D-CDA:A Denoise And Change Detection Approach For Flood disaster location From SAR Images

Runbo Xie[1],Yangguang[1],Yuping Zhang[1],Dongzhe Han[1], Huangmeng[1],Liushuai[1],Wangze Lu[1]

[1]Software Engineering Technology Research Center, School of Information Engineering, Institute of Disaster Prevention, Hebei, China

Key Points:

- We designed a deep learning network for SAR image denoising.
- We combined SAR image denoising and change detection techniques for flood disaster localization.
- We constructed a dataset for flood change detection using SAR images and a dataset for SAR image denoising.

Corresponding author: Yangguang, ygcmd@163.com

Abstract

Flood disasters are among the most devastating natural disasters worldwide. The occurrence of such disasters is often accompanied by strong precipitation and other weather factors, making it more difficult to identify the affected areas. Moreover, Synthetic Aperture Radar (SAR) technology can capture images in 24-hour window and penetrate through clouds and fog. The change detection technology, based on SAR images, is generally utilized to locate disaster-stricken areas by analyzing the differences between pre- and post-disaster images. However, this method mainly faces two challenges: the presence of speckle noise reduces the accuracy of the difference detection and the lack of a suitable SAR dataset for flood disaster change detection. Therefore, this research proposes a novel two-stage approach for locating the flood disaster area, named Denoising-Change Detection Approach (D-CDA). The first stage consists of a nine-layer denoising network with an encoder-decoder structure, called SAR Denoising Network (SDNet). It utilizes a multi-residual block and parallel convolutional block attention module to extract features during the encoding process to suppress the noise component. In the second stage, a novel convolution neural network is proposed to detect the changes between bitemporal SAR images, namely Coordinate Attention Fused Network (CAFNet), which combines the Siamese network and UNet++ as the backbone and fuses multi-coordinate attention modules to enhance the change features. Moreover, a change detection dataset (ZhengZhou Flood - ZZF-dataset) is constructed using Sentinel-1 SAR images based on the flood disaster in Zhengzhou of China in 2021. The simulations verify the effectiveness of the proposed method. The experimental results indicate that the D-CDA achieves favorable detection performance in locating flood disaster areas.

1 introduction

Flood disasters often cause serious social and economic devastations (Munawar et al., 2022). After a flood event, accurate and immediate locating of the flood disaster-stricken area is very important for post-disaster rescue and relief efforts (Priyatna et al., 2023). Flood disaster localization typically poses a significant challenge due to the complex interplay of multiple factors, such as rain, clouds, and fog. Synthetic Aperture Radar (SAR) sensors possess the capability to capture images in 24 hours, penetrate through fog and clouds (Tebaldini et al., 2022), and acquire long-term and large-scale observation data, facilitating the monitoring of dynamic changes in water bodies and precise identification of water body areas. Therefore, SAR images are invaluable for emergency responders and disaster managers during crisis times.

In the past few years, many Change Detection (CD) methods were proposed to detect bitemporal SAR images, and these methods generally contain unsupervised learning methods and supervised learning methods (H. Chen et al., 2019). Most unsupervised learning methods usually consist of three steps: preprocessing, Difference Image (DI) generation, and DI classification with a focus on the last two steps. In the early stages of research, ratio-based methods were usually utilized to generate DI (Bazi et al., 2005), such as the Logarithmic ratio operator (Dekker, 1998), Gaussian ratio operator (B. Hou et al., 2014), and Neighborhood ratio operator (Gong et al., 2011). Moreover, several clustering algorithms were used for the DI classification, such as the K-means clustering algorithm (L. Liu et al., 2019), K-means++ clustering algorithm (Atasever & Gunen, 2021), Fuzzy C-Means (FCM) clustering algorithm (Kumar et al., 2020), etc... As for the flood events, they often result in the destruction and alteration of ground features, rendering traditional DI classification techniques ineffective in accurately demarcating the boundaries of inundated areas.

With the development of unsupervised deep learning, several convolutional neural network-based methods have been proposed to replace the traditional machine learning methods. For instance, Chen et al. proposed a parallel multi-scale spatial pooling Convolutional Neural Network (CNN) to exploit the changed information from the noisy DI (J.-W. Chen et al., 2020). Moreover, Gao et al. introduced a dual-domain network model for change detection of SAR images; the proposed approach combines spatial domain and frequency domain information to improve the accuracy of change detection. However, the mentioned methods were limited to the number of training samples, which affected the accuracy of the CD. Furthermore, references (H. Chen et al., 2022) and (Qu et al., 2021) generate false labels using the FCM clustering algorithm to increase the sample size based on DI and further improve the efficiency of change detection. The use of DI as a benchmark in CD methods, based on SAR images, weakens the preservation of the semantic information within the image, resulting in blurred boundaries and diminished accuracy of CD outcomes.

Most supervised learning methods, based on accurately labeled samples, usually obtain more change information than unsupervised learning methods; hence, they are more suitable for the accurate detection of change areas (Dong et al., 2020)(Gao et al., 2021). Therefore, many scholars proposed supervised CD methods based on SAR. For example, Lanka et al. proposed a CD method for mapping the extent of drought in a lake using supervised classification; thus, pre-processing, thresholding, segmentation, and Random Forest Classification (RFC) steps were applied to estimate the decay in water levels (Lanka & Puli, 2023). Moreover, Ma et al. proposed an approach based on multi-grained cascade forest and multi-scale fusion for SAR images. This approach detects the changed and unchanged areas of the images by using the well-trained multi-grained cascade forest (Ma et al., 2019). With the development of deep learning, many scholars have focused their research on the field of deep learning to improve the accuracy of CD. For instance, Li et al. proposed a non-smooth Nonnegative Matrix Factorization (nsNMF) nonlinear network to build for learning hierarchical, nonlinear, and localized data representations. Meanwhile, an extreme learning machine is integrated into the nonlinear nsNMF model to construct a deep nsNMF network for satisfactory classification (H.-C. Li et al., 2020). For instance, Wang et al. proposed a supervised Principal Component Analysis Network (PCA-Net) method, the pixels near the boundary between two classes are exploited to guide network training (R. Wang et al., 2018). Moreover, paper (Samadi et al., 2019) proposed a supervised deep belief network acting as the architecture and provided a dataset with an appropriate data volume and diversity. The network used the input images and the morphological operator results of the images to train the deep learning model to detect the changes in SAR images.

Based on the study of the aforementioned techniques, several researchers employed CD methods in SAR images to detect water expansion and submerged areas caused by floods (Abijith & Saravanan, 2022). For example, Li et al. integrated the fully-connected conditional random field model with long-range pairwise potential connections for the detection of flood disasters based on Sentinel-1 SAR images (Y. Li et al., 2018). In addition, He et al. constructed a flood dataset based on optical images of Sentinel-2 and the SAR images obtained from Sentinel-1, and proposed a cross-model change detection network to extract the flood area (He et al., 2023). Furthermore, Zhao et al. employ dual-time SAR images for detecting changes in flood disasters and used a transfer learning framework to improve the detection accuracy for flood-affected areas (Zhao et al., 2023a). However, the system-inherent speckle noise in SAR images superimposed the disturbing salt-and-pepper pattern on the real information, which had an undesirable effect on locating disaster areas. Therefore, in the task of locating flood disasters, many studies focused on denoising first and carrying out CD subsequently to improve the accuracy of disaster area positioning (Dong et al., 2022)(Duan et al., 2021).

In recent years, denoising methods primarily are categorized into three types: Spatial Domain Filtering-based Methods (SDFM), Frequency Domain Filtering-based Methods (FDFM), and Deep Learning-based Methods (DLM) (Thakur & Maji, 2022b). SDFM is relatively simple because it processes the grayscale values of the image directly. So, it has been widely used in early SAR image denoising, such as the Frost filter (Frost et al., 1982a), Lee filter (Frost et al., 1982b), Kuan filter (Kuan et al., 1985), and SAR-BM3D (X. Liu et al., 2020). However, SDFM is limited by window size generally, which results in blocking effects and artifacts in the denoised image. To solve these problems, FDFM is introduced to suppress the speckle noise of SAR images through inverse transformation (Parrilli et al., 2011),(S. Q. Liu et al., 2014). Therefore, Liu et al. (S. Liu et al., 2017) converted speckle noise to additive noise using shearlet transform. Reference (Jakka et al., 2019) used discrete wavelet transforms for image fusion and applied an improved gray wolf optimization algorithm for an adaptive optimization of the coefficients. In papers (X. Zhang et al., 2022) and (Chunhua & Fangchao, 2022), an improved wavelet denoising algorithm, based on fast non-local mean filtering and a threshold function, is proposed to improve the effectiveness of denoising. As a result, FDFM performs better in preserving image edge details than SDFM. However, after frequency domain transformation and inverse transformation, information loss often occurs, reducing theregore the denoising accuracy.

