Antonin VOBECKY

+420 739 003 777 ◇ a.vobecky@gmail.com ◇ Prague, Czech Republic
vobecant.github.io ◇ github.com/vobecant ◇ linkedin.com/in/antoninvobecky ◇ Google Scholar

PhD researcher in computer vision specializing in multimodal models and weakly supervised learning.

EDUCATION

2019 – 2025 (*thesis submitted*) **Doctoral Degree (PhD)** in COMPUTER VISION
Faculty of Electrical Engineering, Czech Technical University
ELLIS PhD student
Advisor Josef Šivic (CTU), co-advisors: Patrick Pérez (Kyutai), David Hurych (valeo.ai)

2014 – 2019 Bachelor's (BSc) & **Master's Degree (MSc)** in ARTIFICIAL INTELLIGENCE AND COMPUTER VISION
Faculty of Electrical Engineering, Czech Technical University
Summa cum laude. Merit scholarships.

SELECTED PUBLICATIONS & PROJECTS (FULL LIST)

DRIVE&SEGMENT: **Unsupervised semantic segmentation** of urban scenes via cross-modal distillation.
- first author, published at ECCV'22 [1] (**oral** presentation) and an extended version published in IJCV [2] (IF=14.5) in 2025
- Developed a multimodal unsupervised segmentation approach to train machine learning models.
- Trained a model using multiple modalities: RGB images and 3D points as priors during training.

POP-3D: **Open-vocabulary 3D occupancy prediction from images**
- Published at NeurIPS'23 [3], first author.
- Trained a model using 3D points and images to perform open-vocabulary 3D occupancy prediction from images.
- The model allows for 3D segmentation based on natural language queries from provided RGB images.
- Multi-GPU training on HPC infrastructure (Karolina supercomputer).

Dataset augmentation using generative models
- paper *Artificial dummies for urban dataset augmentation* published at AAAI'21 [4], first author
- pedestrian dataset augmentation using generative adversarial networks to improve the performance of pedestrian detectors

A study of **Test-time Contrastive Concepts for Open-World Semantic Segmentation**
- training-free approaches to single-object segmentation using pre-trained vision-language models
- analysis of large-scale pre-training image-language data, working with LLMs (Mixtral)
- preprint [5], collaboration with Monika Wysoczańska (Warsaw Uni.), Oriane Siméoni (Meta FAIR), and others

Unsupervised object localization: Observing the background to discover objects
- Published at CVPR'23 [6], collaboration with Oriane Siméoni (Meta FAIR), Gilles Puy (valeo.ai) and others
- A lightweight approach to unsupervised object localization by looking for the background in the feature space

MOCA: **Self-supervised representation learning** by predicting masked online codebook assignments
- published in TMLR [7], to be presented as a poster at ICLR'25
- collaboration with Spyros Gidaris, Andrei Bursuc (both valeo.ai) and others
- single-stage and standalone method with good contextual reasoning properties and invariance to image perturbations

Detecting decision ambiguity from facial images
- published at a conference on Automatic Face & Gesture Recognition (FG 2018) [8]
- collected a large-scale dataset from "Who Wants to Be a Millionaire?" TV show, automatic annotation of the dataset
- prepared during an internship with Jiri Matas (CTU VRG)

Ongoing projects: large vision-language model fine-tuning for autonomous driving; LLM fine-tuning on domain-specific data

SELECTED WORK EXPERIENCE

Task Leader for the EXA4MIND EU Project *2023 – now*
- Task leader and machine learning researcher in the EXA4MIND EU project targeted at working with extreme-scale data
- developed and helped to deploy computer vision software to run on hundreds of terabytes of industry videos on HPC

Visiting Researcher at VALEO.AI, Paris *'22,'23,'24*
- Worked with Patrick Pérez (now Kyutai), Andrei Bursuc, and Oriane Siméoni (now FAIR) on research in ML for driving, resulted in several publications, e.g., ECCV'22 [1] NeurIPS'23 [3], or CVPRw'24 [9] and more

