International Journal of Applied Mathematics in Control Engineering 6 (2023) 151-157

Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http:www.ijamce.com

Research on AdvGAN+ Adversarial Samples Generation Method Based on Conditional Information

Yongqiang Zhang[a,b,c], Zewei Ji[a], Xiaohan Sun[a], Jinlong Ma[a,b,c], Weidong Wu[a,*]

[a] *School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China*

[b] *Hebei Technology Innovation Centre of Intelligent IoT, Shijiazhuang, Hebei 050018, China*

[c] *Shijiazhuang Intelligent Communication IoT Industrial Technology Research Institute, Shijiazhuang, Hebei 050018, China*

ARTICLE INFO

Article history:
Received 21 May 2023
Accepted 29 June 2023
Available online 2 July 2023

Keywords:
Adversarial examples
Generative Adversarial Networks
Artificial Neural Networks
Artificial Intelligence

ABSTRACT

Due to the existence of adversarial samples, deep neural networks with higher confidence levels are misjudged, which poses a great threat to the actual deployment of artificial intelligence systems. Research on the generation method of more aggressive adversarial samples can help build a more stable and more safe neural network model. Aiming at the problems of the current generation method of adversarial sample methods, such as poor image quality and slow generation speed, this paper proposes a method of adversarial sample generation based on the generative adversarial network AdvGAN+ on the basis of AdvGAN. Use the generator that generates the adversarial network to add small disturbances to the original data, and then add a condition information as a guide. After adding the conditional information, the generator and discriminator can converge to the target result faster and generate higher quality Image, and the resulting adversarial samples are more aggressive. By comparing with known adversarial sample generation algorithms, experiments on the Mnist dataset show that the proposed AdvGAN+ model has improved aggression.

1. Introduction

Artificial intelligence is a research hotspot in the computer field, among which deep learning (e.g., Lecun et al., 2015) and neural networks (e.g., Li et al., 2020) have made remarkable achievements, and at the same time, security problems of artificial intelligence have also arisen. Studies have shown that deep neural networks (e.g., Hashimoto et al., 2019) are vulnerable to attacked by adversarial samples, which cannot be detected by human beings and greatly affect the normal operation of intelligent systems. Therefore, research on adversarial samples is of great significance for improving the security of neural networks (e.g., Chen et al., 2021).

By adding small perturbations (e.g., Won et al., 2019) to the data set, adversarial samples are formed. The input samples after the perturbations will cause the target model with high confidence (e.g., Dash et al., 2017) to give an incorrect output. This method of deceiving the neural network is called adversarial samples. Adversarial sample research has gained close attention in the field of security. If artificial intelligence system makes misjudgment, it may cause security problems, which pose a great threat to the deployment of artificial intelligence system (e.g., Fischer et al., 2021), such as the identification of pedestrian road signs by autonomous driving, face recognition, security system (e.g., Zhang et al., 2021; Ravi Prakash et al., 2021; Gao et al., 2019). The vulnerability of deep learning to adversarial samples is not unique, and this phenomenon is common in many machine learning models (e.g., You et al., 2021). Therefore, in-depth study of adversarial samples will promote the progress of the whole field of machine learning.

At present, adversarial sample generation methods are mainly divided into two categories: Optimization based method and gradient based method (e.g., Wang et al., 2020). Gradient-based generation methods are common, and many scholars have studied adversarial samples of deep learning models. For example, Szegedy et al. (e.g., Szegedy et al., 2013) added subtle and imperceptible disturbances to artificial images to maximize the training of deep learning classification model, resulting in classification errors of the network model. Goodfellow et al. (e.g., Goodfellow et al., 2015) proposed FGSM (Fast Gradient Sign Method) algorithm based on gradient to generate adversarial samples. Moosvi-Dezfooli et al. (e.g., Moosavi-Dezfooli et al., 2016) proposed a DeepFool attack based on the generation of adversarial samples classifier decision boundary (e.g., Fayed et al., 2021).

