HIGH TECHNOLOGY LETTERS | Vol. 25 No. 1 | Mar. 2019 | pp. 35 ~ 41

doi:10.3772/j.issn.1006-6748.2019.01.005

General image classification method based on semi-supervised generative adversarial networks①

Su Lei(苏 磊)*, Xu Xiangyi*, Lu Qiyu*, Zhang Wancai②**
(* Electric Power Research Institute of State Grid Shanghai Electric Power Company, Shanghai 200437, P. R. China)
(** Beijing State Grid Fidelity Technology Development Co., Ltd, Beijing 100067, P. R. China)

Abstract

Generative adversarial networks (GANs) have become a competitive method among computer vision tasks. There have been many studies devoted to utilizing generative network to do generative tasks, such as images synthesis. In this paper, a semi-supervised learning scheme is incorporated with generative adversarial network on image classification tasks to improve the image classification accuracy. Two applications of GANs are mainly focused on: semi-supervised learning and generation of images which can be as real as possible. The whole process is divided into two sections. First, only a small part of the dataset is utilized as labeled training data. And then a huge amount of samples generated from the generator is added into the training samples to improve the generalization of the discriminator. Through the semi-supervised learning scheme, full use of the unlabeled data is made which may contain potential information. Thus, the classification accuracy of the discriminator can be improved. Experimental results demonstrate the improvement of the classification accuracy of discriminator among different datasets, such as MNIST, CIFAR-10.

Key words: generative adversarial network (GAN), semi-supervised, image classification

0 Introduction

Generative adversarial networks (GANs) are a class of methods which are based on game theory to learn generative models. The aim of GANs is to train generator G to produce samples from the data distribution. The generator's input is a noise vector. With real data distribution, the noise vector can be transformed as generated samples, which are similar to real data samples. Then the generated samples are input into discriminator D which usually attempts to distinguish generated distribution $p_{\text{model}}(x)$ from real samples. It transmits feedback signal to the generator, then the generator's weights are updated. And generator G in turn is trained to deceive the discriminator to recognize the fake data as real.

Recently, the applications of GANs have made great progress. Many applications have shown that they can produce excellent samples[1]. However, the training process of GANs requires balance between the generator and the discriminator. Utilizing gradient descent techniques is typically trained where the techniques are devoted to finding the lowest value of the cost function. However, this value does not satisfy the demand of the GANs, it is hard to find the Nash equilibrium of the game. Usually, the game between the generator and the discriminator may fail to converge and the gradient may descend continuously until vanish.

Image classification tasks have become a hard and time-consuming task. There remains a problem, the data is huge, while useful data is becoming less and less. Most of the data are unknown to us. Therefore, the full use of the unknown data is of vital importance. In this work, the semi-supervised learning is incorporated with generative adversarial networks to improve the sample generation and classification accuracy.

The rest of the paper is structured as follows. In Section 1, related work of image classification and generative adversarial networks are discussed. In Section 2, the proposed method is elaborated. In Section 3, experiments are conducted to demonstrate the effectiveness of the proposed method. In Section 4, conclusions are drawn.

1 Related work

There has been so much work having been carried

① Supported by the National Natural Science Foundation of China (No. 61501457) and National Key Technology R&D Program (No. 2015BAK21B00).
② To whom correspondence should be addressed. E-mail: zhangwancai100067@163.com
Received on Mar. 5, 2018

out to design suitable classifiers to deal with the classification problems in last decades[2-5]. Generally, those methods can be classified into three types, i. e., unsupervised, supervised and semi-supervised methods. Unsupervised methods focus on training models from large unlabeled samples. For the unsupervised methods do not need any labeled data, it can be easily applied in the classification tasks. Many unsupervised methods, such as clustering[6], graph-based method[7], have demonstrated impressive results in classification tasks. However, without the priori knowledge, one cannot ensure the relationship between clusters and classes.

