ARTISTIC NEURAL STYLE TRANSFER FOR IMAGES

Aravind Karrothu[1], Sanapathi Adithya[2], Veswanth Addagarla[3], Rudraraju Sahitya[4]

[1]Assistant Professor, [1234]Dept of CSE, [1234]GMR Institute of Technology, Rajam-India

ABSTRACT

Painting is one of the popular forms of art. To redraw a picture in a certain style in the past required a skilled artist and a lot of work. There are many techniques and studies researching how to turn pictures into beautiful art works. Among these studies, the convolutional neural network is one of the deep learning techniques used in creating artistic images by dividing up and recombining the content and style of the images. This image editing process is called Neural Style Transfer (NST). It's an optimization approach that combines two images known as a content image and a style reference image (say, an artwork from a painter) - so that the output image resembles the content image, but appears to have been painted in a style reference image way. For images classification, Visual Geometry Group (VGG) is most popular algorithm because of its pre-trained with several classes and huge amount of data. The style will be applied to the content image effectively using this model. To get the best art image by computing the high-level features in content image and low-level features along with gram matrices in style image. Optimization algorithm such as Limited-memory BFGS is used to optimize the resultant image. Ultimately, the Artistic neural style transfer gives the AI painted image.

Keywords: VGG, Neural Style Transfer, Transfer learning, Convolutional neural network, Image classification.

1 INTRODUCTION

Numerous studies and methods are being used to investigate how to create artificial artworks from pictures automatically. One such method is called "neural style transfer," which builds on the deep learning of a convolutional neural network (CNN) to produce amusing images by skillfully stylizing ordinary images with the specified visual art style. Digital images can be edited to adopt the appearance or visual style of another image using a set of software techniques called neural style transfer (NST). Deep neural networks are utilized for these image modifications. We need huge convolutional neural networks since we have to deal with pictures. NST can be used to create digital art from photos, such as by giving user-provided photographs the appearance of well-known paintings. Globally, this approach has been employed by designers and artists to create new artworks based on popular images. Leon A. Gatys initially discussed neural style transfer in his article in 2014 [16]. It requires two pictures, referred to as a content picture and a style reference picture and blend them together so the output image looks like the content image, but painted in the style of the resultant image.

A highly artistic use of neural networks is called neural style transfer. NST can be used to "paint" an image in the manner of a piece of art or painting given a content image. Neural Style

Submitted: 16/11/2022 **Accepted**: 05/12/2022

Transfer is significant in the field of computer vision as well. Deep neural networks that can correctly combine and derive feature representations of images. To create a combined image, the content representation and style representation are extracted and rebuilt. The total loss is determined by utilizing a linear combination of the content loss, the style loss, and both separately. Finally use the optimizers to reduce the time and optimize the image.

2 LITERATURE SURVEY

Ma, W., Chen, Z., & Ji, C [1] In this paper, the author proposed a new method that divides one memory-intensive task into several smaller, less memory-intensive tasks. This Style Transfer is a collection of NST methods that render input images using feedforward neural networks. Due to the high dimension of the output layer, these networks are computationally memory intensive. Most mainstream devices cannot directly stylize high-resolution photos due to memory limitations. Fast style transfer includes a pre-trained network for loss estimation and a feedforward neural network for image transformation. It overcomes image resolution limitations so multiple devices can quickly support high-resolution style transfer without the need to retrain models. Only the transmission network is updated. The weights of the pre-trained feature extractor remain constant throughout.

Shih, C. Y., Chen, Y. H., & Lee, [2] In this paper, a framework for creating stylized map images was proposed. Style can be successfully conveyed in photographs using contemporary approaches. This research presents a multi-step map art style transfer framework that can be used to improve results by reducing noise in map art images. The proposed framework creates a new piece of map art that uses reference style and content images. "Initial Phase" and "Refinement Phase" are two steps that combine the proposed system. The first purpose of this map art system is to remove the portrait from the reference map art and then use it for the style transfer stage. The second goal of this map art system is to apply the extracted portrait style to the input image. To achieve this, CNN-based neural style transfer is used as the basis for our style transfer technique, and the first result is adjusted to produce a new map drawing so that the final output map drawing is more like the reference map drawing.

