

# Selected Topics in Visual Recognition using Deep Learning. Homework 3

Kakhaev Ivan, 0860838

November 2020

## 1 Introduction

For this assignment given next dataset: Tiny VOC dataset containing 1,349 training images and 100 test images with 20 object classes. On this amount of data is inevitable to overfit the model. All files and readme, that describes file structure and steps for reproducing the results, available on GitHub<sup>1</sup>.

## 2 Methodology

### 2.1 Data

As you can see from Figure 1 we have heavy class imbalance, where class 15 is person. Horizontal flip will be used as augmentation.

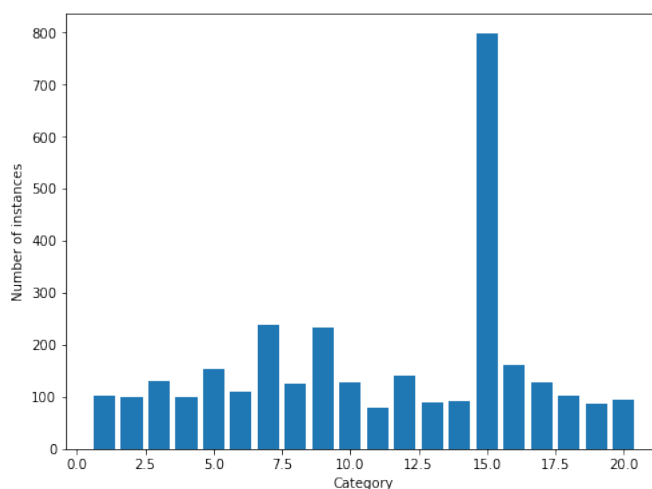


Figure 1: Classes distribution

---

<sup>1</sup><https://github.com/veax-void/TinyPascalSegmentation>

## 2.2 Model architecture

For this project I will go with Mask R-CNN model provided by torchvision library.

## 2.3 Hyperparameters

For optimization of parameters I will use SGD with parameters: lr=0.005, momentum=0.9, weight\_decay=0.0005. Also I used StepLR as learning rate scheduler with parameters: step\_size=3, gamma=0.1.

## 3 Conclusion

This project was tough because of dataset preparation steps. I tried different models but they all have different requirements for target data structure. And find it take some time to get and force to work coco toolkit, but it was interesting. At the end I get mAP  $\approx 0.33$ , is not the best score but I learned a lot along the way.

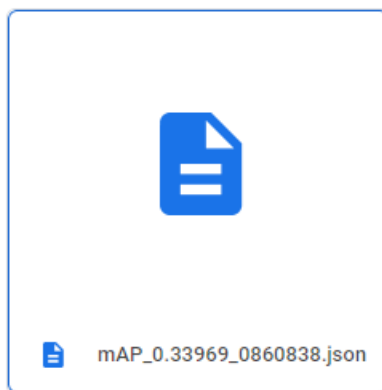


Figure 2: Screenshot of the score



Figure 3: Results