CARTOONIFYING IMAGES USING DEEP LEARNING AND OPENCV IN PYTHON

Filjo T R[1], Ashimna Shamsudeen[2], Rohit Shaji[3],Jithin Davis[4], Nandakrishna K M[5],Vaishnav V Bhatt[6],Ajith P J[7],Bindu Anto[8]

[1] *Student , Computer Engineering, KKMMPTC MALA, Kerala, India.*
[2] *Student , Computer Engineering, KKMMPTC MALA, Kerala, India.*
[3] *Student, Computer Engineering, KKMMPTC MALA, Kerala, India.*
[4] *Student ,Computer Engineering, KKMMPTC MALA, Kerala, India.*
[5] *Student ,Computer Engineering, KKMMPTC MALA, Kerala, India.*
[6] *Student , Computer Engineering, KKMMPTC MALA, Kerala, India.*
[7] *HOD Robotic Process Automation, KKMMPTC MALA, Kerala, India.*
[8] *HOD Computer Engineering, KKMMPTC MALA, Kerala, India.*

ABSTRACT

In the contemporary era of digital art and entertainment, there exists a burgeoning trend that involves the transformation of ordinary photographs into captivating cartoons. This project introduces a pioneering and inventive approach that harnesses a fusion of advanced deep learning techniques and the versatile Python library known as OpenCV. The fundamental methodology can be distilled into two crucial and interlinked steps: the employment of a deep learning-based model, such as a Convolutional Neural Network (CNN), to master the intricate art of edge detection. This initial step serves as the bedrock for capturing the characteristic outlines and distinctive shapes reminiscent of classic cartoon drawings. The next pivotal phase in the process involves the application of a sophisticated color quantization algorithm. The purpose here is to systematically reduce the multitude of distinct colors within the input image, artfully replicating the simplified yet enchanting color palettes that are so quintessential to the world of cartoons. In this intricate journey towards cartoonification, a plethora of diverse color quantization methods are explored and experimented with. The goal is to arrive at a visual aesthetic that not only delights the eye but also evokes the nostalgic charm of traditional hand-drawn animations.As these two essential components of the project, edge detection and color quantization, converge, a transformative metamorphosis takes place. The result is an image that transcends the realm of ordinary photography, metamorphosing into a captivating and artistic cartoon. The marriage of cutting-edge deep learning technology with the versatile capabilities of OpenCV empowers creators and artists to infuse their photographs with a distinct, timeless, and visually appealing cartoon-like quality.This endeavor, at its core, is a testament to the ever-evolving intersection of technology and art, offering a creative avenue that invites exploration and experimentation, leading to the emergence of unique and engaging visual experiences.

Keyword:- *Open CV(COMPUTER VISION) library in Python, Edge detection, Deep learning-based edge detection model, CNN:-Convolutional Neural Network, Color quantization.*

1.INTRODUCTION

Cartoons are often used in numerous situations. Since cartoons are produced creatively, they require sophisticated and precise artistic abilities. It takes a lot of time for the artist to depict cartoons in large quantities for animated films since they must accurately define the cartoon's sketch to achieve a decent outcome. People attempt to turn their own images into cartoons in order to understand how they would seem in a cartoon, which is one of the many ways

that cartoons are used outside of the film business. However, because this process takes so long and uses so many resources, it could be challenging to gather them all and carry it out properly. It takes less time for artists to create films using this cartoonizing process. By using this method, they can achieve the same results as before. However, cartooning is a quick and effective alternative to drawing an image.

2.Milestones

In 2016, The paper titled "A Neural Algorithm of Artistic Style, Leon A. Gatys, Alexander S. Ecker, Matthias Bethge" [1].Humans have mastered the technique of crafting a complex interaction between the content and style of an image in great art, particularly painting, to produce singular visual experiences. There is no artificial system that can match the capabilities of this process, and its algorithmic foundation is unknown. A group of biologically inspired vision models known as Deep Neural Networks has showed near-human performance in challenging visual tasks including object and face identification. Here, we introduce a synthetic system that produces creative images of excellent perceptual quality using a Deep Neural Network. A neural algorithm is provided by the system, which can separate and recombine the information and style of any given image. Our approach provides a road towards an algorithmic understanding of how people create and interpret artistic imagery in light of previous studies utilising FMRI and electrophysiology that have demonstrated startling similarities between performance-optimised artificial neural networks and biological vision. The method presents a distinct class of stimuli that could be used to test particular computational theories about how the human brain interprets creative style.

