Extensions to Neural Style Transfer

Team 4D Tensor
Neha Chawla, Swarnita Venkatraman, Sneha Bandi, Praveen Iyer

1 Related Work

In recent times, though Neural Style Transfer (NST) has achieved good performance, studies highlighting improving its stability, quality, flexibility, and evaluation are still in its early stages. There is not enough literature on work done in incorporating user feedback with NST to personalize the generated image based on user's preferences and doing multiple style transfer using Cascade Style Transfer.

Traditional Convolutional Neural Networks (CNN) hierarchically stack to form content representations. It is possible to form style representations by capturing the correlations of different filter responses within each layer. In (Gatys et al., 2015a) produced the first work investigating the extraction style representations from images. Additionally, much work has been done on exploring the perceptual possibilities of neural style, where novel extensions to the original neural style algorithm have greatly improved the perceptual quality of generated images, as well as introduced a number of new features. These features include multiple style transfer (Dumoulin et al., 2017), color-preserving style transfer (Gatys et al., 2016), as well as content-aware style transfer (Gatys et al., 2017).

To allow for the styles of multiple images to be transferred to a single content image, simple modifications to the style loss function have been proposed in this paper (Makow and Hernandez, 2017). The paper explains two perceptual loss functions of interest: feature reconstruction loss and the style reconstruction loss. Both make use of a pre-trained loss network trained for image classification on the ImageNet dataset. For Feature Reconstruction Loss, the pixels of the output image have similar feature representations as computed by the loss network, and in Style Reconstruction Loss, the style reconstruction loss penalizes differences in style, such as differences in colors and textures (Gatys et al., 2015b), common patterns, etc. This loss formulation selects the style layers and weights independently for each style image enabling the model to weigh overall effects of each style image on the generated image.

The paper (Wang et al., 2020) focuses on leveraging Cascade Style Transfer by combining existing approaches in Serial Style Transfer (SST) architecture to improve quality and Parallel Style Transfer (PST) architecture to improve quality and flexibility. It is one of the first papers to focus on domain independent style transfer across three domains namely artistic, semantic, and photo realistic using one content and one style image. For SST architecture it proposes a linear combination of n methods chosen thoughtfully based on each of their strengths with regards to factors like content fidelity, global color and local patterns with performance levels of High, Medium, and Low. The paper suggests reasonably concluded serial schemes for each of the domains based on their end goal. In PST architecture, a single backbone is shared by all methods and the total loss is optimized in parallel. The proposed ParallelNet architecture uses weighted losses and by increasing the weights of respective losses both global color and local textures of the style image of two different methods can be achieved. A user study was conducted by randomizing various combinations of style and content images where users were shown results of 10 compared methods and SST/PST outperformed the state of the art models for all three domains. Since PST was shown to achieve quality and flexibility to adapt to multiple domains it can be explored in our project by incorporating user feedback for weights adjustment. Besides, the paper also mentions more effective and efficient scheme designs and the combination of SST and PST has not been explored yet.

In this paper (Wang et al., 2021b), 2021 introduce 3 new quantitative factors (the content fidelity, global effects, and local patterns) for evaluating NST are introduced. Using these factors ideas have been touched upon which involve the usage of Cascade Style Transfer (CST) and user feedback. It brings up the point where CST as an architecture is biased more to style transfer than content preservation. This can be an important feature in the case of multi style transfer which we aim to do. Additionally, the paper also discussed how the user feedback can be incorporated using the 3 quantitative evaluation factors by modifying the loss function to include the user preferences. This paper also provides a thorough comparison of traditional multi-objective (MO) networks and CST architectures with average stylization ranking metrics clearly highlighting the superiority of CST over MO networks in this aspect.

"Interactive Artistic Multi-style Transfer" (Wang et al., 2021a) is a research paper that proposes a novel approach for interactive multi-style transfer in artistic images. A new system is introduced that allows users to manipulate the style and content of an input image by selecting different style images and adjusting various parameters. Their method combines feature normalization, multi-layer transformation, and iterative optimization to achieve high-quality style transfer while maintaining the content of the original image. The proposed approach enables users to create customized artistic images with multiple styles, which can be used for a variety of applications such as graphic design, digital art, and visual communication.

