2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC) | 978-1-6654-2463-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/COMPSAC51774.2021.00037

CGAN-IRB: A Novel Data Augmentation Method for Apple Leaf Diseases

Xinbin Yuan
College of Information Engineering
Northwest A&F University
Yangling, China
yuanxinbin@nwsuaf.edu.cn

Cong Yu
College of Information Engineering
Northwest A&F University
Yangling, China
yc@nwsuaf.edu.cn

Bin Liu*
College of Information Engineering
Northwest A&F University
Yangling, China
liubin0929@nwsuaf.edu.cn

Henan Sun
College of Information Engineering
Northwest A&F University
Yangling, China
magneto0617@foxmail.com

Xianyu Zhu
College of Information Engineering
Northwest A&F University
Yangling, China
zxy@nwsuaf.edu.cn

***Abstract*—At present, the identification of apple leaf diseases plays an important role in controlling apple leaf diseases and improving apple yield. CNNs(Convolutional Neural Networks) have been widely used in apple leaf diseases identification, but the training of the CNNs requires a large number of images. The lack of images would make the CNNs hard to generalize. Thus the CNNs are unable to recognize new disease images. Focusing on this problem, this paper proposes a new model named CGAN-IRB(Conditional Generative Adversarial Network with the Improved Residual Block) for data augmentation. Firstly, various improvements have been made based on CGAN to generate high-quality, robust, and specific-category images of apple leaf diseases. Among which the embedding of the residual block has been found to significantly improve the model performance. Then the interpolation algorithm is used instead of deconvolution to increase the image size. Finally, the TTUR(Two-Timescale Update Rule) training strategy is employed and all the convolutional layers of the network are spectrally normalized to stabilize the training of the network. The performance of CGAN-IRB was tested both on image generation and classification tasks. Experiment results show that the images generated by the network possess high quality and robust features, providing a novel solution for the data augmentation of apple leaf diseases. The new GAN-based data augmentation method leads to significant improvements in the classification accuracy of CNNs. In the case of all tested CNNs, the classification accuracy improvements are 11.75% and 2.17% on average over non-augmented and traditional-augmented, respectively. Among them, the classification accuracy of GoogLeNet V2 and ShuffleNet V2 is 99.34% and 99.67%, respectively. The data augmentation approach proposed in this paper can be used more widely in the field of disease identification, solving the problem of insufficient data sets, and can be extended to related fields where data sets are difficult to obtain.**

***Index Terms*—data augmentation, generative adversarial networks, apple leaf disease identification, convolutional neural networks**

*Corresponding author:Biu Liu
E-mail:liubin0929@nwsuaf.edu.cn

I. Introduction

Apple leaf diseases can greatly affect the growth and yield of apples. Therefore, a method is urgently needed to identify apple leaf diseases quickly and timely. In recent years, CNN-based image classification models have shown good results in image recognition and are widely used in crop disease recognition [1–11]. Rather than manually selecting features to feed traditional machine learning classification methods, CNNs provide end-to-end pipelines to automatically extract advanced robust features and thus significantly improve the usability of plant leaf identification. Although the CNN structure is constantly improving [12–24] and bringing higher and better results, the network training still relies on a large amount of data, and the generalization ability of the network is very poor when the amount of data is small, which would further affect the recognition accuracy of the network. However, the acquisition of apple leaf disease images is still a difficult problem. It mostly requires manual image collection and expert annotation, which is time-consuming and laborious.

Data augmentation methods are widely used to solve the problem of insufficient data set, such as image flipping, cropping, rotation, scaling, and deformation based on spatial transformation, adding random noise, and multi-sample synthesis techniques, such as SMOTE, sample pairing, and mix up. The images generated by these methods have the same distribution as the original, so cannot provide more information for the training of the classification model. Thus the improvement of the model is limited, so there is a need for generating different distribution to increase sample diversity. GAN [25] learns from the feature distribution of the original images, and the generated has a distribution similar to it, which can provide more feature information to the classification model. Although GAN has many advantages, the training process converges slowly, and it is easy to collapse. Therefore, many improvements have been made based on the original GAN.

978-1-6654-2463-9/21/$31.00 ©2021 IEEE
DOI 10.1109/COMPSAC51774.2021.00037

CGAN allows the network to generate images of a specific category by adding conditional information, sequentially the disadvantages of too free image generation and too long convergence time are solved. Thus this paper proposes a model based on CGAN, then makes further improvements, and finally generates high-quality images. Therefore, this paper providing a new solution for data augmentation.

