文章编号：1001-9014(2024)06-0775-09

DOI:10.11972/j.issn.1001-9014.2024.06.008

DIFNet: SAR RFI suppression network based on domain invariant features

LYU Wen-Hao[2,3], FANG Fu-Ping[1*], TIAN Yuan-Rong[1]

(1. School of Electronic Science, National University of Defense Technology, Changsha 410073, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. School of physics and optoelectronic Engineering, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China)

Abstract: Synthetic aperture radar (SAR) is a high-resolution two-dimensional imaging radar. However, during the imaging process, SAR is susceptible to intentional and unintentional interference, with radio frequency interference (RFI) being the most common type, leading to a severe degradation in image quality. To address the above problem, numerous algorithms have been proposed. Although inpainting networks have achieved excellent results, their generalization is unclear. Whether they still work effectively in cross-sensor experiments needs further verification. Through the time-frequency analysis to interference signals, this work finds that interference holds domain invariant features between different sensors. Therefore, this work reconstructs the loss function and extracts the domain invariant features to improve its generalization. Ultimately, this work proposes a SAR RFI suppression method based on domain invariant features, and embeds the RFI suppression into SAR imaging process. Compared to traditional notch filtering methods, the proposed approach not only removes interference but also effectively preserves strong scattering targets. Compared to PISNet, our method can extract domain invariant features and hold better generalization ability, and even in the cross-sensor experiments, our method can still achieve excellent results. In cross-sensor experiments, training data and testing data come from different radar platforms with different parameters, so cross-sensor experiments can provide evidence for the generalization.

Key words: synthetic aperture radar, radio frequency interference suppression, domain invariant features, SAR imaging

DIFNet:基于域不变特征的合成孔径雷达干扰抑制网络

吕文浩[2,3], 方付平[1*], 田元荣[1]

(1. 国防科技大学 电子科学学院,湖南 长沙 410073;
2. 中国科学院大学,北京 100049;
3. 国科大杭州高等研究院 物理与光电工程学院,浙江 杭州 310024)

摘要:合成孔径雷达(synthetic aperture radar,SAR)是一种高分辨率的二维成像雷达,但在成像过程中,合成孔径雷达易受到有意和无意的干扰,导致图像质量的严重下降,其中最常见是射频干扰。为了解决上述问题,众多算法被提出,虽然图像修复已经取得了优秀的结果,但是其泛化能力未知,在跨传感器实验中它是否仍然有效需要进一步验证。通过在时频域上对干扰信号分析,本工作发现射频干扰在不同传感器之间具有域不变的特征。因此,本工作重构了损失函数,并提取域不变特征,以改善网络的泛化能力。最终,本工作提出了一种基于域不变特征的合成孔径雷达射频干扰抑制方法,并将射频抑制网络嵌入到合成孔径雷达的成像过程中。所提方法与传统的陷波滤波方法相比,不仅能够消除干扰,还能有效保留强散射目标。同时与PISNet相比,所提方法可以提取域不变特征,具有更好的泛化能力,即使在跨传感器实验中,仍然可以取得优秀的结果。在跨传感器实验中,训练数据和测试数据来自不同的雷达平台,具备不同的雷达参数,因此,跨传感器实验可以为模型的泛化能力提供证明。

Received date: 2024-03-26, **revised date**: 2024-06-20 **收稿日期**:2024-03-26,**修回日期**:2024-06-20

Foundation items: Supported by the National Natural Science Foundation of China (62001489)

Biography: LYU Wen-Hao (2000-), male, Zhengzhou, master. Research area involves image processing, signal processing, and machine learning. E-mail: lvwenhao22@mails.ucas.ac.cn

[*] **Corresponding author**: E-mail: capkoven@mail.ustc.edu.cn

关 键 词:合成孔径雷达;射频干扰抑制;域不变特征;合成孔径雷达成像
中图分类号:O441 文献标识码:A

Introduction

Synthetic aperture radar is an active microwave sensing system that adopts synthetic aperture and pulse compression techniques to acquire high-resolution images[1]. Different waveband SAR systems are suitable for different applications. P-band SAR is commonly used for underground imaging and vegetation penetration. L, S, and C-band SARs are widely used for ocean monitoring and agricultural management. X and Ka-band SARs are often employed for high-resolution imaging[2-3]. In addition, laser-based synthetic aperture radar is gradually receiving widespread attention[4-5]. During imaging, intentional and unintentional interferences often exist, with RFI being a widely common type, and the wide-range and high-intensity RFI significantly degrades SAR image quality[6-7].

To address these issues, numerous interference suppression algorithms have been proposed, broadly categorized into three types: non-parametric methods, parametric methods, and semi-parametric methods [8]. Regarding non-parametric methods, Ref. [9] proposed an eigen-subspace-based filtering approach and this method holds very good compatibility with existing SAR imaging algorithms. Reference [10] proposed a generic subspace model for characterizing a variety of RFI types, and next designed a block subspace filter to remove RFI in SLC data. Parameterization methods often use iterative methods to solve interference parameters, and then filter out interference[11-12], which is often constrained by complex environment. The semi-parametric methods have gradually become mainstream due to its excellent performance, but they still face the drawback of high computational complexity. Common semi-parametric methods include sparse reconstruction[13], variants of robust PCA[14-15], and so on. Deep learning has been widely deployed in various fields due to its excellent performance[16-17], and naturally it is introduced into interference suppression[18-19]. The time-frequency domain radio frequency interference suppression method proposed in Ref. [20] achieved better performance than robust PCA, and the networks proposed in Refs. [18, 20] are collectively referred to as the image inpainting network.

