Artistic Image Generation Using Neural Style Transfer

Ragipati Karthik[1], Anjali Paladi[2], Barre Ruchitha[3], Challa Taruni[4], Chinni Akanksha[5]
[1,2,3,4,5]*Department of Artificial Intelligence and Machine Learning, Malla Reddy University Hyderabad, India*

***Abstract*—Neural Style Transfer (NST) refers to a deep learning method that produces artistic images through a combination of content and style of another image. This paper is an artistic image generating system that works on an arbitrary style transfer model that is built around an already trained Convolutional Neural Network, which is VGG19. The suggested solution supports several invisible styles with effective feedforward inferencing, unlike the conventional techniques that involve retraining whenever the style changes. Content and style features have been mined out of various network layers and optimized loss function is used to trade content integrity and aesthetic appearance. Visual appeal of aesthetically pleasing stylized images achieved through experimental results, but structural integrity is preserved, indicates that neural style transfer is useful in creative and digital art applications.**

***Index Terms*—Neural Style Transfer, Arbitrary Stylization Model, Deep Learning, Convolutional Neural Network, VGG19.**

I. INTRODUCTION

Neural Style Transfer (NST) is a deep learning methodology, which allows creation of the artistic images by integrating the content of another image with the style of another image. It uses Convolutional Neural Networks (CNNs) to isolate and be able to differentiate content and style representations across various layers of an existing network. NST has received a lot of attention over the recent years because of its use in meaningful applications in digital art, visual media, and creative design. The conventional style transfer processes are not always limited to one artistic style, and would have to be retrained on any new style, which is computationally intensive and less scalable. To overcome these shortcomings, this study uses an arbitrary style transfer model that uses a pre-trained CNN, namely, VGG19, and enables the transfer of unseen styles through a series of applications, and without retraining. The proposed method can produce aesthetic visual artistic images with minimal content loss and style loss and maintain the integrity of the content image, which proves that deep learning can be effectively used to generate creative image formats [1].

A. Motivation & Contribution

The conventional image art demands imagination, time, and physical labor so that it is not accessible to all users. Automated methods of artistic generation have been of critical importance with the growing demand of digital art and visual media. The Neural Style Transfer is one such promising solution that can create an environment where machines can learn the visual patterns of artworks based on the image and then apply it to new pieces of information without a human operator. B. Contributions The key findings of this work include the following:

- Generation Framework of artistic images. Creation of a deep learning-based artistic image Neural Style Transfer.
- CNN-Based Features Extraction. Extraction of hierarchical content and style representations using a pre-trained convolutional neural network.
- Optimization Strategy based on losses. Combination of content loss, style loss and total variation loss to enhance smoothness and quality of visuals.
- Arbitrary styles can be supported. Capability to use various artistic styles without re-training the model.
- Live Applications: Creative. Illustration of NST in the field of digital art, media design and visualization of creations.

B. Literature Review

Neural Style Transfer (NST) has been widely studied to be useful in the improvement of the creation of artistic images through the effective separation and recombination of content and style representations. Early CNN-based methods like SAFIN came up with self-attention and enhanced normalization methods to learn global style patterns without altering the structure of the content albeit at a higher computational cost [2]. Models which were trained using contrastive learning achieved better content-style disentanglement and visual consistency across various styles but needed complicated training processes and high-quality samples [3]. Aesthetic-conscious and affinity-based attention networks also increased the quality of visual representations and content preservation at the finer-grained scale at the expense of increased memory and inference time [4], [1]. The recent hybrid CNN Transformer architectures enhanced global context modelling and consistency in style distributions but with a huge computational complexity [5]. Controllable and object-aware style transfer other works were devoted to style transfer, which can be finely-stylized and semantically controlled, but used expensive optimization or accurate segmentation models [6]. Permutational based creative techniques investigated expressive stylization, but lacked structural consistency at times [7]. These works indicate a trade-off between the quality of stylization, control, and efficiency, and in the spirit of encouraging an arbitrary style transfer method that encourages the use of several unknown styles without retraining and maintaining the same level of performance [3][5][14].

