MDDNet: multilevel difference-enhanced denoise network for unsupervised change detection in SAR images.

ZONG, H., ZHANG, E., LI, X., ZHANG, H. and REN, J.

2025

MDDNet: Multilevel Difference-Enhanced Denoise Network for Unsupervised Change Detection in SAR Images

He Zong
College of Information Engineering
Northwest A&F University
Yangling, China
1549680969@nwafu.edu.cn

Erlei Zhang*
College of Information Engineering
Northwest A&F University
Yangling, China
erlei.zhang@nwafu.edu.cn

Xinyu Li
College of Information Engineering
Northwest A&F University
Yangling, China
xinyuli@nwafu.edu.cn

Hongming Zhang
College of Information Engineering
Northwest A&F University
Yangling, China
zhm@nwsuaf.edu.cn

Jinchang Ren
School of Computing and Engineering
Robert Gordon University
Aberdeen, United Kingdom
j.ren@rgu.ac.uk

***Abstract*—Change detection in synthetic aperture radar (SAR) images is a hot yet highly challenging task in remote sensing. Existing unsupervised SAR change detection methods often struggle with inherent speckle noise and insufficiently utilize pseudo-labels, particularly neglecting uncertain areas. In this paper, we propose a multilevel difference-enhanced denoise dual-branch network (MDDNet), comprising representation learning and change detection branches. First, fuzzy c-means clustering is employed to generate pseudo-labels, categorizing the image areas as changed, nochanged, and uncertain. Second, we design a denoise representation loss function in the representation learning branch to maximize the use of pseudo-labels, while mitigating speckle noise. Furthermore, a multilevel difference computation module is proposed to focus on changes in ground objects and capture more comprehensive change information. Experimental results on three public SAR datasets show that the proposed method outperforms six state-of-the-art methods, achieving the best performance with an average overall accuracy of 98.86% and an average Kappa coefficient of 89.36%.**

***Index Terms*—Change detection, denoise representation, multilevel difference computation, synthetic aperture radar (SAR) images**

I. Introduction

Remote sensing change detection is a technique of identifying changes in ground objects by analyzing the images acquired at different times in the same geographical location [1]. In particular, synthetic aperture radar (SAR) images are independent of light conditions, enabling the monitoring of ground objects under all weather and at all times. Due to this property, SAR image change detection technology is extensively applied in various fields including resource monitoring [2], urban planning [3], and disaster assessment [4].

This research was funded by the National Natural Science Foundation of China (No.62376225).

* Corresponding author: Erlei Zhang (erlei.zhang@nwafu.edu.cn)

From the perspective of the availability of ground truth labels, SAR image change detection methods can be broadly categorized into two types: supervised and unsupervised. Supervised methods are based on supervised classification schemes, aiming to train a powerful classifier with the labeled data. However, collecting sufficient labeled data is time-consuming and labor-intensive due to the distinctive imaging mechanism of SAR images [5]. In contrast, unsupervised methods are more practical and challenging, as they do not rely on ground truth labels. Recently, unsupervised SAR change detection methods based on deep neural networks (DNNs) have advanced significantly with the development of deep learning [6]. These DNN-based methods follow a two-step approach. Initially, traditional change detection algorithms such as clustering are employed to generate pseudo-labels, categorizing image areas into three classes: changed, nochanged, and uncertain. Subsequently, a DNN is designed and trained using the generated pseudo-labels. For instance, Gong *et al.* [7] utilized the deep belief network to detect changes that occurred between SAR images. Yi *et al.* [8] incorporated Gabor wavelet into their proposed network to diminish the impact of speckle noise. In [9], self-attention and convolution were combined to effectively capture global semantic information, highlighting important features. Ma *et al.* [10] introduced a feature fusion of information transfer network (FFITN), incorporating an information transfer module to capture salient area knowledge.

While many unsupervised methods have succeeded in their applications, there are still several main drawbacks to them. First, most methods [9]–[13] emphasize the design of diverse DNNs, but often overlook the full utilization of pseudo-labels. They mainly focus on the changed and nochanged areas when using pseudo-labels, neglecting the information in uncertain areas, which can degrade the change detection

performance. Second, some methods [8] [14] [15] incorporate image decomposition, like wavelet transform, into the network to retain low-frequency information in extracted features while losing high-frequency information to mitigate speckle noise interference. However, this may result in losing feature details, leading to numerous false detections. Third, the multilevel features extracted by their networks remain underutilized, failing to capture rich multilevel change information.

