Evolutionary Intelligence (2023) 16:1899–1906
https://doi.org/10.1007/s12065-021-00665-z

SPECIAL ISSUE

Modified generative adversarial networks for image classification

Zhongtang Zhao[1,2] · **Ruixian Li**[3]

Received: 18 December 2020 / Revised: 15 July 2021 / Accepted: 24 August 2021 / Published online: 29 August 2021

Abstract
In the image classification task, the existing neural network models have insufficient ability to characterize the features of the classified objects, which leads to the problem of low recognition accuracy. Therefore, we propose a modified Generative Adversarial Networks (GAN) for image classification. Based on the traditional generative adversarial network, By constructing multiple generation models and introducing collaboration mechanism, the generation models can learn from each other and make progress together in the training process to improve the fitting ability of the model for real data and further improve the classification quality. Finally, a generative adversarial network is designed to generate the occlusion samples, so that the model has good robustness for the occlusion objects recognition. The Top-1 error rate is used as the evaluation index. The experiments are conducted on the public data sets containing Cifar10, Cifar100, ImageNet2012. The comparison experiment results show that the proposed method can improve the feature representation ability of the GAN and improve the accuracy of image classification. The average accuracy is higher than 90% and the error rate is lower than 1.0%.

Keywords Image classification · Generative adversarial networks · Discriminant network · Cooperation learning

1 Introduction

Image classification is an important research direction in the field of computer vision [1–3]. Its main task is to automatically identify the target in the image by computer and assign it to the corresponding category set. Traditional image classification algorithms mainly include three steps: preprocessing, feature extraction and classification. Feature extraction is the most important step in image classification, and the quality of features extracted in this step directly affects the performance of image classification. The features extracted by traditional image classification algorithms are relatively redundant and do not have good generalization, so the researchers have proposed many excellent neural network models [4–6].

At present, due to its unique advantages in image processing, convolutional neural network (CNN) is widely used in image classification tasks. However, due to the multilayer convolution and pooling operation, a large amount of important information is lost, which leads to the insufficient expression ability of the features extracted by the convolutional neural network. To solve these problems, researchers propose a series of feature enhancement modules to improve the robustness of the features. However, most of the feature enhancement modules not only increase the computational load of the original neural network model, but also fail to improve the accuracy.

From the perspective of feature extraction and selection, the existing image classification methods are usually based on two types of feature learning frameworks: the Bag of Words (BOW) model and the deep learning model.

The feature representation of the traditional BOW model [7] is to generate a global representation for the image, and the generation process mainly consists of three steps: manual feature extraction, feature encoding and feature aggregation. First, features are extracted manually on dense grids or sparse points of interest. Then the features are quantified through different coding schemes. Finally, these coded

✉ Ruixian Li
xdwangxd@163.com

Zhongtang Zhao
zhaozt@sina.com

1 School of Intelligent Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450000, China

2 Intelligent Robot and System Advanced Innovation Center, Beijing Institute of Technology, Beijing 100081, China

3 School of Information Systems Engineering, Information Engineering University, Zhengzhou 450000, China

features are combined by feature aggregation to form image-level representation. The common models for BOW feature classification include Support Vector Machine (SVM) and random forest [8, 9]. In addition, the BOW feature learning method combined with the probabilistic theme model also achieves good results. Jin [10] represented the image features as dictionaries through the BOW model, and used the Probabilistic Latent Semantic analysis (PLSA) model to find out the potential themes from a large number of images to classify the images. Ghorai [11] assumed that the potential topic space could be learned from the two modes of vision and text. PLSA was used to learn the data in different models to obtain the corresponding semantic topic distribution, and then the two models were fused through adaptive asymmetric algorithm to obtain better classification effect. Then, Filisbino [12] proposed a learning method combining mixed generation and discriminant model. It used continuous PLSA to model the visual features of images to reduce the impact of cluster granularity on classification performance, and adopted an integrated classifier chain to classify multi-labeled images. However, in the specific task of image classification, manual features are not always the best result.