With the development of remote sensing technology and deep learning method, much more scholars apply CNNs to denoise SAR images. In paper (Chierchia et al., 2017), a novel SAR denoising network deployed logarithmic and exponential transformations in residual CNNs to denoise SAR images. Paper (Thakur & Maji, 2022a) proposed an attention and gradient-based SAR denoising network to remove speckle noise from SAR images and preserve finer details. The gradient information of the noisy image was concatenated with its features, and then an intermediate feature denoising block and two attention blocks were employed to reduce the noise in the feature mapping. Furthermore, reference (Ko & Lee, 2021) proposed a continuous attention module for speckle denoising and usesd a data-driven approach with the gradient descent algorithm to train, where the context blocks were deployed at the minimum scale to capture multi-scale information. In addition, paper (Z. Liu et al., 2020) proposed a spatial and transform domain CNN, which applies a spatial channel attention module to enhance the boundaries between noise-related and non-noise-related features to improve denoising performance. Finally, in reference (S. Liu et al., 2021), the Multiscale Residual Dense Dual Attention Network (MRDDANet) with multi-scale modules of different kernel sizes was presented to extract shallow features from noisy images and map them into a residual dense dual attention network to obtain the deep features of SAR images.

Motivated by the influence of denoising methods, many scholars combine denoising and change detection technology to locate flood-affected areas. For instance, reference (Cao et al., 2019) designed an end-to-end deep denoising model to remove the noise of SAR images and utilized a three-layer CNN to divide the denoised image into change and unchanged regions. Moreover, paper (J. Wang et al., 2022) proposed two parallel paths to denoise images, one path dedicated to estimating the noise distribution of images whereas the other focused on identifying the relationship between noise and signal. Then, the proposed method used a multi-region convolution module to emphasize the changed area at two-time points and obtain the CD results. To sum up, all mentioned methods have effectively improved the accuracy of change detection by applying denoising. Nevertheless, these methods lack sufficient CD training data.

Inspired by the above research, we propose, in this paper, a two-stage approach to locate the flood-stricken area that consists of denoising and CD, named Denoising-Change Detection Approach (D-CDA). In the denoising stage, a nine-layer encoder-decoder network is presented to suppress the speckle noise of SAR images, called SAR Denoising Network (SDNet). The Multi-Residual Block (MRB) and Parallel Convolutional Block Attention Module (P-CBAM) are designed in the encoder branch to extract features, focusing on the noisy and non-noisy information. In the CD stage, a novel deep learning network is described to locate the flood-affected areas by detecting the changes between the pre-disaster and the post-disaster denoising images, called Coordinate Attention Fused Network (CAFNet). It combines the Siamese network and UNet++ as the backbone to enhance the detection efficiency, employs the Coordinate Attention (CA) modules (Q. Hou et al., 2021) to focus on the boundary of the flood-stricken areas in the encoder part, and employs

Multi-level Damage-feature Aggregation Attention Module (MDAAM) (Y. Zhang et al., 2023) to capture global, diversity, and similarity information of the multilevel features. The contributions of our proposed method can be summarized as follows:

1. We propose a two-stage approach D-CDA for locating disaster-stricken areas. In fact, D-CDA incorporates a denoising network and change detection network to remove the unwanted noise and improves therefore the locating accuracy;
2. The denoising network SDNet employs a simple nine-layer encoder-decoder structure that combines MRB and P-CBAM to extract the ground objects and reduce the noise impact using spatial and channel features of SAR images;
3. The change detection method CAFNet aggregates the siamese network, UNet++, CA modules, and MDAAM to extract change information caused by flood disasters and express the irregular edges of flood-affected areas;
4. A ZhengZhou Flood dataset (ZZF-dataset) is simulated based on Sentinel-1 SAR images to verify the performance of the proposed approach. The database contains 1900 pairs of images and covers various landmarks such as cities, bridges, rivers, beaches, etc. . .

2 Dataset

2.1 ZZF-dataset

ZhengZhou of China's Henan Province experienced flooding on July, 20th, 2021 resulting from heavy rainfall. The flood disaster caused 398 deaths, mountain floods, landslides, serious urban waterlogging, and significant damage to subways and underground spaces (Zhao et al., 2023b)(Vekaria et al., 2022). Moreover, the Zhengzhou flood disaster was characterized by a wide range of impact, a large affected area, and obvious flood features, which facilitated CD and flood mapping (Peter et al., 2020). Therefore, this paper collects the SAR images of the ZhengZhou flood based on the Level-1 Ground Range Detected (GRD) image of Sentinel-1 to construct the CD dataset. Compared to the existing images, we select one image 5 days before the flood and one image 7 days after the flood as the pre-disaster and post-disaster images, respectively. The information on the dual-temporal SAR images is shown in Table 1.

Table 1. Data Information

	Pre-temporal image	Post-temporal image
Satellite	Sentinel-1	Sentinel-1
Mode	IW	IW
Instrumen	SAR-C	SAR-C
Product type	GRD	GRD
Spatial resolution	10m	10m
Product level	L1	L1
Ingestion Date	7/15/21 13:57	7/27/21 12:15
Mission datatake id	299935	301281
Upper Left Accuracy	113°07E ,36°22N	113°08E ,36°22N
Lower Right Accuracy	116°18E ,35°12N	116°18E ,35°12N

To obtain a dataset suitable for detecting changes to locate the flood-affected area, the dual-temporal images are first registered using OpenCV, and then they are utilized to generate the Difference Image (DI) by the logarithmic ratio operator to highlight the changing areas, which gives a preliminary observation of flood disaster-stricken. Then, a rigorous process of evaluation is adopted and the DI images are reviewed to determine the edge areas of the affected regions. Finally, Labelme software is used to draw the polygons of the changing areas to generate the label map.

To confirm the quality and reliability of the dataset, we invite five domain experts to evaluate it in terms of content, quality, and relevance to research questions. They were also asked to make subsequent modifications and improvements based on their feedback and suggestions. Meanwhile, an in-depth internal review was performed to ensure the accuracy, consistency, and usability of the dataset. Ultimately, after multiple rounds of evaluation and revision, we successfully label the flood-affected areas. Specifically, the five experts that were invited had a large knowledge in the field, including two professors with over ten years of experience in related fields, two senior researchers with over eight years of research experience in the field, and one technical expert with rich experience in dataset construction and evaluation. To sum up, the labeling process of changes in bi-temporal SAR images is shown in Figure 1 where the white regions represent the flooded areas, and the black regions indicate the non-flooded areas in the label image.

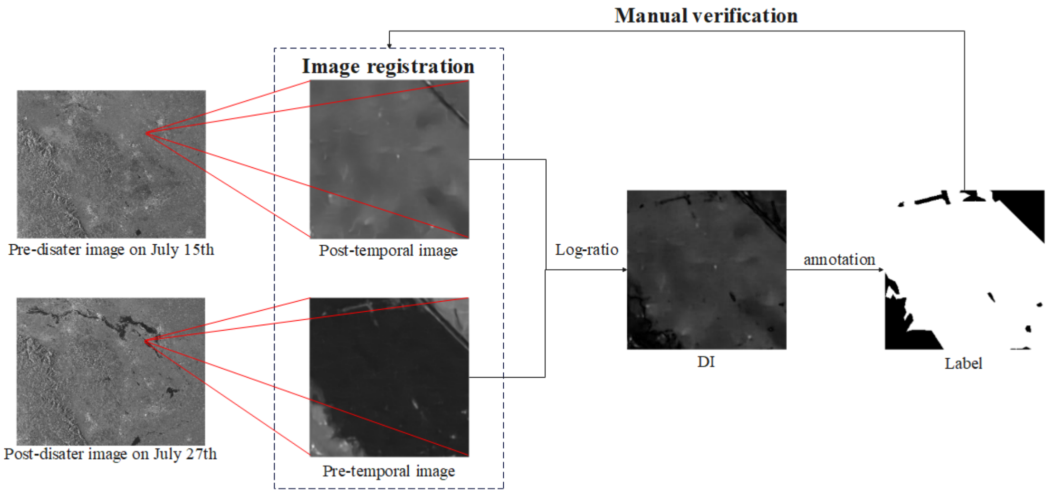

Figure 1. The generation of label image in ZZF-dataset

Moreover, the labeled pre- and post-disaster images are cut into 5000 pairs of 32×32 pixel small size with a certain repetition ratio between small scales by the sliding window. To enlarge the image number, the small images were handled by rotation and flipping adjustment whereas the ZZF-dataset contains 20000 pairs of images where almost 6000 pairs include affected area and the remaining (about 14000) cover various unaffected landmarks such as cities, bridges, rivers, beaches, etc... Moreover, the dataset is divided into a train set, test set, and validation set in a 3:1:1 ratio. Figure 2 shows some representative images of the ZZF-dataset.

