International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 05 May 2024
Page No: 1348 - 1354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0186

Generating Human Face with Dcgan and Gan

Manimegala M [1], Gokulraj V[2], Karisni K[3], Manisha S[4]

[1]Associate Professor, Department of Computer Science &Engineering, Sri Shakthi Institute of Engineering & Technology, Chinniyampalayam, Coimbatore, Tamil Nadu, India.

[2,3,4]Student, Department of Computer Science &Engineering, Sri Shakthi Institute of Engineering &Technology, Chinniyampalayam, Coimbatore, Tamil Nadu, India.

Emails: *manimegalacse@siet.ac.in[1], gokulrajv21cse@srishakthi.ac.in[2], karisnik21cse@srishakthi.ac.in[3], manishas21cse@srishakthi.ac.in[4]*

Abstract

Generative Adversarial Networks (GANs) are prominent in unsupervised learning for their exceptional data-generation capabilities. GANs utilize backpropagation and a competitive process between a Generative Network (G) and a Discriminative Network (D). In this setup, G generates artificial images while D distinguishes real from artificial ones, enhancing G's ability to create realistic images. Deep Convolutional Generative Adversarial Networks (DCGAN) are particularly notable, using a convolutional architecture to produce high-quality human face images. This study trains DCGAN on the CelebFaces Attributes Dataset (CelebA), demonstrating its ability to generate human faces from unlabeled data and random noise. Evaluation is done quantitatively using the Structural Similarities Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) to assess image quality. Additionally, this abstract will compare the effectiveness of GANs and DCGANs in human face generation.

Keywords: *DCGANs: Deep Convolutional Generative Adversarial Networks; GANs: Generative Adversarial Networks; HFG: Human Face Generation; IQE: Image Quality Evaluation; UL: Unsupervised Learning.*

1. Introduction

GANs and DCGANs have transformed the landscape of generating human faces using artificial intelligence. Initially introduced in 2014, GANs employ an adversarial training process with a generator and discriminator to create lifelike facial images. However, conventional GANs often struggle with architectural limitations and training instability, leading to challenges in producing diverse and high-quality faces [1]. In contrast, DCGANs, introduced in 2015, utilize deep convolutional neural networks in both generator and discriminator components, allowing for better capture of intricate facial features and spatial relationships. This architectural improvement enhances feature extraction and spatial modeling, resulting in superior performance compared to traditional GANs. DCGANs exhibit enhanced training stability and convergence, enabling the generation of highly realistic and visually appealing human faces. Their advancements have positioned them as a cornerstone in generating top-tier face images for various applications, including facial recognition systems, virtual avatars, and artistic endeavors, underscoring the vast potential of artificial intelligence in creative image synthesis [2].

2. Literature Survey on Gan and Dcgan for Human Face Generation

2.1. Existing Dataset for Human Face Generation

2.1.1. Label Faces in the Wild (LFW) Dataset

The Labeled Faces in the Wild (LFW) dataset is a benchmark dataset in the field of face recognition, consisting of over 13,000 images of faces collected from the internet [3]. These images capture faces in various poses, lighting conditions, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 05 May 2024
Page No: 1348 - 1354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0186

backgrounds, reflecting the diversity of real-world scenarios. While originally intended for face recognition tasks, the LFW dataset has also been utilized for training GANs and DCGANs for human face generation, enabling researchers to create more realistic and diverse synthetic faces. **(Aouada, D., Cherenkova, k., et al, 2020;) [4]**.

2.1.2. Anime Faces Dataset

The Anime Faces dataset comprises images featuring anime-style characters' faces, characterized by exaggerated features and vibrant colors typical of anime and manga art. With thousands of images from various anime series and illustrations, it serves as a crucial resource for training and evaluating GANs and DCGANs for anime face-generation tasks. Pre-processing techniques, including noise reduction and color normalization, enhance image quality, while augmentation strategies increase dataset diversity. However, copyright issues surrounding character likenesses require careful consideration. Despite challenges, expanding and diversifying the dataset ethically is vital for advancing anime face generation research.). **(Cao, C., Chai, M., Woodford, O., & Luo, L. 2018;) [5].**

2.2. Pre-Processing Techniques for Image Enhancement

2.2.1. Noise Removal

Noise removal techniques, such as Gaussian blur or median filtering, are applied to input images to eliminate unwanted artifacts and enhance image clarity. By reducing noise, the visibility of facial features is improved, leading to more accurate and realistic generated faces. **(Doersch, C. 2016;) [6].**

