

Anchit Gupta

☎ +91 9682045965 | ✉ gupta.anchit.99@gmail.com | 🔗 anchit-gupta-22-ag | 📄 anchitgupta22.github.io | 📍 Noida, IN

PROFILE

Visual Computing master's graduate from Saarland University with applied research experience at DFKI, building multimodal systems that connect video understanding, computer vision, and language models. Experienced in PyTorch-based model development, distributed HPC training, and retrieval-augmented generation (RAG). Seeking ML/AI Engineer or Applied Research roles building multimodal, video-language systems.

EDUCATION

Saarland University <i>MSc. Visual Computing</i> - Thesis: "Report Generation for Eye Surgeries: Bridging Video Understanding and Language Models" - Grade: 1.8 (German scale: 1.0 best, 4.0 passing; "Good")	Saarbrücken, Germany Mar. 2023 – June 2026
University of Petroleum and Energy Studies <i>Bachelor's in Computer Science, specialization in Business Analytics & Optimization</i> Grade: 8.91 (CGPA)	Dehradun, India Aug. 2017 – June 2021

EXPERIENCE

Research Assistant <i>German Research Center for Artificial Intelligence (DFKI)</i> - Contributed to the EEGain project; co-authored a pre-print on an EEG-based emotion recognition framework, training and evaluating PyTorch pipelines on SLURM-managed HPC clusters. - Containerized and deployed ML pipelines on institutional HPC infrastructure using Docker, managing remote compute environments and automating notebook-based workflows with Voila for the NFDI-MatWerk project.	Oct. 2024 – Mar. 2026 Saarbrücken, Germany
Summer Research Intern <i>Rhocron</i> Built a face recognition pipeline for low-resolution imagery, combining YOLOv3 detection with SRGAN-based super-resolution to improve recognition accuracy in noisy, unconstrained conditions.	May 2020 – July 2020 Gurugram, India

PROJECTS

Master's Thesis: Report Generation for Eye Surgeries <i>Python, Docker, Git</i> - Trained a phase-driven pipeline on DFKI's HPC cluster for automated German-language ophthalmic surgery report generation, using 255 paired surgical videos and insurance reports, with a multi-faceted lexical, semantic, and factual-consistency evaluation framework. - Integrated RAG, improving BLEU by $\approx 6\%$ ($65.8 \rightarrow 69.6$) and ROUGE-L by $\approx 7\%$ ($0.71 \rightarrow 0.76$), while cutting hallucination rate to 4.17% ($\approx 56\%$ factually-verified statements). - Found end-to-end multimodal pretraining (I3D/VideoMAE + LeoMistral) failed to align video tokens with limited data, motivating the phase-driven architecture instead.	Sept. 2025 – Mar. 2026
Various training approaches on the Lipophilicity dataset <i>Python, Docker, Git</i> - Implemented PEFT methods (BitFit, LoRA, iA3) from scratch, training 0.5 – 2.4% of model parameters vs. full fine-tuning. - LoRA outperformed iA3/BitFit across all strategies (best R^2: 0.644 vs. 0.551 – 0.602). - Influence-based data selection beat default sampling (R^2 0.738 vs. 0.733), matching full-dataset performance with fewer samples.	Oct. 2024 – Feb. 2025
SSL Methods for Urban Scene Segmentation <i>Python, Docker, Git</i> - Evaluated SSL pretext tasks (rotation, jigsaw) on U-Net/DeepLabv3; jigsaw reached 85.61% pixel accuracy with 20% labeled data (≈ 3 pts of the 88.45% supervised baseline), cutting training time $\approx 46\%$. - Jigsaw generalized better than rotation to foggy conditions, though both degraded significantly under fog.	Apr. 2024 – Aug. 2024

TECHNICAL SKILLS

Programming: Python, C++, C, Java
Machine Learning & Deep Learning: PyTorch, Scikit-learn, XGBoost, Hugging Face Transformers
Computer Vision: OpenCV, Image Segmentation, Video Understanding, Object Detection
LLMs & Multimodal AI: RAG, PEFT (LoRA, BitFit, iA3), Vision-Language Models, Prompt Engineering
Infrastructure & Deployment: Docker, SLURM, HPC Cluster Computing, Linux, Git