KARLO KOLEDIĆ

Computer Vision and Robotics Researcher

🌐 koledickarlo.github.io   ✉ kkoledic@gmail.com

⌨ github.com/koledickarlo   in linkedin.com/in/kkoledic/

RESEARCH INTERESTS

I am an engineer working at the intersection of computer vision and robotics, with focus on 3D perception. My main expertise lies in research and development of novel deep learning architectures that tackle traditional geometric robotics problems, including SLAM, 3D reconstruction and sensor fusion.

EXPERIENCE

Researcher
Faculty of Electrical Engineering and Computing

March 2021 – Present
Zagreb, Croatia

- Developed a novel monocular visual odometry algorithm, fusing self-supervised depth estimation with traditional epipolar geometry optimization, outperforming all state-of-the-art methods on the KITTI dataset
- Created a novel monocular depth estimation method, solving a long-standing problem of generalization for arbitrary camera parameters.
- Researched and utilized novel methods for enhancement of robotic perception, including: diffusion-based NeRF for surround-view driving scenes, world models for active inference, Gaussian splatting for visual SLAM
- Creating and conducting exams and exercises, mentoring students during projects and master theses
- Coursework: Machine Learning, Estimation Theory, Autonomous Mobile Robots, 3D Vision, Sensors and Perception in Robotics

Computer Vision Engineer
Snap Inc.

June 2024 – September 2024
Vienna, Austria

- PhD research internship: computer vision for latest generation of Snap's augmented reality glasses
- As a part of the SLAM team, contributed to localization and mapping efforts, following latest CI/CD practices
- Led the project of deep-learning-based future pose forecasting, resulting in increased accuracy and robustness during warping and rendering
- Effort included all stages of development, including initial research, data acquisition and analysis, model training and model deployment on a resource-constrained device
- Implemented various data streaming, synchronization and quantization techniques, ensuring real-time performance for high-frequency input data

Junior R&D Engineer
Visage Technologies AB

August 2020 – March 2021
Zagreb, Croatia

- Researched and developed various Extended/Unscented Kalman filter methods for state-constrained estimation
- Worked in an autonomous driving team focused on multiple object tracking for a major automobile company

Computer Vision Intern
RealNetworks, Inc.

July 2018 – July 2019
Zagreb, Croatia

- Implemented face detection/face recognition networks in native CUDA/cuDNN/cuBLAS code. Achieved a 2.5 times faster performance than the TensorFlow CUDA implementation
- Implemented face detection algorithm for Android NDK using TensorFlow Lite
- Research on various frequency and movement based techniques for face liveness detection
- Generated synthetic data in MakeHuman/Blender and implemented a face pose estimator in TensorFlow

EDUCATION

Ph.D. in Electrical Engineering and Computer Science
University of Zagreb

Feb 2022 – Present

M.Sc. in Electrical Engineering (Robotics)
University of Zagreb

Sep 2018 – July 2021

- Thesis: Visual-inertial Odometry based on Factor Graph Optimization

Student exchange	Sep 2019 – July 2020
Politecnico di Milano	

B.Sc. in Computer Science	Sep 2015 – July 2018
University of Zagreb	

- Thesis: Detection and Classification of Teeth in Panoramic X-ray Images Using Faster-RCNN Architecture

SKILLS

Programming: Python, C++, Java, MATLAB, bash
Software:: Linux, ROS, Git, Docker, ONNX, LaTeX
Libraries: CUDA/cudNN, Numpy, PyTorch, TensorFlow, Pandas, Darts, CVX, OpenCV, Eigen, GTSAM
Deep Learning: Object Detection/Segmentation/Tracking, Domain Adaptation/Generalization, GAN, Diffusion, Transformers, VLMs, Run-time Optimization/Quantization
Computer Vision: Monocular/Stereo Depth Estimation, Multi-view Geometry, Gaussian Splatting, NeRF, Calibration
Robotics: Sensor Fusion, Visual Odometry, SLAM, Nonlinear Optimization, Model Predictive Control, World Models, Deep Reinforcment Learning
Languages: Croatian (native), English (C1-C2), German (B1)

PUBLICATIONS

GVDepth: Zero–Shot Monocular Depth Estimation for Ground Vehicles based on Probabilistic Cue Fusion
Karlo Koledić, Luka Petrović, Ivan Petrović, Ivan Marković
Project Page | arXiv | PDF
under review

GenDepth: Generalizing Monocular Depth Estimation for Arbitrary Camera Parameters via Ground Plane Embedding
Karlo Koledić, Luka Petrović, Ivan Petrović, Ivan Marković
arXiv | PDF
under review in Applied Soft Computing (**ASC**)

MOFT: Monocular Odometry based on Deep Depth and Careful Feature Selection and Tracking
Karlo Koledić, Igor Cvišić, Ivan Marković, Ivan Petrović
DOI
2023 International Conference on Robotics and Automation (**ICRA**)

Towards Camera Parameters Invariant Monocular Depth Estimation in Autonomous Driving
Karlo Koledić, Ivan Marković, Ivan Petrović
DOI
2023 European Conference on Mobile Robots (**ECMR**)

Enhancing Gaussian Splatting SLAM with Feature-based Tracking
Muhammad Awais, Karlo Koledić, Luka Petrović, Ivan Marković
2025 International Conference on Automation, Robotics, and Applications (**ICARA**)

PROJECTS

F1tenth autonomous racing: Development of Model Predictive Contouring Control algorithm for autonomous racing,
4th fastest time at **IROS 2020**
UWB/IMU sensor fusion: Trajectory estimation during sport activities. Fusion via factor-graph optimization in GTSAM
Drone 6-DoF estimation: Error-state Kalman filter for IMU, optical flow and time-of-flight measurements. Implementation uses the Eigen library, allowing real-time 800Hz estimation on ESP-32 and Arduino MCUs
Aerial image matching: POC for drone image matching via incremental robust feature matching and homography estimation

AWARDS

Rector's Award for best project
Dental State Estimation using Deep Learning
- Analyzed dental images to predict biological traits with the accuracy of 98% and age with a median error of 3 years.
- Assessed dental state via object detection and classification