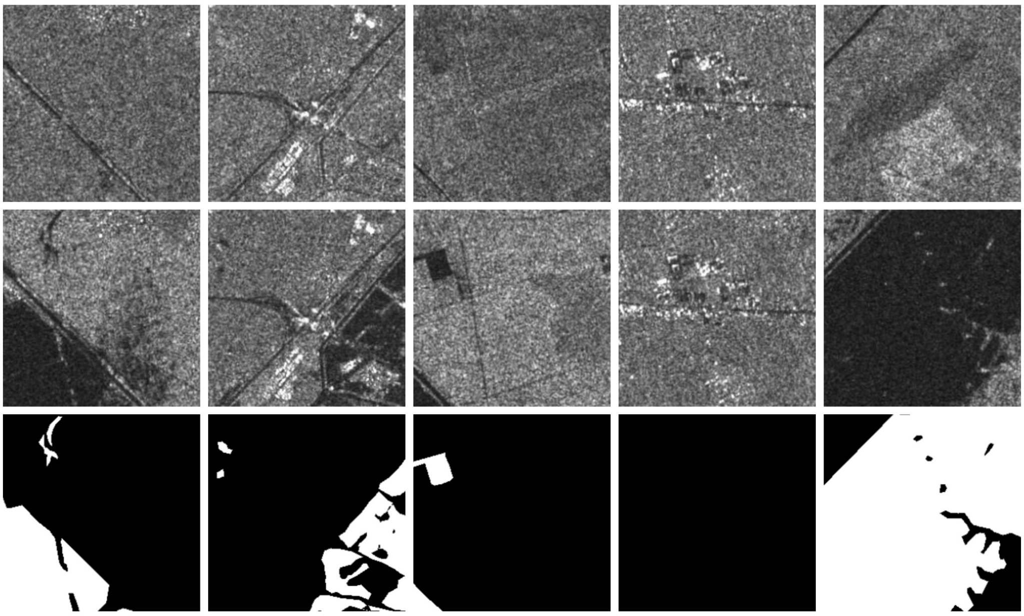

Figure 2. The representative images of ZZF-dataset. The first row are pre-disaster images, the second row are post-disaster images and the third row are label images.

2.2 Simulated denoise dataset

To train and evaluate the proposed denoising method, a dataset with denoising labels is constructed. Considering the difficulty in obtaining effective denoising labels for SAR images, the UC Merced Land Use dataset (abbreviated as UC dataset) (Yang & Newsam, 2010) is an optical remote sensing image dataset to simulate the denoise dataset. It consists of 2100 images, having mainly a size of 256x256 pixels and containing lots of ground objects that are easily affected in flood disasters, such as streams, rivers, bridges, waterfront buildings, cities, farmland, etc...

The construction procedure for the simulated denoise dataset includes four steps: First, we randomly select 1900 images from the UC dataset including streams, rivers, bridges, beaches, and buildings to simulate noise. Then, the high-resolution RGB image in the UC dataset is transformed into a grayscale image with a 10-meter resolution as a label image. To produce this dataset, we take a similar approach to the one presented in (Ko & Lee, 2021), where the simulated speckle noise, formed by the Gamma function, is added to the transformed image to simulate the SAR image as the original one. Finally, the original and labeled images are transformed with cutting, rotation, flipping, contrast adjustment, etc... to enlarge the size of the simulated denoise dataset with 30400 pairs of 64×64 pixel images. Furthermore, the noise formula is constructed as follows:

$$Y = N \cdot X \tag{1}$$

where X is the optical image (denoising), Y is the simulated image, and N is the noise component. If N follows a Gamma distribution (Dong et al., 2020), with a unit mean and variance of 1/L, its probability density function would be defined as follows:

$$p(N) = \frac{L^L N^{L-1} e^{-LN}}{\Gamma(L)}, N \geq 0, L \geq 1 \tag{2}$$

where L represents the number of lines of sight in the multi-view processing. In this method, L is set to four and represents the Gamma function. Figure 3 is a set of example images.

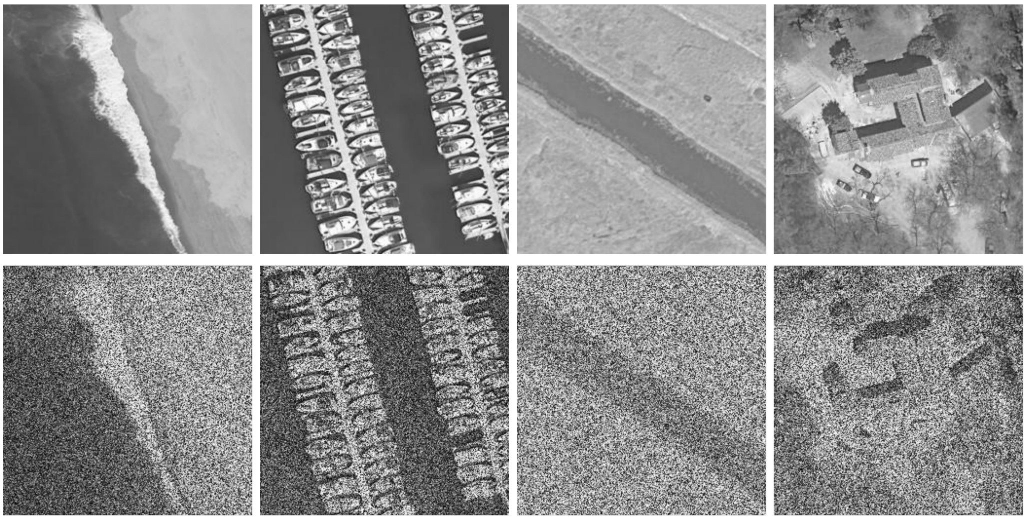

Figure 3. The representative images of the simulated noise image. The top row shows four optical images, and the second row displays the corresponding simulated SAR images.

3 Method

In this paper, we propose a two-stage approach to detect flood-stricken areas, named the D-CDA. In the first stage, a nine-layer encoder-decoder SDNet is proposed to suppress the speckle noise by extracting flood information using MRBs and P-CBAM. In the second stage, we design a CAFNet based on the Siamese network, UNet++, CA, and MDAAM to distinguish between flood pixels and non-flood pixels.

3.1 SDNet

To locate flood-stricken areas in SAR images, speckle noise seriously affects the locating accuracy and efficiency (Cao et al., 2019). Therefore, a denoising network SDNet is developed to reduce the adverse effects of noise. SDNet adopts a typical nine-layer encoder-decoder structure as a backbone, utilizes convolution block, MRB, and long skip connection to extract and concatenate features, and employs P-CBAM to focus on both channel and spatial information of the feature maps to suppress noise and preserve real pixel. The architecture of SDNet is displayed in Figure 4, and the details are shown in Table 2.

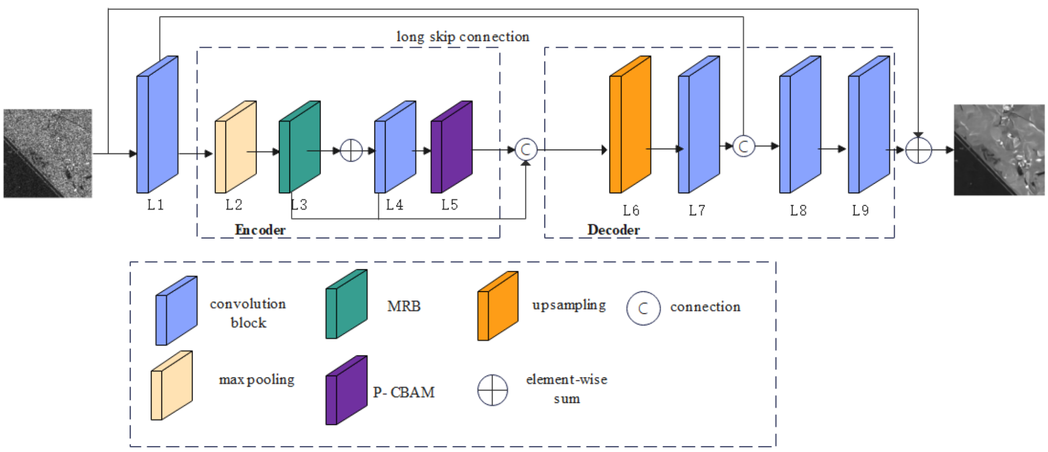

Figure 4. The architecture of SDNet

The encoder part of the SDNet utilizes a typical 3×3 convolution block to extract 128 shallow feature maps and employs a simple max-pooling layer with a size of 2x2 (stride of 1) to reduce the size of the feature map. Then, this latter interacts with our designed MRB to better capture the flood information. Moreover, the P-CBAM is designed to extract both channel and spatial features for suppressing the noise component to restore real information.

In the decoding stage, the feature maps are initially restored to the input size, followed by the usage of three convolution layers to convert feature maps into denoised images. To preserve important features in the reconstruction process and extract deep-level features, a long skip connection is added between the encoder and decoder process. This helps the network learn better feature relationships between its different layers, thereby better capturing the correlations between them to provide better speckle noise removal.

