Background Style Transfer

Zhicheng Ding
zd2212
zd2212@columbia.edu

Qingtian Gong
qg2159
qg2159@columbia.edu

Yuanqing Hong
yh2866
yh2866@columbia.edu

Abstract

Motivated by the demand for applying an artistic style to the background only, we proposed an idea to apply style transfer to certain regions of images. To achieve this, we will combine semantic segmentation and artistic style transfer. We will first apply segmentation to the raw image and separate the humans from the background. Then we will transfer artistic style to the background only.

1. Motivation

After making an investigation in the field of artistic style transfer, we observed that most of the applications transferred the same style to all regions of the image. However, there is a demand for applying the artistic style to specific regions. For example, in traveling photos, people may just want to transfer the style of the background rather than distort the whole image including their portraits.

Inspired by this demand, in this project, we will explore the possibility to apply style transfer to certain regions (e.g. background). To achieve this, we will combine semantic segmentation and artistic style transfer. The system separates the humans from the background using semantic segmentation. Then it transfer artistic style to the background only.

2. Model Statement

The model requires two images as input, one is the target image and the other one is the style image. The model processes the target image using semantic segmentation and then transfers the style to the target image. As a result, the output image combines the content from the target image and style from the style image.

3. Methods

3.1. Semantic Segmentation

3.1.1 Model

We plan to follow the method of DeepLabv3+ [1] for image segmentation. DeepLabv3+ model combines Atrous Spatial Pyramid Pooling (ASPP) module and encoder-decoder structure to obtain rich contextual information as well as capture sharp boundaries, which well fulfills the demand of boundary recognition for style transfer.

3.1.2 Dataset

We will use the PASCAL VOC 2012 [3] dataset to train the model. This dataset contains 20 different classes (we will use people class only). The training data has 11,530 images containing 27,450 ROI annotated objects and 6,929 segmentations.

3.2. Artistic Style Transfer

We will follow the idea of Engstrom [2] to train the style transfer. This method reduces the total amount of parameters in the model to achieve faster style transfers with pretrained style.

4. Evaluation

We will evaluate Semantic Segmentation in terms of mean intersection-over-union (mIOU) of our predicted bounding box and the ground-truth. For the results of style transfer, we will compare the loss between context image and style image.

5. Task Allocation

Zhicheng Ding is responsible to explore and train semantic segmentation model while Qingtian Gong will work for implementing style transfer. Yuanqing Hong is responsible for data pre-processing and the combination of two models.

References

[1] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision – ECCV*, 2018.

[2] L. Engstrom. Fast style transfer. `https://github.com/lengstrom/fast-style-transfer/`, 2016.

[3] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html.