SAR Target Recognition with Generative Adversarial Network (GAN)-based Data Augmentation

2021 13th International Conference on Advanced Infocomm Technology (ICAIT) | 978-1-6654-3188-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICAIT52638.2021.9701974

Jiayuan Kong
Key Laboratory of Radar Imaging and Microwave Photonics (Ministry of Education)
Nanjing University of Aeronautics and Astronautics
Nanjing, China

Fangzheng Zhang*
Key Laboratory of Radar Imaging and Microwave Photonics (Ministry of Education)
Nanjing University of Aeronautics and Astronautics
Nanjing, China
zhangfangzheng@nuaa.edu.cn

***Abstract*—In convolutional neural network (CNN)-based synthetic aperture radar (SAR) target recognition, a large number of SAR image datasets are required to train the CNN. To deal with the problem of insufficient training datasets, a generative adversarial network (GAN) can be applied for data augmentation. In this paper, the performance of SAR target recognition with GAN-based data augmentation is investigated using the MSTAR datasets. The results show that, by applying the GAN-based data augmentation using a proper data augmentation factor, a target recognition accuracy as high as 99.38% can be achieved, which is improved by 2.35% compared to the condition without data augmentation. Thus, the GAN-based data augmentation provides a promising solution to enhancing the accuracy of SAR target recognition.**

***Keywords*—Convolutional neural network (CNN), synthetic aperture radar (SAR), data augmentation, generative adversarial network (GAN), target recognition**

I. INTRODUCTION

Synthetic aperture radar (SAR) is an active, high-precision imaging radar capable of working in all-day and all-weather conditions [1-4]. Target recognition based on SAR images has been investigated widely in the last few decades. Since the SAR images are poorly resistant to interference and has a single color, they are not as clear and legible as the optical images. Traditional SAR target recognition methods are mainly realized based on the statistical and physical features of the image for manual modeling. While, these methods are difficult to be adapted to the complex and variable nature of the SAR target images [5]. In recent years, convolution neural network (CNN) has been widely used in the field of radar target recognition [6, 7]. Compared with traditional target recognition methods, CNN-based target recognition method can well extract the features of the target from the SAR images and achieve a high recognition accuracy [8]. Usually, a large number of training datasets are needed to train the CNN. While, the SAR images are difficult to acquire, which constrains the amounts of training datasets. As a result, the overfitting problem originated from a small amount of training datasets would limit the generality of the CNN-based SAR target recognition [9]. To deal with this problem, many SAR data augmentation methods have been proposed, which are implemented by means of rotation, cropping, and so on [10]. In 2014, the generative adversarial network (GAN) is proposed to generate data samples that are different from but have nearly the same data distributions with the existing datasets [11]. Therefore, GAN provides a feasible solution for data augmentation in SAR target recognition.

In this paper, a new GAN-based data augmentation method for SAR target recognition is proposed and experimentally validated using the MSTAR datasets. Through experimental analysis, the effectiveness of improving the accuracy of SAR target recognition by leveraging GAN-based data augmentation is verified.

This work was supported by the NSFC program (61871214), the NSFC program of Jiangsu Province (BK20180066), and the Jiangsu Provincial Program for High-level Talents in Six Areas (DZXX-005).

II. PRINCIPLE

Fig. 1 shows the schematic diagram of the CNN-based SAR target recognition method. The SAR image is used as the input of a pre-trained CNN model, which extracts the features of the SAR image and finally classifies the image to get the classification result.

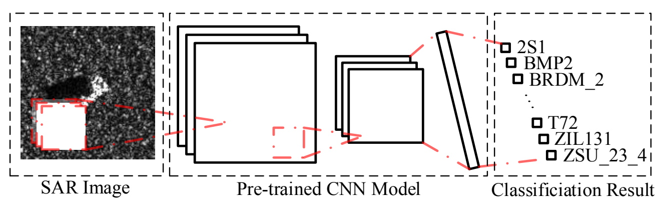

Fig. 1. Schematic diagram of the CNN-based SAR image recognition.

Fig. 2 shows the schematic diagram of the GAN-based data augmentation. The whole process is divided into three parts: GAN model training, data augmentation, and CNN model training. Firstly, the original SAR images and original image labels are used as the training datasets to train the GAN model, with the optimized GAN denoted by "pre-trained GAN model". Then, random tensors and custom labels are input to the pre-trained GAN model to obtain new SAR images denoted by "generated SAR images", which are similar to but not identical to the original SAR images. Then, the generated SAR images and the original SAR images are mixed to get the enlarged training datasets. Finally, the CNN model is trained using the augmented training datasets. The optimized model after training is denoted by "pre-trained CNN model", which is used for SAR target recognition.

