Research on asphalt materials based on machine vision and generate adversarial networks

Wenyin Song[1, 2, *], Haibin Li[1, 2, a]

[1]Shan dong Provincial Academy of Building Research Co.,Ltd,Ji Nan 250100,China;

[2]Shan dong Quality Inspection and Testing Center of Construction Engineering Co. , Ltd. Ji Nan 250000,China;

*Corresponding Author Email: 389036081@qq.com, [a]21063754@qq.com

Abstract. With the rapid development of artificial intelligence and deep learning, computer vision-oriented generative models have been widely used. Among them, the generation of adversarial network has the most far-reaching influence. To solve the problem of dynamic instability in the training of generative adversarial networks, this paper proposes a rapid construction method of generative adversarial networks based on hidden layer characterization. The generation process of the adversarial network is divided into two independent generation processes to generate the representation of the experience hidden layer and the final result respectively, so as to stably generate the training dynamics of the adversarial network and capture more data patterns. This method can effectively and stably generate the training dynamics of the adversarial network. Finally, theoretical analysis proves that this method can stably generate the training dynamics of adversarial network and reduce the difficulty of adversarial training. Large-scale experiments on multiple data sets demonstrate the effectiveness of the proposed method.

Keywords: Generate adversarial networks, train dynamic instability, mode collapse, hidden layer characterization, and feature equalization.

1. Introduction

With the development of deep learning technology, deep learning methods based on cyclic neural networks and convolutional neural networks have achieved remarkable results in the field of data generation, especially in the field of visual data generation. Pixe1RNN is the most representative work in the generation model based on cyclic neural network. It converts the data distribution into the learning and prediction of the pixel sequence, which makes the generation task can be solved by the autoregressive algorithm. Pixe1RNN first assumes that the value of each pixel only depends on meaningful neighbors in the space, and then uses observable given points to model the conditional distribution of each pixel, so as to observe and learn the distribution of real data. However, despite the strong fitting ability of neural network, Pixe1RNN still has the problem of low data sampling efficiency. On the other hand, variational autoencoders based on encoder and decoder structure mainly rely on deep neural networks to complete the distribution of Bayesian inference learning target data. The hidden space of the variational autoencoder is designed to be continuously distributed for random sampling and interpolation, which makes it a more efficient generation model than the previous work. Through the research of the above generation model, it can be found that how to build a good generation model is a very challenging problem. In 2014, Goodfellow I proposed a completely new Generative model, Generative Adversarial Networks (GANs). The first model has become the most influential representative model of deep learning generation model by virtue of its superior feature capture ability and excellent generation performance. Yann LeCun, one of the three deep learning wagons, even called it the most interesting development in AI in the last decade.[1].

2. Construction and training of counter network framework

In In order to stabilize the training dynamics of generative adversarial networks and alleviate the problem of mode collapse, this paper proposes a new rapid construction framework for generative adversarial networks. The previous direct mapping construction process from random white noise to data distribution is decomposed into two independent mapping construction processes from random white noise to hidden layer representation and from empirical hidden layer representation to data distribution. To this end, we train a combination of two completely independent generators and discriminators to generate samples of empirical hidden layer representations and data distributions that are close to hidden layer representations of data distributions represented by autoencoders, respectively. Figure 1 shows the rapid build framework proposed in this article. All components in the framework are parameterized by the neural network. An autoencoder is trained to map the data flow into a hidden manifold, as shown in the blue dashed box. A pair of small-scale generators and discriminators parameterized by a convolutional neural network are then played against each other so that the generator can map random white noise into a hidden layer representation of a true distribution, as shown in the green dashed box. Finally, a large generative adversarial network consisting of a residual network is trained to search for the target transition mapping from the hidden layer representation to the true distribution, as shown in the black dashed box[2].