FGSM and Deepfool are two classical attacks adversarial the sample generation strategy: one-step attack (e.g., Raaijmakers et al., 2019) and iterative attack (e.g., Shi et al., 2020), which are relatively

* Corresponding author.
E-mail addresses: wwd@hebust.edu.cn (W. Wu)
doi:

traditional attack methods. On the basis of the original data, by constantly accessing the information output of the target classification model, constantly computing disturbances, and generating new antagonistic samples through superposition. As a result, the target classification model will misjudge, and the structure and other information of the target classification model need to be constantly obtained to calculate the disturbance, which has weak practicability and mobility (e.g., Biggio et al., 2012). So we propose a Generative Adversarial Networks, GANs based method to generate adversarial samples, by adding a conditional information to guide the whole process of training to fight the production of samples. AdvGAN+ is an improved network model with AdvGAN as the original model. This conditional information can be any type of auxiliary information, such as the category label of the image or the feature distribution information of its classification information (e.g., Xu et al., 2021). By adding guidance information, the quality and speed of adversarial sample generation can be guaranteed.

2. Related works

2.1 Gradient based adversarial sample generation method

At present, there are many methods to generate countermeasure samples. This section will introduce three typical countermeasure sample generation methods based on gradient attack strategy:

(1) Optimization method

CW (carkini&wagner) attack is an attack method based on optimization. It not only ensures high attack accuracy, but also takes into account two aspects of low resistance to disturbance. When the model gives the wrong classification, it is still imperceptible to the human eye. CW attack depends on the initial optimization form of countermeasure samples. The problem of image x finding countermeasure samples is defined as $\text{minimize}\mathrm{D}(x, x+\delta)$, Where x is fixed, and the goal is to find the countermeasure disturbance δ of minimizing the objective function $\mathrm{D}(x, x+\delta)$. The core of generating countermeasure samples is to find countermeasure disturbance, which makes the classification model misjudge. D are some distance measurement functions, such as $\mathrm{L}_0, \mathrm{L}_2, \mathrm{L}_\infty$.

(2) FGSM

Goodfellow et al. (e.g., Goodfellow et al., 2015) proposed a high-dimensional linear space theory (e.g., Lee et al., 2016) to explain the phenomenon of adversarial attacks, and proposed a fast gradient symbol attack algorithm FGSM. The algorithm uses the first-order approximate solution of loss function to generate anti-disturbance. For the original image x , the attacker uses the method of maximum loss function $\mathrm{L} = (x^*, \mathrm{y})$ to generate anti-sample $x^* = x + \eta$, where η represents small disturbance, the samples x^* after small disturbance do not get normal classification results in the target classification model, so as to achieve the purpose of adversarial attack. This method is called FGSM fast symbol attack.

(3) DeepFool

Moosvi-Dezfooli et al. (e.g., Moosavi-Dezfooli et al., 2016) proposed the Deepfool method that uses mathematical geometry ideas to generate adversarial examples. Look for the smallest disturbance r^* to move the sample onto the hyperplane. In the linear dichotomies problem (e.g., Rasche et al., 2017), DeepFool method looks for disturbances perpendicular to the hyperplane, as shown in Figure 1. For a function solving the image dichotomies task $\mathrm{f}_{(x)} = \mathrm{W}^{\mathrm{T}} x + b$, the model correctly predicts the point x_0 to the side $\mathrm{f}_{(x)} > 0$, if the attacker wants to adversarial attack on the point, making the model incorrectly predict the point x_0 to the side $\mathrm{f}_{(x)} > 0$ we need to add a well-designed counter disturbance at that point. The Deepfool method uses mathematical geometry ideas to provide a feasible strategy for the method. The distance to the classification boundary is regarded as the robustness of the classification model to the samples. To generate adversarial samples on this basis, the distance needs to be shortened and overcome. Therefore, Moosvi-Dezfooli et al. took this distance as the key factor to generate anti-disturbance. When the disturbance vector x_0 slowly pushed the point to the decision boundary, the robustness (e.g., Rasche et al., 2017) of the classification model became lower and lower until it was no longer robust and misclassification occurred, at which time the anti-attack of the classification model was completed.

