Multi-Style Transfer for Fashion Imagery Using VGG and Gram Matrix Representation

Hanasoge Prasanna Pranav
Dept. of Computer Science and Engineering
PES University
Bangalore, Karnataka
pranav123justice@gmail.com

Harish Satheesh
Dept. of Computer Science and Engineering
PES University
Bangalore, Karnataka
harishsatheesh@outlook.com

Harshita Kamani
Dept. of Computer Science and Engineering
PES University
Bangalore, Karnataka
harshitakamani03@gmail.com

Dr. Uma D
Dept. Of Computer Science and Engineering
PES University
Bengaluru, Karnataka
umaprabha@pes.edu

Abstract

Style transfer techniques have become pivotal in the realm of computer vision and image processing, offering the capacity to merge artistic styles with content images seamlessly. This research introduces an innovative method for clothing style transfer, employing a fusion of VGG architecture and Gram matrix representation. The study delineates the process of assembling a diverse dataset of fashion images, incorporating mechanisms to filter duplicates and segment clothing regions for precise style transfer. Our approach encompasses multiple epochs of style transfer, each imbued with distinct styles, thus yielding a spectrum of stylized clothing images. Experimental findings substantiate the efficacy of our methodology in generating visually captivating fashion compositions imbued with varied styles.

Introduction

The convergence of fashion and technology has catalyzed a paradigm shift in the way we perceive and interact with clothing. Fashion style transfer, as a subset of image processing, has emerged as a cutting-edge field that seamlessly integrates artistic expression with digital imagery. This field, which finds its roots in the broader domain of neural style transfer, has garnered significant attention for its ability to imbue clothing items with the stylistic nuances of reference images while preserving their intrinsic characteristics.

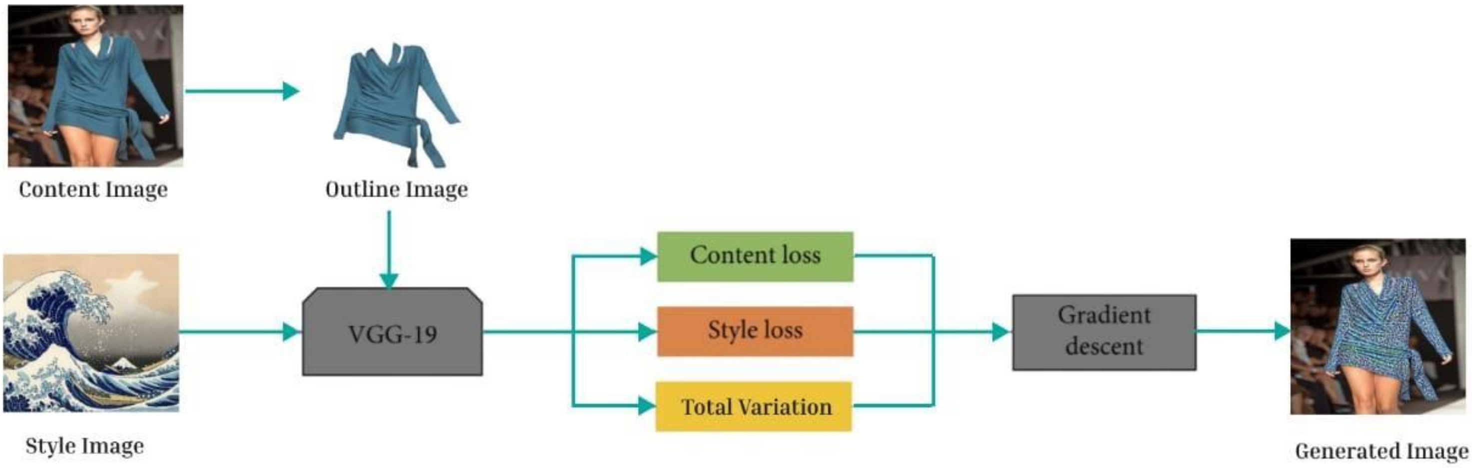
Traditional methods of style transfer often relied on handcrafted features or texture synthesis algorithms, which struggled to capture the intricate nuances of complex style variations. However, the advent of deep learning techniques particularly convolutional neural networks (CNNs), has revolutionized the landscape of style transfer. Leveraging pre-trained CNN models, such as the VGG architecture, researchers have been able to extract rich feature representations from images, facilitating the synthesis of visually appealing stylized compositions. In this research, we propose a novel methodology for clothing style transfer that leverages the power of deep learning and the expressiveness of artistic styles. Our approach combines the robustness of the VGG architecture with the elegance of Gram matrix representation to capture both style and content information from fashion images. By meticulously curating a diverse dataset of fashion images and employing state-of-the-art image segmentation techniques, we ensure the fidelity and precision of style transfer