The main contributions of this paper are summarized as follows:

- **CGAN-IRB-based data augmentation.** The model proposed in this paper obtains a lower FID(Fréchet Inception Distance) [26] score than most models. The generated images have high quality and robust features, which can provide much feature information for the CNN-based classification models, thereby improving the performance of the network.
- **Improving the performance of CNN-based model.** Original data set has a total of 1,193 images of apple leaf diseases. Adding the images generated by CGAN-IRB to the original data set, the scale of the data has been expanded by nearly 13 times. Training on the augmented data set greatly enhances the generalization ability of the trained model and therefore improves the recognition accuracy of the model. Compared with the traditional-augmented and non-augmented, the classification accuracy of CNN-based models is greatly improved.

The rest of this paper is arranged as follows. Section II summarizes related work. Section III describes the research of the CGAN-IRB model and the data augmentation method based on it. Section IV is the exploration of the GAN model and the comparison experiment of the accuracy of the CNN-based classification model. Section V summarizes and looks forward to the full text.

II. RELATED WORK

In [27], Zhao et al. proposed the SaGAAN model to synthesize high-quality hyperspectral samples to assist the classification of high-dimensional hyperspectral data. The self-attention mechanism was applied to capture the long-range dependencies across the spectral bands and further improved the quality of generated samples. In [28], Abdelhalim et al. proposed SPGGAN-TTUR, which has the attention mechanism based on PGGAN, and TTUR strategy was applied to stabilize the training process of generating skin disease images. The model used a procedural training method to generate high-resolution images. These methods all generate high-resolution images and use the attention mechanism to improve the spatial structure of the image, leading to the factor that the network takes up a lot of memory, and the training speed is slow. So it is not suitable in resource-constrained devices.

In [29], Zhu et al. utilized improved CDCGAN for vigor rating of orchid seedlings, which was a significant but labor-intensive task in a modern commercial greenhouse. It performed a bypass connection to CDCGAN, which was similar to the dense block. The dense block was suitable for feature extraction, but when used for image generation, the direct connection of the shallow layer to the latter layer would make the image quality bad. Therefore, the generated images were blurry. In [30], Espejo-Garcia et al. combined advanced transfer learning with data augmentation techniques through GANs for early weeds identification, after that all the tested networks obtained promising results. However, the tomato images generated by their GANs got best FID score at about 90, which is obviously high. These methods have made improvements to the GAN network structure, and do not occupy too much memory. They have achieved good results when used for data augmentation, but the image quality is not high. Therefore, a new method is urgently needed which uses less memory and the resulting image quality is high.

It can be seen that GAN has been widely used in the field of data augmentation to solve the problem of insufficient data sets and to improve the performance of CNN-based classifiers, thus reducing the investment of human labor and material resources. However, the methods mentioned above have their shortcomings in speed, size, and image quality. Thus cannot be used as data augmentation methods generally. Different from the above improvements, this paper takes the residual block as the springboard and proposes a data augmentation method based on CGAN-IRB, which occupies less memory and the FID score of the generated samples is lower.

III. MODEL STRUCTURE

The CGAN-IRB model is proposed to generate 128×128 images of five categories of apple leaf diseases. In the first part, the structure of the generator and discriminator are shown in detail, and the second part explains the training strategy to improve the model performance.

A. CGAN-IRB Model

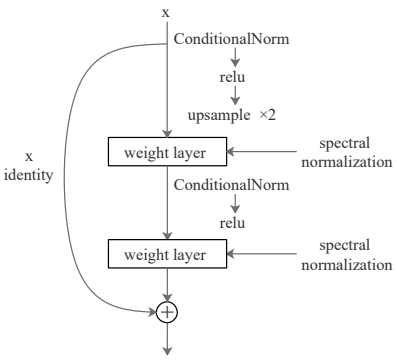

Fig. 1. The basic module of generator.

1) Generator: First of all, the generator needs to establish a mapping from noise to the image. As the network structure of the generator deepens, the non-linear layer increases and the feature mapping becomes more complicated. In this case, it is difficult for the generator to improve the identity mapping based on the gradient information provided by the discriminator alone. However, the residual block changes the initial identity mapping of the network to the residual mapping

and uses a short connection to connect the original identity x to the output layer of the stacked nonlinear layers, which can solve the problems of the difficulty of training and degradation problem due to the deepened network. Therefore, the residual block is used as the basic module of the generator. Then, in this experiment, the generator needs to generate five different categories of apple leaf disease images, so the label information of different categories of images is added as conditional information. Further, to prevent conditional information from fading and disappearing during the propagation process, the strategy of adding conditional information to each basic module is applied in the generator. In addition, in the previous GAN, the size of the image is increased through the deconvolution layer. Using the deconvolution layer in the residual block, when the dimensionality increases, the original identity needs to be deconvolved before it can pass through the short connection. However, the weight of the deconvolution layer is continuously improved during the learning process, which will undoubtedly make the residual learning useless. Therefore, the original identity needs to be smoother and more stable when the size increases. The interpolation algorithm predicts the surrounding point pixels through the neighboring point pixels to upsample the image. This method is fixed, it will not change during the learning process of the network and is more stable. Therefore, in the basic module, the interpolation algorithm is used to replace the deconvolution to increase the image size. Finally, to stabilize the training process of the generator, spectral normalization is performed on each convolutional layer. The overview diagram of the generator is shown in Fig.2.

