Project: Offline mathmatical expression recognizer

Memmber:

- Yang Zhang
- Yisu Peng

DATA:

1. We find three types of data: numbers, English letters and mathmatical symbols.
    - for numbers, we use mnist data:
        - Mnist data can be download directly from the website without polishing too much.
    - for English letters, we find Chars74K(English Hand writting) data:
        - We have not work on this data yet.
    - For mathmatical symbols, we use detexify data:
        - We made serveral scripts to transform detexify data to fit caffe.
        - convert stroke(svg) to png; we made the stroke bold which turned out a better result
        - composited the color of the image
        - resized image to 28x28 (this number is taken from mnist)
        - saved images into lmdb

Deep Leaning Framework:

1. We tried torch, theano and caffe. For torch, both of us were not familiar with Lua so we did not choose it. Theano is written in pure python and focus on fundamental deep learning mathematic while caffe has higher level of abstraction so we decided to use caffe.
2. Built up envioronment to run Caffe.

Experiments:

The first thing we tried is to use original LeNet to do classification on a subset of math symbols (we picked 10 symbols from detexify data). We got very low accuracy, less than random guess. We figured out the problem is that the learning rate is too high that caused the training to be diverging. After we decreased the step size to 0.001, we get about 50% percent accuracy, but it's still not plausible enough.

Then we tried to composite the color from the white background to the black background. We do this because we found that lenet is using this format. We believe this may improve the computational efficiency. We get only little improvement.

The major change we did was to bold the stroke when generating the image from stroke trail data to an

image, we got 86.77% correct rate.

Future Planing:

1. Try to segmentation (split the math formula, maybe have not time for this one, we may focus on neural network)
2. Try to extend the number of classes to 100+
3. Analyse caffe lmdb database data structure because we need to add our data to it