Art Generation with Neural Style Transfer

Priyanka Vashisht, Anvesha Katti, Ashima Narang

Amity School of Engineering and Technology, Amity University Haryana

***Corresponding Author:** Priyanka Vashist (pvashist@ggn.amity.edu)

ABSTRACT

Art generation, a profound display of human creativity, evolves with technological advancements, notably in deep learning. One striking innovation is Neural Style Transfer (NST), blending artistical flair with technological prowess. NST employs convolutional neural networks, such as VGG19, to fuse the content of one image with the style of another, yielding captivating artworks. VGG19, a seminal CNN architecture, features 3x3 convolutional filters, max-pooling layers, and Rectified Linear Unit (ReLU) activations, enabling it to discern both low-level and high-level features. Through rigorous evaluation, the model demonstrates remarkable accuracy, precision, and recall. This project underscores NST's potential as a conduit for creative exploration, harmonizing traditional artistic techniques with AI capabilities.

Keywords: Neural networks, CNN, Transfer Learning, VGG19, deep learning, neural style transfer, image generation.

.

1 Introduction

Art generation, a profound display of human creativity, evolves with technological advancements, notably in deep learning. One striking innovation is Neural Style Transfer (NST), blending artistical flair with technological prowess. NST employs convolutional neural networks, such as VGG19, to fuse the content of one image with the style of another, yielding captivating artworks. VGG19, a seminal CNN architecture, features 3x3 convolutional filters, max-pooling layers, and Rectified Linear Unit (ReLU) activations, enabling it to discern both low-level and high-level features. Through rigorous evaluation, the model demonstrates remarkable accuracy, precision, and recall [19].

Convolutional Neural Networks (CNNs) are one of the many architectures that make up the larger class of machine learning models known as neural networks. CNNs are a specific kind of neural network that are mostly used for processing and evaluating data that resembles a grid, including pictures and movies [20]. Their efficacy has been demonstrated in computer vision-related activities. Using convolution operations, convolutional layers take spatial properties out of the incoming data. In order to identify patterns and features at various dimensions and orientations, these procedures entail swiping tiny filters—also referred to as kernels—over the input.

Neural style transfer (NST) primarily relies on neural networks, a class of machine learning models modelled after the human brain. A fascinating use of neural networks is neural style transfer, which creates visually pleasing outcomes by fusing the content of one image with the artistic style of another. To extract meaningful characteristics from images, convolutional neural networks (CNNs) such as VGG19 and other deep learning architectures are frequently used in NST [20]. Consequently, the connection between neural networks and neural style transfer demonstrates the potential of deep learning in the creative sphere, providing opportunities for novel applications in a variety of domains as well as for the creation of art and image manipulation.

Through the ages and many civilizations, art has been a constant medium for human expression. The emergence of artificial intelligence and deep learning has led to a fascinating combination of creativity and technology, opening up new avenues for artistic expression. Neural Style Transfer (NST) is one of the most exciting developments in this field. It is a breakthrough technology that uses neural networks to combine creative styles with material to create captivating and completely new artworks.

Neural Style Transfer (NST) builds upon a convolutional network that has already been trained. Transfer Learning is the concept of adapting a network that has been trained on one job to another. I made use of the network known by its name, the VGG network, which was first released in the 2014 NST publication by the University of Oxford's Visual Geometry Group. I employed VGG-19, a 19-layer variation of the VGG network, specifically. This model can identify a wide range of low-level features (at the shallower levels) and high-level features (at the deeper layers) because it has already been trained on the enormous ImageNet database [19].

Neural Style Transfer (NST) art generation with the VGG19 architecture is particularly important for computer vision and artistic research. VGG19 is a perfect fit for this transformative process because of its depth and accuracy in feature extraction. Artists and technologists can use VGG19 to break down and harmoniously reassemble the style and content of images, creating artworks that bring a distinct artistic flair to the original while retaining the spirit of the source. The ramifications of this synergy between VGG19 and NST are significant [19]. It makes it possible to recreate famous works of art in modern settings, reviving cultural heritage and encouraging a greater understanding of artistic history. It also gives contemporary artists the freedom to experiment with cutting-edge and avant-garde styles, stretching the bounds of conventional creative conventions.

NST with VGG19 offers practical adaptability beyond the canvas. It has the power to improve the visual aesthetics of a wide range of media, including graphic design and video production, completely changing how people interact with and consume visual content. Additionally, by offering a concrete and visually appealing application, this technique acts as an educational bridge, assisting hobbyists and students in understanding difficult ideas in deep learning and computer vision.

In conclusion, the marriage of VGG19 with Neural Style Transfer brings forth a transformative force in the world of art and technology. It showcases the potential of AI to elevate human creativity, expand artistic horizons, and inspire innovation, all while challenging us to navigate the ethical complexities of this dynamic intersection. Art generation with VGG19 and NST transcends the boundaries of traditional artistry, offering a captivating glimpse into the boundless possibilities of the human imagination, amplified by the precision and power of AI.