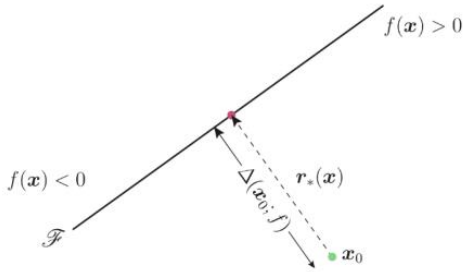

Fig. 1.. Adversarial disturbance in binary classification tasks

2.2 Generative Adversarial Network

In 2014, Goodfellow et al. (e.g., Goodfellow et al., 2014) put forward a peculiar network structure—generative adversarial network (GAN), which assessment the generation model by conducting adversarial training through the generation network and the discriminant network. Figure 2 is the network structure of GAN network.

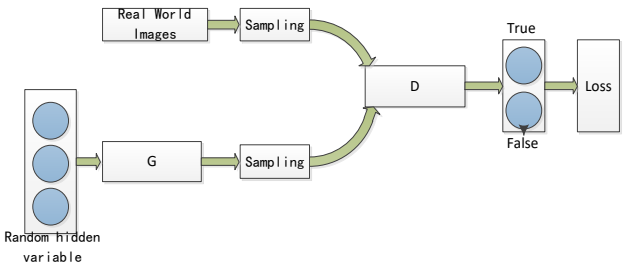

Fig. 2. The network structure of GAN

The learning mode of GAN adopts the unsupervised mode in machine learning (e.g., Kishore et al., 2022), so that the neural network can automatically obtain the structure of the original data, the generate sample is not different from the real data. As a generation mode of unsupervised learning (e.g., Wei et al., 2015), The structure of GAN consists of two parts, including generator model G and discriminator model D. Generator model study the real data distribution of sample data, and generate the generation of the sample with the real data distribution, judging device model responsible for judging the generator to generate samples of truth, it is controlled by the loss function of the two models by constantly adversarial training (e.g., Wei et al., 2019), the resulting data and real data are the same generation. The objective functions of G and D can be in the form of the maximum and minimum value functions:

$$\frac{min}{G}+\frac{max}{D}V(D,G)=E_{X\sim P_{data}(x)}\left[logD_{(x)}\right]+E_{z\times P_z(z)}\left[log\left(1-D(G_{(z)})\right)\right] \quad (1)$$

Where z is the Gaussian random noise (e.g., Deng et al., 2020), Grepresents the generation network, D represents the discriminant network, $P_{data}(X)$ and $P_z(Z)$ represent the probability distributions of the real data and random noise respectively. $x\sim P_{data}$ and $D_{(x)}$ represent the distribution from real data and the extraction from Gaussian random noise respectively. $G_{(z)}$ represents the output vectors of the discriminant network and the generation network.

3. Adversarial sample generation method

3.1 AdvGAN

In 2018, Chaowei Xiao et al. (e.g., Xiao et al., 2018) proposed a novel antagonistic sample generation method AdvGAN.

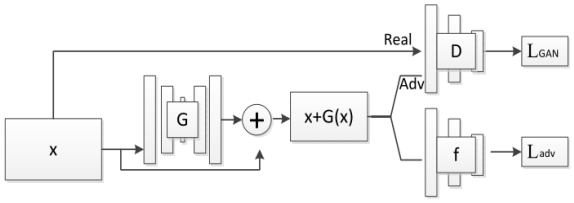

Fig. 3. The network of AdvGAN

As shown in Figure 3, an AdvGAN consists of three parts: a generator, a discriminator, and a target model. The generator accepts the initial input and adds perturbation, and then feeds it into the discriminator, which differentiates it from the initial data.

In classical white box attacks (e.g., Bai et al., 2018), such as optimization-based methods and gradiency-based methods, the attacker needs to conduct adversarial attacks through the architecture and internal functions of the model accessed by white boxes. However, by deploying AdvGAN, after the current feed network (e.g., Lee et al., 2021) is trained, it can directly add adversarial perturbation to the input samples without accessing the model, which improves the practicability and mobility of the adversarial samples.