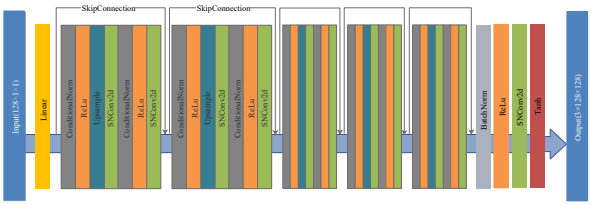

Fig. 2. The overview of generator model.

The input of the generator is a 128-dimensional noise vector, which first passes through a linear layer, and then a stack of five layers of basic modules, each layer doubles the size of the image. Finally, an image with a size of 128×128 is output through a tanh activation layer. The specific parameters of the model are shown in TABLE I.

TABLE I
THE SPECIFIC PARAMETERS OF GENERATOR MODEL.

Name	Type	Input size	Output size
Linear	Linear	128×1×1	512×4×4
Basicblock_G1	Basicblock_G	512×4×4	512×8×8
Basicblock_G2	Basicblock_G	512×8×8	512×16×16
Basicblock_G3	Basicblock_G	256×16×16	256×32×32
Basicblock_G4	Basicblock_G	256×32×32	128×64×64
Basicblock_G5	Basicblock_G	128×64×64	64×128×128
BatchNorm	normalization	64×128×128	64×128×128
ReLu	activation	64×128×128	64×128×128
SNConv2d	convolution	64×128×128	3×128×128
Tanh	activation	3×128×128	3×128×128

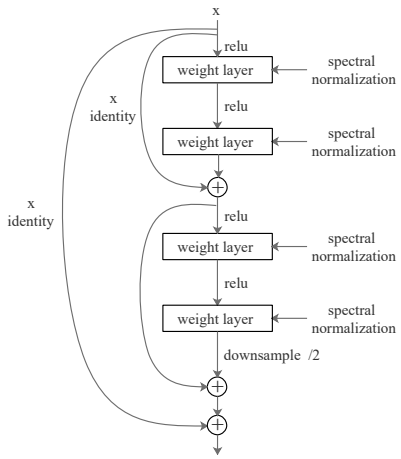

Fig. 3. The basic module of discriminator.

2) Discriminator: The Generator does not directly extract features from the original image data set during image generation, but improves the quality of the generated image based on the gradient information provided by the discriminator. During backpropagation, gradient explosion and gradient disappearance would occur. In both cases, the gradient information would be lost, making it difficult to improve the generator, and even the quality of the generated image would deteriorate. Therefore, in order to better retain the gradient, as shown in Fig.4, short connections are used to improve the original residual block. Then it is used as the basic module of the discriminator. The overview diagram of the discriminator model is shown in Fig.4.

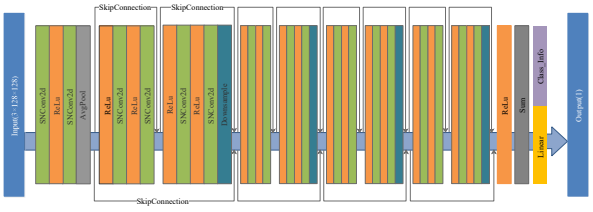

Fig. 4. The overview of discriminator model.

The input of the discriminator is an image with a size of 3×128×128, which is first subjected to dimensionality reduction processing, followed by a stack of four layers of basic modules. After each layer, the image size will be halved. Finally, when passing through the linear layer, condition information is added to output a one-dimensional vector, which represents the gap between the input image category and the real category. The specific information of the discriminator is shown in TABLE II.

TABLE II
THE SPECIFIC PARAMETERS OF DISCRIMINATOR MODEL.