II. PROBLEM STATEMENT

The creation of artistic images maintaining content order and at the same time reproducing the complex artistic styles with the desired precision is a complicated task. The majority of Neural Style Transfer algorithms either distort the structure of images through over stylization or do not reproduce the details of the artistic pieces well. Also, there are methods that involve retraining of new styles or they are very computationally expensive. Thus, the research question of this study is to develop a highly effective system of artistic images generation with Neural Style Transfer that preserves the original content and enables the use of various artistic styles as well as can create a visually appealing outcome with computationally reasonable complexity [8][7].

III. METHODOLOGY

The suggested system adheres to an organized channel of generating artistic images in the Neural Style Transfer manner.

A. Image Input and Preprocessing

As input, a content image and a style image are given. The images are both augmented to the fixed resolution, normalized, and transformed to the form of tensors to make them compatible with the convolutional neural network.

B. CNN Extraction of Features

A pre-trained convolutional neural network helps in extracting hierarchical feature representation of the two images. The content structure is captured by deeper layers and artistic textures and patterns of colours are captured by the intermediate layers.

C. Representation of Content and Style

Directly taken CNN feature maps are used as content representation. Style representation is calculated based on statistical dependencies between feature maps, which makes it possible to capture global artistic trends.

D. Loss Function Computation

The weighted loss function is characterized by three items: content loss which maintains structure, style loss which enforces artistic resemblance and total variation loss which enhances smoothness and minimizes noise.

E. Image Optimization

Gradient-based optimization is used to optimize the generated image in a bid to reduce the overall loss. The image is refined in every iteration to represent the content and style better.

F. Artistic Image Generation

The resulting image will maintain the semantic meaning of the original image but will have the artistic style, resulting in aesthetic and attractive images. The method makes arbitrary transfers of style without retraining possible [1] [9].

IV. SYSTEM ARCHITECTURE

A. Overall Architecture

The proposed image generation system based on Neural Style Transfer is a modular and sequential system. There are major components of the system which include: 1. User Input Module: Takes the content image and the style image that a user gives. 2.Preprocessing Module: Conducts image resizing, normalizing and converting the images to tensors to ensure they are suitable to the deep learning models. 3.Feature Extraction Module: With the help of a trained convolutional neural network, it is used to produce hierarchical content and style feature representations. The style transfer and optimization module is utilized to transfer the target image to the generalized image in the manner of the provided example. ¡—human—¿4. Style Transfer and Optimization Module: the following module is the style transfer and optimization module; it is used to transfer the target image to the generalized image as in the case of the given example. Calculates content loss, style loss and total variation loss, and optimizes the generated image iteratively to make artistic stylization. 5.Output Generation Module: Produces the completed stylized image without altering the structural contents and artistic features. The general pattern of the system working is as:

B. Feature Extraction Layer

The extraction layer of features is realized based on a pre-trained convolutional neural network, i.e. a VGG-based architecture. The layer records valuable visual information in several levels: • Edges, textures and colour patterns are extracted by lower layers. The deeper layers are used to store high level semantic and structural information. These content and style characteristics of the input images are represented separately by these extracted feature maps.

C. Optimization and Loss Layer

The layer is used to direct the synthesis process of the image by computing the following loss functions:

- Content Loss: Retains the spatial information as well as semantics of the content image.
- Style Loss: This is a style that guarantees similarity of the created image with the style image in terms of texture and artistic patterns.
- Total Variation Loss: Smooths noisia and enhances visual smoothness.

These losses are weighted to guide the optimization process to produce a visual, coherent artistic image.