In 2016, The paper titled "Image Style Transfer Using Convolutional Neural Networks,Leon A. Gatys, Alexander S. Ecker"[2].A challenging image processing task is to render the semantic content of an image in many styles. The absence of image representations that directly convey semantic information and enable the separation of image content from style may be argued to be a significant limiting factor for prior methods. Here, we make high level picture information explicit by using image representations built from Convolutional Neural Networks tuned for object detection. We present A Neural Algorithm of Artistic Style, which can deconstruct and reconstitute the style and content of real photographs. By fusing the appearance of multiple well-known works of art with the content of arbitrary photographs, the algorithm enables us to create new images with great perceptual quality. Our findings show how Convolutional Neural Networks can learn deep image representations and show how they may be used for sophisticated image synthesis and modification.

In 2016, The paper titled "Perceptual Losses for Real-Time Style Transfer and Super-Resolution, Justin Johnson, AlexandreAlahi, Li Fei-Fei" [3].Problems involving the transformation of an input image into an output image are taken into consideration. A per-pixel loss between the output and ground-truth images is a common technique used to train feed-forward convolutional neural networks for these types of challenges. Parallel research has demonstrated that high-quality images can be produced by developing and optimising perceptual loss functions based on high-level characteristics retrieved from pretrained networks. In order to train feed-forward networks for image modification tasks, we propose combining the advantages of both methods and using perceptual loss functions.We demonstrate results on image style transfer, where a feed-forward network is trained to quickly and accurately address the optimisation challenge stated by Gatys et al. Although our network is three orders of magnitude faster than the optimization-based method, it nonetheless produces results of a similar quality. We also test single-image super-resolution, where a perceptual loss is substituted for a per-pixel loss to provide visually appealing outcomes.

In 2016, The paper titled "Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, Chuan Li, Michael Wand"[4].In order to train generative networks effectively for texture generation, this research suggests Markovian Generative Adversarial Networks (MGANs). Even while deep neural network methods have recently shown impressive results in terms of synthesis quality, they still require a significant amount of processing resources (minutes of run-time for low-res photos). This efficiency problem is addressed in our paper. We precompute a feed-forward, strided convolutional network that can directly produce outputs of any size and captures the feature statistics of Markovian patches in place of the numerical deconvolution used in earlier research. Such a network may convert images, such as pictures, immediately from brown noise to realistic texture or artistic paintings. We achieve quality that is on par with more current neural texture creation techniques using adversarial training. Our run-time performance (0.25 M pixel pictures at 25 Hz) is noticeably better than that of earlier neural

texture synthesisers (at least 500 times faster), as no optimisation is needed at the creation stage. This concept is used in the synthesis of textures, the transfer of styles, and video stylization.

In 2016, The paper titled "Texture networks: Feed-forward synthesis of textures and stylized images. D.Ulyanov, V. Lebedev, A.Vedaldi, V. Lempitsky"[5]. Deep networks may create stunning textures and stylised images from a single texture example, as Gatys et al. recently shown. However, the optimisation procedure for these solutions is cumbersome and slow. Here, we provide an alternate strategy that shifts the computational burden to a stage of learning. Our method trains compact feed-forward convolutional networks to generate several samples of the same texture of any size from a single texture example and to transfer the artistic style of one image to any other image. The resulting networks are incredibly lightweight and are hundreds of times faster than Gatys et al. at producing textures of equal quality. In a broader sense, our method emphasises the strength and adaptability of generative feed-forward models trained using expressive and complicated loss functions.