In the computer vision field, Teng [13] had proved that CNN performed better than many other methods based on manual features. Li [14] proposed that the image representation learned by CNN on large-scale data sets could be effectively transferred to other visual recognition tasks based on limited training data. Yan [15] proposed an integration framework for classification, named Over-Feat (OF), which achieved better precision than the traditional BOW model. Chaib [16] also used a pre-trained CNN model to extract image features, but used an integrated classifier chain instead of SVM to classify depth features, so that they could learn the association between marked data sets to obtain better performance. Zhu et al. [17] proposed spatial regularization network (SRN), which was used to generate the whole labeled attention graph and capture the underlying relationship between them. Then, a ResNet-101 network was adopted to aggregate the regularized classification result and the original result. Mou [18] developed a circular memory attention module to realize the explicable image classification including two alternate components. You [19] proposed a multi-marked classification model based on the graph convolutional network to establish a directed graph on the objects, in which each node was represented by the marked embedded word. The marked graph was mapped into a group of interdependent object classifiers by training the graph convolutional network.

Deep neural network (DNN) uses an architecture composed of several nonlinear transformations to model the high-level abstraction of visual data and shows good effectiveness in image classification tasks. Teng [20] compared several traditional feature learning methods and used CNN to solve the problem of image classification based on several loss functions. The results showed that the classification performance of CNN was significantly higher than that of the traditional method. To get an effective CNN model, the CNN needs to learn a lot of parameters in training. However, it is very difficult to train a CNN for specific tasks on the limited training data sets. Therefore, using parameter transfer learning to optimize the CNN model has become a widely advanced method. Many works had proved that the parameters in the pre-trained CNN model on diversity ImageNet can be transferred to the new model and extract features of other data sets without sufficient training samples. In other words, parameter transfer learning in training phase is conducive to increasing the data set, which can further optimize the model, solve the problem of sample shortage in disguise and obtain more accurate features. Some researchers adopt semi-supervised learning methods for image classification.

The main idea of semi-supervised learning is to use the information provided by marked data and unmarked data, improve the learning performance in the case of little marked data. The marked data provides data and label joint distribution information, while unmarked data only provides data distribution information. In reference [21], the self-training-based semi-supervised learning method was used for target detection, which could obtain the same effect as the traditional training model with a larger labeled data set. In reference [22], a ladder network method constructed by the self-encoding network was proposed. By adding horizontal connection, noise encoder and corresponding denoising decoder were added into the normal feed-forward network to learn the unlabeled training data. Zhao et al., [23] chose FCN, ResNet, and PSPNet as classifiers. The models were trained by different proportions of training samples from Jingjinji region. Then it used the trained models to predict the results of the study areas. Wang et al., [24] proposed a weakly supervised deep learning framework with uncertainty estimation to address the macula-related disease classification problem from OCT images with the only volume-level label being available. At present, the semi-supervised learning methods mainly include generative model, self-training model, semi-supervised SVM, entropy regularization model and graph-based model [23], etc,. The existing researches results show that compared with supervised learning with labeled data, the learning performance of semi-supervised learning is modified.

In this paper, a semi-supervised image classification method based on modified GAN is proposed. By using a little labeled data and unlabeled data, the discriminator network of GAN can output data category labels, and a semi-supervised classification method based on little labeled samples is realized. Comparing to the supervised learning methods, the performance of the proposed GAN is better than that of other semi-supervised networks. This method can be applied to image classification, medical diagnosis, anomaly detection and image recognition, etc,.

This paper is organized as follows. In Sect. 2, we detailed elaborate the proposed GAN. The experiments for demonstrating the effectiveness of this new image classification method are conducted in Sect. 3. There is a conclusion in Sect. 4.

2 Proposed GAN

2.1 Generative adversarial networks (GAN)

GAN consists of a generator network G and a discriminator network D. Figure 1 is the process of GAN.

Generator G maps the random noise z conforming to a specific distribution P_z(such as Gaussian distribution, uniform distribution, etc.) to the target domain, and uses it to learn the probability distribution P_{data} of real data, so that it can make it generate a sample $G(z)$ that conforms to the real data distribution $P_{data}(x)$ as far as possible. The discriminator D determines whether the input sample comes from the real data x or the generated data $G(z)$, and outputs a probability value $D(\cdot)$ belonging to the real data.

