Proceedings of the 4th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/52/20241115

Data augmentation-based enhanced fingerprint recognition using deep convolutional generative adversarial network and diffusion models

Yukai Liu

Department of Computer Science and Technology, Xidian University, Xi'an, China

yukailiu@ucsb.edu

Abstract. The progress of fingerprint recognition applications encounters substantial hurdles due to privacy and security concerns, leading to limited fingerprint data availability and stringent data quality requirements. This article endeavors to tackle the challenges of data scarcity and data quality in fingerprint recognition by implementing data augmentation techniques. Specifically, this research employed two state-of-the-art generative models in the domain of deep learning, namely Deep Convolutional Generative Adversarial Network (DCGAN) and the Diffusion model, for fingerprint data augmentation. Generative Adversarial Network (GAN), as a popular generative model, effectively captures the features of sample images and learns the diversity of the sample images, thereby generating realistic and diverse images. DCGAN, as a variant model of traditional GAN, inherits the advantages of GAN while alleviating issues such as blurry images and mode collapse, resulting in improved performance. On the other hand, Diffusion, as one of the most popular generative models in recent years, exhibits outstanding image generation capabilities and surpasses traditional GAN in some image generation tasks. The experimental results demonstrate that both DCGAN and Diffusion can generate clear, high-quality fingerprint images, fulfilling the requirements of fingerprint data augmentation. Furthermore, through the comparison between DCGAN and Diffusion, it is concluded that the quality of fingerprint images generated by DCGAN is superior to the results of Diffusion, and DCGAN exhibits higher efficiency in both training and generating images compared to Diffusion.

Keywords: Data Augmentation, Machine Learning, Image Generation, GAN.

1. Introduction

Image recognition is an essential undertaking within computer vision and image processing domains [1, 2]. Its objective is to classify the objects of interest in images or videos. Fingerprint recognition is one specific application of recognition. As a reliable tool for identity verification and security, fingerprint recognition relies on large-scale, high-quality fingerprint datasets for training. However, due to privacy concerns associated with collecting real fingerprints, existing datasets still suffer from limitations, such as data scarcity and imbalance. To overcome these challenges and enhance the accuracy and robustness of fingerprint recognition systems, data augmentation has emerged as a prominent direction, which involves reasonable transformations and expansions of training data to increase its diversity, richness, and the model's robustness.

To tackle these challenges and enhance the precision and resilience of fingerprint recognition systems, prior literature has introduced diverse techniques for data augmentation. Some commonly used models include traditional methods like geometric transformations and noise addition. These techniques aim to enhance the diversity and richness of the training data. In recent years, deep learning models, such as Generative Adversarial Networks (GANs) [3] and Variational Autoencoders (VAEs) [4], have also been employed for data augmentation. Complex data distributions can be learned by these models, allowing them to create synthetic samples with a high resemblance to real data. They offer the advantage of producing more realistic and diverse augmented data, capturing intricate patterns and variations in the images. Despite the progress made in data augmentation techniques, there are still limitations to be considered [5]. Taking the conventional generative model, GAN, as an example, mode collapse is a common issue in GANs. When the generator network fails to effectively learn the entire data distribution, it may fall into a mode collapse scenario, generating a small fraction of samples from the distribution of data. As a consequence, the generated images lack variety and depth. In addition, due to the complexity of the generator network and the optimization challenges during training, GANs may also suffer from mode blurring and distortion issues. However, for fingerprint data that requires high-quality images, image blurring and distortion are unacceptable. Therefore, it is imperative to improve the model or adopt superior generative models to enhance the quality, stability, and diversity of the generated images. Building upon this idea, a notable superior generative model worth mentioning is the Diffusion model [6]. The currently popular Diffusion model, as a top-notch model, has demonstrated excellent performance within multiple application fields, which encompass computer vision, waveform signal processing, time series modeling, and adversarial purification. Moreover, the Diffusion model surpasses GAN in various image generation tasks, capable of generating high-quality and diverse images. Additionally, the Diffusion model is closely related to research fields like robust learning, representation learning, and reinforcement learning.

