Beautifying handwriting using neural networks

Pankesh Bamotra
Language Technologies Institute
Carnegie Mellon University
Pittsburgh, PA 15213
pbamotra@cs.cmu.edu

Sanchit Agarwal
Language Technologies Institute
Carnegie Mellon University
Pittsburgh, PA 15213
sagarwa1@cs.cmu.edu

Abstract

Though we live in the age of emails and SMSs, there is no denying the importance of handwritten messages to help connect with someone at much more personalized level. However, handwritten text might impact the legibility and comprehensibility of the content. Our idea is to beautify the handwritten text by superimposing it with features from an existing typography. This way we retain the handwriting style while enhancing it with stylistic features of a known font. The idea is similar to the one employed by the popular Prisma phone application that creates artistic images by recombining style extracted from one image and content from another.

1 Overview of project idea

1.1 Why the problem is important

The problem is important as it would allow personalized textual communication without compromising on the readability and comprehensibility of the text. The work would also have creative and artistic value. Further, we believe this would be a fun problem to tackle that would advance our understanding of neural networks. Moreover, to the best of our knowledge, such work has not been done before, which also makes it exciting.

1.2 What challenges you need to solve

- We have to figure out the right granularity at which to learn the style and apply it to the handwritten text. We have two choices to explore, viz. character level and word level.
- We need the right model to extract the stylistic features from the known font. Gatys et. all, use VGG-Network in their paper on neural style [1]. We'll have to experiment and see if that works for our task too.
- We need to find the right strategy to mix the extracted style with the both content and style in the handwritten text.

1.3 Which datasets you are planning to use

- Google Fonts dataset (https://github.com/google/fonts)
- 10000fonts.com (http://www.10000fonts.com/catalog/)
- IAM Handwriting Database (http://www.fki.inf.unibe.ch/databases/iam-handwriting-database)

1.4 What metrics you are planning to use to measure your performance

We will use manual evaluation to judge the quality of the output on following three parameters: correctness, legibility and closeness to original text with style from known font.

10807 Class Project (2016).

2 Literature Survey

2.1 Learning Typographic Style - S. Baluja, Google Research

This is a crucial paper in our overall research. Baluja has shown in the paper that it is possible to learn the style features of the fonts from their corresponding images. There are two tasks presented in the paper – the first one is about classifying if a given character exhibits the characteristics of font by using only four characters from that font as features. The other task is a multi-task learning problem where the author tries to learn the other characters of the same font again using only a few characters as basis. Thus we can see that a very small subset of characters is sufficient to learn both discriminative and generative models for representing typographic style. This pair of simultaneously learning the discriminative/generative model allows the author to evaluate the generative model. The accuracy of the discriminative model is 92.1%. The neural network architecture uses convolutional layers, pooling layers, and fully connected layers.

2.2 Separating Style and Content with Bilinear Models - J.B. Tenenbaum, MIT; W. T. Freeman, MERL

This paper presents a framework for solving two-factor tasks using bilinear models for sensory inputs of lesser complexity like handwriting and image poses using efficient algorithms based on the singular value decomposition and expectation-maximization. Bilinear models are able to incorporate the phenomena of mathematical separability while allowing multiplicative interactions between the two factors being referred to as style and content. In general, these two factors can be any two independent factors from the perspective of a perceptual observation.

2.3 A Neural Algorithm of Artistic Style - L.A. Gatys, A.S. Ecker, M. Bethge - University of Tubingen, Germany

This paper proposes a view that a deep neural network can model the complex interplay between the content and style of the images. The task here was to create artistic images with high perceptual quality. The authors have used CNN based approach to obtain a representation of the style of an input image alongwith feature space representations to capture texture information. On similar terms as the previous paper, this paper beings out the fact that representations of content and style in a CNN are separable. That fact can be used to manipulate both representations independently to produce artistic styled images. An important fact the authors mention is that the visually most appealing images are created by matching the style representation up to the highest layers in the network. Typical convolutional layers like max-pooling and average pooling have been discussed in the paper.

3 Milestones

- Milestone 1: Process data
- Milestone 2: Work on CNN or bilinear models based architecture to separate style from content in fonts.
- Milestone 3: Conduct experiments on the above model
- Milestone 4 (towards the end): Work on combining handwritten text with style extracted from fonts.
- Milestone 5: Conduct experiments on the above model
- Milestone 6: Report

References

[1] Shumeet Baluja. *Learning Typographic Style*, arXiv (2016).

[2] Tenenbaum, Joshua B., and William T. Freeman. *Separating style and content with bilinear models*, Neural computation 12.6 (2000): 1247-1283.

[3] Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. *Neural algorithm of artistic style*, arXiv preprint arXiv:1508.06576 (2015).