Universal Style Transfer via Feature Transforms

Rama Rohit Reddy
Suma Reddy Duggenpudi
V. Jaswanth

Aim of the Project

The aim of the project is to develop a method for universal style transfer which enjoys style-agnostic generalization ability with marginally compromised visual quality and execution efficiency.

What is Universal Style Transfer?

Content Image

What is Universal Style Transfer?

Content Image

Style Image

What is Universal Style Transfer?

- Output image preserves some notion of the content but carries characteristics of the style.

What is Universal Style Transfer?

Output Image

Why Universal Style Transfer?

- Style transfer is an important image editing task which enables the creation of new artistic works.
- The key challenge lies in how to extract effective representations of the style and then match it in the content image.

Major Problems with existing techniques

- Optimization based methods can handle arbitrary styles with pleasing visual quality but it takes many iterations to generate good results hence requires high computational costs.

Arbitrary	√
Efficient	×
Learning-free	√

Major Problems with existing techniques

- Feed-forward approaches can be executed efficiently but are limited to a fixed number of styles or compromised visual quality

Arbitrary	×
Efficient	√
Learning-free	×

Selling Point of this paper.

- It is learning free.

Arbitrary	√
Efficient	√
Learning-free	√

Selling Point of this paper.

- Existing feed-forward base techniques would need to be trained on predefined styles and then fine-tuned for new styles.
- The method in paper, is completely independent of the style during training phase.

About the dataset

Mainly there are two datasets deployed for two different purposes. They are :-

1. Microsoft COCO dataset
2. Describable Textures Dataset

Microsoft COCO dataset has 83K images.

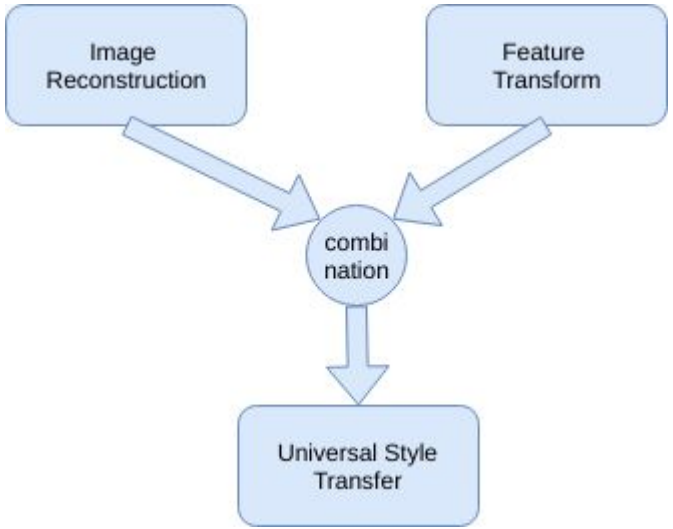
Describable Textures Dataset has around 5640 images organized according to a list of 47 terms which are inspired from human perception.

The Main Task

The main task of this problem is to extract effective representations of the style and then match it to the content image.

Method Outline.

- The style transfer problem is formulated as a combination of two processes, viz. Image Reconstruction and Feature transform using Whitening and Color Transform.



Method Outline.

- The feature transformation matches the statistics of a content image to a style image.
- The reconstruction part is responsible for inverting features back to the RGB space.

