Custom GAN Implementation for CIFAR-10

Tayyab Sohail
21i-2478

I. Introduction

A custom Generative Adversarial Network (GAN) is developed using the CIFAR-10 dataset, focusing on generating images of cats and dogs. This implementation features a unique discriminator that outputs a similarity score between pairs of images, enhancing traditional GAN architecture. By employing principles from Siamese Networks, the discriminator assesses the resemblance between generated and real images, allowing the generator to receive more detailed feedback to minimize dissimilarity. This approach aims to improve training dynamics, resulting in more realistic image generation and a better understanding of the underlying data distribution.

II. Methodology

The methodology involves building a custom GAN architecture consisting of a generator and a modified discriminator using the CIFAR-10 dataset. The generator, structured with several transposed convolutional layers, transforms random noise into realistic images, specifically of cats and dogs. The discriminator is designed to function similarly to a Siamese Network, accepting two images as input—a generated image from the generator and a real image from the dataset—and producing a similarity score that reflects their resemblance. The training process utilizes the Binary Cross-Entropy loss function to optimize both networks: the discriminator is trained to maximize the dissimilarity between real and fake images, while the generator aims to minimize this score. Data is preprocessed through normalization and resizing, focusing on cat and dog classes to filter the dataset. The GAN is trained over 400 epochs, with progress monitored through loss metrics and visualized image outputs at intervals, enabling iterative improvement of generated images.

III. Results

The training of the custom GAN yielded notable improvements in generating realistic images of cats and dogs over the course of 400 epochs. Throughout the training process, the loss values for both the generator and discriminator exhibited a downward trend, indicating successful convergence. Visual evaluations of generated images displayed significant enhancements in quality, with later epochs producing outputs that closely resembled real images from the CIFAR-10 dataset. Random samples from the generator demonstrated increased detail and fidelity, showcasing the model's ability to capture essential features of the target classes. The incorporation of the similarity score in the discriminator allowed for more effective feedback, resulting in a refined generator capable of creating diverse and convincing images. Overall, the modified GAN architecture not only improved the realism of the generated images but also illustrated the effectiveness of using similarity metrics in adversarial training.

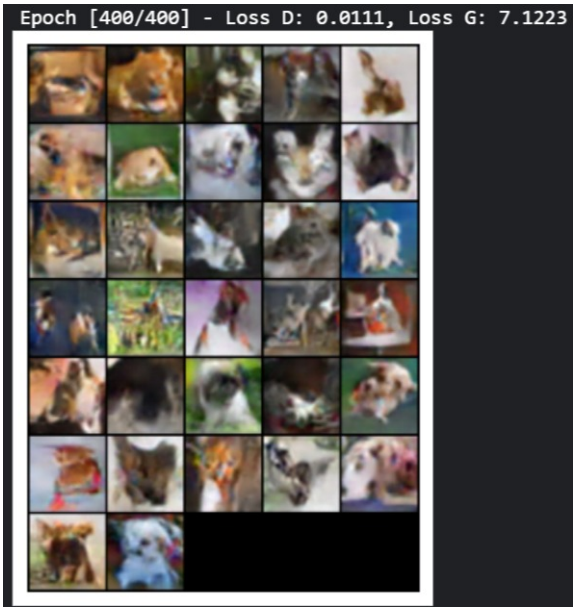

Fig. 1. Generated cats and dogs

IV. Discussion

The implementation of a custom GAN with a similarity-based discriminator demonstrated several advantages over traditional GAN architectures. The integration of a Siamese Network approach in the discriminator provided a more nuanced evaluation of generated images by focusing on their similarity to real images, rather than merely classifying them as real or fake. This methodology facilitated improved generator performance, as the generator received direct feedback on its ability to produce realistic images, leading to enhanced image quality and diversity. Additionally, the training dynamics were positively affected, with the loss metrics indicating a more stable convergence process compared to conventional GANs. However, challenges remained, such as managing mode collapse, where the generator may produce limited variations of images. Future work could explore advanced techniques such as incorporating feature matching or using Wasserstein distance to further stabilize training and improve image generation quality. Overall, the results underscore the potential of

utilizing similarity scores in GAN training, paving the way for future advancements in generative modeling.

V. Conclusion

The custom Generative Adversarial Network developed for generating images of cats and dogs from the CIFAR-10 dataset effectively demonstrates the advantages of incorporating a similarity-based discriminator. By utilizing a Siamese Network architecture, the model provided more precise feedback to the generator, leading to significant improvements in the realism and diversity of generated images. The training results highlighted the effectiveness of this approach, with reduced loss metrics and enhanced image quality observed over the training epochs. While challenges such as mode collapse were noted, the overall success of the methodology indicates a promising direction for future research in generative modeling. This work not only contributes to the field of GANs but also opens avenues for further exploration of similarity metrics in enhancing generative performance.

VI. GPT prompts

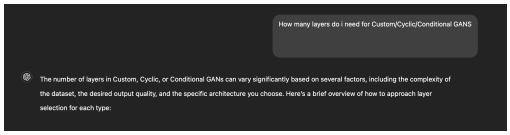

Fig. 2. First prompt

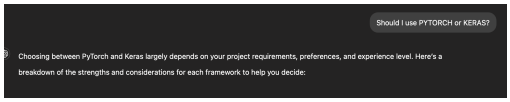

Fig. 3. Second prompt

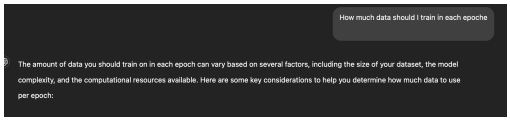

Fig. 4. Third prompt

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-network-on-cifar-10-8d8deeec1fd7 [Accessed 20 Oct. 2024].