$ between $\tilde{\mathbf{X}}$ and $\hat{\mathbf{X}}$ (respect to notations Fig. 1). To do so, the well-known detection algorithm called Reed-Xiaoli [9] is used on the complex SAR data. The principle is to locally estimate a scene's parameters (covariance matrix and mean vector) and to decide if the central pixel belongs to the same class as the neighboring pixels. For each coordinate (k, l), we define the patch $\mathbf{X}^c_{k,l} = \{\mathbf{x}^c_{i,j}\}_{i,j \in \mathcal{B}^e_{k,l}}$, where $\mathcal{B}^e_{k,l}$ denotes the set of indices (i, j) of pixels belonging to $\mathbf{X}^c_{k,l}$ and where the exponent e indicates that the boxcar contains an exclusion window to prevent us from estimating the scene with the test pixel and its close surroundings. Therefore, we can define the corresponding estimates:

$$\hat{\boldsymbol{\mu}}_{k,l} = |\mathcal{B}^e_{k,l}|^{-1} \sum_{i,j \in \mathcal{B}^e_{k,l}} \mathbf{x}^c_{i,j}\,, \tag{1}$$

$$\hat{\boldsymbol{\Sigma}}_{k,l} = \frac{1}{|\mathcal{B}^e_{k,l}| - 1} \sum_{i,j \in \mathcal{B}^e_{k,l}} \left(\mathbf{x}^c_{i,j} - \hat{\boldsymbol{\mu}}_{k,l}\right) \left(\mathbf{x}^c_{i,j} - \hat{\boldsymbol{\mu}}_{k,l}\right)^H, \tag{2}$$

where $\hat{\boldsymbol{\mu}}_{k,l}$ is the Sample Mean Vector estimate of the mean vector $\boldsymbol{\mu}$, $\hat{\boldsymbol{\Sigma}}_{k,l}$ is the Sample Covariance Matrix (SCM) estimate of the covariance matrix $\boldsymbol{\Sigma}$ associated to the complex-valued image patch $\mathbf{X}^c_{k,l}$. Then, the RX anomaly score is evaluated through the following procedure:

$$RX(\mathbf{x}^c_{k,l}) = \left(\mathbf{x}^c_{k,l} - \hat{\boldsymbol{\mu}}_{k,l}\right)^H \hat{\boldsymbol{\Sigma}}^{-1}_{k,l} \left(\mathbf{x}^c_{k,l} - \hat{\boldsymbol{\mu}}_{k,l}\right)\,, \tag{3}$$

From a given intensity denoised image $\mathbf{X}$, this score allows to define a new image $\tilde{\mathbf{X}}$ with less abnormal values:

$$\tilde{\mathbf{x}}_{k,l} = \begin{cases} \mathbf{x}_{k,l} \;\text{ if }\; RX(\mathbf{x}^c_{k,l}) \leq t, \\ m\left(\{\mathbf{x} \in \mathbf{X}_{k,l} \mid RX(\mathbf{x}^c) \leq t\}\right) \;\text{ otherwise,} \end{cases} \tag{4}$$

where the scalar t is a fixed threshold that will determine whether a pixel will be considered abnormal, where $m(.)$ is the median operator along the height and width dimensions, and where vectors $\mathbf{x}$ and $\mathbf{x}^c$ are respectively the intensity and

the complex value of the same pixel coordinates. To explain (4) more precisely, for every pixel considered abnormal by RX detector in patch $\mathcal{B}_{k,l}$, its value is replaced by the median of the pixels in $\mathcal{B}_{k,l}$ that are normal according to RX.

Finally, the new reconstruction loss can be defined through the following equation:

$$L_{rec,RX} = \frac{1}{h\,w\,p} \sum_{i,j} \|\tilde{\mathbf{x}}_{i,j} - \hat{\mathbf{x}}_{i,j}\|_1 \ . \tag{5}$$

2.3. Anomaly detection method

Once the reconstruction process is performed, a change detection between the input image $\mathbf{X}$ and its reconstruction $\hat{\mathbf{X}}$ will highlight badly reconstructed areas. For each pixel, a squared L_2 error is combined with a squared Frobenius norm of the difference of locally estimated covariance matrices. This leads to the anomaly score:

$$A_{k,l} = \lambda \left\| \mathbf{X}_{k,l} - \hat{\mathbf{X}}_{k,l} \right\|_2^2 + (1-\lambda) \left\| \hat{\mathbf{\Sigma}}^{\mathbf{X}}_{k,l} - \hat{\mathbf{\Sigma}}^{\hat{\mathbf{X}}}_{k,l} \right\|_F^2 . \tag{6}$$

where $\hat{\mathbf{\Sigma}}^{\mathbf{X}}_{k,l}$ and $\hat{\mathbf{\Sigma}}^{\hat{\mathbf{X}}}_{k,l}$ are respectively the SCM of $\mathbf{X}_{k,l}$ and $\hat{\mathbf{X}}_{k,l}$ locally estimated in the boxcar $\mathcal{B}_{k,l}$. The parameter $\lambda \in [0,1]$ is a scalar factor used to balance the importance of the two values that are normalized separately between 0 and 1 beforehand.

