Neural Style Transfer on Videos

Abhiraj Chaudhary, Hitesh Goyal, Ashwin Iyer, Yash Tripathi
VIT Chennai, Tamil Nadu, India 600127
Email: abhiraj.chaudhary2019@vitstudent.ac.in,
hitesh.goyal2019@vitstudent.ac.in,
ashwiniyer.u2019@vitstudent.ac.in,
yash.tripathi2019@vitstudent.ac.in

***Abstract*—Human beings have always had a keen interest in art. Over the years, artists have used various styles and have made marvelous pieces of art. Now, we cannot bring back the artists, but, we can use their art and see how they would have made other artistic works. A special application would be to apply such techniques on videos to see how they would look if they were an artistic work of these people.**

1. Introduction

Neural Style Transfer is a concept of Neural Networks which helps applying the style of one image to the other. The features and objects of the original image are re-styled with the style image.

If done right, this transfer can Transform any video to enhance different features of the same, you can make a day image appear as if it was night, increase the realism of animated and graphical objects and even allow for various theme settings in different vision related applications.

The concept of Neural Style Transfer is to use Neural Networks increase the information that can be inferred form , can to be applied on videos. Applying this on videos will increase its scope as it could provide endless material for artistic purposes creating a dynamic filter which can mold to any situation based on the data that is provided to it.

The concept of Neural Style Transfer, to increase its scope, can to be applied on videos. Applying this on videos will increase its scope, usage and probably importance as well.

Neural style transfer is implemented by optimizing the output image to match the content statistics of the content image and the style statistics of the style reference image. These statistics are extracted from the images using a convolutional network. Starting from the network's input layer, the first few layer activations represent low-level features like edges and textures. The intermediate layers of the model are used to get the content and style representations of the image. As you step through the network, the final few layers represent higher-level features—object parts like wheels or eyes.

2. Related Work

2.1. Neural Style Transfer by TensorFlow

This Paper uses deep learning to compose one image in the style of another image. This is known as neural style transfer and the technique is outlined in A Neural Algorithm of Artistic Style. Neural style transfer is an optimization technique used to take two images—a content image and a style reference image and blend them together so the output image looks like the content image, but "painted" in the style of the style reference image.

2.2. Image Style Transfer Using Convolutional Neural Networks

Rendering the semantic content of an image in different styles is a difficult image processing task. Arguably, a major limiting factor for previous approaches has been the lack of image representations that explicitly represent semantic information and, thus, allow to separate image content from style. Here we use image representations derived from Convolutional Neural Networks optimised for object recognition, which make high level image information explicit. We introduce A Neural Algorithm of Artistic Style that can separate and recombine the image content and style of natural images. The algorithm allows us to produce new images of high perceptual quality that combine the content of an arbitrary photograph with the appearance of numerous well known artworks. Our results provide new insights into the deep image representations learned by Convolutional Neural Networks and demonstrate their potential for high level image synthesis and manipulation.

2.3. Very Deep Convolutional Networks For Large-Scale Image Recognition

In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3×3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision.

2.4. Research on Neural Style Transfer Algorithm

Researchers have successfully applied the convolutional neural network (CNN) to style transfer. Since then, Neural Style Transfer (NST) has received widespread attention in both scientific and industrial fields. Researchers in the field of machine vision are constantly proposing ways to optimize image style migration. This paper aims to summarize the history of style transfer before and after the rise of CNN, classify the existing classical and improved algorithms, and compare the results of some of them. Finally, after this study, we put forward some suggestions on the development trend of image style transfer.

3. Method

3.1. Implementation

3.1.1. The model

To implement Neural Style Transfer, we use the VGG19 network architecture, a pretrained image classification network. These intermediate layers are necessary to define the representation of content and style from the images. For an input image, try to match the corresponding style and content target representations at these intermediate layers.

At a high level, in order for a network to perform image classification (which this network has been trained to do), it must understand the image. This requires taking the raw image as input pixels and building an internal representation that converts the raw image pixels into a complex understanding of the features present within the image.