Supervised classifiers, which are widely used in classification tasks, can also be utilized to improve the performance by utilizing the prior information of the class labels. Typical supervised classifiers include support vector machine (SVM)[8,9], artificial neural networks (ANN)[10] and sparse representation-based classification (SRC)[11,12], etc. SVM is a kind of kernel-based method that aims at exploring the optimal separating hyperplane between different classes. ANN is motivated by the biological learning process of human brain, while the SRC stems from the rapid development of compressed sensing in recent years. Although the supervised classifiers are of different use, the classifiers' performance is heavily decided by the number of labeled samples. Therefore, in the supervised learning method, the unlabeled data is of no use. However, unlabeled data is huge in modern society. Making full use of the unlabeled data to do the classification tasks is of vital urgent.

Semi-supervised learning is designed to relax the small sample problems, when it comes to the condition that the labeled samples are not enough. Thus, wealth of unlabeled samples is significant. The semi-supervised methods can be roughly divided into four types:

- Generative models[13] estimate the conditional probability density to obtain labels of unlabeled samples.
- Graph-based methods[14]. This kind of method utilizes labeled and unlabeled samples to construct graphs. And at the same time, it minimizes the energy function, which is aimed to assign labels to unlabeled-samples.
- Low density separation[15]. It is aimed to place boundaries in regions where few labeled or unlabeled data exist. The transductive support vector machine (TSVM)[16] is one of the state-of-the-art algorithms.
- Wrapper-based methods. This kind of method applies the supervised method and labels the unlabeled data iteratively. Self-training[17,18] and co-training[19] algorithms are the commonly used wrapper-based methods.

The traditional semi-supervised learning methods have shown better performance. However, with the development of deep learning, the neural networks can hierarchically obtain high level abstract representation[20,21], which has recently become the main stream in the image processing area, especially in classification tasks. The typical deep architecture includes deep brief network (DBN)[22], convolutional neural networks (CNN)[23], etc. What is mentioned above are all supervised learning frameworks, which require a large number of labeled samples for training.

Recently, generative adversarial networks (GANs) have been applied to the image generation tasks with generative and discriminative convolutional networks successfully. Goodfellow et al.[24,25] proposed the theoretical framework of generative adversarial networks. And this framework can generate images without any supervised information. Later, Radford et al.[26] proposed deep convolutional generative adversarial networks (DCGANs) for unsupervised representation. To solve the situation of gradient vanishing, WGAN[27] is proposed to use the Wasserstein distance instead of the Jensen-Shannon divergence, to make the data set distribution compared with the distribution learned by the generator. Obviously, they show that the sample quality is closely related to the network's convergence and the training rate is really improved. Another direction of image synthesis with GANs is to synthesize images by conditioning on supervised information, such as text or class labels. Conditional GAN[28] is one of the work that develop a conditional version of GANs by additionally feeding class labels into both generator and discriminator of GANs. Info-GAN[29] introduces a new concept, which divides input noise z into two parts, one is the continuous noise signal that cannot be explained, and the other is called C. Where C represents a potential attribute that can be interpreted as a facial expression. In addition, the concept of information theory is added to GAN. Recently, Reed et al.[30] utilize GANs for image synthesis using given text descriptions, enabling translation from character level to pixel level.

2 The general image classification method with semi-supervised GANs

2.1 Adversarial learning

The adversarial learning adopts the idea of game theory and is combined with the unsupervised learning to jointly train the model. The training is formalized as a game in which the generative model is trained to generate outputs to fool the discriminator. Reasonably definition, the models of adversarial learning are trained to compete with each other and can continually improve the output of each model. In generative adversarial networks, the generative model tries to generate images from noise as real as possible, and the discriminative model is equivalent to a binary classifier, which is utilized to differentiate the generated samples and real samples. Based on concealing the discriminator, the generator updates its own weights, and updates the weights of the discriminator by distinguishing real and fake samples at the same time. The adversarial model can be described as the following competition game:

$$\min_G \max_D J(D, G) = E_{x \sim p_{\text{data}}(x)} \log(D(x)) + E_{z \sim p_{\text{noise}}(z)} \log(1 - D(G(z))) \quad (1)$$

where $D(x)$ representing the probability of recognizing is a real image rather than a generated image. $G(z)$ represents the generated images after inputting the noise. In this process, G and D are trained simultaneously:

- The input of the generator is the noise in the distribution of p_{noi}. Through training, the noise distribution can be transformed into the data distribution p_{dat}, which can be guided to generate the samples as real as possible.
- The discriminator's input consists of two parts: the real data and the generated data. The discriminator is devoted to distinguishing the real data and generated data.

In order to solve the minimum and maximum game problem. The gradient of the discriminator network is optimized first in a manner of gradient ascent in each iteration and the parameters of the generated network are updated in gradient descent. Let ω_N represent the neural network N, and the optimization process can be written as follows.

- Let discriminator D fixed, to update the generator G by the following:

$$\omega_G \leftarrow \omega_G - \gamma_G \Delta_G J \quad (2)$$

$$\Delta_G J = \frac{\partial}{\partial \omega_G} E_{z \sim p_{\text{noise}}(z)} \log(1 - D(G(z, \omega_G), \omega_D)) \quad (3)$$

- Let generator G fixed, to update discriminator D by

$$\omega_D \leftarrow \omega_D + \gamma_1 \quad (4)$$

$$\Delta_D J = \frac{\partial}{\partial \omega_D} \{E_{x \sim p_{\text{data}}(x)} \log(D(x, \omega_D)) + E_{z \sim p_{\text{noise}}(z)} \log(1 - D(G(z, \omega_G), \omega_D))\} \quad (5)$$

The above equations are the parameter updating rules of generator and discriminator. Generally, the both networks obey the principle of stochastic gradient descent, which is used to minimize the empirical loss function.

2.2 Semi-supervised learning

The semi-supervised learning needs label information when dealing with the classification tasks. It is halfway between supervised and unsupervised learning. In addition to unlabeled data, the semi-supervised learning is provided with some supervision information, but not necessarily for all examples. To formalize the semi-supervised learning to a more mathematical formulation, a knowledge on $p(x)$ is given that the variable gains through the unlabeled data has to carry information that is useful in the inference of $p(y | x)$. Obviously, if this is not the case, semi-supervised learning will not yield an improvement over supervised learning. Semi-supervised learning scheme makes full use of the unlabeled data, in which there are an abundance of structural information to learn. A more formal description can be given that is a data set $X = x_{i(i \in n)}$, which can be divided into two parts, $X_k = (x_1, x_2, \cdots, x_k)$ labeled $Y_k = (y_1, y_2, \cdots, y_k)$ and the other set $X_m = (x_{x+1}, x_{x+2}, \cdots, x_{k+m})$, whose label is not known. Then the data can be trained to learn the distribution to do the classification tasks.

Learning from the formal description, it is considered that a standard classifier for classifying data point x into one of N possible classes, the main idea of the model is training a network playing both the roles of a classifier performing image classification task as well as a discriminator trained to distinguish generated samples produced by a generator from the real data. To be more specific, the discriminator takes images as input and classifies them into $N+1$ classes. And the true samples are classified into the first N classes and generated samples are classified into the $(N+1)$th class, as is shown in Fig. 1.

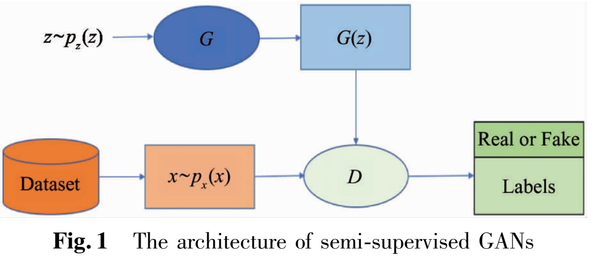

Fig. 1 The architecture of semi-supervised GANs

The discriminator is not only utilized to distinguish real images from generated images, but also to classify images into different classes.