Although image inpainting networks have achieved excellent results, their generalization is unclear, and whether they still work effectively in cross-sensor experiments needs further verification. What's more, in SAR interference suppression, there is a significant issue of incomplete data. Typically, we either obtain clean data or interfered data. Clean data and interfered data lack a corresponding relationship. To solve the above problems, this paper proposes a RFI suppression network based on domain invariant features, which offers the following contributions:

(1) Through time-frequency analysis of interference signals, we find that interference holds domain invariant features between different sensors. Therefore, this paper reconstructs the loss function and extracts the domain invariant features to improve its generalization. What's more, we also found that interference holds global characteristics on time-frequency spectrogram. Therefore, we adopt Transformer as the backbone network, and reduce the computational complexity by limiting the attention mechanism into local windows.

(2) Compared to traditional notch filtering methods, our network avoids mistakenly classifying strong scattering targets, and the proposed method achieves better interference suppression effect. Compared with image inpainting networks, this method holds stronger generalization ability in cross-sensor experiments. Even if the training data and testing data come from different sensors, the algorithm can still achieve excellent results. What's more, our method only requires the interfered data to perform interference suppression. Therefore, this approach can bypass the issue of incomplete data.

The organization of this paper is as follows: Section 1 introduces the signal model and the network. Section 2 presents the experimental results. And Section 3 provides a summary of the total paper.

1 Method

This paper proposes a network based on domain invariant features (DIFNet), and embeds the RFI suppression into SAR imaging process. The overall process is illustrated in Fig. 1, and the algorithm is shown in Table 1. The first step is to locate interfered SAR echoes, and the method is as follows: the SAR imaging algorithm is similar to a linear transformation, which converts SAR echoes into SAR images. Therefore, there is a clear correspondence between SAR echoes and SAR images. In this paper, we first locate the interference area in the SAR images, and then find the corresponding SAR echoes. The second step is to transform signals into time-frequency domain one by one by short-time Fourier transform. The third step is to use the proposed method to suppress interference. The fourth step is to transform signals into original domain by inverse short-time Fourier transform. The fifth step is to transform SAR echoes into SAR images by SAR imaging algorithm. Moreover, interference suppression mainly consists of three steps. The first step is to model interference signals, and then construct training data based on the proposed model. The second step is to separate the interference from aliasing signals. The third step is to convert SAR echoes to SAR images by SAR imaging algorithms.

1.1 RFI signal models

Common RFI can be categorized into narrowband interference, chirp broadband interference and sinusoidal broadband interference[6]. The narrowband interference can be expressed as follows:

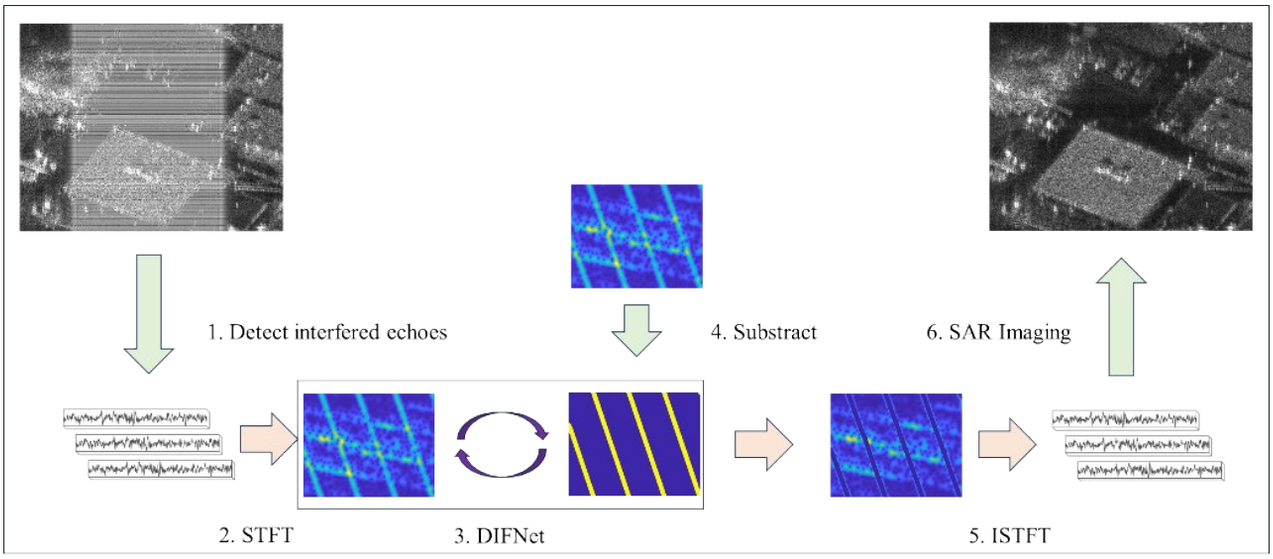

Fig. 1　Flow chart of RFI suppression network based on DIFNet
图1　基于DIFNet的射频干扰抑制网络流程图

Table 1　DIFNet's pipeline
表1　算法流程

Algorithm I: DIFNet's pipeline
1. Detect RFI in SAR images;
2. Perform STFT pulse-by-pulse;
3. Predict RFI by DIFNet;
4. Subtract RFI;
5. Perform ISTFT pulse-by-pulse;
6. Convert SAR echoes into SAR images.