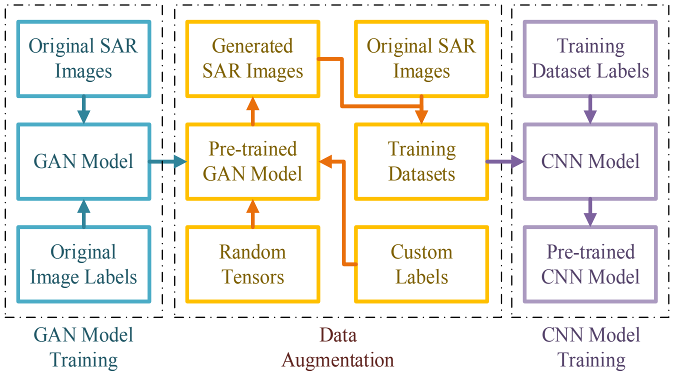

Fig. 2. Schematic diagram of the GAN-based data augmentation method.

III. Experimental Demonstration

A. Datasets Preparation

In our investigation, MSTAR datasets are adopted as the original datasets. The images are uniformly scaled to a size of 128*128. The MSTAR datasets include 10 classes of vehicles as the targets, as shown in Table I. These SAR images are obtained with pitch angles of 17 degrees and 15 degrees, respectively. The total number of SAR images is 3203, and the detailed numbers of images of each target used as the training datasets and the testing datasets are also shown in Table I.

TABLE I. Different Targets and the Number of Images Used for Training and Testing for the Original Datasets

SAR Target	Training Datasets	Testing Datasets
2S1	299	274
BMP2	233	195
BRDM_2	298	274
BTR_60	256	195
BTR70	233	196
D7	299	274
T62	299	273
T72	232	196
ZIL131	299	274
ZSU_23_4	299	274

B. Experiment Environment

The processor used for this experiment is an Intel(R) Core (TM) i7-7700HQ CPU @ 2.80GHz, with 8G of RAM and an NVDIA GeForce GTX 1050 graphics card. The deep learning framework of PyTorch1.7.0 is adopted to construct the deep learning model.

C. Model Selection And Parameter Setting

In our implement, the auxiliary classifier generative adversarial network (ACGAN) [12] is adopted to perform data augmentation, and the ResNet34 model [13] is applied as the CNN model used for SAR target recognition. Compared with the traditional GAN, ACGAN can combine specific image labels to generate specific kinds of images, and ACGAN contains a classifier that can judge the generated images to further improve the quality of the generated images. Fig. 3 shows the schematic diagram of the ACGAN. The random tensors and custom labels are input to the generative network (G) to obtain a set of generated images. Then, the original image labels and original images and generated images are input to the discriminative network (D) to obtain the discriminant scores and classification results, respectively. Finally, the model parameters of the G and D are updated by calculating the total loss from these the discriminant scores and classification results.

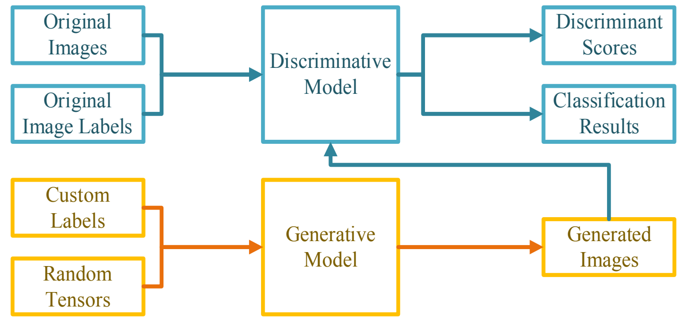

Fig. 3. Schematic diagram of the ACGAN.