3.2 AvdGAN+

AdvGAN+ model is based on AdvGAN. Unlike AdvGAN model, AdvGAN + adds an additional information c to both the generator and the discriminator, it can be any information, such as category label information or classified image feature data of its target model. The network structure of the model is shown in Figure 4, by conveying the condition information c to the discriminant model and the generation model take as a part of the input layer. The AdvGAN+ network model consists of the following parts: a generator G, a discriminator D, and a target classification model f. By inputting the original image into the generator, the generator generates counter disturbances, and then adds adversarial disturbances back to the original image. The image information is sent to the discriminator for authenticity determination, and finally an adversarial sample with high consistency with that real image is generated.

Input the clean sample x into G to generate adversarial perturbation, then send $x+G_{(x)}$ to D to differentiate the adversarial sample from clean image. The purpose of the discriminator D is to inspire the generated sample to be imperceptible from the data in the original class. At the same time, by adding conditional information c to guide the training, the generated results are more realistic. To achieve the purpose of deceiving the target model, the generated input $x+G_{(x)}$ to the target classification model, where the output loss is L_{adv}, which indicates the distance between the target attack and the target category t, and the target loss function is optimized. When the model reaches the optimal value, $G_{(x)}$ is best to adversarial disturbance.

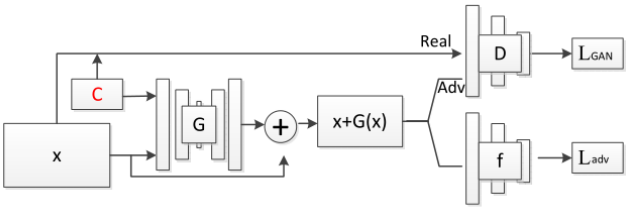

Fig. 4. The network structure of AdvGAN+

The objective loss function is divided into three parts: L_{GAN}, L_{adv}, L_{hinge}, represented by:

$$L=L_{GAN}+\alpha L_{adv}+\beta L_{hinge} \quad (2)$$

$$L_{GAN}=E_x\left[logD_{(x|\,c)}\right]+E_x[\log\left(1-D\left(x+G_{(x|\,c)}\right)\right) \quad (3)$$

$$L_{adv}=El_z(x+G_{(x|\,c)},t) \quad (4)$$

$$L_{\text{hinge}}=E_x\max(0,\|G_{(x|\,c)}\|_2-z) \quad (5)$$

Among them, the purpose of D is to differentiate the disturbance data $x+G_{(x|\,c)}$ from x, and the purpose of optimizing the anti-loss is to encourage the generated data distribution to be close to the original data. The difference from the formula in the original text is that the samples in the generator and the discriminator are both with the addition of information c, this is conducive to the GAN network in the generation process with a goal.

L_{adv} is the misjudgment loss, the purpose is to make the image generated toward the adversarial sample, the effective target attack disturbance image is classified as a t-type target, which represents the distance between the prediction with the target-type t.

L_{hinge} is the hinge loss. Its main purpose is to stabilize the training of GAN. z represents the limit specified by the user and can also stabilize the training of GAN.

4. Experiments

4.1 Experimental environment

The experimental platform is Windows7 operating system, CPU is Intel (R) Xeon (R) CPU E5-2630 v2 @2.60GHz, running memory is 64.0GB. The configuration environment of the experiment is Tensorflow 1.13.1 and Python 3.7. The public Mnist data set is used to verify the feasibility of the generation of adversarial samples.

4.2 Experimental design

The Mnist data set has 10 categories. For these categories, three methods of DeepFool, FGSM and AdvGAN are used for attack testing. LeNet-5 and VGG-16 are used as target classification models.

In the first stage, the GAN network needs to be trained. The purpose is to obtain a trained generator and discriminator, and then add the target classification model f to the network to form a complete

AdvGAN+ model. In the second stage, the conversion generator is started to learn the distribution of real samples, and under the guidance of conditional information, small disturbances are added to generate adversarial samples. In the training process, the target model and the discriminator are trained in turn to ensure that the adversarial samples are generated while also ensuring the authenticity of the adversarial samples.

4.3 Experimental results

In order to evaluate the effectiveness of the attack model AdvGAN+, four sets of comparative experiments were set up, Optimization method, DeepFool, FGSM and AdvGAN were used to train the Mnist dataset to generate adversarial samples, and the results were compared by sending the training results to the target classification model. Three evaluation criteria were used as the indicators of the comparative experiment, the accuracy of the model, the quality of generated samples and the time of generated samples. Compared with the other three attack strategies, these adversarial examples are closer to the distribution of real data.

