International Journal of Research in Engineering, Science and Management
Volume 7, Issue 3, March 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Interactive Art Generation with Style Transfer using Image URL

D. Likitha Sri Sravani[1*], T. Yuva Kishore[2], M. Midhun Reddy[3], B. Shiva Naga Durga Prasad[4]
[1,2,3,4]*Student, Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, India*

***Abstract*: This research study presents style transfer between images using VGG19 model, neural style transfer (NST) algorithms for an innovative art generation. The previous art generation methods use VGG16 model which doesn't predict or create images accurately. As the complexity of images increases, the NST algorithms needs to be trained to improve its performance. To overcome the problems, a convolutional neural network with 19 deep layers is introduced. For the frontend designers and artist, it is a challenging task to create new images for creating web pages. The process involves providing image URLs of two images-content image, style images to transfer styles from one to another. This art generation has emerged as a dynamic and innovative approach for digital art creation which enables individuals to actively participate in artistic process. The architecture adept at preserving content while effectively transferring desired stylistic elements. The aim is to provide a user-friendly interface for the end user which generates many new images in the future. The research study involves tensorflow Hub, which contains pre-processed machine learning algorithms.**

***Keywords*: Neural Style Transfer (NST), Art generation, Image URL, Tensorflow hub, Visual Geometry Group (VGG19), Deep learning approach, style image, content image, Convolutional Neural Network (CNN).**

1. Introduction

The evolution of artistic expression has always kept pace with technological developments. With the use of Convolutional neural network (CNN) common photographs can be transformed into visually captivating artworks by dividing their stylistic features and information [2]. Using deep neural networks such as VGG19 makes it possible to extract complex visual information at different abstraction.

By utilising the VGG19 model capabilities, this research article seeks to explore the field of interactive art generation through style transfer approaches. The features extracted from both the images i.e., content and style images are obtained by passing them through the VGG19 network [13].

The difference between the generated image and the independent output-which combines style and content is measured by a loss function. Three losses are measured in the process which are content loss, style loss and total loss. The goal is to minimize the total loss by updating the pixels of the generated image using optimization algorithms like gradient descent [9], [10]. Style transfer has many uses in real time for example in some military or satellite sensing images where images need to be captured in night vision, but the limitations will only capture only some contents of image, so the transformation of styles can make the image understandable. Understanding consumer's strong desire is the goal of examining the interactive design approach.

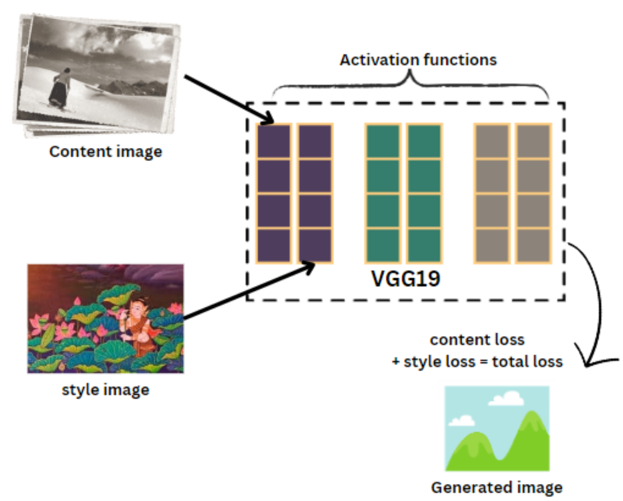

Fig. 1. Style transfer architecture

A. Art Generation Architecture & Process

In general, any art generation process involves utilizing deep learning techniques like Generative Adversarial Networks (GAN's) and Variational Autoencoders (VAE's) [1]. The common process for art generation involves data collection and preprocessing, Model training, Generation, Evaluation and refinement.

The art generation with style transfer architecture involves: Style Transfer Model and User interface (UI).

