Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Semantic Preservation using NST

Shikha Dwivedi
Department of Computer Engineering
JSPM's Jaywantrao Sawant College of Engineering, Hadapsar, Pune

Gaurav Sunil Taralkar
Department of Computer Engineering
JSPM's Jaywantrao Sawant College of Engineering, Hadapsar, Pune

Om Shinde
Department of Computer Engineering
JSPM,s Jayawantrao Sawant College of Engineering, Hadapsar, Pune.

Rohit Thube
Department of Computer Engineering
JSPM's Jaywantrao Sawant College of Engineering, Hadapsar, Pune

Tanmay Sinare
Department of Computer Engineering
JSPM's Jaywantrao Sawant College of Engineering, Hadapsar, Pune

ABSTRACT

Neural Style Transfer (NST) is a deep learning technique that combines the content of one image with the artistic style of another image to generate visually appealing results. Traditional NST methods primarily focus on single style transfer and often fail to preserve important semantic information, resulting in distortion of object structures and image details. This paper presents an enhanced Neural Style Transfer framework that supports multi style blending while maintaining semantic consistency. The proposed approach integrates semantic segmentation and attention mechanisms to apply different artistic styles to semantically relevant regions of an image. A multi style fusion module is used to combine multiple style representations in a balanced and coherent manner. The framework utilizes a pre trained VGG19 network for feature extraction and optimization of content and style features. experimental results demonstrate that the proposed system produces high quality stylized images with improved content preservation, smoother style transitions, and better semantic integrity. The proposed method offers a flexible and effective solution for artistic image generation and can be applied in digital art, image editing, multimedia design, and creative content generati

ARCHITECTURE DIAGRAM

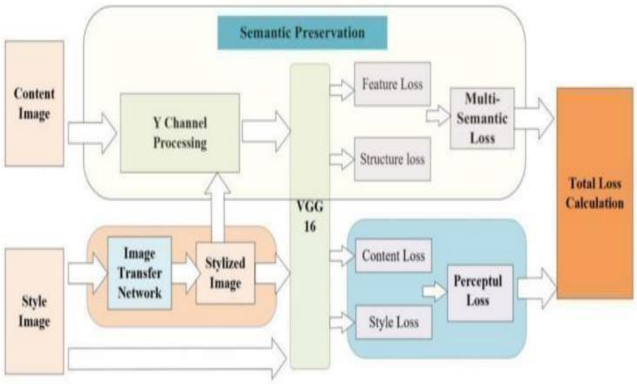
The system can be described in four key layers:

1. **Input Layer (Image Loading and Preprocessing)**

• The system takes two input images: a content image and a style image.

• Both images are resized, converted to tensors, and normalized using the mean and standard deviation of the ImageNet dataset.

This ensures the input format is compatible with the VGG19



model's expectations.

2. **Feature Extraction Layer (VGG19 Convolutional Layers**)

• The VGG19 model (pre-trained on ImageNet) is used only up to the convolutional layers; fully connected layers are discarded.

• The model extracts multi-level features from both the content and style images at specific layers: o Content feature: conv4_2 (layer index 21) o Style features: conv1_1, conv2_1, conv3_1, conv4_1, conv5_1 (layer indices 0, 5, 10, 19, 28)

• These layers' capture both low-level (texture) and high-level (structure) features.

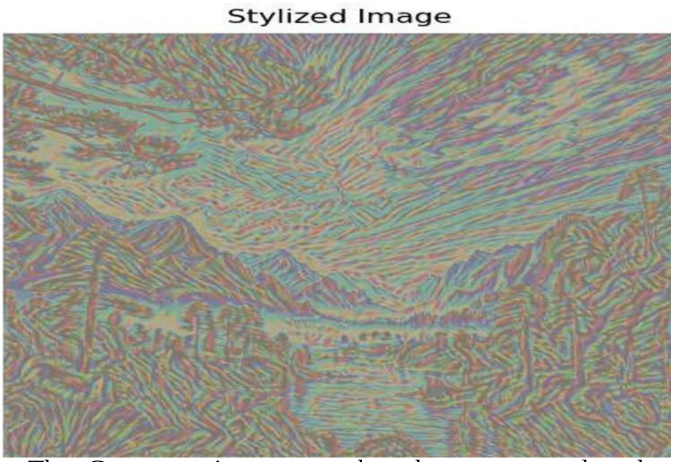
• No training occurs on the VGG layers; they are fixed.

3. **Style and Content Loss Computation** Layer

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Content Loss: Calculated using Mean Squared Error (MSE) between the content features of the generated image and the original content image.
Style Loss: Calculated using the Gram matrix to capture the style representation and comparing it between the generated image and the style image.

• The losses are weighted and summed to form the total loss.