Four methods of adversarial attack were used to train the original data to generate adversarial samples, and then transferred to the unified target model to conduct adversarial attacks. The experimental results shown in Table 1 were obtained. The recognition rate in Table 1 is taken from the average of the results of multiple experiments. In multiple experiments, the same Mnist data set is used to train the three methods. The comparison of three adversarial sample generation algorithms shows that the experimental results show that on the Mnist data set, the attack rate of AdvGAN+ is significantly improved compared to Deepfool and FGSM.

Tab. 1. Adversarial sample comparison experiment results.

Model algorithm	LeNet-5		VGG-16	
	Target model accuracy			
	Original sample	Adversarial sample	Original sample	Adversarial sample
Deepfool	0. 9866	0. 3259	0. 9912	0. 2968
FGSM	0. 9729	0. 2998	0. 9845	0. 2762
AdvGAN	0. 9905	0. 1283	0. 9812	0. 1478
AdvGAN+	0. 9895	**0. 1142**	0. 9856	**0. 1247**

The comparison results are shown in Table 2. The FGSM algorithm is simple and has a small amount of calculation. It only needs one backpropagation and one gradient calculation to generate a disturbance, but the attack target cannot be specified, and there is a possibility that a disturbance with a bad effect will be generated. Deepfool's attack speed is faster than FGSM, and the disturbance is smaller than FGSM, but it intelligently looks for disturbances that make the sample cross the decision boundary. The attack of AdvGAN+ is that the original image adds a small disturbance under the guidance of condition information, so that the generated result is more realistic, and the attack performance is better than the above three algorithms.

The quality of several images generating countermeasure samples and the original images is calculated, and the generation quality of countermeasure samples is compared. SSIM method is used. The calculation results are shown in Table 3.

According to the data in the table, comparing the SSIM values of each group of pictures, the values of the classical adversarial sample generation method are less than the values of advGAN, and the SSIM value can effectively judge the visual quality of the generated image. The larger the value, the higher the image quality. Therefore, the image quality generated by the anti sample generation method of advGAN is higher. The condition information is guided by the category label information of MNIST data sets 1-9, so that each round of generation results proceed in this direction, and the generation results are shown in Fig. 5.

Tab. 2. Generate adversarial sample comparison.

Model algorithm / Optimization method	FGSM	Deelfool	AdvGAN	AdvGAN+
Original sample				
Adversarial sample				

Table. 3. SSIM Statistical table.

type / method		
Optimization method	0.735	0.721
FGSM	0.745	0.723
Deepfool	0.654	0.647
AdvGAN+	**0.825**	**0.827**

Tab. 4. SSIM Statistical Table.

method	Optimization method	FGSM	Deepfool	AdvGAN	AdvGAN+
Time(s)	8.62	5.54	6.48	0.36	**0.14**

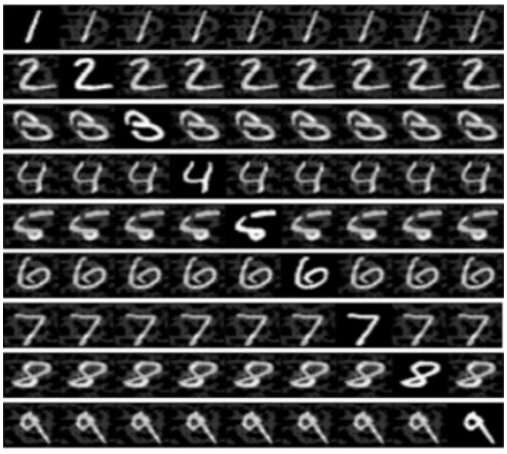

Fig. 5. Generates adversarial sample results of AdvGAN+

Since the GAN network itself has the problem of difficulty in convergence, it is chosen to increase the number of training rounds during training to ensure the true validity of the experiment. The experimental results are shown in Figure 6. After many adversarial trainings, the loss functions of the generator and the discriminator gradually converge, indicating that the generator can generate adversarial samples that the discriminator cannot distinguish between

true and false, which proves that the model training is completed.

