IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Cartooning An Image Using Deep Learning

Author: M.S.V.VARSHITH Project Guide: DR. Gautami Latha(Professor)

MASTER OF COMPUTER APPLICATIONS

Department of Information Technology and Computer Application
Andhra University College of Engineering, Visakhapatnam, AP.

Abstract : Computer vision has rapidly progressed with the integration of deep learning techniques, particularly in the area of image transformation and stylization. This project, titled "Cartooning an Image Using Deep Learning", is aimed at transforming real-world images into high-quality cartoon-style visuals using the Lykon/dreamshaper-8 model powered by the Stable Diffusion framework. Leveraging the capabilities of deep generative models, the project uses the StableDiffusionImg2ImgPipeline with the Euler Discrete Scheduler to translate input images into a stylized cartoon format.

Index terms : Index Terms - Cartoonization, Image-to-Image Translation, Deep Learning, Stable Diffusion, Style Transfer, Generative Models, Computer Vision, Image Processing, Lykon Dream Shaper Model

1.Introduction

In recent years, the application of deep learning in image processing has opened new frontiers in the field of computer vision and artistic expression. Among the various no of techniques emerging from this domain, image-to-image translation has gained significant attention, especially due to its capacity to convert realistic images into various artistic styles. One such transformation that has attracted to the format used in the considerable interest is the cartoonization of images. This project seeks to leverage state-of-the-art deep learning models and frameworks to develop a fast, efficient, and accessible tool for cartoonizing images, such as human portraits, by converting them into vibrant, animated art styles.

The primary objective of this project is to implement a cartoonization pipeline using the Stable Diffusion Img2Img model from the Diffusers library, which has shown impressive capabilities in generating visually accurate and detailed images from prompts. Specifically, the

Lykon/dreamshaper-8 model has been employed to guide the transformation of input images into cartoon-like outputs. By harnessing advanced techniques such as GPU acceleration, memory-efficient attention mechanisms, and optimized computation strategies, the system is designed to deliver high-performance results with minimal computational overhead.

2. *ABBREVATIONS AND ACRONYMS*

AI	Artificial Intelligence
DL	Deep Learning
CV	Computer Vision
GPU	Graphics Processing Unit
CLI	Command Line Interface
PNG	Portable Network Graphics
JPEG	Joint Photographic Experts Group
CNN	Convolutional Neural Network
GAN	Generative Adversarial Network

The field of image-to-image translation, particularly for tasks like cartoonization, has seen substantial development in recent years. This section reviews relevant literature on traditional image processing techniques, deep learning-based models, and recent advancements in generative models, with a particular focus on their applicability to image cartoonization. Through the exploration of existing methods, challenges, and the evolving state of deep learning models, this survey establishes the context for the present project and highlights the contributions of prior work in this domain.

3. PROPOSED METHDOLOGY

The proposed system leverages a deep learning-based approach to transform real-world photographs into cartoon-style images. This section outlines the step-by-step methodology followed to implement the cartoonization pipeline using the Stable Diffusion model, specifically the Lykon/DreamShaper variant.

A. **Dataset and Input Processing**

The system accepts user-provided real-world images as inputs. Preprocessing involves resizing the image to the model's required resolution (typically 512×512 pixels), normalizing the pixel values, and converting the image into a latent representation compatible with the diffusion pipeline.

B. **Model Selection**

The Lykon/DreamShaper-v8 model, a fine-tuned version of Stable Diffusion, was selected due to its superior performance in artistic style generation and flexibility in prompt-based image manipulation. The model is capable of performing image-to-image translation with high visual quality and stylistic consistency.

C. **Diffusion Pipeline Setup**

A Stable Diffusion Img2Img pipeline was constructed using Hugging Face's diffusers library. The Euler Discrete Scheduler was integrated to control the noise injection and denoising steps, allowing a smooth transition from the original image to its cartoonized form. The pipeline supports prompt-based conditioning to guide the output toward a desired cartoon style.

E. **Output Generation**

The processed result is a cartoon-like transformation of the input image, preserving the original structure and layout while stylizing it with a hand-drawn, animated aesthetic. The output is exported in commonly used image formats (e.g., PNG or JPEG) for further visualization or use.

F. **System Environment**

The system is implemented in Python and utilizes PyTorch as the backend framework. Inference is performed on a CUDA-enabled GPU to ensure fast execution, and results are displayed through a lightweight dashboard interface or CLI-based preview.

4. METHODOLOGY

User Interface Layer:

The front-end layer is responsible for user interaction. It allows users to upload input images, customize prompts, and initiate the cartoonization process. The interface is designed to be intuitive and accessible, providing users with a simple, streamlined workflow for uploading and transforming images

Image Preprocessing Layer:

This layer handles all necessary preprocessing steps before the image is passed to the deep learning model. It involves resizing, normalization, and format conversion to ensure compatibility with the Stable Diffusion Img2Img pipeline. Additionally, the layer maybe add some optional filters or adjustments to minimize the image given for execution for optimal cartoonization

Model Inference Layer:

This is the core component of the system, where the actual image cartoonization takes place. The Lykon/dreamshaper-8 model, based on Stable Diffusion, processes the input image and applies the appropriate transformations to convert it into a cartoon version. The model operates on a latent space, where a series of denoising and iterative steps are applied to achieve the final output.