Name	Type	Input size	Output size
SNConv2d	convolution	3×128×128	64×128×128
ReLu	activation	64×128×128	64×128×128
SNConv2d	convolution	64×128×128	64×128×128
AvgPool	pool	64×128×128	64×64×64
Basicblock_D1	Basicblock_D	64×64×64	128×32×32
Basicblock_D2	Basicblock_D	128×32×32	256×16×16
Basicblock_D3	Basicblock_D	256×16×16	512×8×8
Basicblock_D4	Basicblock_D	512×8×8	512×4×4
ReLu	activation	512×4×4	512×4×4
View+Sum	linear	512×4×4	512×1
Linear+Class Info	linear	512×1	1×1

B. Strategies to Improve Network Performance

1) Conditional Normalization: In addition to the differences in the features of the lesions, the general features of the images of different categories of apple leaf diseases are the same. Therefore, training on the five categories of disease image data sets at the same time would make more image features be extracted by the discriminator. But directly training on different categories of data sets would make the lesion features of the generated images bad, so the image category information is one-hot encoded as conditional information and is added to noise vector as the input of the generator, and then conditional information is added to each following basic module, which can make the image features of the generated images more obvious, and the generated images of different categories are more distinguishable. Finally, the discriminator judges the generated images through the category information, and the generator can finally generate images of different categories at the same time.

2) Spectral Normalization: GANs have always been difficult to train and the training process is unstable. The direct reason is that the gradient changes drastically during backpropagation. In [31], Arjovsky et al. proposed a gradient clipping strategy, which artificially restricted the gradient to a certain range. But with that processing, it lost the integrity of the information. In [32], Gulrajani et al. used gradient penalty instead of gradient clipping, which largely solved the problems in WGAN. The gradient penalty made the training of the network more stable by limiting the gradient of the backpropagation, but this also made it largely depended on the support of the generation distribution. In the training process, the generative distribution and its support gradually changed, thus destabilized the effect of such regularization. Moreover, WGAN-GP would consume a lot of computing resources. In contrast, spectral normalization could better solve this problem.

Spectral normalization imposes Lipschitz constraints on the weight of the convolution kernel:

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} \leq M \tag{1}$$

M is Lipschitz constant, x_1, x_2 is the input of the network, and $f(x_1)$, $f(x_2)$ is the output of the network. By limiting M to be less than or equal to one, the parameters of the network will not change too drastically during the training process. Therefore, the gradient during backpropagation is more stable. Spectral normalization achieves the effect of Lipschitz constant constraint by dividing the parameter matrix of the convolutional layer of the network by the spectral norm of its parameter matrix, thus improving the stability of the training process of GAN.

IV. EXPERIMENT EVALUATION

A. Experiment Environment

The experiment was carried out in the Ubuntu 16.04.2 environment, using the PyTorch deep learning framework. The remaining parameters are shown in TABLE III.

B. Data Acquisition

Original data set was obtained from orchards and laboratory. It contains 1,193 images of apple leaf diseases in five categories. The number of images of each disease spot varies. See TABLE IV for details.

TABLE IV
APPLE LEAF DISEASE DATA SETS.

Class	Quantity
Alternaria blotch	422
Brown spot	176
Gray spot	135
Mosaic	435
Rust	215
Total	1193

As shown in Fig.5, the spots of Alternaria blotch are small round spots, brown, and dark brown around the periphery at the beginning. Later, the diseased spots gradually became larger, and most of the leaves and the leaf edge tip became brown in the later stage of the disease. The lesions of the brown spot are characterized by brown in the middle, green at the edges, yellow at the periphery, and many small black spots on the surface. The features of gray spot disease are reddish-brown round or nearly round spots at the beginning and later become gray with small black spots scattered in the center. Mosaic disease spots are irregular, with large dark green and light green color changes, and the edges are not clear. Oily orange-red dots appeared on the leaf surface at the initial stage of rust disease, and then gradually expanded to form orange lesions.

TABLE III
THE HIGH PERFORMANCE UBUNTU 16.04.2 SERVER.

Configuration item	Value
Central processing unit	Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz (X2)
Graphics processor unit	GP100GL [Tesla P100 PCIe 16GB]
Hard Disk	2 TB
Operation system	Ubuntu 16.04.2 LTS (64-bit)
Python	3.7.2
Pytorch	1.6.0

(a) Alternaria blotch (b) Brown spot (c) Gray spot

(d) Mosaic (e) Rust

Fig. 5. Apple leaf disease data sets.

C. Experiment Results and Analysis

1) Evaluation Metrics: In this paper, FID is selected as the evaluation index of GAN performance, and its calculation formula is as follows:

$$FID\,(x,g) = \|\mu_x - \mu_g\|_2^2 + \mathrm{Tr}\left(\Sigma_x + \Sigma_g - 2\,(\Sigma_x\Sigma_g)^{\frac{1}{2}}\right) \quad (2)$$

where μx, Σx, and μg, Σg are the mean and covariance of the sample embeddings from the data distribution and sample distribution, respectively. The calculation of the FID score is to directly extract the features of a batch of images through the inception v3 network, and then calculate the distance between the generated samples and the real data in the feature space, so this index roughly evaluates the model performance from two aspects, which are the diversity of the samples and the robustness of features. That is, if the images generated by the model on each category is the same, it will get a higher FID score. If the quality of the image generated by the model is not high, a higher FID score would also be obtained. Therefore, the higher the quality of the generated image and the richer the diversity, the lower the FID score.