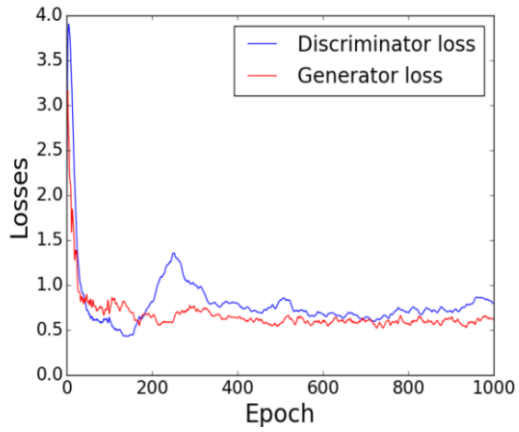

Fig. 6. Network training loss function of AdvGAN+

The experimental results are shown in Figure 7. Compared with the attack performance of advGAN, the advGAN network model with conditional information converges faster and has stronger attack effect on the target network. For the countermeasure samples with the same number of rounds, its recognition rate is lower, which shows that after adding conditional information C, the model is easier to generate real and effective countermeasure samples. Therefore, the countermeasure samples generated by advGAN can obtain a higher attack rate in a shorter time. Through the conclusion analysis and comparison, the time required for each generation of 10 countermeasure samples is analyzed and compared. The results are shown in Table 4.

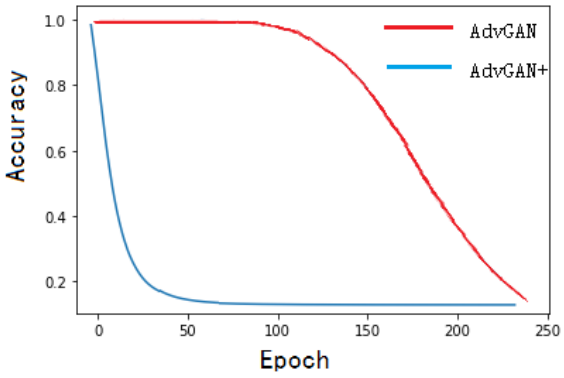

Fig. 7. Adversarial Attack Training Results of AdvGAN+

Among the classical attack methods, the fastest FGSM method takes about 5.54 seconds. The advGAN method based on GAN calculates the disturbance and superimposes it with the original image. It improves by an order of magnitude based on the classical method of generating countermeasure samples, which takes only 0.36 seconds. The advGAN method mentioned in this paper is faster than advGAN because of the addition of condition information c, which proves that the attack method in this paper can generate countermeasure samples faster after training.

4.4 Attacks on real-world sample cases

The above content illustrates the advanced nature of the AdvGAN+ attack algorithm. The algorithm also obtained good results in real-world attacks. As the adversarial training progresses, the generated adversarial samples give wrong judgments in the target classification model. The wrong judgment result is shown in Figure 8. From a visual point of view, the three real traffic signs are turn left, turn left and speed limit, but after adding small interference, the label information of the image itself has changed, YOLOv3 target detection The result of the algorithm is turn left→turn right, turn left→pass, speed limit→turn left, giving a wrong judgment.

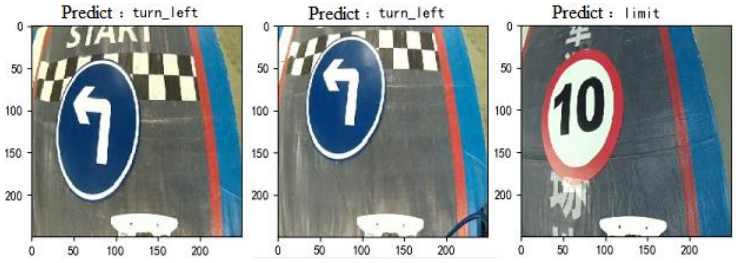

Fig. 8. Examples of real-world attacks

After iterative training, the recognition rate of the classification model with higher confidence is gradually decreasing, and the training results are shown in Table 5.

Table. 5. Real-world attack analysis.

