Assignment 7 - CNN
Advanced Topics in Neural Networks

George Laurențiu APETRIA

General notes

Starting from solving the previous homework, the following general aspects should be stated: the training images consist of 10,000 32x32 color images. In the case of the classification problem, the class to which each image belongs is represented by an integer from the set {0, noOfClasses - 1}. For each image, this integer is called the label. In the present case, there are 10 classes (airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks) represented by the labels 0, 1, 2, ..., 9. The model implemented this time do not use the same training approach. Instead, the ground truth images generated based on the transformations performed through the library functions are used as ”labels”.

I. Model

The constructor receives as input parameters the dimensions of an image. The dimensions of the input data are needed to define the size of the fully-connected layers; the convolutional ones will know how to deduce the size from the tensors it receives as input.

A. Convolutional layer 1

The first parameter refers to the number of pixel values of the input images (i.e. the number of color values, color channels). Since we are working with color images, this parameter has the value 3. The first convolutional layer generates 32 feature maps applying 32 filters of size 3x3.

B. Convolutional layer 2

The first parameter the number of channels - will be generated by the number of output channels provided by the first convolutional layer. Since the first layer generates 32 feature maps, then the number of input channels for the second convolutional layer will be 32. The applied filters will be as in the case of the first convolutional layer of size 3x3.

C. Pooling layer

Through the second convolutional layer, a number of 64 feature maps are provided. Thus, a 2-fold reduction in the dimensions of the generated feature maps will be required.

D. Activation function

The activation function used is LeackyReLU().

E. Other info

In order to prevent the sudden decrease of the values of the pixels in the image, both the normalization of the training instances and a particular way of initializing the weights of the fully-connected layers were used.

II. Loss function

The MSE function is widely used function in regression problems. Considering that in this situation we want to determine the differences between the input characteristics and the expected output (generated image vs ground truth), MSE offers an optimal estimate in this sense. Moreover, the degree of interpretability of the results provided by the MSE function is increased. A lower value produced by the MSE function indicates that the predictions are on average closer to the actual values. On the contrary, if the value of the MSE function is higher, then we understand that the predicted values are not really close to the expected values.

III. Early stopping criteria

The early stopping criteria has the role of preventing unjustified prolongation of the training process under well-defined conditions. These conditions refer to the evolution of the loss in the training process. Thus, if the loss does not decrease considerably during the epochs, then the training ends. In addition, the early stopping criteria also plays an important role in preventing the phenomenon of model overfitting.

IV. Benchmark

The entire training process was performed using both a CPU and a GPU. There are differences between the two approaches, in terms of the required execution time. By using the GPU to run the ”test_inference_time” function, lower execution time was recorded.