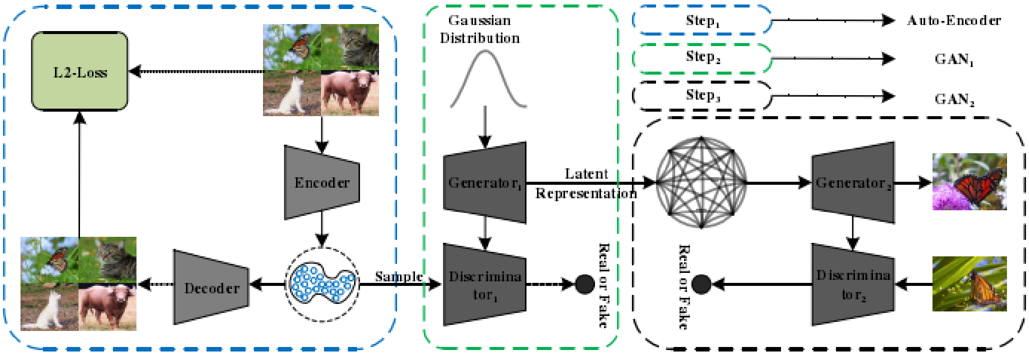

Fig. 1 The basic framework for rapid build methods

2.1. Data embedding in hidden layer space is realized by using autoencoder

A large residual network based autoencoder is introduced into the framework to achieve target transfer mapping from data distribution to reconstructed samples. The average absolute error loss function is used to calculate the reconstruction error between the data distribution and the reconstructed sample. In this process, it can be found that the encoder is trained to encode the data manifold from the real image in the hidden layer space, and the purpose of mapping the data distribution to the coded distribution in the hidden layer space is realized. Correspondingly, the decoder then decodes the encoding distribution in the hidden layer space back into the data manifold. By encoding and decoding the real data, the original data distribution is mapped into the hidden space. With the progress of training, the autoencoder can continuously reduce the reconstruction error between the data distribution and the reconstructed samples, which means that the representation of the data distribution projected into the hidden layer space can be continuously optimized, and ultimately will be the unbiased estimation of its probability distribution. Compared with data distribution, hidden layer representation usually has a lower dimension and a more compact distribution, so it is easier to be learned and encoded by neural networks. In addition, the calculation of the autoencoder is equivalent to the calculation of the encoding map f_θ and the decoding map g_ε[3]::

$$(\nu_{gt}, \chi) \xrightarrow{f_\theta} (\mu_{gt}, \Omega) \xrightarrow{g_\varepsilon} (\nu_{gt}, \chi) \tag{1}$$

Where, f_θ and g_ε is parameterized by convolutional neural network, and $f_\theta : \chi \rightarrow \Omega$ is a data embedding that pushes the probability measure ν_{gt} in feature space R^d to the data distribution μ_{gt} in hidden layer space. In the training process of autoencoder, f_θ is a homeomorphism mapping but an inverse homeomorphism mapping g_ε.

2.2. Construct target transition mapping from random white noise to empirical hidden layer representation

In the generative adversarial network, the primary purpose is to convert random hidden space variables into new samples approximating the data distribution, which requires that the initial input should not have any prior knowledge of the real data. If we directly construct the target transition mapping from the hidden layer representation of data distribution in the hidden space of the autocoder to the data distribution, although we can obtain the optimal result, it is contrary to the goal of generating adversarial network. Therefore, this paper first needs to construct a target transition mapping from random white noise to hidden layer representation. At this stage, the main purpose of the rapid construction framework is to construct a target transition mapping from random white noise ζ to the hidden layer representation μ_{gt} and formalize it into the formula (2), where the empirical hidden layer representation is the projection of the hidden layer representation of the data distribution into the discriminator manifold[4].

$$(\zeta, Z) \overset{g_e}{\rightarrow} (\mu, \Omega) \tag{2}$$

g_e is parameterized by a neural network ,and μ is a distribution whose support set is similar to the topology μ_{gt}.

In the rapid construction framework of generative adversarial network proposed in this paper, in order to save computing costs as much as possible, only a small-scale standard generative adversarial network based on convolutional neural network is used at the current stage to achieve this purpose, considering that the hidden layer characterization only contains very low dimensions. The generator G_1 was used to generate new samples that approximated the low-dimensional hidden layer representation of the data distribution in the hidden space of the autoencoder, while the discriminator D_1 was trained to check whether the distribution of the generated samples approximated the sample obtained from the empirical hidden layer distribution sampling. Because the hidden layer representation is a low-dimensional projection of data distributed in the hidden layer space, it can be fitted and converged quickly. As a result, no performance-related data enhancement was used in generator G_1, and random discarding was introduced in discriminator D_1 to avoid model overfitting. Using the standard generative adversarial network structure, the training process can also be formalized into a maximum and minimum optimization problem.

2.3. Construct target transition mapping from empirical hidden layer representation to data distribution

Different from the previous method of generating adversarial networks directly searching the target transition mapping from random white noise to data distribution in feature space, this paper trained a large-scale generating adversarial network model based on residual network to construct the target transition mapping from the empirical hidden layer representation generated by generator G_1 to data distribution. The formal expression of the process is as follows:

$$(\mu, \Omega) \overset{g_\zeta}{\rightarrow} (\nu_{gt}, \chi) \tag{3}$$

$g_\zeta$is parameterized by neural network. In this phase, as with the standard generative adversarial network, the generator and discriminator are constructed around two loss functions. Where, generator $G_2(u)$ maps the hidden layer representation to the real sample ν_{gt} and discriminator D_2 (x) determines whether the current input sample is derived from the generation distribution generated by generator $G_2(u)$. Based on the principles of game theory, the generator G2(u) and discriminator D2 (x) are trained alternately to meet the confrontation between them. The Settings are exactly as they were originally set, except that the hidden space variable of the input generator G2(u) is replaced by the output of generator G_1. This is also from the side the surface verifies the hypothesis in this paper that the hidden layer representation is a better hidden space variable than random white noise for constructing the target transition mapping to the data distribution. Experiments show that compared with the original generative adversarial network, this process has more stable training dynamics and can converge faster and capture more data patterns in the feature space[5].

