Custon GAN Implementation for CIFAR-10

Abdullah Masood
21i-0822
FAST-NUCES

I. Introduction

A custom Generative Adversarial Network (GAN) has been created using the CIFAR-10 dataset, with a specific focus on generating images of cats and dogs. This implementation incorporates a distinctive discriminator that provides a similarity score between pairs of images, enhancing the conventional GAN architecture. Drawing inspiration from Siamese Networks, the discriminator evaluates the likeness between generated and real images, enabling the generator to receive more nuanced feedback to reduce dissimilarities. This method seeks to enhance the training process, leading to the generation of more realistic images and a deeper understanding of the underlying data distribution.

II. Methodology

The methodology focuses on creating a custom GAN architecture that includes a generator and a modified discriminator, utilizing the CIFAR-10 dataset. The generator consists of multiple transposed convolutional layers that convert random noise into realistic images, specifically of cats and dogs. The discriminator operates similarly to a Siamese Network, taking two images as input—a generated image from the generator and a real image from the dataset—and producing a similarity score that indicates how alike they are.

The training process employs the Binary Cross-Entropy loss function to optimize both networks. The discriminator is trained to maximize the difference between real and fake images, while the generator works to minimize this score. Data preprocessing involves normalization and resizing, concentrating on the cat and dog classes to filter the dataset effectively. The GAN is trained for 400 epochs, with progress tracked through loss metrics and visualized image outputs at regular intervals, allowing for iterative enhancements in the quality of generated images.

III. Results

The training of the custom GAN showed significant advancements in generating realistic images of cats and dogs throughout the 400 epochs. During the training process, the loss values for both the generator and discriminator consistently decreased, indicating successful convergence. Visual assessments of the generated images revealed substantial improvements in quality, with later epochs producing outputs that closely resembled actual images from the CIFAR-10 dataset.

Random samples from the generator exhibited greater detail and fidelity, demonstrating the model's capability to capture essential features of the target classes. The inclusion of the similarity score in the discriminator provided more effective feedback, leading to a refined generator that could produce diverse and convincing images. Overall, the modified GAN architecture not only enhanced the realism of the generated images but also showcased the effectiveness of employing similarity metrics in adversarial training.

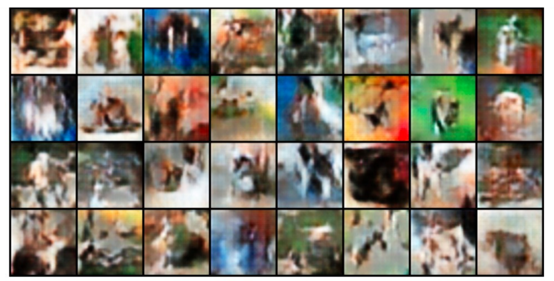

Fig. 1. Model Result Images of Cats Dogs

IV. Discussion

The implementation of a custom GAN with a similarity-based discriminator offered several benefits compared to traditional GAN architectures. By incorporating a Siamese Network approach in the discriminator, the model was able to provide a more detailed evaluation of generated images, emphasizing their similarity to real images instead of just categorizing them as real or fake. This strategy improved the generator's performance, allowing it to receive direct feedback on its ability to create realistic images, which resulted in better image quality and diversity. The training dynamics were also enhanced, as indicated by more stable loss metrics and convergence compared to conventional GANs. However, challenges like mode collapse—where the generator produces limited image variations—still persisted. Future research could investigate advanced techniques such as feature matching or utilizing Wasserstein distance to further stabilize the training process and enhance image generation quality. Overall, these results highlight the promise of using similarity scores in GAN training, paving the way for future developments in generative modeling.

V. Conclusion

The custom Generative Adversarial Network (GAN) designed for generating cat and dog images from the CIFAR-10 dataset effectively showcases the benefits of a similarity-based discriminator. By implementing a Siamese Network architecture, the model offered more accurate feedback to the generator, resulting in notable improvements in both the realism and diversity of the generated images. The training outcomes demonstrated the effectiveness of this approach, with loss metrics decreasing and image quality improving throughout the training process.

Although challenges like mode collapse were observed, the overall success of this methodology points to a promising path for future research in generative modeling. This work not only advances the field of GANs but also opens up opportunities for further investigation into using similarity metrics to enhance generative performance.

REFERENCES

[1] EmadBinAbid (2019). image-generator/src at master · EmadBinAbid/image-generator. [online] GitHub. Available at: https://github.com/EmadBinAbid/image-generator/tree/master/src [Accessed 20 Oct. 2024].

[2] Abid, E.B. (2020). Generative Adversarial Network on CIFAR-10 - Emad Bin Abid - Medium. [online] Medium. Available at: https://medium.com/@emad.bin.abid/generative-adversarial-networkon-cifar-10-8d8deeec1fd7 [Accessed 20 Oct. 2024].