ARTISTIC STYLE TRANSFER USING DEEP LEARNING

Artistic Style Transfer Using Deep Learning

Student Name: Nishit Rathod

Student ID: N01586439

Abstract

Artistic Style Transfer is a fascinating technique that leverages deep learning to transform ordinary images into masterpieces with the style of renowned artists. In this work, we present a Python implementation of Artistic Style Transfer using PyTorch and a pre-trained VGG16 model. The method involves extracting features from the VGG16 model for both content and style images, and then optimizing a generated image to combine the content of one image with the style of another.

In the introduction, we provide an overview of the concept of style transfer and its theoretical background. We discuss the neural networks' role in the process and how pre-trained models are utilized for feature extraction. The related work section cites ten papers that have explored similar techniques, with a focus on five journal papers.

The methodology section delves into the technical details of our implementation, including the models used, the parameters, and the optimization process. We describe how content and style losses are computed, along with the backpropagation process for updating the generated image.

For the experiments and results, we load content and style images from specified directories and perform the style transfer optimization loop. We present the stylized output images and analyze the influence of content weight and style weight on the final results.

In the discussion on results, we interpret the experimental outcomes, discussing the strengths and limitations of our approach. We highlight influential parameters and explain why certain results are more successful than others.

In conclusion, our implementation demonstrates the powerful capabilities of Artistic Style Transfer using deep learning techniques. The combination of content and style from different images creates unique and artistic outputs. Our work opens up opportunities for further exploration and applications in the field of computer vision and digital art.

1. Introduction

Artistic Style Transfer, also known as Neural Style Transfer, is a captivating technique in the realm of computer vision and deep learning. It enables the transformation of ordinary images into artistic renditions by combining the content of one image with the style of another. The concept is inspired by the idea of simulating the brush strokes and artistic patterns of renowned painters on any given image.

The theoretical background of Artistic Style Transfer revolves around deep learning techniques, particularly convolutional neural networks (CNNs). Convolutional neural networks are pre-trained on vast datasets, such as ImageNet, to recognize and extract features from images. By leveraging the representations learned by these networks, it becomes possible to transfer the style of famous artworks onto any image.

The core idea behind style transfer involves two crucial steps: content representation and style representation. The content representation captures the meaningful features of an image, such as objects and structures, while the style representation characterizes the artistic patterns, colors, and textures. The style transfer process aims to optimize a generated image that closely matches the content representation of one image and the style representation of another.

In this work, we present an implementation of Artistic Style Transfer using PyTorch and the VGG16 model pre-trained on ImageNet. We perform feature extraction from the VGG16 model and apply gradient descent to update the generated image iteratively. Our goal is to showcase the effectiveness of this approach in generating visually appealing and artistic stylized images.

