THz imaging enhancement based on Noise2Noise algorithm

Tianhe Wang, Xuejun Huang, Jinshan Ding, Yuhong Zhang

Terahertz (THz) imaging has an outstanding advantage of high resolution due to the high frequency and has promising potential in VideoSAR. However, limited to the THz source power and the air absorption, the THz image usually has a low SNR and is susceptible to noise in remote sensing and imaging. In order to improve the quality of THz images, a THz image enhancement method is proposed based on the noise2noise idea. The THz images are reconstructed with the AFBP algorithm. They are organized as noisy image pairs and filtered with a mask to remove the influence of moving targets. Then, the Noise2Noise network is constructed based on the CNN network and takes the noisy image pair as input and reference. In the training stage, 1000 noisy image pairs are used as the training set and 100 noisy images are used as the test set to verify the performance of the proposed method. The experimental results based on real VideoSAR data demonstrate that the proposed method is capable of suppressing noise and enhancing the THz image.

Introduction: Terahertz (THz) imaging offers some special characteristics such as penetration, safety, broadband and high resolution and has a huge potential in security, nondestructive evaluation and radar remote sensing [?]-[?]. The THz radar is likely to be widely used in the future as its source technology matures and its MMIC-based device develops. Some publications have revealed the efforts of THz radar system development these years, mostly focusing on the detection of concealed weapons, explosives and material defects. A 220-325GHz and 500-750GHz rail-based THz imaging system with two-dimensional aperture synthesis was shown in [?]. Cooper [?] presented the 675 GHz personnel real-time screening radar developed in JPL. Ding [?] presented a 540 GHz imaging system with a duplexer and the experimental results of 3D synthetic aperture imaging. Gu [?] reported a fast imaging system for personnel screening. These researches illustrate the prospect of the THz radar.

The THz radar has the advantage of a very short synthetic aperture due to its high frequency and has promising potential application in VideoSAR. Meanwhile, limited to the THz source power and the absorption of the water and oxygen in the air, THz imaging usually has the disadvantage of a low signal-to-noise ratio (SNR) and is susceptible to noise, especially in remote sensing and imaging. In the traditional methods, filters [?] such as mean filter including non-local mean filter, median filter, gaussian filter et al., are used to filter the noise, but usually degrade the target areas as well. Nowadays, enhancement methods with a deep learning neural net outperform traditional methods. Deep learning with neural networks, such as CNN, can provide an effective and implicit way to acquire many noise features and predict the clean images based on the features. In these methods, a CNN is adopted to learn the nonlinear mapping ability from a noisy image to a clean image and, with enough noisy-clean image pairs, the networks can achieve state-of-the-art results [?, ?]. Unfortunately, clean ground truths in THz imaging are very difficult to obtain practically. Recently, Lehtinen et al. [?] introduced the Noise2Noise training method which only needs pairs of corrupted images as training data. In this paper, a THz image enhancement method is proposed based on the Noise2Noise training strategy and the experiments based on the real VideoSAR data verified the performance.

Terahertz image Enhancement Method: The backprojection algorithm offers more advantages than traditional SAR image algorithms, such as the precise compensation and large integration angle. The accelerated fast backprojection (AFBP) algorithm [?] can provide optimal performance and high efficiency simultaneously. Therefore, the AFBP algorithm is applied to generate the THz SAR images.

THz imaging is very sensitive to noise and the performance may degrade due to the low SNR. A higher pulse repetition frequency (PRF) can lead to a higher SNR but will impose a heavy burden on the hardware system. Alternatively, considering the difficulty to obtain the clean ground truth, the Noise2Noise method is very suitable to enhance the THz images. Lehtinen et al. proved that if each clean ground truth is substituted by an uncorrelated noisy image that has the same corrupted distribution of the input, the parameter set of the network may also achieve optimization with enough training data. Theoretically, in the case of a large number of training data, the networks trained by the clean and noisy references are the same.