The experiment makes various improvements based on CGAN. The CGAN with full convolutional network structure(CDCGAN) is tested, then the attention mechanism was added to the CDCGAN(CDCGAN-SA). Further, the CGAN with the residual block in both the discriminator and generator(CGAN-RB) is tested, then the CGAN with the improved residual block(CGAN-IRB) is tested, and finally, the performance of the network after replacing the interpolation algorithm with deconvolution(CGAN-IRB-Deconv) is verified.

It can be seen from TABLE V that the FID of the Brown spot samples generated by the networks is generally low. This is because the features of the brown spots are simple and easy to extract. The FID score of mosaic samples is generally high because the mosaic disease spots have complex features. The discriminator of CDCGAN is stacked with ordinary convolutional layers, thus the ability to extract features is poor. After adding the attention mechanism, the improvement of the network is small, and it is even lower than the original in the quality of generated Gray spot disease images. This is because after adding the attention mechanism, it will consume a lot of computing resources and storage space, the memory occupancy rate has almost doubled, and the model convergence speed has also slowed down. Besides, the number of Gray spot disease images in the original data set is small, under almost the same training time, the performance of CDCGAN-SA is relatively poor, so the Gray spot disease images generated by CDCGAN-SA have no obvious features and poor quality. The improvement of the network is small, mainly because the improvement brought by the attention mechanism is mostly in the spatial structure of the generated image, but the apple leaf disease images used in the experiment do not have particularly strict spatial structure characteristics, so the improvement brought by attention mechanism is small. To achieve the goal of fast training speed and small memory usage, the residual block is used instead. To avoid too deep the network and increase the convergence speed of the network, 5 residual block layers are employed in the generator and 8 residual block layers in the discriminator. The results show that the effect of the residual block is far superior to the attention mechanism, and the FID score of Brown spot disease is almost reduced to half of the CDCGAN, which shows the strong feature extraction ability of the residual block. Therefore, further improvements are made based on the residual block. It can be seen that the FID score of the generated samples of CGAN-IRB is greatly reduced, and the average is reduced to about 50, which is 28.7% lower than the original residual block. Besides, after the interpolation algorithm is replaced with deconvolution in CGAN-IRB, the network performance is greatly weakened, and the FID score is even lower than CDCGAN.

2) The Over View of Generated Images: As an evaluation index, FID has unavoidable shortcomings. The calculation of

TABLE V
FID OF GAN MODELS.

data set\network	CDCGAN	CDCGAN-SA	CGAN-RB18	CGAN-IRB-Denconv	CGAN-IRB
Alternaria blotch	113.07	112.14	80.39	215.01	71.26
Brown spot	98.12	99.37	50.74	174.67	47.31
Gray spot	132.09	189.21	74.51	159.48	42.29
Mosaic	148.33	131.65	97.37	216.57	68.44
Rust	110.42	94.85	78.69	219.22	42.87
Avg	120.41	125.44	76.34	196.99	54.43

the distance between the original image and the generated image in the feature space cannot reflect whether the spatial structure of the image is satisfactory. Therefore, samples generated by different networks are shown to bring more intuitive feelings.

