Research and Analysis of DCGAN in Different Application Fields

Chongyue Liu

School of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing Jiangsu, 211816, China
202321168012@njtech.edu.cn

Abstract. Deep convolutional Generative Adversarial Network (DCGAN) is an important type of generative adversarial network. It integrates convolutional neural network (CNN) into the adversarial framework and performs well in image generation and data augmentation. This paper reviews and summarizes the basic technical points of DCGAN, including its network structure and key optimization methods. The application fields of DCGAN are mainly introduced: industrial defect detection (bearings, fabrics), agricultural pesticide residue detection and plant disease identification, as well as medical diagnostic image synthesis. Its value lies in solving the problems of small samples and the imbalance of datasets. DCGAN is increasingly used in these application fields due to its outstanding performance. Despite the challenges such as unstable training, model collapse and ethical risks that DCGAN faces, it still has a promising future. In the future, DCGAN will break data bottlenecks, drive industries to shift from experience-driven to data-driven, and become a core engine for multi-field intelligent upgrading.

Keywords: Deep Convolution Generative Adversarial Networks, Convolutional Neural Network, Image Generation.

1 Introduction

Deep learning is a highly dynamic research direction in the sphere of machine learning, which has not yet gained pace of development over the last several years. It can learn the complex patterns and feature representations among big data by building multi-layer neural networks, and has shown to perform excellently in many areas of its application including image recognition, speech recognition and natural language processing. As the technology of deep learning is increasingly becoming more mature, the range of its application is also becoming wider, which offers new ideas and approaches to resolve an ever-greater number of practical tasks.

Goodfellow et al. were the first to propose the use of Generative Adversarial Networks (GAN), which is a significant innovation in the area of deep learning. The main concept of GAN is to figure out how to learn the characteristics of data distribution by playing an adversarial game between the Generator and the Discriminator, thus, creating realistic samples. In this type of adversarial process, the aim of the Generator

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_94

is to produce the samples, which are the closest to the real data and can mislead the Discriminator; the aim of the Discriminator is to differentiate between the real samples and the fake ones produced by the Generator. Such an adversarial mechanism of training allows GAN to identify the sophisticated distribution of data and create samples with a high level of realism.

A variant of GAN is the deep convolutional Generative Adversarial networks (DCGAN). DCGAN has contributed to the primary development of convolutional neural networks (CNNS) into the generator and discriminator of GAN. It makes use of the strong feature extraction properties of CNN effectively improving the quality and diversity of the formed samples. CNNS have certain intrinsic strengths of handling image data in contrast with traditional fully connected neural networks. It can automatically obtain the local features and spatial structure data of images in convolutional layers, and pooling layers, and the main characteristics of images are more effectively captured [1]. DCGAN plays an important role in the unsupervised learning and the image-generative skills. It not only has the capability to produce high-quality images, it can also learn the latent feature representations of data which can be given useful information in a follow-up task.

Theoretically speaking, comprehensive studies regarding the use of DCGAN in other spheres of work are favorable to understanding and developing the theoretical framework of the GAN even more. Through performance studies of DCGAN with various data distributions and task models, issues and limitations of the model structure, training algorithm, loss function, etc can be identified and hence the theoretical underpinning of GAN models become improved and optimized.

In the practical sense, the use of DCGAN in manufacturing, agricultural, and medicine industry has given new ideas and approaches in which the sectors have been developed. As an example, online detection of the method to detect defects in small industrial components, make sure that the safety and functionality of aerospace, precision machine and instrumentation are ensured [2], it is possible at the level of the industrial sector; on-line detection of tomato leaf disease is possible to maintain the agricultural economy [3, 4].

This article seeks to conceptualize the present-day implementation of the DCGAN in manufacturing, healthcare, and agriculture sectors, by critically examining and summarizing the literature in the field. Through sorting and appropriate literature study, we may have a comprehensive knowledge of the definite application techniques, successes and drawbacks of the use of DCGAN in different areas, generalize its benefits and disadvantages and become the reference basis of the overall use and enhancement of DCGAN to more areas.