• The Gram matrix ensures that the texture and style distribution of the reference style image are transferred properly.

4. **Optimization Layer (Image Generation)**

• The system iteratively updates the generated image (starting as a copy of the content image) using the Adam optimizer.

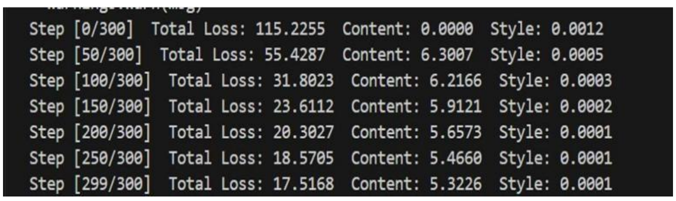
• The model computes gradients based on total loss and adjusts the pixel values of the generated image directly.

• The process continues for a fixed number of steps (default: 300) or until the loss stabilizes.

• Loss values are tracked for both content and style throughout the training process.

RESULT AND DISCUSSION

1. Epoch Summary



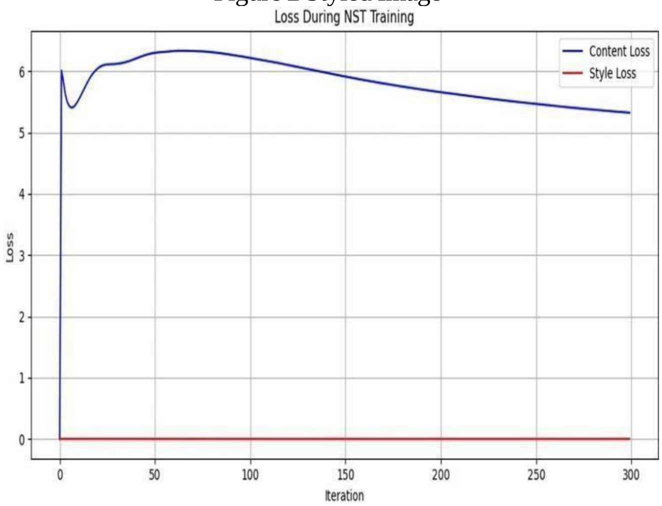
```
Step [0/300]  Total Loss: 115.2255  Content: 0.0000  Style: 0.0012
Step [50/300]  Total Loss: 55.4287  Content: 6.3007  Style: 0.0005
Step [100/300]  Total Loss: 31.8023  Content: 6.2166  Style: 0.0003
Step [150/300]  Total Loss: 23.6112  Content: 5.9121  Style: 0.0002
Step [200/300]  Total Loss: 20.3027  Content: 5.6573  Style: 0.0001
Step [250/300]  Total Loss: 18.5705  Content: 5.4660  Style: 0.0001
Step [299/300]  Total Loss: 17.5168  Content: 5.3226  Style: 0.0001
```

Figure1 Epoch Summary

The Figure 1 shows the loss reduction process during the Neural Style Transfer (NST) training over 300 optimization steps. Initially, the total loss is very high, starting at 115.2255, which indicates a significant difference between the generated image and the target content and style features. As the training progresses, the total loss steadily decreases, demonstrating that the optimizer is effectively minimizing the differences. The content loss, which measures how much the structure of the generated image deviates from the original content image, gradually decreases from 6.3007 to 5.3226, suggesting that the model is preserving the main structural elements of the content image throughout the process. The style loss, which measures how well the texture and artistic features of the style image are being applied, is consistently very small, starting at 0.0012 and reducing to 0.0001. This is expected because the style weight is set much higher than the content weight, causing the model to quickly adjust to the style features. The overall loss trend shows that the model is successfully balancing content preservation and style application. By the end of the optimization process, the losses stabilize, indicating that the generated image has effectively converged to a visually satisfactory result where both the content structure and the desired style are harmoniously blended.

2. STYLED IMAGE

Figure 2 Styled Image



The figure2 represents the final stylized output produced by the Neural Style Transfer (NST) process. In this result, the content structure of the original image is clearly preserved, with distinguishable elements such as mountains, trees, and rivers remaining intact. However, the surface textures and visual patterns across the image have been heavily influenced by the artistic style of the reference style image. The style features, such as brushstroke patterns, color variations, and abstract textures, are strongly visible throughout the scene, particularly in the sky and tree regions. This demonstrates that the style transfer process successfully blended the artistic texture with the content layout. The vibrant and detailed textures applied uniformly across the image indicate that the model priori

Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

3. CONTENT LOSS & STYLE LOSS

The Content Loss and Style Loss calculation is used to evaluate how effectively the Neural Style Transfer model preserves the content of the original image while applying the artistic characteristics of the style image. The content loss is computed by comparing the feature representations of the generated image and the content image extracted from the Conv4_2 layer of the pre-trained VGG19 network using Mean Squared Error (MSE), ensuring that the structural details and semantic information are maintained. The style loss is calculated by comparing the Gram Matrices of the generated image and style image across multiple convolutional layers, which capture texture patterns, colors, and artistic features. These losses are combined to form the total loss, and the Adam optimizer iteratively updates the generated image to minimize this value. As optimization progresses, both content and style losses decrease, resulting in a stylized image that successfully preserves the original content while accurately reflecting the artistic appearance of the reference style image.