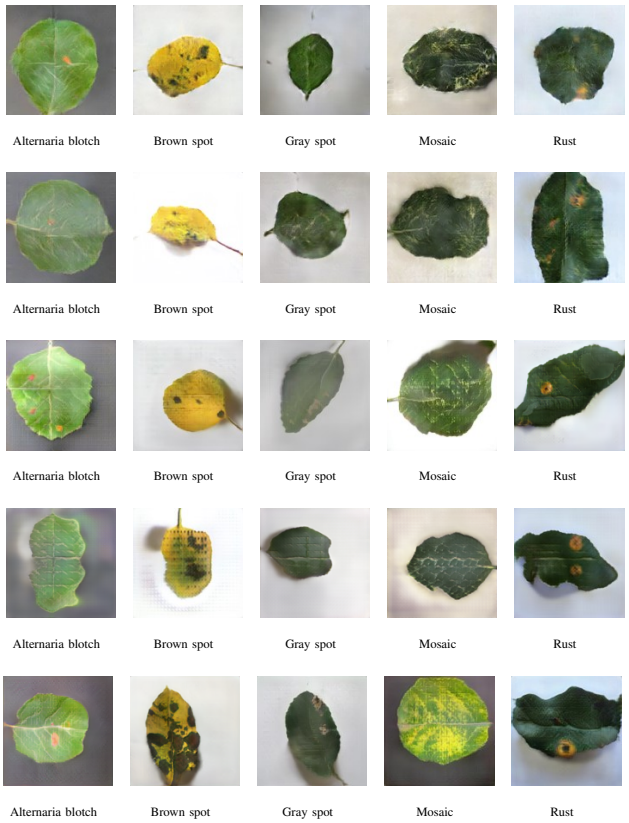

Fig. 6. Apple leaf disease images generated by different GAN models. (a)-(e)CDCGAN; (f)-(j)CDCGAN-SA; (k)-(o)CGAN-RB; (p)-(t)CGAN-IRB-Deconv; (u)-(y)CGAN-IRB.

It can be seen from Fig.6 that under the guidance of conditional information, all samples of different lesions generated by the networks are easier to distinguish. However, in terms of single lesion features, only samples of all categories generated by CGAN-IRB are of high quality and have clear lesion features, while samples of all categories generated by CGAN-IRB-Deconv have high-frequency artifacts. In addition, the features of the Mosaic disease are the most complicated. Except for CGAN-IRB, the Mosaic disease samples generated by other models are blurry. Among them, CGAN-RB has a little light green color change, while CDCGAN and CDCGAN-SA have almost no Discoloration, CGAN-IRB-Deconv has obvious high-frequency artifacts. In contrast, the lesions generated by CGAN-IRB have clear and robust features, with obvious Mosaic features which are large areas of dark green and light green discoloration. The experimental results show that the image features generated by CGAN-IRB are robust and of high quality, and therefore reflect the strong ability of the improved residual block in the discriminator to extract features and the effectiveness of using the residual block in the generator. In CGAN-IRB-Deconv, the appearance of high-frequency artifacts in the samples confirms the previous conjecture that deconvolution is not applicable to the residual block.

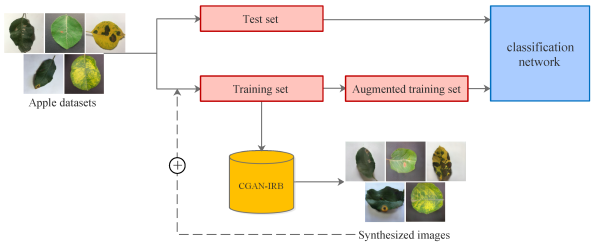

Fig. 7. The flow chart of data augmentation.

D. The Classification Result

The original data set is divided into two parts, one part is used as a basic data set for data augmentation, and the other part is used as a test set to test the trained network. GAN-based data augmentation method and traditional data augmentation method are used to expand the basic data set at a ratio of about 1:14 to obtain a new data set. To confirm that the residual block has indeed improved the classification accuracy of the recognition network, the Apple-CGAN-RB data set is also made for testing. The details of the data set are summarised in TABLE VII.

The new data set is used to train the CNN-based classifier, and the classification accuracy is tested on the original test set. The flow chart of the whole process is shown in Fig.7.

The classic convolutional neural network VGG16 and GoogLeNet V2 is trained on the augmented data set, the ResNet34, ResNet50, ResNet101, and the mobile-end

TABLE VI
THE DETAILS OF THE APPLE LEAF DISEASE DATA SETS.

Class	Original	Traditional	CGAN-RB	CGAN-IRB
Alternaria blotch	422	5895	5895	5895
Brown spot	176	2451	2451	2451
Gray spot	135	1890	1890	1890
Mosaic	435	6089	6089	6089
Rust	215	2998	2998	2998
Total	1193	19323	19323	19323

lightweight network MobileNet V2, ShuffleNet V2 are also trained. During the training process, GoogLeNet V2, MobileNet V2, and ShuffleNet V2 converge faster and show higher classification accuracy in the test set, but ResNet101, ResNet50, ResNet34, and VGG16 still have small fluctuations in the training process, so the central value of the fluctuation range is taken as the final classification accuracy.

It can be seen from TABLE VII and Fig.8 that the accuracy of CNN-based classifiers trained on the original data set is almost less than 90%. After data augmentation, the classification accuracy is generally improved. Compared with the non-augmented, the method of CGAN-RB based data augmentation increases the accuracy of the network by an average of 10.29%. After CGAN-IRB is applied for data augmentation, the accuracy increases by 11.75% on average. Compared with the traditional-augmented, the accuracy of the network has increased by 0.8% and 2.17% respectively. Besides, the recognition accuracy of GoogLeNet V2 and ShuffleNet V2 are as high as 99.34% and 99.67%, respectively, which greatly confirms that the images generated by CGAN-IRB have robust features. This makes it possible to further improve the accuracy of lesion identification while greatly reduces the consumption of material and financial resources.