$$s_{\mathrm{nbi}}(t)=\sum_{n=1}^{N}A_n\cdot\mathrm{rect}\left(\frac{t}{t_r}\right)\cdot\exp\left(\left(j2\pi f_c t+j2\pi f_n t\right)\right),\quad(1)$$

wherein, t_r is the during time, f_c is the carrier frequency, A_n is the amplitude, f_n is the frequency offset, and N is the number of interference signals. The chirp modulation interference can be expressed as:

$$s_{\mathrm{cm}}(t)=\sum_{n=1}^{N}B_n\cdot\mathrm{rect}\left(\frac{t}{t_r}\right)\cdot\exp\left(\left(j2\pi f_c t+j2\pi k_n t^2\right)\right),\quad(2)$$

where, B_n is the amplitude, and k_n is the tuning rate. The sinusoidal modulation interference can be represented as follows:

$$s_{\mathrm{sm}}(t)=\sum_{n=1}^{N}C_n\cdot\mathrm{rect}\left(\frac{t}{t_r}\right)\cdot\exp\left(j2\pi f_c t+j\beta_n\cdot\sin\left(2\pi f_n t\right)\right),\quad(3)$$

where, C_n is the amplitude, β_n is the modulation coefficient and f_n is the modulation frequency. And Eqs. (1-3) can be uniformly expressed as follows:

$$s_{\mathrm{RFI}}(t)=\sum_{n=1}^{N}D_n\cdot\mathrm{rect}\left(\frac{t}{t_r}\right)\cdot\exp\left(j2\pi f_c t+j2\pi k_{\mathrm{RFI}}t^2\right),\quad(4)$$

wherein, D_n is the amplitude, k_{RFI} is the tuning rate. When $(k_{\mathrm{RFI}}\cdot t_r)$ is small, RFI is a narrowband interference, and when $(k_{\mathrm{RFI}}\cdot t_r)$ is large, RFI is a broadband interference. Referring to Eq. (4), it can be found that:

$$f=k_{\mathrm{RFI}}\cdot t\quad.\quad(5)$$

In Eq. (5), it can be seen that the signal holds global characteristics on time-frequency spectrogram, so a Transformer network will work well. Moreover, in cross-sensor experiments, interference signals do not change with radar signals. Therefore, in the signal domain, interference signals hold domain invariant characteristics, which enlightens us to extract the homogeneous characteristics of interference, so that the algorithm may be generalized between different sensors.

1.2　Network

Figure 2 illustrates the DIFNet, which consists of an encoder and a decoder, the input is the interfered image $I\in R^{1\times H\times W}$, and the output is the label $O\in R^{1\times H\times W}$. The input projection layer consists of three CNN layers and a ReLU activation layer, the channels' number is C, and the extracted feature can be expressed as $X_0\in R^{C\times H\times W}$. The encoder consists of an input projection layer, multiple Transform block layers, and down-sampling layers. The decoder consists of an input projection layer, multiple Transform block layers, and up-sampling layers. Internal structures of DIFNet are shown in Fig. 3, the Transformer block consists of a local multi-head attention mechanism layer (Local-MAM) and a CNNs layer. In the Local-MAM, firstly, it divides the input into N windows, then, extracts global information within each window, and finally concatenates all windows. The CNNs consists of three CNN layers and a ReLU activation layer. Both encoder and decoder include L layers, each layer consists of a Transformer block and a down-sampling layer or up-sampling layer. The down-sampling layer will reduce the size of the image by half and double the number of channels, and the up-sampling layer will double the size of the image and reduce the number of channels by half. For a given input $X_0\in R^{C\times H\times W}$, the output feature map of the l-th stage can be represented as $X_l\in R^{2^l C\times\frac{H}{2^l}\times\frac{W}{2^l}}$. And there is a skip

connection between the encoder and decoder. In Fig. 2, in order to better balance performance and computational cost, we set the parameters as follows: $H \times W = 512 \times 512$, $C = 16$, $M \times M = 8 \times 8$, $L = 4$. When the image size is too large, the computational load will increase rapidly, so we set the input image size as $H \times W = 512 \times 512$. Increasing the number of channels C will further increase the amount of extracted information, but in our experiment, we found that the performance improvement is slower when C is further increased. Setting the window size to $M \times M = 8 \times 8$ can effectively balance computation and performance.

1.2.1 Transformer block

The proposed Transformer Block consists of a Local-MAM and a CNNs, and its advantages are as follows: firstly, compared to traditional Transformer, it significantly reduces computational complexity, because our method limits the calculation into a non-overlapping local window. Secondly, the propose block can capture both global and local information by Local-MAM and CNNs. The Transformer block can be represented as follows:

$$X_l = \text{Local} - \text{MAM}\left(\text{LN}\left(X_{l-1}\right)\right) + X_{l-1} \quad , \quad (6)$$

$$X_l = \text{CNNs}\left(\text{LN}\left(X_l\right)\right) + X_l \quad . \quad (7)$$

1.2.2 Local-MAM

For traditional Transformer, due to its global receptive field, its computational cost is particularly high. However, there is a significant amount of redundancy information in images. Therefore, we can limit the attention mechanism to a local window. We first split $X \in R^{C \times H \times W}$, into $M \times M$ non overlapping blocks, and the input data to Local-MAM is $X^i \in R^{C \times M \times M}$. Next, we calculate multi-head attention on each window. The computation is as follows:

$$X = \left\{X^1, X^2, \cdots, X^N\right\},\ N = HW/M^2 \quad , \quad (8)$$

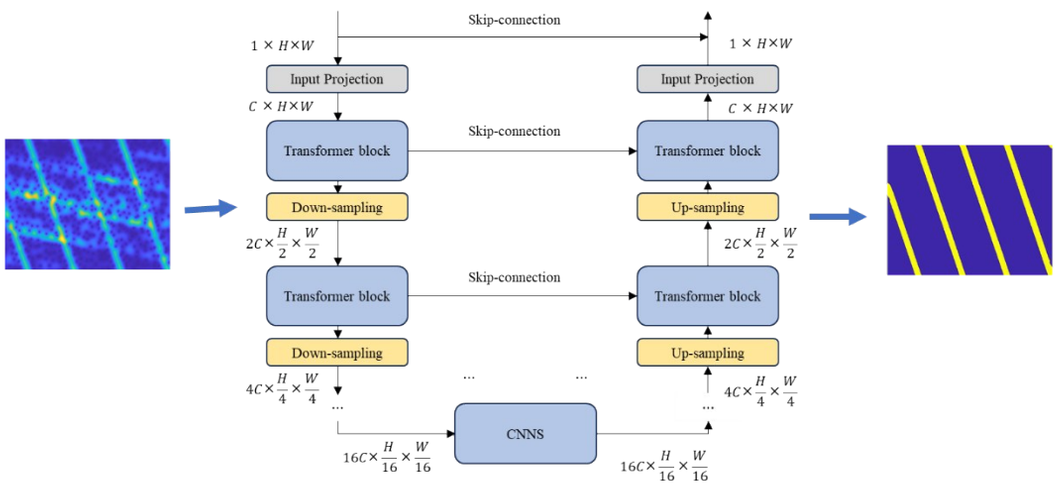

Fig. 2 DIFNet diagram
图2 DIFNet网络图

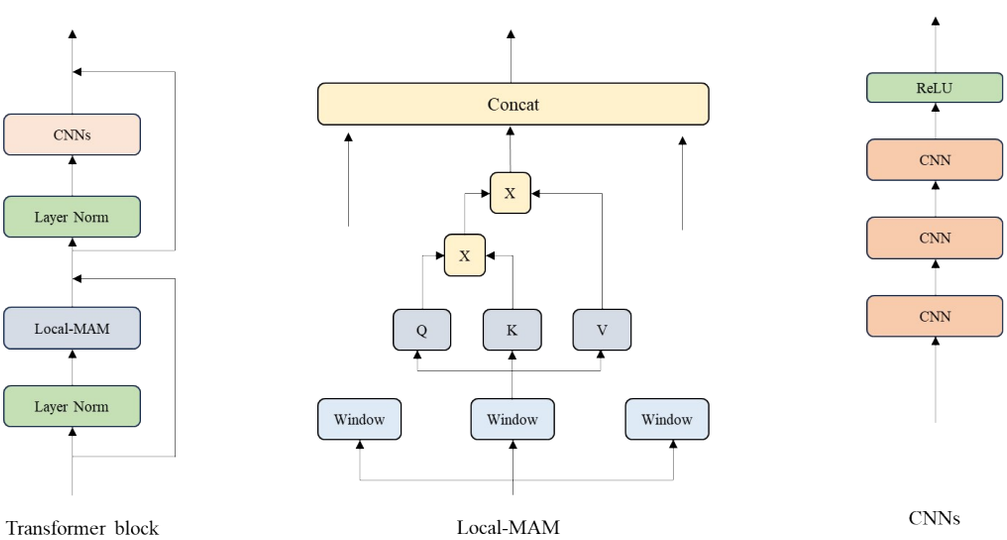

Fig. 3 Internal structures of DIFNet
图3 DIFNet内部结构图

$$Y_k^i = \text{Attention}\left(X^iW_k^Q, X^iW_k^K, X^iW_k^V\right),\ i = 1,\cdots,N. \quad (9)$$

$$\hat{Y}_k = \left\{Y_k^1, Y_k^2, \cdots, Y_k^N\right\} \quad . \quad (10)$$

Lastly, we concatenate the output of all multi-head attention layers, and adopt a linearly projected layer to obtain final result. Similarly, we also introduced a relative position encoding B. The calculation of multi-head attention is as follows:

$$\text{Attention}\left(Q,K,V\right) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}} + B\right)\cdot V. \quad (11)$$

Compared to traditional Transformer, the computational complexity of Local-MAM is reduced from $O\left(H^2W^2C\right)$ to $O\left(M^2HWC\right)$, and M is the window size. Usually, $M \ll \min\left(H, W\right)$, so this method can reduce complexity.

1.2.3 CNNs

For the standard Transformer, it is hard to capture local contextual information due to its equal distance between pixels. Considering the importance of neighboring pixels for image tasks, we introduce a cascade CNNs in the Transform block to capture local information, and the CNNs consists of three CNN layers and a ReLU activation layer.

1.2.4 Loss function

In inpainting network, the loss can be defined as:

$$l = \sqrt{\left(HX - Y\right)^2 + \varepsilon} \quad , \quad (12)$$

H is the network's matrix, X is the input, and Y is the label. In order to extract domain invariant interference features, Y is optimized as follows:

$$Y_{\text{pixel}} = \begin{cases} 1, & \text{pixel} \in RFI \\ 0, & \text{pixel} \notin RFI \end{cases} \quad . \quad (13)$$

Pixel represent pixel value of the image, and in Eq. (13), it can ensure that the measurement distance of RFI does not change in different sensors. Therefore, it can induce the network to learn domain invariant features.

So, the loss can be expressed as follows:

$$l = \sqrt{\left(HX_{\text{pixel}} - Y_{\text{pixel}}\right)^2 + \varepsilon} \quad . \quad (14)$$

In Eq. (12), the loss function contains two constraints. The first part is the interference constraint, and the second part is the target constraint. In Eq. (14), we set the target pixels to 0 and the interference pixels to 1. According to Eq. (5), the interference signal holds domain invariant features, while the target signal varies with radar parameters. Through the above constraints, and the network will learn domain invariant features, and can be migrated between different sensors. Lastly, we subtract the predicted interference from the original time-frequency spectrogram, so the interference can be filtered out.

1.3 SAR imaging

After removing the interference, we need to perform imaging processing on the clean echoes. The radar imaging process is shown in Fig. 4. The radar moves from point A to point B with a speed of v, and during this motion, it continuously detects the target. The vertical distance between the radar and the target is R_0, and the distance equation between the radar and the target is represented as $R\left(\eta\right)$.

