Semantic Segmented Style Transfer

Jihyeon Janel Lee
Julia Wang
Kevin Yang

CS 231N
Spring 2017

Background & Introduction

- One challenge for art historians has been analyzing paintings, recognizing their artists, and identifying their **style and content**.

Figure 1. *Which painting was done by Charles Pissarro or Childe Hassam (both influenced by Claude Monet)?* Deep learning algorithms can identify nuanced connections and similarities.

- Deep learning has not only enabled recognition but also **generation of art** in a particular style.

Tubingen in Germany → in the style of *The Starry Night* → in the style of *The Scream*
Figure 2. Examples of style transfer from Justin Johnson's implementation of *A Neural Algorithm of Artistic Style* by Gatys, et al.

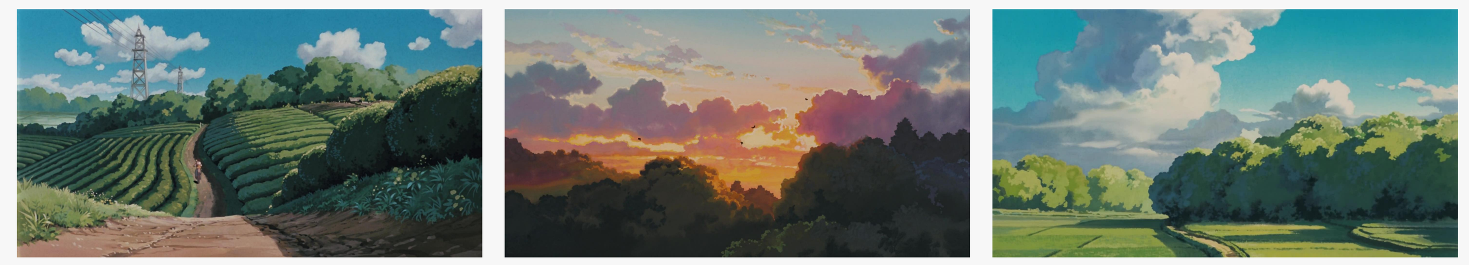
- Recently, there have efforts to develop better style transfer techniques that can process images meaningfully, e.g. **Semantic Style Transfer**.

Figure 3. Example input for *Neural Doodles*, implementation of Semantic Style Transfer (Champandard, 2016), based on the context-sensitive style transfer algorithm, Neural Patches (Li, 2016). Based on annotation and style of input, output takes a "doodle" and produces image of input style.

Problem

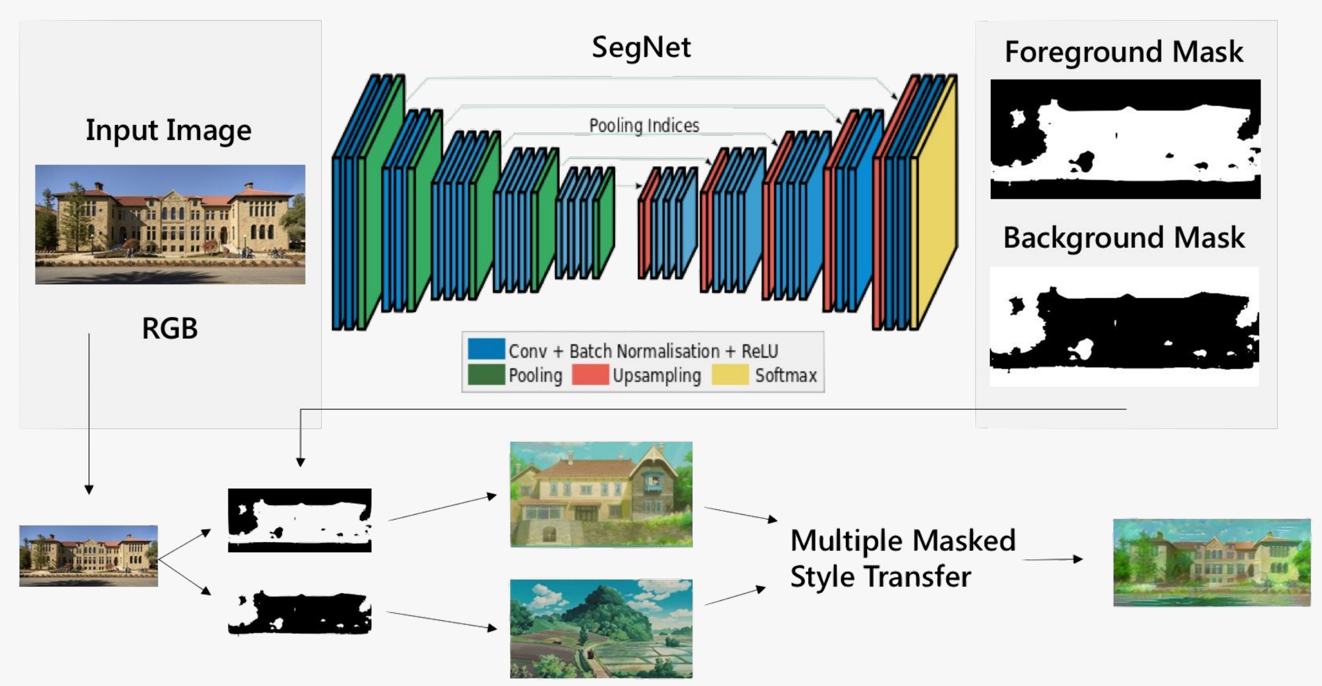
- In animated films, there is often a distinctive style characteristic of a production studio. Overarching question: **How do we capture the essence of that style?**
- We investigate a particular problem to answer this question: **using segmentation to semantically transfer style** to detailed landscapes and buildings from animation.