3. Performance evaluation test

3.1. Experiment with synthetic data sets

The fast construction method proposed in this paper achieves sub-optimal pattern coverage results and optimal generation quality. The generated samples are shown in Figure 2. The left and right pictures show the generated samples of models on MNIST and Stacked-MNIST data sets respectively. In both data sets, the method proposed in this paper can capture all data patterns, and there is almost no pattern mixing or poor quality sample generation. Table 1 reports the results of the proposed rapid build method compared with other efforts.

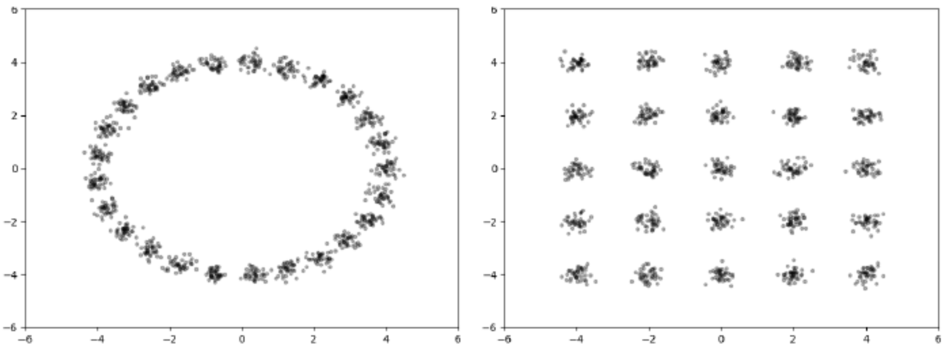

Fig. 2 Benchmarking MNIST and Stacked-MNIST data sets with generated sample examples

The experimental results show that no matter what kind of benchmark model is used, the fast construction method proposed in this paper can not only capture more data patterns, but also significantly improve the generation quality of generated samples. In fact, when the target loss function provided by CTGAN is used to generate the adversarial network, the results close to the theoretical optimal can be obtained. It can be found that the fast training framework achieves optimal results on almost all indexes, especially the two indexes of high quality sample ratio and reverse KL divergence. The fast build method proposed in this paper does not add any extra computational cost, and all the performance improvements come from the characteristic priors[6].

Tab. 1 Generate experimental comparison results with synthetic data sets under baseline Settings

	2D-ring			2D-grid		
	Capture mode number	Sample ratio	Inverse KL divergence	Capture mode number	Sample ratio	Inverse KL divergence
GAN	19.9±0.5	95.3	0.45	17.3±0.8	94.8	0.70
ALI	21.2±0.3	97.4	0.37	24.1±0.3	95.6	0.14
MD	13.2±0.4	34.7	1.97	23.8±0.2	79.9	0.18
BEGAN	22.3±0.2	88.4	0.72	21.9±0.1	84.2	0.77
CTGAN	23.6±0.3	98.3	0.04	23.5±0.5	98.1	0.06
PacGAN	24.7±0.2	96.5	0.05	24.7±0.3	94.2	0.02
PresGAN	25.2±0.1	97.2	0.04	24.8±0.2	94.5	0.03
BourGAN	24.8±0.2	97.9	0.02	24.9±0.1	95.9	0.02
SRGAN	24.8±0.2	97.5	0.02	24.7±0.1	98.4	0.01
AE-OT	24.9±0.1	99.8	0.01	24.9±0.3	99.5	0.09
AE-OT-GAN	24.8±0.2	99.9	0.01	24.8±0.2	99.7	0.01

3.2. MNIST and Stacked-MNIST data set experiment

After quantitative evaluation of the rapid construction method proposed in this paper, real data sets of different sizes are used to evaluate the performance of the method on real images. Firstly, the relatively simple MNIST and Stacked-MNIST data sets were used to evaluate the generation against network generation performance based on the fast build method. On the MIVIST dataset, all models were trained using only 1,000 randomly collected handwritten images and no data-enhancement methods related to generation performance were used. And in the Stacked-MNIST data set, the training data came from 128,000 samples obtained by random sampling and the corresponding training batch size was 64. After randomly sampling 26,000 generated samples of the Stacked-MNIST data set with the proposed rapid construction method, the proposed method obtained 0.05% 0.008 empirical KL divergence, which was significantly improved compared with the benchmark model. Experimental results on two data sets show that the rapidly constructed framework can fully capture all data patterns in the data distribution, and generate samples with clear and sharp contrast between foreground and background. FIG. 3 shows the results of the proposed methods in this paper on MNIST and Stacked-MNIST data sets, with little pattern mixing or poor sample quality generation.