Lin, C. C., Hsu, C. B., Lee, [3] In this paper, the author proposed a modified Universal Style Transfer method to create visual effects for stylized images, which increases the applicability of our technology for real-world image generation. NST is a method to cleverly apply a specified visual art style to everyday images to create interesting images. In this paper, a powerful post-processing framework has been developed using a modified UST approach, image fusion, and color enhancement techniques. The problems are with the color scheme, stroke severity, and image contrast settings. This UST was presented as a solution to these problems and for the ability to colorize the image. In this study, feature transformation operations combined with style transfer are used to recreate the image. The advantage of this approach is the ability to edit photos using any visual style without having to be trained in any predefined styles. Researchers improve the learning or network architecture of CNNs by using several different methods to create additional visual effects. Meanwhile, the results of these techniques do not seem to be satisfactory. Since there is no instrument to measure the effects of artistic style transfer, qualitative assessment must be highly subjective

Choi, H. C. [4] The primary objective of the paper is to conduct a thorough investigation in order to propose a new approach that is more successful in regulating the strength of the desired style. A pre-trained convolutional neural network that is used as a feature extractor for image content and style in image classification. One method is regression specification, which regulates the output style intensity of the stylized image according to the desired characteristic function between the output style strength and the style control parameter. In order to determine the appropriate relationship between style output force and style control parameter, the style transfer network uses intermediate style anchors.

Li, Z., Zhou, F., Yang [5] In this paper, single image super-resolution (SISR) is introduced to achieve smooth textures. CNN-based algorithms have shown better performance on SISR. Several lossy functions such as content loss, style loss, pre-pixel loss, and full variation control are used in this paper. This model was effectively trained as a two-phase training method. The work of the style transfer network is divided into three parts: feature engineering, style transfer, and image reconstruction. Relu3 3 and the VGG16 network were used to compute style and content information.

Dinesh Kumar, R., Golden Julie [6] This article is to generate an image whose properties are more like content images. The model in this article renders the input image to the extent of known works of art. This model mainly consists of the necessary layers that were mostly present in the VGG-16 model. The generated output will be the input for the next successive layers. With the help of these layers only can gather the features for both style and content representation. The result is a minimization of losses between the loaded style and the content representations.

Sheng, J., Song, C., Wang, J [7] In this paper, the author proposed a new Chinese painting style transfer (CPST) algorithm for transfer using the unique properties of ink and wash, which generates Chinese paintings using machine learning techniques. In Chinese painting, fine brush and hand brush are two different techniques. The Chinese Painting Style Transfer (CPST) algorithm claims to correctly transfer both painting techniques and ink tones to natural paintings. This Chinese Painting Style Transfer (CPST) algorithm aims to transfer the reference Chinese painting style to the input image without changing its shape. First, four main limitations are considered when comparing Chinese and Western paintings. The four main limitations are brush strokes, space reservation, diffused ink tone, and yellowing. A convolutional neural network (CNN) is used to incorporate these constraints into the transmission. The CNN is divided into different styles and content layers to faithfully preserve the style of the reference image.

Wang, X, [8] In this paper, a multi-style transfer approach is proposed that combines deep neural network (DNN)-based style transfer algorithms to achieve multiple style transfers in a single image. The author of this article designed a multi-style interactive art rendering using an image cropping tool useful for cropping content images into multiple parts. The user walks through the properties of the style and selects the style and algorithm from the manager. An interactive image cropping tool has been developed to crop content images into multiple parts. Each section has a style image and user selection algorithm. This framework includes six content images for any art style and ten typical art styles. It also allows you to add more art style to this frame.