The style transfer model utilizes a pre-trained deep learning network such as VGG16, VGG19 or other architectures trained specially for style transfer. Usually, the model comprises of convolutional layers used for feature extraction at various levels of abstraction. On the other side an interactive interface (providing image URL's) that allows users to upload their content image, choose or adjust artistic styles and control all the parameters influencing the style transfer process. The user interface (UI) provides real-time feedback, displaying the

*Corresponding author: likitharamesh0309@gmail.com

evolving artwork during the style transfer process. The VGG19 [13] architecture belongs to Visual Geometry Group (VGG) family of models contains 19 layers that takes image input of size of 224*224 pixels with three colour channels (RGB). The CNN contains rectified linear unit (ReLU) activation function to introduce non-linearity. It uses the concept of padding to crop the image and to retain the contents of image without removing the edges.

In the process, the first step is providing input i.e., content, style image, the content image is uploaded by the user and it can be typically a photograph or an image they want to apply artistic styles to and users can select one or more style images which they want to refer to. Next step is Feature extraction in which the content and style images are passed through the pre-trained convolutional neural network (CNN) (VGG19) and features are extracted from them. Now, initialize the generated image that will be iteratively updated to incorporate the desired style while preserving the content. In the optimization loop [9], the features are reconstructed, the content features from the content image and the style features from the style images are compared with the features of the generated image at various layers of the network and loss functions are computed based on the differences between these features to guide the optimization process. The generated image is iteratively updated using optimization algorithms lie gradient descent to minimize the combined loss function. During each iteration, the image gradually transforms to better match the content of the original image with the style of the reference images. Finally, visualization and output artwork can be downloaded or save.

2. Literature Survey

The article discusses the key features and components of Generative adversarial networks (GAN), how generator wins over the discriminator and generates images that never exist. It also highlights on the types of GAN's like vanilla GAN and image to image generation (Cycle GAN). This also examine research on leveraging pre-trained GAN models for transfer learning tasks, domain adaptation and reusability of learned features across different images [1].

This article highlights the significance of Convolutional neural network (CNN) and explore architectural innovations and contributions of LeNet, AlexNet, VGG, GoogleNet, ResNet etc., also highlights their unique features and applications, implementation of CNNs for development on resource constrained devices [2].

This article highlights several challenges faced when applying convolutional neural network (CNN) in real-time like overfitting, vanishing gradients, need for large datasets, adversarial attacks, robustness to variations, addressing these challenges remains an active area of research in the field of deep learning [3].

The exploration of the Neural style transfer (NST) that separates and combines content and style features are represented in this article. It also discusses variants and enhancements of NST, deep learning architecture for NST, loss and optimization methods involved in it [4]. This article showcases how styles are transferred using neural style transfer (NST) algorithms and user interactions, control in NST, Various metrics that includes style consistency, content preservation and perceptual similarity to the reference image and explore discussions on the ethical considerations of NST which has copyright, ownership, and the impact of AI-generated content on art and creativity [5]. The article highlights on Adaptive Instance Normalization (AdaIN) that allows style transfer with more fine-grained control over the intensity of style transfer tasks. It also gives multi-modal and cross-domain applications [6]. The article emphasizes unsupervised image-to-image translation which is explained by cycleGAN. This technique maximizes the performance of interactive art generation, this focuses on cycle-consistency loss, generator-discriminator structure, day-to-night conversion. The image synthesis across different domains such as medical image translation using cycleGAN is the key point in this article [7].

It addresses the challenges and considerations in neural style transfer (NST) which involves computational complexity, Memory and storage requirements, fine-tuning style transfer parameters, interpretability and control, preserving content and semantic information [8]. The article describes the key features and components of optimization techniques which mainly involves gradient descent used to iteratively optimize the generated image to minimize the differences between content and style representations [9]. The article mainly discusses the loss function methodologies and the differences between features of the content and style images including mean squared error (MSE), L1 and L2 regularization and controlling model complexity, local minima and convergence behaviour [10]. This article leverages Tensorflow Hubs pre-trained models for transfer learning, fine-tuning and emphasizing the benefits of using pre-trained modules for downstream tasks [11].

