Image Generation Using Generative Adversarial Network

Anand Upadhyay[1], Shivprakash Vishwakarma[2]
[1]*Department of IT, Thakur College of Science and Commerce, Mumbai, India*
[2]*Department of DS, Thakur College of Science and Commerce, Mumbai, India*

Abstract: In recent years, generative adversarial network has shown promising results in generating high-quality images in various domains. This paper present Methods for generating images using generative adversarial network to generate high quality images. DCGAN (Deep Convolutional Generative Adversarial Network) is a type of Generative Adversarial Network (GAN) used for image generation. The DCGAN architecture consists of two main components: a generator and network and a discriminator network. During training, the generator and discriminator are optimized simultaneously using backpropagation and stochastic gradient descent. The training process continues until the generator produces images that are indistinguishable from real images, as determined by the discriminator. DCGANs have been successful in generating high-quality images in various domains, including faces, objects, and scenes, and have become a popular choice for image generation tasks in machine learning.

Keywords : Deep Learning, Generator, Discriminator, Gradient Descent, Image processing, CNN.

1. INTRODUCTION

Generative Adversarial Networks or GANs for short are an approach to generate modelling using deep learning methods, such as convolutional neural networks. Generative Modelling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities or patterns in input data in such a way that the model can be used to generate or output new examples that plausibly could have been drawn from original dataset. Generative Adversarial Networks (GANs) have emerged as a powerful tool for image generation in recent years. Deep Convolutional GANs (DCGANs), in particular, have shown remarkable success in synthesizing high-quality images with a naturalistic appearance. In this paper, we explore the use of DCGANs for image generation in the context of generating high quality images. The main contributions of this work are generating quality images with minimal loss of pixels.

The use of deep convolutional neural networks in GANs has allowed for the generation of images with increased resolution and detail, while maintaining high levels of variability. However, despite their impressive results, the training process of GANs can be challenging, as it requires balancing the objectives of the generator and discriminator.

A. DATASET:

The dataset is about Fashion MNIST. Fashion-MNIST is a dataset of images of clothing and accessories, developed as a direct drop-in replacement for the original MNIST dataset which contains images of handwritten digits. The dataset was created by e-commerce company Zalando and is intended to be used as a benchmark for image classification tasks.

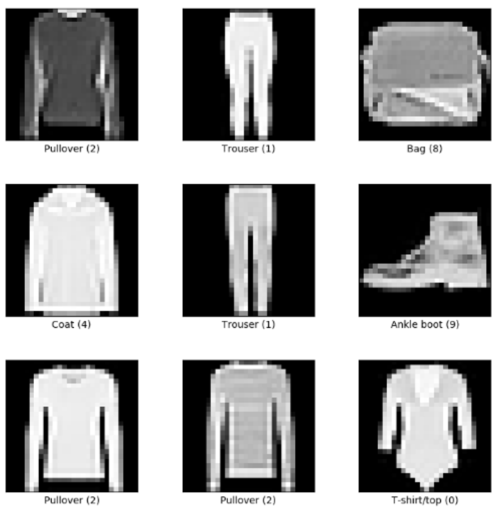
Fashion-MNIST contains 70,000 images in total, with 60,000 images in the training set and 10,000 images in the test set. Each image is 28x28 pixels in size and is grayscale. The images in the dataset depict 10 different types of clothing and accessories, including t-shirts, trousers, bags, shoes, and coats.

The images in the dataset are labelled with one of 10 classes, including T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, and Ankle boot. These classes are intended to be more challenging that the classes in the original MNIST dataset, as the images in Fashion-MNIST are more varied and complex.

Fashion-MNIST is widely used in research and academia as a benchmark dataset for image classification tasks. It is very useful dataset for testing and developing new machine learning models, particularly for image classification and deep learning. Due to the similar structure and image size of the dataset, it is often used as a direct replacement of the original MNIST dataset, and it is a good way to test the performance and generalization of models before applying them to more complex datasets.