A typical problem in training the network with THz images is that there may exist moving targets. The moving targets result in mismatches between the two images in an image pair, which can degrade the performance of the neural networks. Therefore, a mask is applied to filter the area of the moving targets before training the network. First, a mean filter is applied to the noisy-noisy images pair to reduce the influence of noise:

$$\begin{aligned} I_1^m &= mfilter(I_1) \\ I_2^m &= mfilter(I_2) \end{aligned} \tag{1}$$

where I_1, I_2 are two images in the noisy-noisy image pair, $mfilter$ is the mean filter operation. In the differential image ($|I_1^m - I_2^m|$), the pixels have large values in mismatched areas and small values in the matched areas. Thus, a threshold γ is set to choose the pixels in the matched areas :

$$Mask = \begin{cases} 1, & if \quad |I_1^m - I_2^m| \leq \gamma \\ 0, & otherwise \end{cases} \tag{2}$$

The $Mask$ not only removes the moving target areas but also filters the mismatched areas on the edges. In our cases, the $Mask$ has a good performance when $\gamma = 0.03$. Since the two images in an image pair are corrupted by uncorrelated noise, both images can be chosen as input and the other as reference. Therefore, the outputs of the network are denoted by:

$$\begin{aligned} I_1' &= f(I_1, \theta) \\ I_2' &= f(I_2, \theta) \end{aligned} \tag{3}$$

where $f(\cdot)$ is the CNN network, and θ is the parameter set. In the training strategy, the network can be trained by penalizing the loss between its predicted image from a given noisy and a second noisy version of the same image. The loss function that removes the moving targets is defined as:

$$Loss = Mask \cdot (0.5 \left\| I_1' - I_2 \right\|_2^2 + 0.5 \left\| I_2' - I_1 \right\|_2^2) \tag{4}$$

The proposed THz imaging enhancement approach is illustrated in Fig. 1. The CNN net consists of an encoder module and a decoder module, as listed in Table I. The CNN takes a THz image pair as input and produces two output images respectively. The encoder extracts features from the input image, which consists of six convolution layers and four average pooling layers. The average pooling layer performs downsampling of the input image. The decoder module outputs the predicted image from features, which consists of five deconvolution layers and each deconvolution layer is followed by a convolution layer except for the last deconvolution layer. In addition, we use skip connections between the encoder and the decoder module at different resolutions to combine both high-level and low-level features.

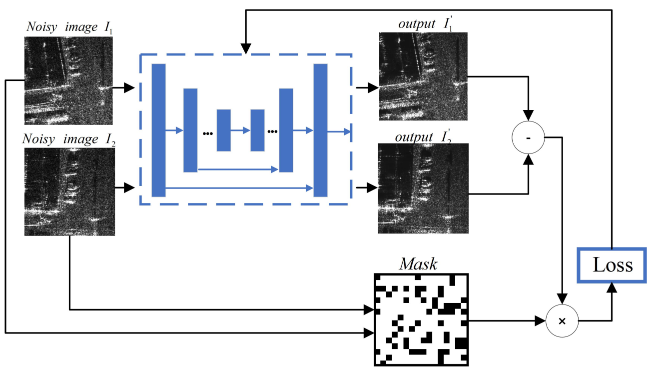

Fig. 1. Network structure of the proposed method

Experiments and Result: In our experiments, the THz radar system works in the spotlight SAR mode and operates at 235 GHz. The image resolution is about 0.2 m. The PRF is set as twice the azimuth bandwidth and the AFBP algorithm is applied to reconstruct the target area. To reconstruct THz images with unrelated noise, the odd number pulses and the even number pulses are used to generate the SAR images respectively. The THz radar data corresponding to one image in the noisy image pair only have

