Neural Image Style Transfer for Handwriting Refinement

Enmao Diao Jiawei Zhou Xu Si Yuting Sun

• Problem Statement/Objective

This project aims to find out an innovative way to leverage neural image style transfer to refine handwriting with any typed style given. More Specifically, we want to focus on transformation between handwriting and LaTeX format. This method should also be able to generalize to refine handwriting with other fonts or even other people's handwriting style. For example, given a handwriting image **x** and its target typed version **y**, we can minimize the objective function that associates with them proposed in neural image style transfer. In this way, we aim to migrate the semantic content of one handwriting image to different typed styles. Our project can be applied to several other applications like handwriting to LaTex markup, signature forgery and calligraphy learning refinement.

• Approach

Generally, there are two types of methods to do neural image style transfer, namely *Descriptive Neural Methods Based On Image Iteration* and *Generative Neural Methods Based On Model Iteration*. For the first method, after properly defining the objective function, we are able to generate a styled image from random noises. As for the latter, we train an image generator network for our handwriting image with respect to the loss function defined in the former method. The main difference is that the second method approximates the optimization procedure by training an image generator network and thus only a single forward pass is required to produce the results, but we are tied to a single style that we use in the training stage. The first method, however, relies on backward and forward passing in a deep neural network which can create high computational cost [1][2]. Our application slightly differs from the standard neural image style transfer. In our situation, we would choose the second approach because we aim to process tremendous handwriting images to a small set of typed style. Thus, we only need to train the image generator network for each style on demand and each handwriting image can be processed without optimizing an objective function. In terms of training samples, typically, we have a large number of handwriting images for training and testing and the styled version can be easily created by standard font compiler.

Specifically, we will use pretrained model like VGG and ResNet as part of our loss network [3]. On the other hand, in our training stage, we will train the image generator network based on generative adversarial networks (GAN) [4]. We can synthesize the training dataset by replacing all individual symbols in im2latex dataset with randomly chosen handwritten symbols taken from Detexify's training dataset. Thus we can have a many-to-one correspondence between handwritten symbols and their typed versions.

• Evaluation

We will use [2] and [5] as our baseline for testing accuracy. Since we do not actually have a quantitative ground truth for our result, we can evaluate the loss function and compare it with [2] and [5] when we test with our synthesized handwriting images. Meanwhile, since our outputs are images, we can also perform qualitative evaluation comparing with the baseline approach. Finally, we will evaluate the computational cost for varying numbers of iterations and image resolutions.

• Collaboration Plan

In this project, our group members are Enmao Diao, Jiawei Zhou, Xu Si and Yuting Sun. Jiawei will first investigate theoretical issues that we are not familiar with like GAN and gram matrix. He will mainly work on literature review and algorithm design with the assistance of the other three. Enmao, Xu and Yuting will first set up hardware support for our deep learning models. We will then focus on data cleaning, model building, model evaluation and visualization. We will collaborate together on the final write-up.

Reference

[1] Y. Jing, Y. Yang, Z. Feng, J. Ye, and M. Song, "Neural style transfer: A review," arXiv preprint arXiv:1705.04058, 2017

[2] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In ECCV, pages 694–711, 2016.

[3] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385 (2015)

[4] Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)

[5] L. A. Gatys, A. S. Ecker, and M. Bethge. A neural algorithm of artistic style. ArXiv e-prints, Aug. 2015.