In 2017, The paper titled "Demystifying Neural Style Transfer, Yanghao Li, Naiyan Wang, Jiaying Liu, Xiaodi Hou"[6].Recent research on neural style transfer has produced some very intriguing findings that have caught the attention of both academics and business. Despite the incredible results, it is still unclear how neural style transfer works, particularly how and why Gramme matrices can be used to describe style. In this study, we analyse neural style transfer as a domain adaptation problem and provide a fresh interpretation of neural style transfer. In particular, we theoretically demonstrate that minimising the Maximum Mean Discrepancy (MMD) with the second order polynomial kernel is equivalent to matching the Gramme matrices of feature maps. We contend that matching the feature distributions between the style images and the generated images is the heart of neural style transfer. We test a number of additional distribution alignment techniques and provide promising findings to support our position even further. This innovative view, in our opinion, bridges these two crucial domains of study and may shed insight on further investigation.

In 2017, The paper titled "A Learned Representation For Artistic Style, 2017 - Vincent Dumoulin, Jonathon Shlens,ManjunathKudlur"[7].The variety of painting techniques offers a wide visual language for the creation of im ages.The extent to which one can comprehend and learn this visual lexicon gauges our knowledge of the more comp lex elements of paintings, if not of images in general.In this research, we look into how to build a single, scaleable d eep network that can accurately capture the artistic sensibility of a variety of paintings.By reducing a painting to a p oint in an embedding space, we show that such a network generalises across a variety of artistic genres. Importantly, this model enables the user to randomly combine the styles discovered from separate paintings in order to experiment with new painting techniques. We hope that our study serves as an important first step in developing detailed models of paintings and opens a window into the organisation of the learnt representation of creative style.

Base paper "Cartoonify Realistic Images and Videos Using OPENCV" -[8]. In today's digital age, we often maintain a continuous online presence where we share photos and comments. To create an attractive profile, we can opt to use our personal photo as a profile picture, create a fun avatar, or even convert our image into a cartoon. This document describes a method for transforming real images and videos into a cartoon format. The main goal of this research is to investigate the feasibility of employing a Generative Adversarial Network (GAN) along with two loss functions: content loss and adversarial loss. This combination aims to produce sharp and clear images during the process of converting realistic images and videos into cartoon versions, with the assistance of K-means clustering. This article includes a comparative analysis of the performance of models developed using deep learning-based generative model architecture and K-means clustering. Additionally, the system evaluates these parameters for both algorithms, with histogram operations conducted in the Python programming language using Convolution blocks and clustering techniques.

Base paper. "Cartoonization of images using machine Learning in 2022" [9]. Image processing entails the application of methods to modify images for improvement or the extraction of valuable information. This falls under the realm of signal processing, where images serve as input, and the output can either be another image or characteristics and features associated with the original image. Key tools for image processing include OpenCV and Scikit-Image, which are Python libraries, as well as NumPy, a library for scientific computing. Additionally, Generative Adversarial Networks (GANs) are commonly employed to learn from extracted representations and generate new images. The primary objective of this approach is to enhance control and adaptability. Generative modeling, an unsupervised machine learning task, focuses on automatically identifying patterns in input data and using these patterns to create plausible new examples. Common image processing algorithms encompass

morphological and mathematical techniques, Fourier transforms, edge detection, and Convolutional Neural Networks (CNNs).

In 2020, The paper titled "Y. Chen, Y.-K. Lai, Y-J. Liu. "cartoonization using white box representation", International Conference on Image Processing," [10]. Image processing involves the application of various operations to manipulate images for enhancement or to extract valuable information. This falls under the realm of signal processing, where images serve as input and the output can be either an improved image or relevant image features. Popular tools for image processing include OpenCV, Scikit-Image, and NumPy, each serving different aspects of the task. Additionally, Generative Adversarial Networks (GANs) are commonly utilized for learning and subsequently generating images that resemble the original dataset. Generative modeling, a form of unsupervised learning, focuses on discovering patterns within input data, enabling the model to create new data examples that align with the dataset. OpenCV and NumPy are fundamental libraries used for real-time computer vision, image processing, and efficient numerical computations, respectively. Image processing algorithms encompass various techniques, including morphological operations, mathematical methods, Fourier transforms for frequency domain analysis, edge detection, and the utilization of Convolutional Neural Networks (CNNs) for advanced image tasks

