A Research on Advanced Noise Suppression and Recognition Technology of SAR Images

Zhiyuan Fang

School of Computers and Artificial Intelligence, Beijing Technology and Business University, Beijing, China

2006010602@st.btbu.edu.cn

Abstract. Synthetic Aperture Radar (SAR) has been commonly utilized in both military and civilian fields in the past decade because of its all-weather, high-resolution, and strong penetration capabilities. However, due to its unique coherent imaging characteristics, speckle noise will be generated in SAR images, which greatly degrades the image quality and hinders subsequent target recognition work. In this paper, several denoising methods will be compared and estimated by indexes such as signal-to-noise ratio (SNR) and other indicators to estimate the model quality. Then the better method will be selected to preprocess SAR images and perform target recognition on the processed images. Based on the classic CNN model, a Squeeze-and-Excitation block will be added into the convolutional layer and train the model using part of the MSTAR dataset SAR images. During training, the SAR image samples are preliminarily divided into 8 classes, and the neural network is trained using the training set and validation set. When the training accuracy, validation accuracy, and validation loss value reach a relatively ideal value, training is stopped. Finally, the model is used to recognize the SAR image categories and compare the average recognition accuracy with the actual image categories. In addition, by comparing the average recognition accuracy before and after noise suppression, the experiments can also verify the performance of the method in SAR image noise suppression.

Keywords: SAR images, speckle noise, target recognition, evaluation model.

1. Introduction

Synthetic Aperture Radar is a type of radar installed on satellites and spacecraft [1]. It takes high-resolution images of the Earth's surface. Due to the different wavelength of the camera sensor, SAR can acquire images in a variety of weather conditions during the day and night, thus reducing the enhancement effect of photography and other optical imaging. The wavelength of SAR is longer than that of camera sensors, up to several meters, so SAR has the ability to penetrate clouds and ignore weather changes, which optical sensors cannot do [2]. Compared to Real Aperture Radar (RAR), SAR improves resolution by synthesizing a virtual large aperture radar.

Due to the imaging mechanism of SAR, coherent imaging will lead to speckle noise in the image, which is different from additive noise such as Gaussian noise. The interference term of multiplicative noise is random and has stronger time-varying characteristics and anti-filtering property, and the deviation caused in information transmission cannot be ignored. This influence will make the result of target detection very unsatisfactory and the recognition accuracy greatly decreased. Therefore, it is very important to select appropriate methods to suppress speckle noise in SAR images [3].

Firstly, this paper compares several noise suppression methods, including classical method (weighted median filtering, truncated median filtering) and frequency domain filtering (wavelet transform). Since wavelet transform has a significant effect on additive noise, it is necessary to pre-process the speckled noise of multiplicative noise. Logarithmic function can be used to process the original SAR image to convert it into additive noise. After preprocessing, wavelet decomposition is performed to obtain high frequency coefficient, threshold value denoising is obtained, high frequency coefficient is obtained after denoising, and finally image is reconstructed together with low frequency coefficient obtained by wavelet decomposition. This method has better denoising effect and can effectively retain edge information [4]. In addition, the morphological filtering method will be introduced, which is based on geometric mathematics to perform operations on the image, various transfers to geometric shapes, involving some knowledge of topology, set, corrosion and expansion are two classical operations. Edge information can be maintained in the case of short processing time. Combined with bitonic filter, it can maintain fast processing speed and robustness in the case of high noise [5]. Using the peak signal to noise ratio, image similarity and other indicators to evaluate, select the best filtering method for SAR image recognition.

Adding attention mechanisms, such as squeezing and excitation networks, to the convolutional layer of a convolutional neural network can provide more feature representations, i.e. get more strong semantic representations from radar data. The target's position similarity and confidence score are utilized to estimate its position, and the knowledge of optical images is used to improve the performance of radar target detection. Effectively integrating the weak model usually results in a more robust model with improved generalization ability [6].

2. SAR Imaging Related Theory

In SAR imaging, the main job of the antenna is to send microwave pulses to the surface of the Earth. SAR is the use of the basic principles of radar to produce images, using the delay of scattered signals. SAR determines the location of the target by measuring the time interval of the signal. Its resolution depends on the pulse width or pulse width. As the pulse width decreases, its resolution increases.

As shown in Fig 1, the carrier flies in a straight line in one direction, while sending out an electromagnetic beam for scanning, and the microwave pulse beam is in a cone shape, the cone shape is determined by the beam height Angle and the antenna length. Each beam forms a radiation band (ellipse) with the ground that can be seen as a collection of spatial elements and then scattered back to the radar to receive radar images.

