Real – Life to Cartoon Image Using Meanshift Based Colour Unification Method

B SRINISHA,N KRISHNAMMAL
Sri Shakthi Institute of Engineering and Technology,
Coimbatore. Tamil Nadu, India,
srinishabalan@gmail.com

Abstract-**Recently, realistic image generation exploitation image process has become a hot topic in machine learning and laptop vision. Realistic pictures may be generated at the element level by applying some image process operations. Learning to come up with colourful cartoon pictures from traditional image isn't solely a remarkable analysis downside however conjointly a possible application in digital recreation. In this work, we have a tendency to investigate the important life image-to-Cartoon Image synthesis drawback by exploitation Selective Gaussian Filter (SGF) and Mean Shift Cluster operation. The new model isn't solely capable of generating cartoon pictures however additionally permits users to point most well-liked filter levels. Experimental results show that the planned work performs higher than existing image-to-cartoon image ways.**

Keywords- Selective Gaussian Filter (SGF), Mean Shift Cluster, Color unification method, Cartoonify image.

I. INTRODUCTION

The image method is also a method for applying operations on an image to enhance it or extract relevant data from it. It's a form of a signal method at intervals additionally the input among the image and conjointly the output is either an image or the image's characteristics/features. Image process is one among the foremost quickly evolving technology nowadays. It's conjointly a serious analysis topic within the fields of engineering and engineering science. Analog and digital image process area unit the 2 sorts of image process strategies used. Arduous copies, like prints and pictures, will take pleasure in analogue image process. Once using these visual tools, image analysts use a spread of informative fundamentals. Digital image method techniques facilitate in a manipulation of the digital photos by exploitation computers. The 3 general phases that everyone forms information ought to bear whereas an exploitation a digital technique's an area unit is preprocessing, improvement, and show information extraction. During this lecture, we are going to point out many basic definitions like image, digital image, and digital image process. Completely different supplies of digital pictures are going to be mentioned and examples for every source are going to be provided. Further, the time from image process to pc vision are going to be coated during this lecture. Finally, we are going to discuss image acquisition and differing types of image sensors.

Grayscale may be a vary of monochromatic shades from black to white. Therefore, a Grayscale image contains solely reminder gray and no color. In image method filters area unit principally accustomed suppress either the high frequencies among the image, i.e. smoothening the image or the low frequencies, enhancing or investigating edges inside the image. A picture is filtered either within the frequency or within the spatial domain. In digital image process, threshold is that the simplest methodology of segmenting pictures. From a grayscale image, threshold is accustomed produce binary pictures. Automatic threshold may be a good way to extract helpful data encoded into pixels whereas minimizing background signal. This can be accomplished by utilizing an electrical circuit to optimize the edge price before changing the initial grayscale image to binary in such some way that the image process approaches area unit used for numerous image level operations. In this time of technology, Cartoons aren't drawn manually. So, we have to convert the images into cartoons easily which helps us to include real-life people in magazines or comics to entertain readers. In order to

accomplish this task image processing approaches are used according to the specific order. We produced cartoonified images by using the machine learning algorithm which produced clear and sharp images using Python.

Fan, Y.P. et al., 2009 generated a method for converting the digital images into cartoonified images to produce cartoon movies, by using mean shift filter, Difference of Gaussian filter, and filtered image and edge map are fused together. Patankar, A.B., et al., 2016 produced a cartoonized image for both the grayscale and colored images by using some algorithms like LBG, KPE, and KMCG. Sikchi, P., et al., 2014 proposed a device using Image processing techniques such as bilateral filtering, contrast stretching, edge detection, and luminance quantization were used to create cartoonified images using Raspberry Pi and OpenCV. Hasan, R., et al., 2017 proposed a model to detect objects and their details and replace them with cartoonified images. Joshi, R. et al., 2021 produced animated images by analyzing the behaviour of surface representation, structure representation, and texture representation with the help of the Generative Adversarial Network (GAN) framework. Patankar, A. et al., 2014 introduced a non-photorealistic rendering technique by using vector quantization and line extraction algorithms along with gradient edge detection method to deliver the cartoon effect. Chen, Y., et al., 2020 proposed a CartoonRenderer framework that uses a single trained model to produce multiple cartoon styles. In nutshell, the proposed Soft-AdaIN method maps the photo into a feature model and renders the feature model back into image space to achieve a high-quality cartoonified image. Rajput, P.S., et al., 2021 designed a Pix2Pix model which trains the model to generate cartoonified images from the real life images. Chen, Y., et al., 2018 proposed CartoonGAN which is easy to use and can generate high quality cartoon images from the real world images. Zhu, J.Y., et al., 2017 proposed a CycleGAN method to cartoonify the images by mapping the images using adversarial loss and coupled with the inverse mapping.