The goal of the generator is to fit "true" data (training samples) and generate "false" data. The goal of the discriminator is to distinguish the true and false data. The network structure of the generator and discriminator is a multi-layer perceptron. Given a real sample set $\{x^1, \cdots, x^n\}$. Suppose that p_x is its data distribution. It randomly samples data from another pre-defined distribution p_z to obtain noise set $\{z^1, \cdots, z^m\}$. Let the input of the generator be z. The output "false" data can be represented as $G(z)$. The input of the discriminator is "true" and "false" data in turn. The output is a one-dimensional scalar, representing the true probability of input. According to the difference of input, they can be expressed as: $D(x)$ and $D(G(z))$. Ideally, $D(x)=1$, $D(G(z))=0$. The network optimization process can be described as a “binary minimax” problem. The objective function is as follows:

$$\min_G \max_D E_{x\sim p_x} \ln D(x) + E_{z\sim p_x} \ln(1 - D(G(z))) \tag{1}$$

If the the data distribution of $G(z)$ is expressed as p_G, so there is a global optimal solution for the “binary minimax” problem, namely the $p_G = p_x$. The generator and discriminator are trained alternately. When updating parameters of generator(discriminator), the parameter of other one is fixed and not updated. In general, the discriminator is better learning ability than the generator. To keep the two in sync, it trains the generator k times and trains the discriminator one more. Through experiments, it is found that the learning ability of generator and discriminator changes with time. Therefore, in the subsequent experiments, this paper designs a dynamic learning method to keep the two in sync by observing the changes of loss function values.

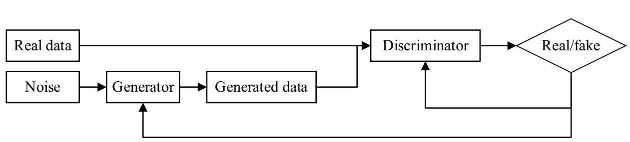


Fig. 1 Process of GAN

2.2 Modified GAN

Generator and discriminator are in a "confrontation" relationship. The ultimate goal is to enable the generator to perfectly fit the data distribution of real samples. Due to the lack of guidance from supervisory information, the fitting process is full of randomness. In practice, limited by the learning ability of the network, it is usually able to fit only a part of the real data distribution, leading to the loss of some modes, namely Mode collapse. As shown in Fig. 2, mode collapse will lead to redundancy of training results, poor image quality and other problems. Through the analysis of real data, it is not difficult to find that there are significant differences between different patterns. For example, men and women in human faces, day and night in scenes, etc., also have connections, such as five-feature structure, object shape and position, etc., which emphasizes differences while ignores connections. The key to solving the problem is to find a balance between them.

Thus, this paper designs the network structure as shown in Fig. 3. The training is synchronized by building two (or more) generators, sharing one input data and one discriminator. The training method is the same as the classic GAN. In addition, generators learn from each other, this step is called "collaboration," which guides each other and makes progress together. The "collaboration" is interspersed with normal training, and the rate can be adjusted according to the actual situation. For example, it trains the generator two times and collaborates one time. From the perspective of data distribution as shown in Fig. 4, classical GAN training can shorten the distance between the real distribution and the generated distribution. Collaborative training can shorten the distance between different generator generation distributions. This approach can not only improve the convergence speed of the model, but also enhance the learning ability of the model and reduce the possibility of mode collapse.