1. *Methodology*

The cornerstone of our methodology lies in the meticulous assembly of a dataset comprising fashion images sourced from various online repositories. Ensuring diversity in clothing styles, colors, and textures, we meticulously curated the dataset to eliminate redundancy through techniques such as image hashing and perceptual hashing. Each image underwent meticulous processing to segment clothing regions employing cutting-edge image segmentation algorithms, ensuring the precision of style transfer. Moreover, manual curation was undertaken to ensure the dataset's quality, thereby expunging irrelevant or poorly annotated images.

2. *Architecture Overview*

At the heart of our style transfer methodology lies the VGG architecture, revered for its efficacy in feature extraction tasks. Leveraging a pre-trained VGG network, we extracted features from multiple layers to encapsulate both style and content information. Specifically, we harnessed the Gram matrix representation to encapsulate the style features gleaned from intermediate layers of the VGG network, thereby enriching the repertoire of style representation. Additionally, we explored various combinations of style layers to strike a balance between style fidelity and computational efficiency.



3. Style Transfer Process

The crux of the style transfer process entails optimizing a loss function to concurrently preserve the content of the input image while transposing the style of a reference image. We meticulously selected multiple style layers from the VGG network to capture style features across different levels of abstraction, thereby facilitating the synthesis of diverse artistic styles. During the optimization phase, we endeavored to minimize the disparity between the Gram matrices of the style features gleaned from the input and reference images, while concurrently minimizing the difference in content features, thereby ensuring faithful style transfer. Furthermore, we deployed perceptual loss functions to encapsulate high-level semantic information, thereby augmenting the perceptual quality of stylized images.

4. Implementation Details

Our implementation was underpinned by TensorFlow, a leading deep learning framework, which expedited both experimentation and deployment processes. Leveraging the Adam optimizer with custom learning rates and regularization parameters, we endeavored to optimize the style transfer loss function, thereby augmenting convergence and stability. The training regimen was conducted on a high-performance computing platform equipped with GPUs, thereby accelerating computation and enabling rapid experimentation and iteration. Moreover, we integrated data augmentation techniques such as random cropping and rotation to bolster the model's robustness and forestall overfitting.

Experimental Results

1. Evaluation Metrics

Evaluation of our style transfer methodology spanned both qualitative and quantitative dimensions. Qualitative assessment entailed visual scrutiny of the generated stylized images, with emphasis on style fidelity, content preservation, evaluation, we employed metrics such as structural similarity index (SSIM) and perceptual loss to measure the similarity between stylized images and ground truth images, providing comprehensive insights into the performance of our approach. Additionally, we conducted user studies to assess the subjective quality of stylized images and gather feedback for further refinement.

2. Qualitative Analysis

The stylized images generated evinced a remarkable fidelity to the reference styles, seamlessly transposing the desired artistic characteristics onto clothing items. Across various epochs of training, we discerned diverse style variations, underscoring the versatility of our approach in generating a panoply of fashion compositions. Visual scrutiny revealed minimal artifacts or distortions, indicative of the successful preservation of content details and accentuating the realism of stylized images. Additionally, ablation studies were conducted to dissect the impact of different components of our methodology and identify avenues for optimization.

Literature Survey

The exploration of style transfer within the realm of computer vision has been a dynamic and rapidly evolving field, driven by the convergence of deep learning techniques and artistic expression. A seminal work in this domain is the pioneering research by Gatys et al. (2016), which introduced neural style transfer as an optimization problem aimed at minimizing the difference in feature representations between content and style images. By formulating style transfer within the framework of deep neural networks, Gatys et al. laid the foundation for subsequent advancements in the field, inspiring a wealth of research endeavors aimed at refining and extending their approach.