To address the aforementioned challenges, we propose a Multilevel Difference-enhanced Denoise dual-branch Network (MDDNet) for unsupervised SAR change detection. In the representation learning branch, a denoising representation loss (DRL) is proposed to maximize the use of pseudo-labels, especially for uncertain areas, while reducing the interference of speckle noise without causing feature loss. In the change detection branch, we introduce a multilevel difference computation module (MDCM) to leverage the multilevel features in our network, capturing comprehensive change information and enhancing the learning of changes in ground objects.

The main contributions can be summarized as follows.

(1) A dual-branch network is designed for unsupervised SAR change detection. In the representation learning branch, we introduce a simple yet effective DRL that maximizes the use of pseudo-labels and improves noise resistance without losing feature details.

(2) We propose MDCM in the change detection branch to capture comprehensive change information and improve the detection of changes in ground objects by leveraging the multilevel features extracted by our network.

(3) The impressive experimental results on three public SAR datasets demonstrate the superiority and practicability of the proposed method compared to six state-of-the-art (SOTA) unsupervised SAR change detection methods.

II. Methodology

The overall diagram of MDDNet is illustrated in Fig. 1, comprising three key parts: (1) pre-classification; (2) representation learning branch; and (3) change detection branch.

A. Pre-classification

Given a pair of bi-temporal co-registered SAR images X and Y of size $H \times W$, capturing over the same geographical area at different times t_1 and t_2, the goal of MDDNet is to generate a change map $\mathrm{CM} \in \mathbb{R}^{H \times W}$ that accurately reflects the changed areas between X and Y. Consistent with other unsupervised DNN-based methods, we first obtain the difference image DI for X and Y using the logarithmic difference operator:

$$\mathrm{DI} = |\log(X) - \log(Y)| \tag{1}$$

Subsequently, fuzzy c-means (FCM) clustering [9] is employed on DI to derive pseudo-labels, categorizing image areas into three classes: changed (Ω_c), nochanged (Ω_n), and uncertain (Ω_u), then guiding the subsequent network learning.

Next, we use a non-overlapping sliding window approach to extract image patches from X and Y as our training samples. In this procedure, the window size is set to $p \times p$, which is also

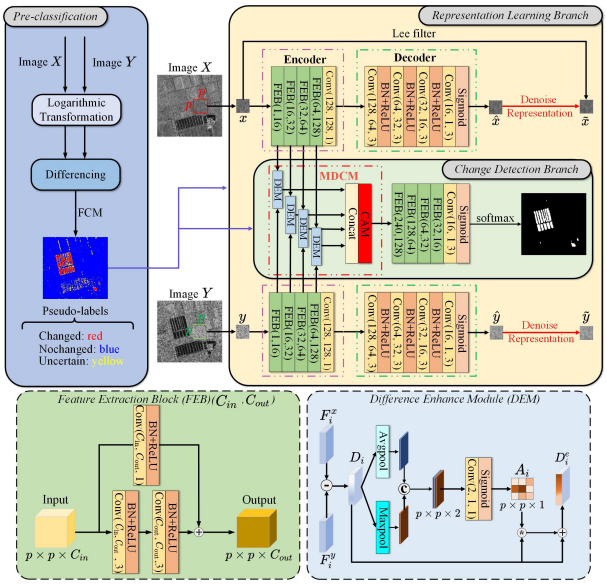

Fig. 1. The overall diagram of the proposed MDDNet. The difference image is obtained by measuring the difference between two logarithmic SAR images X and Y, followed by fuzzy c-mean (FCM) clustering to generate the pseudo-labels, segmenting image areas into three classes (changed, nochanged, and uncertain). In the representation learning branch, image patches x and y located at the same spatial location are input into Encoder and Decoder to obtain $\hat{x}$ and $\hat{y}$. Next, the denoising representation learning is applied to $(\hat{x}, \tilde{x})$ and $(\hat{y}, \tilde{y})$ to learn robust and denoised features, where $\tilde{x}$ and $\tilde{y}$ are the patches filtered by the Lee filter. In the change detection branch, a multilevel difference computation module (MDCM) is employed to fully leverage multilevel features extracted from the Encoder for supervised classification with pseudo-labels. (Here, $p \times p$ is the size of SAR image patches, and a standard channel attention mechanism (CAM) is applied in MDCM to highlight key information in multilevel differences)

the extracted image patch size. Then, corresponding image patches x_i and y_i are extracted, where $i \in \{1, \ldots, N\}$ and N represents the number of patches.