2.2.2. Image Normalization

Image normalization involves standardizing the pixel values of input images to a common scale, ensuring consistent representation across the dataset. Normalization helps mitigate variations in lighting conditions and camera settings, resulting in more uniform training data for GANs and DCGANs **(Bardsley, J.M. 2008;) [7].**

2.2.3. Resizing

Resizing input images to a uniform size is essential for ensuring compatibility with the network architecture of GANs and DCGANs. Resized images facilitate efficient processing and parameter sharing across layers, improving the convergence speed and performance of the generative models. **(Fujimura, K., Yamauchi, T., & Yamaguchi, O. 2019;) [8].**

2.2.4. Histogram Equalization and Contrast Enhancement

Histogram equalization and contrast enhancement techniques adjust the distribution of pixel intensities in input images to enhance image contrast and detail. By improving the visibility of facial features, these techniques enable GANs and DCGANs to generate more visually appealing and lifelike human faces [9].

3. Gan and DCGAN for Face Detection Method

3.1. GAN-based Face Detection Techniques

3.1.1. Advarsarial Training for Face Detection

In GAN-based face detection, adversarial training involves training a generator network to synthesize realistic face images and a discriminator network to differentiate between real and synthetic faces. Through iterative refinement, adversarial training enhances the generator's capability to produce authentic-looking faces and the discriminator's proficiency in discerning fake ones, thus improving the overall accuracy and robustness of face detection systems **(Goodfellow, I., Bengio, Y., & Courville, A. (2016).; Hinton, G.,2012;) [10].**

3.1.2. Data Augmentation with GANs

Utilizing GANs for data augmentation in face detection involves generating synthetic face images with diverse variations in poses, expressions, and lighting conditions. By augmenting face detection datasets, GANs increase diversity, enabling more robust models trained on augmented data to perform effectively in real-world scenarios with varying environmental conditions.

3.1.3. GAN-Based Data Imputation

GANs are employed for data imputation in face detection tasks to fill in missing or occluded regions within facial images. By generating plausible in paintings of absent facial features, GAN-based data imputation methods enhance the completeness of input images, thereby improving the subsequent performance of face detection algorithms **(Ioffe, S., & Szegedy, C. 2015;) [11].**

3.2. DCGAN-Based Face Detection Techniques

3.2.1. Feature Extraction with DCGANs

DCGANs play a crucial role in extracting discriminative features from facial images for face detection tasks. Leveraging their deep convolutional architecture, DCGANs capture hierarchical representations of facial features, facilitating effective feature extraction for subsequent face detection algorithms.

3.2.2. DCGAN-Based Cascaded Detectors

Integration of DCGANs into cascaded face detection frameworks aims to enhance detection accuracy and robustness. By incorporating DCGANs at various stages of the detection pipeline, cascaded detectors effectively localize faces with diverse scales, orientations, and occlusions, thereby improving overall detection performance.

3.2.3. Transfer Learning with Pre-Trained DCGANs

Pre-trained DCGAN models, initially trained on extensive face image datasets, can be fine-tuned for face detection tasks using transfer learning. By leveraging the learned representations from DCGANs, transfer learning facilitates the efficient adaptation of pre-trained models to specific face detection domains, resulting in improved detection performance even with limited labeled data. **(Jaderberg, M., Simonyan, K., 2015;) [12].**

These techniques demonstrate the versatility and effectiveness of both GANs and DCGANs in advancing face detection capabilities, addressing various challenges, and improving the accuracy and robustness of face detection systems across different applications and scenarios in Figure 1.