In this regard, this study employed a variant of GAN, namely Deep Convolutional Generative Adversarial Network (DCGAN) [7], along with the advanced generative model, Diffusion, with the aim of data augmentation. Compared to traditional GANs, DCGAN incorporates key design elements that enhance its performance. These elements include the use of batch normalization layers, convolutional and deconvolutional layers, the Leaky Rectified Linear Unit (LeakyReLU) activation function, and the Adam optimizer. As a result, DCGAN can capture richer and more diverse image features, generate high-resolution images, and better preserve image clarity and authenticity, thereby mitigating the issues of blurring and distortion encountered in traditional GANs. On the different side, Diffusion, an innovative state-of-the-art deep generative model, employs the introduction of continuous Gaussian noise to intentionally distort the training data. Subsequently, it undertakes the task of learning to recover the original data by skilfully reversing this noise process [6]. Diffusion exhibits remarkable image generation capabilities. Therefore, Diffusion is also chosen as the approach used for generating fingerprint data.

2. Method

2.1. Dataset preparation

In this experiment, the Fingerprint Image Dataset from FVC2000_DB4_B was employed. The Fingerprint Dataset for FVC2000_DB4_B is a collection of fingerprint images used for fingerprint recognition research, but it can also be used for data augmentation tasks. Within this dataset, there are 800 high-quality fingerprint images with a size of 160×160 pixels and a resolution of 500DPI, as shown in the sample images in Figure 1. For easy reference, the source of this dataset is also available [8].

In terms of the DCGAN and Diffusion models, this study employed slightly different data preprocessing methods. In the case of DCGAN, this study resized the images to 64×64 pixels for computational convenience. Following that, the images were subjected to normalization, where the pixel values were scaled from the range [0, 1] to [-1, 1], resulting in image data possessing zero mean and unit variance. The utilization of this normalization method plays a role in boosting the stability and convergence speed of the model. Regarding the Diffusion model, this study resized the images to 80×80 pixels and performed random cropping. This involved randomly scaling the images within a given scale range (0.8 to 1.0) and then randomly cropping them to the specified size. This technique enhances the diversity and generalization ability of the data. Finally, this study applied the same normalization process as in DCGAN.

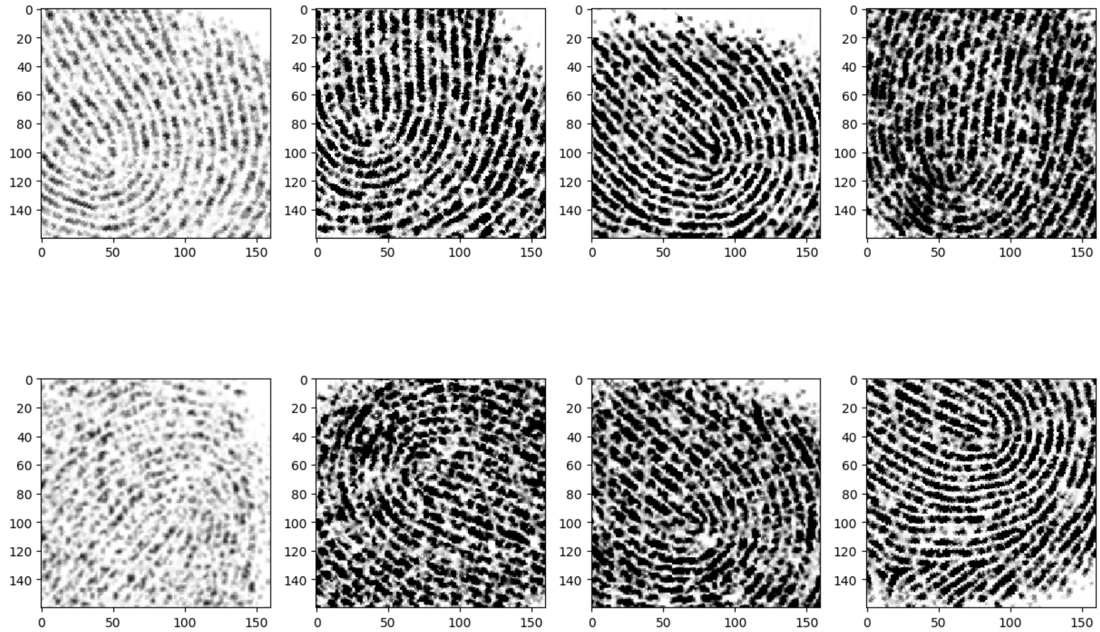


Figure 1. Sample images from the Fingerprint Image Dataset.