Because the input and discriminator networks are shared between generators, the generator distribution may appear to overlap phenomenon. This not only fails to achieve the desired goal, but also creates additional network load. In order to avoid this phenomenon, different network structures and random

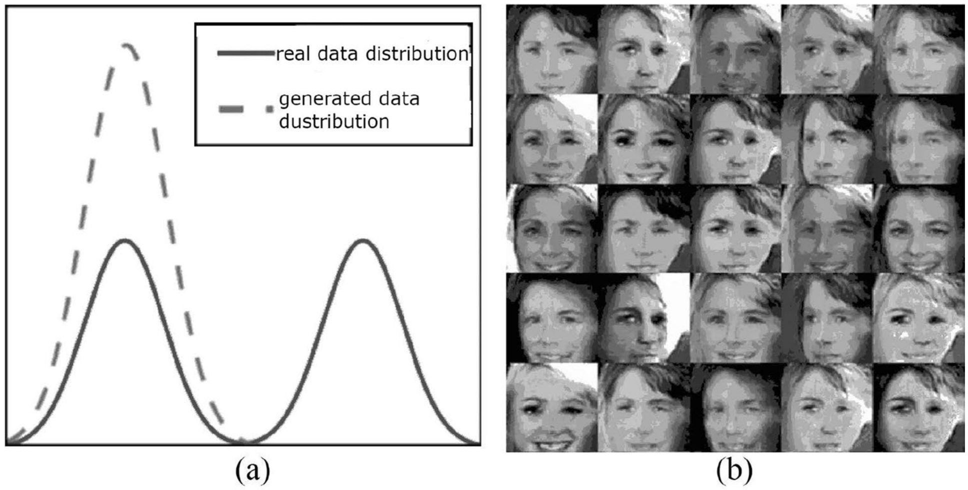


Fig. 2 Mode collapse in GAN. **a** synthetic data distribution cannot fit real data distribution in good shape; **b** mode collapse leads to synthetic data redundancy

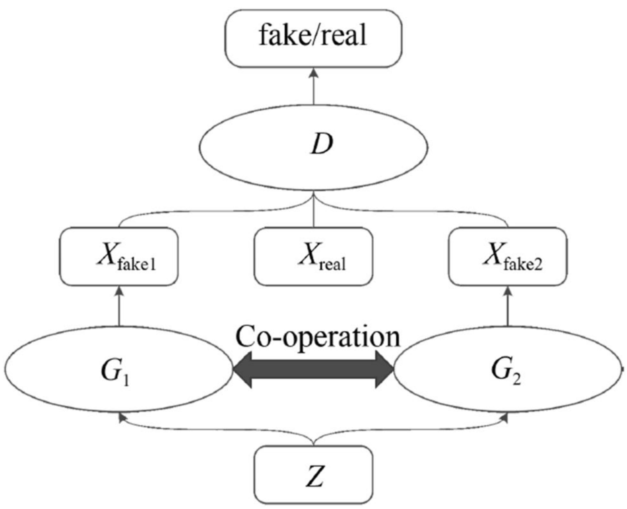


Fig. 3 Proposed GAN

weight initialization are adopted in designing the generator. Overlap problem does not occur in the actual training process, and the results produced by different generators are always different to some extent. The objective function of the discriminator is:

$$\max E_{x\sim p_x} \ln D(x) + E_{z\sim p_z} \ln(1 - D(G_1(z))) + E_{z\sim p_z} \ln(1 - D(G_2(z))) \tag{2}$$

For the generator, $E_{x\sim p_x} \ln D(x)$ is unaffected, so its objective function is:

$$\max E_{z\sim p_z} \ln D(G_1(z)) + E_{z\sim p_z} \ln D(G_2(z)) + \lambda L \tag{3}$$

$$L = -||G_1(z) - G_2(z)||_2 \tag{4}$$

where λ is the constant. The collaboration factor L selects L2-norm to shorten the distance between the generators. $D(G_1(z))$ and $D(G_2(z))$ are the discriminant results of