Method : Image Reconstruction

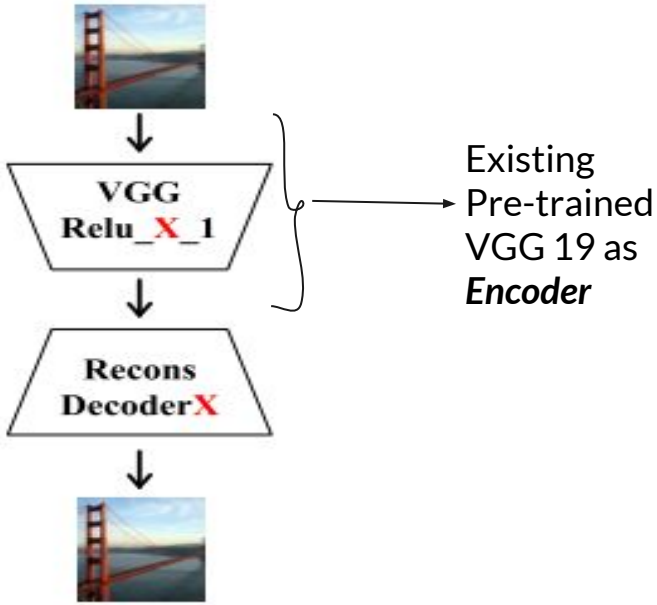
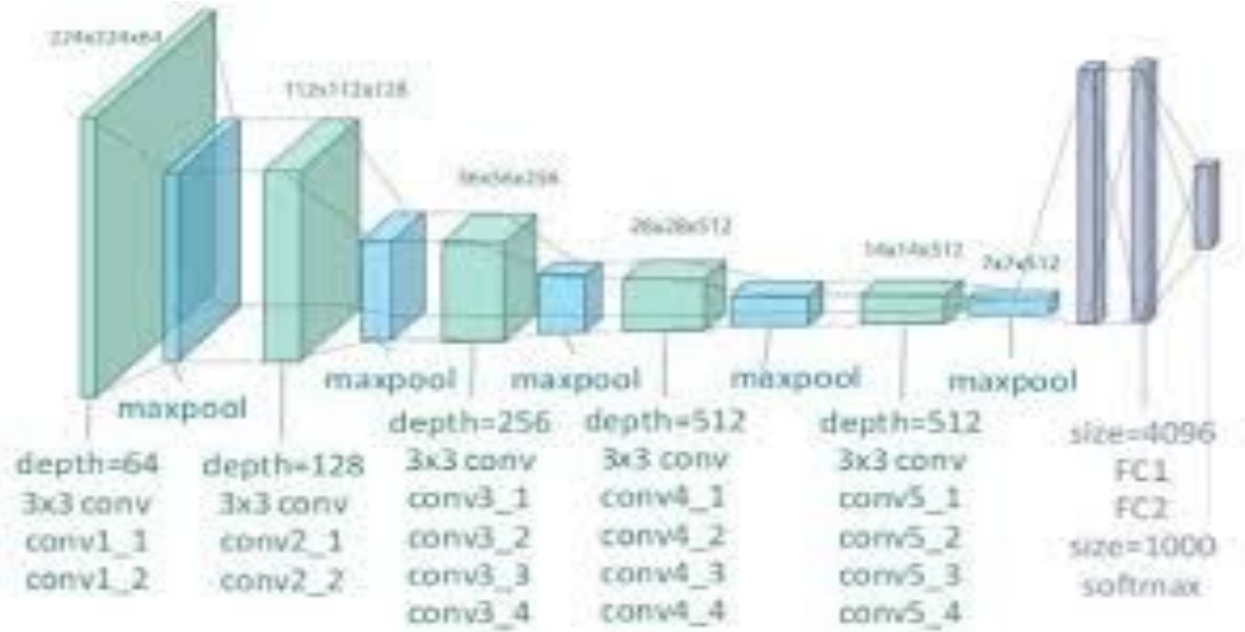


Fig : Architecture for image reconstruction. 'X' represents the layer number.

Encoder (VGG19) architecture



Method : Image Reconstruction (Decoder)

- The Recons Decoder is then trained to reconstruct the Image.
- The decoder is designed to be symmetrical to that of VGG-19 network with the nearest neighbour upsampling layer used for enlarging feature maps.
- Important to note that 5 decoders are trained for reconstruction one for each layer of VGG 19.

Fig : Architecture for image reconstruction. 'X' represents the layer number.

Method : Image Reconstruction (Decoder)

- The pixel reconstruction loss and feature loss are employed for reconstructing an input image. The following is the function :

$$L = \|I_o - I_i\|_2^2 + \lambda \|\Phi(I_o) - \Phi(I_i)\|_2^2$$

Here, I_i and I_o are the input image and reconstruction output, and **Φ** is the VGG encoder that extracts the Relu_X_1 features. In addition, **λ** is the weight to balance the two losses.

- Lastly, after training the decoder, it is fixed and will not be fine-tuned later on. This will be used as a feature inverter.

Fig : Architecture for image reconstruction. 'X' represents the layer number.