2.2. Generation model based on GAN and diffusion model

In this study, DCGAN and Diffusion were employed as methods for generating fingerprint images. DCGAN is a deep learning model based on the GAN framework. DCGAN incorporates several important improvements compared to traditional GANs. The main elements of DCGAN are the Generator and the Discriminator. The Generator receives random noise as input and converts it into synthetic image samples. By receiving real image samples and generated image samples from the Generator, the Discriminator performs classification to discern between real and synthetic images. DCGAN incorporates several improvements compared to traditional GANs. Firstly, the Generator network of DCGAN employs a structure with transpose convolutional layers, gradually upsampling and increasing the number of feature channels to map low-dimensional noise vectors into high-resolution images. The specific Generator structure is shown in Figure 2. Secondly, the Discriminator network utilizes a structure with convolutional layers, using convolution and pooling operations to map input images to a discriminative result representing the probability of being a real image. Finally, DCGAN utilizes LeakyReLU activation function and the Adam optimizer. LeakyReLU allows for the propagation of negative values, enhancing the non-linear expressive power of the network, while Adam optimizer facilitates faster convergence and improved training stability.

The diffusion model, which is influenced by non-equilibrium dynamics, is classified as a latent variable model [9]. Its goal is to generate samples from noise that closely resemble the real data distribution. The training process of the Diffusion model can be outlined as follows: Initially, a forward diffusion process is established, acting as a fixed Markov chain that progressively converts real or generated data into pure noise. This conversion is achieved by iteratively introducing Gaussian noise. This process consists of multiple diffusion steps, each with a noise intensity and a noise map. Subsequently, a reverse diffusion process is defined, which is a learnable neural network that gradually removes noise from pure noise and recovers the original data. This process also consists of multiple diffusion steps, each with a time step. Then, using the forward diffusion process, a training set is generated, including noise intensities, noise maps, and the images after adding noise. The neural network of the reverse diffusion process is trained using this training set to learn how to infer the added noise from the noisy images. Finally, with the trained neural network of the reverse diffusion process, the pure noise is gradually removed step by step to obtain a reasonable and normal image.

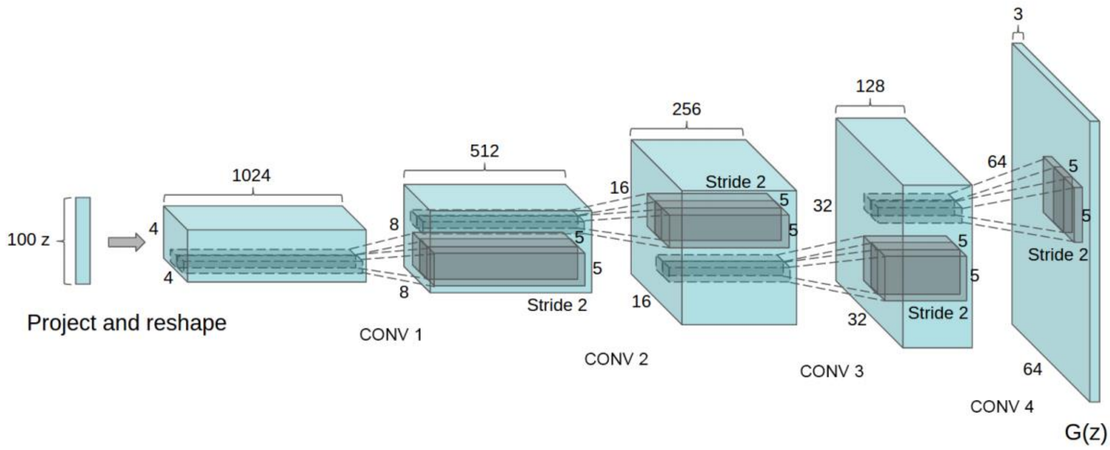


Figure 2. The architecture of the DCGAN generator [7].

