Assignment 3: Bird image classification competition

Henrique Gasparini Fiuza do Nascimento
Ecole Normale Superieure Paris-Saclay
61 Avenue du President Wilson, 94230 Cachan
henrique.gasparini_fiuza_do_nascimento@ens-paris-saclay.fr

Abstract

In this work, we present our approach and results in a image classification challenge on a reduced version of Caltech-UCSD Birds-200-2011[3] with 20 classes of birds. We explored techniques such as data augmentation, hyperparameter tuning, and transfer learning to obtain interesting accuracies in a Kaggle competition.

1. Introduction

The success of deep neural networks for image classification tasks has been first demonstrated on Imagenet [2] and later for problems with small training datasets, using various transfer learning approaches[4].

A very common approach is to pick a state-of-the-art architecture that solves a "large problem" and use the weights of the model trained on the large dataset to initialize the training on the small dataset. Another way of seeing transfer learning is to perceive the outputs from deep learning models (as well as from intermediate layers) as features.

Finally, augmenting image data using affine transformations and optimizing hyperparameters can also be useful.

2. Related work

Image classification with small data is a quite common problem and has been studied in many papers and tutorials[1].

3. Method

We used two distinct types of architectures. In the class *PretrainedNet*, we allowed several architectures with good performance on ImageNet, in particular Residual Networks, and Inception Networks. One can choose how many of their layers will be frozen so that the remaining ones are fine-tuned to the small training dataset.

In the class *ShallowNet*, the frozen pretrained model (ResNet18) is used to compute features which are concatenated to one of the final layers of a shallow convolutional model, being followed by a fully-connected layer.

4. Implementations and performances

Until the deadline, the best results were obtained for shallow models that do not use pretrained models to construct features. The most effective techniques were largely augmenting the dataset (to get 5000 examples by class) - 25.2,% test accuracy - and reducing the learning rate by half - 23.2% test accuracy.

None of the implemented approaches that used pretrained deep neural networks produced interesting results, as all test accuracies were below 10%.

5. Conclusions

In the limited time of this work, transfer learning has not proved to be useful in the experiments, while some more basic techniques such as data augmentation and hyperparameter tuning were important.

We still have reason to believe that a correct use of transfer learning could lead to much better results (e.g. from looking at the competition dashboard).

Could we spend more time working on this problem, the next steps would most likely be training the same models on the largely augmented dataset (which was not done due to time constraints) and exploring other architectures.

References

[1] P. Jay. Almost any image classification problem using pytorch, 2017.

[2] A. Krizhevsky, I. Sutskever, and G. E. I. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012.

[3] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The Caltech-UCSD Birds-200-2011 Dataset. Technical report, 2011.

[4] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson. How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, 2014.