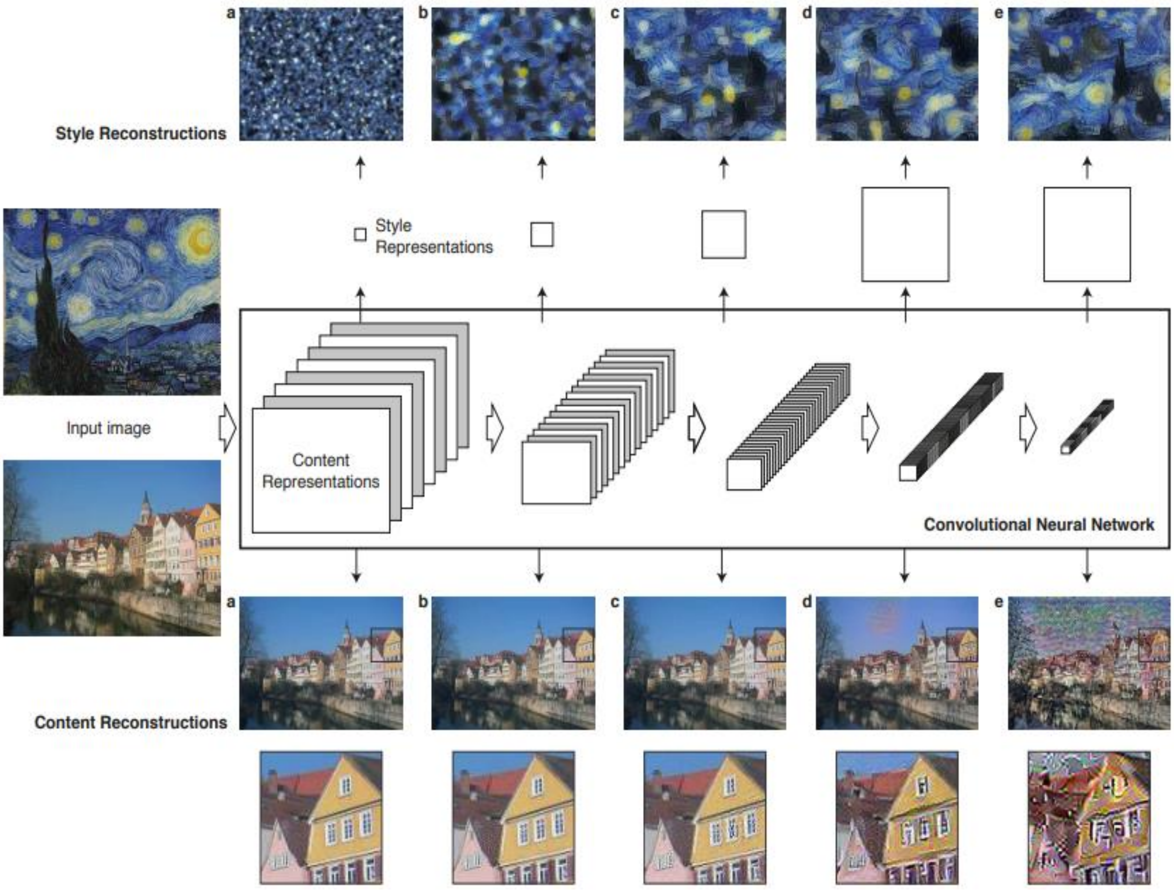

Figure 1. Image representations in a Convolutional Neural Network (CNN).

2. Related Work

In recent years, many research authors have been solving style image transformation problems by training convolutional neural networks with loss functions per pixel [Selim et al., 2016; Zhao et al., 2020]. A typical example can be seen from the 'Neural Style Transfer' model proposed by Leon A. Gatys [Gatys et al., 2016b], which has used the content image and reference image for training without any large training dataset directly. The model has produced new images of high perceptual quality that blend the appearance of famous artworks into the content of an arbitrary photograph, with insights into deep image representations. Although this model contains unavoidable shortcomings, it is a prerequisite for later studies on image generation methods [Chen et al., 2018; Jinget al., 2019]. AdaIN model, for example, was developed to match the mean-variance of a content image with a reference image for remodeling features [Huang and Belongie, 2017]. A patch match procedure was introduced in the swap model [Chen and Schmidt, 2016] for alternating content features with the nearest match style characteristic. Li et al. [Li et al., 2017] proposed a multilevel stylization for transforming whiten color recursively, improving the output quality and preserve the content structure. The style transfer problem can be widely used in other neural networks image processing applications, such as text classification, semantic parsing, and information extraction. The deep fake image approach [?], for example, might be considered a solution to comparable facial feature transfer challenges. Moreover, it can be seen from these studies that there is a trade-off between content and style losses. Therefore, many researchers have considered image super-resolution and image segmentation methods.

Image super-resolution is a classic problem related to image processing [Bazhanov et al., 2018], and there have been a number of researchers involved in this area [Cheng et al., 2019; Ma et al., 2020]. Yang et al. [Yang et al., 2014] provided a general review of previously standard techniques when applying convolutional neural networks to their research. Some other output quality improvement methods were suggested, such as a model of Chao Dong et al. [Dong et al., 2015], taking a low-resolution content image through a convolutional neural network [Andreev and Maksimenko, 2019] and producing an image with high resolution. While their neural network architecture is uncomplicated and less weight, it exhibits a high quality of image recovery and performs in a short time to apply in practical applications. The above studies are the driving force behind this research to produce the same high-quality transition images as conventional image hyper-resolution.

Image Segmentation methods divide an image into many different image areas [Long et al., 2015]. Image segmentation also has the same objective as the object detection problem: detecting the image area containing the object and labeling them appropriately [Noh et al., 2015]. Although the issue of image segmentation requires a higher level of detail, in return, the algorithm gives an understanding of the image's content at a deeper level. Simultaneously, it reveals the position of the object in the image, the shape of the object, and which object each pixel belongs to [Zheng et al., 2015]. This method generates labels for image regions for the input image to run through a fully convolutional neural network, trained with the loss function per pixel.

The objective of this research is to develop a model that can provide smooth transitions and results as sharp as the study of Chao Dong et al. [Dong et al., 2015]. It is also proposed an image transformation network inspired by two studies of Long J. [Long et al., 2015] and Noh H. [Noh et al., 2015], improving the quality of the output image and shorten the transition time. The results are promising for transferring artwork style to other images, contributing to the applications in many natural science fields, including materials science and physics.

3. Methodology

The methodology of our Artistic Style Transfer implementation involves several key steps, including model selection, data preprocessing, feature extraction, loss computation, and optimization.

