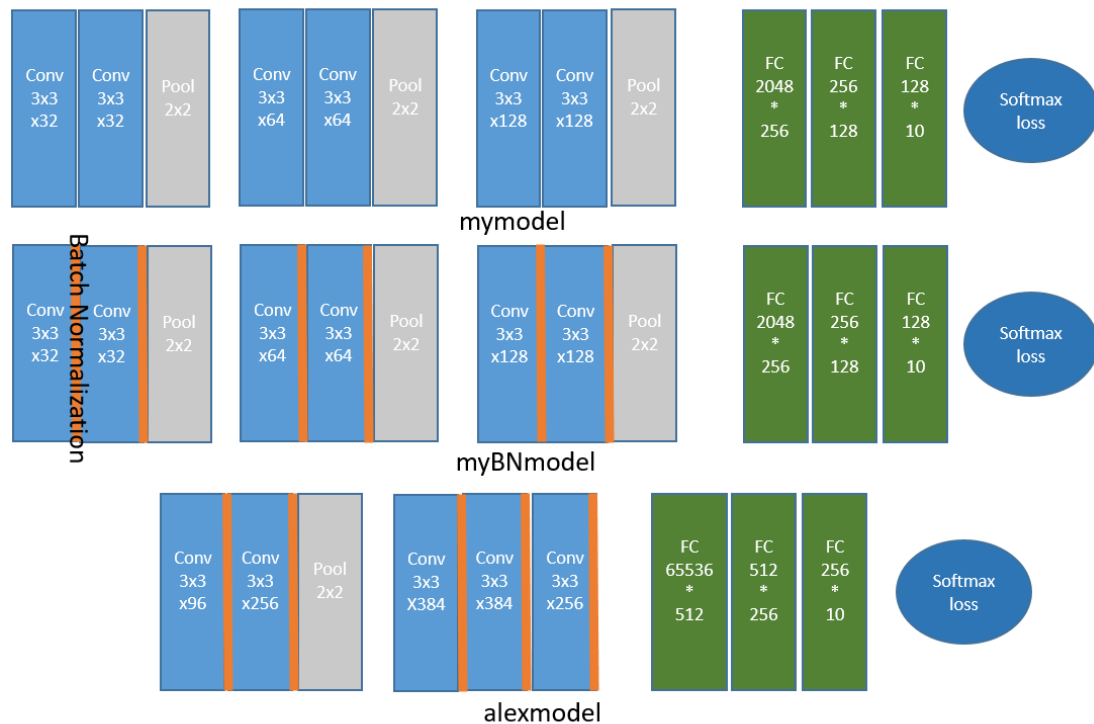


DataScience Hw3 report

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%

```
10000/10000 [=====] - 1s 67us/step
iter: 100/100
Epoch 1/1
490/490 [=====] - 18s 37ms/step - loss: 0.1791 - acc: 0.9403
validation:
1000/1000 [=====] - 0s 60us/step
val loss: 0.128, val acc: 0.967
10000/10000 [=====] - 1s 67us/step
10000/10000 [=====] - 1s 67us/step
testing loss: 0.734, testing acc: 0.818
testing loss: 0.734, testing acc: 0.818
enhoshen@mediag-HP-Z820-Workstation:~/datascience/hw3$
```

mymodel

```
val loss: 0.181, val acc: 0.946
iter: 100/100
Epoch 1/1
490/490 [=====] - 20s 41ms/step - loss: 0.1558 - acc: 0.9482
validation:
1000/1000 [=====] - 0s 106us/step
val loss: 0.237, val acc: 0.933
10000/10000 [=====] - 1s 110us/step
testing loss: 0.775, testing acc: 0.811
testing loss: 0.626, testing acc: 0.834
enhoshen@mediag-HP-Z820-Workstation:~/datascience/hw3$
```

myBNmodel

```
Epoch 1/1
490/490 [=====] - 102s 208ms/step - loss: 0.2201 - acc: 0.9348
validation:
1000/1000 [=====] - 1s 590us/step
val loss: 0.246, val acc: 0.928
10000/10000 [=====] - 6s 598us/step
testing loss: 0.590, testing acc: 0.817
testing loss: 0.535, testing acc: 0.829
Exception ignored in: <bound method BaseSession.__del__ of <tensorflow.python.client.session.Session
object at 0x7fcfe020bc88>
Traceback (most recent call last):
  File "/usr/local/lib/python3.5/dist-packages/tensorflow/python/client/session.py", line 696, in __
del__
TypeError: 'NoneType' object is not callable
enhoshen@mediag-HP-Z820-Workstation:~/datascience/hw3$ CUDA_VISIBLE_DEVICES=0 python3 hw3.py
```

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.