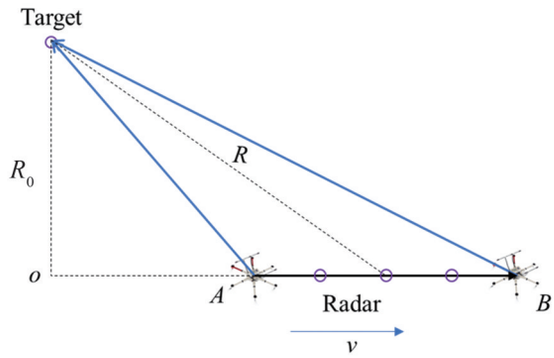

Fig. 4　SAR imaging model
图 4　合成孔径雷达成像模型

In the motion, the clean SAR echoes can be expressed as follows:

$$s_r\left(\tau,\eta\right) = A_0 w_r\left(\tau - 2R\left(\eta\right)/c\right) w_a\left(\eta - \eta_c\right)\exp$$
$$\left(-j4\pi f_c R\left(\eta\right)/c\right)\exp\left(j\pi K_r\left(\tau - 2R\left(\eta\right)/c\right)^2\right), \quad (15)$$

wherein, τ is the fast time, η is the slow time, A_0 is the amplitude, w is rectangle window function, η_c is the Doppler center, $R\left(\eta\right)$ is range function, f_c is the carrier frequency, and K_r is range tuning rate. The range function can be expressed as follows:

$$R\left(\eta\right) = \sqrt{R_0^2 + \left(v\eta\right)^2} \quad , \quad (16)$$

wherein, R_0 is vertical distance between the radar and the target, and v is radar speed. By range compressing, the signal can be expressed as follows:

$$s_{rc}\left(\tau,\eta\right) = A_0 p_r\left(\tau - 2R\left(\eta\right)/c\right) w_a\left(\eta - \eta_c\right)\exp$$
$$\left(-j4\pi f_c R_0/c\right)\exp\left(-j\pi\frac{2v^2}{\lambda R_0}\eta^2\right) \quad , \quad (17)$$

wherein, p is a sinc-function signal. By azimuth Fourier transform, the signal can be expressed as follows:

$$S_{rc}\left(\tau,f_\eta\right) = A_0 p_r\left(\tau - 2R\left(f_\eta\right)/c\right) W_a\left(\eta - \eta_c\right)\exp$$
$$\left(-j4\pi f_c R_0/c\right)\exp\left(-j\pi\frac{f_\eta^2}{K_a}\right) \quad , \quad (18)$$

wherein, f_η is azimuth frequence, and $K_a = \dfrac{2v^2}{\lambda R_0}$ is the azimuth tuning rate. By correcting range cell migration component, the signal can be expressed as follows:

$$S_{rm}\left(\tau,f_\eta\right) = A_0 p_r\left(\tau - 2R_0/c\right) W_a\left(\eta - \eta_c\right)\exp$$
$$\left(-j4\pi f_c R_0/c\right)\exp\left(-j\pi\frac{f_\eta^2}{K_a}\right) \quad . \quad (19)$$

By azimuth compressing, the signal can be expressed as follows:

$$S_{ra}\left(\tau,\eta\right) = A_0 p_r\left(\tau - 2R_0/c\right) p_a\left(\eta\right)\exp$$
$$\left(-j4\pi f_c R_0/c\right)\exp\left(j\pi f_{\eta,\eta}\right) \quad . \quad (20)$$

From the above processing, the target is focused, and we can acquire the clean SAR images.

2 Experiments

In our experiments, training data comes from airborne MiniSAR, with an image size of 512×512, and training data includes a total of 2 048 images. In the training, the maximum epoch is 100, the batchsize is 4, the learning rate is 0.0002, the weight decay is 0.02, the optimizer is AdamW, and iterative loss curve and pixel accuracy is shown in Fig. 5. The interference parameters in the training are shown in Table 3. In the same sensor experiments, there are 512 testing images, and the testing data and training data come from different scenes, so there is no overlap between the training data and testing data. In the cross-sensor experiments, there are 512 testing images. The testing data was captured in the Korean region in 2019 from public Sentinel-1satellites. Therefore, there is no overlap between the testing data and training data. To validate the effectiveness of the proposed method, experiments are conducted on both MiniSAR dataset and Sentinel-1 dataset. The resolution of MiniSAR is 0.1 m, while the resolution of Sentinel-1 is 5×20 m. All training data comes from MiniSAR, and the testing data comes from MiniSAR and Sentinel-1. The radar parameters of training and testing data are shown in Table 2. From Table 2, it can be seen that in the cross-sensor experiments, the training and testing data come from different radars with different parameters. Therefore, the cross-sensor experiments can be used to verify the generalization.

Table 2 Radar parameters of training and testing Data
表2 训练数据和测数数据的雷达参数

Data Parameters	Training data	Same-sensor experiments	Cross-sensor experiments
Source	MiniSAR	MiniSAR	Sentinel-1A
Band	X-band	X-band	C-band
Bandwidth	1.5 GHz	1.5 GHz	100 MHz
Polarization mode	HH	HH	VV/VH

