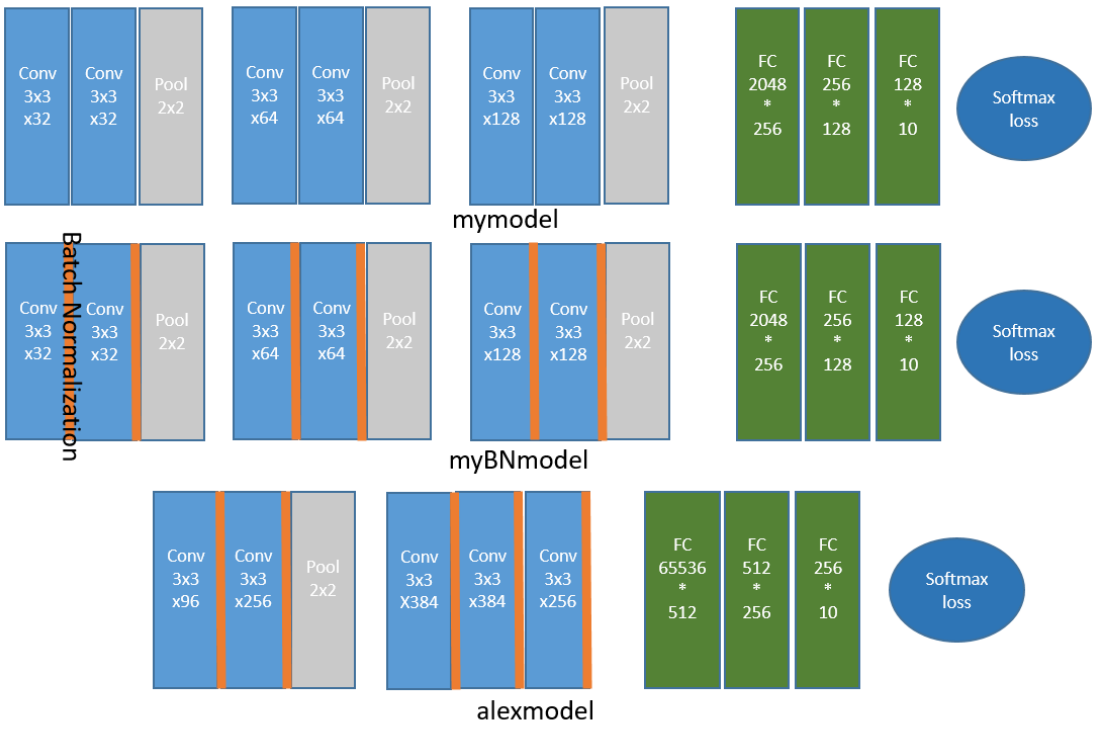
DataScience Hw3 report

R05943011 沈恩禾

Architectures

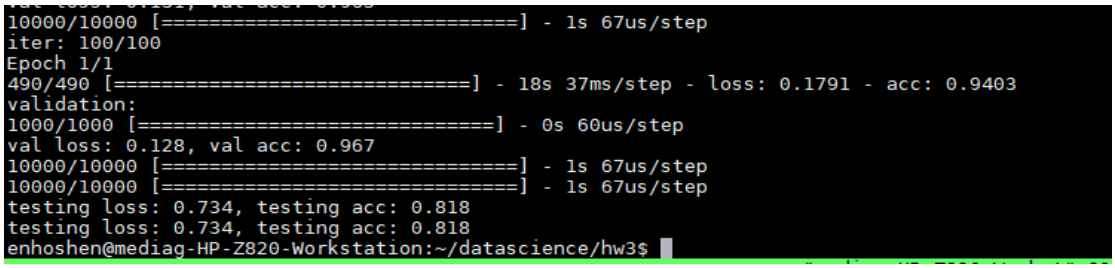


I experimented with 3 models, as fig1 showed, derived from the original CNN model for cifar10 and AlexNet , implemented on Keras with Tensorflow backend.

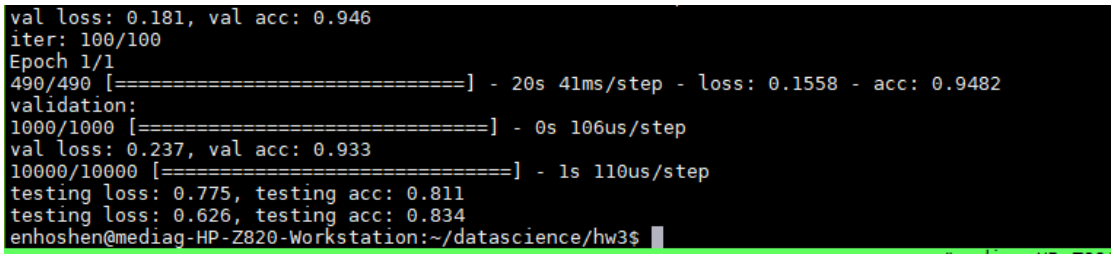
- Relu for activation function after each convolution layers
- Add batch normalization after activation function of each convolution layers for the last two models.
- Adam optimizer is used with initial learning rate 0.0002
- Batch size is 100
- Train 100*1 epochs
- Preprocess the image by flipping horizontally, shifting(0.2 units towards each direction) and rotating (at most 10 unit) using `Keras.preprocessing.image.ImageDataGenerator`

Experiments

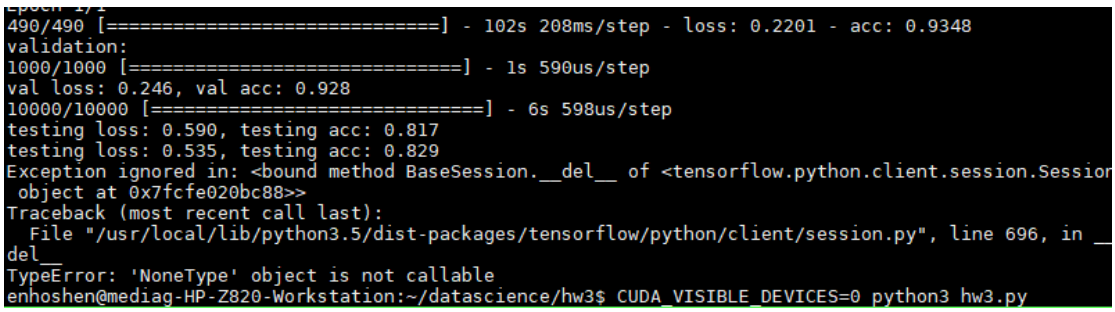
Experiment	Testing acc after 100 epochs
mymodel w/o data preprocessing	<70%
mymodel	81.8%
myBNmodel	83.4%
alexmodel	82.9%



mymodel



myBNmodel



Alexmodel

- Mymodel without data augmentation won't beat the 75% standard, which tells us that the testing set is way harder than the original training set, data preprocessing is promising.
- Mymodel with batch normalization ended up the best model after 100 epochs, and it did show that it converged faster, and better.
- Alexmodel with larger model capacity ended up beat original mymodel, but validation accuracy haven't converge yet, since it's passed the 75% standard I only examine its performance after same epochs.