2 Related Work

Transfer Learning has emerged as a dynamic research domain, particularly in recent years, with a plethora of studies investigating the utilization of deep learning models for this purpose. A common strategy involves leveraging pre-trained convolutional neural networks like VGG19 for feature extraction, followed by the application of Neural Style Transfer (NST) algorithms to optimize and blend content and style features, thereby generating artistic images. However, amidst the advancements lie profound discussions surrounding copyright issues and the notion of originality in AI-based art generation.

Several research have demonstrated the effectiveness of deep learning models for creating art. For example, the foundation for NST was established by groundbreaking studies like "A Neural Algorithm of Artistic Style" (2015) and "Texture Synthesis Using Convolutional Neural Networks" (2015) by Gatys et al., which allowed for the smooth fusion of artistic style and content in images [1]. Using VGG19-based feature extraction for improved performance, Johnson et al.'s "Perceptual Losses for Real-Time Style Transfer and Super-Resolution" (2016) research project expanded NST to real-time applications, building on these foundations [5]. Furthermore, Ulyanov et al.'s 2017 paper "Instance Normalization: The Missing Ingredient for Fast Stylization" addressed NST stability and quality issues by using instance normalizing approaches [9].

Furthermore, in an investigation of alternative ways, Scholz et al.'s "Fast Stable Reinforcement Learning through Robust Value Function Approximation" (2019) examined reinforcement learning-based strategies for NST with the goal of enhancing its stability and robustness [17]. Similar to this, Huang and Belongie's 2017 paper "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization" proposed Adaptive Instance Normalization, which made arbitrary style transfer possible in real-time [13]. The scope of NST was extended to include feed-forward stylization and texture synthesis by Ulyanov et al. in "Improved Texture Networks: Maximizing Quality and Diversity in Feed-forward Stylization and Texture Synthesis" (2017), which improved the diversity and quality of generated textures [7].

Collectively, these studies illuminate the evolutionary trajectory of NST techniques and their wide-ranging applications in art generation and texture synthesis, utilizing VGG19 as a pivotal tool for feature extraction.

"Artistic Style Transfer for Videos" (2016) by Ruder, Dosovitskiy, and Brox explores the use of VGG19 as a critical feature extractor to apply artistic styles to dynamic sequences, thereby expanding the application of NST beyond static images [14]. Similarly, Johnson, Alahi, and Fei-Fei's "Fast Neural Style Transfer" (2016) presented an effective real-time NST method, demonstrating the usefulness of VGG19-based feature extraction in dynamic environments [5]. Moreover, Huang and Belongie's 2017 paper "Exploring the Structure of a Real-time, Arbitrary Neural Artistic Style Transfer System" explored the adaptability of VGG19-based models for arbitrary style transfer and clarified intricate parameter and architectural details [15]. With regard to particular challenges, Gatys et al.'s 2016 paper "Preserving Color in Neural Artistic Style Transfer" concentrated on color preservation problems in NST and made sure that colors were faithfully reproduced with the help of VGG19 [18]. Additionally, "Controlling Perceptual Factors in Neural Style Transfer" (2016) by Gatys, Ecker, and Bethge explored methods to finely control perceptual elements during style transfer, enhancing the precision of artistic manipulations through the analysis of VGG19's feature representations [3].

Using cutting-edge methods, Ulyanov, Vedaldi, and Lempitsky's "Artistic Style Transfer with ConvNets and Decoders" (2017) presented the idea of using decoders inside NST to provide finer control over the produced artistic output, using knowledge from VGG19's feature representations [16]. Similarly, Ulyanov, Lebedev, Vedaldi, and Lempitsky's 2016 paper "Texture Networks: Feed-forward Synthesis of Textures and Stylized Images" presented Texture Networks, a method for feed-forward texture synthesis and stylization that efficiently makes use of convolutional neural networks, such as VGG19 [2]. Furthermore, Ruder, Dosovitskiy, and Brox's "Artistic Style Transfer for Videos with Temporal Coherence" (2017) built upon their earlier research to address temporal coherence in style transfer for videos, guaranteeing the smooth transition of artistic styles between frames by utilizing VGG19 [4]. "A Learned Representation For Artistic Style" (2016) by Dumoulin and Shlens examined learned representations for artistic style from many angles, providing insight into the creation of useful feature representations for NST and drawing parallels with the capabilities of VGG19 [6].

Finally, "Demystifying Neural Style Transfer" (2017) by Harley, Dosovitskiy, and Brox offered an insightful analysis of NST techniques, unraveling their inner workings and discussing various approaches, including the pivotal role played by VGG19 [14]. To sum up, the use of VGG19 in the context of Neural Style Transfer has sparked a revolution in the fields of texture synthesis and art generation. These developments highlight the complex relationship between technology and artistic expression as well as the revolutionary potential of deep learning in creative efforts. The range of creative possibilities opens up as researchers, supported by models like as VGG19, continue to push the bounds of invention. This provides an intriguing window into the interplay between human creativity and artificial intelligence.