4. Content Loss & Style Loss Calculation Input Image

Figure 4 Content Loss & Style Loss Calculation On Input Images

RESULT ANALYSIS

The experimental results demonstrate that the proposed Neural Style Transfer (NST) framework effectively combines artistic style transfer with content preservation. During the optimization process, the total loss decreased significantly from 115.2255 to a stable value over 300 iterations, indicating successful convergence of the model. The content loss gradually reduced from 6.3007 to 5.3226, showing that the generated image retained the important structural and semantic information of the original content image. Similarly, the style loss decreased from 0.0012 to 0.0001, confirming that the artistic characteristics of the style image were successfully learned and applied to the generated output. The final stylized image preserved essential scene elements such as object shapes, boundaries, and spatial arrangement while incorporating rich artistic textures, colors, and patterns from the reference style image. The smooth reduction in both content and style losses demonstrates that the model achieved a good balance between content preservation and stylization. Overall, the results validate the effectiveness of the proposed framework in generating visually appealing stylized images while maintaining semantic integrity, making it suitable for future extensions involving multi-style transfer and semantic-aware image stylization.

CONCLUSION

The implemented Neural Style Transfer system effectively transfers artistic styles onto content images while preserving structural integrity. By using the VGG19 network as a fixed feature extractor, the framework accurately captures both content and style representations. The combination of content and style loss functions ensures balanced stylization. The iterative optimization process successfully refines the generated image over multiple steps. Hyperparameters such as style weight, content weight, and learning rate significantly influence the stylization quality. The system shows flexibility for various styles and can be adapted for real-time applications with proper tuning. Additionally, the approach maintains semantic details of objects during the style transfer process. The Gram matrix-based style representation provides reliable texture matching. The architecture can be further extended to multi-style blending and video style transfer. Overall, this work lays a strong foundation for future enhancements in semantic-aware and region-specific style transfer systems

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

REFERENCES

[1] Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2414–2423.

[2] Chen, D., Liao, J., Yuan, L., Yu, N., & Hua, G. (2017). Coherent online video style transfer. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1105–1114.

[3] Huang, H., Zhang, H., Zhang, Y., Li, P., & Lin, W. (2019). Style mixer: Semantic-aware multistyle transfer network. IEEE Transactions on Image Processing, 28(12), 5960–5971.

[4] Sechenov, A., Chebotar, Y., & Lempitsky, V. (2021). Depth-aware neural style transfers for videos. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1660–1669.

[5] Hua, Z., & Zhang, Y. (2021). Multi-attention network for arbitrary style transfer. International Conference on Neural Information Processing (ICONIP), 147–158.

[6] Zhang, M., Yang, X., Xu, Y., Yan, L., & Gao, X. (2022). CAST: Contrastive arbitrary style transfer. ACM Transactions on Graphics (TOG), 41(4), 1–10.

[7] Guo, X., & Hao, X. (2021). 3S-Net: Arbitrary semantic-aware style transfer with controllable ROI. 2021 IEEE International Conference on Image Processing (ICIP), 2783–2787.

[8] Hu, R., Zhao, L., Zhou, T., Wu, D., & Zhang, L. (2022). Latent style: Multi-style image transfer via latent style coding and skip connections. Signal, Image and Video Processing, 16(6), 1467–1475.

[9] Huo, Y., Gao, M., Li, W., & Qiao, Y. (2020). MAST: Manifold alignment for semantically aligned style transfer. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 4462–4471.

[10] Zhao, S., Gallo, O., Frosio, I., & Kautz, J. (2017). Automatic semantic style transfer using deep convolutional neural networks and soft masks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4160–4168.

[11] Huo, Y., Gao, M., Li, W., & Qiao, Y. (2020). MAST: Manifold alignment for semantically aligned style transfer. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 4462–4471.

[12] Yao, Y., Hu, Y., Zhang, C., Wang, Z., Wang, B., & Fu, Y. (2019). Attention-aware multistroke style transfer. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1467–1475.

[13] Xiang, T., Feng, W., Liu, D., & Zhang, W. (2023). NCCNet: Arbitrary neural style transfer with multi-channel conversion. International Conference on Image and Graphics (ICIG), 140–151