New Style Image Generation using Neural Style Transfer

N K Bansode

Army Institute of Technology, Pune

---***---

Abstract - *Image style transfer using new techniques neural style transfer is method used to immigrate the style of one image and generate new image using contents image. The process of the neural style consists of texture extraction from style image and rendering contents image with the texture. The convolution neural network is a key technique in the process of neural style transfer. Neural style transfer finding many line realistic photo generation, style generation in the clothing, generating artificial art sculpture.*

***Key Words**: Neural Style Transfer, CNN, Content image, Style Image*

1.INTRODUCTION

The generation of the new image requires creative mind and it is work of special skill of the art producing with imagination or reproducing the image with new style or color. Humans have mastered the skill of an art generation of the visual experience through the composing of the complex structure of the content and style of image. Some artist are specialized in the production or generation of the new art[11]. Neural style transfer now days providing important results in the domain of the image art. The image processing techniques are developed over period of the couple decade for object detection and face recognition, the performance of the deep neural network near human performance.[13]. Neural style transfer is implemented by convolutional neural network by transferring style of one image in order to generate new image.

The neural style transfer involves finding out the matching part of the style image and content image in the convolution neural network layer. The gram matrix is generated to represent minimized a special maximized mean discrepancy. The traditional patch based methods, finding the patch of the image and getting analysis and synthesis of the features present in the path used for the matching of source image with the target image. Even Markov Random Field (MRF) used to analyze the feature present in the patch of the image. These non-parametric methods of the image texture analysis suffered from the fundamental problem of the local feature representation. The neural style transfer recently introduced for image style transfer. This neural style transfer based on the newly developed convolution neural networks. The artistic style of the image transfer is based on the gram matrix of the neural layer activation [12].

The generation of the new image performed by using iterative optimization of new image from white noise by matching new activation with gram matrix of the content image and style image. The global features are summarized and gram matrix produced with the specified layer of the convolution neural network. However, in this process a set of different global features are encoded into a same gram matrix due to this gram matrix representation become unsuitable [14]. The gram matrix generation only consider the global feature hence local feature are ignored. The neural style transfer methods are divided based on the use of techniques for style image generation. These methods are based on image optimization-based methods that optimize images online. The model optimization-based methods trains neural network for a specific style [15]. The neural style transfer methods are based on the parametric and non-parametric for image generation.

2. Related Work:

This section, presents the state of art of the neural style transfer work. Gatyes et.al.[1] studied the problem of the art generation using the artificial neural network. Yanghao Li, et al[2] demonstrated demystifying neural style transfer for generation of the new image using content and style image. This indicates that neural style transfer is intrinsically a process of distribution alignment of the neural activations between images. The kernel method based on the illumination analysis [17]. The convolution neural network used gram matrix of the neural activation layer to present the style of the image. The iteratively new image is generated from the initial white noise image by matching the neural activation with the content image and the gram matrix with the style image. Haochen Li [3], described a survey of work of neural style transfer on the video and images. The neural style transfer for the real time and super resolution images studied and implemented in the real time. They also described mixed style transfer on mixed style transfer via image fusion.

Len Du [4], presented new technique of the random weights in the layer l called as *ran*VGG and spatial smoothness is added. The removal of the content loss from the calculation of the total loss becomes easy to perform the work on the style loss. The randomized binary weights performing better as compare to the convolution neural network weights. Haozhi Huang et.al. [5] Explored the idea of the neural style transfer for the video using feedforward convolution neural network. They applied style transfer techniques to video using image frames. The Video generally produces the flicker due to the fame image. The hybrid loss used in the training stage, which combines losses in the spatial and temporal domain. The on fly optical computation is capable of performing real-time video styling. Zhiyuan Hu et.al [6] presented new layer for the image style transfer. The neural style transfer works on the content image and the style image. The content image is modified according to the colour and texture of the style image. They novel method of aesthetics aware model optimizations based style transfer. The multiple reference images are used to decide the colour and texture features. The two-path structure method is use for generation of the aesthetics style image.