Student Researcher at VALEO *2017 – now*
- Research in computer vision for perception and data generation for autonomous driving; selected publications: [10, 11, 12]

SKILLS AND OTHER EXPERIENCE

Extensive **experience** with Python, PyTorch, SciPy, Bash, HPC, and more.
Reviewer for CVPR, ECCV, NeurIPS, IJCV, and others.
Voluntary work in the Czech borderlands; running and co-organizing **youth camps** and formative courses
Proficient in **English**, playing **guitar**, local **football** team founder/manager/player

FULL LIST OF PUBLICATIONS

[1] Antonin Vobecky, David Hurych, Oriane Siméoni, Spyros Gidaris, Andrei Bursuc, Patrick Pérez, and Josef Sivic. Drive&segment: Unsupervised semantic segmentation of urban scenes via cross-modal distillation. In *European Conference on Computer Vision*, pages 478–495. Springer Nature Switzerland Cham, 2022.

[2] Antonin Vobecky, David Hurych, Oriane Siméoni, Spyros Gidaris, Andrei Bursuc, Patrick Pérez, and Josef Sivic. Unsupervised semantic segmentation of urban scenes via cross-modal distillation. *International Journal of Computer Vision*, pages 1–23, 2025.

[3] Antonin Vobecky, Oriane Siméoni, David Hurych, Spyridon Gidaris, Andrei Bursuc, Patrick Pérez, and Josef Sivic. Pop-3d: Open-vocabulary 3d occupancy prediction from images. *Advances in Neural Information Processing Systems*, 2023.

[4] Antonin Vobecky, David Hurych, Michal Uricar, Patrick Pérez, and Josef Sivic. Artificial dummies for urban dataset augmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 2692–2700, 2021.

[5] Monika Wysoczańska, Antonin Vobecky, Amaia Cardiel, Tomasz Trzciński, Renaud Marlet, Andrei Bursuc, and Oriane Siméoni. A study of test-time contrastive concepts for open-world, open-vocabulary semantic segmentation. *arXiv:2407.05061*, 2024.

[6] Oriane Siméoni, Chloé Sekkat, Gilles Puy, Antonin Vobecky, Éloi Zablocki, and Patrick Pérez. Unsupervised object localization: Observing the background to discover objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3176–3186, 2023.

[7] Spyros Gidaris, Andrei Bursuc, Oriane Siméoni, Antonin Vobecky, Nikos Komodakis, Matthieu Cord, and Patrick Perez. Moca: Self-supervised representation learning by predicting masked online codebook assignments. *Transactions on Machine Learning Research*, 2023.

[8] Pavel Jahoda, Antonin Vobecky, Jan Cech, and Jiri Matas. Detecting decision ambiguity from facial images. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 499–503, 2018.

[9] Sophia Sirko-Galouchenko, Alexandre Boulch, Spyros Gidaris, Andrei Bursuc, Antonin Vobecky, Patrick Pérez, and Renaud Marlet. Occfeat: Self-supervised occupancy feature prediction for pretraining bev segmentation networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4493–4503, 2024.

[10] Antonin Vobecky, Michal Uricar, David Hurych, and Radoslav Skoviera. Advanced pedestrian dataset augmentation for autonomous driving. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2019.

[11] Michal Uricar, Jan Ulicny, Ganesh Sistu, Hazem Rashed, Pavel Krizek, David Hurych, Antonin Vobecky, and Senthil Yogamani. Desoiling dataset: Restoring soiled areas on automotive fisheye cameras. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0, 2019.

[12] Michal Uricar, Ganesh Sistu, Hazem Rashed, Antonin Vobecky, Varun Ravi Kumar, Pavel Krizek, Fabian Burger, and Senthil Yogamani. Let's get dirty: Gan based data augmentation for camera lens soiling detection in autonomous driving. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 766–775, 2021.

[13] Julia Skovierova, Antonin Vobecky, Miroslav Uller, Radoslav Skoviera, and Vaclav Hlavac. Motion prediction influence on the pedestrian intention estimation near a zebra crossing. In *VEHITS*, pages 341–348, 2018.