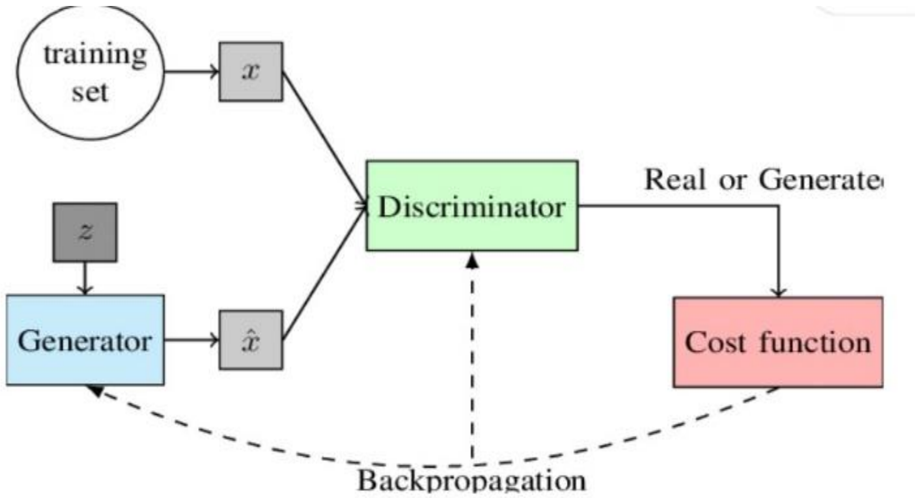

Figure 1 Generator and Discriminator

3.3. Loss Function of Simple GAN

In the paper that introduced GANs the generator aims to minimize the following function while the discriminator aims to maximize it:

- (min_{G} max_{D} V (D, G)) represents the optimization objective of the GAN.
- (V (D, G)) is the value function that the generator and discriminator aim to optimize.
- (E_ {x + p_{text{data}} (x)} [log D(x)]) is the expectation of the discriminator's log probability of assigning real data as real.
- (E_ {z + p_{z}(z)} [log (1 - D(G(z)))]) is the expectation of the discriminator's log probability of assigning fake data generated by the generator as fake.

The generator seeks to minimize this function, while the discriminator aims to maximize it. This adversarial setup fosters the competition between the generator and discriminator, leading to the generation of realistic data by the generator [14].

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 05 May 2024
Page No: 1348 - 1354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0186

$$\min_{\theta_g}\max_{\theta_d}\left[\mathbb{E}_{x\sim p_{data}}\underbrace{\log D_{\theta_d}(x)}_{\text{Discriminator output for real data x}}+\mathbb{E}_{z\sim p(z)}\log(1-\underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}})\right]$$

4. Result and Discuss

The loss function within the simple GAN framework epitomizes the adversarial training dynamic, where the generator and discriminator networks engage in a minimax game. Through this game, the generator refines its ability to generate synthetic data closely resembling real samples, while the discriminator hones its capacity to differentiate between authentic and synthetic data in Figure 2. **(Karras, T., Laine, S., Aila, T. 2019, Liu, J., Yan, Z., & Ouyang, W. 2020;) [13].**

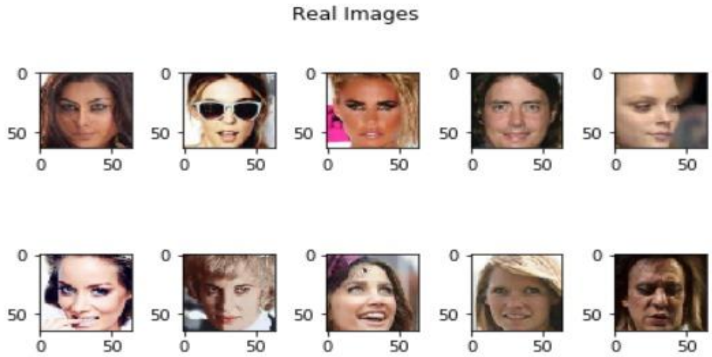

Figure 2 CELEB_DataSet Used in DCGAN and GAN

This minimax objective function, represents the expected discriminator log-probability of accurately classifying real data in Figure 3.

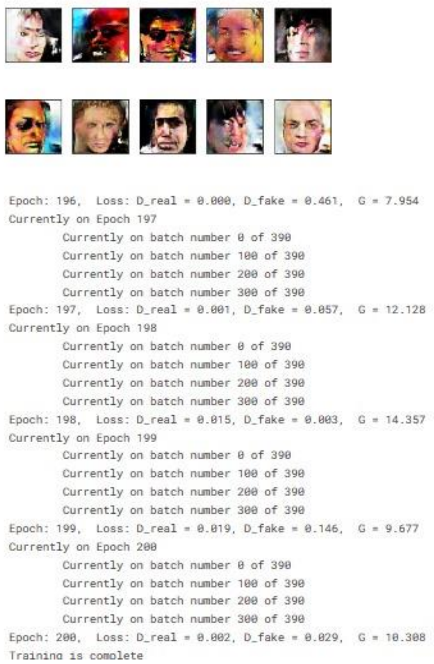

Figure 3 Output Visual Station of DCGAN

Here, the generator endeavors to minimize this term by crafting samples that deceive the discriminator into assigning high probabilities to synthetic data.

Conversely, the second term, denotes the expected discriminator log-probability of inaccurately classifying generated samples as real. The discriminator aims to maximize this term by effectively identifying generated samples as fake Figure 4&5.