Yijun L. et al [7], explored the idea of the universal style transfer via feature transform. The neural style transform is performed using content image and style image. In the research work they have implemented pair of features as whitening and colouring transforming that reflects the a direct matching of the feature covariance of content image to a given style image which is similar to optimization of the gram matrix based on the cost of neural style transfer. A feature transform of the image coupled with the pre-trained general encode-decode network and transforming processing is implemented by simple feed forward neural network. The results shows effectiveness of the feature transform method for the neural style transform and the use of these feature transform for natural textural analysis shown effective results. Gantugs Atarsaikhan et.al [8] explored neural style transfer for guided style transfer for shape stylization. They proposed method for neural style transfer based shape design method. The parametric and non-parametric both features are used for neural style transfer. They proposed method for generation of style without professional skill. The contents images are restricted to the clip art and binary silhouette image and style image can be any arbitrary image. Falong Shen et al.[9], presented novel method of the neural style transfer using meta networks. They designed new method to address the network generation task, which provides network for neural style transfer.

Fujun Luan et. al[10] introduced photo style transfer in transferring the reference image into the new image using style image. The matting laplacian method used for constrain transformations of the input and output images. The new feature such as semantic feature segmentation are used for neural style transfer. The resultant image is a photo realistic in a broad variety of scenarios.

3. Methodology

Style transfer is new technique of the image generation using combination of the content image and style image. The convolution neural network plays very important role in the image processing in the recent years and researches to think to solve image processing using this new technique [1]. Convolution neural network consist of the hierarchical layers of the computational units that processes image in the hierarchical and feed forward manner. Each layer in the convolution neural network extracts features of the image using different filter banks. The output of the each layer is called as feature map.

The style of the image is represented by using feature space, which captures the texture of the input image. The feature space is built on the top of the filters responses. The correlation between different filters responses and spatial extent of the feature map. The style feature produces texturiser version of the input image that captures appearances in terms of color and localised structures.The goal of the style using neural network transfer is to create new image x^* ,from content image x_c and new style for image x_s. The feature extracted from the both images is given by xc and xs in the layer l of the CNN is given by $F^l \ni R^{NlxMl}$, $P^l \ni R^{NlxMl}$, $S \ni R^{NlxMl}$ respectively. Where N^l the number of the feature maps in the layers l and M^l is the height times width of the feature maps.

Let $\bar{p}$ and $\bar{a}$ be original content image and generated in the layer l let P^l and F^l feature map generated by the convolution neural network. The squared loss between these layered images is given by equation (1)

$$L_{Content(\bar{p},\bar{x},l)} = \frac{1}{2} \sum_{i,j} (F^l_{i,j} - P^l_{i,j})^2 \quad (1)$$

The feature correlation given by Gram matrix equation(2)

$$G^l_{i,j} = \sum_k F^l_{ik} \, F^l_{jk} \quad (2)$$

The generation of the new image from the texture of the style image performed using gradient decent from the new white noise image by finding another images that represent style representation of the original image. The minimum distance between gram matrix of the original image and the gram matrix of the generated image. The total loss in the gram matrix is given by equation (3)

$$E_l = \frac{1}{4 N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)2 \quad (3)$$

Where A^l is a original image and G^l generated image in the layer *l*. The total loss is given by

$$L_{Style(\bar{a},\bar{x})} = \sum_{i,=0}^{L} w_l E_l \quad (4)$$

To generate new image which combines contents of the original image and textural of the style image, it is required to minimize square loss between white noise images from the content image. The total loss function due to content and style is given by equation (5)

$$L_{Total\,(\bar{p},\bar{a},\bar{x})} = \alpha L_{Contents}(\bar{p},\bar{x}) + \beta L_{Style}(\bar{a},\bar{x}) \quad (5)$$

4. Experimentations:

The style transfer of the one image to the content image is implemented using convolutional neural network. Various images used for style as well as content images. The result of the experiment implementation is shown in table 2. The model parameters for VGG convolution neural network in shown in table given below.