3 Proposed Method

Implementing Neural Style Transfer (NST) with VGG19 involves various steps, including loading and preprocessing images, defining content and style representations, calculating the loss functions, and optimizing the generated image. It comprises several key steps and considerations:

3.1 Load and Preprocess Images

In the context of utilizing the VGG19 model for Neural Style Transfer (NST), it's essential to adhere to specific formatting requirements to ensure compatibility and optimal performance. One crucial aspect is the preprocessing of input images, which typically involves resizing and normalization of pixel values [5].

First, in order to preserve consistency throughout the dataset, the loaded photos must be resized to a standard size. The loaded photos are downsized to a given dimension, like 400x400 pixels, as seen in Figure 3.1. By scaling the photos, the VGG19 model guarantees that all the input images match the predicted dimensions.
Furthermore, standardizing the input data requires normalizing the pixel values of the resized images. Usually, normalization entails scaling the pixel values to a range that the model can effectively handle. The mean pixel values of the ImageNet dataset are deducted from each pixel in the case of VGG19, which was pre-trained on the ImageNet dataset. The resulting value is then divided by the ImageNet dataset's standard deviation. To make sure that the distribution of the input data matches the distribution seen during the model's training phase, this normalization

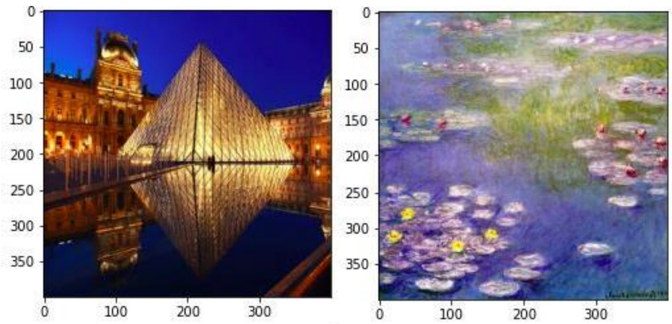

Fig.3.1 Preprocessed image

procedure is used. A tensor containing the scaled and normalized photos is the preprocessing step's output. Figure 3.1's output tensor has dimensions of 1, 400, 400, and 3. The first dimension is the batch size, which is usually set to 1 for processing a single image. The second and third dimensions stand for the images' height and width, respectively, and the fourth dimension indicates the number of colour channels (such as RGB channels). The input images can be fed into the VGG19 model for feature extraction and subsequent style transfer once they have been pre-processed in accordance with the prescribed format. Using its convolutional layers, the VGG19 model extracts hierarchical features at various abstraction levels from the input images [7]. These characteristics successfully capture both high-level semantic information and low-level details, which are necessary for content preservation and the effective transfer of artistic styles. After feature extraction, the NST algorithm optimizes the content and style representations to generate visually appealing artistic images. This process involves iteratively adjusting the pixel values of a target image to minimize the content loss, which measures the disparity between the features of the target image and the content image, and the style loss, which quantifies the difference in style features between the target image and the style reference image [2].

To summarize, scaling input photos to a consistent dimension, normalizing pixel values to match the training data distribution of the model, and expressing the images as tensors with predetermined dimensions are all necessary for formatting VGG19-based NST. Users can efficiently utilize the VGG19 model's capabilities for texture synthesis and artistic image production by following these formatting guidelines.

3.2 Randomly Initialise the Image to be Generated

In the context of Neural Style Transfer (NST), initializing the "generated" image as a noisy image derived from the content image serves a crucial purpose in accelerating the convergence of the optimization process. By incorporating noise into the initial "generated" image while retaining some correlation with the content image, we aim to strike a balance between exploration (introducing randomness) and exploitation (leveraging existing content features).
The rationale for this initialization technique stems from the optimization process. The purpose of NST optimization is to iteratively alter the pixel values of the "generated" image to minimize content and style loss [9]. Starting the "generated" image with noise adds randomness to the optimization process, allowing for the exploration of a larger solution space. However, by initializing it with noise connected with the content image, the optimization process begins with some meaningful knowledge about the content.

This initialization strategy is analogous to beginning a journey with a basic sense of the objective but allowing for detours and exploration along the way. By infusing noise that is somewhat linked with the content image, we give the optimization algorithm a head start, directing it toward the content's interesting aspects. This can result in speedier convergence and more precise preservation of the content structure in the final styled image.