Epochs	100	200	300
Adversarial Samples			
Predicted (turn_right)	71. 73%	12. 79%	3. 269%

The comparison of the generated adversarial samples is shown in Figure 9. The left and right are clean samples and adversarial samples generated by this method respectively. Visually, it seems that there is not much difference between the image and the adversarial sample.

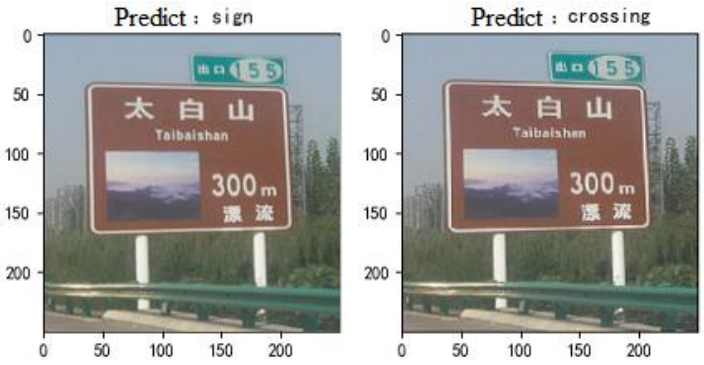

Fig. 9. Comparison of original images and adversarial examples

5. Summary

The safety and robustness of neural network models have received more and more attention, and model misjudgment caused by adversarial samples is an effective detection method. This paper has proposed the AdvGAN+ model, which uses a generative adversarial network to generate adversarial samples, getting rid of the dependence on gradient parameters. The adversarial samples generated by the small disturbance generated by the generator

improve the attack ability compared with the classic gradient-based methods. The attack and defense of adversarial samples is similar to the relationship between a spear and a shield. Only when a more aggressive attack method is developed can the defense method be developed.

The migration of model-generated adversarial samples to other deep learning models is still a research difficulty. The existing adversarial attack methods can only target the target classification model. At the same time, for the adversarial samples generated by GAN, how to improve the robustness of the model during adversarial training is also worthwhile. The direction of continuing research.

Acknowledgements

This research was funded by National Defense Basic Research Plan (grant number. JCKYS2020DC202), Natural Science Found ation of Hebei Province (grant number. F2022208002), Science a nd Technology Project of Hebei Education Department (Key pro gram) (grant number. ZD2021048).

References

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. J. Nature. 521(7553) : 436-438.

Li, H. L., Barnaghi, P., Enshaeifar, S., Ganz, F., 2020. Continual Learning Using Bayesian Neural Networks. J. IEEE transactions on neural networks and learning systems.

Hashimoto, T., Saito, D., Minematsu, N., 2019. Many-to-Many and Completely Parallel-Data-Free Voice Conversion Based on Eigenspace DNN. J. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 27(2).

Chen, H., Zhang, Y., Cao, Y., Xie, J. , 2021. Security issues and defensive approaches in deep learning frameworks. J. Tsinghua Science and Technology, 26(06): 894-904.

Won, J., Seo, S. H., Bertino, E., 2019. A secure shuffling mechanism for white-box attack-resistant unmanned vehicles. J. IEEE Transactions on Mobile Computing.

Dash,J. K., Mukhopadhyay, S., Gupta, R. D., 2017. Multiple classifier system using classification confidence for texture classification. J. Multimedia Tools and Applications, 76(2).

Fischer, S.,WengerA., 2021. Artificial Intelligence, Forward Looking Governance and the Future of Security. J. Swiss Political Science Review, 27.

Zhang, W., 2021. A robust lateral tracking control strategy for autonomous driving vehicles. J. Mechanical Systems and Signal Processing, 150.

Ravi Prakash M L, Chethan Chandra S Basavaraddi, Ananda Babu J, Sapna S Basavaraddi., 2021. Hybrid Neuro-fuzzy Network Based Face Recognition for Occluded Images[J]. Journal of Research in Science and Engineering, 3(5).

Gao, J., Wang, J., Zhang, L., Yu, Q., Huang, Y., Shen, Y., 2019. Magnetic signature analysis for smart security system based on tmr magnetic sensor array. J. IEEE Sensors Journal, PP(8):1-1.