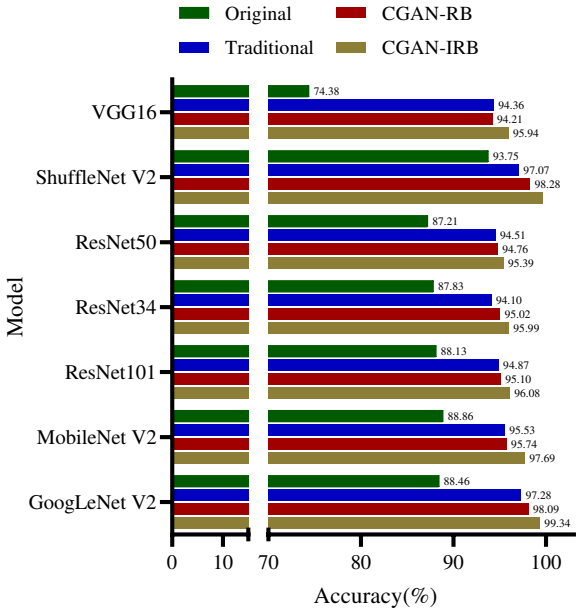

Fig. 8. The bar chart of classification accuracy of CNNs.

E. Feature Visualization

Visualization technology is a useful method to explore how CNNs learn features, which can also help understand the features of the images. Therefore, feature visualization is used to verify that the images generated by CGAN-IRB have robust features.

It can be seen from Fig.9 that the color of the heat map is darker where there is a diseased spot. In the grayscale image, the pixel value of the place where the diseased spot is located is larger, which is clearly highlighted. This confirms that the images generated by the model have strong features.

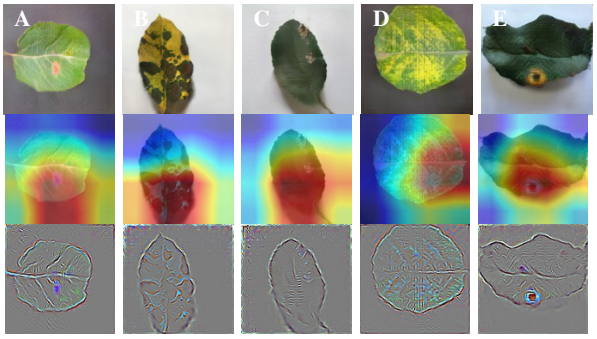

Fig. 9. Grad-CAM visualization. (A)Alternaria blotch; (B)Brown spot; (C)Gray spot; (D)Mosaic; (E)Rust.

V. CONCLUSIONS

This paper proposes CGAN-IRB for data augmentation, which solves the problem of poor generalization ability of the recognition network due to insufficient data sets. The model is based on CGAN, and by introducing an improved residual block, the generated images are high-quality and have robust features, which provides a new idea for the improvement of the GAN model and a novel solution for data augmentation.

TABLE VII
THE CLASSIFICATION ACCURACY(%) OF CNNS.

data set\model	GoogLeNet V2	MobileNet V2	ResNet101	ResNet34	ResNet50	ShuffleNet V2
Original	88.46 ± 0.23	88.86 ± 0.12	88.13 ± 0.29	87.83 ± 0.32	87.21 ± 0.31	93.75 ± 0.02
Traditional	97.28 ± 0.05	95.53 ± 0.07	94.87 ± 0.10	94.10 ± 0.45	94.51 ± 0.41	97.07 ± 0.03
CGAN-RB	98.09 ± 0.03	95.74 ± 0.14	95.10 ± 0.53	95.02 ± 0.32	94.76 ± 0.17	98.28 ± 0.07
CGAN-IRB	99.34 ± 0.05	97.69 ± 0.12	96.08 ± 0.23	95.99 ± 0.34	95.39 ± 0.43	99.67 ± 0.02

The accuracy of the recognition network trained on non-augmented data set is around 90%, which can hardly be used for the classification of apple leaf diseases. After adding the images generated by CGAN-IRB, the size of the original data set is expanded 14 times, the classification accuracy of the CNN-based classifiers is 11.75% higher than non-augmented and 2.17% higher than the traditional-augmented, and the accuracy of GoogLeNet V2 and ShuffleNet V2 is 99.34% and 99.67%, respectively. Therefore, the data augmentation method proposed in this paper can also be used for data set expansion in other areas where data sets are insufficient.

