valeo.ai

4 open internships at valeo.ai in 2023

Valeo.ai

We are a research lab in the center of Paris conducting long-term research on machine learning and computer vision for automotive applications. We tackle a variety of research problems like scene understanding (from images, 3D point clouds or radar data), frugal learning (including self-supervision, few-shot, zero-shot and unsupervised methods), domain adaptation, reliability, explainability, motion forecasting, etc. More information about the team can be found here.

Internships at valeo.ai

We regularly hire as interns a number of MSc students in their final year, to work on research topics. Most of these internships result in paper submissions to top-tier conferences such as CVPR, ECCV or ICCV. Some trainees go on to do a PhD thesis in the lab. Practical information:

- **Location:** valeo.ai, 100 rue de Courcelles, 75017 Paris.
- **Time:** usually spring and summer (typically 6 months).
- **Applicants:** finishing their MSc with a solid background in computer vision and machine learning, particularly in deep learning with strong PyTorch coding skills.

How to apply?

If you wish to apply to one or more internships, send a mail to the above contact people with:

- a cover letter explaining your interest and adequacy for the topic,
- your CV/resume,
- transcripts of your grades from last year, as well as this year when already available.

4 Topics in 2023

1. **Exploiting pre-trained diffusion models for autonomous driving**
    - *Contact: Mickael CHEN and Tuan-Hung VU*

    Abstract: Diffusion models are showing very impressive results and have quickly become ubiquitous in image generation and edition. In particular, models pre-trained on very large datasets embed general knowledge that can be useful in the synthesis of corner cases. Indeed, by bringing unexpected elements into standard driving scenes, we can simulate rare events that would be not observable otherwise. This internship aims to design and validate a method that builds and exploits such synthetic data to improve the training and evaluation of vision models used in autonomous driving pipelines.

2. **Exploiting foundation models for autonomous driving**
    - *Contact: Eloi ZABLOCKI and Oriane SIMEONI*

    Abstract: *The use of foundation models, e.g., CLIP, Flamingo, CoCa, for downstream applications lead to impressive performances. Indeed, learned textual and visual representations contain fine and precise semantic knowledge about concepts at hand which make it now possible to use them in reasoning tasks. Leveraging and adapting*

this knowledge to new domains, e.g., autonomous vehicles, without retraining models from scratch while retaining learnt reasoning capacities is the key question that will be asked during the internship. In particular, the internship will focus on the following questions: 1. How to embed specific driving concepts (e.g., objects, traffic signs, corner case situations) into foundation models? The starting point is to build on textual inversion that learns new token embeddings in the textual space of CLIP given a few images of the concept of interest. 2. How to use these learned driving specific concepts for downstream applications. They may include retrieval, self-supervised semantic segmentation, continual learning, data augmentation, scene edition. 3. How to generalize from static image concepts to video concepts?

3. **Self-supervision for camera-based bird's-eye-view perception in autonomous driving**
    - *Contact: Spyros GIDARIS, Andrei BURSUC* and *Alexandre BOULCH*

Abstract: To plan and drive safely, autonomous cars need accurate 360-degree perception and understanding of their surroundings from multiple and diverse sensors, e.g., cameras, RADARs, and LiDARs. The bird's-eye-view's (BEV) representational space, a.k.a. top-view occupancy grid, recently gained considerable interest within the community with the release of multiple real-world autonomous driving datasets (Waymo Open, nuScenes, ONCE). BEV appears as a suitable and natural space to fuse views from multiple cameras mounted on a vehicle (Philion et al., 2020; Bartoccioni et al., 2022) and to capture semantic, geometric, and dynamic information. However the acquisition and annotation costs for such sensor rigs is significant. Self-supervised representation learning (SSL) (Gidaris et al., 2021; Weng et al., 2019) arises as a promising approach for annotation efficient learning. In SSL, a deep neural network is first pre-trained on a large collection of unlabeled data and then subsequently quickly fine-tuned for a task of interest with few available annotated samples. In this internship we aim to design self-supervised strategies for BEV perception from multiple cameras by leveraging different supervisory signals (Lidar, temporal information) from raw sequences of driving data.

4. **Knowledge distillation between lidar sensors**
    - *Contact: Gilles PUY and David HURYCH*

Abstract: In order to achieve high performance, today's neural networks for point cloud semantic segmentation or 3D object detection require large annotated datasets. Several datasets are available to train these networks: KITTI, nuScenes, Waymo, Once, etc. Yet any network trained on one dataset performs poorly on the others. One of the reasons is that these datasets are acquired with different lidars and networks do not generalize well across lidars. Creating and annotating a dataset for each new sensor is not a scalable option. Instead, in this internship, we want to leverage networks pre-trained independently on each of these datasets and study how to transfer their knowledge to a new network with an increased performance on a new datasets with different lidar setup.