

COL341

Assignment 3 Report

Week 1. 10/10/2019.

We've started approaching this assignment by implementing various well-known neural network architectures and modifying them to find the best performing one.

So far we've tried

- The sequential model from assignment 2, with added dropouts after each block of layers — 35.38% accuracy.
- A modified version of the sequential model from assignment 2 — 43.93% accuracy.
Modifications being
 - Increase the number of layers, i.e. make the model deeper.
 - Dropout after blocks of layers. We added more dropouts in the later parts, less dropouts in the earlier parts of the network.
 - Regularize weights after each convolution layer.
 - Increase patience of early stopping callback.
- ResNet-50 [1] with 3 layer blocks — 26.7% accuracy.
- The previous model was overfitting, so we reduced the number of layers and tried again. This time there was no overfitting, the model stopped training due to early stopping criteria — 28.56% accuracy.

In the coming weeks, we're planning to implement Res2NeXt-29 [2], optimize it as much as we can and then add squeeze-and-extraction blocks [3] to each block of that.

Citations

- [1] <https://arxiv.org/abs/1512.03385>
- [2] <https://arxiv.org/abs/1904.01169v2>
- [3] <https://arxiv.org/abs/1709.01507v4>