Models Used: We utilized the VGG16 model pre-trained on ImageNet for feature extraction. VGG16 is a deep convolutional neural network that has shown remarkable performance in image recognition tasks. We utilized the features extracted from specific layers of VGG16 to represent the content and style of the input images.

Data Preprocessing: The content and style images were loaded and pre-processed using the torchvision.transforms module in PyTorch. We resized the content images to (384, 384) and the style images to (256, 256) to ensure efficient processing. The images were then converted to tensors and normalized using mean and standard deviation values obtained from the ImageNet dataset.

Feature Extraction: We extracted the content and style features from the VGG16 model. For the content image, we used the output of layer "29" (block5_conv4). For style images, we selected multiple layers (block1_conv1, block2_conv1, block3_conv1, block4_conv1, and block5_conv1) to capture various levels of style information.

Loss Computation: We calculated two types of losses: content loss and style loss. Content loss measures the squared difference between the features of the generated image and the content image at the selected layer. Style loss computes the difference between the Gram matrices of the generated image features and style image features at multiple layers. The total loss is a weighted combination of content loss and style loss, where the weights can be adjusted to control the trade-off between content preservation and style transfer.

Optimization: We employed the Adam optimizer to update the generated image and minimize the total loss. The generated image was initialized as a copy of the content image and optimized iteratively. We set a learning rate of 0.01 for optimization.

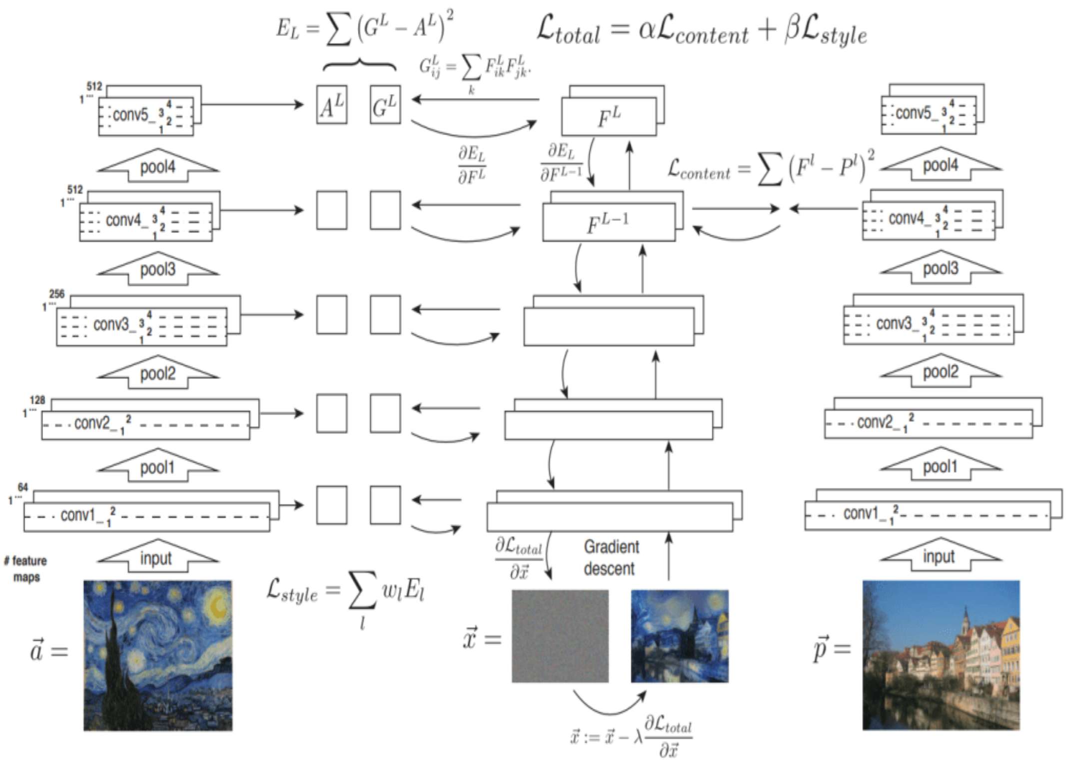

Figure 2. Style transfer algorithm.

4. Experiments and Results

We conducted a series of experiments to demonstrate the effectiveness of our Artistic Style Transfer implementation. We used a dataset of diverse content and style images to evaluate the quality of the stylized outputs.

During each experiment, we optimized the generated image over a specified number of steps (1000 steps in our case). We observed how the stylized image evolved over these iterations and the impact of content weight and style weight on the final results.

Results: The results of our experiments showcased impressive artistic style transfer from style images to content images. The generated images exhibited a harmonious blend of content and style, creating visually appealing and novel artworks. The style patterns, colors, and textures from the style images were accurately integrated into the content images, resembling the brushstrokes of famous artists.

