NeAT: Neural Artistic Tracing for high resolution Style Transfer

[1]Dan Ruta, [1]Andrew Gilbert, [1,2]John Collomosse, [2]Eli Shechtman, [2]Nicholas Kolkin

[1]CVSSP University of Surrey, [2]Adobe Research

Background:

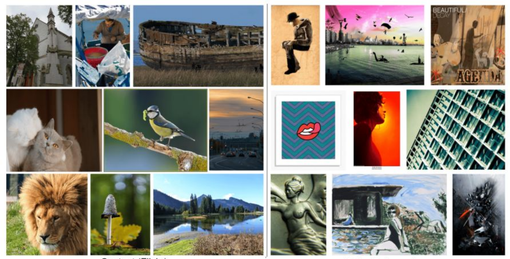
We set out to create a Neural Style Transfer (NST) method for high resolution style transfer, maintaining maximum content structure and details.

We re-formulate feed-forward style transfer as image editing, rather than image generation, resulting in a model which improves over the state-of-the-art in both preserving the source content and matching the target style.

We introduce the BBST-4M dataset, a new, large scale, high resolution dataset of 4M images. As a component of curating this data, we present a novel model able to classify if an image is stylistic. We use BBST-4M to improve and measure the generalization

Contribution #1 - BBST-4M Dataset:

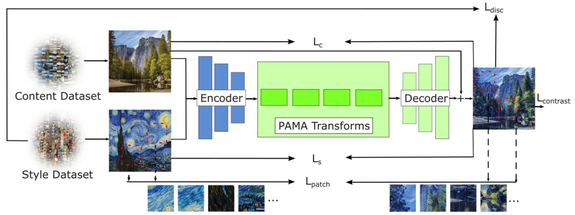
- 4M high resolution image dataset
  - 2M content images
  - 2M style images
- Content images collected from Flickr
- Style images collected from Behance.net
- Images filtered with a stylistic prediction model
  - Trained over a manually labelled subset of the StyleBabel dataset
  - The model can predict if an image is artistic, or not



Content (Flickr) Style (Behance)

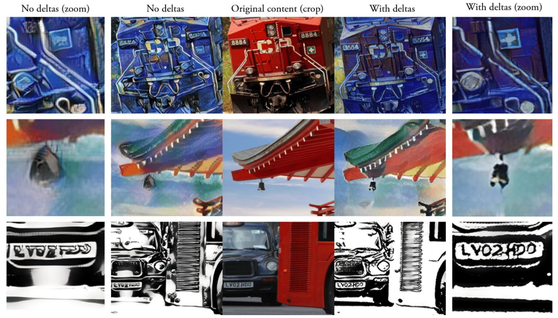
- All images are at least 1024px on the smallest side
- Training with BBST-4M leads to higher generalization
  - 56% preference of NeAT trained on BBST-4M vs WikiArt+MSCoco, on fine-art styles similar to WikiArt
  - 60% preference of NeAT trained on BBST-4M vs WikiArt+MSCoco, on contemporary digital art

Contribution #2 - NeAT style transfer model:

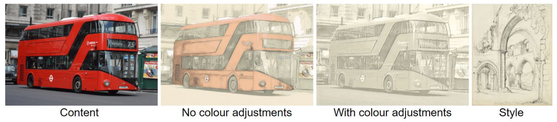


Reformulation of NST as an image editing process through use of the original image as residual

- We train the model to predict pixel deltas over the original content image, instead of re-generating all the content image details from scratch
- We avoid the lossy encoding-decoding process, and focus our model only on learning the stylistic changes required
- We maintain far more content details in our output, when using deltas

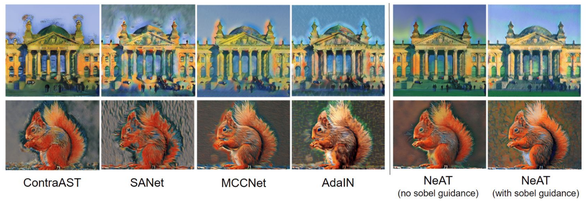


- We use additional colour adjustments on the content image, to avoid strong colours skewing the resulting image



Content No colour adjustments With colour adjustments Style

Contribution #3 - Sobel-guided patch co-occurrence discriminator loss for removing style halos:



Examples of style halos appearing around the edges of complex textures against simple textures

- In our training objective, we include a patch co-occurrence loss
- We make novel changes to this, to guide its patch selection in a non-random way
  - We match patches of simple style texture from the style image with patches of simple texture in the content image
  - Simple areas such as the background sky should have a simple style texture, whereas a complex texture such as a face should have complex style details