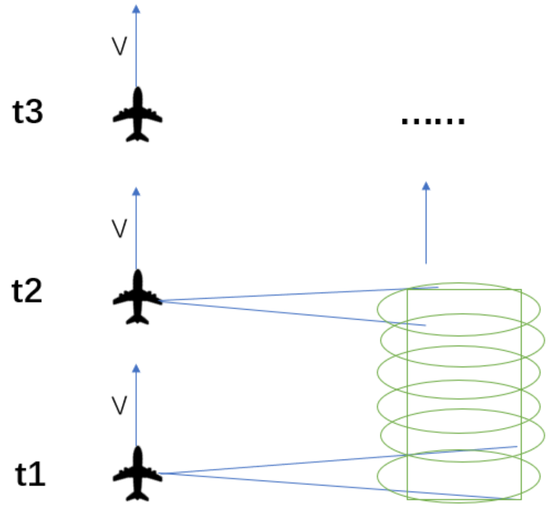

Figure 1. Overhead view of flight on board aircraft

The minimum distance between two adjacent objects, when the distance is distinguishable, it is called spatial resolution. There is a relative motion between the ground target and the radar during imaging, resulting in Doppler effect, which causes the frequency of the returned radar pulse to fluctuate. The frequency that the observer receives is greater than the frequency that the source emits when the distance between the source and the observer gets smaller, and the frequency that the source emits is lower when the distance between the source and the observer gets smaller. The higher the velocity of the wave source, the greater the Doppler shift effect. It is the use of Doppler effect in SAR image to measure the phase delay and frequency drift information at multiple points, synthesize a pulse, improve the resolution.

3. Typical Methods and Applications

3.1. Speckle Noise Model

Widely accepted model of coherent speckle noise is as follows:

$$I(x,y)=R(x,y)N(x,y) \tag{1}$$

In the above formula, (x,y) represents the direction coordinate and distance coordinate of SAR image scattering unit space. $I(x,y)$ represents image intensity (received image), $R(x,y)$ represents pure SAR image, $N(x,y)$ is independent of $R(x,y)$ and represents noise data.

Actually, the radar receives $I(x,y)$ after imaging, and needs to obtain $N(x,y)$ that is close to the ideal noiseless situation, and $R(x,y)$ is randomly varied.

The evaluation indexes commonly used to remove speckle Noise include ENL, PSNR, Signal to Noise Ratio (SNR), Structural Similarity (SSIM), and etc.

$$ENL=\frac{\mu^2}{\sigma^2} \tag{2}$$

ENL can measure the smoothness of A uniform region, μ representing the mean, σ representing the standard deviation, in different regions A and B with the same mean, if the degree of fluctuation of B data is smaller, σ will also become smaller, and the larger the ENL, the smoother the region.

3.2. Simple Morphological Theory

Corrosion and expansion are two important operations in morphology theory. The expansion operation, also known as Minkowski addition, is performed by sliding the structural element across the entire image and placing it on each pixel to ensure that the pixel of the structural element is equal to at least one pixel in the corresponding pixel of the image, thus preserving the value of this pixel and achieving boundary expansion.

The expansion operation process of a binary image is roughly shown in Fig 2.

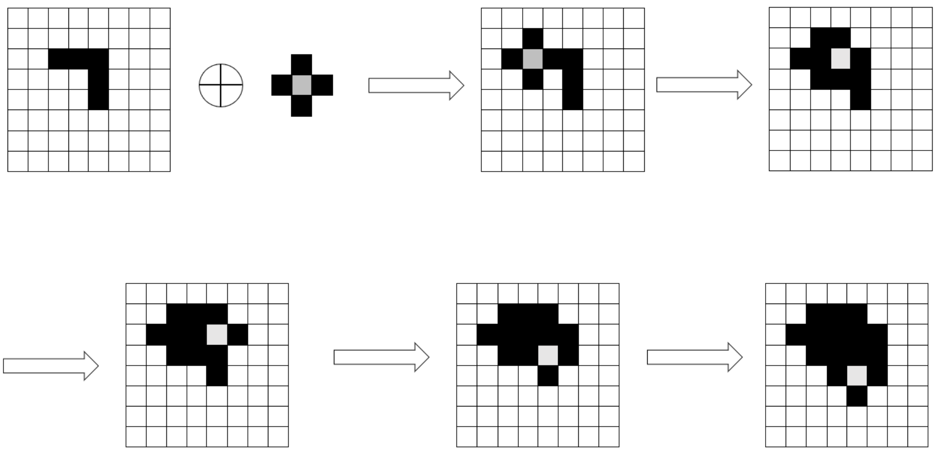

Figure 2. The expansion of binary image

As shown in Fig 2, the set of "T" elements is expanded by the set of cross-shaped elements. Scan all the pixels of the "T" element set, traverse the structure element in the image, and carry out the "and" operation on the structure element and its covered pixels to extend outwards. Using this characteristic, the holes and cracks in the original image can be filled effectively by expansion operation processing.

