



HAL
open science

Apprentissage auto-supervisé guidé statistiquement pour la détection d'anomalies en imagerie SAR

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

► To cite this version:

Max Muzeau, Chengfang Ren, Datcu Mihai, Jean-Philippe Ovarlez. Apprentissage auto-supervisé guidé statistiquement pour la détection d'anomalies en imagerie SAR. GRETSI 2023, Aug 2023, GRENOBLE, France. hal-04248060

HAL Id: hal-04248060

<https://hal.science/hal-04248060v1>

Submitted on 18 Oct 2023

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.

Apprentissage auto-supervisé guidé statistiquement pour la détection d'anomalies en imagerie SAR

Max MUZEAU^{1,2} Chengfang REN¹ Datcu MIHAI³ Jean-Philippe OVARLEZ^{1,2}

¹SONDRA, CentraleSupélec, Université Paris-Saclay, 91192 Gif-sur-Yvette, France

²DEMR, ONERA, Université Paris-Saclay, 91120 Palaiseau, France

³University Politehnica of Bucharest (UPB), Romania

Résumé – La détection d'anomalies dans les images acquises par un radar à synthèse d'ouverture (SAR) suscite beaucoup d'intérêts en télédétection. Cependant, la rareté des données anormales et le manque d'annotations conduit la plupart des algorithmes dans ce domaine à être non supervisés. Pour résoudre ce problème, cet article propose une nouvelle fonction de coût prenant en compte des informations sur les statistiques des données d'entrées. Elle sera appliquée à un autoencodeur adversarial (AAE) dont le but est de reconstruire les données d'entrée sans les anomalies après les avoir encodées dans un vecteur latent. Pour cela, l'AAE est entraîné sur une fonction de perte guidée par le détecteur Reed-Xiaoli (RX) remplaçant tous les pixels de l'image d'entrée jugés anormaux par les valeurs normales environnantes. De plus, les anomalies sont largement sous-représentées par rapport au reste de l'image, ce qui induit une mauvaise reconstruction des anomalies après le traitement encodage/décodage par l'AAE. Finalement, un détecteur de changement est appliqué entre l'image d'entrée et de sortie de l'AAE pour délimiter les anomalies. Une étude expérimentale sur des images SAR en bande L montre que la fonction de coût proposée améliore les performances de détection.

Abstract – Anomaly detection in Synthetic Aperture Radar (SAR) images is an essential topic in remote sensing. However, the task is challenging due to the scarcity of anomalous samples and the lack of annotated data, leading most algorithms in this field to be unsupervised. This article proposes a new loss that adds prior statistical information to address the issue. It will be applied to an adversarial autoencoder (AAE) whose goal is to reconstruct the input data without anomalies after encoding them in a latent vector. To achieve this, the AAE network is trained on a loss function guided by the Reed-Xiaoli (RX) detector, which replaces any pixels in the input image deemed too abnormal with normal surrounding values. In addition, anomalies are largely under-represented compared to the rest of the input image, which led to a poor reconstruction of the anomalies after the encoding/decoding process by the AAE. Finally, a change detection is applied between AAE's input and output images to delimit the anomalies. An experimental study on L-band SAR images shows that the proposed loss function improves the detection performance.

1 Introduction

La télédétection est un domaine d'étude qui suscite un fort intérêt en raison du nombre croissant de satellites en orbite qui fournissent une grande quantité de données, en particulier dans le cas des images SAR. Toutefois, le manque d'annotations constitue un problème majeur dans ce domaine. L'apprentissage statistique à base de réseaux de neurones nécessite généralement un ensemble conséquent de données bien annotées. En l'absence de telles annotations, nous devons utiliser des algorithmes auto-supervisés qui peuvent apprendre de manière indépendante. Cela représente un défi car il n'est pas toujours simple de trouver une tâche d'auto-supervision appropriée.