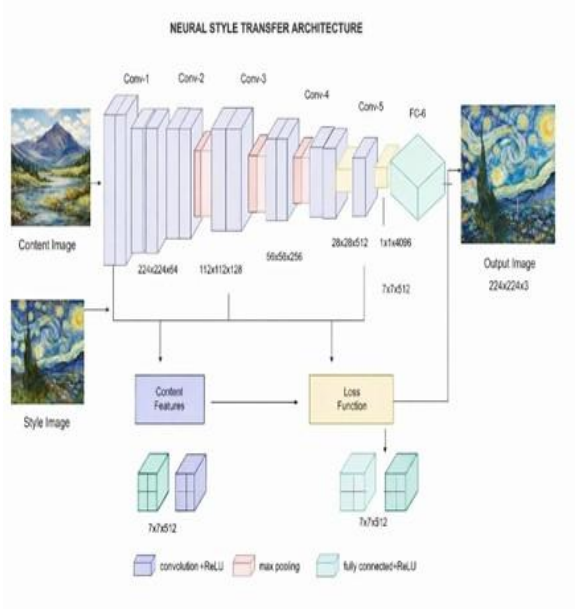

Fig. 1. Architecture

D. Output Layer

Once the image is optimized, it is put back in the image format. The resulting output image has managed to combine the artistic style without compromising the original content layout and visual clarity.

E. Scalability and Modularity

The modular structure permits the independent modification of the preprocessing, feature extraction and optimization modules. Future extensions to this design include real time style transfer, video stylization, multi-style learning and style intensity control by the user.[8][12].

V. DATASET DESCRIPTION

The diversity and quality of the images used in the process of experimentation has a dominant impact on the effectiveness of an artistic image generation system based on Neural Style Transfer (NST). The set of publicly available images used in this work is represented by real-world photographs to represent the content and artistic paintings to represent the style. The visual qualities represented in these images are

extensive; they allow the model to produce artworks that are richly visually varied.

A. Dataset Characteristics

The sample is a set of two sets of independent images:

- Content Image: this consists of natural scenes, objects and photographs that focus on human beings and give clear spatial and structural details.
- Style Images: The style images are artistic paintings having unique textures, brush strokes, and colour patterns. All images are contained in the conventional image formats and differentiate in terms of resolution, artistic sophistication and compositions of the image.

This variety means that the process of style transfer can be considered in the context of various artistic and architectural conditions.

B. Data Distribution

Since the Neural Style Transfer algorithm does not need labelled data, the dataset is not distributed in classes. Rather, there is a focus on preserving diversity in the content and artistic expression. Different combinations of content and style images are applied to test how the model will be able to generalize artistic patterns across the visual contexts.

C. The strategy to be used in data preprocessing will involve

Each image is subjected to a similar preprocessing pipeline so that they can be compatible with the deep learning framework. This involves scaling images to a predetermined input resolution, normalization of pixel intensity values and transformation of images into tensor representations. These measures can stabilize the optimization process and maintain significant visual characteristics that are required to extract the style effectively and retain the content.

D. Data Utilization

Instead of dividing data into training and test sets, the dataset is exploited by the means of experimental work that involves trial and error on the combinations of content and style images. The method enables qualitative assessment of the quality of stylization, content retention and artistic consistency in a series of visual situations [4].

VI. TRAINING CONFIGURATION AND MATHEMATCAL FORMULATION

This part gives a description of optimization setup and mathematical formulation as applied to artistic image generation with Neural Style Transfer. As opposed to traditional supervised learning, the NST framework is based on the idea of perceptual loss functions based on deep convolutional features that are optimized using iterations.

A. Optimization Configuration

This configuration was selected because it is the least expensive system to maintain in the long run. ¡—human—¿A. Optimization Configuration: The reason this configuration was chosen is that it is the least costly system to maintain over the long run. This system has an optimization-based methodology in which the output image is successively refined to reduce a weighted sum of both content and style losses. The generated image parameters are optimized and a pre-trained convolutional neural network is invoked with a purpose of extracting features only. The important configuration parameters are:

- Optimizer: Adam
- Learning Rate: 0.01
- Number of Iterations: 500–1000 • Content Preservation (alpha): Content preservation.
- Style Weight ( ): regulates artistic stylization.
- Total Variation Weight: Enhances the visual smoothness. The optimization process is brought to convergence and a visual stability is achieved in the generated image.