B. Representation Learning Branch

Through the aforementioned steps, we have obtained image patches as training samples. Here, assuming a pair of image patches x and y of size $p \times p$ are extracted from X and Y.

In the representation learning branch, self-supervised learning networks comprising Encoder and Decoder are established to process x and y. The network structure and parameter settings are detailed in Fig. 1. Initially, x and y are fed into the Encoder to extract features $F_x \in \mathbb{R}^{p \times p \times 128}$ and $F_y \in \mathbb{R}^{p \times p \times 128}$, where the channel dimension is 128.

Recent methods, such as [10] and [13], often neglect uncertain areas during network training because these areas cannot be simply utilized for feature constraints or classifier training to enhance the network's discriminatory capability like the changed and nochanged areas with accurate labels. Nevertheless, we can utilize the information from uncertain areas to enhance the network's representation capability. Additionally, incorporating image decomposition into the network to mitigate the impact of speckle noise may lead to feature information loss, which requires a more effective solution. Therefore, we introduce DRL in the image domain to naturally

address these concerns. Motivated by [16], we first apply the classic Lee filter [17] to denoise image patch x to obtain filtered image patch $\tilde{x}$. Similarly, we can obtain $\tilde{y}$ for y. Then, the Decoder maps the features F_x and F_y back to the original image domain, obtaining $\hat{x}$ and $\hat{y}$. Subsequently, DRL is employed to further utilize all image areas in pairs $(\hat{x},\tilde{x})$ and $(\hat{y},\tilde{y})$, especially uncertain areas, allowing the features extracted by Encoder to accurately reflect geographical information with reduced speckle noise. It is defined as

$$\mathcal{L}_{\mathrm{DR}}(\Omega_c,\Omega_n,\Omega_u)=\sum_{(i,j)}(\Omega_c+\Omega_n+\Omega_u)(\|\hat{x}_{(i,j)}-\tilde{x}_{(i,j)}\|_1+\|\hat{y}_{(i,j)}-\tilde{y}_{(i,j)}\|_1) \tag{2}$$

By introducing $\mathcal{L}_{\mathrm{DR}}$, we maximize the utilization of all image areas, including uncertain areas while minimizing the impact of speckle noise without losing feature information.

C. Change Detection Branch

In the change detection branch, we introduce MDCM to utilize the multilevel features F_i^x and F_i^y extracted by four cascaded Feature Extraction Blocks (FEBs) in the Encoder, where $i\in\{1,2,3,4\}$. Simply subtracting F_i^x and F_i^y to obtain difference information lacks attention to changes, which may hinder the detection of changed ground objects. Therefore, we propose the Difference Enhance Module (DEM) (depicted in Fig. 1) to enable the network to filter out irrelevant changes and focus on truly changed areas. Specifically, We initially derive D_i by computing $\left|F_i^x-F_i^y\right|$, where $|\cdot|$ denotes the absolute value operation. Subsequently, we calculate the attention map $A_i\in\mathbb{R}^{p\times p\times 1}$ using the formula:

$$A_i=\mathrm{Conv}(\mathrm{Concat}(\mathrm{Avgpool}(D_i),\mathrm{Maxpool}(D_i))) \tag{3}$$

where Avgpool and Maxpool denote the average-pooling and max-pooling operations, while Conv represents 1×1 convolutional layer followed by a Sigmoid function. Now we can enhance D_i to obtain more resilient difference information as

$$D_i^e=D_i\times A_i+D_i \tag{4}$$

where D_i^e is the enhanced D_i. Next, $\{D_1^e,D_2^e,D_3^e,D_4^e\}$ are concatenated along the channel dimension and then processed with a standard channel attention mechanism (CAM) from [18] to highlight critical information from multiple levels of feature differences

$$D^e=\mathrm{CAM}(\mathrm{Concat}(D_1^e,D_2^e,D_3^e,D_4^e)) \tag{5}$$

where D^e denotes the final difference information. Then D^e is fed into the subsequent network layer to generate probability map $P\in\mathbb{R}^{p\times p}$ for classification using pseudo-labels Ω_c and Ω_n. The cross-entropy loss function is used as the change detection loss to classify image pixels, calculated as

$$\mathcal{L}_{\mathrm{CD}}(\Omega_c,\Omega_n)=-\sum_{(i,j)}(y_{(i,j)}\log(p_{(i,j)})+(1-y_{(i,j)})\log(1-p_{(i,j)})) \tag{6}$$

where $y_{(i,j)}\in\{0,1\}$ is the pseudo-label of pixels, and $p_{(i,j)}\in[0,1]$ denotes the softmax probability of pixels in P. If $y_{(i,j)}=1$, then $y_{(i,j)}\in\Omega_c$; otherwise, $y_{(i,j)}\in\Omega_n$. It is worth noting that pixels in Ω_u do not contribute to $\mathcal{L}_{\mathrm{CD}}$.

