

中文字体识别与重建

Chinese Calligraphy Font Classification and Transformation

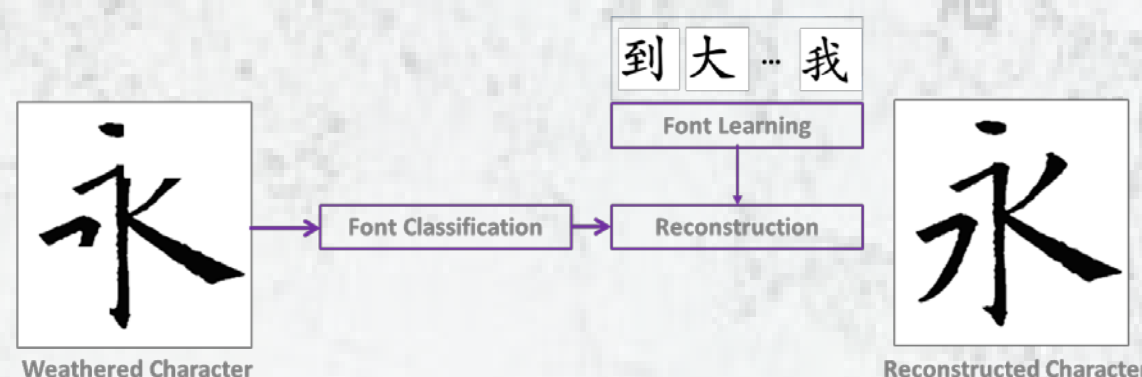
Motivation

- Many great works of ancient Chinese calligraphy were carved onto stone tablets. However, after centuries of weathering, many of the characters on these stones have been rendered incomplete.
- Inspired by this problem, we sought to answer the following question: is it possible to apply machine learning techniques to recognize the font style of a given calligraphy character, and ultimately, to reconstruct the same character in its original style?

Challenges

- Pictorial in nature, Chinese characters are significantly more complicated to process than their English counterparts.
- Classification of the font style of a Chinese character relies on our ability to extract the relevant discriminative features. However, recognizing these distinctive features is a highly non-trivial problem.

