UNIVERSITÀ DEGLI STUDI DI PADOVA | DIPARTIMENTO MATEMATICA

University of Padua
Department of Mathematics "Tullio Levi-Civita"
Master's Degree in Computer Science

Project of Vision Cognitive and Services: Instance Segmentation - SOLOv2 and Deep Snake

Eleonora Signor	eleonora.signor@studenti.unipd.it
Francesco Bari	francesco.bari.2@studenti.unipd.it

Proposal - January 2022

Project content

The goal of our project is to study and compare alternative techniques to Mask R-CNN, for solving instance segmentation tasks.
First we will study Mask R-CNN, to understand its strategies and potentials, and then we will focus on the comparison of two alternative approaches: SOLOv2 and Deep Snake. That both seem to have better performances than Mask R-CNN. Our activity will be divided into 3 steps:

1. Study of the papers (Mask R-CNN, SOLOv2 and DeepSnake), with identification of the differences in approach;
2. Fine-tuning of the models. We will use the models present in:
   - https://github.com/zju3dv/snake/;
   - https://github.com/aim-uofa/AdelaiDet/.
3. Evaluation of the models and subsequent report writing.

The datasets we will use to carry out phase (2) will be:

- Cityescape, for the evaluation of urban scenes, in order to assess systemicity;
- WildDash, contains difficult images and counter-productive camera characteristics, which should allow us to make some considerations about the robustness of the models.

We have chosen to deal with the Urban Street Scenes as a sub-context.

For phase (3) we plan to carry out a quantitative assessment using at least the IoU and AP indices. We use all datasets for evalutation (with even MS COCO for test the general behaviour of the models), and fine-tuning only those related to the Urban Street Scenes sub-context.
We will also investigate the possibility of changing the backbone and head architectures in SOLOv2 and Circular convolution in Deep Snake.
Any consideration we make will always take into account the official results obtained by Mask R-CNN.