Hong-an, [9] In this paper, the author proposed an improvised method to mine on image style transfer mapping relations by adding L1 loss, which minimizes the difference between the input image and the output image and improves the effect of image styling. Therefore, it is important to change the image style in order to use the information effectively. The main contribution of this paper is the use of the VGG (Virtual Geometry Group) network model to find information about the content and style characteristics of an image. L1 loss and increased perceptual loss demonstrated improvement in the structural similarity between resulting and artistic images.

Zhao, H. H., Rosin [10] In this paper, the author proposed a new image synthesis method for transferring image styles. Neural image style transfer has two types of image feature representation methods used in style transfer based on deep learning and local approaches. Global style information helps eliminate patch transfer errors, such as transferring the mustache, mouth, and eyes to the wrong places. Local style loss helps preserve detailed styles better. By combining this method, patterns are transferred and artifacts are reduced. In order to incorporate these constraints into the transfer, a pre-trained VGG (virtual geometry group) mesh is useful for generating feature maps.

Wang, Z., Zhao [11] In this paper, the author proposed an improved method for evaluating and improvising the quality of NST stylization using various deep learning techniques. In this work, its clearly observed a few challenges, i.e., firstly, decompose design transfer quality into different quantifiable factors. Furthermore, two new methods of using factors to increase the quality of stylization are presented. The first is to use requirements to combine existing methods to improve strengths. The second is to optimize the factors to obtain better requirements.

Wang, Z., Yang [12] In this article, the author examines the CNN-based transfer of artistic styles and explores the key reasons for this transfer. This method is intended to detect that a lost photo of a stylized result is due to distortion occurring in both the content preservation phase and the style transformation phase; therefore, here he proposed a photographic style transfer method capable of improving photographic stylized results. A total loss function is used to reconstruct content details and avoid the geometric mismatch problem, and a fused model with an edge-retaining filter is used to minimize artifacts. Qualitative evaluations show that this approach successfully suppresses bias while obtaining stylized results.

Tuyen, N. Q., Nguyen, [13] In this paper, the author proposed a complex multimodal style transmission network, called deep correlation multimodal style transmission. Image representation plays a key role in stylization, and by creating alternative image representation strategies, different stylization results can be generated. Compared to other correlations, the images generated by the Gram matrix are more effective in balancing performance, content preservation, and style customization. Since there is no instrument to measure the effects of artistic style transfer, qualitative assessment must be highly subjective.

Cheng, M [14] The author of this paper suggested a technique for image stylization that preserves and emphasizes structures. They successfully maintain layout structures while including artistic elements under the guidance of a global texture extraction network and a local texture refining network. Iterative optimization methods or modern neural-style transfer approaches that train feed-forward neural networks have shown impressive results. It considers two variables, the first of which is the global structure that is depicted.

Virtusio [15] In this research, the author developed a single-style input style transfer method with a hybrid human-artificial intelligence (H-AI)-inspired design that takes into account human control over the stylization process and allows for a variety of outputs. The various perceptual characteristics present in a single style image serve as the inspiration for this strategy.

3 METHODOLOGY

3.1 Proposed approach

In general, neural style transfer (NST) combines the appearance of one picture with the information contained in another. Before the creation of CNNs, the issue was tough to solve since it was challenging to extract texture information using hand-crafted features. The capacity to distinguish between an image's style and content using a convolutional neural network is known as neural style transfer (CNN). You have countless options when using neural style transfer to add a whole distinct style to an existing image. Any style is acceptable, including cubism, surrealism, impressionism, and original works**.** The goal of this neural style transfer is to produce a picture that combines the "contents" of one image with the "style" of another image. In order to compare how closely the created image matches the desired creative output, we must capture both the style and the content. A very artistic use of neural networks is called neural style transfer. NST may be used to paint a picture in the manner of a piece of art or painting. The content picture would be stylized in a number of ways, such as color scheme, image patterning, or brush strokes. Basically, NST turns any regular image into a piece of art. Convolutional Neural Networks (CNN) are a subclass of Deep Neural Networks that perform in image processing applications. To create a merged image, the content representation and style representation are extracted and rebuilt. The total loss is determined by utilizing a linear combination of the content loss, the style loss, and both separately. It employs two small-number parameters, alpha and beta, which are optimized to produce outcomes.