3.1.1 Multi-Residual Block

In the decoding stage, the feature maps are initially restored to the input size, followed by the usage of three convolution layers to convert feature maps into denoised images. To preserve important features in the reconstruction process and extract deep-level features, a long skip connection is added between the encoder and decoder process. This helps the network learn better feature relationships between its different layers, thereby better capturing the correlations between them to provide better speckle noise removal.

In SDNet, the objective is to enhance the effect of feature maps to suppress the influence of noise on SAR images and to design a residual structure as shown in Figure 5, which consists of a Residual Channel Attention Block (RCAB) (Xie et al., 2019) module to focus more on high-frequency features while using a 3×3 convolutional block RCBA as shown in Figure 6. Therefore, the equation of the multi-residual blocks is represented as follows:

$$MRB\left(F_{\mathrm{RB}}\right) = F_{\mathrm{MRB}} + \left(W_{3\times 3} * RCAB\left(F_{\mathrm{MRB}}\right) + b\right) \tag{3}$$

where the convolution block consists of a 3x3 convolution layer with a bias variable b, the operator represents the convolution operation, and MRB represents a multi-residual block.

Table 2. The details of SDNet. Input size and output size are in the form of C×W×H, where C is the number of channels, W and H are the width and height of the image, respectively.

	Layer	Blocks	Input size (C×W×H)	Output size (C×W×H)
ENCODER	L1	Convolution (3×3)	1×64×64	128×64×64
	L2	Max-pooling (2×2)	128×64×64	128×32×32
	L3	MRB	128×32×32	128×32×32
	L4	Convolution (1×1)	128×32×32	128×32×32
	L5	CBAM	128×32×32	128×32×32
DECODER	L6	Up-sampling	128×64×64	128×64×64
	L7	Convolution (3×3)	128×64×64	128×64×64
	L8	Convolution (1×1)	128×64×64	128×64×64
	L9	Convolution (3×3)	128×64×64	1×64×64

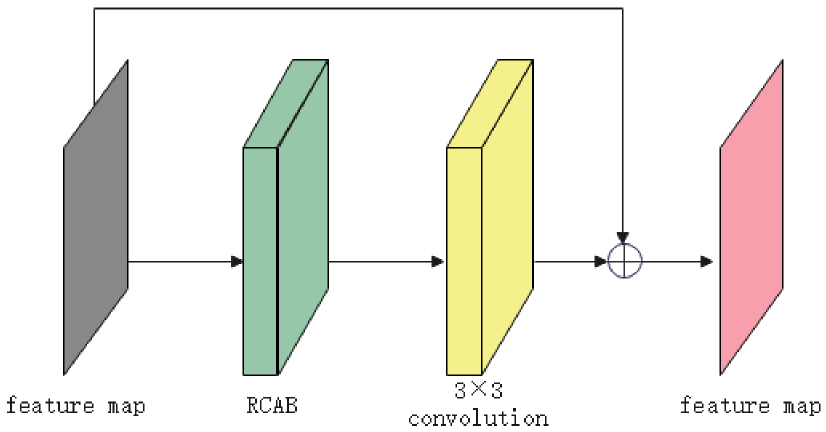

Figure 5. The structure of MRB

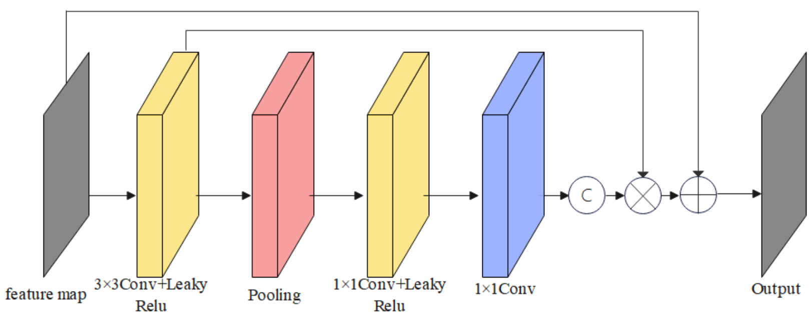

Figure 6. The structure of RCAB

3.1.2 P-CBAM

To focus on flood feature information and suppress speckle noise, this paper designs an attention module as shown in Figure 7. The input feature maps are simultaneously inserted into the Channel Attention Mechanism (CAM) and the Spatial Attention Mechanism (SAM). The feature maps obtained from both the CAM and SAM are multiplied to obtain the spatial weights, which reduce the location differences between the feature maps at different levels, focusing on the important regions and features, and reducing the influence of noise and interference.

On one hand, the SAM performs the weighting process on the feature maps in the spatial dimension, giving higher weights to focus on flood information that has a larger proportion in the calculation of the overall feature representation; thus, it ignores irrelevant noise information in the input image. On the other hand, CAM performs the weighting process on the feature maps in the channel dimension, highlighting useful feature maps and reducing the impact of useless feature maps on network attention. Moreover, it focuses on extracting important feature information while considering optimal positions, enabling the network to pay more attention to critical details within the input image.

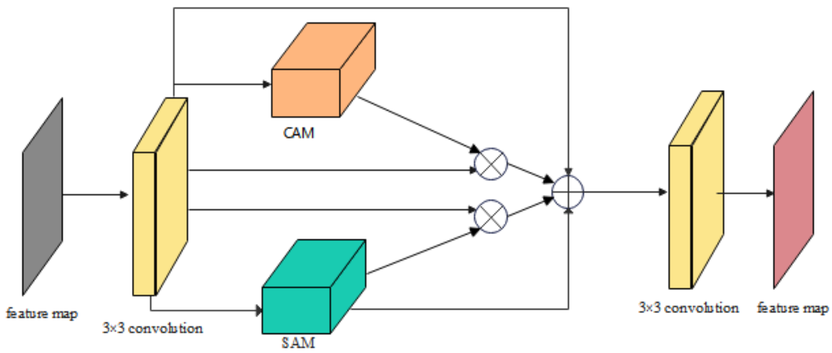

Figure 7. The structure of parallel CBAM

As a result, the SAM obtains the attention feature map by adaptively fusing the features within the local receptive field. This process helps the network to pay more attention to useful features while suppressing noise signals. The SAM structure is shown in Figure 8.

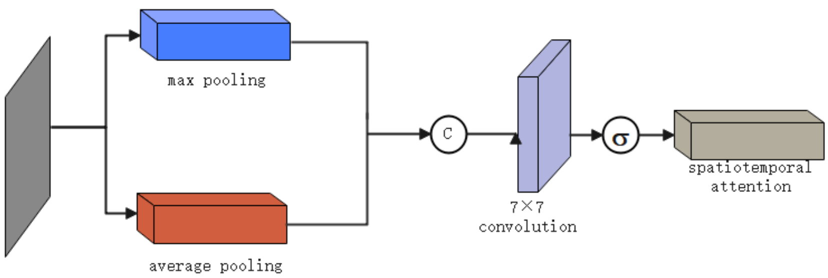

Figure 8. The structure of SAM

As for the CAM, it is used to extract features between different channels where the CAM structure is shown in Figure 9. This attention mechanism is integrated within the SDNet method, developed in this paper, without increasing the computational burden.

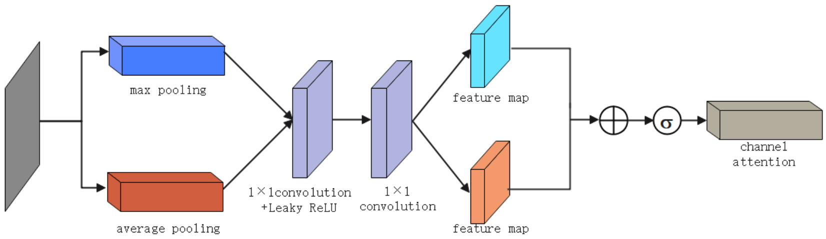

Figure 9. The structure of CAM

3.1.3 loss function

This paper adopts the loss function proposed in (P. Wang et al., 2017), which is formulated as follows:

$$L_{\mathrm{TV}} = \sum_{i=1}^{W}\sum_{j=1}^{H}\sqrt{\left(\hat{x}^{i+1,j}-\hat{x}^{i,j}\right)^2+\left(\hat{x}^{i,j+1}-\hat{x}^{i,j}\right)^2} \tag{4}$$

$$L_{\mathrm{DG}} = \log_{10}\left(\sum_{i=1}^{W}\sum_{j=1}^{H}\frac{\|\hat{x}_{i,j}-x_{i,j}\|_2^2}{\|y_{i,j}-x_{i,j}\|_2^2}\right) \tag{5}$$

$$L(\theta) = L_{\mathrm{DG}} + \lambda_{\mathrm{TV}}L_{\mathrm{TV}} \tag{6}$$

The loss function consists of two parts: the Total Variation (TV) loss (Lattari et al., 2019) and the Discriminative Gradient (DG) loss (Shen et al., 2020) where x and y represent the ground truth image and the denoised image, and the noise image, respectively. Moreover, i and j denote the pixel coordinates, and W and H represent the number of pixels in the width and the height dimensions, respectively. The value of is set to 2e-4.

