CSC7343

HW2 Due Apr. 2nd (23:59 central time)

Include the writeup and all your code in a zip file and upload it in moodle on or before the due time. (Use one .py file for the experiment on one model variant. Name the files as minivgg.py, var1.py, var2.py and var3.py.)

Implement a mini-vgg network and several of its variants. Train and test the models on the cifar-10 dataset. Investigate and compare the performance of the variants to the original model. Write half to one page to 1) summarize your experiment results and discuss 2) the classification performance of the models as well as 3) their size (# of parameters) and 4) the computation time to train them.

Mini-VGG network structure (**all conv/fc layers except the last one use relu activation**):

Layer	Type(window size) – n filters
1	Conv3 – 64
2	Conv3 – 64
3	Maxpool – 2x2
4	Conv3 – 128
5	Conv3 – 128
6	Maxpool – 2x2
7	Conv3 – 256
8	Conv3 – 256
9	Maxpool – 2x2
10*	fully Connected 512
11	soft-max

* Note your need a reshape layer before this layer to reshape the data

- Implement the mini-vgg model and report its performance (in terms of test accuracy, same below).
- **Variant 1**: Remove the maxpool layers. Using stride=2 in the conv layer before the maxpool to achieve similar size reduction. Would the performance improve? Report the performance of the variant.
- **Variant 2**: Add a few dropout layers in the model. Would the performance improve? Try 2 different ways to add the dropout layers. Describe the ones you tried and their performance.
- **Variant 3**: Remove layers 9 and 10. Add two layers of (1, 1) convolution: conv (1, 1) – 128; conv (1, 1) – 10. Then add "GlobalAveragePooling2D" to merge feature maps before pass them to layer 11. This is an all-convolutional structure (no fully connected layers). Report the performance of this variant.