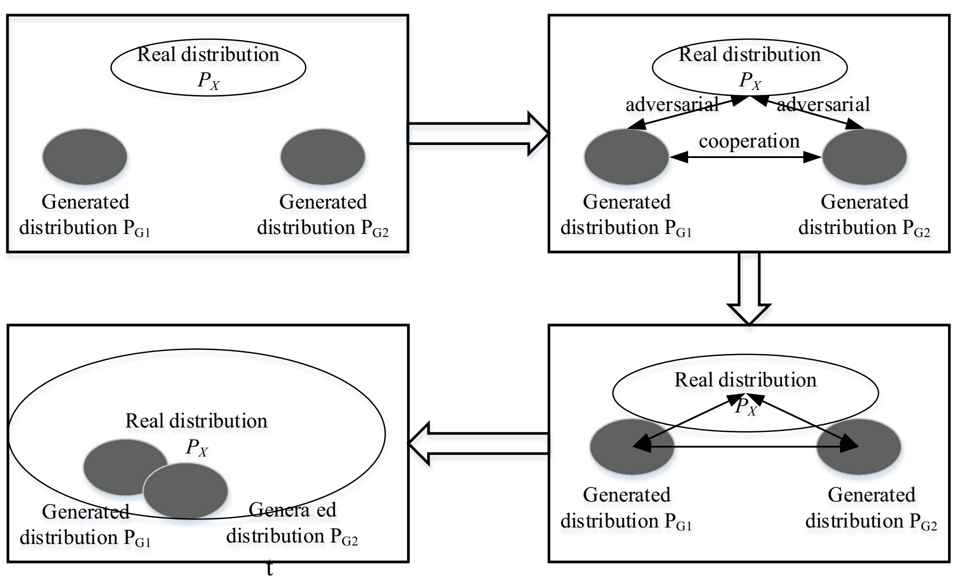


Fig. 4 Fitting process in proposed network

generated data by generator G1 and G2, respectively. Defining the parameters,

$$s = D(G_1(z)) - D(G_2(z)) \quad (5)$$

When s > 0, G1 has a higher score in the results obtained by discriminator D, that is, $G_1(z)$ has a higher image truth degree. So the distance between G2 and G1 should be shortened, which can be done by fixing G1 parameter, calculating the collaboration factor L, and punishing the network connection weight of G2. When s < 0, it is completely different. G2 should be fixed and G1 should be punished. The severity of punishment is related to the size of s. In this way, the generator with a higher score is judged to have an attractive force on the generator with a lower score. Due to the randomness of the network, G1 and G2 are trained alternately and assisted each other. Finally, they are converged to the real data distribution. To sum up, such a network structure is called "cooperative GAN".

3 Experiments and Analysis

In this paper, the proposed GAM is evaluated on two common benchmarks: Cifar and PASCAL VOC 2012. Comparison experiments are conducted under Linux16.04 operating system, CPU (32cores, 2.1 GHz) and GPU TTAN Xp 1060. Pytorch is the deep learning framework.

In this section, Cifar10 and Cifar100 data sets are used to evaluate the effectiveness of the proposed algorithm. Each data set contains 60,000 color images with size of 32 × 32, 50,000 images are used for training and 10,000 for testing. Data enhancement is done using the usual methods: clipping and random horizontal flipping.

In this section, Cifar10 and Cifar100 data sets are used to evaluate the effectiveness of the proposed algorithm. Each data set contains 60,000 color images with size of 32 × 32, 50,000 images are used for training and 10,000 for testing. Data enhancement is done using the usual methods: clipping and random horizontal flipping.

To ensure the fairness of the experiments, the Stochastic Gradient Descent (SGD) method [26] is adopted for the experiments. Momentum is set as 0.9. The training batch is set as 128. The testing batch is set as 100. The weight attenuation is set as 0.0005. The learning rate is 0.1. The

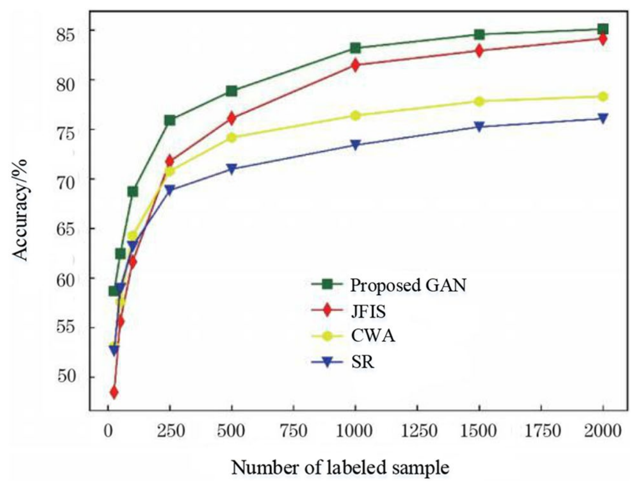


Fig. 5 The trend of average classification accuracy with different labeled samples