Le cas de la détection d'anomalies, c'est-à-dire la détection de motifs qui diffèrent du fond de manière locale, est traité ici pour les images SAR polarimétriques. Ce domaine de recherche a été largement étudié dans les images hyperspectrales [1], les images médicales [2], et la vision industrielle [3]. Des recherches ont été menées dans le domaine du SAR [4], mais pas autant que dans les autres domaines susmentionnés, bien qu'il y ait de nombreuses applications, telles que la détection des marées noires, le sillage des navires ou l'archéologie. Cela est principalement dû au manque d'annotations. L'un des algorithmes les plus utilisés pour la détection d'anomalies

est l'autoencodeur ou sa version *adversarial* appelée AAE [5]. Comme les données SAR sont corrompues par un bruit important appelé le speckle [6], un prétraitement dit de "despeckling" est nécessaire pour filtrer ce bruit [7, 8]. En outre, l'objectif d'un autoencodeur est de reconstruire l'image le plus fidèlement possible, ce qui n'est pas l'objectif de la détection d'anomalies.

Dans cet article, nous proposons une solution au problème du manque de données annotées pour l'imagerie SAR en introduisant une nouvelle fonction de coût conçue pour la détection d'anomalies avec des réseaux de type encodeur-décodeur. Pour ce faire, nous intégrons une information *a priori* fournie par un détecteur statistique. Dans notre cas, le détecteur RX [9] génère une probabilité d'anormalité pour chaque pixel avant l'apprentissage. Ensuite, la fonction de coût est guidée par cette valeur, un seuil est fixé pour définir ce qui doit être considéré comme une anomalie ou non. Chaque pixel considéré comme anormal est remplacé par une valeur "normale" basée sur les pixels environnants afin de forcer une mauvaise reconstruction de ces zones anormales.

2 Méthode proposée

Un algorithme d'apprentissage profond auto-supervisé développé explicitement pour la détection d'anomalies en im-

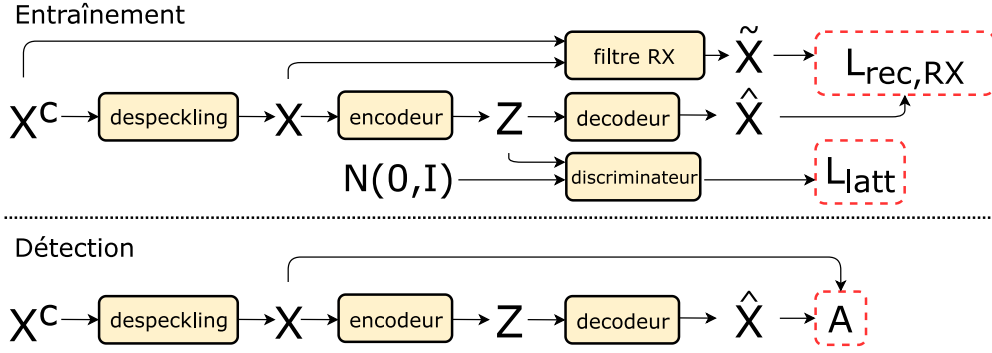


Figure 1: Architecture de l'AAE guidé par le détecteur RX pour la phase d'entraînement et de détection.

agerie SAR [8] est utilisé afin de tester la méthode proposée. La fonction de perte standard L_1 est remplacée par la nouvelle perte proposée dans la section (2.2). Une erreur pixel par pixel et une distance entre matrices de covariance sont calculées pour déterminer le score d'anomalie. Notons $\mathbf{X}^c \in \mathbb{C}^{h \times w \times p}$ l'image SAR polarimétrique à valeurs complexes de taille $h \times w$ et de profondeur p , et $\mathbf{X} \in \mathbb{R}^{h \times w \times p}$ la même image en log intensité après le processus de despeckling expliqué dans [10]. La notation $\mathbf{x}_{i,j}$ caractérise le vecteur de taille p , extrait de la matrice \mathbf{X} à la position (i, j) . Une vue d'ensemble de l'architecture est présentée dans la Fig. 1. Chaque bloc sera expliqué dans les sous-sections suivantes.

2.1 Réseau utilisé pour la détection d'anomalies

L'algorithme de détection d'anomalies par apprentissage profond se décompose en deux étapes distinctes. Tout d'abord, il y a un prétraitement par filtrage du speckle, ensuite, un AAE est utilisé pour faire une prédiction de l'image d'entrée sans les zones anormales.

Les images SAR contiennent un bruit de speckle important. Il corrompt les observations et, lorsqu'il s'agit de détecter des anomalies, il crée de nombreuses fausses alarmes. C'est pourquoi nous devons éliminer ce bruit au préalable. Il est essentiel de noter que les pixels SAR sont à valeur complexe, mais qu'avec l'algorithme de despeckling, la phase est perdue. Nous travaillerons alors avec les valeurs en intensité, mais le SNR gagné en éliminant le bruit est plus important que l'information que nous perdons en éliminant la phase. Ceci est réalisé avec le réseau [10] qui est pré-entraîné pour débruiter chaque polarisation des images SAR en bande L de l'ONERA.

