24-789: Deep Learning for Engineers Assignment 1 Programming Report 2

Angelos Mavrogiannis

February 2020

1 CNN Model

The model is based on the work introduced in [1] and the code implementation on [2]. The inputs of the network are $32x32$ images, with the per-pixel mean subtracted.

- The first layer is a 3x3 convolution layer.

- Then, a stack of 6n layers with 3x3 convolutions is used on the feature maps of sizes {32, 16, 8} respectively, with 2n layers for each feature map size.

- The numbers of filters used are {16, 32, 64} respectively.

- The subsampling is performed by convolutions with a stride of $s = 2$.

- The network ends with a global average pooling, a 10-way fully-connected layer, and a *softmax*.

- There are totally $6n + 2$ stacked weighted layers.

Network Architecture

output map size	32x32	16x16	8x8
of layers	$1 + 2n$	$2n$	$2n$
of filters	16	32	64

2 Hyperparameters

- Learning Rate $\eta = 0.001$

- Number of Epochs: 20

3 Results

The accuracy of the model on the CIFAR-10 images is **82.52%**

References

[1] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 770–778.

[2] becauseofAI. *322 CNN ResNet$_C$ifar10Tutorial*. 2019. URL: https://pytorch-tutorial.readthedocs.io/en/latest/tutorial/chapter03_intermediate/3_2_2_cnn_resnet_cifar10 (visited on 02/20/2020).