Postal-address interpretation, bank-check processing, and tablet notetaking applications each attempt alphanumeric character recognition. In each of the mentioned examples, characters are interpreted against a fairly constant background. But what if the background is not so consistent? What if we need to interpret text in photos with wildly varying backgrounds - text from magazines, photos of signs, photos of license plates, etc.? And what if these photos are of poor quality? Can the characters still be identified correctly? With the use of machine learning algorithms, recognition of obscure text is possible.

Using a dataset of alphanumeric characters from Google Street View Images, I will employ several classification techniques, such as KNN, SVM, and MLP (and possibly other techniques as they are presented in lecture) to assign a label to several thousand characters. I will then compare the results of each of the classification methods. The possible characters in the dataset are ‘A-Z’, ‘a-z’, and ‘0-9’. Each image file contains a single dominant character although it may be partially obstructed, blurry, sloppy, or dull. This dataset can be found and downloaded from Kaggle.com.

Kaggle offers a short tutorial on the Julia programming language and a tutorial into the implementation of the KNN algorithm in Julia. Therefore, I intend to process some of my data using Julia. I will likely also utilize the Caffe deep learning framework.

References


http://caffe.berkeleyvision.org/