2 Challenges

Existing challenges with multiple style transfer include increased training time due to a separate network tied to one style. Adding multiple styles can lessen the impact of each individual style, especially when the styles are somewhat different (Makow and Hernandez, 2017). Higher style loss exists when blending two images compared to having just one style image (Makow and Hernandez, 2017). Cascade Style Transfer is shown to improve the stylistic quality and flexibility. (Wang et al., 2021b)(Wang et al., 2020) Style is known to be preserved more in Cascade Style Transfer while in traditional Multi Objective Networks content is known to be preserved more. (Wang et al., 2021b) Exploring possibilities of performing multiple style transfer using cascade style transfer to exploit its benefit of preserving style gives us room to incorporate multiple styles. Besides, existing literature on a combined methodology has not been found yet. Since PST was shown to achieve quality and flexibility to adapt to multiple domains, it can be explored with Multiple Style Transfer by incorporating user feedback for weights adjustment. More effective and efficient scheme designs and the combination of SST and PST can also be explored.(Wang et al., 2020)

Two style images might have contradicting styles that can lead to unsatisfactory output. The ordering of style images in the cascade style transfer will also matter leading to vastly different outputs for different ordering of input. There is no clear expected output or ground truth for NST. This means that for evaluation we rely on qualitative methods which can be challenging. Designing more fine grained metrics might also be needed to improve the quality of results.

References

Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. 2017. A learned representation for artistic style. In *International Conference on Learning Representations*.

Leon A. Gatys, Matthias Bethge, Aaron Hertzmann, and Eli Shechtman. 2016. Preserving color in neural artistic style transfer. *CoRR*, abs/1606.05897.

Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2015a. A neural algorithm of artistic style. *Journal of Vision*, abs/1508.06576.

Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2015b. Texture synthesis using convolutional neural networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'15, page 262–270, Cambridge, MA, USA. MIT Press.

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Aaron Hertzmann, and Eli Shechtman. 2017. Controlling perceptual factors in neural style transfer. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3730–3738.

Noah Makow and Pablo Hernandez. 2017. Exploring style transfer: Extensions to neural style transfer. *Stanford University Report*.

Xiaohui Wang, Yiran Lyu, Junfeng Huang, Ziying Wang, and Jingyan Qin. 2021a. Interactive artistic multi-style transfer. *International Journal of Computational Intelligence Systems*, 14.

Zhizhong Wang, Lei Zhao, Haibo Chen, Zhiwen Zuo, Ailin Li, Wei Xing, and Dongming Lu. 2021b. Evaluate and improve the quality of neural style transfer. *Computer Vision and Image Understanding*, 207:103203.

Zhizhong Wang, Lei Zhao, Qihang Mo, Sihuan Lin, Zhiwen Zuo, Wei Xing, and Dongming Lu. 2020. Cascade style transfer. *Under review as a conference paper at ICLR 2020*.

A Division of Labor

In the project proposal, Sneha worked on writing the introduction and outlining the objective and goals of the project. Swarnita worked on identifying existing work and key ideas for novelty that will potentially make the project successful. Neha worked on identifying what difference the project would make if implemented successfully, highlighting both positive and negative connotations. Praveen worked on identifying potential challenges that can be encountered while working and possible solutions to overcome them. The computation costs estimation, division of work and expected timeline was collectively discussed amongst all team members.

For the project pitch, the slides and content to be presented was handled by all four team members. The person chosen to deliver the pitch was Sneha.

In the survey report, each team member contributed to the Related Work section by reading and summarizing one research paper. Praveen worked on user feedback mechanism related works, Swarnita worked on cascade style transfer architecture research, Sneha worked on multiple style techniques and Neha worked on exploration of interactive styling. Each team member also contributed two challenges to the 'Challenges' section that were identified from the respective papers and earlier from the proposal. All team members worked on contributing a set of references to the survey report and the final addition to Overleaf Latex format.