3. EXPERIMENTS

An ablation study is realized to test the apport of the new loss. The architecture is the same as the one described in [8]

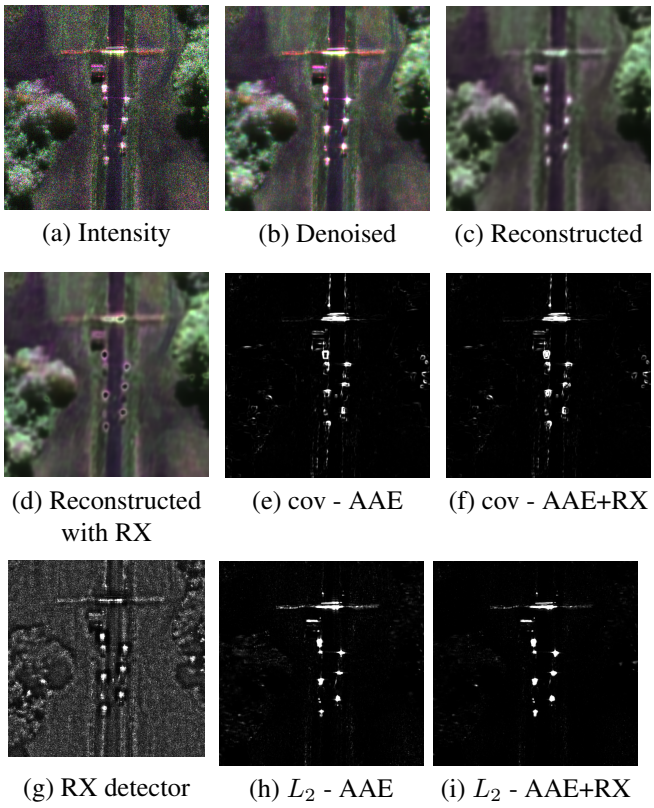
3.1. Dataset, training, and hyperparameters

The dataset is a full polarimetric X-band SAR image of size $4800 \times 30000 \times 4$ acquired by SETHI, the airborne SAR developed by the ONERA [11]. The resolutions of these images are about 20 cm in both azimuth and range domains for the four polarization channels ($p = 4$).

To train the AAE, we use 163785 patches of size 64×64 grouped in batches of size 128, so we get 1279 batches for one training epoch with a total of 40 epochs. To update the weights, we use the optimizer Adam [12] with a cyclical learning rate [13] that goes linearly from 10^{-3} to 10^{-2} in 2558 batches.

For the RX supervision, the SCM and SMV described in (2) and (1) are estimated with a boxcar of size 39×39 and an exclusion window of size 31×31. For (4), the threshold t is fixed at 0.04, and the boxcar is of size 64×64. For the anomaly detection score defined in (6), the SCM is estimated with a boxcar of size 11×11, and λ is equal to 0.68.

For every result, the operation $\tilde{x} = \min(x, s)$ is used to highlight a better dynamic range. For a given percentage p, the threshold s is fixed such that $p\%$ of the pixels are above it.



(a) Intensity (b) Denoised (c) Reconstructed

(d) Reconstructed with RX (e) cov - AAE (f) cov - AAE+RX

(g) RX detector (h) L_2 - AAE (i) L_2 - AAE+RX

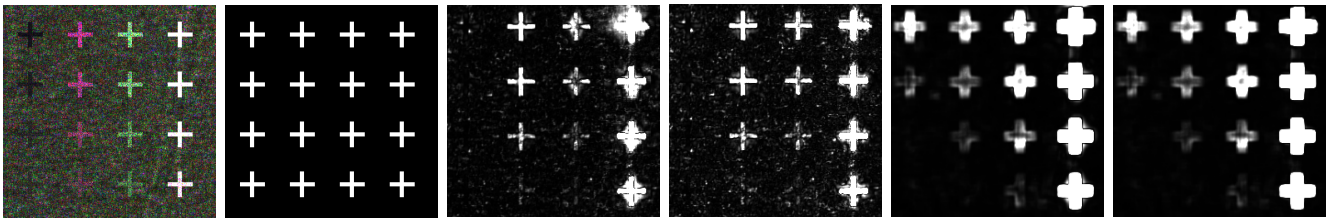
Fig. 2. Qualitative evaluation of the apport of the guided RX loss on real data. L_2 and cov are the two parts of (6)

The comparison of the same detection test (Fig. 2 (e), (f), for example) is always displayed with the same percentage.