Original

Relu1_1
prediction

Original

Relu3_1
prediction

Original

Relu5_1
prediction

Method : Feature Transform using Whitening and Color Transform

Method : Whitening Transform

- By whitening transformation, we can effectively disassociate the features of the style of the image. It is important to note that reconstruction from the features which are subjected to whitening transformation would preserve the content but removes any information related to style.

Original

Whitened Image

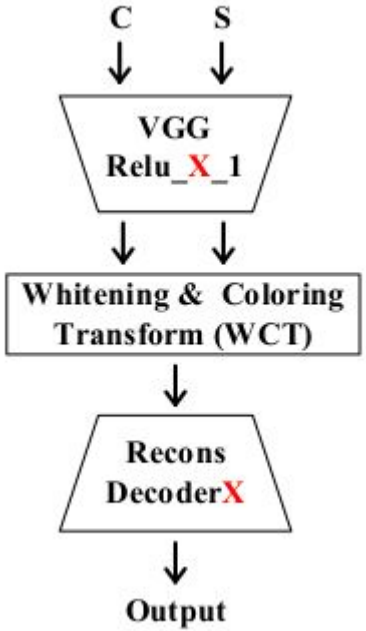
Method : Colouring Transform

- Next by using coloring transform, we will associate the style that we derive from the style image to these features of the style of the image.

Method : Whitening and Colouring Transforms

- WCT is the process of disassociating the current style of the input image and associating the style of the style image with the input image.
- WCT is basically mathematics which plays a central role in transferring the style characteristics from style image while still preserving the content.