Table 1 Average classification accuracy with different labeled samples

Labeled data	SR	CWA	JFIS	Proposed GAN
25	57.23	59.52	53.67	**65.79**
50	65.24	61.71	61.82	**67.17**
100	67.31	67.97	67.14	**73.01**
250	72.24	76.19	76.85	**79.44**
500	76.88	77.37	79.51	**83.67**
1000	80.01	80.19	87.07	**87.88**
1500	79.44	80.34	87.22	**88.67**
2000	79.77	81.72	88.05	**90.26**

Bold values indicate best achieved results

Table 2 Top-1 error rate with different methods

Method	Parameter size/10^6	Top-1 error rate/%
SR	11.29	25.74
CWA	17.98	24.56
JFIS	21.34	22.38
Proposed GAN	**24.51**	**19.14**

Bold values indicate best achieved results

Table 3 Results on Cifar data set

Class	SR	CWA	JFIS	Proposed GAN
airplane	93.3	94.2	95.6	**98.1**
automobile	84.6	87.2	89.5	**92.4**
bird	82.3	83.6	85.7	**92.7**
cat	90.5	92.4	93.7	**96.8**
deer	83.3	86.4	89.6	**92.5**
dog	87.9	91.2	92.5	**96.7**
frog	87.1	88.2	88.7	**92.4**
horse	90.3	92.7	95.6	**96.8**
ship	91.3	93.6	94.5	**97.2**
struck	89.3	92.5	94.1	**96.6**

Bold values indicate best achieved results

Table 4 Results on PASCAL VOC 2012 data set

Class	SR	CWA	JFIS	Proposed GAN
aero	95.6	96.2	96.9	**98.2**
bike	84.3	86.5	87.2	**89.8**
bird	79.5	84.6	89.3	**92.7**
blt	58.9	59.6	62.8	**71.4**
boat	87.5	89.6	92.3	**95.6**
bus	91.5	92.6	93.3	**95.7**
car	86.9	89.4	91.3	**94.6**
cat	90.5	91.6	93.7	**96.1**
chair	68.7	70.2	72.3	**81.5**
cow	74.7	88.1	89.5	**92.1**
dog	89.6	92.5	94.6	**97.1**
hrs	89.7	91.6	92.3	**97.6**

Bold values indicate best achieved results

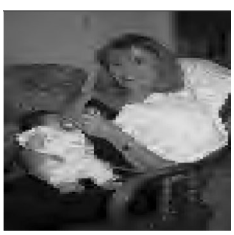
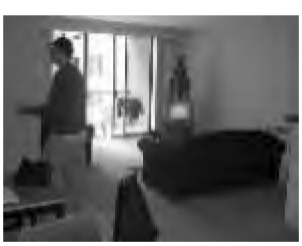
Table 5 Comparison of average accuracy, average error, running time value

Class	SR	CWA	JFIS	Proposed GAN
average accuracy/%	92.1	97.5	97.9	**98.2**
average error/%	5.6	4.2	3.8	**2.7**
Running time/s	2.95	1.78	1.23	**0.56**

Bold values indicate best achieved results

final result is the five-testing average value. The comparison methods are SR [27], CWA [28], JFIS [29]. Table 1 shows the average classification accuracy of the four algorithms with different labeled samples. Figure 5 shows the trend of average classification accuracy with different labeled samples.