The proposed MDDNet can be trained by minimizing $\mathcal{L}_{\mathrm{total}}$, which can be written as

$$\mathcal{L}_{\mathrm{total}}=\mathcal{L}_{\mathrm{DR}}+\mathcal{L}_{\mathrm{CD}} \tag{7}$$

After training MDDNet, all the image patches of X and Y are input into it, and the final CM can be obtained from the change detection branch.

III. Experimental Results and Analysis

A. Datasets and Evaluation Metrics

To validate our proposed method, we conduct experiments on three public SAR datasets, as depicted in Fig. 2(a)-(c). The Bern dataset [20] includes two images with a size of 300×412, acquired in April and May 1999 by the European Remote Sensing Satellite-2. The Red River dataset [21], comprising two images from August 1996 and August 1999 captured by the ERS-2 satellite, each with a size of 512×512 pixels. Finally, the Farmland dataset [19] contains two images sized 306×291, collected by the RADARSAT-2 satellite in June 2008 and June 2009. The two images in this dataset are single-look and four-look, which means different noise levels, increasing the difficulty in identifying changes.

In quantitatively measuring the performance of different methods, six common evaluation metrics [22] are employed in our experiments, including false positive (FP), false negative (FN), overall error (OE), overall accuracy (OA), Kappa coefficient (KC), and computation time (CT). The lower the value of FP, FN, OE, and CT, the higher the value of OA and KC, and the better the performance of one method.

B. Implementation Details

In our experiments, the image patch size p is set to 32. The Adam optimizer [23] with an initial learning rate $\mathrm{lr}=0.0001$, $\beta=\{0.9,0.999\}$ is used for network optimization across all datasets. The training epochs and batch size are set to 100 and 16, respectively. To verify the superiority of our proposed method, we compare it against six SOTA unsupervised SAR change detection methods including SAFNet [19], DDNet [11], FFITN [10], CAMixer [9], TSPLR [13], and WBANet [12]. All the experiments are conducted using the PyTorch library on a workstation equipped with an NVIDIA GeForce RTX 3090 GPU.

C. Performance Comparison

Visual comparisons of change maps generated by different methods across three datasets are presented in Fig. 2, while the corresponding quantitative evaluation metrics are provided in Table I. As illustrated in Fig. 2, comparative methods misclassify more changed areas as nochanged (highlighted in green), leading to a high FN. In contrast, MDDNet detects more truly changed areas with the lowest FN, effectively extracting rich change information and demonstrating superior performance. Notably, MDDNet can also accurately detect changes with a few noisy points in the Farmland dataset, which has much

