International Journal of Research in Engineering, Science and Management
Volume 4, Issue 7, July 2021
https://www.ijresm.com | ISSN (Online): 2581-5792

Cartoonizer: Convert Images and Videos to Cartoon-Style Images and Videos

S. Rajatha[1*], Anusha Shrikant Makkigadde[2], Neha L. Kanchan[3], Sapna[4], K. Janardhana Bhat[5]
[1,2,3,4]*Student, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*
[5]*Associate Professor, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*

***Abstract*: The process of converting real-life high-quality pictures and videos into practical cartoon images and videos is known as cartoonization. The saved model decomposes uploaded images and videos into three different cartoon depictions as surface representation, structure representation, texture representation, which further instructs the network optimization to generate cartoon image. It helps to sleek the image, filter the qualities, transforming it to sketches, and translating the output from a domain to another. The extracted outputs are fed to a Generative Neural Networks (GAN) framework, which helps to improve our problem making the solution more flexible and varied, where GAN stands for Generative Adversarial Network is used to transform uploaded images (snapshots) to the finest cartooned image. Using the loss function and its two types named as Adversarial loss and Content Loss, we gained a flexible as well as a clear edge defined images.**

***Keywords*: Animation, Generative Adversarial Network (GAN), Image processing, Video processing.**

1. Introduction

Cartoons are frequently used in various kinds of implementations. Cartoon is an artwork that has been widely applied in diverse sections. Cartoons are artistically made and it requires effective and clear human artistic skills. Modern cartoon animation artists make use of collection of sources to create content. While sketching cartoons in an animated movie, it gets time-consuming for the artist as they need to define the sketch of the cartoon properly to get fine output. As we know that animation plays a vital role in the modern world of entertainment. To resolve the problem faced by artists we have created a program with the help of GAN which converts both images and videos into an animated image and video. This procedure is called as cartoonization. The variety of cartoon arts require context-specific assumptions and prior knowledge to develop algorithms. For instance, some cartoon arts concentrate more on global palette themes and some concentrate on scarce and clean color blocks in imaginative expression. To assure all the requirements of converting real images, we have taken help of Generative Adversarial Network (GAN). This consumes less time to convert a real image to a cartoon image. Not only the images can be converted, we can also transform the real videos into cartoon one. Hence, not only time is saved but also fine work is generated within an enough amount of time. This will offer a great opportunity to the animation industry to create more animated movies and clips.

2. Literature Review

[1] It observes the cartoon painting behaviour and consulting artists, it proposes to separately identify three white-box representations from images: the surface representation that contains a smooth surface of cartoon images, the structure representation that refers to the scarce color-blocks and raze global content in the celluloid style workflow, and the texture representation that reflects high frequency texture, contours, and details in cartoon images. A Generative Adversarial Network (GAN) framework is used to learn the extracted representations and to cartoon images. The learning objectives of this method are separately based on each extracted representations. This enables the approach to meet artists' requirements in different styles and diverse use cases. Qualitative comparisons and quantitative analysis, as well as user studies, have been conducted to validate the effectiveness of this approach. At the end, the ablation study demonstrates the influence of each component in the framework.

[2] In this paper a solution to transforming photos of real-world scenes into cartoon style images, which is valuable and quite challenging in computer vision and graphics. This solution belongs to learning based methods, which have recently become popular to styling images in artistic forms such as painting. However, existing formats do not make satisfactory outputs for cartoonization, due to the fact that cartoon arts have individual characteristics with high level simplification and abstraction, and cartoon images tend to have clear borders, smooth color pigmentation and relatively simple textures, which exhibit significant challenges for texture descriptor-based loss functions used in existing methods. The paper proposes Cartoon GAN, a generative adversarial network (GAN) framework for cartoonization. The model takes photos and cartoon images that are unpaired for training, which is easy to use. Two novel losses suitable for cartoonization are proposed as a semantic content loss, which is evaluated as a sparse regularization in the high-level feature maps of the VGG

*Corresponding author: rajathasalian1999@gmail.com

network to cope with considerable style variation between photos and cartoons, and uses an edge-promoting adversarial loss for preserving clear edges. It then initiates an initialization phase, to improve the convergence of the network to the target manifold. This method is also more efficient to test and train than existing methods. Experimental results prove that this model is able to generate high quality cartoon images from real-world images (i.e., following specific artists' styles and with clear edges and smooth shading) and outdraws state-of-the-art methods.