5. Results

The experiment of the neural style transfer performed on the content images using combination style image. The new generated images as neural style transfer. The results of the image style transfer is given table 2 below:

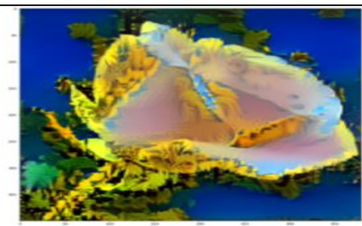
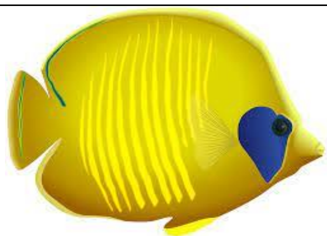
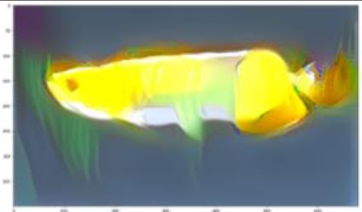
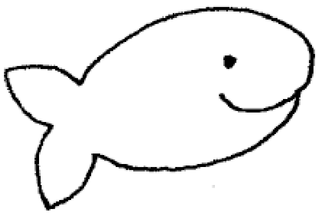
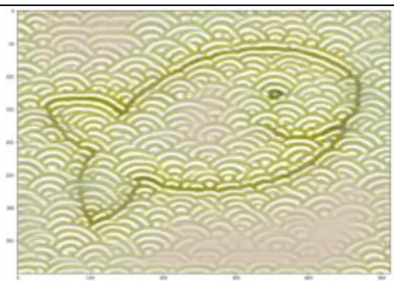
Table 1 : Model Parameters

Sr. No	Layer Type	Output Shape	Parameters
1	input_1 (Input Layer)	[(3, 400, 961, 3)]	
2	block1_conv1 (Conv2D)	[(3, 400, 961, 64)]	1792
3	block1_conv2 (Conv2D)	[(3, 400, 961, 64)]	36928
4	block1_pool (MaxPooling2D)	[(3, 200, 480, 64)]	0
5	block2_conv1 (Conv2D)	[(3, 200, 480, 128)]	73856
6	block2_conv2 (Conv2D)	[(3, 200, 480, 128)]	147584
7	block2_pool MaxPooling2D)	[(3, 100, 240, 128)]	0
8	block3_conv1 (Conv2D)	[(3, 100, 240, 256)]	295168
9	block3_conv2 (Conv2D)	[(3, 100, 240, 256)]	590080
10	block3_conv3 (Conv2D)	[(3, 100, 240, 256)]	590080
11	block3_pool(MaxPooling2D)	[(3, 50, 120, 256)]	0
12	block4_conv1 (Conv2D)	[(3, 50, 120, 512)]	1180160
13	block4_conv2 (Conv2D)	[(3, 50, 120, 512)]	2359808
14	block4_pool (MaxPooling2D)	[(3, 25, 60, 512)]	0
15	block5_conv1 (Conv2D)	[(3, 25, 60, 512)]	2359808
16	block5_conv2 (Conv2D)	[(3, 25, 60, 512)]	2359808
17	block5_conv3 (Conv2D)	[(3, 25, 60, 512)]	2359808

18	block5_conv3 (Conv2D)	[(3, 25, 60, 512)]	2359808
19	block5_pool	[(3, 12,	0

	(MaxPooling2D)	30, 512)]	

Table : Neural Style Transfer Result

Sr. No.	Content Image	Style Image	Neural Style Transfer
1			
2			
3			
4			

6. Conclusion

In this paper, neural style transfer method applied to generate new image from the contents of the original image by using texture of the style image. The input image of the size 224 x 224x3 . The convolutional neural network of VGG 16 used to extract the features of the contents image and style images at different layers. The gram matrix error between the contents image and style image is minimized to generate the new image with the given style image. The experimentation on the different contents images and style images is effective using neural style transfer. The neural style transfer considers only global feature maps of the both contents and style image. The use of both global and local feature maps may improve style transfer process into a new image.