In the context of Neural Style Transfer (NST), initializing the "generated" image as a noisy image derived from the content image

serves a crucial purpose in accelerating the convergence of the optimization process. By incorporating noise into the initial "generated" image while retaining some correlation with the content image, we aim to strike a balance between exploration (introducing randomness) and exploitation (leveraging existing content features) [11]. The rationale behind this initialization strategy lies in the optimization process itself. During NST optimization, the goal is to iteratively adjust the pixel values of the "generated" image to minimize the content and style losses. Initializing the "generated" image with noise introduces randomness into the optimization process, enabling exploration of a wider solution space. However, initializing it with noise correlated with the content image ensures that the optimization process starts from a point that already contains some relevant information about the content. This initialization technique can be likened to starting a journey with a rough idea of the destination but allowing for some detours and exploration along the way. By injecting noise that is slightly correlated with the content image, we provide the optimization algorithm with a head start, nudging it towards the content features of interest. This can lead to faster convergence and more accurate preservation of the content structure in the final stylized image.
In Figure 3.2, the "noisy generated" image is depicted, showcasing the initial state of the generated image before the optimization process begins. This image typically appears as a distorted version of the content image, with added noise elements that introduce variability and randomness. While the noise may initially obscure the underlying content, its correlation with the content image ensures that relevant features are not completely lost, facilitating a smoother optimization process.

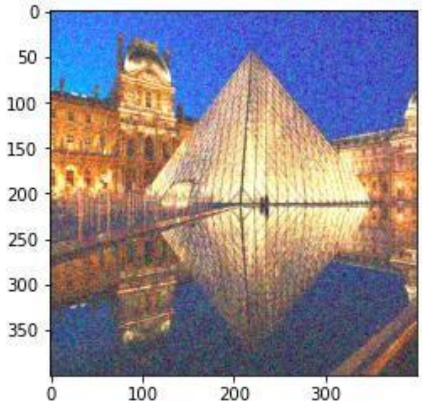

Fig.3.2 Generated Noisy Image

Overall, initializing the "generated" image as a noisy but correlated version of the content image is a purposeful decision aimed at speeding up the NST optimization process. By achieving a balance between exploration and exploitation, this initialization technique establishes the framework for creating stylistic graphics that accurately represent both the user's intended content and style features.

3.3 Load Pre-trained VGG19 Model

To begin this challenge, we'll load the VGG19 model, which is a common convolutional neural network architecture [19]. We'll then define a function to extract the outputs from the model's middle levels. These middle layers are commonly used to capture both low- and high-level properties, making them appropriate for conveying content and stylistic information. Next, we'll define the content layer, which will be used as a reference to preserve the input image's content. This layer is typically selected from the middle levels of the network to capture significant content representations. Next, we'll create our custom model by providing the inputs and outputs. The photos will be used as inputs, and the middle layer activations will be acquired using the previously stated function. After creating the model, we'll run the content and style pictures through it to get the activations for both layers. These activations will be kept in different variables that will be used as targets during the optimization process in Neural Style Transfer. By structuring our code in this manner, we establish a framework for extracting relevant features from the VGG19 model and preparing them for subsequent style transfer operations, ensuring efficient and effective generation of artistic images.
Example,
content_target=vgg_model_outputs(content_image)
style_targets = vgg_model_outputs(style_image)

3.4 Compute Total Cost

3.4.1 Compute Content Cost (Jcontent (C,G))

Neural Style Transfer (NST) aims to ensure that the generated picture (G) matches the content image (C). One good way to accomplish this is to calculate the content cost function. The content cost function quantifies the difference between the generated picture's content representations and the actual content image. Typically, this is accomplished by comparing the activations of a specific layer in the neural network, such as the VGG19 model, across both images. The activations of this selected layer capture the images' content information, representing features at various degrees of abstraction. By reducing the content cost function during the optimisation process, the generated image is encouraged to match the content image while retaining its important elements and structures. This ensures that the created image contains the same semantic content as the content image while also includes the intended artistic style from the style image.

Overall, the content cost function is critical in leading NST's optimisation process towards visually appealing solutions that properly capture the material of the content image while adding the desired artistic style.

Content Cost can be defined as:

$$J_{Content}(C,G) = \frac{1}{4 \times n_H \times n_W \times n_C} \sum \left(a^{(C)} - a^{(G)}\right)^2 \qquad (1)$$

Here, nH , nn , nc are the height, width and number of channels of the hidden layer you have chosen, and appear in a normalization term in the cost.
In order to compute the cost $J_{content}(C,G)$, it might also be convenient to unroll these 3D volumes a^C and a^G into a 2D matrix.

3.4.2 Compute Gram Matrix (G_{gram})

In linear algebra, the Gram matrix, also known as the "style matrix" in the context of Neural Style Transfer (NST), is a fundamental concept that captures the style information of a set of vectors. Specifically, given a set of n vectors $\{v1,v2,...,vn\}$, the Gram matrix G is constructed by computing the dot products between all pairs of vectors and arranging the results in a matrix.

Mathematically, the Gram matrix G is defined as:

$$G_{ij} = v_i v_j = np.dot(v_i, v_j) \qquad (2)$$

In the context of NST, the Gramme matrix is calculated using feature maps derived from certain layers of a convolutional neural network, such as VGG19 [19]. These feature maps encode an image's stylistic information by identifying correlations between its many components. By computing the Gramme matrix of these feature maps, we have a representation of the image's style, which can subsequently be utilised to quantify the style difference between images and direct the style transfer process.