This smoothing loss is a noise reduction approach that, if it is a near match to the original signal, minimizes the overall variation of the signal while keeping significant features like edges. Because the Smoothing loss effectively counts the amplitude of all the fluctuations occurring in the image, we utilize this loss as our final loss. Before putting the content and style images into the pretrained network, we must use normalization to automatically rescale and remove the mean of the photos. In order to compare these content and style images to the produced image, we send them to different layers of the pretrained network. We compute the total loss, which is the weighted sum of the content loss and the style loss as compared to the produced image. Until convergence is attained, the optimizers are utilized to minimize this overall loss and update the image.

3.2 Feature Losses

The characteristics that a neural network discovers inside a picture are dependent on loss. We must define a loss function that will optimise the losses toward the intended result in order to obtain the desired image. Here, we'll apply the idea of per-pixel losses.

1) Content Loss:

The difference between the content features of a base picture and the content features of an image that was created with a new style is known as the content loss in neural style transfer.

2) Style Loss:

The difference between the produced picture's lower-level characteristics and the base image is measured as "style loss." elements like colour and texture, for instance.

While content loss comes from upper levels, style loss comes from all layers. To ensure that there is a clear distinction between the style picture and the created image, it penetrates the deepest levels possible. We don't want the original artwork to become worthless or lose its true significance, after all.

3.3 Total Loss:

You must determine the overall loss after accounting for both style loss and content loss.

The sum of the content loss and the style loss is known as the total loss. You may comprehend how the optimizer finds a picture that has the look of one image and the substance of another by determining the total loss.

The loss function we minimize is,

$$L_{total}(p,a,x)= \alpha*(L_{content}(p,x))+ \beta*(L_{style}(a,x))$$

These weights, which can be adjusted for different outcomes, are for the content reconstruction and the style reconstruction, respectively. This is the “Combined Loss”.

3.4 Normalization:

In Style Transfer and some generative networks, normalization are used to control the style of the output image. We add one more block to automatically rescale and remove the mean of the images going through VGG19. While training neural network on image dataset, all images are their mean removed, and values are rescaled according to the standard deviation. This is a common processing which makes the neural network convergence easier. The only implication is that once we want to use this network, we have to normalize the input image according to the coefficients. Normalize the activation outputs by $(x-\mu)/\sigma$ and Scale it by multiplying γ and then shift it by adding β.

$$N(x)=\gamma * \left(\frac{x-\mu(x)}{\sigma(x)}\right) + \beta$$

x : activation output, $\mu(x)$,$\sigma(x)$ are the representation of styles, γ is the standard deviation of target style representation, β is the mean of target style representation. Therefore, by controlling γ and β in the normalization, we can control the style of the output.

3.5 Denoising Techniques:

Removing various types of noise from an image is a crucial step in image processing. During picture storage, transmission, and capture are the main times that noise in an image is introduced. Determining the original image content is essential for effective performance in a variety of applications, including image restoration, visual tracking, image registration, image segmentation, and image classification. As a result, image denoising is key in these applications. Image denoising has actually been studied for a very long time and is a classic problem. Denoising techniques are used to reduce the noise of the image in order to increase the quality of generated output image. Some of the techniques are linear smoothing, median filtering, Smoothing loss or Total variation regularization.