In the nine-layer SDNet, long skip connections capture features across different layers which enhances the ability to discriminate noise information. Moreover, the encoder progressively reduces the spatial dimensions of the feature maps to capture global contextual information. In the decoder, when remapping the encoded features back to the original input space, leveraging the global contextual information is performed to better reconstruct the details and local the structures. The P-CBAM pays more attention to the flood information in SAR flood images and ignores noise. As the MRB combines low-level features with high-level features, it better handles noise problems in complex images and compensates for the shortcomings of shallow networks in removing image noise. Thus, MRB prioritizes flood information.

3.2 Change Detection Phase

In this paper, we design an encoder-decoder structure network by modifying the Siamese network and UNet++; this structure is named CAFNet. The CA architecture is shown in Figure 10 whereas the convolution block, used in the modification, is illustrated in Figure 11 and Table 3.

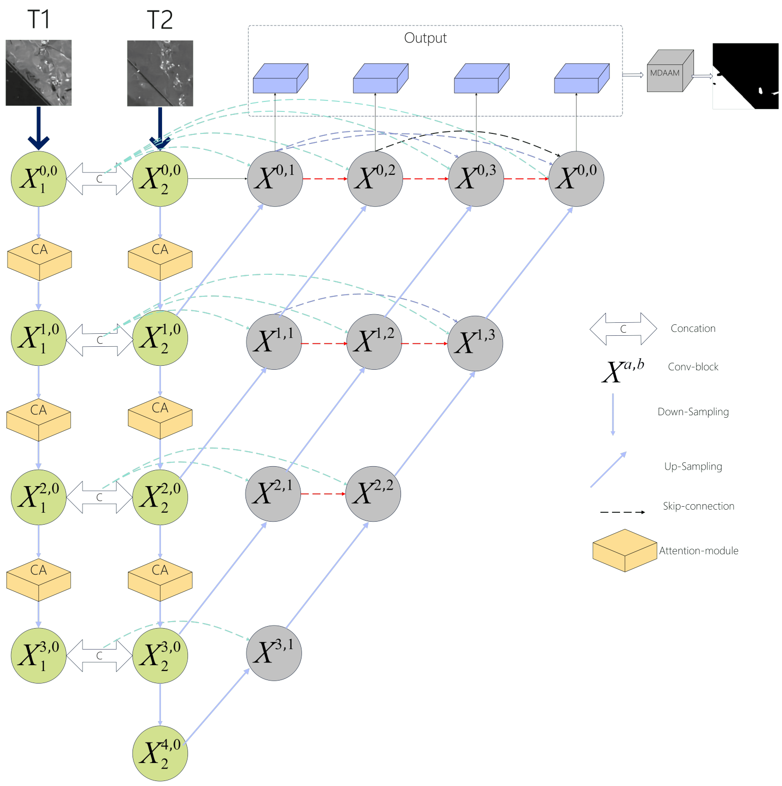

Figure 10. The structure of CAFNet

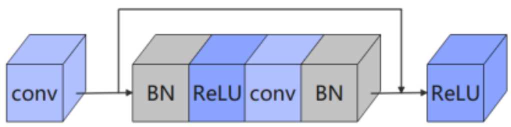

Figure 11. The structure of Convolution Block

The main goal of the encoding stage is to extract fuse feature maps from the input images. Due to the problem of sample imbalance in the ZZF-dataset, the change detection phase network takes two dual-temporal SAR remote sensing images, T1 and T2, as input and processes them through the two share-parameter branches. The T1 branch passes through four convolution block and three down-sampling while the T2 branch goes through five convolution block and four down-sampling operations. The convolution is used to extract features between the two images, helping to improve the robustness of the samples. A connection structure is used to fuse the two features between the convolution blocks of the two branches and outputs a fuse feature map to ensure information integrity.

Moreover, the SAR flood disaster areas exhibits strong spatial continuity and contains rich information. To capture the contextual relations in SAR images, the CA mechanism is used as it incorporates this contextual information into feature maps. CA uses two one-dimensional global pooling operations to aggregate the feature input along the vertical and horizontal directions, resulting in two separate direction feature maps. Then, these two-direction feature maps, with specific directional information, are encoded into two attention feature maps, each capturing correlations along a long spatial direction. Therefore,

Table 3. Details of CAFNet. Input size and Output size are in the form of C×W×H, where C is the number of channels of the images, W and H are the width and height of the image respectively.

	Layer	T1	T2	Input size (C×W×H)	Input size (C×W×H)
ENCODER	L1	Conv X1(0,0)	Conv X2(0,0)	1×32×32	32×32×32
	L2	CA	CA	32×32×32	32×32×32
	L3	Down-sample	Down-sample	32×32×32	32×16×16
	L4	Conv X1(1,0)	Conv X2(1,0)	32×16×16	64×16×16
	L5	CA	CA	64×16×16	64×16×16
	L6	Down-sample	Down-sample	64×16×16	128×8×8
	L7	Conv X1(2,0)	Conv X2(2,0)	128×8×8	128×8×8
	L8	CA	CA	128×8×8	128×8×8
	L9	Down-sample	Down-sample	128×8×8	128×4×4
	L10	Conv X1(3,0)	Conv X2(3,0)	128×4×4	256×4×4
	L11		Down-sample	256×4×4	256×2×2
	L12		Conv X2(4,0)	256×2×2	512×2×2
DECODER	L13		up-sample	512×2×2	512×4×4
	L14	Conv X(3,1)		512×4×4	256×4×4
	L15	up-sample		256×4×4	256×8×8
	L16	Conv X(2,2)		256×8×8	128×8×8
	L17	up-sample		128×8×8	128×16×16
	L18	Conv X(1,2)		128×16×16	64×16×16
	L19	up-sample		64×16×16	64×32×32
	L20	Conv X(0,0)		64×32×32	32×32×32
	L21	MDAAM		32×32×32	1×32×32

the location information is preserved in the generated attention feature map. Finally, the two attention feature maps are merged into one after multiplication. The structure of the CA attention mechanism is shown in Figure 12.

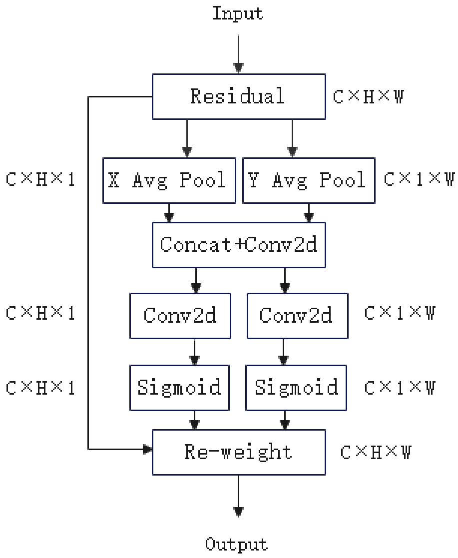

Figure 12. The structure of the CA

The CA module captures in a better way the spatial and texture information of flood disaster areas to refine the irregular edge information at the boundary between the changing and non-changing areas and improve the accuracy of change detection. In addition, since the CA module adaptively learns the relationships between different positions, it improves the robustness of the model and makes the network more adaptable to SAR flood disaster images.

In the decoding stage, the feature maps are restored using the same size through continuous up-sampling. To maintain the same resolution of the feature maps as well as the integrity of the detailed information, this paper adds a dense skip connection mechanism between the encoder and decoder. After down-sampling the dual-temporal images, the fuse features are transmitted to the decoder through skip connections to compensate for the loss of the deep features in the decoder. Finally, after passing through the decoder, the network generates four feature maps with different semantic levels, having the same size as the original image, and the prediction results are delivered through MDAAM to automatically integrates and refine the global contextual semantic and localization information in different-scale feature layers. While changing the detection task, the number of unchanged pixels is often far greater than the number of changed pixels. Furthermore, to reduce the impact of sample imbalance, this paper adopts a mix loss function in the change detection phase network; this latter is formulated as follows:

$$\mathcal{L} = \mathcal{L}_{wce} + \mathcal{L}_{dice} \tag{7}$$

$$\mathcal{L}_{wce} = \frac{1}{\mathrm{N}} \sum_{\mathrm{P}} \left[-\mathrm{w_c R_p} \log\left(\mathrm{P_p}\right) - \mathrm{w_u}\left(1 - \mathrm{R_p}\right) \log\left(1 - \mathrm{P_p}\right) \right] \tag{8}$$

$$\mathcal{L}_{dice} = 1 - \frac{2 \cdot Y \cdot softmax(Y)}{Y + softmax(Y)} \tag{9}$$

where represents the weight cross-entropy loss, and is the Dice coefficient loss. In addition, and represent the changed and unchanged classes, respectively. Finally, and indicate the predict and ground truth labels of a pixel (i, j).