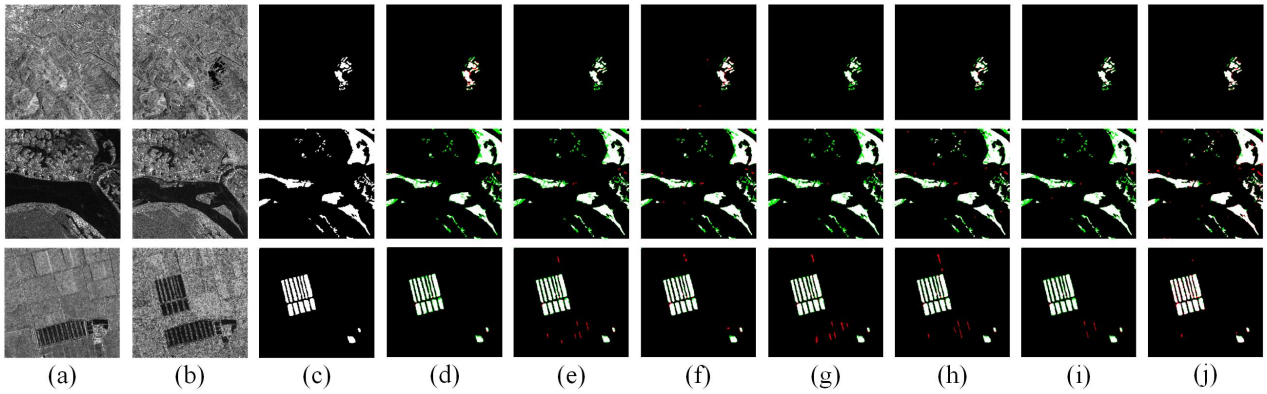

Fig. 2. Visual comparison of the change maps generated by different change detection methods across three SAR datasets, where the first to third rows are the Bern, Red River, and Farmland datasets, respectively. (a) SAR image X. (b) SAR image Y (c) Ground truth image. (d) Result by SAFNet [19]. (e) Result by DDNet [11]. (f) Result by FFITN [10]. (g) Result by CAMixer [9]. (h) Result by TSPLR [13]. (i) Result by WBANet [12]. (j) Result by the proposed MDDNet. (The correctly classified changed and nochanged areas are shown in white and black colors. Red indicates FP, and green indicates FN)

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT METHODS ACROSS THREE SAR DATASETS. THE BEST RESULTS ARE IN BOLD, AND THE SECOND-BEST RESULTS ARE UNDERLINED.

Method	Results on Bern Dataset						Results on Red River Dataset						Results on Farmland Dataset					
	FP	FN	OE	OA(%)	KC(%)	CT(s)	FP	FN	OE	OA(%)	KC(%)	CT(s)	FP	FN	OE	OA(%)	KC(%)	CT(s)
SAFNet [19]	228	227	455	99.49	80.06	720.12	**192**	9493	9685	96.31	80.90	1968.78	**50**	1010	1060	98.81	88.31	681.72
DDNet [11]	41	342	383	99.58	80.73	180.42	336	8149	8485	96.76	83.63	181.66	290	809	1099	98.77	88.38	152.93
FFITN [10]	210	**167**	377	99.58	83.77	372.12	660	7205	7865	97.00	85.10	1405.82	182	637	819	99.08	91.39	399.48
CAMixer [9]	**31**	450	481	99.47	74.31	**62.22**	225	9477	9702	96.30	80.88	115.91	438	738	1176	98.68	87.82	**56.12**
TSPLR [13]	78	280	358	99.60	82.82	110.69	694	7197	7891	96.99	85.06	458.68	355	705	1060	98.81	88.97	144.12
WBANet [12]	46	310	356	99.61	82.40	83.62	252	9015	9267	96.46	81.87	155.19	141	976	1117	98.75	87.83	75.61
MDDNet	147	171	**318**	**99.65**	**85.91**	78.52	2169	**3662**	**5831**	**97.78**	**89.69**	**110.81**	499	**263**	**762**	**99.14**	**92.47**	63.81

stronger speckle noise, highlighting its denoising capability. As reported in Table I, the proposed MDDNet achieves the highest OA and KC. Specifically, MDDNet demonstrates notable enhancements in KC compared to the second-best method across all datasets: 2.14% (Bern), 4.59% (Red River), 1.08% (Farmland). Additionally, the computational time costs of all methods are also presented in Table I using the CT metric. Our method achieved nearly the lowest CT, with only a bit time-consuming compared to CAMixer. Its CT is acceptable given the good detection results with the highest OA and KC obtained by our proposed MDDNet, highlighting its practicality. Overall, the experimental results across all datasets demonstrate the efficacy and practicality of MDDNet in both quantitative and qualitative evaluations.

D. Ablation Study

We conduct ablation experiments to verify the validity of the proposed MDCM and DRL. Three variants are designed for comparison: (1) a basic network without MDCM or DRL, in which the output of the final FEB is fed into the change detection branch; (2) a network with MDCM; (3) a network combining both MDC and representation loss (RL), where RL denotes reconstruct the original patch instead of the filtered patch; and (4) our complete method with MDCM and DRL.

The contributions of each component are presented in Table II. Compared to single-level features, leveraging multilevel features enables a more comprehensive extraction of change information, resulting in enhanced performance. The utilization of RL maximizes the use of pseudo-labels, enhancing the OA and KC values across all datasets. Additionally, our complete method incorporates denoising within the basis of RL, effectively reducing the impact of speckle noise without losing feature information, thereby achieving superior change detection performance. Overall, the effectiveness of MDDNet can be demonstrated through the analysis provided above.