Fig. 4 shows the network structure of G model and D model in our ACGAN. The input of the G model is a 100-dimensional random tensors, which is concatenated with a 10-dimensional custom labels to obtain a 110-dimensional input tensors. The 110-dimensional tensors are passed through the fully connected (FC) layers, the batch normalization (BN) layers and the ReLU layers to obtain tensors with a channel of (640*8*8, 1, 1). The output tensors are reshaped to tensors having a channel of (640, 8, 8). Then, a series of transpose convolution (Trans-Conv) layers, BN layers and ReLU layers are used to obtain tensors with a channel of (1, 128, 128). Finally, the generated images are obtained by nonlinear activation of the Tanh layer. Meanwhile, the one-dimensional SAR images with a size of 128*128 are input to the D model. After a series of convolution (Conv) layers, BN layers and LeakyReLU layers, tensors with one channel of (640, 8, 8) are obtained. Then, the tensors are reshaped to tensors having a channel of (640*8*8, 1, 1). Finally, two channels of (128, 1, 1) and (10, 1, 1) are obtained after different FC layers, respectively. The tensors with a channel of (128, 1, 1) are activated by Sigmoid nonlinearity to the final discriminant scores, and the tensors with a channel of (10, 1, 1) are output as the classification results. The size of the convolutional kernel of the ACGAN model is set to 4, the step size is set to 2, and the padding is set to 1. Thus, the size of the tensors become half of that of the original tensors for each convolutional operation, and the size of the tensors become twice of that of the original tensors for each transpose convolutional operation.

The training parameters of the ACGAN model are as follows. The number of samples selected for one training batch-size is set to be 64. The optimizer used for ACGAN model is adaptive moment estimation (Adam). The loss function is the cross entropy (CE) loss function. The learning rate is set to 2e-4 and the momentum betas are set to (0.5, 0.999).

When training the ResNet34 model, the training epoch is set to 50, and the number of samples selected for one training batch-size is 64. The learning rate is set to 1e-4, and the random seed is set to 20. The optimizer uses Adam, and the loss function is also the CE loss function.

with higher performance is applied, the time for generating images with the ACGAN can be further shortened.

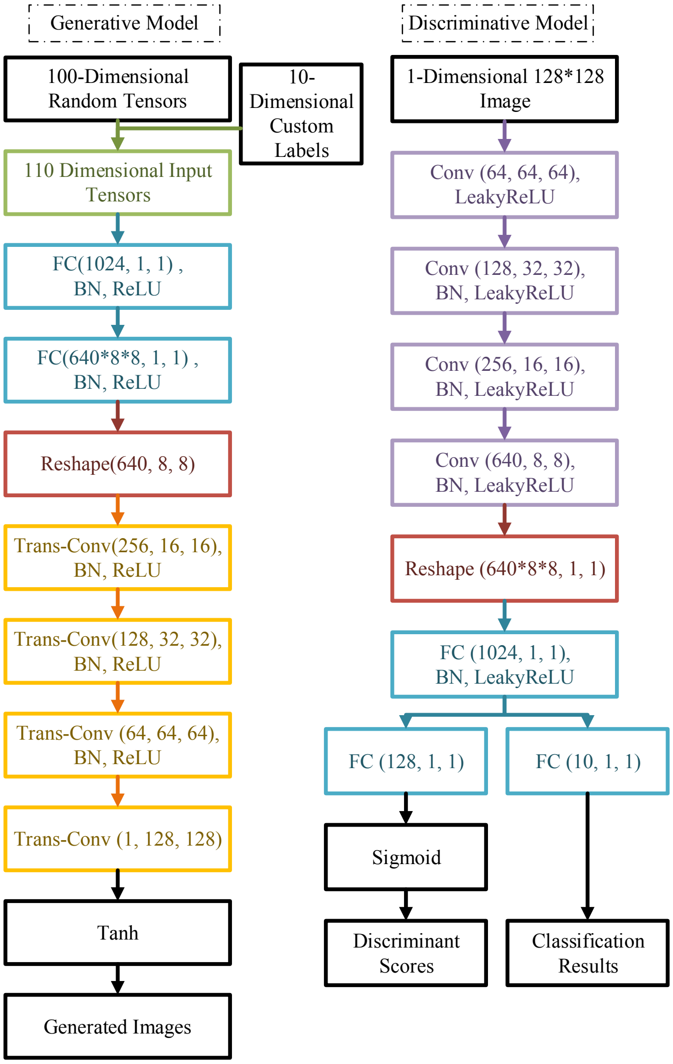

Fig. 4. The network structure of G model and D model.