To detect changes in the SAR flood disasters image, CAFNet utilizes the CA attention mechanism to improve the accuracy of the change detection by focusing on the relationship between the positions in the input images. This method is applied to help the network in highlighting important features in the adoption of a position-aware approach to feature learning. Moreover, the CA attention mechanism explores the spatial correlation among different regions to effectively improve the representation of the input data and enhance the accuracy of the change detection, especially for complex scenes with small variations.

4 Experiment

4.1 experimental environment

This paper conducts experiments on an NVIDIA 3090 GPU using the PyTorch framework to train and test the models. The operating environment for both stages of the network is CUDA 11.4 and PyTorch 1.7. In the SDNet, the batch size is set to two, and the Adam optimizer is used with an initial learning rate of 1e-4. The learning rate decreases by 1e-5 per iteration, and the total number of iterations is set to 20. In the localization stage network, the batch size is set to 16 whereas the Adam optimizer is used with an initial learning rate of 1e-4. The learning rate decreases by 5e-5 per iteration, and the total number of iterations is set to 50.

4.2 evaluation criterion

In the denoising stage, this paper use two evaluation metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The formulas for calculating these metrics are as follows:

$$PSNR = 10 \cdot \log_{10} \frac{(x_{\max})^2}{MSE(x,\hat{x})} \tag{10}$$

$$SSIM = \frac{(2\mu_{\hat{x}}\mu_x + c_1)(2\sigma_{\hat{x}x} + c_2)}{(\mu_{\hat{x}}^2 + \mu_x^2 + c_1)(\sigma_{\hat{x}}^2 + \sigma_x^2 + c_2)} \tag{11}$$

where xˆ and x are the denoised image and the label image; xmax is the maximum value of x; xˆ and x are the mean of xˆ and x; xˆ and x are the standard deviations of xˆ and x; ˆxx is the covariance between xˆ and x; and c1 and c2 are constants for computational stability.

The evaluation standards, used in the localization stage experiment through this work, mainly include six parts: mean Intersection over Union (mIoU), Precision, Recall, F1 score, Kappa coefficient, and Pixel-wise Correct Classification (PCC). Precision refers to the ratio of true positive samples among all samples predicted as positive. Its formula is as follows:

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{12}$$

The recall represents the proportion of true positive samples among all real positive samples. It is calculated as follows:

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{13}$$

The mIOU is the overlap between two regions, represented by the following equation:

$$mIOU = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} \tag{14}$$

The F1 score is a metric that considers both precision and recall and consists of the harmonic mean of precision and recall. Moreover, it comprehensively considers the relationship between both variables, and a higher value indicates a more effective experimental method. It is calculated as follows:

$$F1 = \frac{2\mathrm{Precision} * \mathrm{Recall}}{2\mathrm{Precision} + \mathrm{Recall}} \tag{15}$$

The PCC represents the percentage of correctly classified pixels in the overall pixel count in CD. A higher value indicates a higher accuracy in change detection, and it is calculated as follows:

$$PCC = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN}} \tag{16}$$

The Kappa coefficient is a metric used for consistency tests; it measures whether the model's predicted results are consistent with the actual classification results. Its calculation is based on the confusion matrix and ranges between -1 and 1 (it is usually greater than zero). Its expression is as follows:

$$Kappa = \frac{\mathrm{P_o} - \mathrm{P_e}}{1 - \mathrm{P_e}} \tag{17}$$

where TP represents the true positive samples, TN represents the true negative samples, FN represents the false negative samples, and FP represents the false positive samples. Moreover, PCC is the sum of the correctly classified samples for each class, divided by the total number of samples, which represents the overall classification accuracy. represents the proportion of agreement, which indicates the percentage of algorithmic predictions that match the results of the manual annotations. Finally, Pe is the probability of consistency in CD results.

4.3 experimental result

In the denoising stage, experiments are conducted to compare our proposed method with U-Net (Lattari et al., 2019), the Spatial and Transform Domain Convolutional Neural Network (STD-CNN) (Z. Liu et al., 2020), and the SAR-CAM (Ko & Lee, 2021), as shown in Table 4. Referring to this table, our proposed method achieves a higher PSNR and SSIM for denoised images where the values are 28.89 and 0.836, respectively, compared to the other methods.

Table 4. Comparison of denoising methods

Method	PSNR (dB)	SSIM
U-Net	24.95	0.4955
STD-CNN	28.29	0.7867
SAR-CAM	28.6	0.8092
Ours (SDNet)	28.57	0.8307

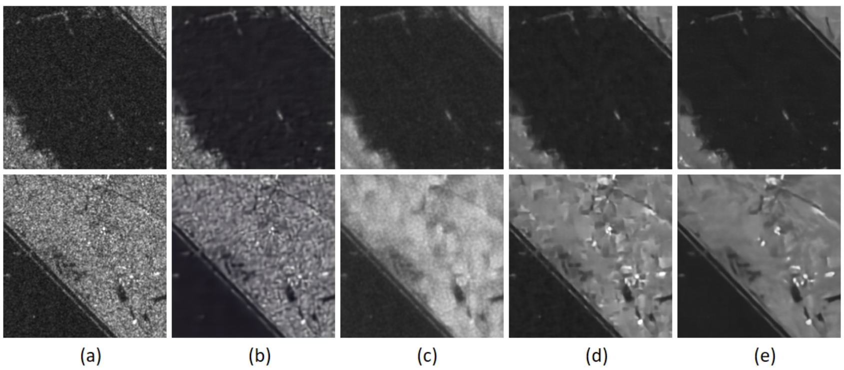

Figure 13. Experimental Comparison result in SAR image . From left to right: (a) original image, (b) U-Net, (c) STD-CNN, (d) SAR-CAM, (e) Our proposed network.

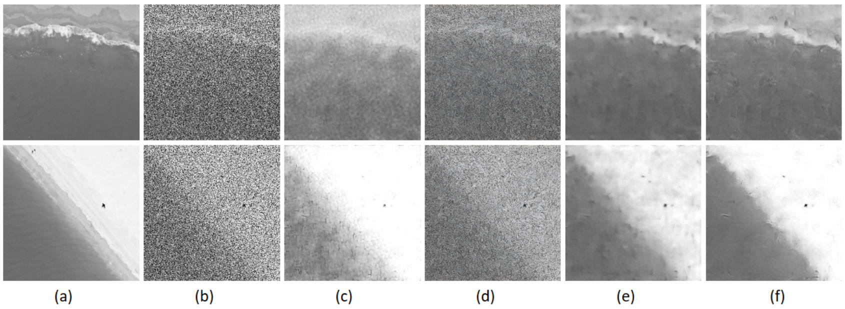

Figure 14. Experimental Comparison result in simulated noise dataset. From left to right: (a) original image, (b)noise image, (c) U-Net, (d) STD-CNN, (e) SAR-CAM, (f) Our proposed network.

In the SDNet, PSNR achieve 28.57 and SSIM achieve 0.8307. Compare with the aforementioned methods, the denoise images by SDNet are smoother and effectively refine the edge information of flood areas while maintaining denoising performance. And highlights the flood characteristics within the images. This indicates that the MRB and parallel CBAM structures in SDNet can effectively extract flood-related information while disregarding noise. We demonstrate the effectiveness of the proposed SDNet in this study by perform the aforementioned comparative experiments. Experimental comparison result as shown in Figs. 13-14.

We introduce to verify the effectiveness of adding a denoising stage before the localization stage and incorporating the CA mechanism for change detection tasks, this paper conduct ablation experiments. The experiments include backbone, SDNet+backbone, CAFNet and D-CDA, as shown in Table 5.

Table 5. Details of CAFNet. Input size and Output size are in the form of C×W×H, where C is the number of channels of the images, W and H are the width and height of the image respectively.

Method	Precision	Recall	F1	PCC	Kappa	mIOU
backbone	0.9031	0.8920	0.8975	0.9590	0.8719	0.8041
SDNet+backbone	0.9114	0.9023	0.9072	0.9634	0.8843	0.8224
CAFNet	0.8965	0.9174	0.9068	0.9620	0.8830	0.8488
D-CDA	0.8935	0.9310	0.9120	0.9681	0.8860	0.8732

The mIOU of the backbone reaches 0.8041, indicating that the change detection backbone is suitable for locating flood-disaster-stricken areas. Moreover, the addition of the CA mechanism to the backbone, the proposed CAFNet, and the mIoU increases by 0.0447. This indicates that the CA mechanism effectively integrates the positional information of the pixels to improve the accuracy of locating flood disaster-stricken areas. More importantly, after incorporating the denoising stage, the overall localization accuracy of F1 improves by 0.0154, and mIOU improves by 0.0691. This demonstrates the effectiveness of the proposed two-stage approach for flood disaster-stricken areas localization.

We aim to validate the superiority of the proposed method as we compare it with DDNet (Qu et al., 2021) and SAFNet (Gao et al., 2021) methods in the experimental section. In addition, this method aims to verify the effectiveness of the SDNet; therefore, we add SDNet before DDNet and SAFNet. Subsequently, we compare the results with our proposed method. Since the aforementioned deep learning methods are designed for change detection of a single pair of SAR images, we use the test sets of the constructed dataset as input, and the experimental results are shown in Figure 15.