3.6 VGG 19 Model

VGG 19 Model - It is referred to as VGG and is a common deep convolutional neural network (CNN). The innovative object recognition model's foundation is the VGG architecture. VGG 19 model has a total of 19 layers, of which 16 are convolutional, 3 are fully connected, 5 are max poll, and 1 is soft max. The foundational layer of a CNN is a convolutional layer. It includes a number of filters, the parameters of which must be learned during model training. Feed-forward neural network is the only type of fully connected layer. It creates the network's last connected layer. The maximum elements from the region of the feature map covered by the filter are selected by the pooling operation that uses the max pool layer. Through the mathematical procedure known as SoftMax, a vector of numbers is converted, and each value's probability is inversely proportional to its position inside the vector. In this the present model of Neural Style Transfer can be achieved by using the pre-trained VGG19 model. Neural style transfer is predicated on the hypothesis that a CNN, trained for a computer vision task, can learn to separate its style representation from its content representation (e.g., image recognition task).

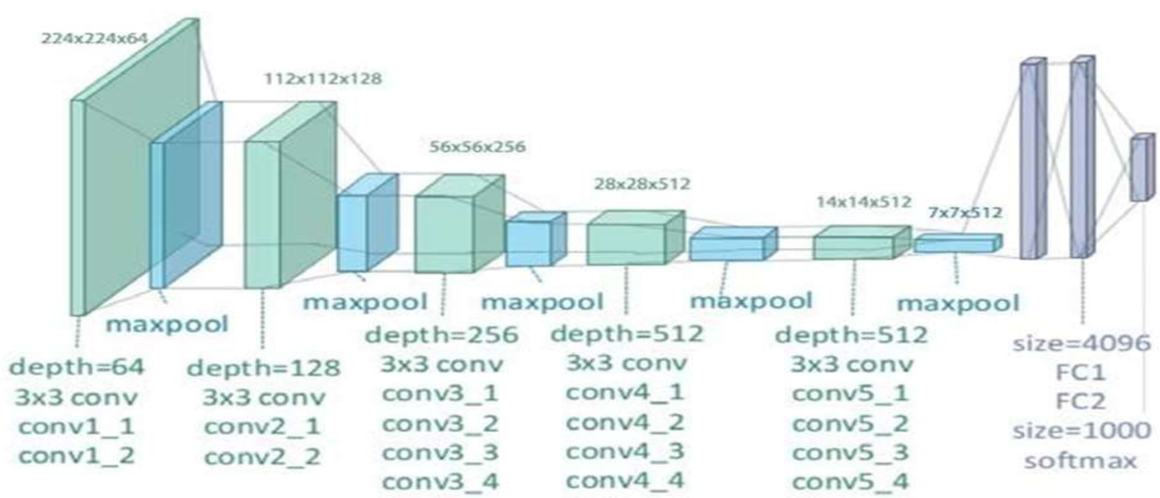

Figure 1: The Architecture of VGG 19

3.7 Optimization in neural network

We have the idea of loss in deep learning, which informs us of how poorly the model is performing right now. We must now teach our network to perform better using this loss. We basically need to take the loss and strive to reduce it because a lower loss indicates that our model will perform better. Optimization is the process of minimising (or maximising) any mathematical statement.

Here we take two different types of optimizers and we will see how they exactly work to minimize the loss function:

1. Adam
2. LBFGS (Limited Memory BFGS).

ADAM (Adaptive Moment Estimation): Adam can be viewed as a fusion of momentum-driven stochastic gradient descent and RMSprop. Each parameter's adaptive learning rates are calculated by Adam.

LBFGS: This algorithm is a member of the Quasi-Newton family of techniques. These procedures, which are based on Newton's technique for locating stationary points of functions, are used to locate local extrema of functions.

4 RESULTS

Optimizers	Style Loss	Content Loss
LBFGS	0.924	4.738
Adam	18.964	5.555
Lbfgs with denoising	0.628	3.977

SSIM	MSE
71.0144	1208.023

Content image image	Style image	Resultant

5 CONCLUSION:

This article provides an overview of the different methods and a detailed comparison table for style transfer to identify the advantages of their methodologies, limitations of their work, and future work. Most evaluations of neural style transfer methods are qualitative, as judging image quality is primarily subjective. The most common approach is side-by-side image comparison and another common method is user study (harmonization), i.e., only the image quality will be judged based on user feedback.