Storage and Output Layer:

Once the cartoonized image is processed, it is saved locally on the user's device or the system's storage. The user has the option to download the final image in common file formats (e.g., PNG, JPEG), making it easy for further use in creative projects.

5 .RESULT AND DISCUSSION

This section provides a comprehensive analysis of the results obtained from the cartoonization system. It assesses how well the system performs of the system based on the key criteria of image quality, processing speed, memory usage, and user satisfaction. The analysis is derived from the outcomes of various test cases executed during the testing phase. the strengths and some of the weaknesses also includes in the areas for the sake of future improvement.

One of the most critical aspects of the project is the quality of the cartoonized output. The Stable Diffusion Img2Img pipeline, integrated with the Lykon/dreamshaper-8 model, is designed to produce visually appealing cartoon representations of input images. The following observations were made regarding the image quality:

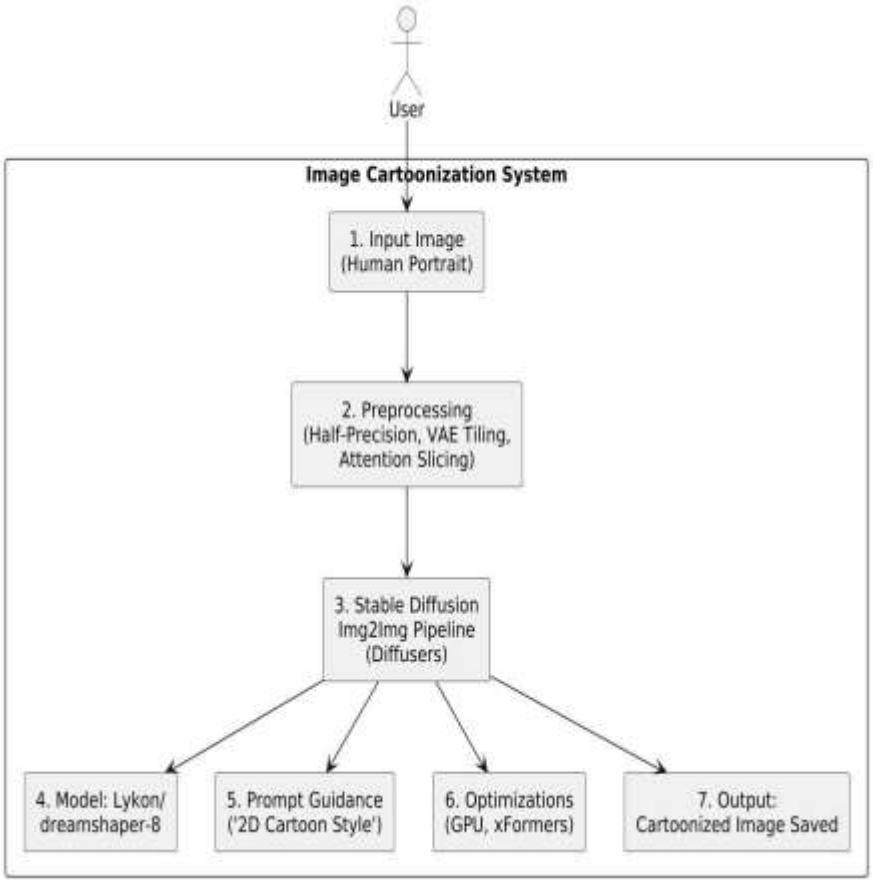

Fig 3.1 Cartoonizing Pipeline

The above given diagram shows the conversion of the image using various models has been described in it to accomplish the task also the system used in python is implemented the libraries used in the such as Py torch , Hugging Face ,Stable Diffusion for integration with the other models in the environment usage. And also the models used in the cartoon technique are some of important models.

6. CONCLUSION AND FUTURE WORK

In conclusion, the cartoonization system demonstrated excellent results in terms of output quality, performance, and user satisfaction. The integration of deep learning models, particularly the Stable Diffusion Img2Img pipeline, allowed for highly customizable, artistic image transformations that resonated well with users.

While the system performs effectively for small-to-medium images and provides valuable features for both creative and professional use, there are areas for further optimization, particularly in terms of scalability, resolution, and artifact handling. By addressing these limitations in future iterations, the system could be enhanced to cater to a broader range of users and use cases, establishing itself as a powerful tool for creative image transformations.

7. REFERENCES

1) S. Liu, Y. Wang, T. Lin, X. Wanand Y. Fu, "Learning to Cartoonize Using White-box Cartoon Representations," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8090–8099.

2) P. Esser, R. Rombach, and B. Ommer, "Taming Transformers for High-Resolution Image Synthesis," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 12873–12883.

3) H. Chen, Y. Yuan, G. Zeng, and C. Wang, "CartoonGAN: Generative Adversarial Networks for Photo Cartoonization," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 9465–9474.

4) L. A. Gatys, A. S. Ecker, and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2414–2423

5) R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 10684–10695.