B. Neural Style Transfer

Let the content image be denoted as C, the style image as S, and the generated image as G. The objective of Neural Style Transfer is to generate an image G that preserves the content of C while adopting the artistic style of S.

Total Loss Function

The total loss function is defined as:

$$L_{total}(G) = \alpha L_{content}(C,G) + \beta L_{style}(S,G) + \gamma L_{tv}(G) \quad (1)$$

where:

- Lcontent measures similarity between content features of

C and G

- Lstyle measures similarity between style features of S and G

- Ltv represents total variation loss
- α,β,γ are weighting factors
- Content Loss

Content loss preserves the spatial structure of the content image and is computed as the squared difference between feature activations:

$$\text{Lcontent} = \frac{1}{2}\sum_{i,j}\left(F_{ij}^{G} - F_{ij}^{C}\right)^{2} \quad (2)$$

where FG and FC represent feature maps of the generated and content images extracted from a selected convolutional layer.

Style Loss

Style loss captures texture and artistic patterns using Gram matrices:

$$G_{ij} = \sum_{k} F_{ik}F_{jk} \quad (3)$$

The style loss is computed as the mean squared error between Gram matrices of the style and generated images across multiple layers.

C. Optimization Process

In the process of optimization:

- The image that is generated undergoes the feature extraction network.
- The computation of content, style, and total variation losses are carried out.
- The overall loss is backpropagated.
- The optimizer moves the pixel values of the generated image.

This looping process goes on until the image created is effective at balancing the content and artistic stylization [13] [11].

Fig. 2. Home Page of the Neural State Transfer

Fig. 3. User Interface for Artistic Image Generation

VII. DEPLOYEMENT AND RESULTS

The Neural Style Transfer model uses material and style elements retrieved with the help of a pre-trained CNN to create artistic images. The experimental results indicate the possibility of successful contents preservation and visually pleasing style transfer which has a smooth texture that mixes well hence proving that the proposed framework is reliable.

VIII. CONCLUSION

has a consistent and steady performance in experimentation. The model is successful in isolating the content structure and maintaining it and artistic textures and patterns of the style image are transferred. The stylized images productively created through the use of deep convolutional feature extraction and loss-based optimization are visually appealing and of high quality. In general, the suggested framework is an effective and stable method of artistic image creation, which justifies the effectiveness of the methods of neural style transfer.

IX. FUTURE SCOPE

Future improvements can involve adoption of faster and real time style transfer models in order to minimize validity

Fig. 4. Input and Output Visualization of Neural State Transfer

cost. It is possible to add the support of multiple styles at once, and style intensity under the control of the user. Advances can include video style transfer, more-resolution image generation and incorporation of transformer- based architectures to enhance the global consistency of style and visual quality.

REFERENCES

[1] Y. Li et al., "Understanding neural style transfer," in Proceedings of IJCAI, 2020. [Online]. Available: https://www.ijcai.org/Proceedings/2020/0309.pdf
[2] Y. Deng et al., "A survey of neural style transfer," IEEE Transactions on Visualization and Computer Graphics, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9093291
[3] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. [Online]. Available: https://ieeexplore.ieee.org/document/9144355
[4] X. Huang and S. Belongie, "Arbitrary style transfer in Realtime with adaptive instance normalization," IEEE International Conference on Computer Vision (ICCV), 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8619169
[5] J. Park and C. Lee, "Safin: Style-aware feature instance normalization," IEEE Access, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9710408
[6] L. Sheng et al., "Avatar-net: Zero-shot style transfer by feature decoration," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9157057
[7] J. Wang et al., "Frequency-aware feature extraction for neural style transfer," Signal Processing: Image Communication, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0923596522002486
[8] X. Wang et al., "Contrastive learning for neural style transfer," IEEE Transactions on Multimedia, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10041771
[9] L. Zhang et al., "Aeanet: Attention-enhanced neural style transfer," IEEE Access, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10298457