Table 1: Parameter for the Noise2Noise network

Module	specifications
Encoder	3x3 convolution, 64 filters, stride 1
	2×2 average pooling layer
	3x3 convolution, 128 filters, stride 1
	2×2 average pooling layer
	3x3 convolution, 256 filters, stride 1
	3x3convolution, 256 filters, stride 1
	2×2 average pooling layer
	3x3 convolution, 512 filters, stride 1
	3x3 convolution, 512 filters, stride 1
	2×2 average pooling layer
Decoder	3x3 deconvolution, 512 filters, stride 2
	3x3 convolution, 256 filters, stride 1
	3x3 deconvolution, 256 filters, stride 2
	3x3 convolution, 128 filters, stride 1
	3x3 deconvolution, 128 filters, stride 2
	3x3 convolution, 64 filters, stride 1
	3x3 deconvolution, 64 filters, stride 2
	3x3 convolution, 32 filters, stride 1
	3x3 convolution, 1 filter, stride 1

one $PRT(PRT = 1/PRF)$ interval with the data corresponding to the other one, meaning the two images are almost the same except for the noise and moving targets. Therefore, only the $Mask$ is applied to remove moving targets and no further correction is required. The THz images are divided into two data sets including 1000 image pairs randomly selected in the training set and 100 images in the test set.

Two frames of the results are shown in Fig. 2. Fig. 2(a) and (c) are the results before enhancement, and Fig. 2(b) and (d) are the corresponding enhanced results. It can be observed that noise is suppressed clearly, especially in the top-left shadow area. As indicated by the red boxes in Fig. 2, four target areas in two frames are chosen to evaluate the performance of the proposed method. The enlarged results are shown in Fig. 3. The first row and the second row are the original and the enhanced results respectively. The comparison in Fig. 3 demonstrates that the noise is suppressed in the enhanced images. The noise is almost eliminated perfectly in the shadow areas while completely retaining the targets. The image entropy (IE), ENL [?], SNR (defined as $SNR = 10log(max(x)^2/\sigma(x)^2)$) and average gradient (AG) [?] are used to evaluate the images quantitatively and the results are listed in Table 2. Table 2 shows that the resulting images are lower in image entropy and average gradient but higher in ENL and SNR. This indicates that the proposed method can suppress the noise and enhance the image contrast, which is consistent with the previous analysis.

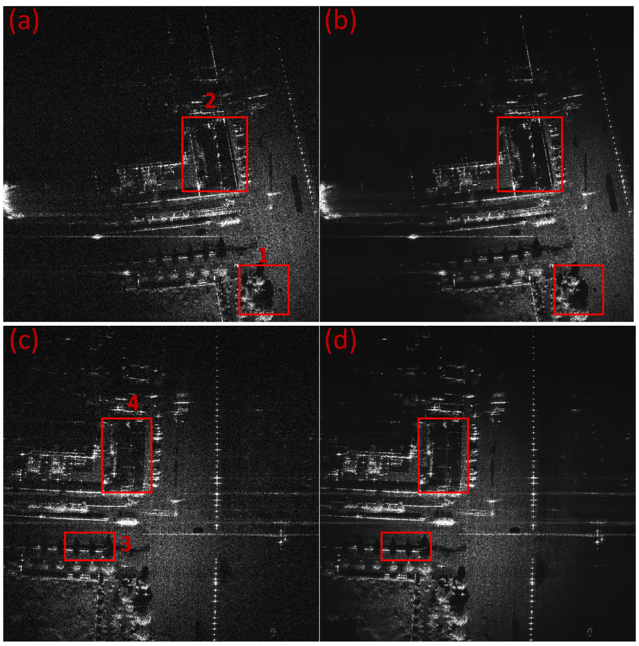

Fig. 2 *(a) and (c) are two frames of the VideoSAR, (b) and (d) are the results of the proposed method.*

Conclusion: In this letter, a THz image enhancement method is presented based on the Noise2Noise idea to improve the THz image quality. The THz images in the 235GHz band are reconstructed with the AFBP algorithm and reorganized as noisy image pairs. A mask filter is used to remove the impact of mismatches caused by the moving targets. The Noise2Noise network is built with CNN networks and the noisy image pairs are used as input and reference. In the experiments, 1000 noisy image pairs are used to train the network and 100 images are used to test the performance. Experimental results have shown the proposed method can suppress the noise and enhance the contrast of THz images and is suitable for shadow detection in VideoSAR.

