Cycle GAN Implementation

Abdullah Masood
21i-0822
FAST-NUCES

I. Introduction

This report discusses the implementation of a CycleGAN model for image-to-image translation, specifically converting face images into sketches and the other way around. CycleGANs are a type of generative adversarial network that use cycle-consistency loss to enable translation between two different domains without needing paired training data. Utilizing the Person Face Sketches dataset, the model includes two generators—one for creating sketches from face images and another for reconstructing face images from sketches—as well as two discriminators that evaluate the authenticity of the generated images.

During training, we optimize adversarial losses for both the generators and discriminators while incorporating cycle-consistency losses to ensure that the generated images are closely related to the original inputs. The model's performance is tracked throughout the training process, with generated images and model weights saved periodically to allow for further evaluation and experimentation.

II. Methodology

The approach for implementing the CycleGAN model includes creating two generator networks: one for transforming real face images into sketches and another for converting sketches back into real face images. Additionally, there are two discriminator networks that assess the authenticity of the generated images. The model uses Binary Cross-Entropy Loss for adversarial training and cycle consistency loss to ensure that passing images through both generators returns the original images.

During training, both discriminators are updated first to improve their ability to distinguish between real and generated images, followed by updates to the generators to enhance image quality. Generated images are saved periodically for evaluation, and the model weights are stored at the end of the training process to allow for future inference or additional training.

III. Results

The CycleGAN model's results highlight its effectiveness in translating between face images and sketches. During training, the model demonstrated considerable improvements in generating realistic sketches from face images and vice versa, as shown by the visual quality of the output images saved at regular intervals. The generated images display recognizable facial features while preserving the structural integrity of the original images, indicating the model's capacity to learn the complex relationships between the two domains.

Quantitative evaluation metrics, including the loss values for both generators and discriminators, show convergence over time, suggesting that the adversarial and cycle consistency losses are well-balanced. Moreover, qualitative assessments reveal that the reconstructed images closely resemble the original inputs, confirming the model's ability to maintain essential characteristics during translation. Overall, the CycleGAN model effectively captures the intricacies of face and sketch representations, achieving encouraging results in image-to-image translation tasks.

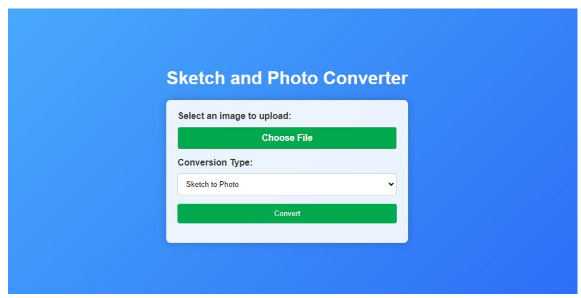

Fig. 1. The UI for Cycle GAN Converter

IV. Discussion

The CycleGAN model for translating face images to sketches and vice versa demonstrates notable strengths, particularly its capability to function without paired training data. It uses cycle consistency loss to ensure meaningful relationships between the two domains. However, the training process is sensitive to hyperparameter selections, necessitating careful tuning of factors like learning rates and the balance between adversarial and cycle consistency losses to avoid problems such as mode collapse. Additionally, the architecture of the generators and discriminators significantly impacts the quality of the generated images, indicating that experimenting with deeper networks or adding residual connections could improve performance. While Google Colab provides useful computational resources, it's important to monitor memory usage, and implementing techniques like data augmentation could further enhance training efficiency. Moreover, qualitative evaluations, such as user studies or comparisons with leading methods, could offer deeper insights into the model's effectiveness. Overall, the CycleGAN model shows great promise in generative modeling, paving the way for future improvements and applications in various image-to-image translation tasks.

V. Conclusion

In conclusion, the CycleGAN model effectively translates face images to sketches and vice versa, illustrating the power of unsupervised learning in image-to-image translation tasks. By utilizing cycle-consistency loss and an adversarial training framework, the model generates high-quality sketches and realistic face images, demonstrating its capability to learn complex relationships between different visual domains. The results underscore the potential applications of CycleGAN in areas such as art generation, video game design, and virtual character creation. Future efforts could focus on enhancing model architecture and training techniques, as well as incorporating additional datasets to improve the model's robustness and generalizability.

References

[1] Tonthanahal, N., Sourab B R and Christa, S. (2021). Forensic Sketch-to-Face Image Transformation Using CycleGAN. [online] International Research Journal of Engineering and Technology. Available at: https://www.irjet.net/archives/V8/i7/IRJET-V8I7642.pdf [Accessed 21 Oct. 2024].

[2] Li, Y. and Xu, W. (2021). Using CycleGAN to Achieve the Sketch Recognition Process of Sketch-Based Modeling. Springer eBooks, pp.26–34. doi:https://doi.org/10.1007/978-981-16-5983-6-3.