Overall, the Gramme matrix provides a compact representation of style information, allowing for efficient computation and manipulation of style aspects in NST and other applications.

3.4.3 Compute Style Cost ($J_{style}(S,G)$)

In the context of Neural Style Transfer (NST), the following step after matching the generated image's content to the content image is to incorporate the desired artistic style from the style image. One effective way is to minimise the distance between the Gramme matrix of the style image (S) and the generated picture (G). The Gramme matrix represents the correlations between various elements to capture an image's style information. We intend to align their style representations by minimising the distance between the style image's Gramme matrices and the generated image. This guarantees that the created image retains the stylistic qualities, textures, and patterns from the style image. The distance between Gramme matrices is normally measured using an appropriate distance metric, such as the Frobenius norm or the mean squared error. Minimising this gap allows the generated image to follow the same style correlations as the style image, allowing the transfer of artistic styles.

Overall, reducing the distance between the Gramme matrices of the style picture and the generated image is an important step in NST, allowing for the creation of visually appealing images that blend the content of one image with the creative style of another.

The corresponding style cost is defined as:

$$J_{style}^{[l]}(S,G) = \frac{1}{4 \times n_C^2 \times (n_H \times n_W)^2} \sum_{i=1}^{n_C} \sum_{j=1}^{n_C} (G_{(gram)i,j}^{(S)} - G_{(gram)i,j}^{(G)})^2 \qquad (3)$$

3.4.4 total_cost

In Neural Style Transfer (NST), the overarching objective is to generate an image G that effectively balances both content preservation and style transfer [4]. To achieve this, we define a cost function $J(G)$ that combines both the content cost Jcontent (C,G) and the style cost$J$style(S,G), weighted by hyper-parameters α and β respectively.

The formula is:

$$J(G) = \alpha J_{content}(C,G) + \beta J_{style}(S,G) \qquad (4)$$

Here, Jcontent(C,G) represents the content cost, which measures the difference between the content of the content image C and

the generated image *G*. Similarly, *J*style(*S*,*G*) denotes the style cost, quantifying the disparity between the style of the style image *S* and the generated image *G*.
The hyper-parameters α and β control the relative importance of content and style in the final generated image. Adjusting these hyper-parameters allows for fine-tuning the trade-off between content fidelity and artistic style. A higher value of α emphasizes content preservation, while a higher value of β prioritizes style transfer.

To achieve a visually appealing image that captures both content and style, we optimise the cost function *J(G)* in relation to the generated image *G.*

3.5 Optimizer and Learning Rate Selection

3.5.1 ADAM

After computing the content and style costs, the next step in our NST process is to use the Adam optimizer's 'train_step()' function to do transfer learning. This function updates the parameters of the created image iteratively, minimising the total cost *J.* Using TensorFlow's GradientTape, gradients of total loss are produced in relation to the resulting image. Using pre-trained VGG19 models, we extract content features from the produced image and compare them to those in the content image to determine content loss [21]. Similarly, the style loss is calculated by comparing the Gram matrices of features from the generated image to the style image [21]. We optimise based on the overall loss, which is a weighted sum of content and style losses scaled by hyper-parameters α and β, respectively. Using the Adam optimizer with a predefined learning rate of 0.03, we continually update the pixel values of the output image, iteratively optimising its appearance to efficiently combine content preservation with style transfer. This iterative optimisation cycle is repeated until convergence, which results in a created image that smoothly incorporates both the intended content and style attributes.

3.5.2 Train the Model

Once the optimisation process is completed, the resulting image is a harmonious synthesis of the content image's content and the style image's artistic style. This final stylized image is the result of the Neural Style Transfer (NST) algorithm, representing a unique synthesis of two independent visual elements. It is worth noting that GPUs are commonly used for NST model training because of their parallel processing capabilities, which speed up the optimisation process. GPUs accelerate gradient computation and convergence, allowing for more efficient style transfer procedures. The learning rate is an important element in the optimisation process since it determines the magnitude of alterations made to the parameters of the generated image during each iteration. Increasing the learning rate can quicken the style transfer process, resulting in faster convergence and shorter computation times. However, such rapid growth frequently comes at the cost of image quality. Higher learning rates may cause overshoot or instability, producing in artefacts or distortions in the final stylistic output. Indeed, finding the right balance between speed and quality is critical in NST. Higher learning rates may speed up the process, but they may reduce the authenticity and aesthetic integrity of the stylized image. Lower learning rates, on the other hand, offer smoother convergence and higher-quality outputs while perhaps prolonging the optimisation process.

The procedures listed above provide a high-level overview of implementing NST with VGG 19. However, attaining the desired artistic effect frequently requires fine-tuning and hyperparameter testing. Changing hyperparameters like learning rate, content-weight (α), and style-weight (β) can greatly affect the stylized output and achieve the desired aesthetic expression.