2 Technical Fundamental Analysis of DCGAN

2.1 Fundamental principles and structure of GAN

The main idea of GAN has been put on adversarial training between the generator and the discriminator. The primary objective of the generator is to take as input a random noise distribution, usually a random low dimensional vector, and process the input via a sequence of neural network layers into samples that appear to have the same distribution as actual data. The discriminator in its turn will determine the values of the input samples and decide between the real and fake ones belonging to the real dataset and those belonging to the fake one which are being created by the generator respectively.

The discriminator and the generator will face each other bitterly during the training process. The generator aims at producing more realistic samples to fool the discriminator and make it misclassify the produced samples to be actual ones; and the discriminator is constantly enhancing its capacity to detect, meaning to correctly resolve the produced samples. It can be considered as a minimax game problem, and its objective function is considered to be:

$$\min_{G}\max_{D} V(D,G) = \mathbb{E}_{x\sim p_{\mathrm{data}}(x)}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))] \quad (1)$$

Where, $p_{data}(x)$ is the distribution of actual data, $p_z(z)$ is the distribution of noise of the generator input, $D(x)$ is the distribution of the probability that the discriminator will classify the real sample x as real, and $D(G(z))$ is the probability that the discriminator will classify the $G(z)$ sample the generator creates on noise z as a real sample.

2.2 DCGAN Improvement and Network Structure

Structural Interpretation. DCGAN possesses some specific features in the networks design. Its generator and discriminator are based on CNN architecture which uses to the fullest potential the efficient powers of convolutional neural networks in image processing and extracting its features.

It is the job of the generator to transform the low-dimensional random noise into high-resolution images. It is typically made of a stack of transposed convolutional layers (also called deconvolution layers). The imprinted convolutional layer eventually enhances the picture image resolution through up-sampling of the input feature mapping.

The role of the discriminator is to detect which image is input then decide whether this is a real or fake image created by the generator. It will be a sequence of convolutional layers, will down-sample the input image by acting on it with convolution operations, will extract in stages the image features, and the roles of the image will be reduced.

Key improvements. Regarding the network structure, DCGAN employed convolutional layers as opposed to the fully connected layers in classic GANs. With a classic single fully connected neural network, the size of the network in terms of the number of network layers grows at an exponential rate, the number of network parameters grows exponentially as well, resulting in an increase in both the computational load and the likelihood of overfitting. Nonetheless, through convolutional layers, the displayed parameters are greatly shrunk by weight sharing, which enhances the computational efficiency.

Another significant enhancement of DCGAN is the introduction of the Batch Normalization technology. This is due to the fact that Batch Normalization balances the input of every layer, and it standardizes the average of the input to 0 and the standard to 1. It may be helpful to make the network training procedure more predictable, allow the model to learn at a larger learning rate, and faster convergence rate, as well as, be less susceptible to parameter changes.

2.3 DCGAN Optimization Strategy

In the process of training DCGAN, when you wish to increase the effect of training and the performance of the model, you should recognize an appropriate optimization algorithm. Adam optimizer is an optimization algorithm that is based on the merits of the Adagrad and RMSprop algorithm. It is able to dynamic control the learning rate, assigning various learning rates to various parameters, thus attaining good performance in the training process. The Adam optimizer adapts the learning rate of every parameter by estimating the first-order and second-order moment of the gradient, and in such a way, the model is able to converge almost quickly during the initial phase of the training and alternatively update the parameters more consistently in the latter phase of training.