[3] Auto-painter: Cartoon image generation from sketch by using conditional generative adversarial networks (GAN). Recently, real-life image generation using deep neural networks has become a trending topic in machine learning and computer vision. Images can be generated at the pixel level by learning from a large cluster of images. Learning to produce colorful cartoon images from black-and-white sketches is not only an interesting research problem, but also a potential application in digital application. This paper refers to the sketch-to image synthesis problem by using conditional generative adversarial networks (cGAN). It proposes the auto-painter model which can automatically generate compatible colors for a sketch. The new model is not only capable of sketching hand-draw sketch with appropriate colors, but also allowing users to indicate preferred colors. Experimental results on two sketch datasets show that the auto-painter performs better that existing image-to-image methods.

[4] Semantics-driven portrait cartoon stylization: This paper proposes a systematic framework for transforming an input human portrait image into an artistic cartoon style. Compared to the previous work of non-naturalistic rendering (NPR), this method exploits the portrait semantics for enhancing and manipulating the cartooning art, based on a semantic grammar model. The referred framework consists of two phases such as a portrait parsing phase to constrain and recognize facial features in a hierarchic manner, and further calculate the portrait saliency with the facial features and a cartoon stylizing phase to abstract and cartoonize the portrait according to the parsed semantics and saliency, in which the corners and structure (edges/boundaries) of the portrait are rendered in two phases. In the experiments, they test the model with different varieties of human portraits like daily photos, and studio photos, identification photos and find satisfactory results. A quantitative estimation of subjective preference is presented as well.

[5] Generative adversarial nets: This paper proposes a new framework for approximating generative models by way of an adversarial process, in which they simultaneously train two models such as, G a generative model that captures the data distribution, and D a discriminative model that approximates the probability that a sample came from the training data rather than G. The training method for G is to maximize the probability of D making a mistake. This framework corelates to a minimax two-player game. In the space of arbitrary functions G and D, a unique result exists, with G recovering the training data distribution and D equal to 12 everywhere. In the case where G and D are defined by multilayer perceptron, the overall system can be trained with backpropagation. Experiments reveals the potential of the framework through quantitative and qualitative evaluation of the generated samples.

[6] Automatic cartoon colorization based on convolutional neural network: This paper proposes with automated cartoon colorization. This is a tough issue, since it is an ill posed problem that usually needs user intervention to achieve superior quality. Prompted by the recent successes in natural image colorization based on deep learning techniques, they inquired the colorization problem at the cartoon domain using Convolutional Neural Network. To the best observation, no existing studies or research papers address this problem using deep learning techniques. Here they inquired a deep Convolutional Neural Network based automated color filling method for cartoons.

3. Proposed Method

Generative Adversarial Networks (GAN) is an algorithm used for the implementation of cartoonization. GAN consists of two phases i.e., a generative model and a discriminative model. The generative model produces new instances of data that features the training data. The discriminator is the model used for testing the featured data and comparing it with the image from the Generator. Discriminator decides whether the output image is real or fake. These generator and discriminator both are neural networks and they run in contention with each other in the training phase. The steps are repeated numerous times so that the generator and discriminator both get a better output after the repetition of steps.

Now coming to the project, images are decomposed into three representations the surface representation, the structure representation, and the texture representations, and to extract corresponding representations three independent modules are introduced. A GAN framework with a generator G and two discriminators namely Ds aims to distinguish between surface representation extracted from model outputs and cartoons, and Dt is used to distinguish between texture representation extracted from outputs and cartoons. Weight for each component can be tuned in the loss function, which allows users to control the result style and adapt the model to diverse different use cases.

The unique trend point was putting together an architecture which would allow us to work for videos along with images at minimal cost. Videos are nothing but a collection of images in bits or frames and we have to cartoonize each frame. We have used two techniques in our project to bring down the video inference time. Bring down the resolution of image to 480p. This essentially decreased the load per frame without any noticeable change in the quality. This made it reach 1 second/image inference time. Decrease the frame rate of the video, video was downplayed from 30 fps to 15 fps which extremely reduced our video computation time. Even while downplaying resolution and reducing the video frames per second, cartoonization was taking 2.5 minutes for a 10 second video. It was beyond normal considering user experience. And hence, converting a video into a cartoon was quite difficult experience.

