

# 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 (class 15 is a person). Horizontal flip and random crop will be used as augmentation to prevent overfitting.

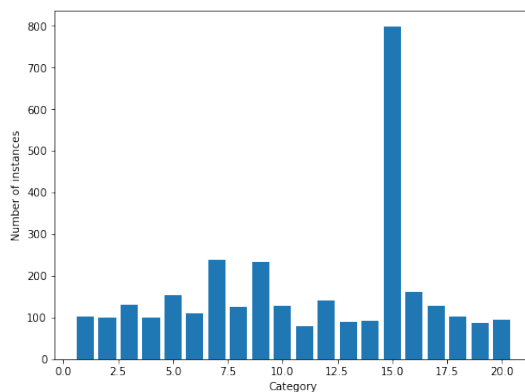


Figure 1: Classes distribution

---

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

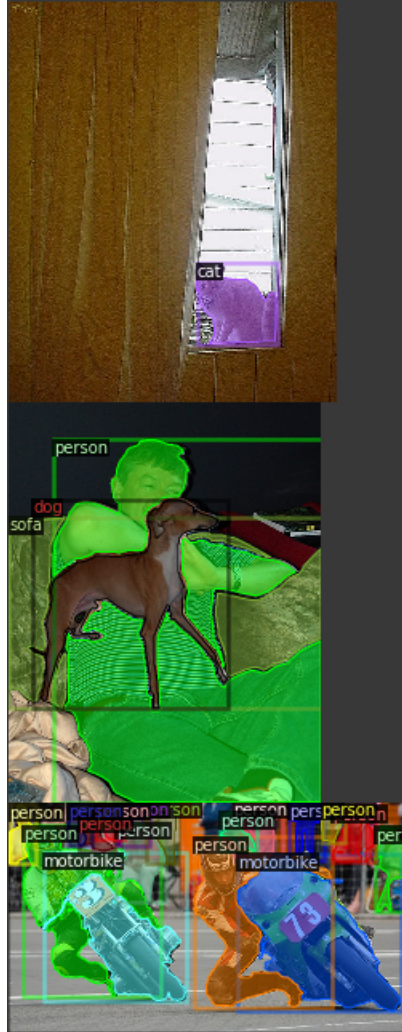


Figure 2: Sample from training dataset

## 2.2 Model architecture

For the model I will use Mask-RCNN with ResNet50 as the backbone. Detectron2 will be used as a training and evaluation framework.

## 2.3 Hyperparameters

For base learning rate was used  $25e-5$  (*BASE\_LR*). For the maximum amount of iterations was used 4000 (*MAX\_ITER*).

## 2.4 Training pipeline

Before training read Readme.md that contained required dependencies and install them.

For the training just run 'python train.py', model will be stored in ".outputs" folder.

## 2.5 Inference

For your data prediction run 'python test.py' and "submission.json" will be stored in ".outputs" folder.

## 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  $\approx 0.31$ , is not the best score but I learned a lot along the way.



Figure 3: Screenshot of the score

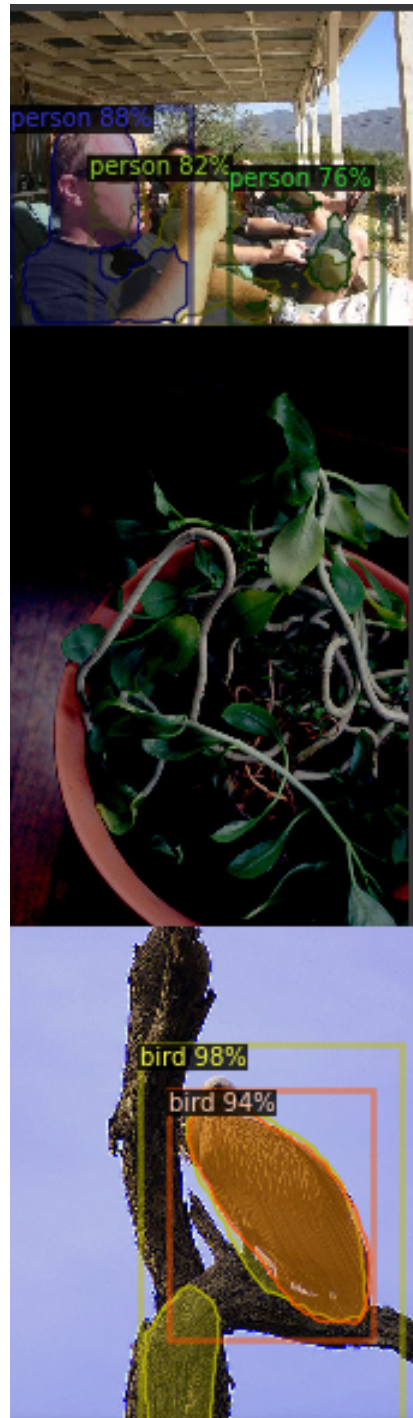


Figure 4: Results