You, W. L., Choi, J. W., & Shin, E. H., 2021. Machine learning model for diagnostic method prediction in parasitic disease using clinical information. J. Expert Systems With Applications, 185.

Wang, L., Chen, W., Yang, W., Bi, F., Yu, F. R., 2020. A state-of-the-art review on image synthesis with generative adversarial networks. J. IEEE Access, PP(99):1-1.

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., et al., 2013. Intriguing properties of neural networks. J. Computer Vision and Pattern Recognition.

Goodfellow, I. J., Shlens, J., Szegedy, C.,2015. Explaining and harnessing advers aria examples. J. ICLR. 1412. 6572.

Moosavi-Dezfooli, S. M., Fawzi, A., Frossard, P., 2016. DeepFool: a simple and accurate method to fool deep neural networks. C. Computer Vision & Pattern Recognition. IEEE, 2574-2582.

Fayed, H. A., Atiya, A. F., 2021. Decision boundary clustering for efficient local svm. J. Applied Soft Computing Journal,110.

Raaijmakers, S., 2019. Artificial Intelligence for Law Enforcement: Challenges and Opportunities. J. IEEE security & privacy, 17(5):74-77.

Shi, Y., Han, Y., Zhang, Q., Kuang, X., 2020. Adaptive iterative attack towards explainable adversarial robustness. J. Pattern Recognition, 105(3):107309.

Biggio, B., Akhtar, Z., Fumera, & G., et al. 2012. Security evaluation of biometric authentication systems under real spoofing attacks. J. IET biometrics, p. 11-14.

Xu, Y., Zhu, L., Yang, Y., Wu, F., 2021. Training robust object detectors from noisy category labels and imprecise bounding boxes. J. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society,37(3).

Lee, J. H., Oh, S. Y., 2016. Feature selection based on geometric distance for high-dimensional data. J. IET Electronics Letters, 52(6):473-475.

Rasche, C., 2017. Rapid contour detection for image classification. J. IET Image Processing, 12(4):532-538.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al., 2014. Generative Adversarial Nets. C. InternationalConference on Neural Information Processing Systems. Cambridge: MIT Press.

Kishore, K. R., Rao, K. S., 2022. A novel approach to unsupervised pattern discovery in speech using convolutional neural network. J. Computer Speech & Language,71.

Wei, H., Chen, H., Li, Y., 2015. Research and implementation of K-means algorithm based on Hama. J. Journal of Physics: Conference Series, 1883(1).

Wei, W., Yi, H , Hu, H., Ruan, C., Chen, D., 2019. Regional attention generative adversarial networks. Electronics Letters, 55(8).

Deng, J., Luo, G., Zhao, C., 2020. Uct-gan: underwater image colour transfer generative adversarial network. IET Image Processing, 14(12).

Xiao, C., Li, B., Zhu, J. Y., He, W., Liu, M., Song, D., 2018. Generating adversarial examples with adversarial networks.

Bai, K., Wu, C., Zhang, Z., 2018. Protect white-box aes to resist table composition attacks. IET Information Security, 12(4), 305-313.

Lee, C. C., Rahiman, M., Rahim, R. A., Saad, F., 2021. A deep feedforward neural network model for image prediction. Journal of Physics: Conference Series, 1878(1), 012062 (5pp).

Yongqiang Zhang born in 1981, Associate Professor and a master supervisor. He graduated from Anhui University of Technology with a master's degree in computer application technology and studied at Army Engineering University with a doctor's degree. Main research areas: AI, IoT, electromagnetism, complex network.

Zewei Ji born in 1998, studied in Hebei University of Science and Technology with a major in computer technology. The main research directions are: computer vision.

Xiaohan Sun born in 1998, studied in Hebei University of Science and Technology with a major in computer technology. The main research directions are: computer vision.

Jinlong Ma born in 1981, Associate Professor, received the Ph.D. degree in information and communication engineering from the Harbin Institute of Technology. His research interests include information spreading dynamics in complex networks and data analysis of online social networks.

Weidong Wu born in 1973, Lecturer, and a master. Main research areas: embedded development, IoT.