Assignment #2

Image Based Biometry 2022/23
Faculty of Computer and Information Science
University of Ljubljana

Basic Detection

I. INTRODUCTION

Your task is to set up two popular detection methods: Haar-cascades [1] and YOLO (v5) [2], [3], and evaluate them on an ear detection task using existing weights – this means you do not need to utilize GPUs for this task.

II. MATERIALS

- **Literature**
  - The original paper is a good start: https://www.merl.com/publications/docs/TR2004-043.pdf [1].
  - YOLO (v4) paper is available here: https://arxiv.org/pdf/2004.10934v1 [3] to get a general idea of how YOLO works.
- **Code**
  - Haar-cascades: https://docs.opencv.org/3.4/d2/d99/tutorial_js_face_detection.html (Essentially it is a three-liner – loading up an image, an XML and calling a prediction).
  - YOLO v5 for which the supplied weights are compatible: https://github.com/ultralytics/yolov5.
  - Please follow the PEP8 guidelines (https://peps.python.org/pep-0008/) and comment a lot.
- **Test data** Use the whole data available below (1000 samples of 100 classes).
- **Report** Please, begin with the template available here: https://www.overleaf.com/read/ppbyvnbqfyrt and limit yourself to two pages.
- **Support Files** (data, trained model and trained cascades) are available here: https://tinyurl.com/ibba2

III. THINGS YOU NEED TO DO

- Set up Viola-Jones/Haar cascade detector, load up XMLs and images supplied here and predict detection on ear images. Based on the supplied ground-truths compute accuracy.
- Set up YOLO v5, load up weights supplied here, run predictions and compute accuracy compared to ground-truths (be careful to use the correct format of annotations).
- Fine-tune VJ parameters to achieve as high results as possible.
- Report IoU, precision-recall scores over all thresholds (step of 0.01) and report appropriate qualitative results.
- Submit the code (without the data) and the report in PDF.

IV. QUESTIONS

To help you with the report and programming goals, try addressing the points below in the report.

- **4 pts** VJ predictions.
- **3 pts** Optimization of VJ parameters.
- **5 pts** YOLO predictions.
- Quantitative analysis:
  - **2 pts** Report IoU for different VJ parameters and the YOLO "baseline".
  - **4 pts** Report precision-recall over all thresholds for different VJ parameters and the YOLO "baseline".
- **2 pts** Qualitative analysis
  - Show a set of 10 representative failed VJ predictions and try to explain them.
  - Show a set of 5 representative best VJ predictions and compare them to YOLO predictions.

V. SUBMISSION

There is one deadline:

- November 11: submission of your VJ and YOLO evaluation code (without the supporting YOLO scripts, only your scripts) and a two-page report (PDF) on Eučilnica: https://tinyurl.com/ibba2submit.

REFERENCES

[1] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in *Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1. IEEE, 2001, pp. I–I.

[2] "Ultralytics – YOLOv5," https://github.com/ultralytics/yolov5, accessed: 2022-10-25.

[3] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.