Cartoonify the Real-Life Image Based on Mini Batch Color Clustering

Ms.B.Srinisha ,Mrs.N.Krishnammal,
Sri Shakthi Institute of Engineering and Technology,
Coimbatore. Tamil Nadu India,
srinishabalan@gmail.com

Abstract-**Recently, creating a realistic image using image processing has become an integral part of machine learning and computer vision. Realistic images in pixel size can be created using certain image processing operations. Learning to make colourful cartoon images from ordinary film or an image is not only an interesting research problem, but also has a potential application for digital entertainment. In this work, we will explore the problem of cartoon imagery from real life image using Laplacian filters and mini-batch k-means clustering operation. The new model not only has the ability to create cartoon images, but also allows users to specify the desired filter size. Experimental results show that the proposed work performs better than existing image-to-cartoon image methods.**

Keywords- Laplacian Filter, Mini Batch K-means Cluster, Sobel edge detection method, Cartoonify image.

I. INTRODUCTION

Digital image processing deals with processing an digital images through a digital computer with the help of algorithms. It is useful in developing a system that performs pre-processing of images. This system gets digital images as its input and pre-processes them using a highly efficient algorithm and finally produces an output as an image. Digital images are pre-processed to analyse and extract information. The extracted information is grouped as radiometric and geometric corrections. Radiometric corrections include corrections like unwanted sensors, sensor irregularities, conversion of data, and noise. Geometric corrections include geometric distortions conversion of data to coordinates on the earth's surface.

In digital image processing, images are represented in grayscale to extract features descriptors that simplify an algorithm and reduce computational conditions. The grayscale of an image represents the amount of light that carries intensity information. In digital image processing, threshold is a process that converts the color image into a black and white image or a grayscale image. Threshold is a type image segmentation which changes the pixel of image into easier format to analyse an image.

Shu, Y., et al., 2021 proposed a CVPR 2018 method to generate a multi style cartoon images based on the generative adversarial network (GAN) from the real world images. Li, X., et al., 2019 proposed a scGAN model which uses an attentive adversarial network to highlight the facial regions and disregards the low-level details. scGAN model improves the content in cartoon portraits to generate different cartoon styles. Wu, H., et al., 2021 designed a deep learning based method to generate the cartoon texture for regular and cartoonic images and also produced both the cartoon and regular texture for natural textures. Caleb, M.D.T., et al., 2021 proposed a CartoonGAN, GAN framework for cartoonization where the videos are sliced into images for training and to generate the high quality cartoon images from the real world images with clear edges and smooth shading. Enayat, S. et al., 2021 proposed a model which accepts both the image and the video as input and generates the cartoonified images and videos by extracting the edges, grey images and median blur with bilateral filter technique.

Li, S., et al., 2020 generated three different styles of cartoon images by using the factors like abstraction shading and crisp edges in terms of visual quality and quantity. Zhao, X., et al., 2021 developed LWAnimeGAN method which trains the photo2cartoon codes and it

converts image into line drawing, adds colors by using the color hints. This method is useful for cartoonists. Gao, C., et al., 2017 proposed a descriptor based model to generate cartoon artistic style in two ways of cartoonized images are cartoon portraits and cartoon sketch by using the guided filter. Pervaiz, I., 2013 proposed a cartoon system which works on the facial images to generate the cartoon images by using the unique features like lips, eyes, eyebrows, hair, ear, and some additional accessories and caricature is formed. Swaroop, K.S., et al., 2020 proposed a method which converts the real time images into cartoon style using the fully connected network. Dai, D, et al., 2014 introduced a dataset 21,302 textures and design a set of texture features, such as textureness, homogeneity, repetitiveness, and irregularity using ETH method and generated high quality textures by using Example-based texture synthesis (ETS) dataset.Wu, H., et al., 2020 introduced a deep learning method which containtwo modulescartoon textures and regularization textures .cartoon textures contain cartoonization methods to generate cartoonic and regular. Regularization module to generate a regular natural texture into a regular cartoon texture by using ImageNet Dataset. Karras, T.,et al., 2019 introduced an alternative generator architecture for generative adversarial networks, borrowing from style transfer literature a new architecture leads to an automatically learned, unsupervised separation of high-level attributes (e.g., pose and identity when trained on human faces) and stochastic variation in the generated images by using FFHQ Dataset

In the proposed work, the image-to-Cartoon Image synthesis problem by using Laplacian Filter and Mini Batch K-means Cluster is investigated. The new model is not only capable of generating cartoon images, but also allows users to indicate preferred filter levels. The proposed work contributes towards the improvement of the same by the following highlights.

- Improved frechet inception distance (FID) value.
- High information extraction can be made in the image to cartoon image generation
- Improved result image quality.