In contrast to the expansion operation, the corrosion operation is a subtraction of a vector. The corrosion operation can eliminate the boundary pixels of the target object, make the boundary shrink inward, filter out the size of the object smaller than the structural element, and finally achieve the size removal of the object. When there is a small connection between two objects, it can be separated by corrosion.

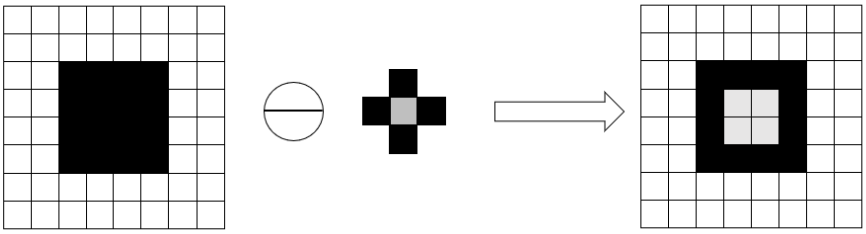

Figure 3. The corrosion of binary image

According to Fig 3, corrosion operation can reduce the edges of objects in the image. Although the corrosion operation is a vector subtraction between sets, it is different from Minkowski's subtraction, so the expansion operation and corrosion operation are not inversely related to each other, and can be used in cascade. The process of corrosion followed by expansion is called open operation, which can effectively eliminate small objects, separate small targets, and keep the boundaries of large objects smooth without significantly changing their area. The operation of expanding before corroding is called the closed operation and helps to remove small black areas.

Corrosion and expansion operations can be used to remove small specks and noise from the original image, in addition to image enhancement (applied in computer vision, photography) and feature extraction (applied in pattern recognition, neural network training). Enhance edges and details without changing image resolution, and extract features from local areas.

However, morphological reconstruction depends very much on the selection of structural elements, and the results may produce large fluctuations, and involve a lot of calculations, which requires high hardware and software of the equipment.

3.3. Truncation Algorithm

The truncation algorithm used in this paper is as follows:

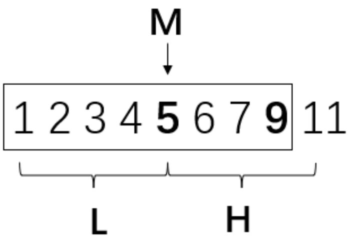

Figure 4. Truncation process

Firstly, the values around the element are arranged in ascending. Then it is necessary to find the minimum, median M, and maximum, where the difference between the median and the minimum is L, and the difference between the maximum and the median is aH. As shown in Fig 4, L is 4, H is 6, and M is 5. Next, compare the size relationship between L and H. If L<H is the case in the figure, the first element that is greater than (or equal to) L+M=9 is found in the order of arrangement (ascending order), and then truncated. After truncation, element 11 in the figure is excluded, the length of the queue of elements is shortened, and finally the so-called filtered value 4 is found in the median of the new queue. If L>H, find elements less than (or equal to) M-H in descending order, and then truncate and find the median in the similar way of L<H.

The changes before and after truncated median filtering are shown in Fig 5.

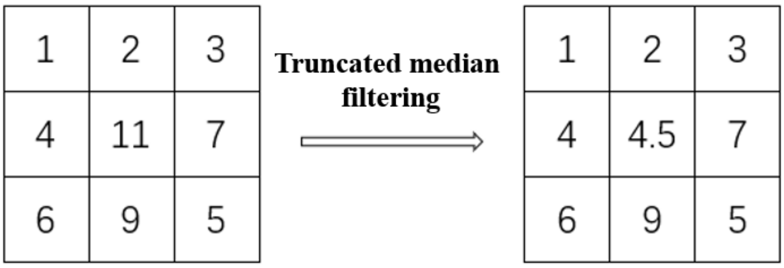

Figure 5. Truncated median filtering operation

3.4. Squeeze-and-Excitation Block

In this net, input layer is the first part of the CNN architecture, and the second part is composed of the combination of convolutional layer and pooled layer, and the third part is composed of a fully linked multi-layer perceptron classifier, the convolution kernel is both, the first two parts each have 32 filters, the last part has 64 filters, and the fully connected layer has 128 features.

The core operation is divided into three parts: squeeze, excitation and reweight.