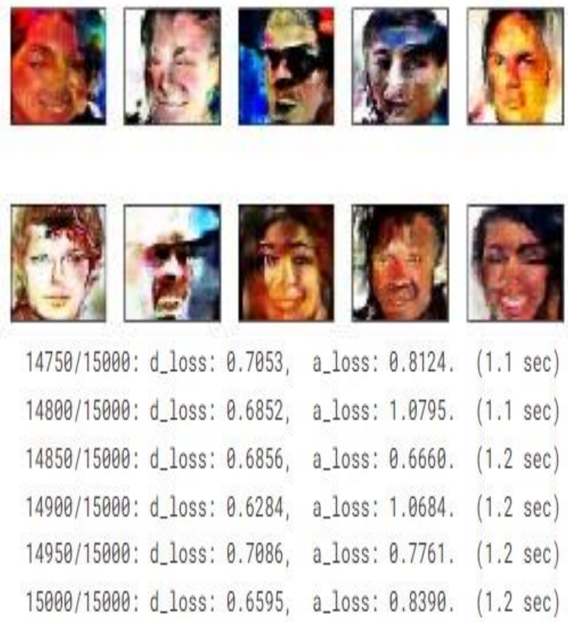

Figure 4 Output Visual station of GAN

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 05 May 2024
Page No: 1348 - 1354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0186

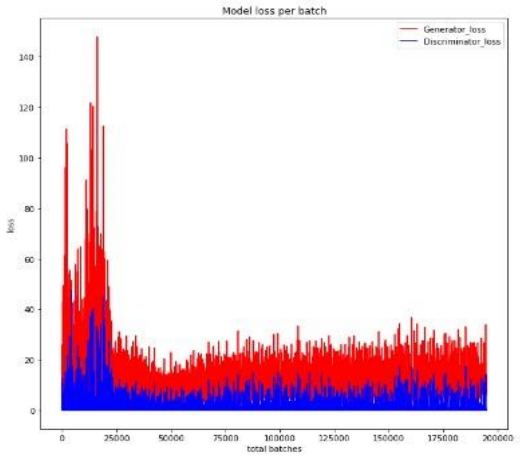

Figure 5 **Loss Curve DCGAN**

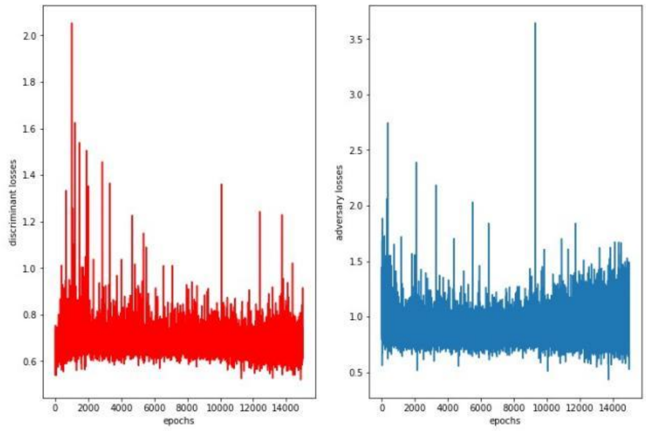

Figure 6 **Loss Curve GAN**

This interplay between the generator and discriminator fosters a dynamic equilibrium. Over time, the generator refines its output to mimic real data distributions, while the discriminator enhances its discriminatory prowess in Figure 6. In the ideal scenario, this process culminates when the generator produces samples indistinguishable from real data, and the discriminator cannot differentiate between real and fake samples beyond chance**. (Mirza, M., & Osindero, S. 2014;) [15].**

However, practical GAN training is often fraught with challenges. Factors suh as architectural design, hyperparameters, and dataset nuances significantly influence training dynamics. Issues like mode collapse, training instability, and gradient vanishing can impede convergence and sample quality. Successful GAN training demands meticulous tuning and monitoring to ensure stable convergence and mitigate overfitting or underfitting risks.