Table 6. Comparison of different methods on the ZZF-Dataset

Method	Precision	Recall	F1	PCC	Kappa	mIOU
DDNet	0.7932	0.7925	0.7929	0.8666	0.5990	0.6070
SDNet+DDNet	0.7933	0.9008	0.8436	0.8587	0.6774	0.7341
SAFNet	0.9365	0.6195	0.7469	0.7773	0.831	0.5113
SDNet+SAFNet	0.8983	0.8315	0.8638	0.9122	0.8032	0.6989
Ours (CAFNet)	0.8965	0.9174	0.9068	0.9620	0.8830	0.8488
Ours (D-CDA)	0.8935	0.9310	0.9120	0.9681	0.8860	0.8732

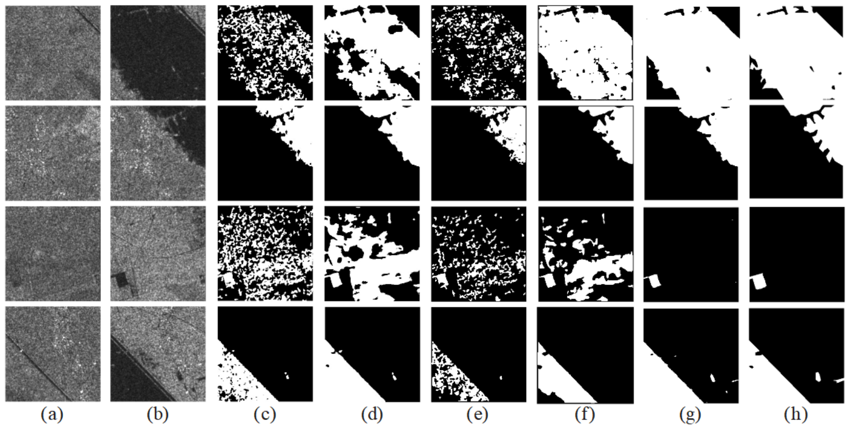

Figure 15. Experimental Comparison result. From left to right: (a) predisaster image, (b) predisaster image, (c) ground truth, (d) DDNet, (e) SDNet+DDNet, (f) SAFNet, (g) SDNet+SAFNet, (h) Our proposed method.

From the experimental results displayed in Table 6, it is seen that the method proposed in this paper improves the various evaluation metrics when been compared to DDNet and SAFNet methods. After adding the denoise stage design, proposed in this paper, to the DDNet and SAFNet networks, the final results have a significant improvement in the recall rate, F1 score, and IOU. In more detail, in the D-CDA proposed in this article, F1 reaches 0.9120, PCC reaches 0.9681, Kappa reaches 0.8860, and mIOU reaches 0.8732. Compare to other methods, D-CDA has a significant improvement. Therefore, it is concluded that the denoising stage, designed in this paper, has a good effect on improving the image quality and removing speckle noise, and the two-stage network, proposed in this paper, has an excellent performance in flood disaster detection accuracy.

From the experimental results displayed in Figure 16, for the first and second row, the DDNet (d) and SAFNet (f) without the SDNet are directly affected by speckle noise, resulting in incomplete detection of flooded areas and many false and missed detections. After adding the SDNet to these methods (g)(e), the detection of disaster areas becomes more regular and continuous. In the third row, due to the different divergence of the gradients in the first-phase SAR images, i.e., the different distribution of gray value changes between pixels in the SAR images and undamaged areas may be falsely classified as damaged areas. In the fourth row, the results detected by our proposed method are more continuous and a more complete detection of flood disaster-stricken areas is obtained. Therefore, the denoise stage improves the image quality and reduces the impact of noise on subsequent detection. Moreover, our proposed method has strong noise resistance and anti-interference ability, which improve the accuracy of CD tasks. Finally, the loss curve of the CAFNet in our proposed method is shown in Figure 16.

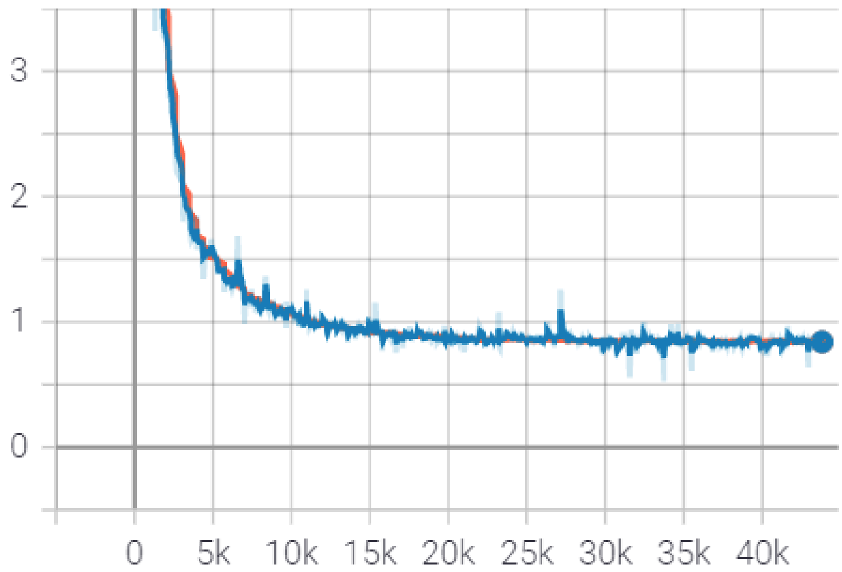

Figure 16. Loss curve of CAFNet.

5 Conclusion

In this study, we propose a two-stage approach for locating flood disaster-stricken areas in Sentinel-1, named D-CDA. In the first stage of this approach, SDNet with a nine-layer encoder-decoder structure is designed. In the encoder structure, P-CBAM and MRB are used to extract features effectively and suppress noise in SAR flood disaster images. In the second stage, the Siamese network and UNet++ are used as the backbone network and the CAFNet is implemented for CD. In the CAFNet, the CA mechanism is incorporated after each down-sampling stage to extract the positional information about features.

We conducted a series of experiments on the ZZF-dataset through ablation experiments and compared the results with other advanced methods. The effectiveness of the proposed method was demonstrated through these experiments. Our empirical evidence indicates that the D-CDA has greatly improved the localization accuracy of flood disaster-stricken areas.

As for the future work, we will attempt to combine multi-source remote sensing data and ground observation data to develop more refined and efficient flood disaster monitoring methods and provide more scientific and reliable support for disaster prevention, reduction, and post-disaster reconstruction. Although the efficiency of our proposed method in this paper is relatively low, in future work, we will attempt to improve the computational speed of the network.

Open Research Section

The code and dataset are available at https://github.com/Luvinori/D-CDA.

References

Abijith, D., & Saravanan, S. (2022). Assessment of land use and land cover change detection and prediction using remote sensing and ca markov in the northern coastal districts of tamil nadu, india. *Environmental Science and Pollution Research*, *29*(57), 86055–86067.

Atasever, U. H., & Gunen, M. A. (2021). Change detection approach for sar imagery based on arc-tangential difference image and k-means++. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Bazi, Y., Bruzzone, L., & Melgani, F. (2005). An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal sar images. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(4), 874–887.

Cao, X., Ji, Y., Wang, L., Ji, B., Jiao, L., & Han, J. (2019). Sar image change detection based on deep denoising and cnn. *IET Image Processing*, *13*(9), 1509–1515.

Chen, H., Jiao, L., Liang, M., Liu, F., Yang, S., & Hou, B. (2019). Fast unsupervised deep fusion network for change detection of multitemporal sar images. *Neurocomputing*, *332*, 56–70.

Chen, H., Zhang, H., & Feng, Z. (2022). Feature learning and change feature classification based on variational auto-encoder for sar change detection. In *International conference on intelligence science* (pp. 409–416).

Chen, J.-W., Wang, R., Ding, F., Liu, B., Jiao, L., & Zhang, J. (2020). A convolutional neural network with parallel multi-scale spatial pooling to detect temporal changes in sar images. *Remote Sensing*, *12*(10), 1619.

Chierchia, G., Cozzolino, D., Poggi, G., & Verdoliva, L. (2017). Sar image despeckling through convolutional neural networks. In *2017 ieee international geoscience and remote sensing symposium (igarss)* (pp. 5438–5441).

Chunhua, W., & Fangchao, W. (2022). A sar image wavelet denoising method based on improved threshold function. , *39*(5), 39–44.

Dekker, R. (1998). Speckle filtering in satellite sar change detection imagery. *International Journal of Remote Sensing*, *19*(6), 1133–1146.

Dong, H., Jiao, L., Ma, W., Liu, F., Liu, X., Li, L., & Yang, S. (2022). Deep shearlet network for change detection in sar images. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–15.