Each two-dimensional feature channel is compressed to a real number in the squeeze operation, which has feature compression along the spatial dimension and an output dimension that nearly matches the input feature channel number. It demonstrates how responses are distributed globally via diverse routes.

An excitation operation is a periodic neural network mechanism that functions likewise to a gate. In order to explicitly represent correlations between feature channels, parameters are learned and used to construct weights for each feature channel.

After each feature channel has been selected, the Reweight procedure uses the Excitation output weight to determine its significance. Then, by multiplying each channel by its prior feature channel, the original features' channel dimensions are re-calibrated [7].

4. Experimental Results

4.1. Datasets

SAR works in the X-band (frequencies between 8 and 12 GHz, is a microwave in the electromagnetic spectrum). Resolution of SAR images are 0.3m×0.3m and pixel size is 128×128. Azimuth varies from 0 to 360 degrees This paper MSTAR is primarily focused on target identification and classification and uses eight classes of military vehicles, including bulldozers, tanks, trucks, artillery and armored carriers [8, 9].

4.2. Results

This experiment used 30 images in BTR_60 in MSTAR dataset were processed by morphological reconstruction filter, adaptive weighted median filter, truncated median filter, Gaussian filter and wavelet transform, and the average values of MSE, SNR, PSNR and SSIM were calculated. The calculation results and filtered images are shown in Table 1 and Fig 6.

Table 1. Performance comparison of several denoising filters on SAR images of BTR_60 armored vehicles

Filter type	MSE	SNR	PSNR	SSIM
Morphological reconstruction	146.83	19.36	26.46	0.95
Adaptive weighted median	178.48	18.51	26.61	0.92
Truncated median	265.11	16.79	23.90	0.90
Gaussian	337.01	15.75	22.85	0.87
Wavelet transform	152.68	19.02	27.34	0.94

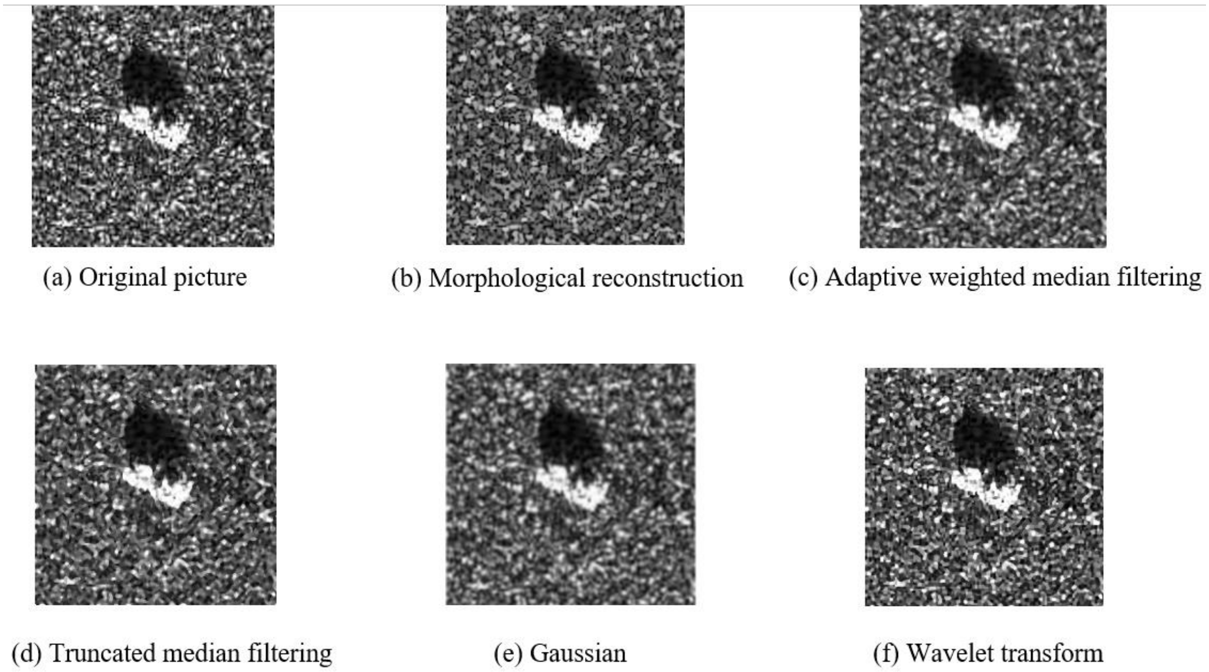

Figure 6. Comparison of denoising results of SAR images of BTR_60 armored vehicles

