Deep Learning Project Proposal – Spring-24 – NYU – Image Synthesis With Generative Adversarial Networks (GANs) Using CIFAR-10

Ark Pandey[1*], Durga Avinash Kodavalla[1*], Priyangshu Pal[2*]

[1]New York University, Tandon School of Engineering, Department of Electrical and Computer Engineering (ECE)
[2]New York University, Tandon School of Engineering, Department of Computer Science and Engineering (CSE)

Problem Statement

The demand for generating realistic images is increasing across various industries, driven by the growing use of technologies like augmented reality, virtual environments, and digital content creation. This prompts the need for advanced deep-learning solutions. This project leverages Generative Adversarial Networks (GANs) to synthesize diverse and photorealistic images from scratch trained on CIFAR-10 dataset. The project aims to gain a deeper understanding of GAN models, their training dynamics, and the challenges in generating high-quality images. The outcome of this project will contribute to understanding the image synthesis technology and have practical applications in various fields, meeting the current need for realistic image generation and paving the way for future innovations.

Literature Survey

Ian J. Goodfellow et al.'s paper "Generative Adversarial Nets,"[2] published in 2014, introduced the groundbreaking framework of Generative Adversarial Networks (GANs). GANs introduce a game-theoretic framework where two neural networks, a generator and a discriminator, are trained simultaneously in an adversarial manner. Despite their success, traditional GAN training methods often face challenges such as mode collapse, training instability, and limited image quality, particularly when generating high-resolution images. Karras et al. proposed the "Progressive Growing of GANs" technique[3], which aims to cope with these challenges by incrementally growing both the generator and discriminator networks during training. This technique offers a promising solution to longstanding challenges in generating high-resolution, realistic images.

Dataset

The CIFAR-10 dataset[1] contains 60000 32x32 color images. The dataset has 10 classes, each having 6000 images and these classes are completely mutually exclusive. Training a GAN on CIFAR-10 requires the generator to learn patterns and textures to generate realistic images without demanding a lot of compute resources, making it a suitable benchmark for evaluating the capabilities of GAN architectures for the project.

Model

GANs consist of two neural networks, a discriminator and a generator which are trained simultaneously through a mini-max game network:

- **Generator:** Generate photorealistic images indistinguishable from real images. It has convolutional layers, batch normalization, ReLU activation functions, and tanh activation function in the output layer.
- **Discriminator:** Distinguish between real and fake images generated by the generator. It is similar to generator architecture with leaky ReLU activation, and sigmoid activation in the output layer.

Deliverables

1. **Train the models:** Training a GAN involves a unique process where two neural networks, the generator and the discriminator, are trained alternatively in a mini-batch stochastic gradient descent manner.
2. **Use different methods for experimentation:** This involves exploring various architectures, data augmentations, loss functions, optimization techniques, regularization, etc., to improve performance and generate high-quality outputs.
3. **Model evaluation:** Performance of trained GAN models is evaluated by visually inspecting quality of generated images and metrics such as the Frechet Inception Distance (FID) & Inception Score (IS) for quantitative evaluation by using pre-trained ResNet18 model as classifier.

References

[1] Krizhevsky, A. (2009). Learning multiple layers of features from tiny images.

[2] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. (2014). Generative Adversarial Networks.

[3] T. Karras, T. Aila, S. Laine, and J. Lehtinen, Progressive Growing of GANs for Improved Quality, Stability, and Variation. 2018.

*These authors contributed equally.