Deep Convolutional Neural Networks

Donnie Kim[1] and Martin Zhou[2]

1 Department of Bioinformatics and Computational Biology, GSBS MDACC, Houston, TX
2 Department of Computer Science, Rice University, Houston, TX

Abstract

SVHN is is a real-world challenging image dataset and we built a convolutional neural network (CNN) architecture that classifies different digits. We achieved an accuracy of 100%/97.7%/95.59% on training/validation/test set.

Problem

Can you **determine the digit** from house numbers in Google Street View House Number (SVHN) images?

Data Set

Each SVHN image is 32 by 32 in pixel and has a digit label ranging from 1 to 10.

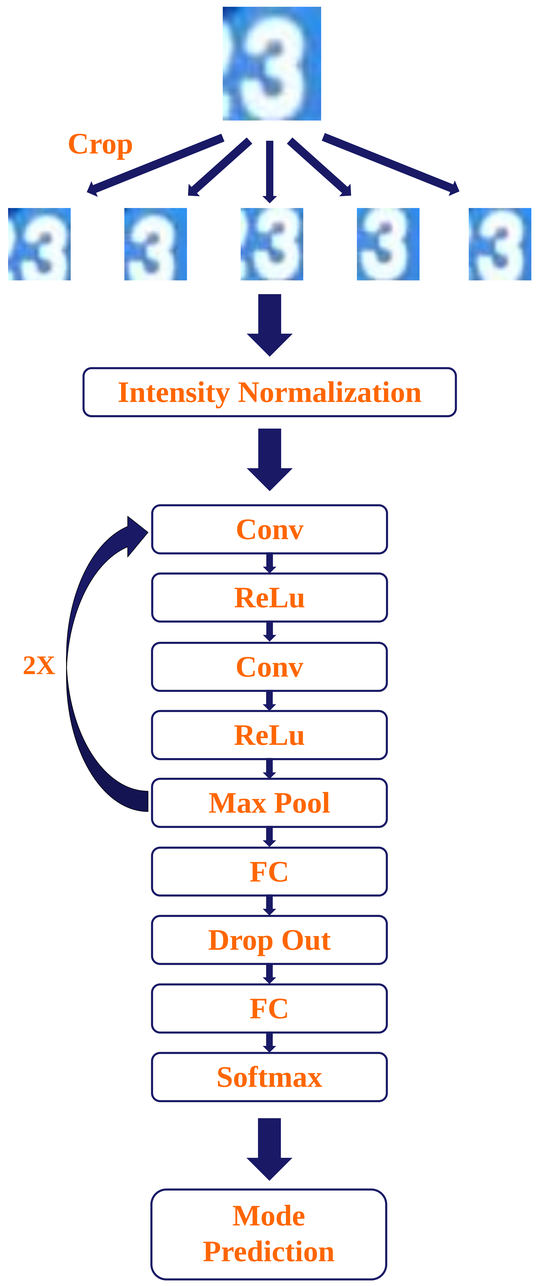
- Training dataset = 73,257
- Test dataset = 26,032
- Extra dataset = 531,131

CNN & Experiment

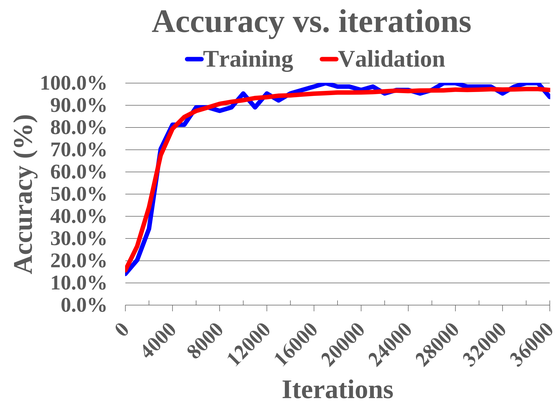
We used MNIST convolutional neural network (CNN) architecture in Tensor Flow tutorial as our initial model and developed from there. The following questions were addressed:

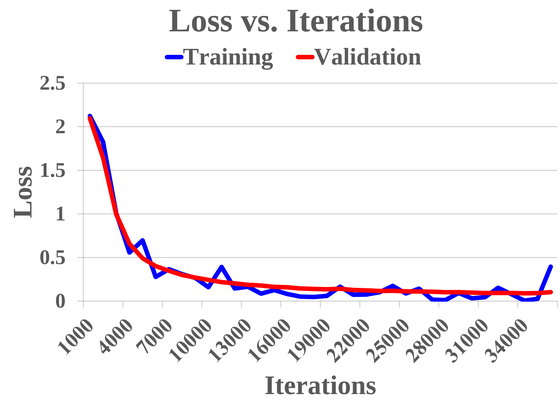
1. What **filter size** to use?
2. How many **number of filters**?
3. How **deep** should the architecture be?
4. Impact of **learning rate**?
5. Impact of **Drop out rate**?

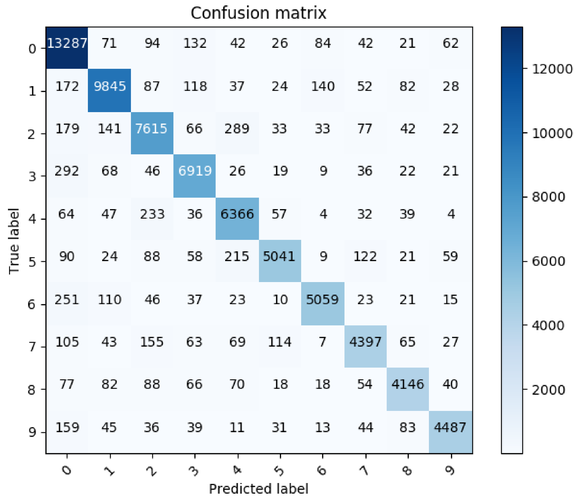
Classification Pipeline



Results







Analysis & Conclusions

- More the number of filters, better the accuracy
- Having a smaller filter size increases accuracy
- The depth of the CNN can only improve the accuracy so much and quickly saturates.
- More the data, better the accuracy.
- Training dataset => Extra dataset => Extra dataset with 5 fold crop
- With our best model, we achieved an accuracy 100%/ 97.7%/ 95.59% accuracy on train, validation, and test sets.