EE243: Advanced Computer Vision
Assignment #5

Ali Jahanshahi (862029266)

May 2019

1 ImageNet classification using deep CNNs

In this assignment, we are asked to train a Convolutional Neural Network (CNN) from scratch. We used the CIFAR-10 dataset, which contains 32×32 images divided into 10 categories. The training set contains 50, 000 images and the test set contains 10, 000 images.

Figure 1 shows the architecture of our custom CNN which is reported by `pytorchsummary` module. We used $Dropout$ layer after $Convolution$ layer (0.2) and between $Linear$ layer and classifier layer (0.5). We achieved 75% accuracy after 25 epochs.

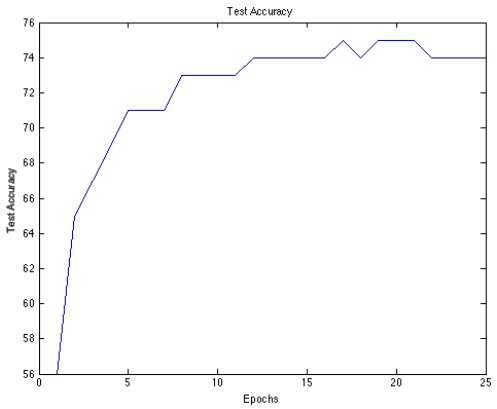
Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 16, 32, 32]	1,216
ReLU-2	[-1, 16, 32, 32]	0
Dropout-3	[-1, 16, 32, 32]	0
BatchNorm2d-4	[-1, 16, 32, 32]	32
MaxPool2d-5	[-1, 16, 16, 16]	0
Conv2d-6	[-1, 32, 16, 16]	12,832
ReLU-7	[-1, 32, 16, 16]	0
Dropout-8	[-1, 32, 16, 16]	0
BatchNorm2d-9	[-1, 32, 16, 16]	64
MaxPool2d-10	[-1, 32, 8, 8]	0
Conv2d-11	[-1, 64, 8, 8]	51,264
ReLU-12	[-1, 64, 8, 8]	0
Dropout-13	[-1, 64, 8, 8]	0
BatchNorm2d-14	[-1, 64, 8, 8]	128
MaxPool2d-15	[-1, 64, 4, 4]	0
Conv2d-16	[-1, 128, 4, 4]	204,928
ReLU-17	[-1, 128, 4, 4]	0
Dropout-18	[-1, 128, 4, 4]	0
BatchNorm2d-19	[-1, 128, 4, 4]	256
MaxPool2d-20	[-1, 128, 2, 2]	0
Linear-21	[-1, 100]	51,300
Linear-22	[-1, 10]	1,010

Total params: 323,030
Trainable params: 323,030
Non-trainable params: 0

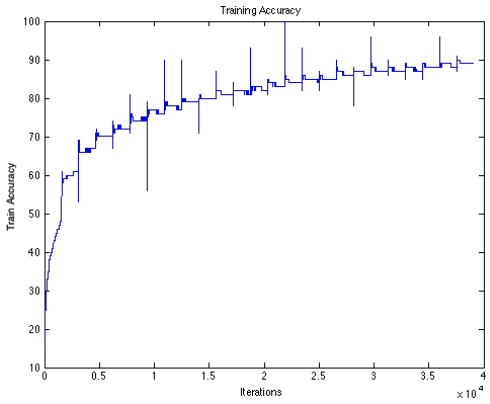
Input size (MB): 0.01
Forward/backward pass size (MB): 1.00
Params size (MB): 1.23
Estimated Total Size (MB): 2.24

Figure 1: CNN architecture obtained from `pytochsummary`.

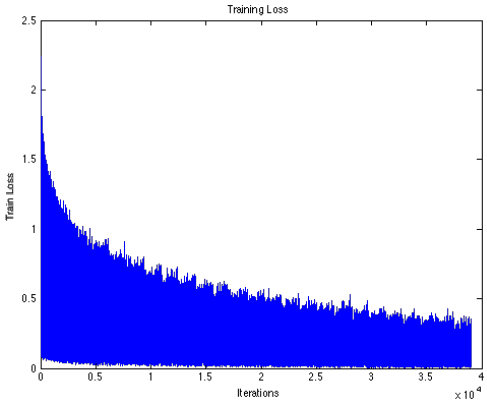
Figure 2 shows training loss, train accuracy and test accuracy. Comparing Figure 2a and Figure 2b, we can infer that the model is over-fitted, because the training accuracy is higher than the test accuracy.



(a) Test accuracy.



(b) Training accuracy.



(c) Training loss.

Figure 2: Training and test accuracy, and training loss of the CNN model in Figure 1

.