Table 6 Image classification sample

Image	Manual labeled	SR	CWA	JFIS	Proposed
	Bicycle Bottle Person	Bottle Person	Bottle Person	Dining table Person	Bottle Dining table Person
	Chair Dining table Sofa	Chair Dining table	Dining table Sofa	Chair Dining table Sofa	Chair Dining table Sofa
	Bottle Chair Sofa Person	Chair Sofa Person	Bottle Sofa Person	Chair Sofa Person	Bottle Chair Sofa Person
	Bottle Chair Dining table Person	Bottle Person	Dining table Person	Bottle Dining table Person	Bottle Dining table Person
	Chair Dining table Person Sofa TV	Person Sofa TV	Chair Dining table Person Sofa	Chair Person Sofa TV	Chair Dining table Person Sofa TV

It can be seen from the above results that proposed GAN can achieve the same classification performance as JFIS with only a few labeled data. Proposed GAN significantly improves the classification performance and is better than the other two models. Table 2 shows the results of Top-1 error rate.

It can be seen from the data in the last column in Table 2 that the proposed GAN in this paper has the lowest Top-1 error rate comparing to other methods. Proposed can effectively improve the classification performance of the network by adding fewer parameters, which shows that the generalization of the module in this paper is good.

Table 3 and Table 4 are the classification results on the Cifar and PASCAL VOC 2012 data set with the four methods, which also show the better performance with the proposed GAN.

Table 5 displays the average accuracy, average error, running time values. As can be seen from Table 5, the average accuracy, average error, running time of proposed GAN are 98.2%, 2.7% and 0.56 s respectively, which is higher than that of SR, CWA, JHIS.

We also give the image classification sample as shown in Table 6. It can be seen that the classification results of the proposed model in this paper are better than other methods in most cases. The “sofa” in the second image, the “Dining table” in the fourth image and the “Chair” in the fifth image have occlusion phenomenon, the proposed GAN can correctly recognize these objects, which indicates that the proposed GAN has stronger robustness for the recognition of occlusion objects. Additionally, the proposed GAN recognizes the small object “bottle” in the third image indicating that the proposed GAN improved the ability of recognizing small objects. By further comparison, it can be found that for the first image, the “dining table” predicted by the new model is closer to the original semantics of the image than the artificially labeled “bicycle”. It shows that even if the classification of the model is not consistent with the manual label, it can reflect the image semantics correctly to some extent.

4 Conclusions

In this paper, an improved GAN model is proposed to solve the problem that the feature representation ability extracted by the existing neural network model is insufficient and the image classification and recognition accuracy is not high. By constructing multiple generators and introducing cooperative mechanism, they learn from each other and make progress together. The image classification quality can be significantly improved, network convergence speed can be accelerated. It improves the learning efficiency and reduces the possibility of mode collapse. The experiment results on the open public data set show that the new GAN model has better performance in image classification than other advanced models. In the future, more advanced deep learning methods will be applied in image classification.

Acknowledgements This work was supported by the Talent Training Joint Fund project (U1504609) of National Natural Science Foundation of China----Henan Government; General Project of Higher Education Reform Research and Practice in Henan Province (2017SJGLX400); Intelligent Robot and System Advanced Innovation Center Open Fund project in Beijing Institute of Technology (2018IRS09); Training Program for Young Backbone Teachers in Henan University of Higher Education (2017GGJS111).

