Birds Classification Challenge

Rguig Mohamed Amine
Université Paris-Dauphine

November 29, 2022

Abstract

The CUB-200-2011 bird dataset is a challenging dataset containing 200 different species of birds. We present a method for detection and classification for a subset of the dataset, developed for a Kaggle competition organized for the 2022 Object Recognition class of the ENS Mathematique, Vision, Apprentissage Master. The goal was to obtain the highest possible score on a test dataset to classify 20 different bird species.

1 Introduction

The bird dataset given present many challenges, as the birds are not centered in the pictures, and are often in a an envirement such as on brach of trees or on grass, evens on hands. So in order to preprocess this data, we had to crop the images using FasterRCNN pretrained on the YOLO dataset, after that we proceeded to augment the dataset via the Library IMGAUG, and finetune our model which is a stacked model of Resnet152 and InceptionV3 and finally for training, inspired by the paper, we train our model first on the cropped and augmented images and then on the original images.

1.1 Preprocessing

We used FasterRCNN Resnet50 and FasterRCNN Mobilenetv3 from Pytorch in order to detect birds, we proceeded then to crop the bounding boxes of which of these two models gave the best probability and above 0.9 , in order to have a perfect crop of the bird, After that in order to augment the cropped images, we used non-orientation augmentation in this file such a gaussian bluring and much more. we decided to only augment by 2 images the classes of which the model didn't perform well.

1.2 Training

For the training, we fine-tuned two existing networks that were trained on the ImageNet dataset, in order to avoid training the whole network from scratch on our small dataset due to overfitting concerns. We froze the weights of the first few layers that extract universal features and dropped the softmax layer of the two networks. We rain the last three bottlenecks of the ResNet152 and the last nception module of the InceptionV3 network. We then proceed to train our model on the cropped-augmented datasets, catch the best model at some small epoch, and then retrain our model on original images in order to take into count of the context of which birds are.

1.3 Evaluation

The model performed better on some classes than other classes on some epochs, so we decided to change the evaluation file for it to consider two checkpoints and compare their performance on each classes to generate the kaggle file .

Conclusion

Finally, before evaluation we catch two checkpoints that performed 96/103 on the validation set, and we generate the kaggle file combining them

References

[1] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. CoRR, abs/1512.00567, 2015.

[2] K. Akash, Das, Sourya Dipta. Bird Species Classification Using Transfer Learning with Multistage Training.Springer Singapore

[3] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. CoRR, abs/1512.00567, 2015