2.3 Model design

2.3.1 Generator

In generative adversarial networks, the generator network is a density network whose job is to produce realistic-looking samples. The generator accepts random noise vectors as inputs. Besides noise vector z, the generator also needs to learn the distribution of real images to fit noise distribution, which is the process of generation tasks. The network learns pixel features of the real images to fit noise and generate images as real as possible. For the model design of generator, four fractionally-strided convolutional layers and a fully connected layer are used. The fractionally-strided convolutional layer is utilized to extract the object's shape through the learned features. The fully connected layer is utilized to map the learned distributed feature representation to the sample space. The architecture is shown as Fig. 2.

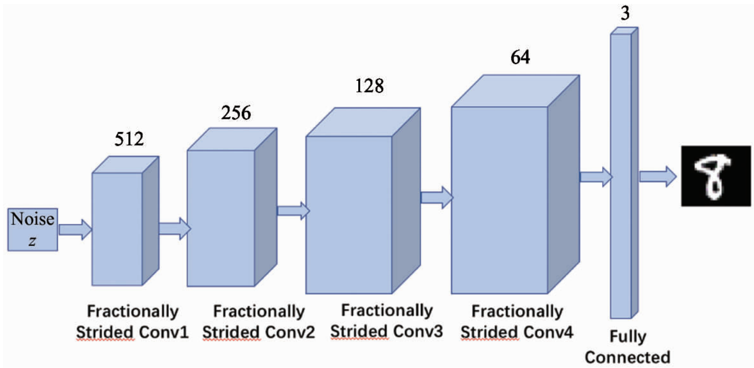

Fig. 2 The architecture of generator

2.3.2 Discriminator

The discriminator tries to figure out whether an image come from the training set or the generator network. Different from the generator, this process is a down-sampling process. The purpose is to extract features from a certain model and then classify the images based on the features. In this process, the most important step is feature extraction, whose aim is to find the suitable features that can distinguish images most. For the model design of discriminator, it is not only utilized to distinguish the real images from the generated images, but also to perform the multi-classification tasks. A fully connected layer and four convolutional layers are used. Finally, a classifier layer is used to perform the classification tasks. The architecture is shown in Fig. 3.

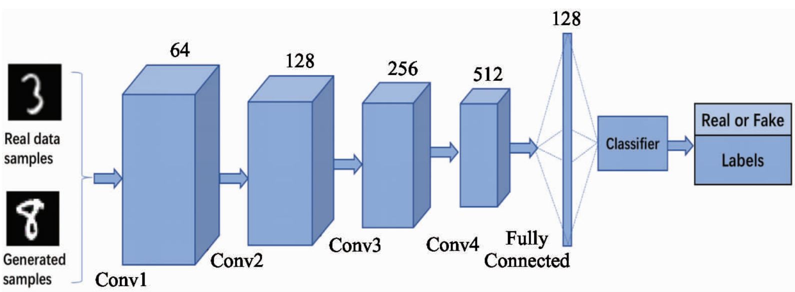

Fig. 3 The architecture of discriminator

2.3.3 Design of loss function

As mentioned in Section 2.2, samples from the generated images are simply added to the training dataset, labeling them with the class "$N+1$". Thus, the dimension of our classifier can be extended to "$N+1$". Therefore, the probability can be expressed as fake as $p_{\text{model}}(N+1 \mid x)$, equal to $G(z)$ in the original GANs. It can be also learnt from the unlabeled data by maximizing the $p_{\text{model}}(y \in \{1, \cdots, N\} \mid x)$. Then the loss function can be represented as follows:

