MVA 2019: Computer Vision and Object Recognition

Amin DHAOU
amin.dhaou@ens-paris-saclay.fr

November 26, 2019

Abstract

In this paper, we proposed different approaches using Convolutional Neural Network (CNN) in order to solve a bird species classification's problems. The data used to train and test our model is a subset of the Caltech-UCSD Birds-200-2011 dataset. We first computed and trained a CNN from scratch and then we performed Transfert Learning. Several techniques such as data augmentation, ensembles and fine tuning parameters were used to improve our techniques.

1 Introduction

During the last few years, CNNs has been really powerful in tasks such as image recognition and image classification. That's why we chosed to perform different The Dataset we worked on is a subset of the Caltech-UCSD Birds-200-2011 dataset contening 20 classes of bird species out of the 200 on the original dataset. The data was split into 1083 training images and 103 validation images

2 Approach overview

2.1 Preprocessing

The images provided has different shapes and the birds in each image is more or less exposed. The bird could be in the background of the image behind leaves or in the foreground. The CNN should first detect the bird and then determine it's species. Because of the small amount of data, we had to perform data augmentation which computes transformation on the original data set such as random rotation, random flip and cropping, and thus increase the number of training and validation data to learn more.

2.2 Models

Our first model was computed from scratch and didn't perform well. It wasn't complex enough to get an accuracy higher than 50 % on the validation dataset. Since the data set is small and our CNN didn't worked, we used pretrained models on the "Imagenet" Data Set (which contain a bird category). We used VGG-16, Densenet161, Resnet152 [2] and trained their last few layers which are the classifications ones in order to fit to our problem. We also tried ensembles, combining resnet152 and resnet101 to improve our score. Finally, We used Inception-v3 pretrained on another Data set called Inaturalist[2]. We also wanted to use bird detector with YOLO but we didn't get time to compute it.

2.3 Fine tuning

To increase our score, we tried several techniques such as learning rate decay, weight decay and dropout in order to reduce over-fitting and modifying the batch size.

3 Results

With the pre-trained models VGG16, Densenet161 and Resnet152 we got scores of 88, 89 and 91 % on the validation dataset. Finally, with Inception v3 trained on Inaturalist, we got 0.83 %.

References

[1] https://pytorch.org/docs/stable/torchvision/models.html
[2] https://github.com/richardaecn/cvpr18-inaturalist-transfer