VI. ACKNOWLEDGEMENT

We are thankful for the valuable opinions of reviewers,which helps us a lot to improve the quality of this draft.This work is supported by the Key Research and Development Program of Shaanxi under grants No.2021NY-138 and No.2019ZDLNY07-06-01, by the Fundamental Research Funds for the Central Universities under grant No. 2452019064 and by China's National College Innovation and Entrepreneurship Training Program of Northwest A&F University (No.S202010712305).

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[3] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[5] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[6] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[7] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2017, pp. 1–12.

[8] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[9] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492–1500.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[11] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6848–6856.

[12] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision*, 2018, pp. 116–131.

[13] M. Agarwal, S. K. Gupta, and K. K. Biswas, "Development of efficient cnn model for tomato crop disease identification," *Sustainable Computing: Informatics and Systems*, vol. 28, pp. 1–12, 2020.

[14] G.-L. Tian, C. Liu, Y. Liu, M. Li, J. Zhang, and H. Duan, "Research on plant diseases and insect pests identification based on cnn," *Iop Conference*, vol. 594, pp. 1–6, 2020.

[15] K. Leong and L. Tze, "Plant leaf diseases identification using convolutional neural network with treatment handling system," in *proceedings of 2020 IEEE International Conference on Automatic Control and Intelligent Systems*, 2020, pp. 39–44.

[16] J. Boulent, S. Foucher, J. Théau, and P. L. St-Charles, "Convolutional neural networks for the automatic identification of plant diseases," *Frontiers in Plant Science*, vol. 10, pp. 1–15, 2019.

[17] S. Srdjan, A. Marko, A. Andras, C. Dubravko, and

S. Darko, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence Neuroscience*, vol. 2016, pp. 1–11, 2016.

[18] G. G. and A. Pandian J, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers Electrical Engineering*, vol. 76, pp. 323–338, 2019.

[19] A. Anagnostis, G. Asiminari, E. Papageorgiou, and D. Bochtis, "A convolutional neural networks based method for anthracnose infected walnut tree leaves identification," *Applied Sciences*, vol. 10, pp. 1–27, 2020.

[20] S. Verma, A. Chug, and A. P. Singh, "Application of convolutional neural networks for evaluation of disease severity in tomato plant," *Journal of Discrete Mathematical Sciences Cryptography*, vol. 23, no. 1, pp. 273–282, 2020.

[21] E. A. H. Mora, V. A. G. Huitron, A. E. R. Mata, and H. R. Rangel, "Plant disease detection with convolutional neural networks implemented on raspberry pi 4," in *proceedings of 2020 IEEE International Autumn Meeting on Power, Electronics and Computing*, 2020, pp. 1–23.

[22] S. K. Bharathi, K. K. Savitha, and A. Nandhu, "Classification of rice plant disease using convolution neural network," *Test Engineering Management*, vol. 83, pp. 14 585–14 589, 2020.

[23] S. S. Hari, M. Sivakumar, P. Renuga, S. Karthikeyan, and S. Suriya, "Detection of plant disease by leaf image using convolutional neural network," in *proceedings of 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking*, 2019, pp. 1–5.

[24] P. Bharali, C. Bhuyan, and A. Boruah, "Plant disease detection by leaf image classification using convolutional neural network," pp. 194–205, 11 2019.

[25] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, pp. 2672–2680, 2014.

[26] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are gans created equal? a large-scale study," in *proceedings of Advances in neural information processing systems*, 2018, pp. 700–709.

[27] W. Zhao, X. Chen, J. Chen, and Y. Qu, "Sample generation with self-attention generative adversarial adaptation network (sagaan) for hyperspectral image classification," *Remote Sensing*, vol. 12, no. 5, pp. 1–14, 2020.

[28] I. S. A. Abdelhalim, M. F. Mohamed, and Y. B. Mahdy, "Data augmentation for skin lesion using self-attention based progressive generative adversarial network," *Expert Systems with Applications*, vol. 165, pp. 1–13, 2021.

[29] F. Zhu, M. He, and Z. Zheng, "Data augmentation using improved cdcgan for plant vigor rating," *Computers and Electronics in Agriculture*, vol. 175, pp. 1–10, 2020.

[30] B. Espejo-Garcia, N. Mylonas, L. Athanasakos, E. Vali, and S. Fountas, "Combining generative adversarial networks and agricultural transfer learning for weeds identification," *Biosystems Engineering*, vol. 204, pp. 79–89, 2021.

[31] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *proceedings of International conference on machine learning*, 2017, pp. 214–223.

[32] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *proceedings of Advances in neural information processing systems*, 2017, pp. 5767–5777.