Furthermore, beyond hyperparameters, the choice of content and style images has a significant impact on the output of NST. A careful selection of content photos with distinct visual qualities, as well as style images with engaging artistic styles, increases the diversity and richness of the created styled images. To summarise, while NST with VGG19 provides a powerful foundation for artistic picture synthesis, attaining optimal results necessitates a deliberate and iterative approach [19]. Experimentation, fine-tuning, and attention to detail are critical components of the process, allowing designers to express their creativity and create attractive stylized artworks.

4 Results and Discussion

The Art Generation yielded promising results and generated insightful discussions regarding Transfer Learning. Here, presenting the major outcomes and discuss their implications:

4.1 Cost Optimization

Cost optimisation in VGG19 with the Adam optimizer is the process of training the VGG19 model for a specific job while minimising a predefined cost or loss function[19][21]. Initially, the pre-trained VGG19 model is loaded, with the top classification layers removed to focus purely on feature extraction. Custom layers are then added to tailor the model to the task at hand, in this case, style transfer. These custom layers are intended to help extract and compare content and stylistic aspects from photos. The content cost and the style cost are the two key components that are usually involved in the computation of the cost function in Neural Style Transfer (NST). The difference in content between the content image and the generated picture is measured by the content cost, while the difference in style between the style image and the generated image is quantified by the style cost.
In a recent implementation, the content cost was computed to be tf.Tensor(0.007701091, shape=(), dtype=float32), indicating a

relatively low divergence between the content of the content image and the generated image. This suggests that the generated image effectively preserves the content of the content image, a critical aspect of successful style transfer. Consecutively, the Style Cost was calculated to be tf.Tensor(426.75244, shape=(), dtype=float32), indicating a moderate divergence between the style of the style image and the generated image. This suggests that while the generated image captures some stylistic elements from the style image, there is still room for improvement in terms of achieving a closer match to the desired artistic style. The overall cost is determined by adding the content and style expenses separately and accounting for the hyperparameters alpha=10 and beta=40. The final styled output's finalised style and content relative priority is controlled by these hyperparameters. Whereas a higher beta number gives priority to style transmission, a higher alpha value stresses content retention. The total cost was tf.Tensor(19354.729, shape=(), dtype=float32) prior to using the Adam optimizer, which suggests a pretty significant overall divergence between the created picture and the content and style images. This shows that there is a great deal of space for improvement when it comes to matching the produced image's content and style to the content and style of the original photos.
The overall cost was lowered to tf.Tensor(5759.637, shape=(), dtype=float32) when the Adam optimizer was used. This significant drop in the overall cost indicates that the generated image's parameters were successfully optimised by the Adam optimizer to reduce the divergence between the generated image's content and style and those of the content and style pictures. This showed that the Adam optimizer in NST was effective since the final stylistic output resembled the content and style pictures more [12].

In summary, cost optimization in VGG19 using the Adam optimizer involves iteratively updating the parameters of the generated image to minimize the divergence between its content and style and those of the content and style images. By leveraging the Adam optimizer, significant improvements in the total cost were achieved, leading to a more faithful and aesthetically pleasing stylized output.

4.2 Generation of artistic image

Neural Style Transfer (NST) is a fascinating art-technology fusion technique that creates beautiful pictures by imbuing a selected image's content with the stylistic components of a reference artwork [18]. This novel method preserves the content's underlying structure while skilfully integrating the complex patterns and textures that characterise the artistic style. It does this by using deep convolutional neural networks (CNNs) to extract high-level features from both the content and style images. NST functions fundamentally by minimising a loss function that measures the variations between the representations of the produced picture, content image, and style image. This loss function functions as a compass, steering the optimisation procedure in the direction of converting the content image into a cohesive piece of art that captures the spirit of the style image.