2.3. Implementation details

In the specific implementation of DCGAN, the generator and discriminator loss functions employ cross-entropy loss. The optimization is performed using the Adam optimizer with a learning rate set to 0.0002. Training is carried out over 1,000 epochs with a batch size of 128.

When implementing Diffusion, the optimization process aims to maximize the variational lower bound of the maximum likelihood probability. The Adam optimizer is used for training with a learning rate of 0.00015. The model is trained with a batch size of 24 for 1000 epochs.

3. Results and discussion

The outcomes of fingerprint image generation using DCGAN and Diffusion are presented in Figure 3 and Figure 4, respectively. In general, the generated images by DCGAN and Diffusion clearly depict fingerprints. This demonstrates that both of these generative models are capable of effectively performing the data augmentation task and generating diverse high-quality images.

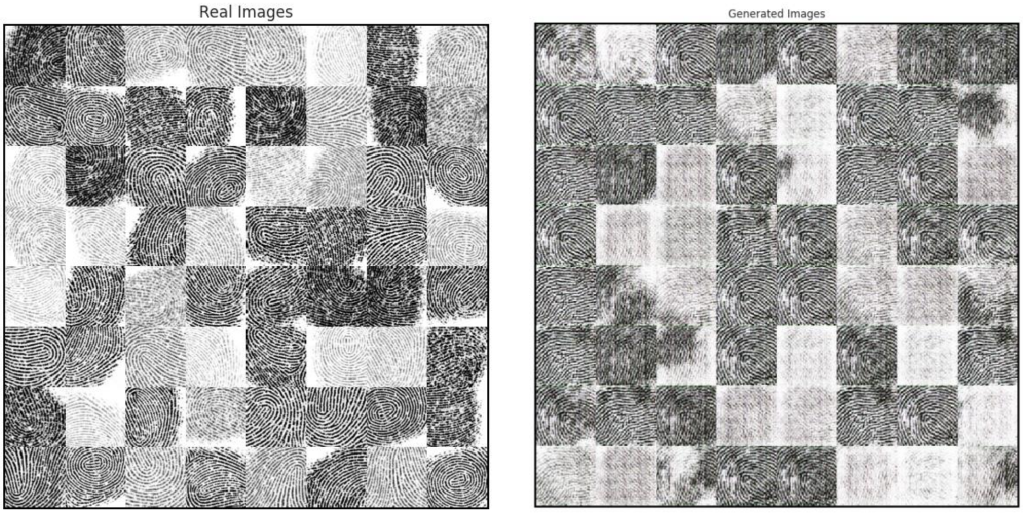
In the generated fingerprint images by DCGAN, some fingerprints have darker colors while others have lighter colors. Additionally, there are instances where certain images exhibit a division with half of the fingerprint having darker colors and the other half having lighter colors. This is because the training dataset consists of fingerprint samples with diverse color distributions, and the DCGAN model learned and reflected this diversity during image generation.

The generated fingerprint images using Diffusion are visualized in Figure 4. Comparing them with the sample images in Figure 1 and the images generated by DCGAN, it can be observed that the generated images by Diffusion only display a portion of the fingerprint. Furthermore, unlike DCGAN, most of the generated images by Diffusion do not have blank areas. This is due to the reason that Diffusion applies scaling during the preprocessing stage of the sample images, allowing the Diffusion model to better capture fingerprint features while eliminating the influence of blank areas at the edges of the sample images.

In the generated fingerprint images by Diffusion, there are also some issues with noise. The generated noise may be due to the model overfitting the noise present in the training dataset. To address this problem, further attempts can be made by using a larger training dataset to train the model or applying regularization techniques to reduce overfitting.

DOI: 10.54254/2755-2721/52/20241115

Comparing the Diffusion model with the DCGAN model reveals the following findings. 1) Compared to the images generated by Diffusion, the ones generated by DCGAN exhibit greater clarity and stability. 2) The iteration and sampling speed of the Diffusion model is slower, resulting in DCGAN being more efficient than Diffusion in both training and image generation processes. However, there have been experiments aimed at accelerating the generation process of the Diffusion model to mitigate its shortcomings [10].