This is also a reason why convolutional neural networks are able to generalize well: they're able to capture the invariances and defining features within classes (e.g. cats vs. dogs) that are agnostic to background noise and other nuisances. Thus, somewhere between where the raw image is fed into the model and the output classification label, the model serves as a complex feature extractor. By accessing intermediate layers of the model, you're able to describe the content and style of input images.

3.1.2. The origin

In 2015, Gatys et al. combined neural network and style transfer, which officially opened the prelude of Neural Style Transfer. Gatys et al. applied the gram matrix to different local feature maps extracted by the VGG-19 network and calculated the correlation between features to form a statistical model. The computation of style features is performed on all convolution layers. A part of local features is directly taken as the content, and the content features are for a convolution layer. Finally, combine the content features and style features of the picture together to form a new picture.

Firstly, a white noise image is generated, and iteration is carried out according to the loss function. Then, gradient descent algorithm is used for back propagation and continuous optimization to obtain the minimum loss. The loss function used by Gatys et al. is the sum of content loss and style loss, both of which have their own parameters.

3.1.3. Some abstracted mathematics

Given a picture $\vec{p}$, the feature map is obtained by calculating in the neural network. Each layer can get n feature maps, which is determined by the number of filters. The feature map is vectorized and the resulting vector is finally put into the matrix F. The element F_{ij} represents the activation response of the i filter in the l layer at position j. Specify the feature representation of a layer l, and generate a style-transferred image $\vec{x}$, so that the feature representation of the layer P_l is equal to the original feature representation of F_l.The loss function is defined as follows:

$$L_{context}(\vec{p}, \vec{x}, l) = 1/2 \sum_{i,j} (F^l_{i,j} - P^l_{i,j})^2$$

3.2. Results

Currently the important result is the images of the implementation. Below, we have a total of 7 images. The first image is a certain frame of the original video and is named *original image*. Following that, there are 6 more images. The first three (by numbering) are the styled frames of the original frame and the next three are the style images used to apply those styles.

Figure 1. Original Image

Figure 2. Styled Image 1

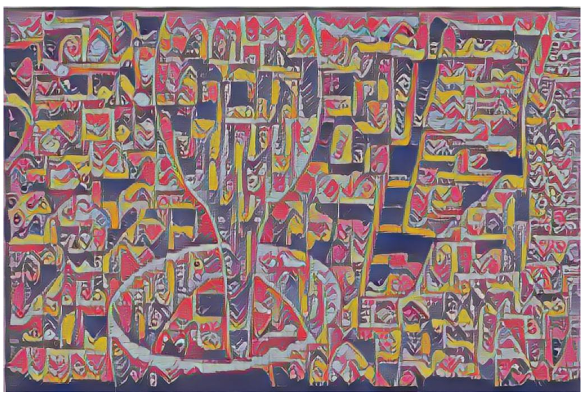

Figure 3. Styled Image 2

Figure 4. Styled Image 3

Figure 5. Style Image 1

Figure 6. Style Image 2

Figure 7. Style Image 3

4. Conclusion

The images of the video show how well the NST can be applied on a video. Some images are better applied than others and that is mainly because of the smoothness of the style image. So, a smooth NST is dependent on the style image as much as the original video. If there are too many features in the style, the video will come out as excessively noisy.

5. Future Work

The only issue with the current implementation is that it is not temporally dependent. Because of that, there is a good chance of having rough frame transitions which can make it less appealing. To fix this, the motion estimation techniques can be added as a dependency to the Style transfer network making sure that two consecutive frames are styled in a similar manner allowing from smoother videos.

6. References

[1] Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 2414-2423).

[2] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4.

[3] Yao C, Li Y, Qi Y. Research on Neural Style Transfer Algorithm. InIOP Conference Series: Materials Science and Engineering 2019 Jul 1 (Vol. 569, No. 5, p. 052107). IOP Publishing.

[4] Neural style transfer by TensorFlow: www.tensorflow.org/tutorials/generative/style_transfer