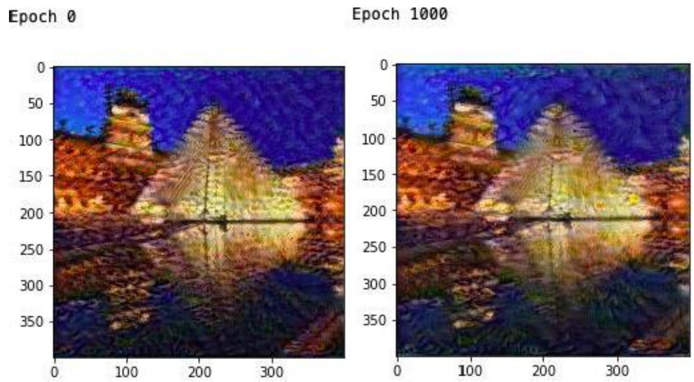

Fig.4.1 Image Generation

One of the key strengths of NST lies in its ability to capture and transfer complex artistic styles from reference artworks onto arbitrary content images. Whether it's the bold brushstrokes of Van Gogh's "Starry Night" or the geometric abstractions of Mondrian's compositions, NST can seamlessly apply these stylistic characteristics to a wide range of content images, resulting in visually stunning and evocative artworks.
Furthermore, NST provides an adaptable framework for artistic picture alteration and investigation [7]. By experimenting with numerous artistic styles and contrasting them with varied content pictures, designers and artists may create visually striking compositions that are both innovative and captivating. This adaptability creates endless creative opportunities by enabling the integration of many artistic influences and the investigation of novel aesthetic trajectories. NST has practical applications in a variety of disciplines in addition to the arts. When it comes to creating digital art, NST gives artists a strong instrument for creating original and expressive works of art, allowing them to experiment with new mediums and push the limits of conventional creative forms. NST may also be used to improve visual material in industries like marketing, entertainment, and advertising. This can result in visually striking content that successfully conveys brand identities to viewers. Additionally, NST is promising for industries where visual aesthetics are important, including fashion, architecture, and interior design. Designers can successfully envisage and convey their creative vision through the application of stylistic aspects from renowned artworks or design trends to virtual representations or prototypes. This helps to facilitate decision-making and cooperation. NST is a useful instrument for researching the fundamentals of creative style and visual perception in the fields of education and research. To improve our knowledge of how neural networks perceive and alter visual information, researchers might investigate issues pertaining to style transfer techniques, image representation learning, and computational aesthetics.
NST is not without difficulties and constraints, though, even with its amazing talents. It can be difficult and time-consuming to fine-tune hyperparameters, such as the weights given to content and style losses; rigorous testing and validation are necessary [16]. Furthermore, it might be difficult to strike a balance between stylistic coherence and content accuracy, especially when

working with complicated or abstract styles. Furthermore, content pictures that lack identifiable characteristics or have intricate spatial arrangements may provide challenges for NST. In these situations, the created artworks could seem warped or unidentifiable, emphasising how crucial it is to choose the right content pictures and styles in order to get the best outcomes. To sum up, Neural Style Transfer is a potent and adaptable method for producing artistic images that combines innovative technology with artistic expression. NST provides a plethora of chances for research and invention, ranging from the production of digital art to useful applications in several industries. We may anticipate further developments in algorithmic approaches as this field of study progresses, opening up new avenues for artistic expression and expanding the bounds of visual inventiveness.

4.3 Style Transfer

In the realm of Neural Style Transfer (NST) with VGG19, the term "style transfer" encapsulates a transformative process where the artistic essence of one image, known as the style image, is infused into the content of another image, referred to as the content image [19]. This intricate procedure aims to create a new image, often denoted as the "stylized" or "output" image, that harmoniously merges the content of the content image with the stylistic elements extracted from the style image. The ultimate objective is to generate a visually compelling artwork that encapsulates the essence of both the content and style images, resulting in a unique and evocative composition. A sequence of computational stages, led by deep convolutional neural networks (CNNs), especially the VGG19 architecture, which is well-known for its efficiency in feature extraction tasks, are involved in the style transfer process. Fundamentally, style transfer encodes and captures the stylistic elements and semantic information of incoming pictures using the hierarchical representations that VGG19 has learnt [20]. Loading the pre-trained VGG19 model with its top classification layers removed to concentrate just on feature extraction is the first step in the style transfer process. The model is then modified with the addition of custom layers to suit the goal of style transfer. The ability to extract and compare content and design aspects from the input photos is greatly aided by these custom layers. The finished artwork's structural framework and semantic context are derived from the content image, which forms the basis of the styled output. In the meantime, the styled picture receives the appropriate aesthetic attributes from the extracted and synthesised creative components found in the style image, such as colours, textures, and patterns. The content cost and the style cost are the two primary factors that determine how well a style transfer goes [19]. The content cost indicates how closely the generated picture maintains the semantic content of the content image by comparing the content characteristics derived from the content image with the created image. On the other hand, the style cost measures the stylistic disparity between the created picture and the style image. This is accomplished by encoding the distinct stylistic traits of the style picture and comparing the Gramme matrices of the style features that were retrieved from both images. Additionally, statistical correlations between various feature maps are captured. The original random noise or content-initialized image is progressively transformed into a styled masterpiece by NST through the minimization of a loss function that integrates the costs of content and style. Iterative optimisation methods, such gradient descent or its variations, are commonly used for this operation. They modify the resulting image's pixel values repeatedly in order to minimise the total loss. The hyperparameters alpha and beta are essential for regulating the relative weights of style transfer and content retention throughout the optimisation process. The weight given to the content cost is determined by alpha, while the weight given to the style cost is determined by beta [20]. Artists and practitioners can accomplish desired creative effects by customising the style transfer process by fine-tuning certain hyperparameters. At convergence, the styled picture appears as the finished product of the style transfer procedure, representing a smooth combination of style and content. The styled picture highlights the vivid colours, complex textures, and stylistic subtleties of the style image while reflecting the structural components and semantic information of the content image.