B. DATA ACQUISTION :

The Fashion-MNIST dataset is a dataset of images of clothing items, such as shirts, pants, and shoes. It was created by researchers at Zalando, a fashion e-commerce company, as a replacement for the commonly used MNIST dataset, which contains images of handwritten digits. The images in the Fashion-MNIST dataset are 28x28 grayscale images, and each image is labelled with one of 10 different clothing item categories. The dataset is publicly available and can be downloaded from the project's website. It's widely used in research papers and is a common benchmark for image classification tasks.



Categories of Clothes in Fashion-MNIST dataset

C. DATA PIPELINE

A data pipeline is a set of processes and tools used to automate the flow of data from one or more sources to one or more destinations. It typically involves several stages, such as data extraction, transformation, and loading.

Data pipelines are used to automate the movement of data from source systems to target systems, and they are important component of many data-intensive applications and systems. The goal of a data pipeline is to ensure that data is moved from source to target in an efficient, reliable, and secure manner, while also maintaining the integrity of the data throughout the process.

In this research the data pipelines consist of Reloading the dataset in new variables and running the dataset through the scale images pre-processing step and catch the dataset for that batch with the help of cache function then shuffle it up and batch it into 128 images per sample and with the help of prefetch we reduces the likelihood of bottlenecking.

D. ARCHITECTURE OF DCGAN:

The DCGAN architecture consists of two main components : the generator network and the discriminator network.

The generator network is designed to generate new images from random noise inputs. It uses a series of transposed convolutions to upscale the input noise, resulting in an image with the desired resolution. The generator also uses batch normalization and Leaky ReLU activation functions to prevent the network from becoming too complex and to stabilize the training process.

The discriminator network is designed to distinguish real images from generated ones. It uses convolutional layers to extract features from the input images and determine whether the images are real or fake. The discriminator network also uses leaky ReLU activation functions and dropout layers to prevent overfitting and improve the network's performance.

The two networks are trained together, with the generator trying to produce images that fool the discriminator, the discriminator trying to correctly identify real and fake images. The training process continues until the generator produces images that are indistinguishable from real ones, and the discriminator can no longer accurately identify the difference between real and fake images.

The architecture of DCGAN in image generation uses a combination of deep convolutional neural network and generative adversarial networks to generate realistic images from random noise inputs. The generator and discriminator networks work together to produce high-quality images, making DCGAN powerful tool for image generation tasks.

2. LITERATURE REVIEW

Generative Adversarial Networks (GANs) are a powerful and versatile class of neural network architectures that have been used for wide range of Image Generation tasks. GANs consists of two networks: a generator network that produces new images and a discriminator network that attempts to distinguish the generated images from real images. The two networks are trained together in an adversarial process, where the generator tries to produce images that the discriminator cannot distinguish from real images.

The earliest work on GANs was presented in a 2014 paper by Goodfellow at al. titled “Generative Adversarial Networks.”. In this paper, the authors propose a simple yet effective architecture of GANs, consisting of a deep convolutional generator network and a deep convolutional discriminator network. They also propose a training algorithm for GANs, which alternates between updating the generator and discriminator networks using stochastic gradient descent.

One of the key challenges in training GANs in ensuring that the generator and discriminator networks converge to a Nash equilibrium, where the generator produces images that are indistinguishable from real images, and the discriminator is unable to distinguish between real and generated images. Early work on GANs focused on improving the stability and convergence of the training process. For example, the technique of “Wasserstein GANs” (WGANs) proposed by Arjovsky et al in 2017, use Wasserstein distance instead of the traditional Jenson-Shannon divergence to stabilize the training of GANs.

Another line of research has focused on the quality of the generated images. One of the most popular architectures for high-quality image generation is “Progressive GANs”(PGANs) proposed by Karrras et al. in 2018. PGANs use a multi-scale architecture that progressively increases the resolution of the generated images during training, resulting in highly realistic images. Another important architectures is “Style-based GANs” (StyleGANs) proposed by Karras et al. in 2019, which is uses a separate generator network to control the style of the generated images, resulting in fine-grained control over the appearance of the generated images.

