Machine Learning – Homework 2

Andrea Gasparini
1813486

December 2020

Contents

1   Introduction

The goal of this Homework was to solve an image classification problem with objects typically available in a home environment.

- *accent plate*
- *breadsticks*
- *cereals box*
- *cocoa drink bottle*

- *nitrile gloves*
- *paper bag*
- *pears*
- *plastic fork*

2   Dataset

The dataset contains **8221** samples and it's based on a 8 classes subset of the RoboCup@Home-Objects dataset. For each class we have a folder containing all the related samples.
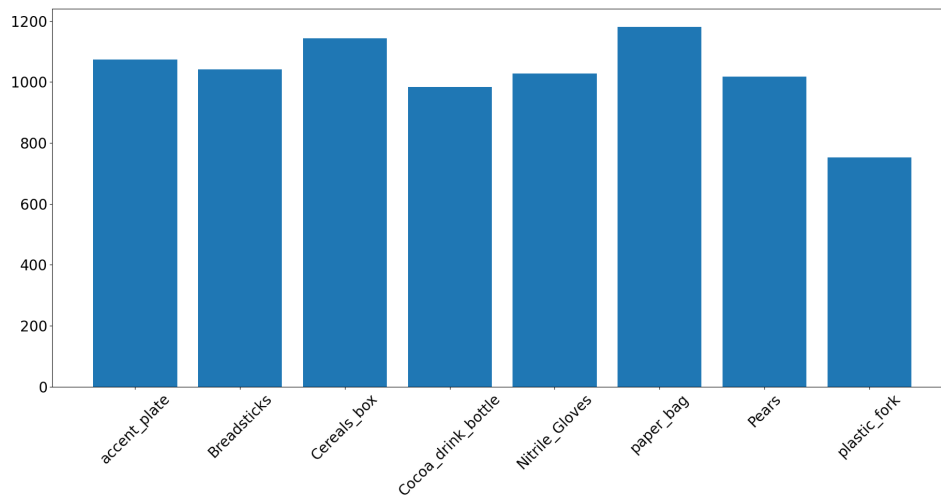

Figure 1: Distribution of the dataset's 8 classes

Figure 2: 10 random samples from the dataset

3 Preprocessing

The first thing I did is the dataset splitting in two subsets for training and validation, respectively 70% and 30% of the original one.

3.1 Data augmentation

Data augmentation includes techniques used to generate *new* training samples from the original ones by applying random jitters and transformations, but at the same time ensuring that the class labels of the data are not

changed. These kind of techniques are useful to reduce overfitting and they usually provide an increase in testing accuracy, at the expense of a slight worsening in training accuracy. In our case I obtained an increase in the performance by 5 to 10% applying flipping, rotations and zooming.