$$\begin{aligned} L &= -\alpha E_{x, y \sim p_{\text{data}}(x, y)}[\log p_{\text{model}}(y \mid x)] \\ &\quad - \beta E_{x \sim G(z)}[\log p_{\text{model}}(y = N+1 \mid x)] \\ &= \alpha L_{\text{supervised}} + \beta L_{\text{unsupervised}} \end{aligned} \tag{6}$$

where, $L_{\text{supervised}}$ and $L_{\text{unsupervised}}$ is calculated as follows:

$$L_{\text{supervised}} = -\alpha E_{x, y \sim p_{\text{data}}(x, y)}[\log p_{\text{model}}(y \mid x, y < N+1)] \tag{7}$$

$$L_{\text{unsupervised}} = -\beta\{E_{x\sim p_{\text{data}}(x)} \log(1 - p_{\text{model}}(y = N+1 \mid x)) \; E_{x\sim G(z)}[\log p_{\text{model}}(y = N+1 \mid x)]\} \quad (8)$$

where α and β are the weight coefficients, which need to satisfy constraint $\alpha + \beta = 1$.

The cross-entropy loss consists of two parts: $L_{\text{supervised}}$ and $L_{\text{unsupervised}}$. $L_{\text{supervised}}$ loss function is given the real data with the negative log probability of the label, while $L_{\text{unsupervised}}$ loss function seems like the standard GAN loss, considering the loss of noise and real data.

3 Experiments

3.1 Data preparation

All experiments are performed in Tensor Flow[31] on a workstation with a Titan X GPU. In the experiments, the MNIST and CIFAR-10 datasets are used. The MNIST dataset contains 60 000 labeled images of digits. Its training set consists of numbers written by 250 different people, 50% of whom are high school students and 50% of the data are from the staff of Census Bureau. The test set also has the same proportion of handwriting digital data.

CIFAR-10 is a small, well studied dataset of 32 × 32 natural images whose number is also 60 000. This dataset is divided into 10 categories. The class of each contains 6 000 images. There are 50 000 for training, and another 10 000 for testing. The ten categories are independent of each other.

These two datasets are used to study semi-supervised learning, as well as to examine the visual quality of samples that can be achieved. For the discriminator, besides the convolutional networks, the weight normalization is also used. And the generator is a network with 4 deep convolutional layers and each layer has a batch normalization.

3.2 Experimental settings

For the number of the two datasets is the same, same experimental settings are used. 50 000 images as training set, another 10 000 are utilized as test set. For the training dataset, it is considered the setups in four conditions, considering the setups with 0, 1 000, 2 000, 3 000, 5 000, 8 000 labeled samples. Correspondingly, the remaining training data size is 50 000, 49 000, 48 000, 47 000, 45 000, 42 000, composed of generated images, as unlabeled data.

Through semi-supervised training with the same epochs, the network can somehow improve the quality of the generated images, as shown in Fig. 4 and Fig. 5.

Fig. 4 For the MNIST dataset, the left is the generated samples through semi-supervised training, the right is the generated samples through unsupervised training

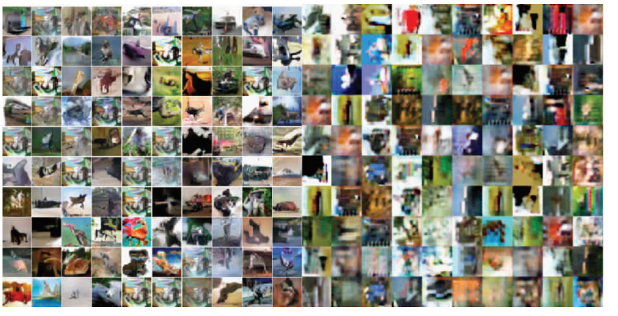

Fig. 5 For the CIFAR-10 dataset, the left is the generated samples through semi-supervised training, the right is the generated samples through unsupervised training

Experimental results also show that through the use of semi-supervised learning, the convergence speed is improved compared with unsupervised learning, as shown in Table 1.

