Class-Based Styling: Real-time Localized Style Transfer with Semantic Segmentation

Lironne Kurzman, David Vazquez, Issam Laradji

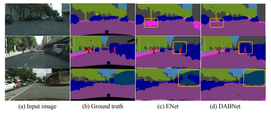
Motivation

- Style transfer allows us to create beautiful artwork by mapping the style of one image onto another, while preserving the content
- We propose a framework performing real-time localized class-based styling, and show that it achieves high-quality stylized images with real-time performance.

Related Work

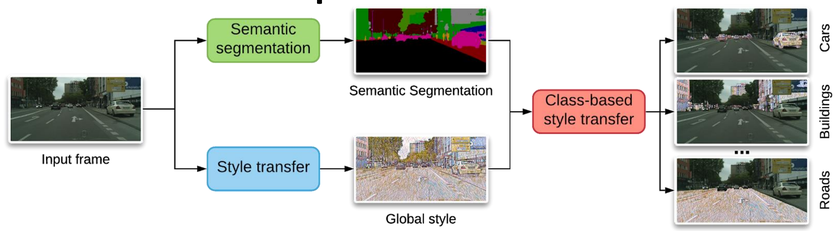
Gatys et al. [3] popularized style transfer through high-quality stylized creations by utilizing the use of deep convolutional neural networks. Unfortunately, this method is too slow to be used in real-time.

When Johnson et al. [6] extended Gatys et al. work using a feed-forward network, inference became substantially quicker and video style transfer became possible.



Real-time semantic segmentation is a trade-off between accuracy and inference speed. Li et al. [7] proposed a Depthwise Asymmetric Bottleneck module. They extract local and contextual features jointly, improving the speed without compromising accuracy.

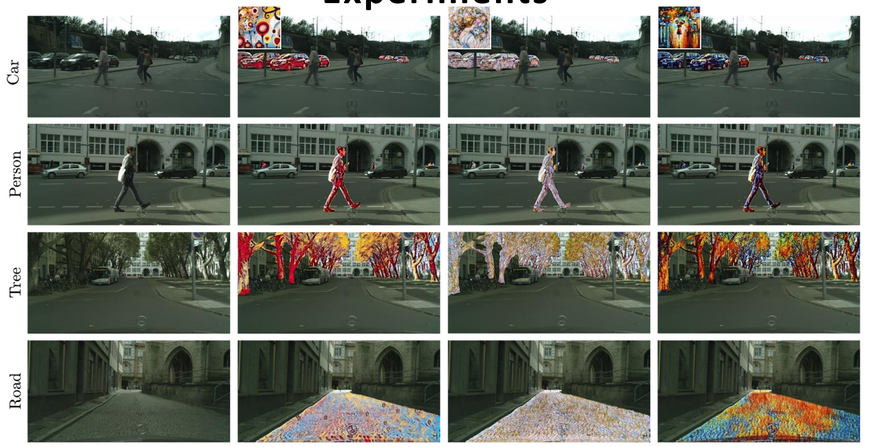
Proposed Method



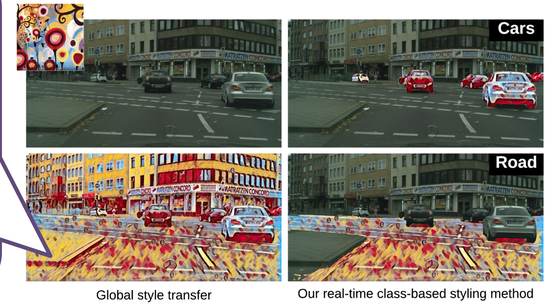
- **Pre-Processing Phase –** CBS extracts a binary mask R_c for object class c from a fast segmentation network [7]. Note that the segmentation network is pretrained on segmenting object class c. Simultaneously, CBS transforms I to a styled image T_S using a fast styling method (we used the method by Johnson et al. [6]).
- **Processing –** Using the binary mask R_c, CBS extracts the background I_b from the unstyled image I and extracts the foreground T_f from the styled image T_S. Finally, CBS adds T_f and I_b to obtain a target image U that contains an unstyled background and a styled object class c.

$$U(c,S) = (R_c * T_S) + (1 - R_c) * I$$

Experiments



we ran CBS on Tesla P100 GPUs. For the Cityscapes [2] 1024 × 2048 frames, CBS is able to process them at a speed of 16 FPS or 60.83 ms / image, making it feasible to run localized style transfer in real-time.



Global style transfer — Our real-time class-based styling method

In contrast, Castillo et al. [1] would take at least 15 seconds for a single image due to the Gatys et al. [3] method required at inference time. This makes it unsuitable for real-time style transfer. Thus, CBS can be used in more applications where artistic work requires real-time performance.

Summary

- We achieve real-time localized style transfer by simultaneously segmenting and styling each frame.
- Real-time stylizing of each class can pave the way to more appealing works of arts, or in advertisements (e.g. product styling.)

Future Work-

- we plan to extend this for 3D scenes by applying localized style transfer to 3D objects

SCAN ME

Github repo.

References

[1][Castillo et al. 2017] - Son of zorn's lemma: Targeted style transfer using instance-aware semantic segmentation.
[2][Cordts et al. 2016] - The cityscapes dataset for semantic urban scene understanding.
[3][Gatys et al. 2015] - A neural algorithm of artistic style.
[4][He et al. 2017] - Mask r-cnn.
[5][Jing et al. 2017] - Neural style transfer: A review.
[6][Johnson et al. 2016] - Perceptual losses for real-time style transfer and super-resolution.
[7][Li et al. 2019] - Dabnet: Depth-wise asymmetric bottleneck for real-time semantic segmentation.
[8][Paszke et al. 2017] - Automatic differentiation in PyTorch.
[9][Simonyan et al. 2014] - Very deep convolutional networks for large-scale image recognition.

lironkurz@gmail.com, dvazquez@elementai.com, issam.laradji@gmail.com