In the proposed work, the image-to-Cartoon Image synthesis problem by using selective Gaussian Filter (SGF) and Mean Shift Cluster is investigated. The new model isn't solely capable of generating cartoon pictures, however additionally permits users to point most well-liked filter levels. The proposed work contributes towards the improvement of the same by the following highlights.

- High customization can be made in the image to cartoon image generation
- The information extraction is high
- Improved cartoonifing process.

In this paper, the proposed method had been documented in upcoming section and it have been experimented with benchmark datasets to prove the mean shift based colour unification model. Finally, the results are concluded in the result analysis section.

II. METHODOLOGY

In this paper, we examined the real life image-to-cartoon Image synthesis problem by using Selective Gaussian Filter (SGF) and Mean Shift Cluster operation. In this paper, the color unification method is proposed in detail.

In this method, the color unification model has been developed based on the mean shift. Input images are given into the model and the Gaussian filter blurs the input image by using the Gaussian function, which also known as Gaussian smoothing. Where the adaptive threshold based edges automatically finds the threshold value of the image and also selects the edges which are stronger than the threshold. Mean shift clustering groups the data points have a high density in a region and adds up all the individual kernels generate a probability surface depending on the kernel bandwidth parameter and perform a masking. After masking,

we obtain the cartooned image. Figure 1 shows the system architecture of the proposed colour unification system.

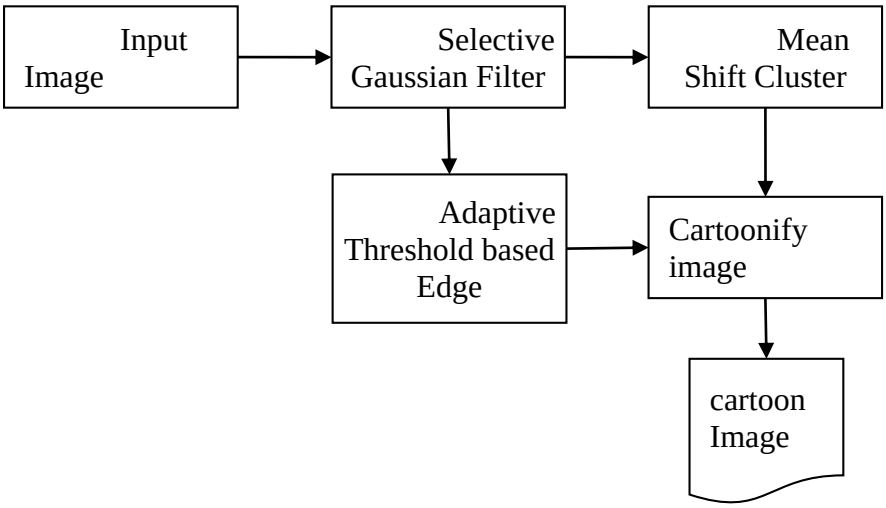

Figure 1: System Architecture of the proposed Color Unification method

A. Selective Gaussian Filter

In a Gaussian pyramid, successive images are weighed and measured using a Gaussian blur (Gaussian average). Each pixel with a local average corresponds to a neighbouring pixel at the bottom of the pyramid. This technique is used exclusively in texture synthesis. In image process, Gaussian blur (also called Gaussian softening) is that the results of blurring a picture by a Gaussian perform. This is a wide used impact in graphics software system, sometimes to cut back image noise and detail. The visual result of this blurring a technique may be a swish blur the same as viewing a picture through a clear screen, as distinct it's totally different from the bokeh result to create by out-of-focus lenses or the shadow of an associate degree object below a traditional illumination. The Gaussian smoothing is additionally used as a pre-processing step in laptop vision algorithms to reinforce image structures to totally different sizes. The operation of the selective Gaussian blur is defined as

$$G(d,\sigma) = \frac{1}{(2\pi\sigma)^{n/2}} e^{-\frac{x^2}{2\sigma^2}} \quad (1)$$