(a) **(b)**

Figure 3. (a) Real fingerprint images. (b) DCGAN-generated fingerprint images.

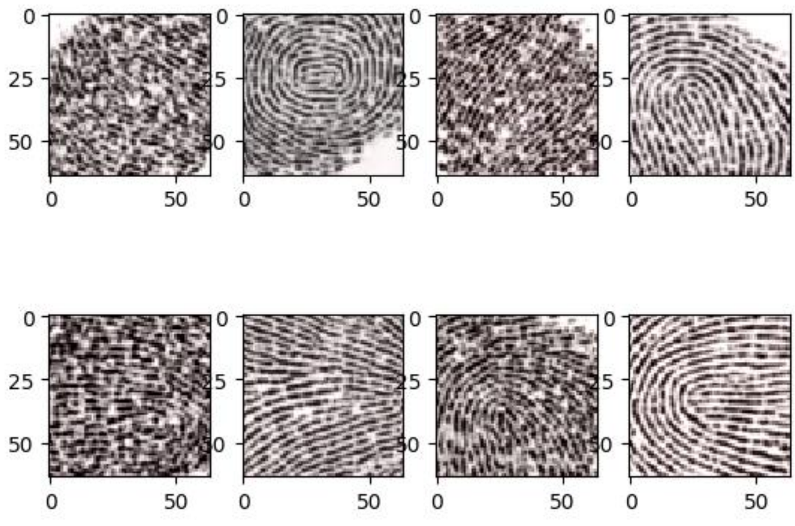


Figure 4. Diffusion-generated fingerprint images.

4. Conclusion

In this work, the generative models DCGAN and Diffusion were proposed for data augmentation tasks on a fingerprint dataset. The fingerprint images generated by both of these models exhibit superior quality compared to traditional GANs. Specifically, DCGAN generated images that captured the diversity of fingerprint images in the dataset, producing a variety of distinct images. Meanwhile, Diffusion captured the local features of fingerprints, generating images that alleviated the impact of

DOI: 10.54254/2755-2721/52/20241115

blank areas in the sample images. Additionally, a comparison between DCGAN and Diffusion reveals that DCGAN outperforms Diffusion in terms of image quality and efficiency. Through the comparison with DCGAN, the shortcomings of the Diffusion model were also exposed. The images generated by Diffusion had more noise, and Diffusion required more time for both training the model and generating images compared to DCGAN. In the future, further experiments can focus on improving the Diffusion model, for example, by adopting stable diffusion to generate higher quality images.

References

[1] Li Z et al 2019 CLU-CNNs: Object detection for medical images Neurocomputing 350 53-59.

[2] Qiu Y et al 2022 Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control 72 103323.

[3] Goodfellow I Pouget-Abadie J Mirza M Xu B Warde-Farley D Ozair S and Bengio Y 2014 Generative adversarial nets Advances in neural information processing systems 27.

[4] Kingma D P and Max W 2013 Auto-encoding variational bayes arXiv preprint arXiv:1312.6114.

[5] Salimans T Goodfellow I Zaremba W Cheung V Radford A and Chen X 2016 Improved techniques for training gans Advances in neural information processing systems 29.

[6] Sohl-Dickstein J Weiss E Maheswaranathan N and Ganguli S 2015 Deep unsupervised learning using nonequilibrium thermodynamics In International conference on machine learning (pp 2256-2265) PMLR.

[7] Radford A Metz L and Chintala S 2015 Unsupervised representation learning with deep convolutional generative adversarial networks arXiv preprint arXiv:1511.06434.

[8] Kaggle 2020 Fingerprint Dataset for FVC2000_DB4_B https://www.kaggle.com/datasets/peace1019/fingerprint-dataset-for-fvc2000-db4-b.

[9] Ho J Ajay J and Pieter A 2020 Denoising diffusion probabilistic models Advances in neural information processing systems 33 6840-6851.

[10] Song J Meng C and Ermon S 2020 Denoising diffusion implicit models arXiv preprint arXiv:2010.02502.