4. Implementation

Fig. 1 represents Flowchart for cartoonizer webapp. Firstly, the user clicks on the button select the image or video. After the system opens window for local storage. System checks for image/video formats allowed. Then the selected image or video will be uploaded to convert into cartoon image/video. After the successful upload of image/video the system processes input and produces the output. After the cartoonized output gets displayed on screen, the user may download the image/video by clicking on download button. Downloaded result will be stored into the local storage. Finally, the user can come out of the system.

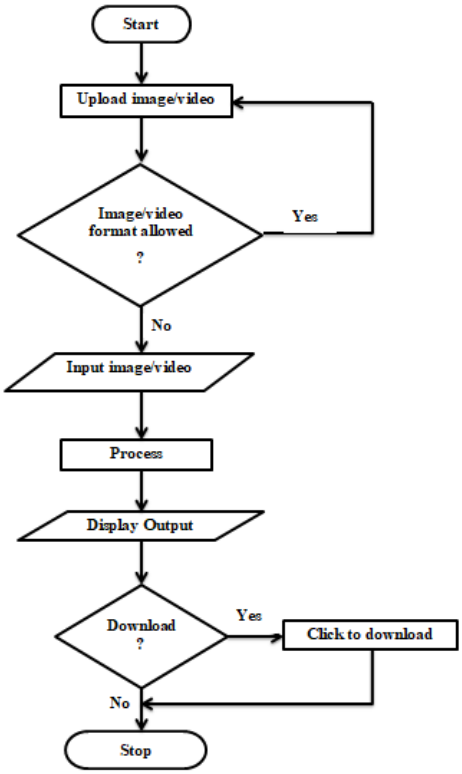

Fig. 1. Flowchart for workflow of webapp

For videos, the video being uploaded must have size less than or equal to 30mb. If the video is longer than 15sec, then it is trimmed to 15sec and converted to cartoon-style video. Audio will be added to the cartoonized video. First 15sec of the video will be considered.

5. Results

The below represents the project results, the first image consists of the real-life image and later after using Cartoon GAN with the help of OpenCV we get the cartoonized image as below. Screenshots of original video and cartoonized video at same time is also given.

Fig. 3. Real and Cartoonized image of a cat

Fig. 4. Screenshots of real and cartoonized video while playing at time 00:12

6. Conclusion

In this paper, we proposed a white-box controllable image cartoonization framework with the help of GAN, which can generate high-quality cartoonized images and videos from real-world photos and videos. Images are decomposed under three cartoon representations surface, structure and texture representation. Image processing modules such as openCV, PIL,etc are used to extract three representations for network training, and result styles could be controlled by adjusting the weight of each representation. Extensive qualitative and quantitative experiments, as well as user studies, have been conducted to validate the performance of the method.

7. Future Work

In future work we would like to focus more on converting longer videos, high quality videos/images and bigger file size. We would also like to work on adding feature that enables user to share the result on various social platforms.

References

[1] Xinrui Wang, Jinze Yu, “Learning to Cartoonize Using White-box Cartoon Representations”, Byte Dance, The University of Tokyo, Style2Paints Research.

[2] Y. Chen, Y.-K. Lai, Y.-J. Liu, "Cartoon GAN: Generative Adversarial Network for photo cartoonization", International Conference on Image Processing, 2018.

[3] Zengchang Qin, Zhenbo Luo, Hua Wang, " Auto-painter: Cartoon Image Generation from Sketch by Using Conditional Generative Adversarial Networks”, International Conference on Image Processing, 2017.

[4] Ming Yanga, Shu Lina, Ping Luoa, Liang Lina, Hongyang Chao, “Semantics-Driven Portrait Cartoon Stylization”, Sun Yat-Sen University, Guangzhou 510275, PR China Lotus Hill Research Institute, Wuhan 430074, PR China.

[5] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversarial Nets.”

[6] Domonkos Varga, “Automatic Cartoon Colorization Based on Convolutional Neural Network,” CBMI '17 Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, June 2017, pp. 1–6.