References

1. Gu X, Angelov PP (2018) Semi-supervised deep rule-based approach for image classification. Appl Soft Comput 68:53–68
2. Yin SL, Zhang Y, Karim S (2018) Large scale remote sensing image segmentation based on fuzzy region competition and gaussian mixture model. IEEE Access 6:26069–26080
3. Asif AL, He H, Shafiq M, Khan A (2018) Assessment of quality of experience (QoE) of image compression in social cloud computing. Multiagent and Grid Systems 14(2):125–143
4. Kieffer B, Babaie M, Kalra S,. Tizhoosh HR (2017) Convolutional neural networks for histopathology image classification: Training vs. Using pre-trained networks, 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, 1–6
5. Karim S, Zhang Y, Asif AL, Muhammad RA (2017) Image processing based proposed drone for detecting and controlling street crimes. In 2017 IEEE 17th International Conference on Communication Technology (ICCT). 1725–1730
6. Yin SL, Li H, Liu DS, Karim S (2020) Active contour modal based on density-oriented birch clustering method for medical image segmentation. Multimedia Tools and Applications 79:31049–31068
7. Ayadi W, Elhamzi W, Charfi I, Atri M (2018) A hybrid feature extraction approach for brain MRI classification based on Bag-of-words. Biomed Signal Process Control 48:144–152
8. Kundegorski ME, Akcay S, Devereux M, Mouton A, Breckon TP (2016) On using feature descriptors as visual words for object detection within X-ray baggage security screening, 7th International Conference on Imaging for Crime Detection and Prevention (ICDP 2016), Madrid, 1–6
9. Yin SL, Bi J (2019) Medical image annotation based on deep transfer learning. J Appl Sci Eng 22(2):385–390
10. Jin B, Hu W, Wang H (2012) Image classification based on plsa fusing spatial relationships between topics. IEEE Signal Process Lett 19(3):151–154
11. Ghorai M, Chanda B (2015) An image inpainting method using pLSA-based search space estimation. Mach Vis Appl 26(1):69–87
12. Filisbino TA, Simao LB, Giraldi GA, Thomaz CE (2017) Combining deep learning and multi-class discriminant analysis for granite tiles classification, 2017 workshop of computer vision (WVC). Natal 2017:19–24
13. Teng L, Li H, Shahid K (2019) DMCNN: a deep multiscale convolutional neural network model for medical image segmentation. J Healthcare Eng 2019:8597606
14. Li P, Chen Z, Yang LT, Gao J, Zhang QC, Deen MJ (2019) An Incremental deep convolutional computation model for

feature learning on industrial big data. IEEE Trans Industr Inf 15(3):1341–1349
15. Yan Y, Zhu Q, Shyu M, Chen S (2016) A Classifier Ensemble Framework for Multimedia Big Data Classification, 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI), Pittsburgh, PA 615–622
16. Chaib S, Yao H, Gu Y, et al., (2017) Deep feature extraction and combination for remote sensing image classification based on pre-trained CNN models, In: International Conference on Digital Image Processing
17. Zhu F, Li H, Ouyang W, Yu N, Wang X (2017) Learning Spatial Regularization with Image-Level Supervisions for Multi-label Image Classification, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI 2027–2036
18. Mou L, Zhu XX (2020) Learning to pay attention on spectral domain: a spectral attention module-based convolutional network for hyperspectral image classification. IEEE Trans Geosci Remote Sens 58(1):110–122
19. You Y, Zhao Y (2019) A human pose estimation algorithm based on the integration of improved convolutional neural networks and multi-level graph structure constrained model. Pers Ubiquit Comput 23(3–4):607–616
20. Teng L, Li H, Yin SL, Sun Y (2019) Modified krill group-based region growing algorithm for image segmentation". Int J Image Data Fusion 10(4):327–341
21. Wang Y, Yue J, Dong Y et al (2016) Review on kernel based target tracking for autonomous driving. J Inform Process 24(1):49–63
22. Jiang W, Luo X (2019) Research on unsupervised coloring method of chinese painting based on an improved generative adversarial network. World Sci Res J 5(11):168–176
23. Zhao X, Gao L, Chen Z, et al., (2019) Large-scale Landsat image classification based on deep learning methods[J]. APSIPA Transactions on Signal and Information Processing 8
24. Wang X et al (2020) UD-MIL: uncertainty-driven deep multiple instance learning for OCT image classification. IEEE J Biomed Health Inform 24(12):3431–3442
25. Yin S, Li H (2020) Hot region selection based on selective search and modified fuzzy c-means in remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13:5862–5871
26. Goodfellow IJ, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial networks. Adv Neural Inf Process Syst 3:2672–2680
27. Dundar T, Ince T (2019) Sparse Representation-Based Hyperspectral Image Classification Using Multiscale Superpixels and Guided Filter. IEEE Geosci Remote Sens Lett 16(2):246–250
28. Li P, Chen P, Xie Y, Zhang D (2020) Bi-modal learning with channel-wise attention for multi-label image classification. IEEE Access 8:9965–9977
29. Dornaika F (2020) Joint feature and instance selection using manifold data criteria: application to image classification. Artif Intell Rev 54:1735–1765