Une fois le speckle filtré, un AAE est utilisé. Le principe est de compresser un patch d'une image en un petit vecteur contenant le plus d'information possible. Le vecteur est ensuite décodé en une estimation du patch d'entrée. En se basant sur l'hypothèse selon laquelle une anomalie se produit rarement, le réseau n'apprendra pas à l'estimer avec précision. La détection de patterns anormaux se fait en comparant l'image en entrée de réseau et son estimation.

Pour entraîner une AAE, il y a deux fonction de coût : une fonction de coût basée sur la reconstruction et correspondant à la norme L_1 entre un patch et son estimation. Puis une autre dans l'espace latent calculée par un discriminateur pour contraindre les vecteurs encodés par le réseau à suivre une distribution gaussienne.

En pratique, même si les anomalies sont rarement présentes dans l'ensemble de données d'apprentissage, le réseau sera, dans certains cas, capable de les reconstruire approximative-

ment. La nouvelle fonction de coût guidée par un détecteur statistique permettra de résoudre ce problème.

2.2 Fonction de coût guidée statistiquement

Dans le cadre de la détection des anomalies, un autoencodeur efficace n'est pas défini par la qualité de sa reconstruction. La difficulté consiste à forcer une mauvaise restitution des anomalies tout en préservant une bonne reconstruction des zones normales. C'est pourquoi la fonction de coût standard pour la reconstruction $L_{rec} = \|\mathbf{x} - \hat{\mathbf{x}}\|_1$ n'est pas adaptée à une telle tâche. Pour ajouter des informations *a priori* sur ce qui doit être restauré ou non, L_{rec} entre \mathbf{X} et $\hat{\mathbf{X}}$ est remplacé par la nouvelle fonction $L_{rec,RX}$ entre $\tilde{\mathbf{X}}$ et $\hat{\mathbf{X}}$ (par rapport aux notations Fig. 1). Pour ce faire, un algorithme de détection bien connu, appelé Reed-Xiaoli [9], est appliqué sur les données SAR complexes. Le principe consiste à estimer localement les paramètres d'une scène (matrice de covariance et vecteur de la moyenne) et à décider si le pixel central appartient à la même classe que les pixels voisins. Pour chaque coordonnée (k, l) , nous définissons le patch $\mathbf{X}_{k,l}^c = \{\mathbf{x}_{i,j}^c\}_{i,j \in \mathcal{B}_{k,l}^e}$, où $\mathcal{B}_{k,l}^e$ désigne l'ensemble des indices (i, j) des pixels appartenant à $\mathbf{X}_{k,l}^c$ et où l'exposant e indique que le boxcar contient une fenêtre d'exclusion pour nous empêcher d'estimer la scène avec le pixel test et son environnement proche. Nous pouvons donc définir les estimations correspondantes :

$$\hat{\boldsymbol{\mu}}_{k,l} = |\mathcal{B}_{k,l}^e|^{-1} \sum_{i,j \in \mathcal{B}_{k,l}^e} \mathbf{x}_{i,j}^c, \quad (1)$$

$$\hat{\boldsymbol{\Sigma}}_{k,l} = \frac{1}{|\mathcal{B}_{k,l}^e| - 1} \sum_{i,j \in \mathcal{B}_{k,l}^e} (\mathbf{x}_{i,j}^c - \hat{\boldsymbol{\mu}}_{k,l}) (\mathbf{x}_{i,j}^c - \hat{\boldsymbol{\mu}}_{k,l})^H.$$

où $\hat{\boldsymbol{\mu}}_{k,l}$ et $\hat{\boldsymbol{\Sigma}}_{k,l}$ sont respectivement la moyenne empirique (SMV) et la matrice de covariance empirique (SCM) associée au patch d'image à valeurs complexes $\mathbf{X}_{k,l}^c$ et où $|\mathcal{B}_{k,l}^e|$ indique le cardinal du patch $\mathbf{X}_{k,l}^c$. Le score d'anomalie RX est ensuite évalué selon la procédure suivante :

$$RX(\mathbf{x}_{k,l}^c) = (\mathbf{x}_{k,l}^c - \hat{\boldsymbol{\mu}}_{k,l})^H \hat{\boldsymbol{\Sigma}}_{k,l}^{-1} (\mathbf{x}_{k,l}^c - \hat{\boldsymbol{\mu}}_{k,l}), \quad (2)$$