Dong, H., Ma, W., Wu, Y., Zhang, J., & Jiao, L. (2020). Self-supervised representation learning for remote sensing image change detection based on temporal prediction. *Remote Sensing*, *12*(11), 1868.

Duan, H., Dong, X., You, S., & Han, S. (2021). A deep learning denoising framework based on ffdnet for sar image change detection. In *2021 ieee 11th international conference on electronics information and emergency communication (iceiec) 2021 ieee 11th international conference on electronics information and emergency communication (iceiec)* (pp. 1–4).

Frost, V. S., Stiles, J. A., Shanmugan, K. S., & Holtzman, J. C. (1982a). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence*(2), 157–166.

Frost, V. S., Stiles, J. A., Shanmugan, K. S., & Holtzman, J. C. (1982b). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence*(2), 157–166.

Gao, Y., Gao, F., Dong, J., Du, Q., & Li, H.-C. (2021). Synthetic aperture radar image change detection via siamese adaptive fusion network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10748–10760.

Gong, M., Cao, Y., & Wu, Q. (2011). A neighborhood-based ratio approach for change detection in sar images. *IEEE Geoscience and Remote Sensing Letters*, *9*(2), 307–311.

He, X., Zhang, S., Xue, B., Zhao, T., & Wu, T. (2023). Cross-modal change detection flood extraction based on convolutional neural network. *International Journal of Applied Earth Observation and Geoinformation*, *117*, 103197.

Hou, B., Wei, Q., Zheng, Y., & Wang, S. (2014). Unsupervised change detection in sar image based on gauss-log ratio image fusion and compressed projection. *IEEE journal of selected topics in applied earth observations and remote sensing*, *7*(8), 3297–3317.

Hou, Q., Zhou, D., & Feng, J. (2021). Coordinate attention for efficient mobile network design. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 13713–13722).

Jakka, T. K., Reddy, Y. M., & Rao, B. P. (2019). Gwdwt-fcm: change detection in sar images using adaptive discrete wavelet transform with fuzzy c-mean clustering. *Journal of the Indian Society of Remote Sensing*, *47*, 379–390.

Ko, J., & Lee, S. (2021). Sar image despeckling using continuous attention module. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 3–19.

Kuan, D. T., Sawchuk, A. A., Strand, T. C., & Chavel, P. (1985). Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE transactions on pattern analysis and machine intelligence*(2), 165–177.

Kumar, J. T., Reddy, Y. M., & Rao, B. P. (2020). Whda-fcm: wolf hunting-based dragonfly with fuzzy c-mean clustering for change detection in sar images. *The Computer Journal*, *63*(2), 308–321.

Lanka, S. S., & Puli, K. K. (2023). Change detection mapping using landsat synthetic aperture radar images. In *2023 international conference on communication, circuits, and systems (ic3s)* (pp. 1–6).

Lattari, F., Gonzalez Leon, B., Asaro, F., Rucci, A., Prati, C., & Matteucci, M. (2019). Deep learning for sar image despeckling. *Remote Sensing*, *11*(13), 1532.

Li, H.-C., Yang, G., Yang, W., Du, Q., & Emery, W. J. (2020). Deep nonsmooth nonnegative matrix factorization network with semi-supervised learning for sar image change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, *160*, 167–179.

Li, Y., Martinis, S., Plank, S., & Ludwig, R. (2018). An automatic change detection approach for rapid flood mapping in sentinel-1 sar data. *International journal of applied earth observation and geoinformation*, *73*, 123–135.

Liu, L., Jia, Z., Yang, J., & Kasabov, N. K. (2019). Sar image change detection based on mathematical morphology and the k-means clustering algorithm. *IEEE Access*, *7*, 43970–43978.

Liu, S., Lei, Y., Zhang, L., Li, B., Hu, W., & Zhang, Y.-D. (2021). Mrddanet: A multiscale residual dense dual attention network for sar image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–13.

Liu, S., Liu, M., Li, P., Zhao, J., Zhu, Z., & Wang, X. (2017). Sar image denoising via sparse representation in shearlet domain based on continuous cycle spinning. *IEEE transactions on geoscience and remote sensing*, *55*(5), 2985–2992.

Liu, S. Q., Hu, S. H., Xiao, Y., & An, Y. L. (2014). Bayesian shearlet shrinkage for sar image de-noising via sparse representation. *Multidimensional Systems and Signal Processing*, *25*, 683–701.

Liu, X., Lu, X., & Ma, L. (2020). An improved synthetic aperture radar image denoising method based on block-matching and 3d filtering. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *43*, 319–323.

Liu, Z., Lai, R., & Guan, J. (2020). Spatial and transform domain cnn for sar image despeckling. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Ma, W., Yang, H., Wu, Y., Xiong, Y., Hu, T., Jiao, L., & Hou, B. (2019). Change detection based on multi-grained cascade forest and multi-scale fusion for sar images. *Remote Sensing*, *11*(2), 142.

Munawar, H. S., Hammad, A. W., & Waller, S. T. (2022). Remote sensing methods for flood prediction: A review. *Sensors*, *22*(3), 960.

Parrilli, S., Poderico, M., Angelino, C. V., & Verdoliva, L. (2011). A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(2), 606–616.

Peter, B. G., Messina, J. P., Lin, Z., & Snapp, S. S. (2020). Crop climate suitability mapping on the cloud: A geovisualization application for sustainable agriculture. *Scientific Reports*, *10*(1), 15487.

Priyatna, M., Wijaya, S. K., Khomarudin, M. R., Yulianto, F., Nugroho, G., Afgatiani, P. M., ... Hussein, M. A. (2023). The use of multi-sensor satellite imagery to analyze flood events and land cover changes using change detection and machine learning techniques in the barito watershed. *Journal of Degraded & Mining Lands Management*, *10*(2).

Qu, X., Gao, F., Dong, J., Du, Q., & Li, H.-C. (2021). Change detection in synthetic aperture radar images using a dual-domain network. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Samadi, F., Akbarizadeh, G., & Kaabi, H. (2019). Change detection in sar images using deep belief network: a new training approach based on morphological images. *IET Image Processing*, *13*(12), 2255–2264.

Shen, H., Zhou, C., Li, J., & Yuan, Q. (2020). Sar image despeckling employing a recursive deep cnn prior. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(1), 273–286.

Tebaldini, S., Manzoni, M., Tagliaferri, D., Rizzi, M., Monti-Guarnieri, A. V., Prati, C. M., ... Mazzucco, C. (2022). Sensing the urban environment by automotive sar imaging: Potentials and challenges. *Remote Sensing*, *14*(15), 3602.

Thakur, R. K., & Maji, S. K. (2022a). Agsdnet: Attention and gradient-based sar denoising network. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Thakur, R. K., & Maji, S. K. (2022b). Sifsdnet: Sharp image feature based sar denoising network. In *Igarss 2022-2022 ieee international geoscience and remote sensing symposium* (pp. 3428–3431).

Vekaria, D., Chander, S., Singh, R., & Dixit, S. (2022). A change detection approach to flood inundation mapping using multi-temporal sentinel-1 sar images, the brahmaputra river, assam (india): 2015–2020. *Journal of Earth System Science*, *132*(1), 3.

Wang, J., Gao, F., Dong, J., Du, Q., & Li, H.-C. (2022). Change detection from synthetic aperture radar images via dual path denoising network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 2667–2680.

Wang, P., Zhang, H., & Patel, V. M. (2017). Sar image despeckling using a convolutional neural network. *IEEE Signal Processing Letters*, *24*(12), 1763–1767.

Wang, R., Zhang, J., Chen, J., Jiao, L., & Wang, M. (2018). Imbalanced learning-based automatic sar images change detection by morphologically supervised pca-net. *IEEE Geoscience and Remote Sensing Letters*, *16*(4), 554–558.

Xie, C., Wu, Y., Maaten, L. v. d., Yuille, A. L., & He, K. (2019). Feature denoising for improving adversarial robustness. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 501–509).

Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th sigspatial international conference on advances in geographic information systems* (pp. 270–279).

Zhang, X., Li, M., Liu, X., Zhou, Z., Wang, X., & Xu, Z. (2022). An improved wavelet denoising algorithm for sar interferogram using fast non-local means filtering. *Geocarto International*, *37*(27), 18600–18617.

Zhang, Y., Yang, G., Gao, A., Lv, W., Xie, R., Huang, M., & Liu, S. (2023). An efficient change detection method for disaster-affected buildings based on a lightweight residual block in high-resolution remote sensing images. *International Journal of Remote Sensing*, *44*(9), 2959–2981.

Zhao, B., Sui, H., & Liu, J. (2023a). Siam-dwenet: Flood inundation detection for sar imagery using a cross-task transfer siamese network. *International Journal of Applied Earth Observation and Geoinformation*, *116*, 103132.

Zhao, B., Sui, H., & Liu, J. (2023b). Siam-dwenet: Flood inundation detection for sar imagery using a cross-task transfer siamese network. *International Journal of Applied Earth Observation and Geoinformation*, *116*, 103132.