7. REFERENCES

[1] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, " A Neural Algorithm of Artistic Style", arXiv:1508.06576v2 [cs.CV] 2 Sep 2015

[2] Yanghao Li, Naiyan Wang, Jiaying Liu, Xiaodi Hou, "Demystifying Neural Style Transfer", Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI- 2017)

[3] Haochen Li, A Literature Review of Neural Style Transfer, Princeton University Princeton University, Princeton NJ 08544

[4] Len Du , "How Much Deep Learning does Neural Style Transfer Really Need?" An Ablation Study, WACV, Computer Vision Foundation 2020.

[5] Haozhi Huang, Wenhan Luo, , Wenhao Jiang, Xiaolong Zhu, Zhifeng Li, Wei Liu, Hao , Lin Ma Wang, "Real-Time Neural Style Transfer for Videos", CVPR 2017, Computer Vision Foundation.

[6] Zhiyuan Hu, Jia jia, Bei Liu, Yaohua Bu , Jianlong Fu "Aesthetic-Aware Image Style Transfer", ACM , 2020

[7] Yijun L, Jimei Yang, Xin Lu, Ming-Hsuan Yang , Chen Fang, , Zhaowen Wang "Universal Style Transfer via Feature Transforms",

[8] Gantugs Atarsaikhan, Brian Kenji Iwana, Seiichi Uchida, "Guided neural style transfer for shape stylization", PLOS ONE, https://doi.org/10.1371/journal.pone.0233489 June 4, 2020

[9] Falong Shen, Shuicheng Yan, Gang Zeng, " Neural Style Transfer via Meta Networks", CVPR 2018, Computer Vision Foundation,

[10] Fujun Luan, Sylvain Paris, Eli Shechtman, Kavita Bala, "Deep Photo Style Transfer",

[11] J. R. Gardner, Y. Li, K. Q. Wein,berger, M. J. Kusner, K. Bala, and J. E. Hopcroft. P. Upchurch, , Deep manifold traversal: Changing labels with convolutional features. CoRR, abs/1511.06421, 2015

[12] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image analogies. In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pages 327–340. ACM, 2001

[13] C. Castillo, S. De, X. Han, B. Singh, A. K. Yadav, and T. Goldstein, "Son of zorn's lemma: Targeted style transfer using instance aware semantic segmentation," in IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2017, pp. 1348– 1352

[14] A. Gupta, J. Johnson, A. Alahi, and L. Fei-Fei, "Characterizing and improving stability in neural style transfer," IEEE International Conference on Computer Vision, 2017, pp. 4067–4076.

[15] K. Sunkavalli, W. Matusik, and H. Pfister M. K. Johnson. Multi-scale image harmonization. ACM Transactions on Graphics (TOG), 29(4):125, 2010.

[16] S. Azadi, M. Fisher, V. Kim, Z. Wang, E. Shechtman, and T. Darrell, "Multi-content gan for few-shot font style transfer," n Computer Vision and Pattern Recognition, 2018

[17] Adelson, J. R. , E. H. & Bergen Spatiotemporal energy models for the perception of motion. JOSA A 2, 284–299 (1985).

[18] Elgammal, A. & Lee, C.-S. Separating style and content on a nonlinear manifold. In Computer Vision and Pattern Recognition, 2004. CVPR 2004.

[19] Kummerer, Bethge, M. Deep Gaze I: Boosting Saliency Prediction ¨ with Feature Maps Trained on ImageNet.

[20] Yprianidis, J. E., Collomosse, J., Wang, T. & Isenberg, T. State of the Art: A Taxonomy of Artistic Stylization Techniques for Images and Video. Visualization and Computer (2013).