Where, σ is that the standard deviation of the Gaussian, d is that the degree of blurring and n is the number of dimensions. The given input image has processed with this Gaussian blur operation in order to reduce the micro level textured information.

B. Adaptive Threshold based Edge

A sensitivity threshold, nominal as a numeric scalar for any methodology, or two-element vector for the 'Canny' and 'approxcanny' ways that. Edge ignores all edges that aren't stronger than threshold. For more information about this parameter, in the figure 1, the filtered image by Selective Gaussian filter.

If you are doing not to specify a threshold, or if you specify the associate degree an empty array ([]), then edge chooses the worth or values mechanically by exploitation the 'log' and 'zero

cross' strategies. If you specify the edge worth zero, then the output image has closed contours as a result of it includes all the zero-crossings within the input image.

The 'Canny' and 'approxcanny' ways use two thresholds. The edge disregards all edges with edge strength below the lower threshold, and preserves all edges with edge strength on high of the higher threshold. You'll be able to specify a threshold as two-element vector of the shape [low high] with low and high values within the vary [0 1]. You'll be able to conjointly specify a threshold as a numeric scalar. But, in the experiment automotive evaluation of threshold value by using zero cross method has adaptively calculated. So the Gaussian blurred image further involved with this adaptive edge detection process. Thus, this extracted edges also utilized for cartooning the image in future.

C. Mean Shift Cluster

Meanshift is falling beneath the class of a agglomeration rule in distinction to unattended learning that assigns the information points to the clusters iteratively by shifting points towards the mode. As such, it's additionally referred to as the Mode-seeking algorithmic program. Mean-shift algorithmic program has applications within the field of image process and pc vision. In contrast to the favour K-Means cluster algorithmic program, mean-shift doesn't need specifying the amount of clusters before. The amount of clusters is set by the algorithmic program with relevancy the information. Mean-shift builds upon the conception of kernel density estimation is kind KDE. Here in the cartoonify process, it is applicable that number of clusters to be finalized based on the image not by the prior assignment process. Imagine that the pre-processed image data was sampled from a probability distribution. KDE could be a technique to estimate the underlying distribution additionally referred to as the likelihood density perform for a collection of knowledge. It works by putting a kernel on every purpose within the knowledge set. There area unit many alternative styles of kernels, however the foremost well-liked one is that the mathematician kernel and it's used here. Adding up all of the individual kernels generates a likelihood surface example density perform. counting on the kernel information measure parameter used, the resultant density perform can vary. So, the given image knowledge are clustered supported the dynamic cluster range and crammed with average distribution worth among the cluster region.

D. Cartoonify Image

Let’s mix the 2 specialties. This may be done victimization MASKING. We have a tendency to perform bitwise and on 2 pictures to mask them. Remember, pictures AR simply numbers. Therefore that’s however we have a tendency to mask edged image on our “BEAUTIFY” image. This finally CARTOONIFY our image!

The on top of strategies have shown the method of cartoonification of the input victimization the Selective Gaussian filter, adaptation threshold primarily based edge and mean shift cluster. By utilizing the tactic, the experimental analyses are performed and comparative results AR shown within the following sections…

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. DATASET DESCRIPTIONS

In this paper, we had included the images from all the three datasets that are ETH Synthesizability dataset, ImageNet dataset and Flickr-Faces-HQ (FFHQ) dataset. ETH Synthesizability dataset (Dai, D et al., 2014) which include 21,302 texture synthesis samples.

