

SUMMARY

Computer Vision Researcher specializing in **Computer Vision and Face Recognition** for **in-cabin automotive systems**. Delivered **production-grade ML solutions** at **Mercedes-Benz R&D** and published at **IROS and NeurIPS**. Experienced in applying **Vision-Language Models (VLMs)** and **Large Language Models (LLMs)** across multimodal perception, fine-tuning, and retrieval-augmented systems. Expertise in translating research into scalable systems, with demonstrated impact including **multi-million-euro cost savings**.

TECHNICAL SKILLS

Languages & Tools: Python, GCP, Git, Weights & Biases (W&B)

ML & DL Frameworks: PyTorch, Hugging Face Transformers, PEFT/LoRA, TRL

Computer Vision: Face Recognition, CNNs, Vision Transformers, Object Detection, Semantic & Instance Segmentation, Depth Estimation, Contrastive Learning, Knowledge Distillation, Incremental & Open-World Learning, IR-to-RGB Colorization

LLMs, VLMs & Retrieval: Vision-Language Models, Multimodal RAG, Two-Stage Retrieval, Cross-Encoder Reranking, QLoRA Fine-Tuning, ChromaDB

EXPERIENCE

• Mercedes-Benz Research and Development India

Bengaluru, India

Computer Vision and Deep Learning Engineer

April 2022 – Present

Member of the Interior-Sensing (IS) Algorithm team, driving development of **computer vision solutions for in-cabin automotive systems**. Contributed to **safety-critical and luxury applications** including **face recognition, depth estimation, and weight prediction**. Owned the pipeline from **proof-of-concept (PoC) to production-ready systems**, translating research ideas into deployable solutions under real-world constraints.

Core responsibilities

- Led end-to-end development of production-grade **in-cabin face recognition systems**, from data strategy and model development to deployment, achieving stringent accuracy and robustness targets.
- Architected a **dual-level authentication framework** enabling both low-latency personalization and high-assurance security for privacy- and payment-sensitive use cases.
- Developed **cross-domain face recognition** and IR-to-RGB colorization solutions to bridge RGB/NIR modality gaps and reduce dependency on large-scale RGB data collection.
- Built deep learning models for occupant analysis, including IR-based **weight estimation**, replacing hardware sensors with software-driven solutions.
- Collaborated across software, quantization, validation, and platform teams, while driving GCP-based ML infrastructure initiatives and contributing to patentable innovations.

Achievements

- Achieved **state-of-the-art performance in automotive face recognition**, outperforming industry-leading solutions deployed in airport-grade security systems
- Introduced an **innovative IR-to-RGB colorization approach**, removing reliance on costly data collection and annotation, delivering **€1.5M+ cost savings** and reducing pipeline turnaround time from **210 days to 2 days (99% reduction)**

• Celstream Systems Pvt. Ltd.

Bangalore, India

Software Engineer

Sep 2014 – Sep 2016

- Built the product's main JavaScript UI console and dynamic data-visualization modules (IgniteUI), and led migration of the in-house application from Adobe Flash to a JavaScript environment.
- Developed REST APIs in Java and data adapters powering live data-visualization modules.

PATENTS & INTELLECTUAL PROPERTY

1. **In-Cabin Face Authentication System** (Filed in Germany and India)
Implemented Patent Award; Inventor lead; Enabling secure multi-level authentication in automotive systems
2. **Intelligent In-Vehicle Hyper-Personalization** (Filed in Germany and India)
Co-Inventor; Enhancing user experience by assimilating signals, memory, and preferences

PERSONAL PROJECTS

- **DocLens — Local Multimodal RAG for Technical Documents** [GitHub](#)
Python · sentence-transformers · ChromaDB · Ollama (Qwen3 / Qwen-VL) · PyMuPDF
Highlights: Recall@5 $\approx$ 0.93 (reranked) · MRR $\approx$ 0.74 · 100-question gold set · 100% local
 - Built a fully-local multimodal RAG assistant for technical PDFs with two-stage retrieval (bi-encoder recall plus cross-encoder reranking) and page-level citations, running entirely on-device with no external APIs.
 - Implemented multimodal caption-and-embed: figures extracted, captioned by a local VLM, embedded into a unified index, and passed back to the vision model for image-grounded answers.
- **Nutrition Coach — Vision-Language Model for Calorie Estimation** [GitHub](#)
PyTorch · Hugging Face Transformers · PEFT/LoRA · TRL · bitsandbytes · Qwen3-VL · Kaggle T4 GPU
Highlights: 81.4 kcal MAE ($\sim$ 34% below baseline) · 100% JSON schema compliance
 - Fine-tuned Qwen3-VL-4B with QLoRA (4-bit, LoRA) to estimate calories and macros from food images, outputting structured JSON at 100% schema compliance on held-out data.
 - Ran a controlled experiment comparing two label sources: real lab-measured data (Nutrition5k) produced genuine estimation while synthetic class-average labels (Food101) degraded accuracy below the un-tuned baseline — a data-quality finding with direct modeling implications.

AWARDS & RECOGNITION

- **Kaizen (Operational Excellence) Recognition** Mar 2026
Selected at department level for innovation leading to significant cost and efficiency gains
- **World Quality Week Recognition** Nov 2025
Presented the benefits, optimizations, and achievements of our solutions
- **Customer Appreciation Recognition** Sep 2025
Recognized by Mercedes-Benz client team; [View Post](#)
- **Impact Award (Silver)** Feb 2025, 2023
Recognized for delivering high-impact ML solutions
- **Team Excellence Award (Silver)** Oct 2024
Awarded for collaborative delivery of production-grade systems

PUBLICATIONS

- **Deepak Singh** et al., “New Vehicles on the Road? No Problem - We’ll Learn Them Too”, 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 2022. [\[Link\]](#)
- **Deepak Singh** et al., “ORDER: Open World Object Detection on Road Scenes”, NeurIPS Machine Learning for Autonomous Driving (ML4AD) Workshop, 2021. [\[Link\]](#)
- **Deepak Singh** et al., “Evaluation of Detection and Segmentation Tasks on Driving Datasets”, International Conference on Computer Vision and Image Processing (CVIP), Communications in Computer and Information Science (CCIS), Vol. 1567, Springer, 2021. [\[Link\]](#)

EDUCATION

- **International Institute of Information Technology** Hyderabad, India
Master of Science by Research in Computer Science, Specialization in Artificial Intelligence Aug. 2018 – Dec. 2021
Master’s Thesis (Graduate Research Scholar, CVIT): Benchmarked SOTA detection and segmentation models on Cityscapes, BDD, and IDD for autonomous driving. Designed ORDER (Open-World Object Detection on Road Scenes) with Feature-Mix, focal regression loss, and curriculum learning to improve unknown-object recognition and small-object detection, achieving superior results on BDD and IDD.