In this paper, the proposed method had been documented in upcoming section and it have been experimented with benchmark datasets to prove the mini k means cluster 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 Laplacian Filter and Mini Batch K-means Clustering operation. In this paper, the color clustering method is proposed in detail.

In this method, the color clustering model has been developed based on the mini batch k-means cluster. Input images are given into the model and the Laplacian filter blurs the input image by using the Convol 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. Mini Batch K-means Clustering groups the data points has high density in a region and adds up all of the individual kernels generates a probability surface depending on the kernel bandwidth parameter and perform masking. After masking, we obtain the cartoonified image. Figure 1 shows the system architecture of the proposed color unification system.

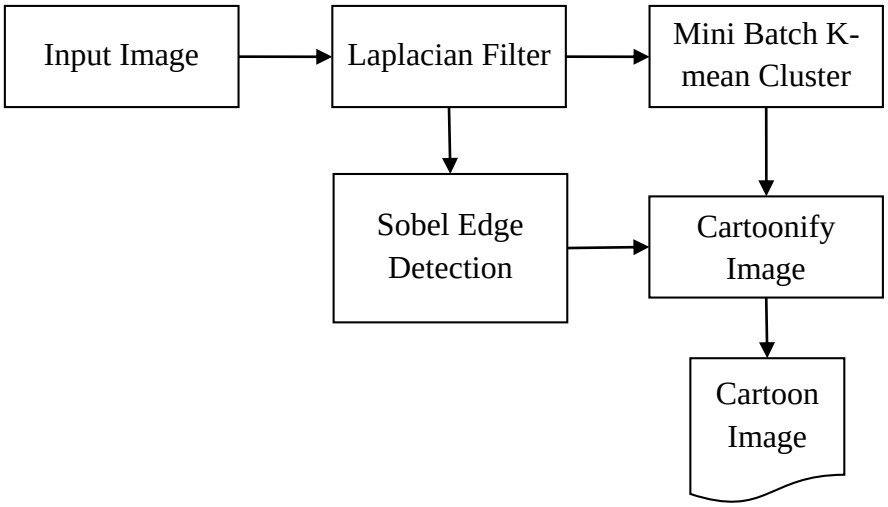

Figure 1: System Architecture of the proposed Mini Batch k-means Clustering method

A. Laplacian Filter

The second derivatives of a picture are computed using a Laplacian filter, which measures the rate at which the first derivatives change. This determines whether a change in neighbouring pixel values is caused by an edge or is part of a continuous progression. Negative values in a cross pattern, centred within the array, are common in Laplacian filter kernels. The corners have either a 0 or a 1 value. It's possible for the centre value to be negative or positive. The Laplacian L(x, y) of a picture with I(x, y) pixel intensity values is calculated as follows:

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

A convolution filter may be used to compute this.

B. Sobel Edge Detection

This work provides a technique for edge identification on pictures with White Gaussian noises that combines the Sobel edge detection operator with soft-threshold wavelet de-noising. Many edge detection approaches have been presented in recent years. The generally used approaches of combining mean de-noising with the Sobel operator, or median filtering and the Sobel operator, are ineffective in removing salt and pepper noise. In this study, we first reduce noise with a soft-threshold wavelet, then perform edge detection on the picture with the Sobel edge detection operator. This approach is most commonly used on photos with White Gaussian noises. When compared to typical edge detection approaches, the method suggested in this paper has a more visible influence on edge detection.

C. Mini Batch K-means Clustering

Clustering techniques have faster convergence times and are better suited to handling large amount of data. To cluster enormous datasets, the mini batch k-means clustering technique can be employed instead of the usual k-means approach. It outperforms the k-means clustering technique since it does not cycle over the full huge dataset. It creates random batches to store the data in memory. Then, throughout each cycle, these random batches of data are gathered which updates the clusters. This approach lowers the computational cost of cluster detection. The clustering and updating of data using Mini Batch K-means (Xiao, B., et al., 2018) is separated into the following steps: (1) divides huge datasets into several batches at random, treating each tiny batch as a separate cluster. (2) Determine the number of cluster clusters to be created at the start. (3) assign tiny data batches to the cluster centre closest to them. (4) Iteratively update the cluster centre until the cluster centre no longer changes. Each micro batch updates the clusters using a

convex mix of the prototype values and the data, with a decreasing learning rate as the number of iterations increases. The inverse of the number of data assigned to a cluster during the procedure is this learning rate. Because the influence of incoming data diminishes as the number of iterations grows, convergence can be observed when no changes in the clusters occur for multiple iterations in a row. The empirical results suggest that it can save a significant amount of time while sacrificing some cluster quality, but no comprehensive study of the algorithm has been conducted to determine how the characteristics of the datasets, such as the number of clusters or their size, affect partition quality.