According to WordNet hierarchy, ImageNet dataset (Wu, H., et al., 2020) contains 14,197,122 image samples. FFHQ dataset (Karras, T., et al., 2019) is a high quality image dataset of human faces contains 70,000 PNG images at 1024×1024 resolution and contains variations in image background. In FFHQ dataset images are crawled from Flickr and it was automatically aligned and cropped by using dlib.

B. EXPERIMENTAL SETUP

This implementation has been carried out on a PC with a 2.3 GHz Intel Core I3, 4 GB RAM, 1 TB Hard Disk and Windows 7 operating system. Using Spyder IDE which allows us to write and execute the code in python. It also allows us to import image processing packages in order to cartoonify the images.

C. RESULT ANALYSIS

We implemented our proposed cartoonifing images using the mean shift based color unification method. In this paper, we had pre-processed the images using Selective Gaussian filter which adds the blur effect to the images using Gaussian function. Adaptive threshold based edge discards all the edges that has weaker threshold and selects the edges that has stronger threshold in the range of [0 1], using the canny and approx canny methods. Mean shift clustering algorithm assigns the data points and clusters the highest density of the data point in the region. Finally, cartoonifies the image using masking. In this paper, we had included the images from all the three datasets that are ETH Synthesizability dataset, ImageNet dataset and Flickr-Faces-HQ (FFHQ) dataset. ETH Synthesizability dataset has 21,302 texture samples. According to WordNet hierarchy, ImageNet has 14,197,122 images. FFHQ high quality dataset has 70,000 PNG images at 1024×1024 resolution and contains variations in image background. Table 1 shows the number of images presented in the datasets. Figure 2 shows the results of gray scale image, threshold image, Gaussian blur, mean shift cluster, cartoonify image along with its input image.

Table 1: Number of images presented in the datasets.

Dataset	No. of images
ETH Synthesizability dataset	21,302
ImageNet dataset	14,197,122
Flickr-Faces-HQ dataset	70,000

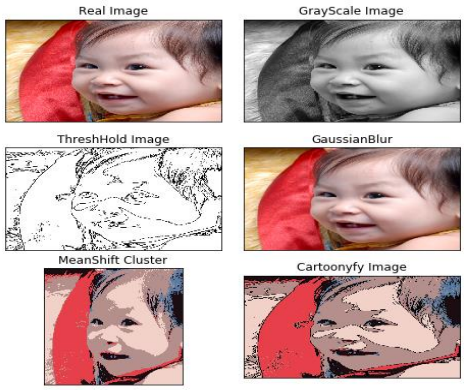

Figure 2: Results of gray scale image, threshold image, Gaussian blur, mean shift cluster, cartoonify image along with its input image.

Table 2: Statistical comparison with the different methods on the mean FID score calculation

Metrics	Nature texture	CartoonGAN (Chen, Y., et al., 2018)	CycleGAN (Zhu, J.Y., et al., 2017)	Proposed method

FID to cartoon texture	182.32	157.09	143.22	111.52
FID to Nature texture	100.11	85.73	69.54	33.32

Table 2 shows the statistical comparison with the different methods on the mean FID score calculation. We also performed the quantitative comparisons for all CartoonGAN (Chen, Y., et al., 2018), CycleGAN (Zhu, J.Y., et al., 2017), and proposed methods by collecting the mean frechet inception distance (FID) value, which is commonly used to calculate the distance between feature vector for both the source image distribution and target image distribution. Figure 3 shows the graphical analysis of performance measure among all three methods. As shown in Table 2 and Figure 3, we can also observe that our method generally achieves the lowest mean FID value, clearly indicating the superiority of meanshift based colour unification method in image cartoonization.

