Object recognition and Computer Vision 2019/2020
Assignment 3: Image classification

Khaoula Belahsen
ENS Paris-Saclay
ENSTA Paris
khaoula.belahsen@ens-paris-saclay.fr

Abstract

This presents a summary of the work done on Assignment 3 from the RecVis course at the MVA master. It shows the ideas implemented to achieve the high accuracy fot the Caltech-UCSD Birds-200-2011 image classification.

1. Introduction

Transfer learning [1] is a well-known method responsible for the growth of deep learning. It revolves around the idea of using the information learned by a trained model for another model which needs to be trained on the different data for a similar task. It is very much used in object recognition, where it uses pre-trained models on larger datasets, for example ImageNet. Since the images available for our challenge overlap with the ones in ImageNet, I used corresponding pre-trained model.

2. Approach

Instead of building models from scratch, I relied on tranfer learning. The datasets we had were composed of 20 classes, instead of the 1000 processed by the ImageNet pre-trained models. I used some approaches to improve my learning : data augmentation and freezing/unfreezing of layers from the pretrained models.

2.1. Data augmentation

The data-sets for the training and validation contained roughly a total of 1190 images divided unevenly across the 20 classes. With some classes being clearly under represented. In a goal of preventing over fitting and improving the generalisation capacity of my model, I played with different data augmentation, introducing different ones for the train and validation sets using Horizontal flips, crops and ColorJitters.

2.2. Pre-trained models

I looked at deep neural networks pre-trained on ImageNet VGG, ResNet, which have already learned features like curves and edges in there early layers, that are relevant and useful to most of the classification problems, and especially ours.

So I opted for keeping the first layers intact and unfreezing the last layers in a progressive manner towards the top. I tried several VGGs 11 and 16 and moved to deeper ResNets.

2.3. Chosen model

I was able to obtain my best accuracy on the test set at value of 0.8 using the ResNet101. At first, I only changed the fully connected layers to adapt it the the 20 classes we have instead of 1000 from the ImageNet. I progressively unfroze the layers on top. I tried to improve the Optimizer for the back propagation on our smaller dataset by adding a momentum of 0.9 and a scheduler for the learning rate.

3. Improvements

By lack of time, a few elements could not be implemented but could still improve my score :

- Look at assembling the prediction from different models
- Analyse false positive and true negatives
- work on the classes that have few to one representatives, and making the neural network see them more often at each epoch in order to tackle the unbalance.

References

[1] Jason Brownlee. A gentle introduction to transfer learning for deep learning, 2017.