Mushroom classification with VGG-16

Lorenzo Vainigli
lorenzo.vainigli@studio.unibo.it
matr. 0000842756

Course of Machine Learning
Laurea Magistrale in Informatica
University of Bologna
A.Y. 2020-2021

Abstract. ...

1 Introduction

The purpose of this project is to study and build a classifier that is able to recognize images of mushorooms and categorieze them.
The model was trained on the database from [1], that contains about 100,000 images and 1,500 categories of mushorooms. VGG-16 was proposed by [2] an it is a powerful covolutional neural network able to reach an accuracy of 92.7% on the ImageNet database [3].

2 Methods

In this section we describe the process of building this project. We begin with an analysis of the dataset to take a look at the data that we will use for train, validation and test.
Then we describe the steps followed to obtain the model for our purpose. Finally we describe the phases of train, validation and test.

2.1 Software

- Google Colab and Google GPUs.
- Pandas and Tensorflow.

2.2 Dataset analysis

The image dataset is composed by an hierarchical structure of folders named with the category name of the images that they contains. Precisely each folder name has the format `ID_super-category_category`, so we use this information to associate these two properties to the images.
There are some annotation files but we ignore it, since our purpose is to associate each image of mushroom to its category and no more.

2.3 Transfer learning from VGG-16

Even if VGG-16 has achieved very good results, its architecture is very complex and it has a lot of parameters that needs to be tuned during training. This leads to a very expensive work if someone want to train VGG-16 from scratch. Instead, we want to use the pre-trained model, at least for convolutional layers, while we want to train the deep layers at the end of the architecture of VGG-16.

2.4 Image processing

- `ImageDataGenerator`
- Image augmentation

2.5 Train, validation and test

- Description of the training with different number of classes and different number of samples per class.
- Time spent on training.

3 Results

– Description of different scores (accuracy, precision, recall, f1-score) resulted from the training with different number of classes and different number of samples per class.

Classes	Samples per class	MobileNetV2		Xception	
		Top-1	Top-5	Top-1	Top-5
3	414	89%	-	89%	-
10	340	77%	98%		
20	255	65%	95%		

Classes	Samples per class	MobileNetV2			Xception		
		Precision	Recall	F1-score	Precision	Recall	F1-score
3	414	89%	89%	89%	90%	89%	89%
10	340	78%	77%	77%			
20	255	66%	65%	65%			

4 Discussion

– Analysis and comparison of results.

5 Conclusions

Further directions

References

[1] The Fifth Workshop on Fine-Grained Visual Categorization. *Fungi Classification Challenge*. 2018. URL: https://sites.google.com/view/fgvc5/competitions/fgvcx/fungi.

[2] Karen Simonyan and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition". In: *arXiv 1409.1556* (Sept. 2014).

[3] Jia Deng et al. "ImageNet: a Large-Scale Hierarchical Image Database". In: June 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.