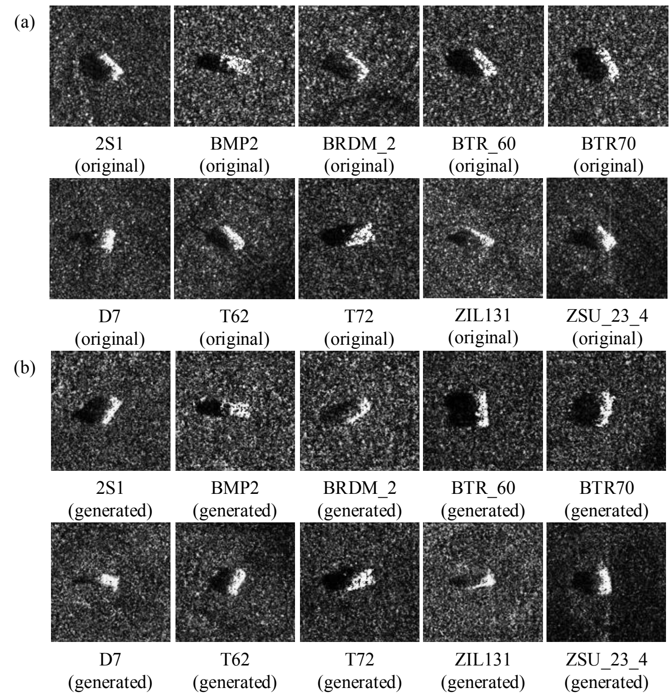

Fig. 5. (a) Selected images from the original SAR image datasets, and (b) the generated SAR images using the ACGAN model after 150 epochs of training.

D. Experiment Results

In the experiment, the image labels of the ten classes of targets (2S1, BMP2, BRDM_2, BTR_60, BTR_70, D7, T62, T72, ZIL131, ZSU_23_4) are specified to be 0, 1, 2, …, 9, respectively. The original SAR images and the SAR image labels are used to train the ACGAN model. After 150 epochs of training, the features of generated SAR images by the ACGAN are nearly the same as the features of original images. The obtained ACGAN model, i.e., the pre-trained ACGAN model, can be applied to generate SAR images of different targets. Fig.5 (a) shows one selected image of each target from the original SAR image datasets. Fig.5 (b) shows the generated SAR images using the ACGAN model after 150 epochs of training. As can been seen, the generated SAR images are as clear as the original SAR images, and they have very similar features compared with the original images. It should be noted that, the generated SAR images are not exactly the same as the original SAR images. For example, the azimuth of the each target in the images may be different, and the coherent speckle noise of the SAR images is also different. Here, generating the same amount of images with that of the original training datasets takes 160 s. If a computer

As a comparison, the SAR target recognition using original datasets without data augmentation is first investigated. The measured accuracy considering all the 10 classes of targets is 97.03%. Then, the pre-trained ACGAN is used to generate SAR images, of which the total number is the same with that of the original training datasets for each target (the data augmentation factor is 1). Here, the data augmentation factor is defined as the ratio between the number of images generated by the ACGAN and the number of images in the original training datasets. The generated SAR images are combined with the original training images to obtain the new SAR training datasets, which is 2 times of the number of original training datasets. The testing datasets remain unchanged. When using the new training datasets to train the ResNet34 model, the final testing accuracy for SAR target recognition reaches 98.80%. Compared with the target recognition without data augmentation, the accuracy is improved by 1.77%. The results verify that the GAN-based data augmentation can effectively improve the target recognition accuracy while avoiding the overfitting problem after expanding the training datasets.

Finally, the target recognition performance with different data augmentation factors is investigated. Specifically, the data augmentation factor changes from 2, 3, 4, …, 9, corresponding to the case the new SAR image training datasets are expanded to 3, 4, …, and 10 times of the original training datasets, respectively. For each cases, the ResNet34 model is retrained, and the target recognition accuracy is recorded, as shown in Table II. In Table II, the results corresponding to the cases of augmentation factor being 0 (without augmentation) and 1 are also included. The relationship between the data augmentation factor and the target recognition accuracy is also plotted in Fig.

6. As shown in Table II and Fig.6, when the data augmentation factor is 2, the highest accuracy of 99.38% is achieved, which is 2.35% higher than the accuracy without data augmentation. When the data augmentation factor are greater than 2, although the target recognition accuracy are improved compared with the case without data augmentation, the accuracy is reduced gradually. This is caused by the fact, we use the ACGAN to generate SAR images with features approaching the those of the original images. Due to the unideal performance of the ACGAN, as the number of generated images increases, the new training datasets would inevitably possess features that deviate from the original datasets, which results in a reduction of the target recognition accuracy. Therefore, to achieve the best SAR target recognition performance, the data augmentation factor should be properly controlled.

In our demonstration, the GAN-based data augmentation is implemented based on the original SAR images. It should be noted that, the GAN-based data augmentation can be used together with other data augmentation methods, such as flipping or rotating the images. By using the combined data augmentation is expected to further improve the recognition accuracy.

