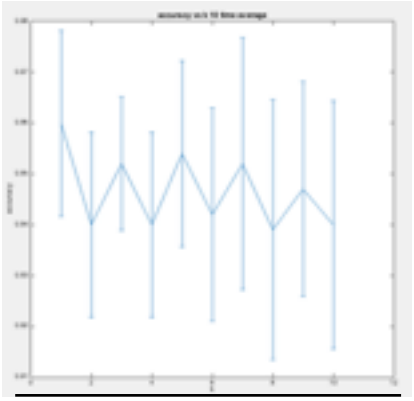
James W Peach

COMP24111- Ex2 - Machine learning
Handwritten Digit Recognition

Evaluation of algorithms

The two algorithms we implemented were the k-nearest neighbour algorithm and the perceptron.

There are lots of different metrics that you can use to compare these algorithms not just accuracy. the speed of training and the speed of testing, the memory usage again during training and during testing. one metric that i was interested in was the accuracy across all the digits not just the 3 6 8's



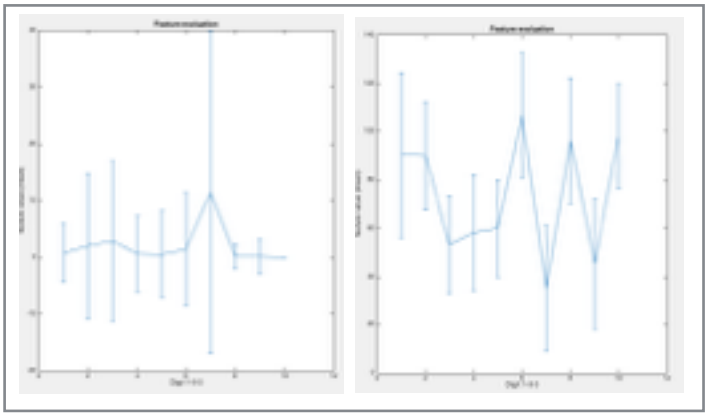
Accuracy of k - over 10 iterations.

Evaluation of features

In order to increase speed and efficiency and to enable more accurate classification I must generate increasing amount of features to test against. however too many features can hinder rather than help.

So I came up with a way of curating my features to improve on my classification.

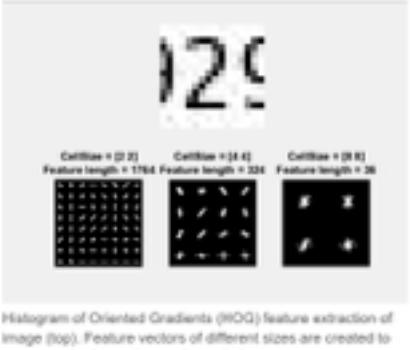
I first decided upon some very simple feature sets for example splitting the image into quadrants and having one feature per quadrant with the mean value of the pixels. I then ran my feature extraction for every example in the data and I plotted an error bars graph of the result for every feature. The graphs clearly depict how the mean and standard deviation changes for each digit in both the features. I can discard the feature in the left graph because I can see that the values are fairly similar for each digit and in order to maximise the accuracy then you need as little overlap between the features as possible.



Histogram of oriented gradients - Feature extraction (Bonus)

One of the popular ways of extracting features from images for object detection is the HOG feature extraction technique. this algorithm extracts information about the direction of the gradients in the image. this is particularly useful as this mimics the way that the human mind must distinguish between the different digits. In my testing of the features as above the HOG feature extraction proved much better in recognising and differentiating multiple digits with fewer features than I could do before.

The implementation is used within the licence provided in the code from the Matlab code sharing website.
(http://www.mathworks.co.uk/matlabcentral/fileexchange/46408-histogram-of-oriented-gradients--hog--code-using-matlab)



Feature extraction from digits (Mathworks) Computer vision