5. Future Work

The advancement of Generative Adversarial Networks (GANs) and Deep Convolutional Generative Adversarial Networks (DCGANs) will continue to unfold across various fronts. Novel architectures and model enhancements are anticipated to be developed, exploring innovative structures and attention mechanisms to improve performance, stability, and scalability. Tackling the persistent challenge of mode collapse will remain a priority, with research efforts aimed at devising robust training techniques and diversity-promoting objectives. Additionally, enabling fine-grained control over generated outputs and facilitating semantic manipulation and attribute conditioning will be crucial for practical applications. Exploring multi-modal generation techniques to enrich diversity and creativity, along with leveraging GANs and DCGANs for unsupervised representation learning, holds promise for discovering meaningful latent representations from unlabeled data. Extending the application of generative models to real-world scenarios across diverse domains, while addressing challenges related to data privacy and ethical considerations, will be pivotal. Moreover, interdisciplinary collaboration among researchers from various fields is expected to foster creativity, encourage cross-pollination of ideas, and accelerate progress toward addressing complex challenges in generative modeling. Through concerted efforts in these directions, GANs and DCGANs are poised to unlock new frontiers and revolutionize numerous domains with their transformative capabilities.

Conclusion

In both Generative Adversarial Networks (GANs) and Deep Convolutional Generative Adversarial Networks (DCGANs) have revolutionized the field of generative modeling, particularly in the context of image generation tasks such as human face synthesis. GANs, with their adversarial training paradigm, have proven to be highly effective in generating diverse and realistic data samples. The adversarial training process between the generator and discriminator networks enables GANs to learn complex data distributions and produce high-quality samples. However, GANs may suffer from training instability and mode collapse, where the generator fails to capture the full diversity of the target distribution. **(Nguyen, A., Yosinski, J., & Clune, J. 2016.;).**

On the other hand, DCGANs extend the GAN framework by incorporating deep convolutional networks, which are particularly well-suited for image data. DCGANs leverage the hierarchical representations learned by convolutional layers to generate more detailed and visually appealing images, especially in tasks such as human face synthesis. Additionally, DCGANs offer improved stability during training and better convergence properties compared to traditional GANs. **(Oord, A. van den, Kalchbrenner, N., & Kavukcuoglu, K. 2016.;).**

In terms of performance for human face generation, DCGANs have demonstrated superior results compared to basic GAN architectures. The use of deep convolutional networks enables DCGANs to capture intricate facial features and produce more realistic faces with finer details. Furthermore, DCGANs exhibit better convergence properties and are less prone to mode collapse, resulting in more stable training and higher-quality generated images. Therefore, in the context of human face generation and similar image synthesis tasks, DCGANs are generally considered better than traditional GANs due to their ability to generate more realistic and visually appealing images with improved stability and convergence. However, it's important to note that the choice between GANs and DCGANs depends on specific application requirements, computational resources, and desired trade-offs between model complexity and performance.

References

[1]. Aouada, D., Cherenkova, K., Gusev, G., et al. (2020). 3D sparse deformation signature for dynamic face recognition. 2020 IEEE international conference on image processing, IEEE, pp. 2835-2839.

[2]. Bardsley, J.M. (2008). An efficient computational method for total variation-penalized Poisson likelihood estimation. Inverse Problems & Imaging, 2(2), 167.

[3]. Cao, C., Chai, M., Woodford, O., & Luo, L. (2018). Stabilized real-time face tracking via a learned dynamic rigidity prior. ACM Transactions on Graphics, 37(6), 1-11.

[4]. Doersch, C. (2016). Tutorial on variational autoencoders. arXiv:1606.05908.

[5]. Eksioglu, E. M. (2017). Deepfake detection using recurrent neural networks. 2017 IEEE 15th international symposium on intelligent systems and informatics (SISY), IEEE, pp. 195-200.

[6]. Fujimura, K., Yamauchi, T., & Yamaguchi, O. (2019). Semi-implicit Eulerian fluid-based sound synthesis with deep generative models. 2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp. 8225-8229.

[7]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[8]. Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine, 29(6), 82-97.

[9]. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167.

[10]. Jaderberg, M., Simonyan, K., Zisserman, A., et al. (2015). Spatial transformer networks. Advances in neural information processing systems, 28, 2017-2025.

[11]. Karras, T., Laine, S., Aila, T. (2019). A style-based generator architecture for generative adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4401-4410.

[12]. Liu, J., Yan, Z., & Ouyang, W. (2020). Benchmarking deep fake detection in the wild. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 9142-9151.

[13]. Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv:1411.1784.
[14]. Nguyen, A., Yosinski, J., & Clune, J. (2016). Synthesizing the preferred inputs for neurons in neural networks via deep generator networks. Advances in neural information processing systems, 29, 3387-3395.
[15]. Oord, A. van den, Kalchbrenner, N., & Kavukcuoglu, K. (2016). Pixel recurrent neural networks. arXiv:1601.06759.