Table 1 The time-consuming contrast of semi-supervised learning scheme and unsupervised learning scheme after the same epochs on MNIST dataset and CIFAR-10 dataset

Methods	Datasets	Time (mins)
Semi-supervised learning	MNIST	87.7
	CIFAR-10	103.2
Unsupervised learning	MNIST	96.1
	CIFAR-10	115.3

From the results, it can be concluded that for the same dataset, semi-supervised learning can learn to converge faster and get higher efficiency, which shows an advantage among unsupervised learning method.

As for the parameters, RMSProp optimizer is used to optimize the parameters. The learning rate of the generator and discriminator is 2×10^{-4}. In each mini-batch, the parameters of the discriminator is updated once, and the generator's parameters are updated twice at the same time.

3.3 Experimental results

The proposed method is validated from theory and experiment. On one hand, the semi-supervised learning scheme can enhance the visual quality of the generated images. On the other hand, through semi-supervised learning scheme, the classification accuracy is improved. The results are shown in Table 2 and Table 3.

Table 2 The contrast of semi-supervised classification accuracy with different number of labeled images and supervised classification accuracy on MNIST dataset

Number of labeled images	Classification accuracy
0	0.892
1000	0.895
2000	0.904
3000	0.917
5000	0.925
8000	0.930
Supervised training	0.968

Table 3 The contrast of semi-supervised classification accuracy with different number of labeled imagesand supervised classification accuracy on CIFAR-10 dataset

Number of labeled images	Classification accuracy
0	0.750
1000	0.761
2000	0.785
3000	0.792
5000	0.801
8000	0.813
Supervised training	0.864

From the results on two datasets, it can be seen that with the increase of the labeled data, the classification accuracy is improved gradually. The semi-supervised training scheme performs better than unsupervised training scheme on image classification tasks. And to some extent, the accuracy can approach to the supervised classification results, which shows the generalization of the semi-supervised learning scheme.

4 Conclusions

In this paper, semi-supervised learning scheme is applied to generative adversarial networks to perform classification tasks, which perform a better performance than the normal classifiers. Besides, through semi-supervised learning scheme, unlabeled data can be used to enlarge the sampling space. And the generated image quality is improved, either. The performance of the model among different datasets is evaluated, such as MNIST and CIFAR-10. Results show the effectiveness of the proposed model through the classification accuracy. It is hopeful to develop and find a more rigorous theoretical understanding on semi-supervised learning scheme in future work.

References

[1] Zhang X Y, Shi H C, Zhu X B, et al. Active semi-supervised learning based on self-expressive correlation with generative adversarial networks [J]. *Neurocomputing*, 2019. doi:10.1016/j.neucom.2019.01.083

[2] Zhang X Y, Wang S P, Yun X C. Bidirectional active learning: a two-way exploration into unlabeled and labeled dataset[J]. *IEEE Transactions on Neural Networks and Learning System*, 2015, 26(12): 3034-3044

[3] Zhang X Y, Wang S P, Zhu X B, et al. Update vs. upgrade: modeling with indeterminate multi-class active learning[J]. *Neurocomputing*, 2015, 162: 163-170

[4] Zhang X Y. Simultaneous optimization for robust correlation estimation in partially observed social network[J]. *Neurocomputing*, 2016, 205: 455-462

[5] Zhang X Y. Interactive patent classification based on multi-classifier fusion and active learning[J]. *Neurocomputing*, 2014, 127: 200-205

[6] Zhong Y F, Ma A L, Zhang L P. An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(4):1235-1248

[7] Zhu W, Chayes V, Tiard A, et al. Unsupervised classification in hyperspectral imagery with nonlocal total variation and primal-dual hybrid gradient algorithm[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(5):2786-2798