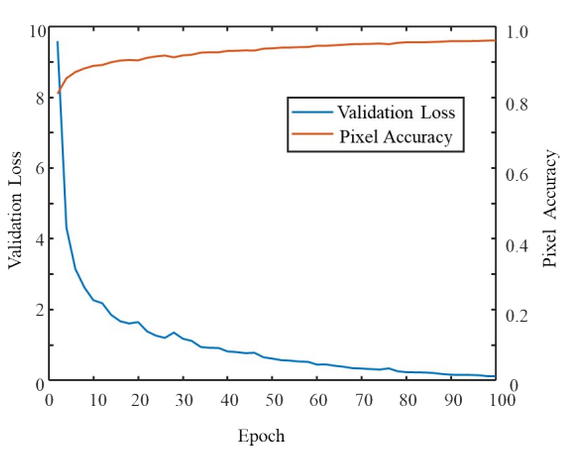

Fig. 5 Iterative loss curve of the network
图5 网络的迭代损失曲线

2.1 Evaluation metrics

To reasonably evaluate the test results, this paper adopts pixel accuracy (PA), intersection over union (IoU), PSNR[18] and ME[12] as evaluation metric. PA and IoU are used to evaluate the DIFNet, and PSNR and ME is used to evaluate image quality. PA is defined as follows:

$$PA = \frac{TP + TN}{To} \quad , \tag{21}$$

where, true positive (TP) represents the true positive instances, that is, the number of instances where the model predicts a positive class and the actual label is also positive. True negative (TN) represents the number of true negative examples, that is, the number of instances where the model predicts a negative class and the actual label is also negative. *To* is the total pixel numbers. IoU is defined as follows:

$$IoU = \frac{A \cap B}{A \cup B} \quad , \tag{22}$$

where, A represents the predicted interference area, and B represents the actual interference area. PSNR is usually used to evaluate the quality of images, and it is shown as follows:

$$\begin{cases} \mathrm{PSNR}\left(X,\hat{X}\right) = 20\log_{10}\dfrac{\mathrm{Max}_X}{\sqrt{\mathrm{MSE}}} \\ \mathrm{MSE}\left(X,\hat{X}\right) = \dfrac{1}{\mathrm{HW}}\sum_{i=0}^{W-1}\left\| X\left(i,j\right) - \hat{X}\left(i,j\right)\right\|^2 \end{cases} , \tag{23}$$

$\hat{X}$ is the filtered image, and X is the label, *MSE* is root-mean-square, and H and W represent pixel numbers. It can be seen that PSNR represents evaluation index of noise level. The larger PSNR, the better filtering performance.

ME is defined as follows:

$$\mathrm{ME} = \mathrm{Ent}\left(\hat{X}\right)\mathrm{Mean}\left(\hat{X}\right) \quad , \tag{24}$$

$\mathrm{Ent}\left(\hat{X}\right)$ is the entropy, $\mathrm{Mean}\left(\hat{X}\right)$ is the mean value. A smaller entropy indicates that the pixel values of the image are concentrated within a smaller range. A smaller mean value indicates a lower amplitude, suggesting that most of the interference has been filtered out. Therefore, a smaller ME indicates a better result.

2.2 Same-sensor experiments

In the MiniSAR experiments, the interference parameters are shown in Table 3. The interference is divided into two types: narrowband interference and broadband interference. The bandwidth of narrowband interference is less than 30 MHz, and the bandwidth of broadband interference is between 30 MHz and 150 MHz. The signal-to-interference ratio of both interferences is -15~0 dB, with 2 narrowband interference sources and 3~5 broadband interference sources.

The filtering results on the time-frequency spectrogram are shown in Fig. 6. In Fig. 6(b), it adopts constant false alarm rate (CFAR) method to filter out interference. Some strong scattering points have high intensity, so they may be mistaken as interference, as marked in the red boxes. Comparing Fig. 6(b) and Fig. 6(d), it can be observed that for the traditional notch filtering

Table 3　Interference parameters in MiniSAR
表 3　MiniSAR 中的干扰参数

Bandwidth / Parameters	Narrowband	Broadband
Interference bandwidth	<30 MHz	30 MHz~150 MHz
SIR	−15 dB~0 dB	−15 dB~0 dB
Interference source	2	3~5

Table 4　Same-sensor Result
表 4　同传感器实验结果

Methods / Indicators	Interfered image	Notch filtering	PISNet	DIFNet
PA	/	94.95%	/	96.55%
IoU	/	74.41%	/	80.40%
PSNR/dB	11.36	21.49	25.05	23.51
ME	0.24	0.09	0.04	0.04

method, some strong scattering targets are mistakenly filtered out, but for DIFNet, these strong scattering targets are well preserved. Wherein, the image restoration network PISNet achieved the best result. Figure 7 shows the imaging results, the horizontal axis represents the azimuth direction and the vertical axis represents the range direction. The size of the image is 512×512. The interference area includes the total range direction, and the interference area roughly ranges from the 100-th pixel to the 400-th pixel in the azimuth direction. It can be seen that the proposed algorithm preserves more details and produces a cleaner filtering result comparing with traditional notch filter. The evaluation results are presented in Table 4, it can be seen that the proposed method achieves a 1.6% improvement in PA, a 5.99% improvement in IoU, a 2.02 dB improvement in PSNR, and a 0.05 decrease in ME compared to the traditional algorithm. PISNet still achieved the best result.

2.3　Cross-sensor experiments

The interfered dataset is obtained from Sentinel-1, captured in the Korean region on 2/16, 2019. The image area is cropped to a size of 512 × 512, as shown in Fig. 9(a). The time-frequency spectrogram is shown in Fig. 8, and the filtering results and performance indicators are shown in Fig. 9 and Table 5. In the Sentinel-1 testing dataset, the training data still comes from MiniSAR. And in this cross-sensor experiment, PISNet does not work, so we do not present its results, while our method can still suppress the interference, demonstrating its good generalization. Since notch filter relies on intensity differences to filter out interference, and at the starting position of the interference in Fig 8(b), the interference power is relatively lower, the lower-intensity interference is not detected as shown in the red boxes, resulting in some residual components. Comparing Fig. 8(b) and Fig. 8(c), it can be observed that the proposed method acquires a better filtering result. Similarly, comparing Fig. 9(b) and Fig. 9(c), it can be seen that it is difficult to filter out low-intensity residual interference for traditional notch filter, while our method can effectively filter out the residual interference. Comparing with traditional notch filter, our method has achieved a 1.89% improvement in PA, a 2.60% improvement in IoU, a 0.15 decrease in ME.