6 FUTURE SCOPE:

After analyzing all reference articles, some gaps are identified in the main topic, i.e., neural style transfer. It is obvious that the main thing of this research is to improve image quality. To improve image quality, you modified the super-resolution style transfer network with both content information and style information to perform a style transformation. Li et al, 2020 noted that the improvement of quality factors, these quality factors can directly serve as a general measure of quality for the transfer of style and impressive visual effects must be achieved by applying some "loss adjustment" method for various fields such as computer vision and deep learning. Using various noise reduction techniques helps to minimize the noise in the resultant image as well as optimizers to minimize total loss and to increase the image quality by gathering more content information from the image.

6 REFERENCES

[1] Ma, W., Chen, Z., & Ji, C. (2020). Block shuffle: a method for high-resolution fast style transfer with limited memory. IEEE Access, 8, 158056-158066.
[2] Shih, C. Y., Chen, Y. H., & Lee, T. Y. (2021). Map art style transfer with multi-stage framework. Multimedia Tools and Applications, 80(3), 4279-4293.
[3] Lin, C. C., Hsu, C. B., Lee, J. C., Chen, C. H., Tu, T. M., & Huang, H. C. (2022). A Variety of Choice Methods for Image-Based Artistic Rendering. Applied Sciences, 12(13), 6710.

[4] Choi, H. C. (2020). Unbiased Image Style Transfer. IEEE Access, 8, 196600-196608.
[5] Li, Z., Zhou, F., Yang, L., Li, X., & Li, J. (2020). Accelerate neural style transfer with super-resolution. Multimedia Tools and Applications, 79(7), 4347-4364.
[6] Dinesh Kumar, R., Golden Julie, E., Harold Robinson, Y., Vimal, S., Dhiman, G., & Veerasamy, M. (2022). Deep convolutional nets learning classification for artistic style transfer. Scientific Programming, 2022.
[7] Sheng, J., Song, C., Wang, J., & Han, Y. (2019). Convolutional neural network style transfer towards Chinese paintings. IEEE Access, 7, 163719-163728.
[8] Wang, X., Lyu, Y., Huang, J., Wang, Z., & Qin, J. (2021). Interactive Artistic Multi-style Transfer. International Journal of Computational Intelligence Systems, 14(1), 1-13.
[9] Hong-an, L., Zheng, Q., Qi, X., Yan, W., Zheng, W., Li, N., & Tang, C. (2021). Neural Network-Based Mapping Mining of Image Style Transfer in Big Data Systems. Computational Intelligence and Neuroscience: CIN, 2021.
[10] Zhao, H. H., Rosin, P. L., Lai, Y. K., Lin, M. G., & Liu, Q. Y. (2019). Image neural style transfer with global and local optimization fusion. IEEE Access, 7, 85573-85580.
[11] Wang, Z., Zhao, L., Chen, H., Zuo, Z., Li, A., Xing, W., & Lu, D. (2021). Evaluate and improve the quality of neural style transfer. Computer Vision and Image Understanding, 207, 103203.
[12] Wang, L., Wang, Z., Yang, X., Hu, S. M., & Zhang, J. (2020). Photographic style transfer. The Visual Computer, 36(2), 317-331.
[13] Tuyen, N. Q., Nguyen, S. T., Choi, T. J., & Dinh, V. Q. (2021). Deep correlation multimodal neural style transfer. IEEE Access, 9, 141329-141338.
[14] Cheng, M. M., Liu, X. C., Wang, J., Lu, S. P., Lai, Y. K., & Rosin, P. L. (2019). Structure-preserving neural style transfer. IEEE Transactions on Image Processing, 29, 909-920.
[15] Virtusio, J. J., Ople, J. J. M., Tan, D. S., Tanveer, M., Kumar, N., & Hua, K. L. (2021). Neural style palette: A multimodal and interactive style transfer from a single style image. IEEE Transactions on Multimedia, 23, 2245-2258.
[16] Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.