[8] Vladimir C. The nature of statistical learning theory[J]. *IEEE Transactions on Neural Networks*, 1997, 8(6): 1564

[9] Kuo B C, Ho H H, Li C H, et al. A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification [J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2013, 7(1):317-326

[10] Braspenning P J, Thuijsman F, Weijters A. Artificial neural networks: an introduction to anntheory and practice [J]. *Lecture Notes in Computer Science*, Springer, 1995, (931)

[11] Chen Y, Nasrabadi N M, Tran T D. Hyperspectral image classification via kernel sparse representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(1):217-231

[12] Zhang Y X, Du B, Zhang L P, et al. Joint sparse representation and multitask learning for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(2):894-906

[13] Li J, Bioucas-Dias J M, Plaza A J. Semi-supervised hyperspectral image classification using a new (soft) sparse multinomial logistic regression model[C]. In: Proceedings of the 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, WHISPERS 2011, Lisbon, Portugal, 2011. 1-4

[14] Belkin M, Niyogi P, Sindhwani V. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples[J]. *Journal of Machine Learning Research*, 2006, 7:2399-2434
[15] Joachims T. Transductive inference for text classification using support vector machines[C]. In: Proceedings of the 16th International Conference on Machine Learning (ICML 1999), Bled, Slovenia, 1999. 200-209
[16] Bruzzone L, Chi M, Marconcini M. A novel transductive SVM for semisupervised classification of remote-sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(11-2):3363-3373
[17] Dopido I, Li J, Marpu P R, et al. A semi-supervised self-learning forhyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51, 4032-4044
[18] Aydemir M S, Bilgin G. Semi-supervised hyperspectral image classification using small sample sizes[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 14: 621-625
[19] Zhang X, Song Q, Liu R, et al. Modified co-training with spectral and spatial views for semi-supervised hyperspectral image classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6):2044-2055
[20] Zhang X Y, Shi H C, Li C S, et al. Learning transferable self-attentive representations for action recognition in untrimmed videos with weak supervision[C]. In: Proceedings of the 33d AAAI Conference on Artificial Intelligence, Hawaii, USA, 2019. 1-8
[21] Cui Z Y, Li Z K, Wu S, et al. Dressing as a whole: outfit compatibility learning based on node-wise graph neural networks[C]. In: Proceedings of the 30th World Wide Web Conference, San Francisco, USA, 2019. 1-11
[22] Chen Y S, Zhao X, Jia X P. Spectral-spatial classification of hyperspectral data based on deep belief network [J]. *IEEE Journal of Selected Topics in Applied Earth Observationsand Remote Sensing*, 2015, 8(6):2381-2392
[23] Li Y, Zhang H K, Shen Q. Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network[J]. *Remote Sensing*, 2017, 9(1):67
[24] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]. In: Proceedings of the Annual Conference on Neural Information Processing Systems, Montreal, Canada, 2014. 2672-2680
[25] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training GANs[C]. In: Proceedings of the Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 2016. 2226-2234
[26] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv: 1511.06434, 2015
[27] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [J]. arXiv: 1701.07875, 2017
[28] Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs[C]. In: Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, Australia, 2017. 2642-2651
[29] Chen X, Duan Y, Houthooft R, et al. Infogan: interpretable representation learning by information maximizing generative adversarial nets[C]. In: Proceedings of the Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 2016. 2172-2180
[30] Reed S, Akata Z, Yan X C, et al. Generative adversarial text to image synthesis[C]. In :Proceedings of the 33nd International Conference on Machine Learning, New York City, USA, 2016. 1060-1069
[31] Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems[J]. arXiv:1603.04467v1, 2016

Su Lei, born in 1981. He received his M. S. degree in Power Electronics and Power Transmission major of Shanghai Jiaotong University. Now he is a senior engineer in State Grid Shanghai Electric Power Company Electric Power Research Institute. His research interests include data mining and power transmission equipmeat online monitoring.