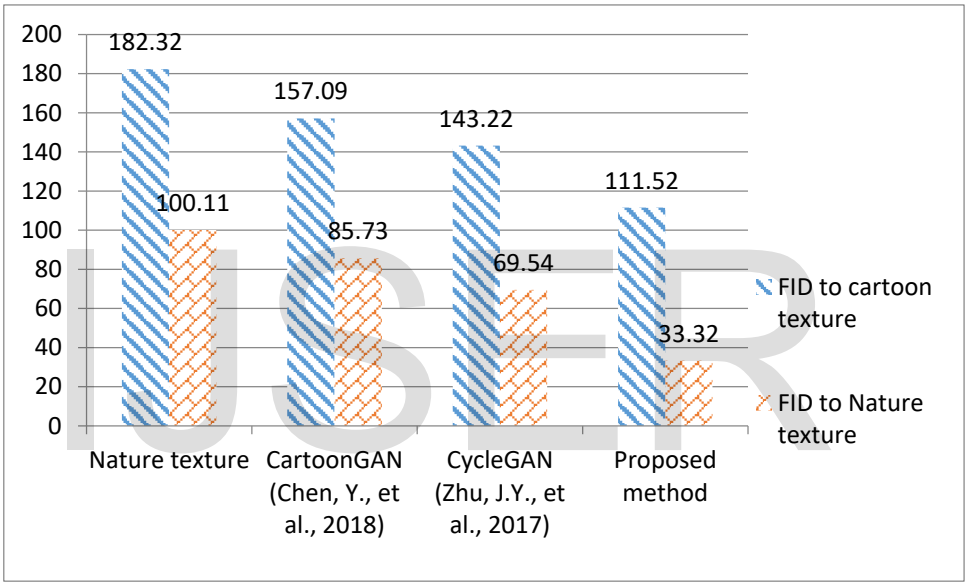

Figure 3: Graphical analysis of performance measure among all three methods.

IV. CONCLUSION

In this work, we investigate the image-to-Cartoon Image synthesis problem by using Selective Gaussian Filter (SGF) and Mean Shift Clustering. The new model is not only capable of generating cartoon image, but also allowing users to indicate preferred filter levels. Experimental results show that the proposed work performs better that existing image-to-image methods images are just numbers. therefore that's however we have a tendency to mask edged image on our "BEAUTIFY" image. This finally CARTOONIFY our image with high FID score.

REFERENCES

[1] Fan, Y.P. and Huang, S.X., 2009. A New Method for Image Cartoonization Based on Mean-Shift and DOG. Coal Technology, 28, pp.160-161.

[2] Hasan, R., Shaffer, P., Crandall, D. and Apu Kapadia, E.T., 2017. Cartooning for enhanced privacy in lifelogging and streaming videos. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 29-38).

[3] Patankar, A.B., Kubde, P.A. and Karia, A., 2016, August. Image cartoonization methods. In 2016 International Conference on Computing Communication Control and automation (ICCUBEA) (pp. 1-7). IEEE.

[4] Chen, Y., Chen, M., Song, C. and Ni, B., 2020, January. CartoonRenderer: An Instance-based Multi-Style Cartoon Image Translator. In International Conference on Multimedia Modeling (pp. 176-187). Springer, Cham.

[5] Chen, Y., Lai, Y.K. and Liu, Y.J., 2018. Cartoongan: Generative adversarial networks for photo cartoonization. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 9465-9474).

[6] Dai, D., Riemenschneider, H. and Van Gool, L., 2014. The synthesizability of texture examples. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3027-3034).

[7] Joshi, R. and Asoba, S., 2021. CONVERSATION OF REAL IMAGES INTO CARTOONIZE IMAGE FORMAT USING GENERATIVE ADVERSARIAL NETWORK.

[8] Karras, T., Laine, S. and Aila, T., 2019. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4401-4410).

[9] Patankar, A. and Tawde, P., 2014. Cartoonization Using LBG And KPE Vector Quantization.

[10] Rajput, P.S., Satis, K., Dellarosa, S., Huang, W. and Agba, O., 2021. cGANs for Cartoon to Real-life Images. arXiv preprint arXiv:2101.09793.

[11] Sikchi, P., Beknalkar, N. and Rane, S., 2014. Real-Time Cartoonization Using Raspberry Pi. IJCAT International Journal of Computing and Technology, 1(6).

[12] Wu, H., Yan, W., Li, P. and Wen, Z., 2020. Deep texture exemplar extraction based on trimmed T-CNN. IEEE Transactions on Multimedia.

[13] Zhu, J.Y., Park, T., Isola, P. and Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision (pp. 2223-2232).