TABLE II. TEST ACCURACY OF SAR DATASETS WITH DIFFERENT DATA AUGMENTATION FACTORS

Data augmentation factor	Recognition accuracy (%)
0 (Without augmentation)	97.03
1	98.80
2	99.38
3	98.97
4	99.05
5	99.13
6	98.89
7	99.09
8	98.64
9	98.52

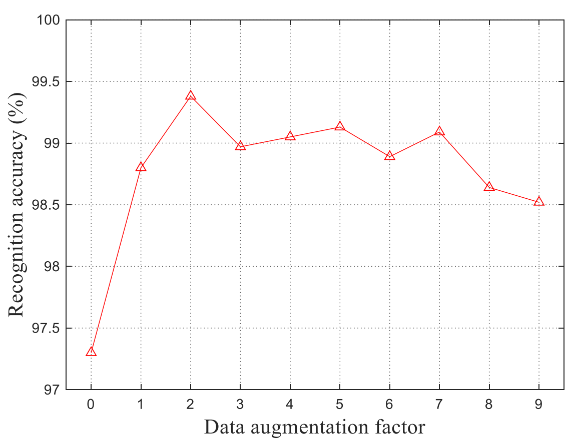

Fig. 6. The relationship between the data augmentation factor and the target recognition accuracy.

IV. CONCLUSION

We have investigated the performance of SAR target recognition with GAN-based data augmentation. Results verify the feasibility of the GAN-based data augmentation in improving the accuracy of SAR target recognition. Specifically, based on the MSTAR datasets, a high target recognition accuracy of 99.38% is achieved using the GAN-based data augmentation, which is 2.35% higher than the accuracy without data augmentation. In addition, the investigation results reveal that, to achieve the best SAR target recognition performance, the data augmentation factor, i.e., the amount of training data generated by the GAN, should be properly controlled. This paper provides new ideas for the field of SAR target recognition with data augmentation.

REFERENCES

[1] Y. Jin, “Multimode remote sensing intelligent information and target recognition: Physical intelligence of microwave vision,” Journal of Radars, vol. 8, no. 6, pp. 710–716, 2019.

[2] F. Zhang, Q. Guo, Z. Wang, P. Zhou, G. Zhang, J. Sun, and S. Pan, “Photonics-based broadband radar for high-resolution and real-time inverse synthetic aperture imaging,” Optics Express, vol. 25, no. 14, pp. 16274-16281, 2017.

[3] X. Ye, F. Zhang, Y. Yang, D. Zhu, and S. Pan, "Photonics-based high-resolution 3D inverse synthetic aperture radar imaging," IEEE Access, vol. 7, pp. 79503-79509, 2019.

[4] F. Zhang, Q. Guo, Y. Zhang, Y. Yao, P. Zhou, D. Zhu and S. Pan, “Photonics-based real-time and high-resolution ISAR imaging of non-cooperative target,” Chinese Optics Letters, vol. 15, no. 11, pp. 112801, 2017.

[5] W. Guo, Z. Zhang, W. Yu, “Perspective on explainable SAR target recognition,” Journal of Radars, vol. 9, no. 3, pp. 462–476, 2020.

[6] Y. Liu, Y. Zhou, X. Liu, F. Dong, W. Chang, Z. Wang, “Wasserstein GAN-Based Small-Sample Augmentation for New-Generation Artificial Intelligence: A Case Study of Cancer-Staging Data in Biology,” Engineering, vol. 5, no. 1, pp. 156-163, 2019.

[7] J. Li, F. Zhang, Y. Xiang, and S. Pan, “Towards small target recognition with photonics-based high resolution radar range profiles,” Optics Express, vol. 29, no. 20, pp. 31574-31581, 2021.

[8] A. Krizhevsky, I. Sutskever, G. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, pp.1097-1105, 2012.

[9] G. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors,” Computer Science, vol. 3, no. 4, pp. 212-223, 2012.

[10] S. Yun, D. Han, S. Chun, S. Oh, Y. Yoo, J. Choe, “CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features,” 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 6022-6031, 2019.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, “Generative adversarial nets,” Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, pp. 2672-2680, 2014.

[12] A. Odena,C. Olah, J. Shlens, “Conditional image synthesis with auxiliary classifier GANs,” Proceedings of the 34th International Conference on Machine Learning - Volume 70, pp. 2642-2651, 2017.

[13] K. He; X. Zhang; S. Ren; J. Sun, “Deep Residual Learning for Image Recognition,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.