Assignment 3: Image classification

Marine Mercier
marine.mer98@gmail.com

1. Dataset

The dataset contains 1702 images, including 517 images for testing. 20 classes are considered. The size of the images varies a lot, and sometimes the bird occupies only a small portion of the image.

1.1. Preprocessing

Resizing the image is a necessary preprocessing step in order to use batch learning. I noticed that larger images gave better results, but the training and computation is slower. I decided to make all images square with size (375x375). I tried different techniques while preserving the aspect of the images. To begin with, I applied a padding and a Center or Random crop on images. Then, I decided to use a BirdCropper to preserve only the bird on the image. For this purpose, I used a pretrained Resnet on the COCO dataset, and use the given bounding box to crop the image around the bird.

Since I am using models pretrained on ImageNet, I could normalize data with the same normalization as the ImageNet dataset. I also tried to normalize data with the statistics of the dataset, but the output was not as good.

1.2. Data Augmentation

I used classic data augmentation to improve my results, like random horizontal flip or random rotation. I also tried adding Gaussian Noise, but it wasn't effective.

2. Approaches

2.1. Convolutional Neural Networks Models

Transfert learning enables to benefit from the knowledge gained from previous trainings. Therefore I used pretrained model and added a fully connected (FC) layers and dropout to avoid overfitting. I tested three different ResNet architectures : Resnet50, Resnet101 and Resnet152. Resnet50 overfits less than bigger architectures, but I got the best results with heavier models thanks to early-stopping, a technique to limit overfitting. I trained the models in three different ways : with all layers unfrozen, with all layers frozen except the FC layers, with most layers frozen except the FC layers and two deep layers which gave the better results.

I used Adam optimizer and tested different learning rate schedulers, OneCycleLR gave the better results. I also used weight decay and gradient clipping to avoid exploding gradients.

2.2. Cross Validation and Bagging

Cross validation allows to test the stability of the model. I cut the available data in four training set and trained the same model four times using a different validation datasets each time. The four models are then combined by voting.

I also combined different models and did bagging using the voting rule. Each model is trained using the previous cross validation technique, and they are combined to give the final output. But Resnet101 got better results alone than bagged with different architectures.

2.3. MixCut augmentation

The CutMix augmentation strategy cut and paste patches among training images, the ground truth labels are also mixed proportionally to the area of the patches. It replaces regional dropout while working as a regularisation for the model.

3. Results

My best result was obtained with a bagging of four models given by the cross validation on Resnet101 with BirdCropping applied. Surprisingly, Models trained with BirdCrop and MixCut don't perform better than the ones trained with just BirdCrop.

3.1. Possible improvements

Using iNaturalist thanks to semi supervised learning would have been useful since it is close to our dataset.