À partir d'une image débruitée en intensité \mathbf{X} , ce score permet de définir une nouvelle image $\tilde{\mathbf{X}}$ avec moins de valeurs anormales :

$$\tilde{\mathbf{x}}_{k,l} = \begin{cases} \mathbf{x}_{k,l} & \text{si } RX(\mathbf{x}_{k,l}^c) \leq t, \\ m(\{\mathbf{x} \in \mathbf{X}_{k,l} \mid RX(\mathbf{x}^c) \leq t\}) & \text{sinon,} \end{cases} \quad (3)$$

où le scalaire t est un seuil fixe qui détermine si un pixel sera considéré comme anormal, où $m(\cdot)$ est l'opérateur médian sur

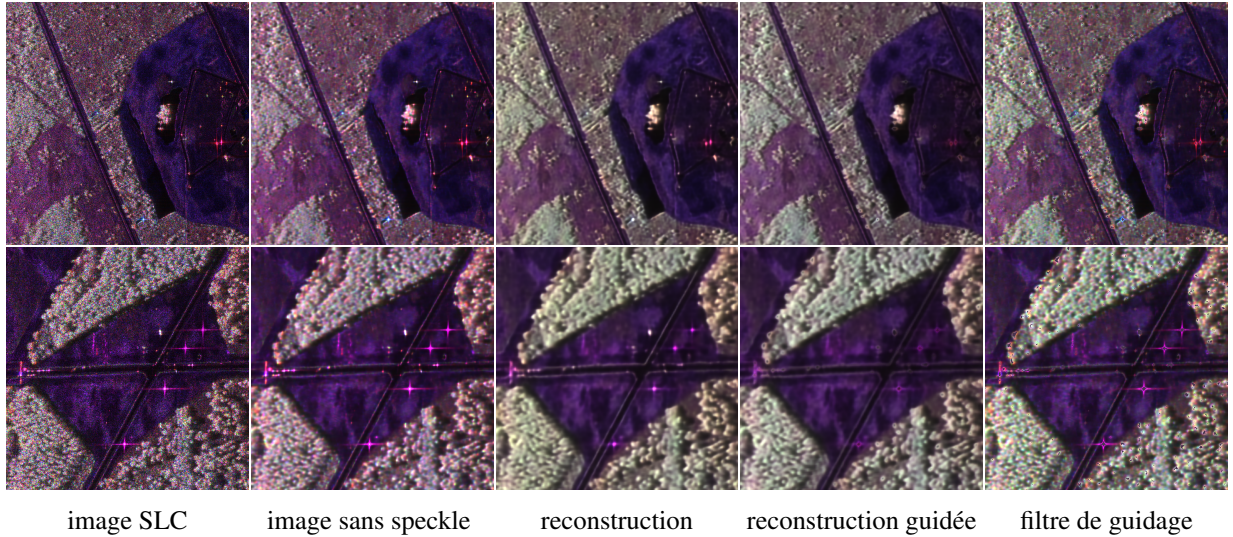


Figure 2: Evaluation de l'impact de la fonction de coût guidée par RX pour la reconstruction

les dimensions hauteur et largeur et où les vecteurs \mathbf{x} et \mathbf{x}^c sont respectivement l'intensité et la valeur complexe des mêmes pixels aux mêmes coordonnées. Plus précisément (3), pour chaque pixel considéré comme anormal par le détecteur RX dans le patch $\mathcal{B}_{k,l}$, sa valeur est remplacée par la médiane des pixels dans $\mathcal{B}_{k,l}$ qui sont normaux selon RX. Enfin, la nouvelle erreur de reconstruction peut être définie à l'aide de l'équation suivante :

$$L_{rec,RX} = \frac{1}{hwp} \sum_{i,j} \|\tilde{\mathbf{x}}_{i,j} - \hat{\mathbf{x}}_{i,j}\|_1. \quad (4)$$