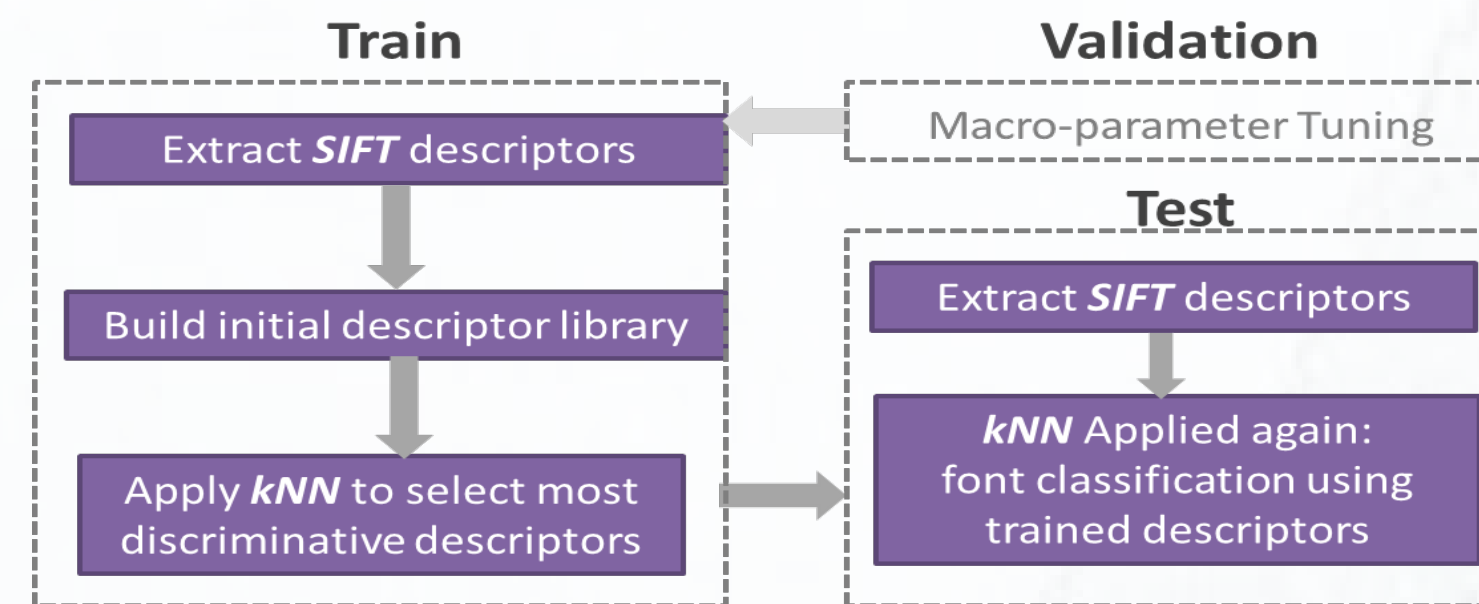
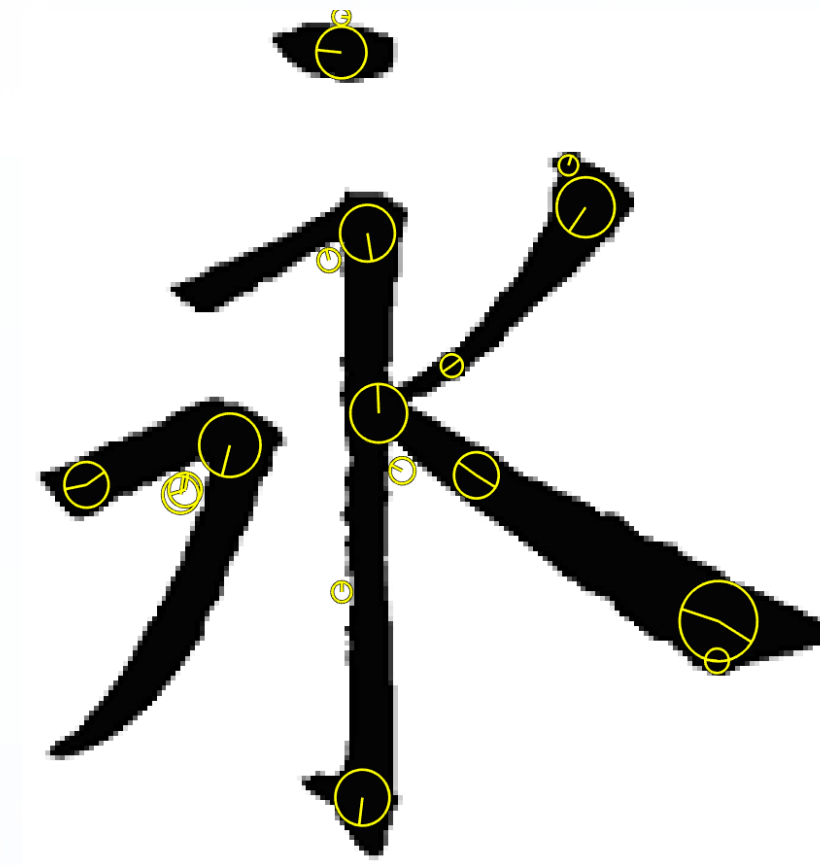
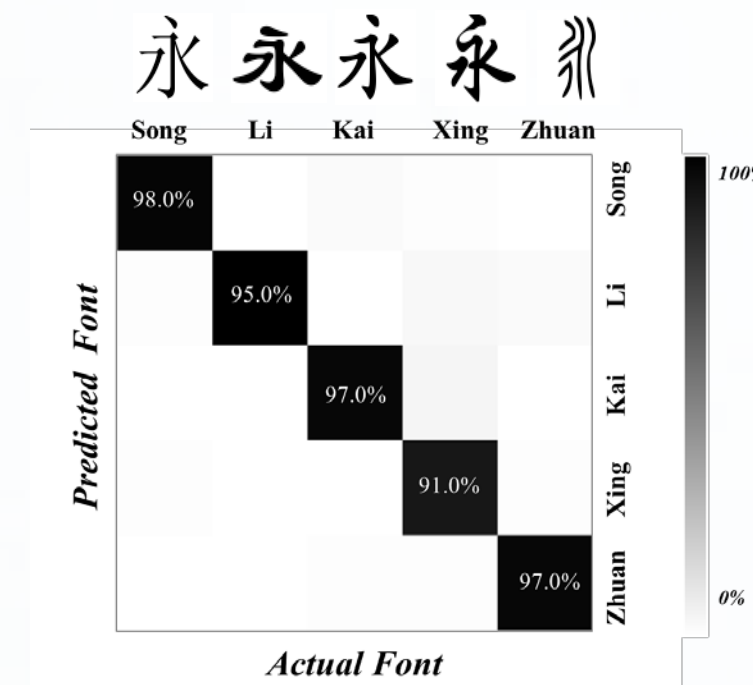
Methodology



