Vishal Vishnu Kagade

kagadevishal@gmail.com
(+49)15758730898
https://www.linkedin.com/in/vishal-kagade/
https://github.com/Vishalkagade
Address: Haunwöhr. Straße 99, 85051
Ingolstadt, Germany

SUMMARY

- **AI Engineer** with 3+ years of experience in building real-time perception systems and GenAI applications across automotive, and accessibility domains.
- Skilled in **Python, PyTorch**, and LLM frameworks like **LangChain and LangGraph**. Experienced in deploying **multimodal** deep learning models (camera, LiDAR, RADAR) and RAG pipelines for user-centric AI solutions.
- Proven track record of designing end-to-end AI initiatives from deep model development to cloud deployment using Docker, Streamlit, and CI/CD pipelines.

EDUCATION

Oct 2021 – May 2025 **Technische Hochschule Ingolstadt** Ingolstadt, Germany
Master of Engineering in International Automotive Engineering, GPA: 2.0/4 (German Grading System, 1.0 = Best)

Aug 2016 – May 2019 **Savitribai Phule Pune University** Pune, India
Bachelor of Technology Mechanical Engineering, GPA: 1.4/4 (German Grading System, 1.0 = Best)

WORK EXPERIENCE

Nov 2024 – Present **SmartAIs GMBH** Munich, Germany
AI and Computer Vision Engineer

- Developed advanced **stereo vision-based** obstacle detection systems leveraging **C++ and Unity C#** for real-time deployment on mobile devices, surpassing LiDAR accuracy.
- Developed a **Vision Transformer-based** stereo depth refinement model in **Python** and **PyTorch**, integrating monocular depth estimation with LiDAR supervision to outperform conventional stereo methods.
- Built an **onboarding RAG application** for visually impaired users using LangChain and LangGraph, integrating semantic retrieval (ChromaDB), prompt design, and vector-based indexing.
- Deployed the solution using **Streamlit** and evaluated user interaction for real-world usage, demonstrating measurable usability improvement.
- Applied **LLM chaining logic** for stepwise interaction design, simulating multi-agent collaboration patterns and showcasing GenAI capability beyond basic QA tasks.

May 2024 – Dec 2024 **Infineon Technologies** Munich, Germany
Master Thesis Student

- Built a real-time **Multi-Modal deep learning** algorithm that fuses RADAR and camera data for simultaneous **depth estimation and object detection** using a single backbone, achieving 37 FPS on A100 GPU systems.
- Implemented a **novel** vision **transformer**-based cross-sensor association network with a **CNN-based** decoder in PyTorch framework, reducing depth estimation error by 8% over state-of-the-art methods and significantly improving spatial feature alignment.
- **Deployed** a real-time depth estimation model on **Hugging Face** with an interactive Gradio-based demo, served via FastAPI, and containerized the entire application using **Docker** for deployment and reproducibility.
- Leveraged **MLOps** tools such as Weights & Biases, MLflow for efficient training, monitoring, and fine-tuning of deep learning models.

Nov 2023- Apr 2024 **Infineon Technologies** Munich, Germany
Computer Vision Working Student

- Designed and tested **ADAS functions** for vehicle tail-light detection and maneuvering action analysis to enhance driving safety on edge computing devices.

- Developed and validated traffic light detection algorithms using **meta-learning with deep learning** models including YOLOv5, SSD, DETR, and Faster R-CNN.
- Built a **robust dataset** for various AI applications including detection, **segmentation**, and motion planning, and developed a parallel processing pipeline for dataset preprocessing over AWS S3 bucket.
- Collaborated with a cross-organizational team on the "Althena project" for ADAS feature development, utilizing frameworks such as GitHub and **Weights & Biases**, integrated with **CI/CD** pipelines.

Jan 2023 – Oct 2023 **Fraunhofer IKS** Munich, Germany

Computer Vision Working Student

- **Researched and developed** 2D multi-object tracking using ByteTrack with Kalman filtering and enhanced association methods, benchmarking performance against state-of-the-art DeepSort.
- Trained a **deep descriptive** neural network to reduce identity switches and improve overall tracking performance.
- Developed an **advanced evaluation metric** for video object detection, accounting for detection accuracy across consecutive frames.

ACADEMIC PROJECTS

Hydranets-The one multitasking algorithm [Github]

- Implemented a **multitasking algorithm** for classification and regression on the UTK Face dataset using PyTorch.
- Developed a real-time encoder-decoder model for predicting semantic segmentation and depth maps, featuring a lightweight **Mobile Net**-based encoder and a **Retina Net**-inspired decoder.
- Evaluated **3D segmentation** maps by leveraging metric depth maps and segmentation outputs with the Open3D library.

RAG application for app onboarding for blind people [Github]

- Built an onboarding RAG application for visually impaired users using the **LangChain** framework and evaluated its real-time usability.
- Deployed the solution on **Streamlit** for demo, with a Chroma vector store for semantic retrieval, ensuring robust performance through cross-testing of retrieval strategies, indexing configurations.
- Developed multiple agent workflow prototypes (parallel, sequential, iterative, conditional) using LangGraph and OpenAI function calling.

Deep Learning and Edge Computing (Master Semester Project) **[Github]**

- Developed a **Single Shot Detector** for car and number plate detection from scratch using PyTorch and TensorFlow and integrated it with **Google Coral Board** and imported it to **TensorFlow Lite** using ONNX.
- Performed image data augmentation using computer vision libraries and employed various machine learning algorithms and AI neural network architectures, such as VGG16, Mobile-NET, for image classification.

Vision Language Models (VLM) Bootcamp [Certificate]

- Completed the OpenCV **Vision Language Model** Bootcamp with a **96%** score, gaining practical skills in applying Vision Language Models (VLMs) for image understanding and description generation.
- Worked hands-on with models like **CLIP** for zero-shot image classification and **Qwen2.5-VL** for image captioning and object detection, enhancing multimodal AI capabilities.

SKILLS, INTERESTS & OTHERS

Programming Languages: Python, C++
Frameworks and Libraries: PyTorch, TensorFlow, LangChain, LangGraph, OpenCV, FastAPI, Streamlit
Technologies: LiDAR, RADAR, Camera, Embedded AI
Deployment & MLOps: Docker, Gradio, AWS (S3), Weights & Biases, CI/CD
Soft Skills: Quick Lerner, Good Communicator, Collaboration, Independent
Languages: English (Fluent), German (B1)