Helin (Henry) Cao

tumcaohelin@gmail.com +49 1742463372 LinkedIn Google Scholar GitHub Personal Website

About Me

PhD in Robotics with expertise in autonomous driving perception, generative models, and 3D scene understanding, with a proven track record of publications in **top-tier robotics conferences**, reinforced by practical experience through **industry collaborations and internships**, and as a contributing member of the **World Champion Team** in domestic service robots.

Education

University of Bonn, Germany 10/2022 – Present

Ph.D. Center for Robotics Bonn & Lamarr Institute

- *Focus: Towards Spatial Intelligence: 3D Semantic Occupancy Prediction with Vision and LiDAR*
- *Advisor: Prof. Dr. Sven Behnke (ELLIS Fellow, Editor of IEEE T-RO & RA-L)*

Technical University of Munich, Germany 10/2018 – 05/2022

M.Sc Electrical and Computer Engineering

- *Thesis: 3D Semantic Scene Completion for RGB-D Scans*
- *Advisor: Prof. Dr. Matthias Nießner & Prof. Dr.-Ing. Eckehard Steinbach*

Publications

1. **SLCF-Net: Sequential LiDAR-Camera Fusion for Semantic Scene Completion using a 3D Recurrent U-Net**
   ***Helin Cao**, Sven Behnke*, in **ICRA 2024**
2. **DiffSSC: Semantic LiDAR Scan Completion using Denoising Diffusion Probabilistic Models**
   ***Helin Cao**, Sven Behnke*, in **IROS 2025**
3. **OC-SOP: Enhancing Vision-Based 3D Semantic Occupancy Prediction by Object-Centric Awareness**
   ***Helin Cao**, Sven Behnke*, in **SMC 2025**
4. **SWA-SOP: Spatially-aware Window Attention for Semantic Occupancy Prediction in Autonomous Driving**
   ***Helin Cao**, Rafael Materla, Sven Behnke*, in **SMC 2025**

Competitions

RoboCup@Home 2023, Bordeaux, France **4th Place among 15 teams worldwide**

Led R&D of object perception stack **Team NimbRo**

- Investigated multiple approaches and selected Mask DINO for fine-grained instance segmentation, enabling precise point cloud extraction for downstream tasks.
- Pre-trained on COCO and fine-tuned with real-world data (lighting, arrangement, density), enhancing robustness and adaptability across diverse scenarios.
- Built and deployed an end-to-end pipeline (data collection, augmentation, annotation, fine-tuning, deployment), enabling timely delivery of task-specific models within hours during the competition.
- Collaborated with teammates on manipulation to accomplish tasks such as Serving Breakfast, Cleaning the Table, and Storing Groceries.
- Contributed to system integration, including communication debugging, ROS2 migration, and hardware setup.

RoboCup@Home 2024, Eindhoven, Netherlands **Champion among 17 teams worldwide**

- Extended the perception stack with a new feature for open-vocabulary instance segmentation (Grounding-DINO + NanoSAM), enabling recognition of arbitrary user-specified objects and improving adaptability in GPSR tasks.
- Collaborated with teammates on speech to accomplish GPSR/EGPSR tasks, leveraging GPT-based task planning to translate natural language instructions into robot commands and execute arbitrary user requests.

Industry Experience

Valeo, Remote, Germany 06/2022 – 09/2022

Research Collaboration mentored by Dr. Norman Müller

- Built a city-scale semantic occupancy map using the SemanticKITTI dataset.
- Evaluated state-of-the-art baselines on the semantic scene completion task, analyzing performance under complex urban conditions.
- Contributed experimental insights to Valeo's autonomous driving research collaboration.

AI Digi+ Solutions GmbH, Munich, Germany 11/2020 – 04/2021

Robotics Engineer Internship mentored by Prof. Dr. Vahid Salehi

- Developed an autonomous mobile robot prototype *AICab2*, focusing on navigation and system integration.
- Hardware integration: Mecanum omnidirectional wheels, RPLIDAR, Arduino, Jetson Nano.
- Software deployment: ROS Navigation stack, Cartographer SLAM.
- *AICab2* was delivered to the *BMW Group Research & Innovation Centre (FIZ)* for proof-of-concept validation.

Skills

- Programming: Python, C/C++, Shell, LaTeX
- Tools: PyTorch, ROS, Matlab/Simulink, Slurm, Container, OpenCV, Blender, Git
- Hardware: PAL Robot Tiago++, Jetson Orin Nano, Velodyne LiDAR, Realsense RGB-D camera
- Language: Chinese (native), English (C1), German (B2)