Classification

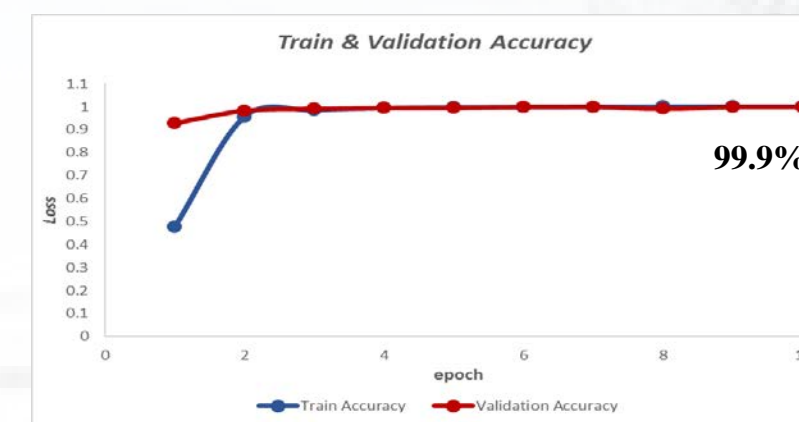
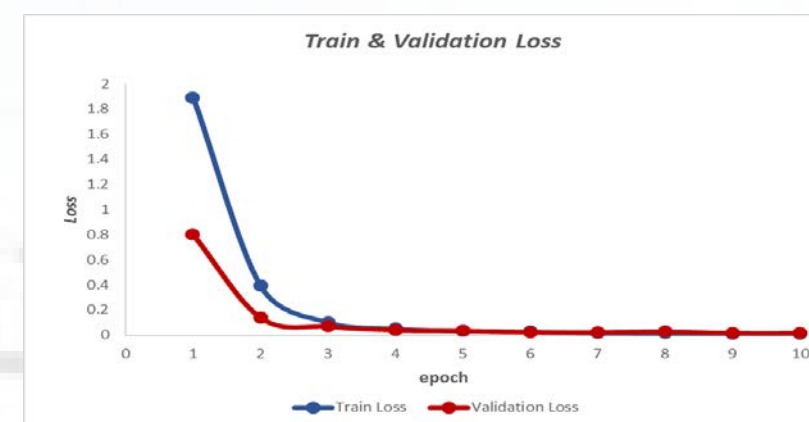
◆ Approach 1: SIFT-kNN

- 1000 Training Characters
- Choose $k_1 = 3, k_2 = 3$



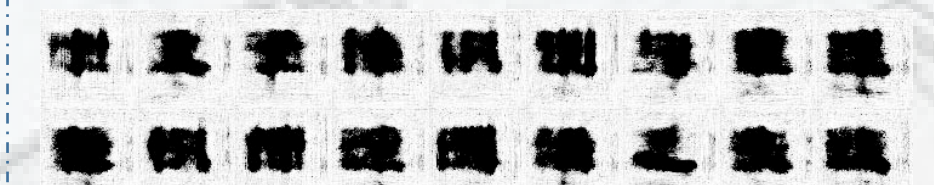
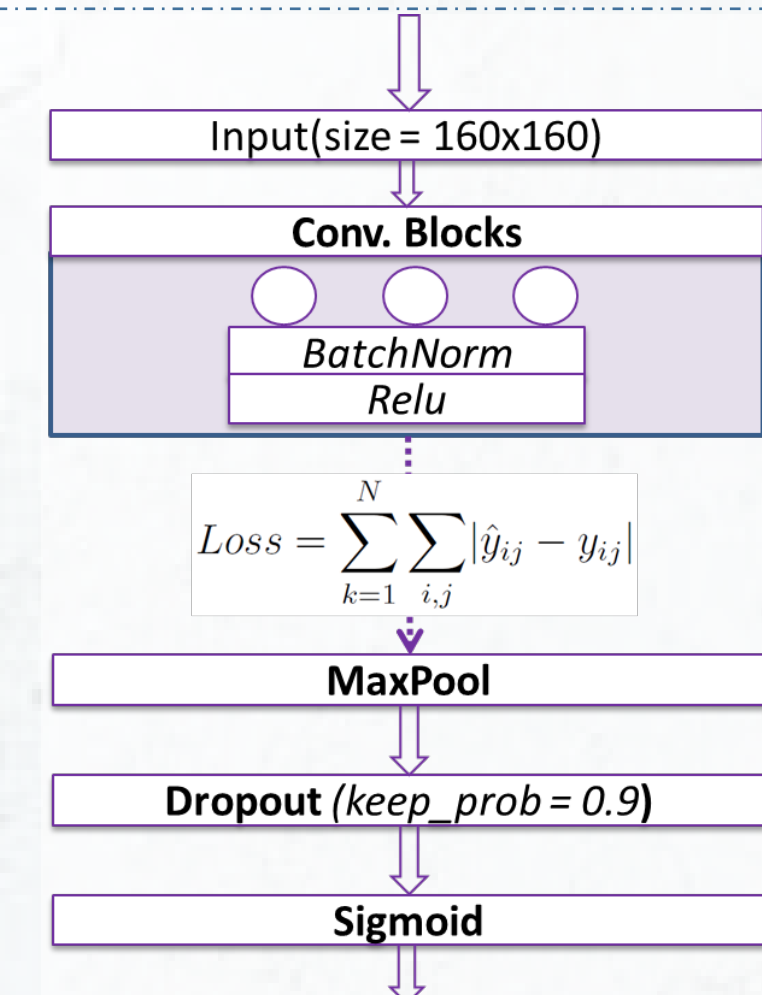
◆ Approach 2: Keras-CNN

- 500 Training Characters
- 1 layer neural network, 32 output filters with size 5x5



Transformation

中文字体识别与重建
卷积神经网络之实践



中文字体识别与重建
卷积神经网络之实践

中文字体识别与重建
卷积神经网络之实践



Stanford CS229 2016-Autumn

Li Deng Liyi Wang Zhaolin Ren

<https://github.com/dengl11/Chinese-Calligraphy-Character-Reconstruction>