The author highlights the Visual Graphics Group (VGG16) model and its architecture it also includes max-pooling, non-linearities, exploration of the trade-off between depth and computational efficiency, evolution of deep learning for computer vision [12].

This article delves into the architecture of VGG19 that builds upon the VGG16 resulting in increased depth and complexity, it has 16 CNN, 3 fully connected layers, padding, pre-trained models learned on ImageNet datasets, it also discusses ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [13]. The implementation of hyperparameters and control mechanisms are discussed in this article which involves content emphasis, blending multiple styles, iterative rendering [14]. This research article addresses how the Depth Extraction Generative Adversarial Network (DE-GAN) innovatively integrates a multi feature extractor, combining U-net, multi-factor extractor and MiDas depth estimation network [15].

3. Proposed System

To design an intuitive and user-friendly interface that allows user to upload their own content images and select style reference images, including options for users to adjust parameters, server-based backend system capable of handling image processing tasks and neural network computations by utilizing Tensorflow to build the style transfer models. In real-time feedback and rendering process, provide immediate visual feedback to users displaying the evolving artwork during the style transfer process. The user can provide their own output image size from the content and style images generated.

The implementation of error handling mechanisms provides guidance to users in case of invalid inputs or errors during the interaction process which supports various formats and resolution for output.

4. Software Implementation

In interactive art generation with style transfer, the software process typically involves a series of steps to enable user interaction and neural network-based image manipulation. Initially, the user interacts with a graphical user interface (GUI) that allows them to upload their desired content image and select or upload reference style images. Behind the scenes, the backend system uses deep learning models, often based on convolutional networks (CNNs) pre-trained on image classification tasks (VGG19), the image is loaded and pre-processed to extract content and style features from the uploaded images. Load and convert to float numpy array, add batch dimension and normalize to range [0,1] Using padding technique and tensorflow hub, image is cropped and contents of image retains from both content and style images. These features are then manipulated according to user-defined parameters, such as style intensity or content emphasis, using algorithms that perform the style transfer.

The software orchestrates this process, combining user input, neural network computations and iterative rendering to facilitate an interactive and creative experience, ultimately generating stylized art based on the user's preferences and interactions.

5. Result and Discussion

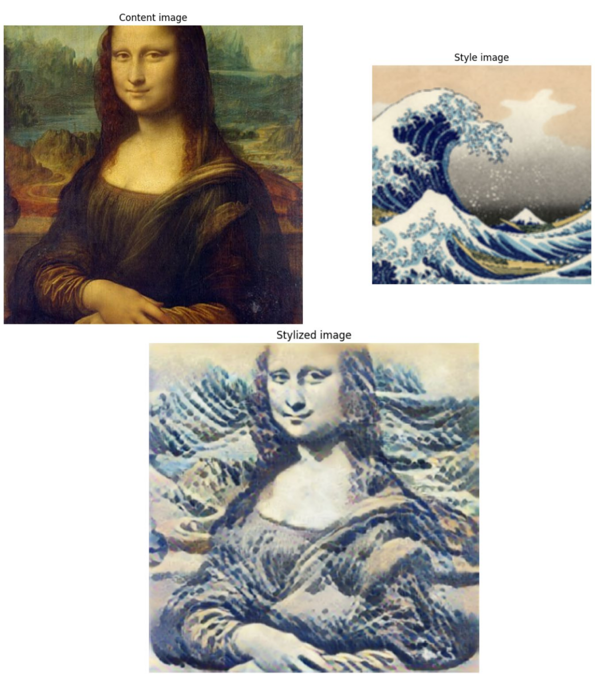

Fig. 2. Generated image

Therefore, the system displays the generated artwork in real-time, allowing users to see the evolving output as they interact with the interface, resulting in visually appealing and creative outputs. The visual outputs demonstrated a commendable fusion of styles onto diverse content, boasting high fidelity to the selected artistic influences while retaining clarity and coherence. Furthermore, while the system managed to produce diverse outputs, there remained opportunities for enhancing the variability of generated artworks.