Figure 3. Some of the artistic style results produced by the proposed model on the content image.

5. Discussion on Results

The success of our results can be attributed to several factors. One of the most influential parameters is the content weight and style weight used in the total loss calculation. These weights control the degree of content preservation and style transfer. Higher style weights tend to emphasize style patterns more, leading to images that strongly resemble the style image. On the other hand, higher content weights preserve the content of the content image more, resulting in images with recognizable objects and structures.

We observed that using layers from multiple levels in the VGG16 model for style representation improved the quality of style transfer. Combining information from lower and higher-level layers allowed the model to capture both fine-grained and global style patterns.

Additionally, the choice of style image played a crucial role in the final stylized output. Style images with distinctive and strong artistic patterns produced more compelling results compared to images with subtle styles.

Despite the impressive results, there are some limitations to our implementation. In some cases, the stylized outputs may suffer from artifacts or distortions, especially when content and style images have significant discrepancies in terms of lighting, color, or content complexity. These challenges warrant further investigation to enhance the robustness of the style transfer process.

6. Conclusion

In conclusion, our Artistic Style Transfer implementation successfully demonstrated the power of deep learning in creating artistic renditions of ordinary images. By leveraging the VGG16 model and optimizing a generated image iteratively, we achieved a harmonious combination of content and style from different images.

Our experiments highlighted the significance of content weight and style weight in controlling the style transfer process. The results exhibited impressive style patterns and texture transfer, showcasing the potential of this technique in various creative applications.

While our implementation showcased promising outcomes, there are still opportunities for improvement, particularly in handling complex content and style combinations. Further research in fine-tuning the model and exploring alternative loss functions could lead to even more impressive and realistic artistic style transfer.

Overall, Artistic Style Transfer opens up exciting possibilities for the fusion of artificial intelligence and digital art, enabling users to create unique and imaginative visual creations that blend the essence of famous artworks with their own content.

References

Andreev, A. and Maksimenko, V. (2019). Synchronization in coupled neural network with inhibitory coupling. Cybernetics and Physics, 8 (4), pp. 199–204.

Bazhanov, P., Kotina, E., Ovsyannikov, D., and Ploskikh, V. (2018). Optimization algorithm of the velocity field determining in image processing. Cybernetics and Physics, 7 (4), pp. 174–181.

Chen, D., Yuan, L., Liao, J., Yu, N., and Hua, G. (2018). Stereoscopic neural style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6654–6663.

Chen, T. Q. and Schmidt, M. (2016). Fast patch-based style transfer of arbitrary style. arXiv preprint arXiv:1612.04337.

Cheng, M.-M., Liu, X.-C., Wang, J., Lu, S.-P., Lai, Y.- K., and Rosin, P. L. (2019). Structure-preserving neural style transfer. IEEE Transactions on Image Processing, 29, pp. 909–920.

Dong, C., Loy, C. C., He, K., and Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38 (2), pp. 295–307.

Gatys, L. A., Bethge, M., Hertzmann, A., and Shechtman, E. (2016a). Preserving color in neural artistic style transfer. arXiv preprint arXiv:1606.05897.

Gatys, L. A., Ecker, A. S., and Bethge, M. (2016b). Image style transfer using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2414–2423.

Huang, X. and Belongie, S. (2017). Arbitrary style transfer in real-time with adaptive instance normalization. pp. 1501–1510.

Li, Y., Fang, C., Yang, J., Wang, Z., Lu, X., and Yang, M.-H. (2017). Universal style transfer via feature transforms. arXiv preprint arXiv:1705.08086.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. pp. 3431–3440.

Ma, W., Chen, Z., and Ji, C. (2020). Block shuffle: A method for high-resolution fast style transfer with limited memory. IEEE Access, 8, pp. 158056–158066.

Noh, H., Hong, S., and Han, B. (2015). Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE international conference on computer vision, pp. 1520–1528.

Selim, A., Elgharib, M., and Doyle, L. (2016). Painting style transfer for head portraits using convolutional neural networks. ACM Transactions on Graphics (ToG), 35 (4), pp. 1–18.

Yang, C.-Y., Ma, C., and Yang, M.-H. (2014). Single image super-resolution: A benchmark. In European conference on computer vision, Springer, pp. 372–386.

Zhao, H.-H., Rosin, P. L., Lai, Y.-K., and Wang, Y.-N. (2020). Automatic semantic style transfer using deep convolutional neural networks and soft masks. The Visual Computer, 36 (7), pp. 1307–1324.

Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., and Torr, P. H. (2015). Conditional random fields as recurrent neural networks. In Proceedings of the IEEE international conference on computer vision, pp. 1529–1537.