+1(548)333-3485
Waterloo, Ontario
khannaaiig@gmail.com

DHEERAJ KHANNA

Portfolio: dkhanna511.github.io
Github: dkhanna511
Linkedin: dheeraj-khanna-05

Education

MASc in Systems Design Engineering, *University of Waterloo, Canada* **Sep 2022 — Mar 2025**
B.Tech in Electronics and Communication, *Guru Gobind Singh Indraprastha University, India* **Aug 2016 — Sep 2020**

Skills

Languages & Platforms	Python, C++, C, Linux, Git, Docker, SSH, CI/CD
Frameworks & Libraries	PyTorch, LibTorch, CUDA, TensorFlow, Keras, ONNX, Scikit-learn, NumPy, Pandas, OpenCV, Pillow, Jupyter Notebook, Plotly, Seaborn, Matplotlib, Weights & Biases
Technologies	Perception, Computer Vision, Machine Learning, Object Detection, Segmentation
Hardware	VLP-16 LiDAR, FLIR Pointgrey Cameras, Raspberry Pi 3B, Arduino, NVIDIA Jetson

Work Experience

Computer Vision Researcher | **VIP Lab, University of Waterloo ; ATS Automation** | *Supervisor: Prof. John Zelek* **Jan 2023 — Present**

- Leading research on Visual Place Recognition (VPR) for aerial imagery; exploring transformer-based, cross-view, and multimodal architectures (CNNs, DINOv2, ViTs) to extract robust features for UAV-based geo-localization.
- Designed a novel prediction model integrating Mamba state-space model and self-attention mechanisms with advanced data association strategies and YOLO-X for real-time multi-object tracking, achieving 3-7% improvement over other methods on complex motion scenarios.
- Created a novel automatic bounding box annotation pipeline for videos with multiple objects using Point Tracking, Segment Anything, and YOLO-v8, achieving the annotation speed of 10 FPS. (*View Paper*)
- Developed and containerized a multi-object tracking pipeline to estimate velocities of hundreds of assembly parts in a high-throughput industrial setting, enhancing production efficiency and enabling scalable deployment.

Computer Vision Research Engineer | **LENS Corporation, India** **Feb 2022 — May 2022**

- Migrated a latent fingerprint extraction and matching pipeline from Python to C++ using Libtorch to enable faster GPU and CPU-based deployment, optimizing performance and reducing latency.
- Programmed custom signal processing functions, including Fourier Transforms and Gabor Filtering, in C++ from scratch to process fingerprint data, overcoming constraints in existing matrix and signal processing C++ libraries.

Perception Research Engineer | **Autonomous Vehicle Project, IIIT Delhi, India** | *Supervisor: Dr. Saket Anand* **Oct 2020 — Feb 2022**

- Built and deployed a multi-sensor calibration system integrating 2 FLIR Pointgrey cameras and 3 VLP-16 LiDARs using a checkerboard target; implemented LiDAR–LiDAR alignment with Iterative Closest Point (ICP), achieving < 4° rotational and <10 cm translational error across modalities using surface normal and Euclidean metrics.
- Developed and optimized a real-time lane detection pipeline in C++ and TorchScript; projected lane boundaries onto HD maps to identify driveable regions, maintaining 12+ FPS on NVIDIA Jetson Xavier. (View Project)

Research Intern | **Microsoft Research India** | *Supervisor: Dr. Akshay Nambi* **Feb 2020 — Sep 2020**

- Redesigned and scaled the Automated Driver License Testing (ALT) project for Regional Transport Offices (RTO) across India. The scaling led to the deployment of the project in 10+ cities in India with 99% accuracy in automated driving test results. (*View Project*)

Projects

Image Enhancement and Object Detection in Rainy Weather Conditions | *Deep Learning, GANs, PyTorch* **Jan 2023 — Apr 2023**

- Implemented deep learning models for image enhancement in rainy weather, integrating GAN-based denoising with morphological transformations and optimized loss functions, improving visual clarity by increasing PSNR by 5.8%.

Multi-Sensor Calibration and Sensor Fusion for Autonomous vehicle | *Photogrammetry, ROS, Sensors* (*View Project*) **Oct 2020 — Feb 2022**

- Created an automated calibration pipeline using ROS for synchronized data capture from multiple sensors; applied RANSAC and DBSCAN for point cloud filtering and visualized alignment in RViz.
- Refined sensor pose estimates using Linear Algebra, 3D geometry, and Statistical ML techniques; integrated calibrated outputs into the autonomous vehicle's perception stack to support tasks like lane detection, HD map projection, and real-time localization.

RecipeDB | *Pandas, Data Visualizations* (*View Project*) **Jun 2019 — Sep 2019**

- Curated and integrated recipe data from 7 online sources, performed multi-level statistical analysis across 22 regions, and programmed interactive visualizations using Pandas, Matplotlib, and Plotly to present insights on regional dietary and nutritional patterns.

Publications

- **SportMamba: Adaptive Non-Linear Multi-Object Tracking with State Space Models for Team Sports**, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW-2025). (View Paper)
- **Attention-Mamba for Multi-Object Tracking**, Conference on Robots and Vision (CRV), 2025. (View Paper)
- **POPCat: Propagation of Particles for Complex Annotation Tasks**, Conference on Robots and Vision (CRV), 2024. (*View paper*)
- **Recipedb: A resource for exploring recipes**,Database Journal (Oxford), pp. baaa077, Nov 2020. (*View Paper*)
- **NTIRE 2019 Challenge on Video Super-Resolution:Methods and Results**, Co-author of team paper, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW-2019). (*View paper*)