TABLE II
ABLATION STUDIES OF THE PROPOSED MDDNET

Method	Bern		Red River		Farmland	
	OA	KC	OA	KC	OA	KC
Basic Network	99.41	79.21	96.89	85.36	98.43	86.93
with MDCM	99.57	81.81	97.19	86.48	98.68	88.69
with MDCM, RL	99.58	83.51	97.52	88.25	98.93	90.76
MDDNet	**99.65**	**85.91**	**97.78**	**89.69**	**99.14**	**92.47**

IV. CONCLUSION

In this paper, we propose MDDNet, a novel unsupervised SAR change detection network that includes representation learning and change detection branches. In the representation learning branch, we design DRL to maximize the use of pseudo-labels and reduce the impact of speckle noise while maintaining feature integrity. Furthermore, we propose MDCM in the change detection branch to extract rich change information by leveraging multilevel features. The remarkable experimental results on three datasets demonstrate the effectiveness of MDDNet over six SOTA unsupervised methods.

REFERENCES

[1] G. Cheng, Y. Huang, X. Li, S. Lyu, Z. Xu, H. Zhao, Q. Zhao, and S. Xiang, "Change detection methods for remote sensing in the last decade: A comprehensive review," *Remote Sensing*, vol. 16, no. 13, 2024.

[2] T. K. Thakur, D. K. Patel, A. Bijalwan, M. J. Dobriyal, A. Kumar, A. Thakur, A. Bohra, and J. A. Bhat, "Land use land cover change detection through geospatial analysis in an indian biosphere reserve," *Trees, Forests and People*, vol. 2, p. 100018, 2020.

[3] M. Che and P. Gamba, "Bi- and three-dimensional urban change detection using sentinel-1 SAR temporal series," *GeoInformatica*, vol. 25, no. 4, pp. 759–773, 2021.

[4] E. Hamidi, B. G. Peter, D. F. Muñoz, H. Moftakhari, and H. Moradkhani, "Fast flood extent monitoring with sar change detection using google earth engine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–19, 2023.

[5] J. Wang, F. Gao, J. Dong, S. Zhang, and Q. Du, "Change detection from synthetic aperture radar images via graph-based knowledge supplement network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1823–1836, 2022.

[6] M. Zhang, L. Liu, Z. Lei, K. Ma, J. Feng, Z. Liu, and L. Jiao, "Multiscale spatial-channel transformer architecture search for remote sensing image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[7] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 1, pp. 125–138, 2016.

[8] W. Yi, S. Wang, N. Ji, C. Wang, Y. Xiao, and X. Song, "Sar image change detection based on gabor wavelets and convolutional wavelet neural networks," *Multimedia Tools and Applications*, pp. 1–14, 2023.

[9] H. Zhang, Z. Lin, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Convolution and attention mixer for synthetic aperture radar image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[10] J. Ma, D. Li, X. Tang, Y. Yang, X. Zhang, and L. Jiao, "Unsupervised sar image change detection based on feature fusion of information transfer," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[11] X. Qu, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Change detection in synthetic aperture radar images using a dual-domain network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[12] J. Xie, F. Gao, X. Zhou, and J. Dong, "Wavelet-based bi-dimensional aggregation network for sar image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[13] S. Fang, C. Qi, S. Yang, Z. Li, W. Wang, and Y. Wang, "Unsupervised sar change detection using two-stage pseudo labels refining framework," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[14] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in sar images based on convolutional-wavelet neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1240–1244, 2019.

[15] R. V. Fonseca, R. G. Negri, A. Pinheiro, and A. M. Atto, "Wavelet spatio-temporal change detection on multitemporal sar images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 4013–4023, 2023.

[16] T. G. Kamod, P. P. Rege, and S. Kulkarni, "Denoise auto-encoder based speckle reduction for risat-1 sar imagery," in *Proceedings of the 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2021, pp. 216–221.

[17] J.-S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Computer Graphics and Image Processing*, vol. 17, no. 1, pp. 24–32, 1981.

[18] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[19] Y. Gao, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Synthetic aperture radar image change detection via siamese adaptive fusion network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 10 748–10 760, 2021.

[20] K. Zhang, X. Lv, H. Chai, and J. Yao, "Unsupervised sar image change detection for few changed area based on histogram fitting error minimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–19, 2022.

[21] H. Dong, W. Ma, L. Jiao, F. Liu, and L. Li, "A multiscale self-attention deep clustering for change detection in sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.

[22] G. Chen, Y. Zhao, Y. Wang, and K.-H. Yap, "Ssn: Stockwell scattering network for sar image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.