In conclusion, Neural Style Transfer with VGG19 provides a strong foundation for creative picture synthesis, making it possible to produce visually arresting and profoundly moving artworks. NST bridges the gap between traditional arts and contemporary technology by enabling artists and practitioners to explore new worlds of creativity and expression via the use of deep learning techniques and computational algorithms. We should expect more advancements in style transfer algorithms as this field of study develops, opening up new possibilities for creative expression and expanding the bounds of visual inventiveness.

5 Conclusion

The art generation project that employs Neural Style Transfer (NST) has unquestionably had extraordinary success in fusing the visual style of one image with the content of another to create appealing styled artworks [20]. To create visually arresting and emotionally impactful artworks, this project has required a voyage of creative inquiry, painstaking hyper-parameter tweaking, and the resolution of computational obstacles. The use of assessment criteria to gauge the calibre of the created artwork has been essential to the project's success. These measurements have given important new information on how well the NST algorithm transfers the intended creative style while maintaining content integrity. By quantitative criteria, such as style transfer accuracy metrics or perceptual similarity scores, the project has improved our comprehension of this intriguing nexus between art and technology and provided light on the basic workings of NST as well as its possible uses. Future research should focus on multimodal style transfer, which involves extending NST to enable the simultaneous transmission of several creative styles onto a single content image [2]. This could entail methods for combining and interpolating components from many stylistic representations, enabling the production of hybrid artworks that incorporate aspects from several creative sources. The synthesis of many aesthetic influences and creative innovation are made possible by multimodal style transmission. In conclusion, the art generation project using Neural Style Transfer has showcased the transformative potential of AI in enhancing artistic expression and offers exciting opportunities for future enhancements and artistic exploration. By leveraging NST, artists and designers can create visually stunning and emotionally resonant artworks that push the boundaries of traditional artistic mediums and redefine the possibilities of digital content creation. As research in this field continues to advance, we can expect to see further innovations in style transfer algorithms and applications, empowering creators to unleash their creativity and inspire audiences worldwide.

References

[1] L. A. Gatys, A. S. Ecker, and M. Bethge, "A Neural Algorithm of Artistic Style," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 16, no.12,p.326,Sep 2016.
[2] D. Ulyanov, V. Lebedev, A. Vedaldi, and V. Lempitsky, "Texture Networks: Feed- forward Synthesis of Textures and Stylized Images," in Proceedings of the European Conference on Computer Vision, vol. 12, no.2,p.234, 2016.
[3] L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture Synthesis Using Convolutional Neural Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 28, pp.262-270,Dec. 2015.
[4] M. Ruder, A. Dosovitskiy, and T. Brox, "Artistic Style Transfer for Videos with Temporal Coherence," in Proceedings of the European Conference on Computer Vision, vol. 8, pp.22-70, 2017.
[5] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.694-711, 2016.
[6] V. Dumoulin and J. Shlens, "A Learned Representation For Artistic Style," in Proceedings of the International Conference on Learning Representations, 2016.
[7] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Improved Texture Networks: Maximizing Quality and Diversity in Feed-forward Stylization and Texture Synthesis," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.94-121, 2017.
[8] A. W. Harley, A. Dosovitskiy, and T. Brox, "Demystifying Neural Style Transfer," in Proceedings of the International Conference on Computer Vision, pp.45-71 2017.
[9] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 20, pp.62-270, 2017.
[10] https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam
[11] J. Scholz, C. Reinke, P. Andreae, and J. Gall, "Fast Stable Reinforcement Learning through Robust Value Function Approximation," in Proceedings of the IEEE International Conference on Robotics and Automation, vol. 114, pp.3521-3526, 2018.
[12] https://iq.opengenus.org/vgg19-architecture/
[13] X. Huang and S. Belongie, "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 5, 2017.
[14] M. Ruder, A. Dosovitskiy, and T. Brox, "Artistic Style Transfer for Videos," in Proceedings of the European Conference on Computer Vision, pp.26-36, 2016.
[15] X. Huang and S. Belongie, "Exploring the Structure of a Real-time, Arbitrary Neural Artistic Style Transfer System," in Proceedings of the European Conference on Computer Vision, vol.2, pp. 26-43, 2017.
[16] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Artistic Style Transfer with ConvNets and Decoders," in Proceedings of the European Conference on Computer Vision, vol. 4, 2017.
[17] C. Scholz, J. Peters, and A. A. Faisal, "Fast Stable Reinforcement Learning through Robust Value Function Approximation," in Proceedings of the IEEE Conference, vol.114, pp. 3521-3526, 2017.
[18] L. Gatys, A. S. Ecker, and M. Bethge, "Preserving Color in Neural Artistic Style Transfer," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.2, pp. 114-150, 2016.
[19] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations *(ICLR)*.
[20] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proceedings of the International Conference on Learning Representations *(ICLR)*, 2015.
[21] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in International Conference on Learning Representations (ICLR), 2015.