Final Model


C
S
VGG
Relu_X_1
Whitening & Coloring
Transform (WCT)
Recons
DecoderX
Output

Transformation MATH

Results

Content Image

Style Image

Results

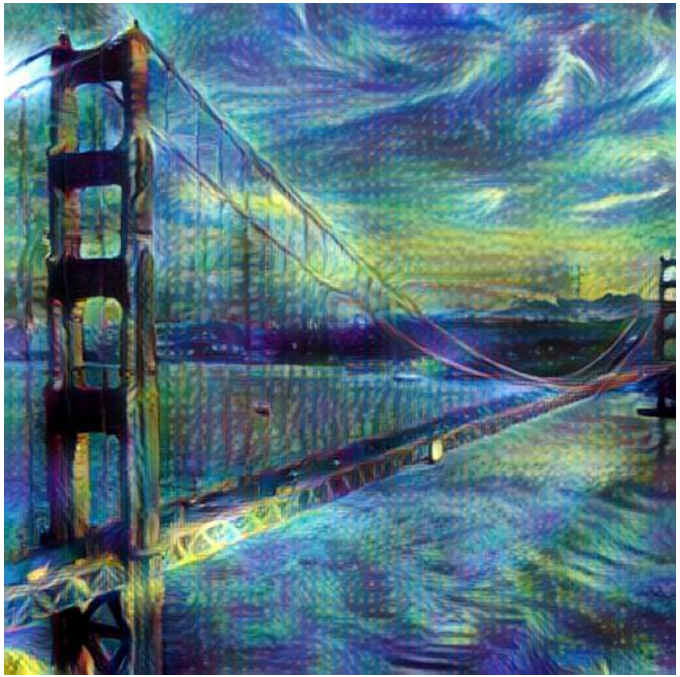
Relu1_1 Output

Relu2_1 Output

Results

Relu3_1 Output

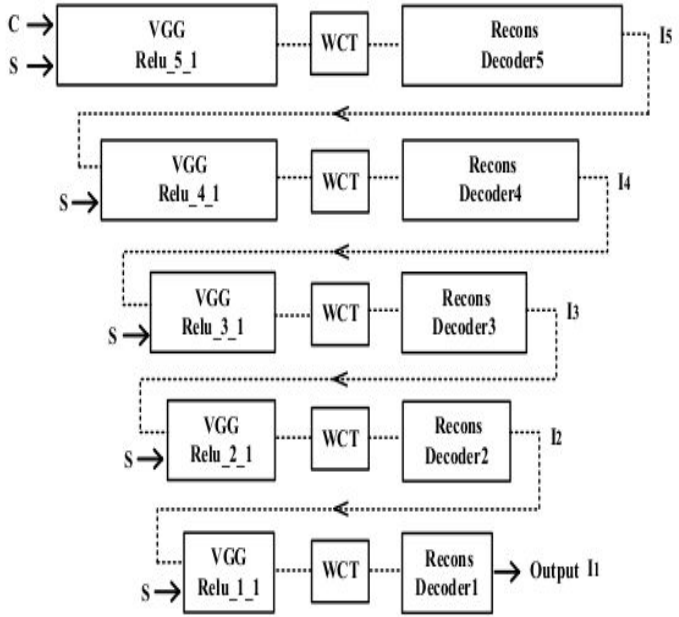
Relu4_1 Output



Multi-level stylization

- From the images in the previous slides, we can notice that:
  - High layer features capture more complicated local structures. For eg:- style structures.
  - Lower layer features carry more low-level information. For eg:- style colours.
- The authors proceeded to use the feature from all layers instead of sticking to just one in order to capture all the useful information.