D. Cartoonify Image

In this paper, we combined two special functionality like cartoonifying and masking. In catoonifying process, it converts the normal natural image texture into cartoonified image texture. In masking process, it combines two pictures using the bitwise operation with the pixels values of an image and adds the drawing style at the border of the image. Finally, our proposed model cartoonifies the image.

The Laplacian filter, sobel edge detection, and mini batch k-means clustering algorithms demonstrated the process of cartoonization of the input. The experimental analyzes were carried out using the procedure, and the comparison findings are provided in the following section.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. DATASET DESCRIPTIONS

We employed three datasets in this study: the ETH Synthesizability dataset, the Flickr-Faces-HQ (FFHQ) dataset, and the ImageNet dataset. To train and test our model, we used a total of 14,288,424 pictures. Whereas 21,302 texture synthesis samples come from the ETH Synthesizability dataset (Dai, D., et al., 2014), 14,197,122 image samples come from the ImageNet dataset (Wu, H., et al., 2020), and 70,000 high quality human face images with variations in the image background with 10241024 resolution come from the FFHQ dataset (Karras, T., et al., 2019), all of which can be cropped automatically with the help of dlib and aligned.

Table 1: Number of images presented in the datasets.

Dataset	No. of images
ETH Synthesizability dataset	21,302

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 used the Mini Batch K-means Clustering to implement our planned cartoonifing images. In this research, we used a Selective Laplacian Filter to pre-process the photos before applying the blur effect with the Convol function. Using the canny and approxcanny approaches, an adaptive threshold based edge discards those edges with weaker thresholds and chooses those with greater thresholds in the range of [0 1]. The mean shift clustering technique allocates data points to regions and groups the data points with the highest density. Finally, use masking to cartoonize the image. We used photos from all three datasets in this paper: the ETH Synthesizability dataset, the ImageNet dataset, and the Flickr-Faces-HQ (FFHQ) dataset. There are 21,302 texture samples in the ETH Synthesizability dataset. ImageNet has 14,197,122 pictures, according to the WordNet hierarchy. The FFHQ high quality collection comprises 70,000 PNG pictures with a size of 1024x1024 pixels and a variety of image backgrounds. The number of photographs in each dataset is listed in Table 1. Gray scale picture, threshold image, Gaussian blur, small batch k-mean cluster, cartoonify image, and their input image are shown in Figure 2.

ImageNet dataset	14,197,122
Flickr-Faces-HQ dataset	70,000

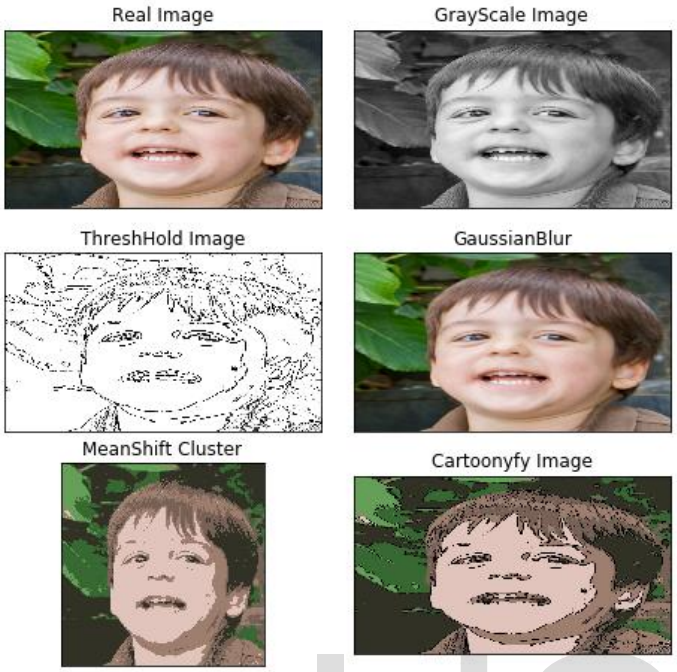

Figure 2: Results of gray scale image, threshold image, Gaussian blur, mini batch k means cluster, cartoonify image along with its input image.

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	192.32	167.09	163.22	131.52
FID to Nature texture	120.11	100.73	80.54	90.32

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 mini batch k mean based colour unification method in image cartoonization.