In the cross-sensor experiments, the training data and testing data come from different radar platforms with different radar parameters. From the above experimental results, it can be seen that image inpainting network does not even work, but our method can still acquire excellent results. The above results demonstrate that our method holds good generalization.

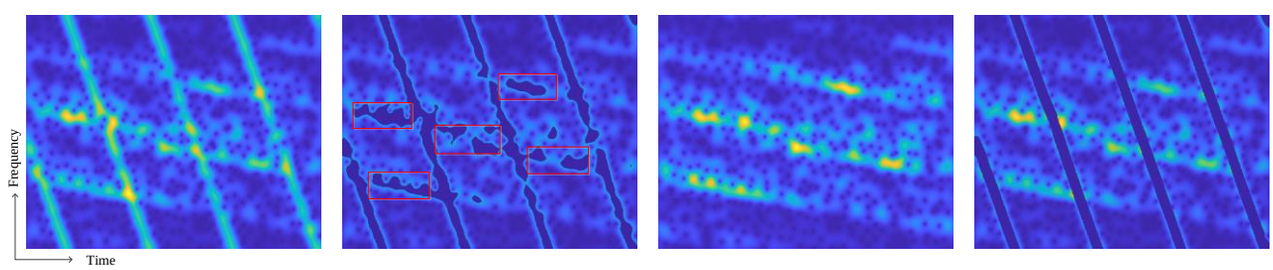

Fig. 6　Time frequency spectrogram: (a) interfered time-frequency spectrogram in same-sensor experiment; (b) time frequency spectrogram by notch filtering; (c) time frequency spectrogram by PISNet; (d) the time-frequency spectrogram by DIFNet
图 6　同传感器实验的时频谱: (a) 被干扰时频谱; (b) 由陷波滤波所得时频谱; (c) 由 PISNet 所得时频谱; (d) 由 DIFNet 所得时频谱

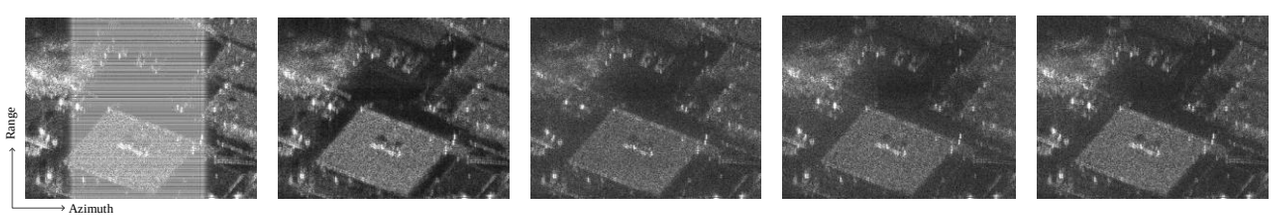

Fig. 7　Suppression results in MiniSAR: (a) interfered image; (b) label; (c) result by notch filtered; (d) result by PISNet; (e) result by DIFNet
图 7　MiniSAR 实验结果: (a) 被干扰图像; (b) 标签; (c) 陷波滤波所得结果; (d) PISNet 所得结果; (e) DIFNet 所得结果

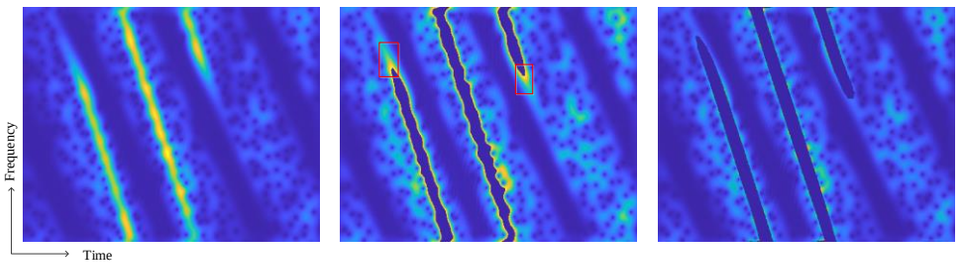

Fig. 8 Time-frequency spectrogram in cross-sensor experiment:(a) interfered time-frequency spectrogram; (b) notch filtering result; (c) DIFNet filtering result
图 8 跨传感器时间-频率谱:(a) 被干扰时频谱;(b) 陷波滤波结果;(c) DIFNet滤波结果

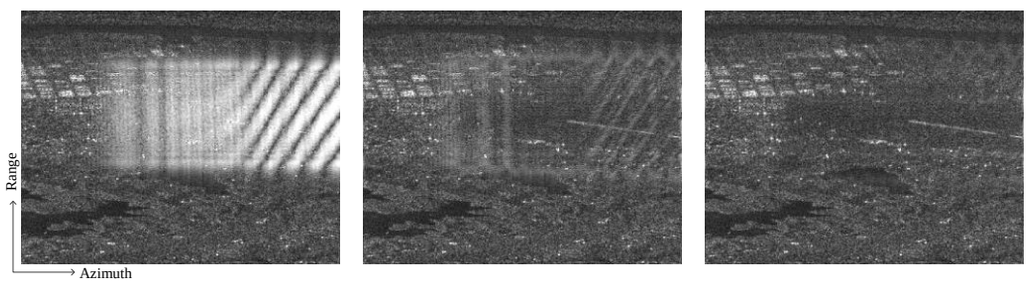

Fig. 9 Cross-sensor results in Sentinel-1:(a) interfered image, ME=3.40; (b) notch filtering image, ME=2.34; (c) DIFNet filtering image, ME=2.19
图 9 Sentinel-1跨传感器实验结果:(a) 被干扰图像,ME=3.40;(b) 陷波滤波图像,ME=2.34;(c) DIFNet滤波图像,ME=2.19