There have also been many papers that propose GANs for specific image generation tasks. For example, “Super-Resolution GANs” (SRGANs) proposed by Ledig et al. in 2016 use GANs to generated high-resolution images from low-resolution inputs. “Image-to-Image Translation GANs” (I2I GANs) proposed by Isola et al. in 2016 use GANs to translate images from one domain to another, such as converting photographs of day to night or changing the season. “Text-to-Image GANs” (T2I GANs) proposed by Reed et al. in 2016 use GANs to generate images from text descriptions, allowing users to generate images of objects or scenes described in text.

One of the limitations of GANs is that the generated images are often difficult to interpret and understand. To overcome this limitation, there have been several papers that propose methods for visualizing and understanding the internal representations of GANs. For example, “Visualizing and Understanding Generative Adversarial Networks” (XU et al. 2017) proposed a method for visualizing the internal representations of GANs using t-SNE, while “Understanding and Improving Interpolation in Generative Adversarial Networks” (Berthelot et al. 2018) proposed a method for visualizing the internal representations of GANs using interpolation.

3. ALGORITHM AND TECHNIQUES

So for this research we have DCGAN(Deep Convolutional Generative Adversarial Network) is a popular deep learning architecture used for generating images. It consists of two main components: a generator network and a discriminator network. The generator network generated synthetic images by transforming a low-dimensional random noise vector into a high-dimensional image, while the discriminator network evaluates the realism of the generated images. The two networks are trained simultaneously in an adversarial manner. The generator network is trained to generate images that can fool the discriminator network, while the discriminator network is trained to accurately identify real and fake images. The training process continues until the generator network generated synthetic images that are indistinguishable from real images.

The generator network in a DCGAN typically uses transposed convolutional layers to upsample the input noise vector into an image, and it uses batch normalization and activation functions to ensure stability during training. The discriminator network, on the other hand, uses convolutional layers to extract features from the input image, and it uses fully connected layers to make the binary classification prediction.

The training process of a DCGAN involves updating the parameters of both the generator and discriminator networks by alternating between the two. The generator tries to generate images that are indistinguishable from real images and the discriminator tries to correctly classify the images as real or fake. The goal is to find a Nash Equilibrium, where the generator is generating images that the discriminator cannot accurately distinguish from real images.

4. METHODOLOGIES

DCGAN is a deep learning-based image generation algorithm that uses generative adversarial networks(GANs). The main idea behind DCGAN is to use convolutional neural network (CNNs) to generate new images from a random noise vector, and to train the generator network using a discriminator network that classifies images as either real or fake:

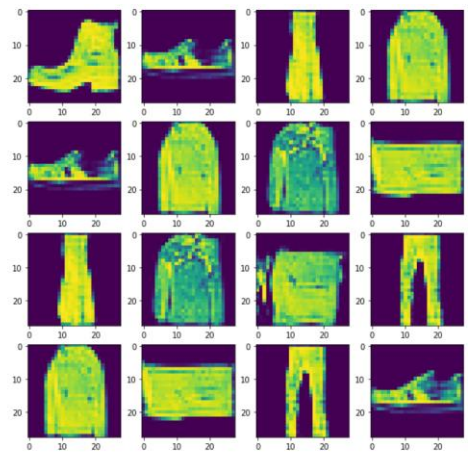
1. Input: A dataset of real images is used to train the DCGAN.
2. Generator Network: The generator network takes a random noise vector as input and generates an image using transposed convolutions and batch normalization.
3. Discriminator Network: The discriminator network takes an image as input and outputs a single scalar value indicating whether the image is real or fake.
4. Training: The generator and discriminator networks are trained simultaneously. During the training process, the generator network tries to produce images that fool the discriminator network, while the discriminator network tries to accurately identify whether an image is real or fake.
5. Output: After training, the generator network can generate new images that are similar to the real images in the training dataset.
6. Evaluation: The quality of the generated images can be evaluated by comparing them with real images from the training dataset, and by observing how well they can fool the discriminator network.

Note: In the DCGAN algorithm, the generator network and discriminator network are both CNNs that are trained using backpropagation and gradient descent. The training process is an iterative process that continues until the generator network produces images that are indistinguishable from real images.