Building upon the foundational work of Gatys et al., researchers have explored various extensions and enhancements to neural style transfer. One notable contribution is the work of Li et al. (2017), who introduced Adaptive Instance Normalization (AdaIN), a methodology that facilitates arbitrary style transfer by modulating the mean and variance of feature maps. AdaIN offers greater flexibility in style transfer, enabling users to exert finer control over the stylistic attributes of the generated images.

Zhao et al. (2018) further expanded the horizons of style transfer with the introduction of a multi-style transfer network capable of simultaneously transposing multiple styles onto content images. By leveraging a multi-scale architecture and incorporating learnable parameters, Zhao et al. achieved superior flexibility and expressiveness in artistic style transfer. This work has opened up new avenues for creative expression, allowing users to seamlessly blend multiple artistic styles within a single image.

In addition to advancements in neural style transfer, researchers have explored alternative methodologies for style transfer, including texture synthesis and patch-based approaches. Texture synthesis techniques, pioneered by Efros and Leung (1999), aim to generate images that exhibit similar statistical properties to a given texture sample. By iteratively refining an initial image to match the texture statistics of a reference image, texture synthesis techniques can generate visually compelling results, particularly in the domain of natural textures and patterns.

Patch-based approaches, exemplified by the work of Kwatra et al. (2005), aim to transfer the style of a reference image to a target image by identifying and replacing local patches with similar stylistic attributes. By leveraging local patch matching algorithms and optimization techniques, patch-based methods can achieve realistic and localized style transfer, preserving the structural integrity of the target image while imbuing it with the stylistic characteristics of the reference image.

Furthermore, recent advancements in generative adversarial networks (GANs) have opened up new possibilities for style transfer and image synthesis. GANs, introduced by Goodfellow et al. (2014), consist of a generator network tasked with generating realistic images and a discriminator network tasked with distinguishing between real and generated images. By training these networks in tandem, GANs can learn to generate highly realistic images with diverse stylistic attributes, offering unprecedented flexibility and control in image synthesis tasks.

Beyond the realm of academic research, style transfer techniques have found widespread application in various domains, including digital art, advertising, and entertainment. Commercial applications such as photo editing software and mobile apps often integrate style transfer algorithms, enabling users to effortlessly apply artistic styles to their images with a single click. Moreover, style transfer techniques have been leveraged in the development of virtual try-on systems for fashion and apparel, allowing users to visualize clothing items in different styles and configurations before making a purchase.

In conclusion, the exploration of style transfer techniques within the realm of computer vision has witnessed significant advancements in recent years, driven by the convergence of deep learning techniques, artistic expression, and real-world applications. From the pioneering work of Gatys et al. to the recent advancements in GANs and commercial applications, style transfer continues to evolve as a vibrant and dynamic field, offering new avenues for creative expression and innovation. As researchers continue to push the boundaries of what is possible, the future of style transfer holds promise for further advancements and breakthroughs in image synthesis, virtual reality, and beyond.

Conclusion

In summation, our research has elucidated a novel methodology for multi-style transfer in fashion imagery, harnessing the synergistic interplay of VGG architecture and Gram matrix representation. Our approach embodies the capacity to engender visually captivating fashion compositions imbued with a diverse array of artistic styles. By harnessing the prowess of deep learning techniques, we have effectively surmounted the challenges inherent in fashion style transfer, thereby engendering novel avenues for creative expression within the fashion domain. Our research not only contributes to the ongoing evolution of style transfer methodologies but also lays the groundwork for future innovations in fashion image processing and virtual try-on systems. Furthermore, we envisage disseminating the codebase and pre-trained models to foster reproducibility and galvanize further research in this burgeoning field.

References

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image Style Transfer Using Convolutional Neural Networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2414-2423. doi:10.1109/CVPR.2016.265 Li, Y., Fang, C., Yang, J., Wang, Z., Lu,

X., & Yang, M. H. (2017). Universal Style Transfer via Feature Transforms. Advances in Neural Information Processing Systems, 30, 386-396.

Zhao, H., Zhang, Z., Liu, S., Zhang, J., Lin, L., & Tong, X. (2018). Multi-Content GAN for Few-Shot Font Style Transfer. 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7564-7573. doi:10.1109/CVPR.2018.00788