2.3 Méthode de détection d'anomalies

Une fois le processus de reconstruction effectué, une détection des changements entre l'image d'entrée \mathbf{X} et sa reconstruction $\hat{\mathbf{X}}$ met en évidence les zones mal reconstruites. Pour chaque pixel, une erreur L_2 quadratique est combinée à une norme de Frobenius au carré de la différence des matrices de covariance estimées localement. On obtient ainsi le score d'anomalie :

$$A_{k,l} = \lambda \|\mathbf{X}_{k,l} - \hat{\mathbf{X}}_{k,l}\|_2^2 + (1 - \lambda) \|\hat{\Sigma}_{k,l}^{\mathbf{X}} - \hat{\Sigma}_{k,l}^{\hat{\mathbf{X}}}\|_F^2, \quad (5)$$

où $\hat{\Sigma}_{k,l}^{\mathbf{X}}$ et $\hat{\Sigma}_{k,l}^{\hat{\mathbf{X}}}$ sont respectivement la SCM de $\mathbf{X}_{k,l}$ et $\hat{\mathbf{X}}_{k,l}$ estimés localement dans le boxcar $\mathcal{B}_{k,l}$. Le paramètre $\lambda \in [0, 1]$ est un scalaire utilisé pour équilibrer l'importance des deux valeurs qui sont préalablement normalisées entre 0 et 1.

3 Résultats expérimentaux

On évalue ici l'apport de la nouvelle erreur de reconstruction. L'architecture est la même que celle décrite dans [8].

3.1 Dataset, entraînement et hyperparamètres

Le jeu de données est une image SAR polarimétrique en bande L de taille $1000 \times 15000 \times 4$ acquise par SETHI, le système d'acquisition aéroporté développé par l'ONERA [11]. La résolution de ces images est d'environ 1 mètre en azimut et en range pour les quatre canaux de polarisation ($p = 4$).

L'architecture du réseau est basée sur celle décrite dans [8]. Néanmoins, nous n'utilisons pas des données en bande X mais

en bande L. A cause de cela, des modifications sont apportées pour mieux convenir aux données. Pour entraîner l'AAE, nous utilisons 57224 patches de taille 32×32 regroupés en batchs de taille 128, de sorte que nous obtenons 447 batchs pour une époque avec un total de 140 époques. Pour mettre à jour les poids, nous utilisons l'optimiseur Adam [12] avec un taux *learning rate* cyclique [13] qui passe linéairement de 10^{-3} à 10^{-2} en 862 batchs.

Pour la supervision RX, les paramètres SCM et SMV décrits à l'Eq. (1) sont estimés avec un boxcar de taille 31×31 et une fenêtre d'exclusion de taille 16×16 . Pour (3), le seuil t est fixé à 0.04, et le boxcar est de taille 32×32 . Pour le score de détection des anomalies défini dans (5), la SCM est estimée avec un boxcar de taille 3×3 , et λ est égal à 0,33.

Pour chaque résultat de détection, l'opération $\tilde{x} = \min(x, s)$ est utilisée pour obtenir une meilleure dynamique. Pour un pourcentage n donné, le seuil s est fixé de telle sorte que $n\%$ des pixels lui soient supérieurs. La valeur de n est égale à 0.2.

3.2 Étude des résultats

Pour évaluer les avantages de l'erreur de reconstruction (4), une comparaison entre les reconstructions avec et sans cet ajout est présentée dans la Fig. 2. On remarque tout d'abord qu'entre l'image originelle et celle sans speckle, les structures sont préservées. Lorsque le réseau est entraîné sans le guidage, les zones définies comme "normales", telles que de la végétation ou de la terre, sont reconstruites fidèlement. On aperçoit un léger flou de l'image causé par la fonction de coût L_2 ainsi qu'une atténuation des polarisations HH et VV dans la végétation, ces erreurs de reconstructions sont moins importantes que celles des anomalies, ce qui ne dégrade pas la qualité de détection. Comme il n'y a quasiment aucun bâtiment dans les données d'entraînement, les forts rétrodiffuseurs sont considérés comme des zones "anormales". Ces zones sont, pour la plupart, atténuées par le réseau lors de la reconstruction, mais elles restent présentes. Lorsque le réseau est guidé avec le détecteur RX, les zones classifiées comme *normales* sont reconstruites exactement comme s'il n'y avait pas de guidage, alors que les pixels classifiés comme anormaux sont remplacés par les valeurs des pixels environnants. Quand on compare la reconstruction avec et sans guidage dans la Fig. 2, on remarque