Multi-level stylization Architecture



(c) Multi-level stylization

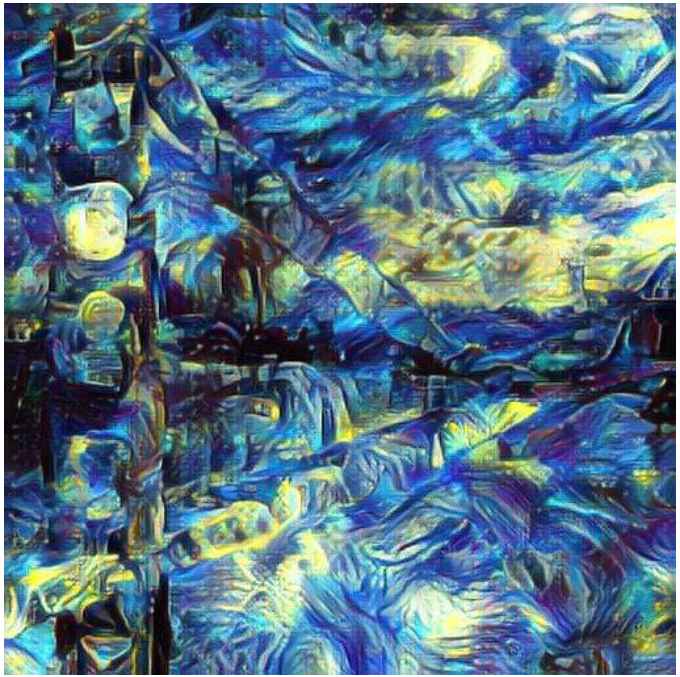
Results

Content Image

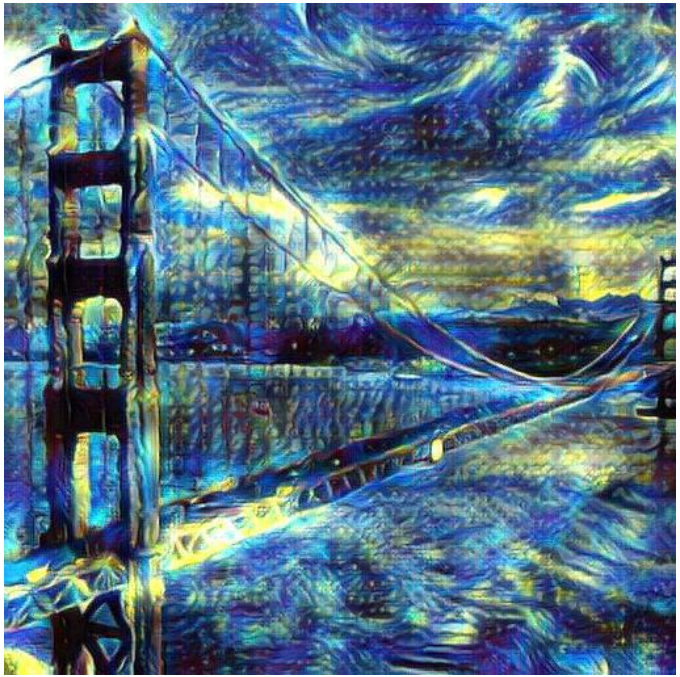
Style Image

Results

5 layers Output



4 layers Output



Results

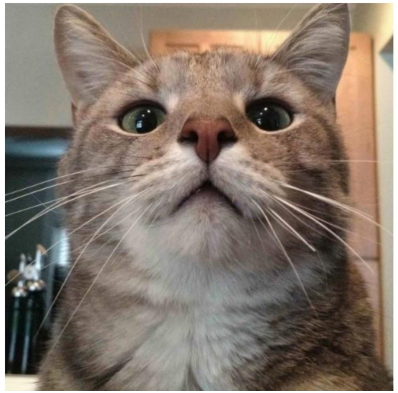
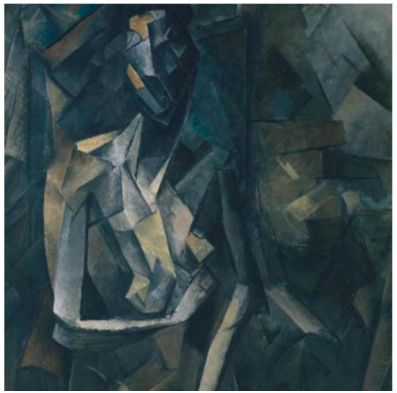
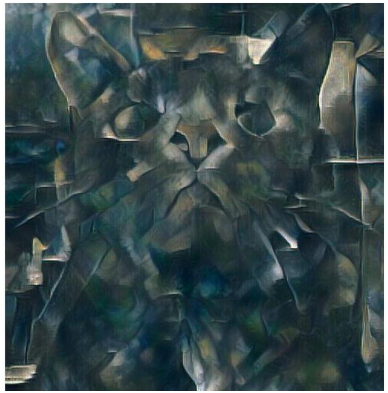
3 layers Output

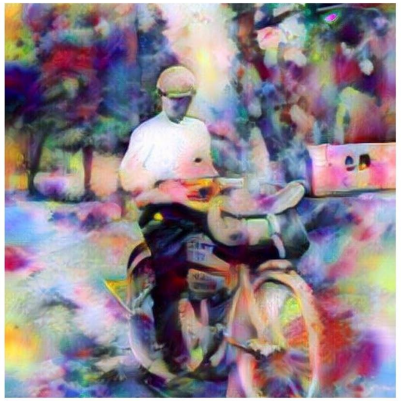
2 layers Output

Results



Results

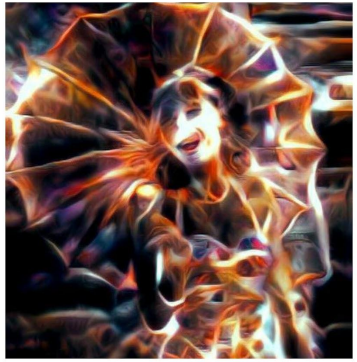




Results



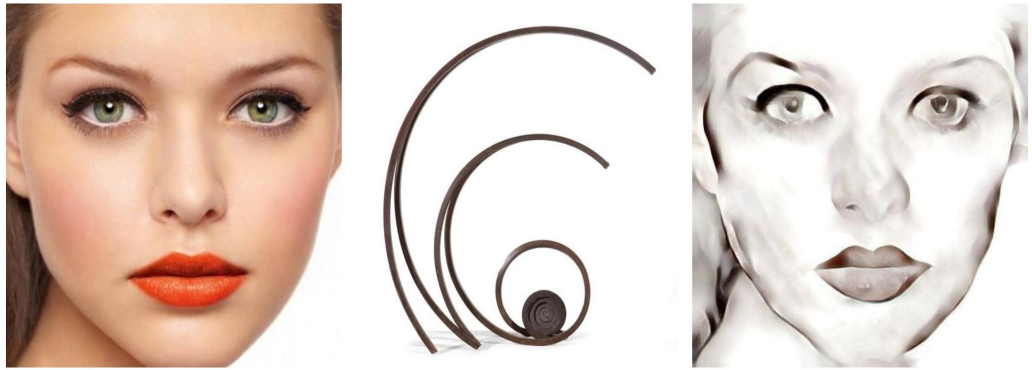




Demo

Conclusion

And hence, this algorithm is efficient as its learning involves free and efficient process because it has no loops of optimization which may take many iterations to generate good results. It is not a style specific network as it does not include style factor while training.



Some Results Generated

Thanks!