The hyperparameters are highly critical to the training output. Among the hyperparameters, there is the learning rate. In the event of excessive learning rate, the model can miss out on the optimum solution throughout training hence failing to converge. In case the learning rate is too slow, the learning model will move very slowly and will consume a lot of training time and computing power. As a rule, the relatively high learning rate may be tried in the first place in order to see what the training process of the model will be like. The learning rate may be slowly lowered in case it has been discovered that the model is unstable or the model failed to converge. Another notable hyper parameter is the batch size which means the size of the number of samples that are fed into the model during each training period. Bigger batch processing will be able to utilize additional sample data to update the parameters and thus enhance the effectiveness and stability of the training, but will also consume more memory. Training with small batch sizes can be done with small memory, although often they can cause more instability. As such, the batch size must be proportionately changed with regards to hardware resources and training effects.

3 Applications to DCGAN in Multiple Areas

3.1 Industrial field

Firstly, bearings in industrial equipment may guarantee the safe and effective performance of equipment. Thus, their faults are difficult to identify and diagnose, and there is an acute need to do so. A new SA-SN-DCGAN model was developed by Zhong et al. to generate synthetic images in case of data scarcity of deep networks [5]. SA is a variation of the attention mechanism that is integrated into the generator, and analyzes features, imposing weights and choosing sensitive information, thus, enhancing the quality of the image. SN is a normalization weight method allowing to fit the Lipschitz constraint, and can be applied to stabilize the training process.

Secondly, fabric defect identification is a challenging topic owing to the variety of fabric defect and absence of defect samples. Wei et al. suggested a DCGAN based multi-stage unsupervised fabric defect detection approach [6]. Such a procedure comprises three phases (GAN training, encoder training, and classifier training). The initial two steps make the model rebuild the images in test. The residual map that represents the defect can be obtained when we subtract the reconstructed image with the original image. They have furthermore added a classifier training step in order to generate likelihood maps of the test images. Lastly, they combined the residual map with the likelihood map, and also made further threshold segmentation to the fused residual map. The results of the experiment proved the effectiveness of such approach in terms of F-score index.

Thirdly, to address the issue of limit sample in detecting a defect made on the surface coating of plastic parts, an enhanced algorithm of DCGAN, an algorithm that is grounded on the DCGAN is presented by Yin et al [7]. The algorithm does not alter the computational effort of the network, but rather makes the generated features richer and improves the existing activation operation, which subsequently reduces the gradient disappearance or explosion occurring during the training of the network. Subsequently, adds the deep convolutional residual blocks to the current generator architecture to generate rather high-resolution generated images, compared with GAN and DCGAN. The findings indicate that the improved DCGAN is more improved in the effect of image sample generation of painting surfaces of plastic parts significantly.

3.2 Agricultural field

Firstly, individuals are very much concerned with pesticide residues in Hami melons which is attributed to the overuse of pesticides. Hence, a quick and non-destructive means of pesticide residue level detection of Hami melons is urgently required to be proposed. In this study, Tan et al. made use of short-wave infrared hyperspectral imaging (SWIR-HSI) to determine the presence of pesticides residues on Hami melons [8]. They suggest a data augmentation technique, i.e., an enhanced DCGAN to enlarge the spectral data of Hami melons with various pesticide residue. Decision Tree (DT), Random Forest (RF) and Support Vector Machine (SVM) were used to test the effectiveness of the improved DCGAN. It has been found that the accuracy rates of them have been improved by 13.13, 7.50 and 11.25 respectively. These study findings also suggest that the hybrid of short-wave infrared hyperspectral imaging and a data augmentation regime rooted in a more advanced deep convolutional generative adversarial network is highly promising in pesticides residue detection in Hami melons.

Secondly, agriculture development can be impacted by the pest attacks on crops and plants. Automatic detection through image processing methods can give quick and precise results as compared to farmers or other professionals who may use the naked eye to observe the plants and then diagnose the diseases. The novel model of recognizing plant diseases is presented by Mahamod et al., as it is centered on leaf image classification, it uses DCGAN, and a classifier, which consists of a multi-layer perceptron (MLP) neural network, trained using the pseudo-inverse learning autoencoder (PILAE) algorithm [9]. The deep convolutional generative adversarial networks address two tasks: (1) synthesizing images based on a small number of categories to overcome the imbalance of the data set; (2) extracting deep features of the whole images in the data set.