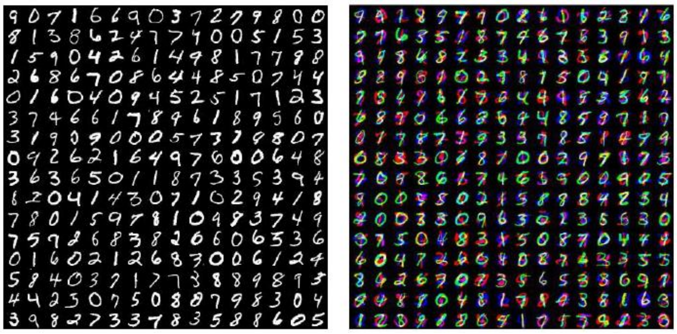

Fig. 3 The quick-build framework on the MNIST and Stacked-MNIST data sets

3.3. CelebA and CeleBA-HQ data set experiments

CelebA and CeleBA-HQ datasets are also used to further evaluate the effectiveness of the proposed generative adversarial network rapid construction method on large-scale real datasets. Complete training set images were used on both CelebA and CeleBA-HQ data sets without adding any additional data enhancement methods related to generation performance. The difference is that the face image in CelebA is uniformly converted to a pixel value of 128 x 128 while the pixel value in Ce1bA-HQ is 256 x 256 when the image data is processed. The baseline loss function used on the two data sets was BEGAN, and the corresponding Frechet Inception Distance scores were 14.7 and 6.8, respectively. FIG. 4 and FIG. 5 show the results generated by the rapid build method on the CelebA data set and the CeleBA-HQ data set respectively[7].

Fig. 4 OThe result sampling of the adversarial network fast construction framework on the CelebA dataset is generated

The target results include a variety of different facial features, such as men and women, straight and curly hair, long and short hair, and different hair colors, with rich patterns. The results show that the rapid construction method captures complex data patterns in the data space and fits them well. People with different skin colors, decorations and appearance can be realistically depicted. In addition, thanks to the unified HD images in the data set, the results show that the rapid construction method of generative adversarial network proposed in this paper can depict the portrait features of the characters very vividly on the CelebA-HQ data set, with excellent generation quality, which is difficult to distinguish compared with the real image.

Fig. 5 The result sampling of the adversarial network rapid construction framework on the CelebA-HQ dataset was generated

In the above experiments,due to the low dimension of the hidden layer representation,the training of the autoencoder and the generation of the empirical hidden layer representation requires only a very low computational cost.The experimental results fully show that the fast construction method can stably generate the training dynamics of the adversarial network and capture more data patterns in the feature space with almost the same computational cost as the standard model.

4. Conclusion

In this paper, we propose a rapid construction framework for generative adversarial networks based on hidden layer representation, which uses empirical hidden layer representation instead of random white noise as the input of generator. In this process, the complete generation process is divided into three stages: realizing data embedding in the hidden layer space by using self-coding, constructing target transfer mapping from random white noise to representation of experience hidden layer, and constructing target transfer mapping from representation of experience hidden layer to data distribution. Finally, the performance of the proposed rapid construction framework is verified by large-scale experiments, and the effectiveness of the rapid construction method is proved.

References

[1] Rikiya Y, Mizuho N, Kinh R, et al. Convolutional neural networks: An overview and application in radiology[J]. Insights Into Imaging, 2018, 9(4): 611-629.

[2] Dongsheng A, Yang G, Min Z, et al. Ae-ot-gan: Training gans from data specific latent distribution[C]. European Conference on Computer Vision, 2020, 548-564.

[3] Metz L, Poole B, Pfau D, et al. Unrolled generative adversarial networks[C]. 5th International Conference on Learning Representations, Toulon, France, 2017, 1-1.

[4] Martin A, Soumith C, Leon B. Wasserstein generative adversarial networks [C] . International con-ference on machine learning, Toulon, France, 2017, 214-223.

[5] Guojun Q.Loss-sensitive generative adversarial networks on lipschitz densities[J]. International Journal of Computer Vision, 2020, 128(5): 1118-1140.

[6] Zihang D, Zhilin Y, Fan Y, et al. Good semi-supervised learning that requires a bad gan[J]. Advances in Neural Information Processing Systems, 2017, 30: 1-1.

[7] Jeff D, Karen S. Large scale adversarial representation learning[J]. Advances in Neural Informa-tion Processing Systems, 2019, 32: 1-1.