Report for Homework 1

Giovanni Giliberto
Department of Economics and Business, University of Catania
Knowledge Discovery Course
giovanni.giliberto18@gmail.com

26th of April, 2022

1 Dataset Description

The dataset that is going to be used is called "STL-10" and it can be found at the following link: https://cs.stanford.edu/ acoates/stl10/.
It is composed by a total of 13000 images, 8000 will be used as train and validation sets and 5000 as test set. Images are 96x96 pixels but their size have been reduced to 48x48 pixels in order to increase training speed. There are a total of 10 classes: airplane, bird, car, cat, deer, dog, horse, monkey, ship, truck. Augmentation has been introduced in order to increase the diversity of data available for the training model. Those are the transformations utilised:

- **Random Rotation by 20 degrees**
- **Random Horizontal flip with 0.5 probability**
- **Random Vertical flip with 0.5 probability**

The images have been then transformed into tensors and after that they have been normalized with mean=0.5 and standard deviation=0.5

1.1 Split

The dataset has been divided into 3 different sets:

- **training:** 7200 images
- **validation:** 800 images
- **test:** 5000 images

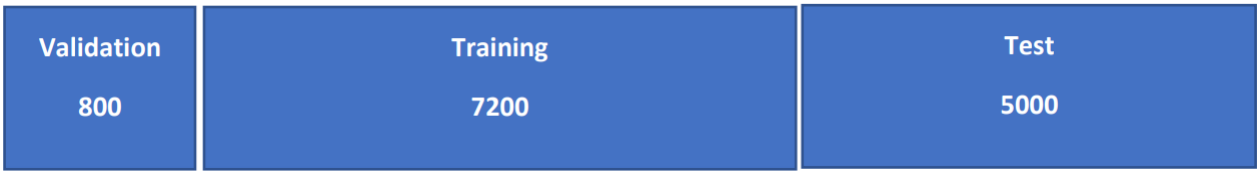

Figure 1: Dataset splits

2 Model Description

The Network is composed by 4 different convolutional layers, each of them followed by a ReLU function and a Batch Normalization layer, then a Fully Connected layer and the Classification layer are added at the end of the model. For each layer the kernel size has been set to 3, padding equal to 0 and stride equal to 1. In layers 2, 3 and 4 a Max pooling layer is added with kernel size equal to 2 and stride equal to 2.

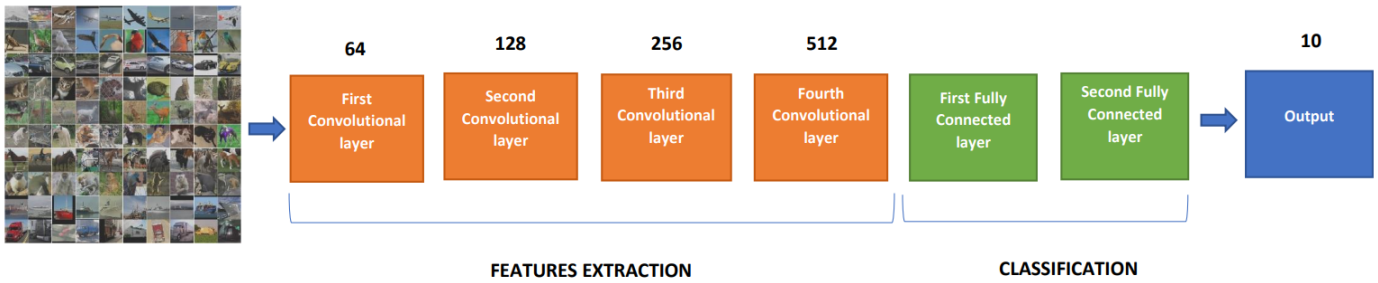

Figure 2: Model Representation

3 Training procedure

For the training procedure it has been utilised the Stochastic gradient descent (SGD) optimizer with a learning rate equal to 0.01. The criterion loss adopted is the Cross Entropy Loss. The number of epochs has been set to 25 as it has been noted that further epochs do not bring any valuable increase in performance and accuracy. The size of each batch is set to 64.