3.3 Medical field

Firstly, breast cancer imaging is the key of detecting the disease as fast as possible and accurately as well. Nonetheless, the unavailability of annotated data of mammograms is a significant issue when constructing deep learning models. Shah et al. came up with an innovative approach which overcomes the limitation of data with the application of DCGAN [10]. The primary goal is to produce artificial mammograms, recreate the natural patterns that are typical of actual data as well as to improve the dataset. The experimental outcomes demonstrate that the given synthesizing approach can reproduce various views of the breast in an accurate way.

Secondly, white blood cells do tell a lot regarding the health condition during a physical examination. Nevertheless, currently the available white blood cell datasets in the medical sector contain limited volumes of data with an uneven distribution of classes. To equalize the number of datasets of white blood cells, Hartanto et al. suggested a way of synthesizing synthetic images with classical data augmentation and DCGAN [11]. After that, a classification model is selected, the ResNet50 model, which is capable of reaching great accuracy when classifying medical images, is trained on balanced datasets. Performance indicators used as a measure of classification include accuracy, precision, recall rate and Fl-score. The findings indicate that the accuracy rate of this model is 82.5%.

Thirdly, one of the crucial methods of diagnosing disorders in the breathing issues is the chest X-ray imaging and the creation and evaluation of AI-based medical hardware can afford even more numerous, artificially generated data. Akhil et al provide a new technique which generates synthetic chest Xray images that simulate real radiography data with a high level of accuracy using DCGAN, a type of generative model architecture [12]. Their approach is to rely on a solution that involves a two-step process, a discriminator network as well as a generator. The discriminator is developed to distinguish between the authentic and the false images whereas the generator is trained to draw the correct images of the chest. With rigorous training program these two networks keep on upgrading the quality of the images they generate. The synthetic X-ray images can be considered valuable resources in elaborating on small datasets, which will enhance the development of robust and accurate deep-learn algorithms in pulmonary disease detection as well as classifications.

4 Challenges and Prospects of the DCGAN

The core challenges faced by DCGAN run through the entire process of training, deployment, and ethics, rooted in the inherent characteristics of GANs and the limitations of convolutional operations. Technically, the "minimax game" of GANs makes it difficult to balance the generator and discriminator. The cross-entropy loss is prone to causing gradient vanishing and "mode collapse". The mode collapse rate on the CIFAR-10 dataset exceeds 35%. At the same time, the number of parameters in its convolutional operations reaches tens of millions. The delay in generating 256×256 images in edge scenarios exceeds 500ms. In applications, the time-consuming annotation of medical 3D CT and the large differences in agricultural samples lead to data scarcity. The deviation between training and scene data will also reduce the performance of downstream models. Ethically, the model easily memorizes details of training data, leading to privacy leakage. The improvement of generation quality brings risks of abuse such as fake quality inspection certificates and pathological reports, impacting the industry's trust system.

The development prospects of DCGAN focus on three directions: algorithm optimization, application expansion, and ethical norms. In terms of technological innovation, the Wasserstein distance loss of WGAN reduces the mode collapse rate to below 12%. Spectral normalization increases the SSIM of medical image synthesis by 18%. The progressive architecture improves the rationality of 512×512 image details in the LSUN dataset by 30%. In terms of application expansion, the model fused with Transformer generates MRI to CT images with a DSC of 0.89. Combining with reinforcement learning can optimize the structure of metamaterial chips. After migration through domain adaptation technology, the accuracy of wheat grain recognition reaches 92%. The "cloud training - edge inference" model also solves the problem of edge deployment. In ethical protection, the 3PC framework and federated learning ensure data privacy. Digital watermarking technology realizes the traceability of generated content. Industry norms also clarify the application boundaries. In the future, it will integrate with quantum computing to achieve a leap in performance, strengthen multimodal generation capabilities, form a virtuous cycle through improved norms, promote the transformation of multiple fields from "experience-driven" to "data-driven" through data augmentation, and become the core engine for industrial intelligent upgrading.

