Assignment 3: Bird Image Classification

Denys Sikorskyi

ENS - Master M2 MVA

denys.sikorskyi@ens-paris-saclay.fr

Abstract

During the competition the goal was to create a model to distinguish 20 types of birds. The training was performed on a subset of the Caltech-UCSD Birds-200-2011

1. Introduction

The modern approach to the Image Classification tasks are Deep CNN. They have outperformed previous generation algorithms by huge margin. For this task, we will use modern models that were trained on the ImageNet dataset. The training will occur with transfer learning.

2. Data preprocessing

2.1 Dataset Description

The dataset consists of 1082 training images and 103 validation images. 517 images are left for getting accuracy score. There 20 types of the birds, meaning 20 different folders for each type. The test set consists only from one folder with mystery category which we have divide into 20 different types.

2.2 Data preprocessing

There is some imbalance in distribution of types, so we have changed some photos from val_images to train_images and vice versa. We will resize all images to (299, 299) as the highest number of pixels that one of the future models requires. Next, we will random rotation of 45 degrees in order to fight problem of relatively small dataset to prevent overfitting of the models. For the same reason, we do random horizontal flip with probability of 0.6.

3. Model creation

During the modeling we used several models, starting with relatively simple and small and finishing with complex ones. Namely, ResNet-18[1], InceptionV3[2], DenseNet201[3] and EfficientNetV2[4]. We have used batch size of 64 and different optimizers. Sometimes the SGD optimizer was used, sometimes RMSprop. For each we have frozen transferred layers and build final layers during the modeling processing using accuracy as a hint. Also, we have implemented ReduceLROnPlateau - each time we haven't increased validation accuracy for 5 times straight we have decreased the learning rate to achieve higher accuracy.

3. Results

The best result was given by DenseNet201 with 87% accuracy on validation and 80.65% on the 517 testing images. Other models showed worse result. The closest to the winner was EfficientNetV2 with just 80% on validation, but 68% on testing images.

Model	Val_accuracy	Test_accuracy
ResNet-18	76%	63.23%
InceptionV3	76%	63.87%
DenseNet201	85%	80.65%
EfficientNetV2	75%	68.38%

4. Conclusion

The highest accuracy that we have achieved is 80.65% on the test images. This score can be increase by increasing number of images in the training and validation datasets, because some of the models have shown much better performance at similar task, but with much bigger datasets. Also, due to small and uneven distribution of types, we can see such difference between validation and test accuracies.

References

[1] ResNet

[2] InceptionV3

[3] DenseNet201

[4] EfficientNetV2