Table 5 Cross-sensor Result
表5 跨传感器实验结果

Parameters / Indicators	Interfered image	Notch filtering	DIFNet
PA	/	92.67%	94.56%
IoU	/	78.63%	81.23%
ME	3.40	2.34	2.19

3 Conclusions

SAR is widely deployed as a high-resolution imaging radar, but it is susceptible to intentional and unintentional RFIs. For image inpainting networks, although they have acquired excellent results, their generalization is unclear. To address this problem, through time-frequency analysis to interference signals, we find that interference holds domain invariant features between different sensors, so we propose a SAR RFI suppression network based on domain invariant features. Compared to traditional notch filtering methods, the proposed method acquires better interference suppression performance. Furthermore, in the cross-sensor experiments, the training data and the testing dataset are from different radars with different resolutions, and the image inpainting networks do not work, but our method can still acquire excellent results. The above demonstrates that our method holds good performance and generalization. Moreover, this method can inspire self-supervised learning, as the segmented time-frequency spectrogram forms a masking task, which can be repaired by self-supervised networks.

Acknowledgment

All data is from open-source data. Thanks for the open-source data from the European Sentinel-1 satellite. https://search.asf.alaska.edu/#/. Also, thanks for the open-source data from Sandia National Laboratories. https://www.sandia.gov/radar/pathfinder-radar-isr-and-synthetic-aperture-radar-sar-systems/.

References

[1] Reigber A, Scheiber R, Jager M, *et al*. Very-high-resolution airborne synthetic aperture radar imaging: signal processing and applications[J]. *Proceedings of the IEEE*, 2013, **101**(3): 759-783.
[2] Adeli S, Salehi B, Mahdianpari M, *et al*. Wetland monitoring using SAR data: a meta-analysis and comprehensive review [J]. *Remote Sensing*, 2020, **12**(14): 2190.
[3] Bonano M, Manunta M, Pepe A, *et al*. From previous C-band to new X-band SAR systems: assessment of the DInSAR mapping improvement for deformation time-series retrieval in urban areas [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **51**(4): 1973-1984.
[4] Zhang Hong-Yi, Li Fei, Xu Wei-Ming, *et al*. Improved differential synthetic aperture lidar for vibration suppression[J]. *Journal of Infrared and Millimeter Waves* (张鸿翼,李飞,徐卫明,等. 经过改进的差分合成孔径激光雷达对振动的抑制[J]. **红外与毫米波学报**), 2015, **34**(05): 576-582.
[5] Wang B N, Zhao J Y, Li W, *et al*. Research on high-resolution imaging technology of array laser synthetic aperture radar [J]. *Journal of Radar* (汪丙南,赵娟莹,李威,等. 阵列激光合成孔径雷达高分辨成像技术研究[J]. **雷达学报**), 2022, **11**(06): 1110-1118.
[6] Joy S, Nguyen L H, Tran T D. Joint down-range and cross-range RFI suppression in ultra-wideband SAR [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, **59**(4): 1-14.
[7] Zhao B, Huang L, Zhou F, *et al*. Performance improvement of deception jamming against SAR based on minimum condition number [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, **10**(13): 1039-1055.
[8] Tao M, Zhou F, Zhang Z. Wideband interference mitigation in high-resolution airborne synthetic aperture radar data [J]. *IEEE Transac-*

tions on Geoscience and Remote Sensing, 2015, **54**(1): 74-87.
[9] Zhou F, Wu R, Xing M, *et al*. Eigensubspace-based filtering with application in narrow-band interference suppression for SAR [J]. *IEEE Geoscience and Remote Sensing Letters*, 2007, **4**(1): 75-79.
[10] Yang H, Li K, Li J, *et al*. BSF: block subspace filtering for removing narrowband and wideband radio interference artefacts in single look complex sar images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, **60**: 1-16.
[11] Bollian T, Osmanoglu B, Rincon R, *et al*. Adaptive antenna pattern notching of interference in synthetic aperture radar data using digital beamforming[J]. *Remote Sensing*, 2019, **11**(11):1346.
[12] Tao M, Li J, Su J, *et al*. Interference mitigation for synthetic aperture radar data using tensor representation and low-rank approximation[C]. 33rd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS). 2020: 1-4.
[13] Chen S, Lin Y, Yuan Y, *et al*. Suppressive interference suppression for airborne SAR using BSS for singular value and eigenvalue decomposition based on information entropy[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, **61**: 1-11.
[14] Huang Y, Zhang L, Li J, *et al*. A novel tensor technique for simultaneous narrowband and wideband interference suppression on single-channel sar system[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(12):9575-9588.
[15] Huang Y, Chen Z, Wen C, *et al*. An efficient radio frequency interference mitigation algorithm in real synthetic aperture radar data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, **60**: 1-12.
[16] Wang Z, Cun X, Bao J, *et al*. Uformer: A general U-shaped transformer for image restoration[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2022: 17662-17672.
[17] Liu Z, Lin Y, Cao Y, *et al*. Swin transformer: hierarchical vision transformer using shifted windows[C]. 18th IEEE/CVF International Conference on Computer Vision (ICCV),IEEE, 2021: 9992-10002.
[18] Fan W, Zhou F, Tao M, *et al*. Interference mitigation for synthetic aperture radar based on deep residual network [J]. *Remote Sensing*, 2019, **11**(14): 1654.
[19] Fang, F, Li H, Meng W, *et al*. Synthetic-aperture radar radio-frequency interference suppression based on regularized optimization feature decomposition network[J]. *Remote Sensing*, 2024, **16**(14): 2540.
[20] Shen J, Han B, Pan Z, *et al*. Learning time-frequency information with prior for SAR radio frequency interference suppression [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, **60**: 1-16.