The results of data comparison show that the morphological reconstruction filtering method is superior to the other four methods, and there is little difference between the morphological reconstruction filtering method and the wavelet transform filtering method. Adaptive weighted median filter, truncated median filter and Gaussian filter have little difference in PSNR and SNR, and the MSE of Gaussian filter is significantly higher than that of other methods. However, all SSIM values are above 0.9 except the Gaussian filter. It confirms a trend that the traditional methods such as the Gaussian filter, mean filter and median filter are not satisfactory, but the neural network and morphological filter have better performance

Table 2 and Table 3 respectively show the average recognition accuracy rate of CNN for various military targets in MSTAR data set before and after denoising (300 samples):

Table 2. Average recognition accuracy of SAR images (before denoised)

Types	Accuracy
BTR_60	92.68%
2S1	90.28%
BRDM_2	84.35%
D7	79.85%
SLICY	93.42%
T62	82.41%
ZIL131	87.40%
ZSU_23_4	90.12%

Table 3. Average recognition accuracy of SAR images (after denoised)

Types	Accuracy
BTR_60	92.76%
2S1	91.04%
BRDM_2	86.96%
D7	82.34%
SLICY	94.15%
T62	83.79%
ZIL131	87.77%
ZSU_23_4	91.50%

By comparing Table 2 and Table 3, the average recognition accuracy of images of all kinds of vehicles is slightly improved. It can be seen that pre-processing before target recognition of a large number of samples is indispensable. In general, the effect of morphological filtering on SAR image noise suppression pre-processing is positive.

5. Conclusion

This paper discusses the effect of four nonlinear filtering methods on SAR image noise suppression and uses the CNN model added squeeze and Excitation block to train and try to target recognition and classification of SAR images in the MSTAR dataset. In the aspect of denoising, the morphological reconstruction filtering method has the best performance, the structure similarity is ideal and the image resolution is maintained. In terms of target recognition, after 100 rounds of training, the average accuracy of eight categories of recognition has reached more than 90%, the highest is 93.42%, and the recognition performance of the D7 category is poor, only slightly more than 80%. Compared with the average target recognition accuracy before and after denoising, the average target recognition accuracy after denoising is obviously higher, and the highest average recognition accuracy is as high as 94.15%. The recognition results before and after denoising also show that morphological reconstruction is more effective in suppressing speckle noise. The lack of background correlation between the training and test images in the MSTAR database, and the fact that all test images are known a priori, affects the fine-tuning process of each method and thus the true effectiveness of each technique. Therefore, you need to create a fixed benchmark dataset from the MSTAR dataset or try using another dataset in experiments.

References

[1] Franceschetti G, Lanari R. Synthetic Aperture Radar Processing, Second Edition. CRC Press, Inc. 2016.

[2] Singh, P., Diwakar, M., Shankar, A., Shree, R., & Kumar, M. A Review on SAR Image and its Despeckling. Archives of Computational Methods in Engineering, 2021: 4633-4653.

[3] Li D, Yu W, Wang K, et al. Speckle noise removal based on structural convolutional neural networks with feature fusion for medical image [J]. Signal Processing: Image Communication, 2021, 99: 116500.

[4] Chunhua Wang, Fangchao Wang, Wavelet denoising method of SAR image based on improved threshold function [J]. Microelectronics and computers, 2022, 39 (5): 39-44.

[5] G. Treece, "Morphology-Based Noise Reduction: Structural Variation and Thresholding in the Bitonic Filter," in IEEE Transactions on Image Processing, vol. 29, pp. 2020: 336-350,

[6] Sun P, Niu X, Sun P, et al. Squeeze-and-excitation network-based radar object detection with weighted location fusion [C] // Proceedings of the 2021 International Conference on Multimedia Retrieval. 2021: 545-552.

[7] CAI ZONG, PEIRUI BAI, QINGYI LIU, et al. An Efficient Model for Dorsal Hand Vein Recognition Based on Combination of Squeeze-and-Excitation Block and Vanilla ResNet [C]. // Proceedings of 2nd international conference on artificial intelligence, robotics, and communication: ICAIRC 2022: 2nd international conference on artificial intelligence, robotics, and communication (ICAIRC 2022), 25-27 November 2022, Fuzhou, China. 2023: 215-224.

[8] Y. Gu, J. Tao, L. Feng and H. Wang, "Using VGG16 to Military Target Classification on MSTAR Dataset," 2021 2nd China International SAR Symposium (CISS), Shanghai, China, 2021: 1-3.

[9] S. Jacob, J. Wall and M. S. Sharif, "Analysis of Deep Neural Networks for Military Target Classification using Synthetic Aperture Radar Images," 2023 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakheer, Bahrain, 2023: 227-233.