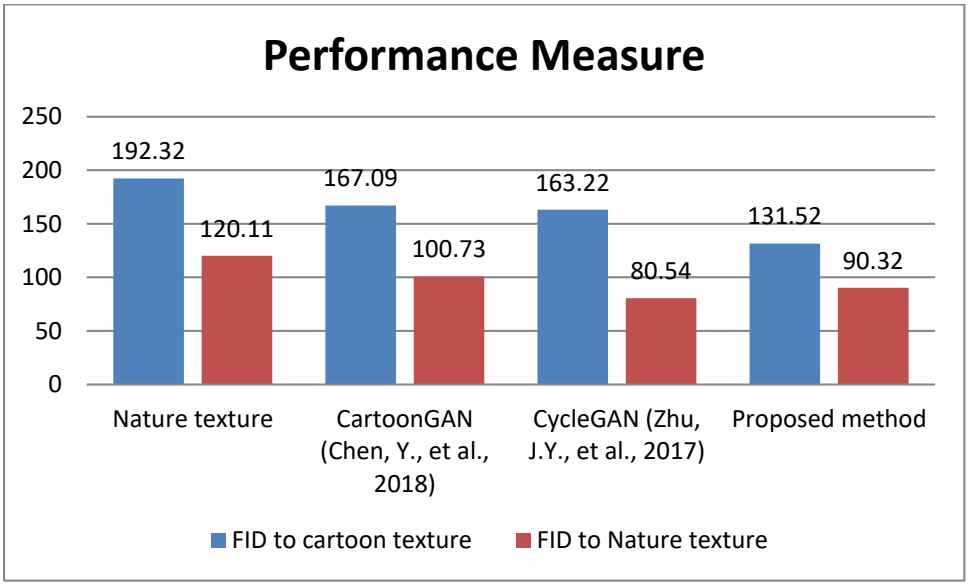

Figure 3: Graphical analysis of performance measure among all three methods.

IV. CONCLUSION:

In this paper, we use the Laplacian filter and Mini Batch k-means Clustering to examine the image-to-cartoon picture synthesis problem. Not only does the new model generate cartoon images, but it also allows users to specify their desired filter settings. The suggested technique outperforms existing image-to-image algorithms, according to experimental data. Images are merely numbers. So that's how we conceal our "BEAUTIFY" image's bordered image. Finally, our image is CARTOONIFYED with a high FID score.

REFERENCES

[1] Shu, Y., Yi, R. and Liu, Y.J., 2021, July. Cartoon Your Life. In *2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (pp. 1-2). IEEE.

[2] Li, X., Zhang, W., Shen, T. and Mei, T., 2019, July. Everyone is a cartoonist: Selfiecartoonization with attentive adversarial networks. In *2019 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 652-657). IEEE.

[3] Wu, H., Li, Y., Liu, X., Li, C. and Wu, W., 2021. Deep texture cartoonization via unsupervised appearance regularization. *Computers & Graphics*, *97*, pp.99-107.

[4] Caleb, M.D.T., Josphineleela, R. and Mary, S.S.C., 2021. REAL TIME VIDEO CARTOONIZATION USING GAN BASED ON MACHINE LEARNING.

[5] Enayat, S. and Khan, I.R., 2021, May. A Novel Method of Transforming Real Images & Videos into Cartoon Images & Videos. In *International Conference on Image Processing and Capsule Networks* (pp. 732-737). Springer, Cham.

[6] Li, S., Wen, Q., Zhao, S., Sun, Z. and He, S., 2020, October. Two-stage Photograph Cartoonization via Line Tracing. In *Computer Graphics Forum* (Vol. 39, No. 7, pp. 587-599).

[7] Zhao, X., Zhou, Y., Wu, J., Xu, Q. and Zhang, Y., 2021, January. Turn Real People into Anime Cartoonization. In *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (pp. 387-393). IEEE.

[8] Gao, C., Sheng, B. and Shen, R., 2017, November. Structure-Preserved Face Cartoonization. In *International Conference on Neural Information Processing* (pp. 622-631). Springer, Cham.

[9] Pervaiz, I., 2013. *Face Cartoonization and Caricature* (Doctoral dissertation, University of Management and Technology).

[10] Swaroop, K.S., Kumar, S. and Sowjanya, A., 2020, December. Real-Image Transformation into a Caricature

Image Using Neural Network. In *International Conference on Advanced Informatics for Computing Research* (pp. 670-679). Springer, Singapore.

[11] Xiao, B., Wang, Z., Liu, Q. and Liu, X., 2018. SMK-means: an improved mini batch k-means algorithm based on mapreduce with big data. *Comput. Mater. Continua*, *56*(3), pp.365-379.

12) Adler A, Yaniv I, Samra Z, Yacobovich J, Fisher S, Avrahami G, Levy I (2006) *Exserohilum*: an emerging human pathogen. Eur J ClinMicrobiol Infect Dis 25(4):247–25

13)S Su, G Wong, W Shi, *et al***Epidemiology, genetic recombination, and pathogenesis of coronaviruses**Trends Microbiol, 24 (2016), pp. 490-502

14)AmjadAlmahairi, SaiRajeswar, Alessandro Sordoni, Philip Bachman, and Aaron Courville. Augmented cyclegan: Learning many-to-many mappings from unpaired data. arXiv preprint arXiv:1802.10151, 2018.

15) 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).

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