Datasets



- We sampled ***Studio Ghibli films*** at 5 fps using ffmpeg, particularly *My Neighbor Totoro* and *Spirited Away*, to serve as our style images.
- We obtained pictures of Stanford campus and scenery as training input images.

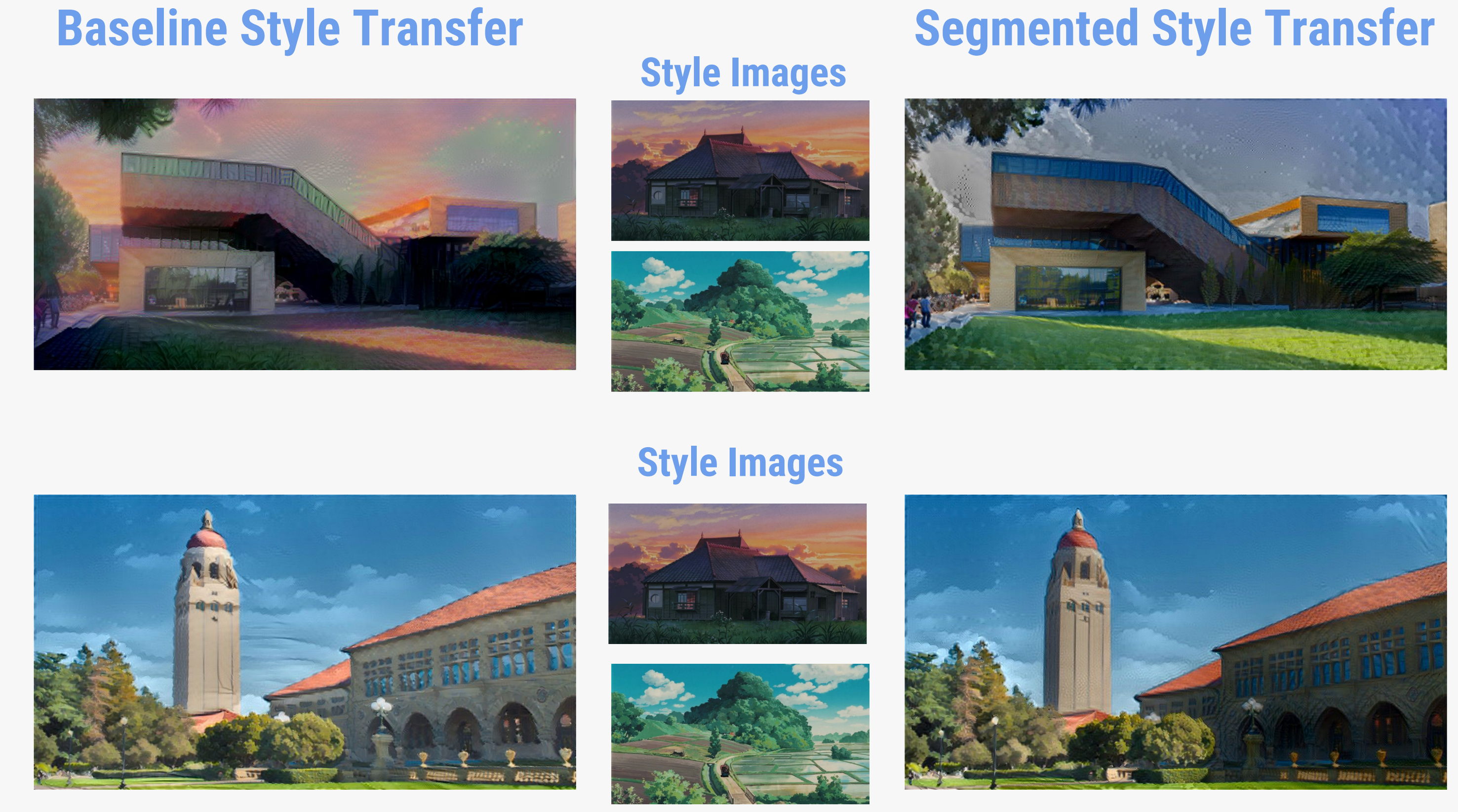
Method



Pipeline of Model:

1. Segment image into foreground and background using mask.
2. Select most suitable movie frame based on content loss.
3. Apply style transfer separately to foreground/ background from the two separate images using our modified version of the style transfer algorithm.
4. Output stylized image with different styles in the foreground/background.

Results



Evaluation: Less distortion of defining edges of buildings and trees after segmentation. Can preserve more semantic meaning with color control.

Cons: Still unclear in what the essence of the animation style is, possibly consider texture as well. Difference between movie frame and input convolutes output, need more data and better selection of style images.

Conclusion & Future Work

- We implemented style transfer semantically by segmenting the input and style images into foreground and background.
- There are many applications of this kind of selective transfer, although the evaluation can be subjective.
- In the future, we hope to use more advanced segmentation techniques to detect and localize objects to preserve more semantic meaning.
- We hope to use more sophisticated methods to identify frames in movies for background/ foreground.