5. RESULT

Image Generation Using GAN. The dataset was divided into train and test part. Where the dataset consists of 60,000 images and test consists of 10,000 images. Then both model is trained simultaneously. The model is trained at 1500 epochs as the learning rate is set at 0.0001 for generator model and 0.00001 for discriminator model with Adam Optimizer and Calculating the loss of Generator and Discriminator with BinaryCrossEntropy.

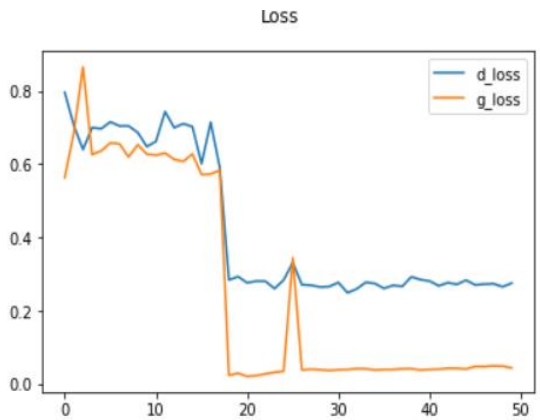
IMAGE GENERATED BY GAN MODEL



The Gradient Descent plays important role in generating the images as it used SGD to optimize the generator and discriminator networks by updating their parameters based on the gradient of their respective loss functions. The generator's loss function is defined as the negative of the discriminator output when applied to the generated image, and the discriminator's loss functions is defined as the binary cross-entropy between the output of the discriminator for real and fake images and the corresponding ground-truth labels.

At each iteration of the training process, a single random sample from the training dataset is selected, and the gradient of the loss functions with respect to the parameters are estimated using backpropagation. The parameters are then updated in the direction that reduces the loss, using the gradients estimated from the randomly selected sample

PERFORMANCE OF GENERATOR AND DISCRIMINATOR IN IMAGE GENERATION



6. CONCLUSION

In Conclusion, the use of Deep Convolutional Generative Adversarial Networks (DCGANs) for image generation has shown promising results in recent years. DCGANs are able to learn to rich and diverse distribution of images, and they can be used to generate a wide range of images, from realistic photographs to abstract art. The use of convolutional layers in the generator and discriminator networks allows DCGANs to capture the spatial structure of the images, which is essential for many image generation tasks.

Overall, the results of this research demonstrate the potential of DCGANs for image generation, and they suggest that DCGANs may be useful tool for a wide range of image generation tasks. Further research is needed to explore the full potential of DCGANs and to develop new techniques for improving their performance and stability. Nevertheless, the results of this research provide a solid foundation for future work in this area.

REFERENCE

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair,S., & bengio, Y. (2014). Generative Adversarial Nets. In Advances in Neural information processing systems.

[2] Lecun, Yann, Yoshua Bengio, and Geoffery Hinton."Deep Learning." Nature521, no.7553(2015): 436-444.

[3] Liu, Ming-Yu, and Oncel Tuzel." Coupled Generative adversarial networks."Advances in neural information processing systems (2016).

[4] Radford, A., Metz, L., & Chintala,S.(2015). Unsupervised representation learning with deep convolutional generative adversarial networks.

[5] Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks.

[6] Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2019). A style-based generator architecture for generative adversarial networks.

[7] Liu, M.Y., Breuel, T., Kautz, J., & Bulskov, D. (2015, June). Unsupervised image-to-image translation networks. In Proceedings of the IEEE international conference on computer vision.

[8] Kim, T., Cha, M., Kim, H., & Choe, Y. (2018). Learning to Discover Cross-Domain Realtions with Generative Adversarial Networks.

[9] Dosovitskiy, A., & Brox, T. (2016). Generating images with perceptual similarity metrics based on deep networks.

[10] Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images.

[11] Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative Adversarial Text to Image Synthesis.

[12] Karras, T., Laine, S., & Aila, T. (2019). Progressive growing of GANs for improved quality, stability and variation. In International Conference on Learning Representaitions..

[13] Singh, N.K. and Raza, K., (2021). Medical image generation using generative adversarial networks. A review. Health informatics: A computational perspective in health.