However, the computational efficiency allowed near-real-time generation, enhancing user engagement. This study highlights the potential of interactive style transfer in art generation, suggesting avenues for refining algorithms and user interfaces to achieve even more compelling and diverse artistic outputs.

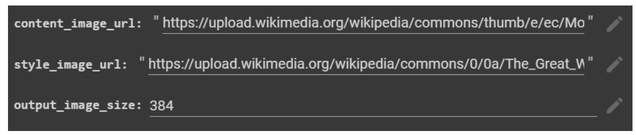

Fig. 3. User input

This shows options for users to adjust parameters and provide input images i.e., content image URL, style image URL and the output image size.

6. Conclusion

In summary, the creation of interactive art generation systems that make use of style transfer techniques is an intriguing combination of cutting-edge technology and artistic expression. These technologies enable people to freely explore their creativity by fusing their own visual content with a variety of artistic styles in real-time, by utilizing neural networks and pre-trained models, users are able to adjust parameters dynamically, see instantaneous visual artwork transformations. These systems convergence of artistic innovation, computational efficiency and user interaction opens up new creative directions. Users can interact with the creative process, try out different styles and ultimately produce customized artworks that transcend different mediums.

References

[1] Aggarwal, A., Mittal, M., & Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. *International Journal of Information Management Data Insights*, *1*(1), 100004.

[2] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, *9*(4), 611–629.

[3] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53–53.

[4] Liu, X. (2021). Research on Neural Style Transfer. *Journal of Physics. Conference Series*, *2079*(1), 12029.

[5] J. Liao, "A Study on Neural Style Transfer Methods for Images," *2022 2nd International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR)*, Xi'an, China, 2022, pp. 60-64.

[6] Huang, X., & Belongie, S. (2017). Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization.

[7] J. Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 2242-2251.

[8] A. Singh, V. Jaiswal, G. Joshi, A. Sanjeeve, S. Gite and K. Kotecha, "Neural Style Transfer: A Critical Review," in *IEEE Access*, vol. 9, pp. 131583-131613, 2021, doi: 10.1109/ACCESS.2021.3112996.

[9] Mehmood, F., Ahmad, S., & Whangbo, T. K. (2023). An Efficient Optimization Technique for Training Deep Neural Networks. *Mathematics (Basel)*, *11*(6), 1360.

[10] Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., & Chavez-Urbiola, E. A. (2023). Loss Functions and Metrics in Deep Learning.

[11] Yoo, Taewon & Chun, Minki & Bae, Yunjung & Kwon, Soohyun & Jung, Hyunggu. (2022). Exploring the Community of Model Publishers on TensorFlow Hub. 63-67.

[12] J. Tao, Y. Gu, J. Sun, Y. Bie and H. Wang, "Research on vgg16 convolutional neural network feature classification algorithm based on Transfer Learning," *2021 2nd China International SAR Symposium (CISS)*, Shanghai, China, 2021, pp. 1-3.

[13] Sudha, V. & Ganeshbabu, Dr. (2020). A Convolutional Neural Network Classifier VGG-19 Architecture for Lesion Detection and Grading in Diabetic Retinopathy Based on Deep Learning. Computers, Materials & Continua. 66. 827-842.

[14] Wang, X., Lyu, Y., Huang, J. *et al.* Interactive Artistic Multi-style Transfer. *Int J Comput Intell Syst* 14, 187 (2021).

[15] Han, X., Wu, Y., & Wan, R. (2023). A Method for Style Transfer from Artistic Images Based on Depth Extraction Generative Adversarial Network. *Applied Sciences*, *13*(2), 867.