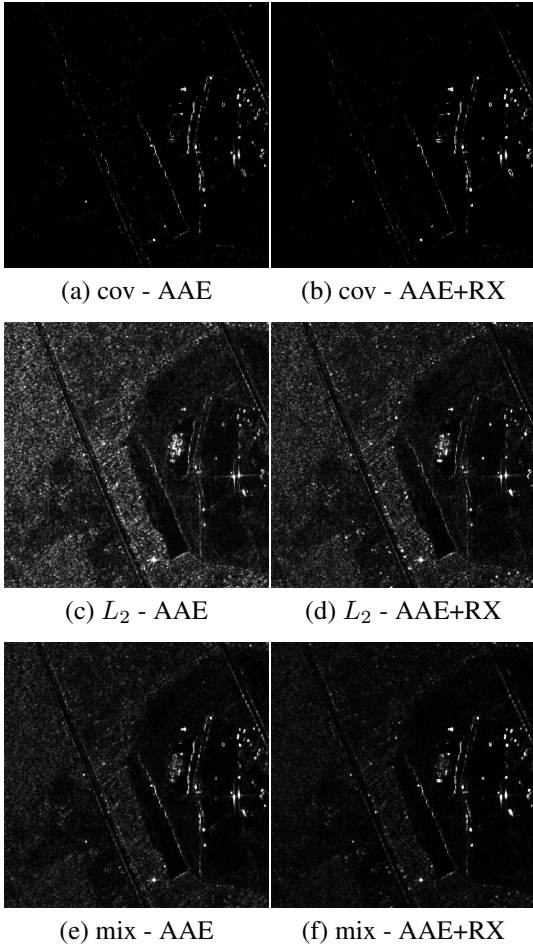


Figure 3: Evaluation qualitative de l'apport de la perte guidée par RX (L_2 et cov représentent les deux termes de (5)).

que la "connaissance" du détecteur, par rapport à ce qui doit être considéré *anormal* ou non, a été transférée au réseau de neurones. Les zones masquées dans le filtre de guidage sont moins fidèlement reconstruites, ce qui est le résultat souhaité.

Comme on l'observe dans Fig. 3 avec les résultats de l'erreur L_2 , les zones "anormales" sont reconstruit de manière beaucoup moins fidèle que les zones "normales". La détection effectuée grâce à la différence entre matrice de covariances permet un résultat avec très peu de bruit, mais présente des trous lorsque l'objet détecté est trop énergétique. À l'inverse l'erreur L_2 permet de mieux faire ressortir les erreurs de reconstruction mais le résultat est plus bruité. C'est pourquoi, un compromis entre ces deux méthodes de détection est mis en place, elles se complètent en corrigeant leurs défauts mutuels.

Rajouter un guidage au réseau de neurones permet au réseau d'avoir un résultat plus confiant. Les zones considérées comme anormales sans le guidage le sont encore plus avec et, à l'inverse, le bruit de la carte de détection est moins marqué. Cette amélioration permet de combler en partie les trous présents dans la carte de détection avec les matrices de covariance. Une autre méthode permettant de les combler pourrait être basée sur l'usage d'une fenêtre plus grande pour estimer ces matrices.

4 Conclusion

Cet article propose une nouvelle méthode d'apprentissage pour les réseaux d'encodage-décodage afin de détecter les anomalies dans les images SAR. Pour ce faire, on injecte des informations statistiques *a priori* pour l'erreur de reconstruction. Le réseau proposé peut remplacer un motif anormal par sa zone environnante tout en préservant la même qualité de reconstruction pour le reste de l'image. Nous utilisons ici le détecteur RX pour guider le réseau. Les expériences montrent une amélioration de la détection des anomalies dans l'image SAR. D'autres méthodes de supervision basées sur la physique des images SAR ou sur différentes propriétés statistiques seront étudiées pour affiner la stratégie d'apprentissage.

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, p. 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)*, pp. 9659–9669, IEEE Computer Society, 2021.
- [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*, pp. 1–6, 2022.
- [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*, vol. 3, pp. 440–449, 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)*, pp. 1–5, 2019.
- [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)*, pp. 464–472, IEEE, 2017.