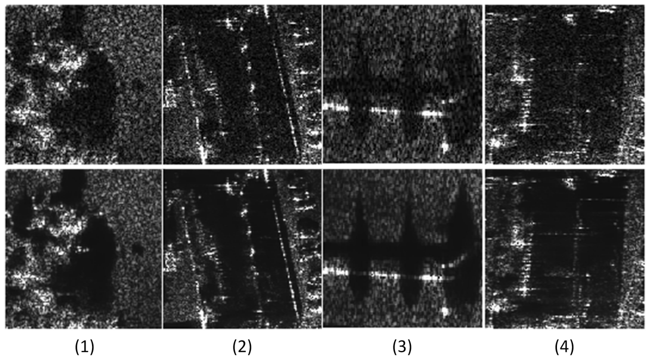

Fig. 3. Enlarged images of the four target areas in Fig. 2

Table 2: result comparison

	original image				Resulting image			
	img1	img2	img3	img4	img1	img2	img3	img4
IE	7.07	6.69	6.68	6.75	6.58	5.76	6.03	5.94
ENL	1.49	0.96	1.46	1.03	1.71	1.04	1.75	1.11
AG	0.23	0.15	0.25	0.17	0.19	0.13	0.18	0.14
SNR	14.9	15.1	17.0 4	15.2	15.87	15.9	18.2	16.1

Tianhe Wang, Xuejun Huang, Jinshan Ding (National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, People's Republic of China)
E-mail: ding@xidian.edu.cn
Yuhong Zhang (School of Electronic Engineering, Xidian University, Xi'an 710071, People's Republic of China)

References

1 S. Saqueb, N.K. Nahar and K. Sertel: "Fast two-dimensional THz imaging using rail-based synthetic aperture radar (SAR) processing," *ELECTRONICS LETTERS*, 2020, **56**, (19), pp.988-990
2 K. B. Cooper, R. J. Dengler, et al.: "THz Imaging Radar for Standoff Personnel Screening",*IEEE Transactions on Terahertz Science and Technology*, 2011, **1**, (1), pp.169-182
3 J. Ding, M. Kahl, et al.: "THz 3-D Image Formation Using SAR Techniques: Simulation, Processing and Experimental Results," *Terahertz Science and Technology, IEEE Transactions on*, 2013, **3**, (5), pp.606-616
4 S. Gu, C. Li, X. Gao, et al.: "Three-Dimensional Image Reconstruction of Targets Under the Illumination of Terahertz Gaussian Beam: Theory and Experiment,"*IEEE Transactions on Geoscience and Remote Sensing*, 2013,**51**, (4), pp.2241-2249
5 J. Ji and Y. Li: "An Improved SAR Image Denoising Method Based on Bootstrap Statistical Estimation with ICA Basis", *Chinese Journal of Electronics*, 2016, **25**, (4), pp.786-792
6 P. Wang, H. Zhang,and V. M. Patel: "SAR Image Despeckling Using a Convolutional Neural Network," *IEEE Signal Processing Letters*, 2017, **24**, (12), pp.1763-1767
7 J. Zhang, W. Li, and Y. Li: "SAR Image Despeckling Using Multiconnection Network Incorporating Wavelet Features," *IEEE Transactions on Geoscience and Remote Sensing Letters*, 2020, **17**, (8), pp.1363-1367
8 J. Lehtinen, J. Munkberg et al.: "Noise2Noise: Learning image restoration without clean data,"*International Conference on Machine Learning*,2018, pp.2965-2974
9 L. Zhang, H.L. Li, et al.: 'A Fast BP Algorithm With Wavenumber Spectrum Fusion for High-Resolution Spotlight SAR Imaging' *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, 2014,**11**, (9), pp.1460-64
10 S.Y. Xia, Y. Huang, et al.: 'Adaptive anisotropic diffusion for noise reduction of phase images in Fourier domain Doppler optical coherence tomograhpy'.*Biomedical Optics Express*, 2016,**7**, (8), pp:2912-2926
11 Tang, S., C. Shen and G. Zhang: 'adaptive regularized scheme for remote sensing image fusion',*Frontiers of Earth Science*, 2016,**10**, (2), pp.236-244