In 2020, The paper titled "Learning to Cartoonize Using White-box Cartoon Representations,"[11].The approach is based on insights gained from observing how cartoon paintings are created and consulting with artists. It involves separating images into three distinct representations: the smooth surface, which captures the overall cartoon image appearance, the structure representation, which emphasizes the simple color blocks and flattened content in a cartoon style, and the texture representation, which highlights the fine details and textures in cartoon images. To achieve this transformation, the authors employ a Generative Adversarial Network (GAN) framework to learn and create cartoonized images.The learning objectives of this method are tailored to each of these extracted representations, providing control and adjustability. This adaptability allows the approach to cater to different artistic styles and various use cases. To validate the effectiveness of their approach, the authors conducted qualitative and quantitative comparisons and user studies, all of which demonstrated that their method outperforms previous techniques. Additionally, they conducted an ablation study to assess the impact of each component within their framework.

In 2017, The paper titled "Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle consistent adversarial networks. " [12].Image-to-image translation is a category of problems in the fields of vision and graphics. Its primary aim is to understand how to convert an input image into an output image by learning from a dataset of matching image pairs. However, in many cases, it's not possible to have such paired training data. We introduce a method for learning how to transform an image from one domain (X) to another (Y) even when we don't have paired examples. Our objective is to develop a mapping function, denoted as G: $X \rightarrow Y$, so that the images produced by G(X) are indistinguishable from those in Y, as assessed by an adversarial loss. Because this mapping can be highly uncertain, we also introduce an inverse mapping, F: $Y \rightarrow X$, and add a cycle consistency loss to ensure that F(G(X)) approximates X and vice versa. We've conducted qualitative evaluations on various tasks that lack paired training data, such as style transfer, object manipulation, seasonal changes, and photo enhancement. Furthermore, we've performed quantitative comparisons against several existing methods, showcasing the superiority of our approach.

3.CONCLUSIONS

In conclusion, this paper is about the area of digital artwork and image transformation which has been an important advancements with the fusion of deep learning techniques and the OpenCV library in Python. The venture mentioned right here gives a singular method to transform regular pix into captivating cartoons. It leverages the power of Convolutional Neural Networks (CNNs) for part detection to capture the different outlines and shapes of cool animated film drawings, observed with the aid of a shade quantization procedure to simplify the colour palette, paying homage to traditional caricature patterns. This technique draws suggestion from the wider area of neural style switch, as highlighted in numerous research papers. These papers have explored the interplay between content material and style in pictures and feature established the potential of deep mastering in generating creative and perceptually attractive visuals. This project offers a comprehensive framework for transforming images into captivating cartoons using deep learning and OpenCV in Python. The proposed method not only simplifies complex

images but also allows for creative customization, making it a valuable tool for artists, designers, and enthusiasts seeking to infuse a touch of whimsy into their digital creations.

4. REFERENCES

[1]. A Neural Algorithm of Artistic Style, 2016 - Leon A.Gatys, Alexander S. Ecker, Matthias Bethge
[2]. Image Style Transfer Using Convolutional Neural Networks, 2016 - Leon A. Gatys, Alexander S. Ecker, Matthias Bethge
[3]. Perceptual Losses for Real-Time Style Transfer and Super-Resolution, 2016 - Justin Johnson, Alexandre Alahi, Li Fei-Fei
[4]. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, 2016 - Chuan Li, Michael Wand
[5]. Texture networks: Feed-forward synthesis of texturesand stylized images, 2016 - D. Ulyanov, V. Lebedev, A. Vedaldi, V. Lempitsky
[6]. Demystifying Neural Style Transfer, 2017 - YanghaoLi, Naiyan Wang, Jiaying Liu, Xiaodi Hou
[7]. A Learned Representation For Artistic Style, 2017 -Vincent Dumoulin, Jonathon Shlens, Manjunath Kudlur
[8].Cartoonify Realistic Images and Videos Using OPENCV
[9]. Cartoonization of images using machine Learning
[10]Y. Chen, Y.-K. Lai, Y-J. Liu. "cartoonization using white box representation", International Conference on Image Processing, 2020
[11] Learning to Cartoonize Using White-box Cartoon Representations,2020
[12] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle consistent adversarial networks. 2017,In Proceedings of IEEE