5 Conclusion

This article describes the fundamental principles of DCGAN and its practical applications in industry, agriculture, and healthcare, while also analyzing its challenges and prospects. DCGAN innovatively combines the adversarial mechanism of CNN with that of GAN, and is supplemented by batch normalization and optimized activation functions, making it more proficient in high-quality image generation and data augmentation.

In practical applications, DCGAN has proposed good solutions to the problems existing in real life. In the industrial field, it can conduct precise defect detection on bearings, fabrics and plastic components. In the agricultural field, through the expansion of spectral and image data, it can be used for pesticide residue detection and plant disease identification. In the medical field, it can synthesize breast X-rays, white blood cell images and chest X-rays, alleviating the problem of data scarcity in disease diagnosis. These applications have confirmed the potential of DCGAN in promoting technological progress in multiple industries.

However, DCGAN still faces some challenges, including unstable training and mode crashes. In the future, we should focus on algorithm optimization, such as integrating Wasserstein Distance and attention mechanisms, expanding multimodal and edge deployment scenarios, and strictly regulating people's ethics. I believe that with the continuous innovation and standardized application of technology, DCGAN is expected to further break through the data bottleneck, drive various industries to transform from experience-driven to data-driven, and become the core engine for intelligent upgrades in multiple fields.

References

1. Fang, W., Zhang, F., Sheng, V. S., et al.: A Method for Improving CNN-Based Image Recognition Using DCGAN. Computers, Materials & Continua, 57(1). (2018)
2. Gao, H., Zhang, Y., Lv, W., et al.: A deep convolutional generative adversarial networks-based method for defect detection in small sample industrial parts images. Applied Sciences, 12(13): 6569. (2022)
3. Wu, Q., Chen, Y., Meng, J.: DCGAN-based data augmentation for tomato leaf disease identification. IEEE access, 8: 98716-98728. (2020)
4. Devi, Y. S., Kumar, S. P.: DR-DCGAN: A deep convolutional generative adversarial network (DC-GAN) for diabetic retinopathy image synthesis. Webology (ISSN: 1735-188X), 19(2). (2022)
5. Zhong, H., Yu, S., Trinh, H., et al.: Fine-tuning transfer learning based on DCGAN integrated with self-attention and spectral normalization for bearing fault diagnosis. Measurement, 210: 112421. (2023)
6. Wei, C., Liang, J., Liu, H., et al.: Multi-stage unsupervised fabric defect detection based on DCGAN. The Visual Computer, 39(12): 6655-6671. (2023)
7. Yin, X., Hou, B., Huang, Y., et al.: Image enhancement method based on improved DCGAN for limit sample//2022 14th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). IEEE, 376-379. (2022)
8. Tan, H., Hu, Y., Ma, B., et al.: An improved DCGAN model: Data augmentation of hyperspectral image for identification pesticide residues of Hami melon. Food Control, 157: 110168. (2024)
9. Mahmoud, M. A. B., Guo, P., Wang, K.: Pseudoinverse learning autoencoder with DCGAN for plant diseases classification. Multimedia Tools and Applications, 79(35): 26245-26263. (2020)
10. Shah, D., Ullah Khan, M. A., Abrar, M.: Reliable breast cancer diagnosis with deep learning: DCGAN-driven mammogram synthesis and validity assessment. Applied Computational Intelligence and Soft Computing, (1): 1122109. (2024)
11. Hartanto, C. A., Kurniawan, S., Arianto, D., et al.: Dcgan-generated synthetic images effect on white blood cell classification//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 1077(1): 012033. (2021)
12. Akhil, M. C. S., Sharma, B. S. S., Kodipalli, A., et al.: Medical image synthesis using DCGAN for chest X-ray images//2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS). IEEE, 1: 1-8. (2024)