4 Performance analysis

The model featuring 4 convolutional layers plus 2 fully connected layers reported those values:

- **Train loss= 0.3507**
- **Train Accuracy = 0.8859**
- **Validation loss= 1.1675**
- **Validation Accuracy= 0.6250**
- **Test loss= 1.3058**
- **Test Accuracy= 0.6003**

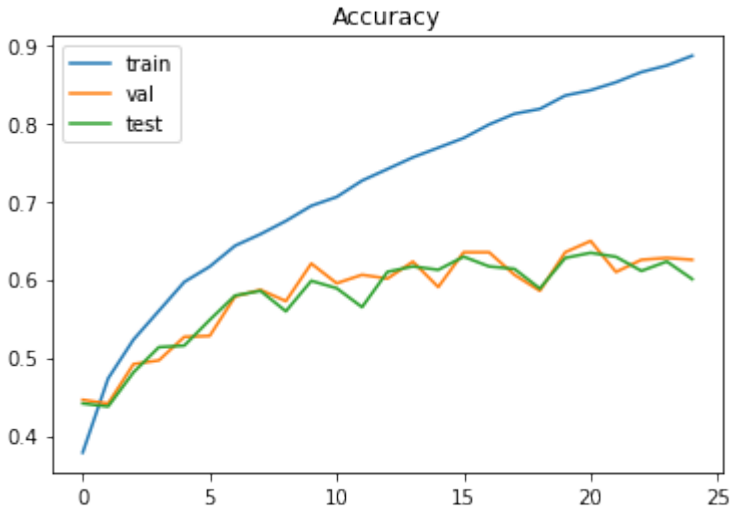

Figure 3: Accuracy

The confusion matrix shows the following results for each class:

- **Airplane** 90 percent of accuracy

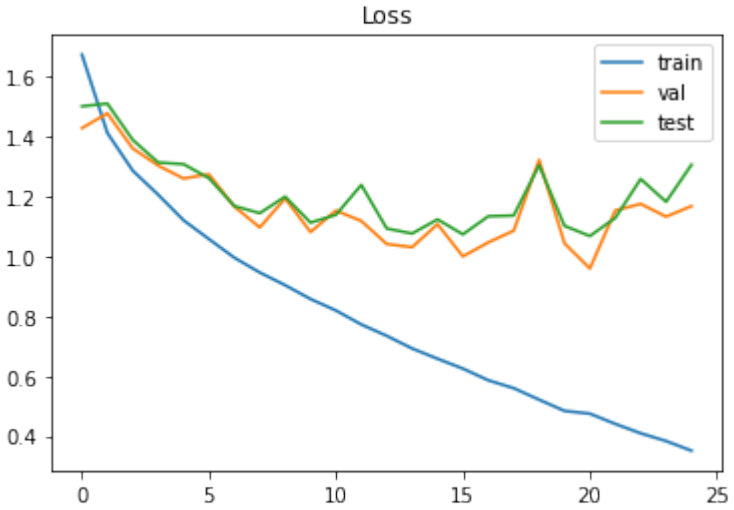

Figure 4: Loss

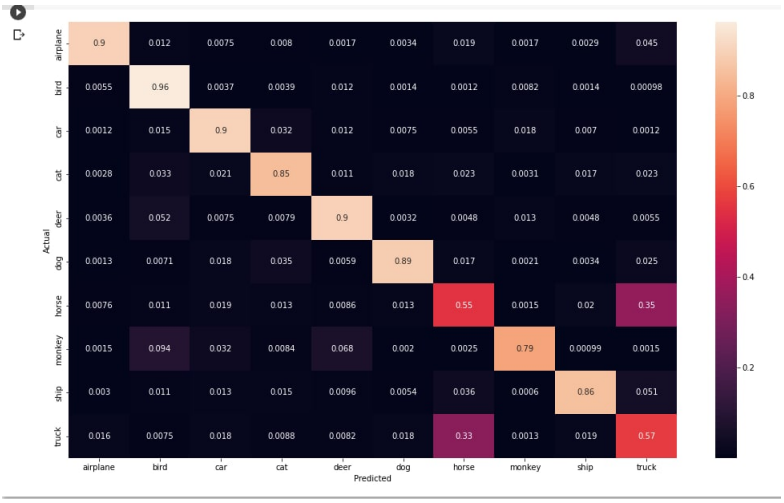

Figure 5: Confusion Matrix

- **Bird** 96 percent of accuracy
- **Car** 90 percent of accuracy
- **Cat** 85 percent of accuracy
- **Deer** 90 percent of accuracy
- **Dog** 89 percent of accuracy
- **Horse** 55 percent of accuracy
- **Monkey** 79 percent of accuracy
- **Ship** 86 percent of accuracy
- **Truck** 57 percent of accuracy

5 Ablation Studies

Several experiments were conducted utilizing incremental number of layers in order to assess the evolution of the accuracy of the model when adding an additional layer. The following results were obtained:

- **1st Conv layer + FC + Classifier:** Test Accuracy= 0.40
- **1st + 2nd Conv layers + FC + Classifier:** Test Accuracy= 0.44
- **1st + 2nd + 3rd Conv layers + FC + Classifier:** Test Accuracy= 0.53
- **1st + 2nd + 3rd + 4th Conv layers + FC + Classifier:** Test Accuracy= 0.60