Class-based Neural Style Transfer using Semantic Segmentation

Prithwiraj Paul A59020278 Unay Shah A59015777

Problem:

Style transfer is a popular method in Computer Vision to make creative images and artistic renditions in particular styles. It can help visualize how an image would look if painted by certain artists, treated with similar filters or taken in particular conditions. It is very common to apply a single style to the entire image, with little research in assigning different styles in different regions. We try to build a system that can assign different styles to different elements in the same image. We first try to efficiently and correctly segment all elements in the image, followed by assigning a particular style to each element type. An extended scope for this project could be building a fast network that can carry out the operations on a live video.

Motivation:

Class-based style transfer can be used widely to create visually appealing art, photo editing and fashion. It could be used to modify particular elements in the entire image. Separating each element before style transfer gives an artist more freedom on the styling of the region. Specific regions of a style image could be used to modify certain elements of the target image. Multiple styles can be used, from multiple artworks belonging to the same artist, to create a compilation of various techniques used across paintings. It can also help visualize how different artists would collaborate to create a single image. Segmentation before style transfer can also help stylize only certain regions in the image and leave out the rest as captured, to highlight those regions. This may be further extended to process certain regions through a separate model and extract information from them.

Approach:

There are two tasks in the project in the following order: Semantic segmentation with object classification and Neural Style Transfer(NST).

- Semantic Segmentation: The paper we will be referring to has used DABNet, and we plan to initially use the same. We will then try to use other state-of-the-art fast segmentation models to improve the speed of the pipeline with Cityscapes or COCO datasets.

- NST: We will first try to implement this using Fast Neural Style, as mentioned in the paper. We will then try to use a state of the art pre-trained image classification model like VGG16 or ImageNet, with Gram matrix, and stylize each segment. Once this is implemented optimally, we will try to retrain the model on WikiArt dataset and use these new weights to compare results with pre-trained weights. This will be an extended scope of the project.

- We could train on different neural networks such as VGG, ResNet, Inception, DenseNet for neural style transfer and make a comparison study

We plan to first implement the architecture proposed in the paper “Class-Based Styling: Real-Time Localized Style Transfer with Semantic Segmentation”, followed by using other models and datasets to compare results. This paper tries to combine both steps of class-based style transfer into a single step. We can try to place the steps sequentially to allow different styling for each object.

Datasets:

- Cityscape: The Cityscapes Dataset focuses on semantic understanding of urban street scenes. It provides semantic, instance-wise, and dense pixel annotations for 30 classes grouped into 8 categories (flat surfaces, humans, vehicles, constructions, objects, nature, sky, and void).

- COCO:The COCO (Common Objects in Context) dataset is a large-scale image recognition dataset for object detection, segmentation, and captioning tasks. It contains over 330,000 images, each annotated with 80 object categories and 5 captions describing the scene.

- WikiArt: WikiArt contains paintings from 195 different artists. The dataset has 42129 images for training and 10628 images for testing.

References:

- https://arxiv.org/pdf/1908.11525.pdf
- https://www.jeremyjordan.me/semantic-segmentation/
- https://github.com/IssamLaradji/CBStyling
- http://cs231n.stanford.edu/reports/2017/pdfs/416.pdf
- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9336718

Timeline:

5 May: Project proposal
7 May: Dataset finalization
10 May: Dataset pipeline
20 May: Initial models prepared for segmentation and style transfer
25 May: Combine the 2 models
4 June: Improvements on models and combined models
8 June: Full project and report submission