3.2. Ablation study

To assess the benefits of the new loss (5), a comparison between the reconstructions with and without it is presented in Fig. 2. Images (c) and (d) are almost identical except for the bright points, which is the initial objective. Without supervision, a hole can be observed in the center of (e), which makes only the detection partial. This problem is corrected through supervision, as shown in (f). We also observe that the L_2 metric (i) contains less noise and better defines anomalies with supervision than (h). Finally, the detection with the RX in (g) shows that our method contains less noise.

A quantitative evaluation of anomaly detection is performed on synthetic data with exact labels. This evaluation is illustrated in Fig. 3 where a real SAR image on which we added test patterns based on real reflectivities is displayed. The value of λ and the size of the boxcar set above for Eq. (6) are optimal for both networks. We based the detection performances on the area under the receiver operating characteristic curve (AUC) for the synthetic anomaly image. Because an important part of the improvement lies in the L_2 error, its results are displayed in addition to the anomaly score. The



(a) test image (b) label (c) L_2 - AAE AUC : 0.780 (d) L_2 - AAE+RX AUC : 0.793 (e) A - AAE AUC : 0.877 (f) A - AAE+RX AUC : 0.881

Fig. 3. Quantitative evaluation of the apport of the guided RX loss on synthetic anomalies

L_2 metric yields better results with the RX-guided network, especially for the bright scatterers at the right of the image, where the test patterns are better defined. These improvements are highlighted in the final anomaly score in (e) and (f), where the results are less spread out on the edges.

4. CONCLUSION

This article proposes a new learning scheme for encoding-decoding networks to detect anomalies in SAR images. This is done by injecting *a priori* statistical information for the reconstruction loss. The proposed network can replace an abnormal pattern with its surrounding area while preserving the same reconstruction quality for the rest of the image. We proposed here to use the RX detector to guide our neural network. The experiments show an improvement in the SAR anomaly detection score. Due to the lack of annotated data in SAR, we have quantitatively evaluated the detection performance of this network through a synthetic anomaly dataset. Further methods of supervision based on the physics of SAR images or different statistical properties could be investigated to refine the learning strategy.

5. REFERENCES

[1] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Deep low-rank prior for hyperspectral anomaly detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.

[2] C. Baur, S. Denner, B. Wiestler, N. Navab, and S. Albarqouni, "Autoencoders for unsupervised anomaly segmentation in brain MR images: a comparative study," *Medical Image Analysis*, vol. 69, pp. 101952, 2021.

[3] C. L. Li, K. Sohn, J. Yoon, and T. Pfister, "CutPaste: Self-supervised learning for anomaly detection and localization," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE Computer Society, 2021, pp. 9659–9669.

[4] S. Mabu, S. Hirata, and T. Kuremoto, "Anomaly detection using convolutional adversarial autoencoder and one-class SVM for landslide area detection from synthetic aperture radar images," *Journal of Robotics, Networking and Artificial Life*, vol. 8, pp. 139–144, 2021.

[5] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *preprint arXiv:1511.05644*, pp. 1–16, 2015.

[6] J. W. Goodman, "Some fundamental properties of speckle," *JOSA*, vol. 66, no. 11, pp. 1145–1150, 1976.

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

[8] M. Muzeau, C. Ren, S. Angelliaume, M. Datcu, and J.-P. Ovarlez, "Self-supervised learning based anomaly detection in synthetic aperture radar imaging," *IEEE Open Journal of Signal Processing*, pp. 1–9, 2022.

[9] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE transactions on acoustics, speech, and signal processing*, vol. 38, no. 10, pp. 1760–1770, 1990.

[10] 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, 2021.

[11] R. Baqué, P. Dreuillet, and H. Oriot, "Sethi: Review of 10 years of development and experimentation of the remote sensing platform," in *International Radar Conference (RADAR)*, 2019, pp. 1–5.

[12] D. P Kingma and J. Ba, "Adam: A method for stochastic optimization," *preprint arXiv:1412.6980*, pp. 1–15, 2014.

[13] L. N. Smith, "Cyclical learning rates for training neural networks," in *IEEE winter conference on applications of computer vision (WACV)*. IEEE, 2017, pp. 464–472.