Project:

Student Project Proposal (suitable as semester or Master project)
Supervisors: Michael Villamizar & Dr. Jean-Marc Odobez
Place: Idiap Research Institute, Martigny (part-time work at EPFL possible)
Keywords: Data annotation, panoptic segmentation, neural networks

Project description

This semester/master project takes place within the framework of the Biped project, funded by the Swiss Innovation Agency and involving the Biped-AI company and the Idiap Research Institute. The aim of the project is to design and develop novel computer vision and machine learning algorithms to improve the assisting system commercialized by Biped-AI which is used by people with reduced visibility to understand their environment and walk safely and navigate through it. More specifically, the algorithms and softwares similar to those used in autonomous vehicles are exploited to analyse the video and depth streams of three wide-angle cameras mounted on a shoulder-worn harness worn by the users. They can filter risks based on trajectories, and warn the user in real-time with short sounds played in bone-conduction headphones, see Figure 2.

Figure 2: Biped project

One important part of the Biped project consists of collecting video data from users walking in indoor and outdoor environments which will be further annotated and later be used to train neural networks aiming at understanding the environment, esp. for detecting obstacles, people, objects (cars, bicycles, benches, etc), as well as recognizing the main scene elements like the ground or walls, all this in order to identify a safe pathway.

Hence, to train these models, the collected data has to be annotated with both semantic information and an instance label, i.e. more precisely, each image pixel is labeled with one semantic class (e.g., pedestrians) and one instance whenever appropriate (e.g., person#1), see example shown in Figure 3. Note that being able to infer this information from one image is known as panoptic segmentation.

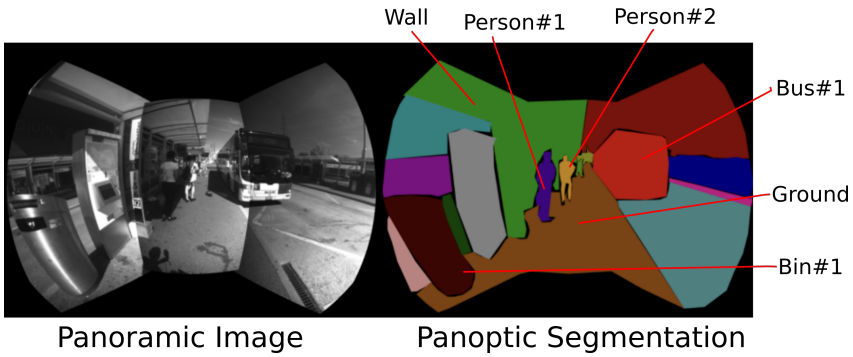

Figure 3: Left: panoramic image built from the three camera images. Right: panoptic segmentation of the image on the left.

However, performing data annotation is a difficult and time-consuming task, especially if the amount of collected data is very large. To reduce human labor and annotation times, one strategy is to rely on pre-trained neural network models to create pseudo-labels for our target dataset by running these models on images. As a result, the image of video data are pre-annotated automatically with "pseudo labels", and the human labor is mainly focused on validating these labels and refining them when they are obviously wrong.

ANoter strategy is to use segmentation models like the Segment Anything Model [1] which, using a few points, can generate a highly accurate segmentation the object of interest.

Main tasks and goals

In this project, the main goal is to build tools to help do the annotation at scale of the data collected within the project, following the above approach. Working in close collaboration with a research scientist on the project as well as with the partner company, you will:

- Study the panoptic segmentation principle and the main deep networks in the state of the art doing this task such as EfficientPS [2], Mask2Former [3], as well as considering or Segment Anything Model (SAM) [1] to be acquainted with the models;
- Apply these models on a small set of relevant images and evaluate their suitability to annotate relevant items (obstacles, people, cars, ground, and other elements);
- Define an annotation protocol and develop a graphical interface using existing tools for labelling project images, e.g. by assigning or refining annotations suggested by the pre-trained networks;
- Apply the protocol and tool to a dataset of project images.

The project is suitable for one or several semester project students or/and a master project.

Practical information

Prerequisites: Good command of Python, basics of Linux, deep learning background, basics of computer vision.

Tools you will use: Pytorch, Git, OpenCV

Dates: Available immediately.

Contacts: If you are interested or you have any questions, please contact **Jean-Marc Odobez** (odobez@idiap.ch).

References

[1] SAM: segment anything model. https://segment-anything.com/.

[2] EfficientPS. http://panoptic.cs.uni-freiburg.de/.

[3] Mask2Former. https://github.com/facebookresearch/Mask2Former.