

Dhruv Makwana

AI Engineer | Deep Learning Researcher

📞 (+91) 99042-78934 | ✉️ dmakwana503@gmail.com | 🏠 dhruvmakwana.github.io | 🌐 dhruvmakwana | 📄 dhruvmakwana | 🎓 dhruvmakwana

Skills

Languages: Python, C++

ML/DL Frameworks: PyTorch, TensorFlow, HuggingFace, ONNX

GenAI & LLMs: LLM Finetuning & Alignment (DPO, GRPO, RLHF, PEFT/LoRA), RAG, Vector DBs (Milvus, FAISS), LLM Evaluation (RAGAS, LLM-as-Judge), LangGraph, CrewAI, AutoGen

Model Optimization: Quantization, Pruning, TFLite, TensorRT, Mixed Precision, Distributed Training (DDP, FSDP)

Deployment & Infra: AWS, Docker, FastAPI, vLLM

Computer Vision: Object Detection, Image Generation, Multimodal

Work Experience

Centific Global Solutions

Hyderabad, India

AI Engineer

Mar. 2024 - Present

- Reduced officer case-review time by ~85% (~75-90 min to ~12-15 min) by redesigning a law enforcement pipeline that pairs deterministic techniques (timestamp OCR, identity tracking, native-text parsing) with narrowly-scoped reasoning-model calls across video, audio, and document evidence, deployed on-premise for data privacy compliance.
- Achieved 0.1 WER on Spanglish (Spanish+English code-switching) ASR for law enforcement by generating synthetic conversational training data from expert-identified critical word lists and pretraining on novel concatenated English-Spanish corpora with noise augmentation before domain-specific fine-tuning.
- Built a real-time multi-camera security video analytics pipeline combining deterministic rule-based alerting with scoped VLM judgment and a correlating decision-layer agent, achieving 66.7% recall at 80% precision on unattended-package detection and 89.5% correct aggression/weapon alerts in live evaluation.
- Designed and trained SurgQ-VLM, a two-stage (SFT → GRPO) surgical video QA model with a novel bidirectional question-conditioned dual-stream compressor, reaching 0.779 keyword accuracy and a 10x BLEU-4 gain over zero-shot on a sealed 21,840-question benchmark.
- Built an LLM-powered data structuring pipeline for a food delivery client using GPT-4 and LangChain to normalize raw product data into schema-compliant structured outputs, with a GPT-4-as-judge quality gate that auto-approved ~90% of rows and reduced manual review effort by 85%.

Ignitarium Technology Solutions

Bangalore, India

AI Engineer

Dec. 2021 - Jan. 2024

- Developed a 6DoF pose estimation pipeline for robotic arm pickup of small industrial parts, resolving severe data scarcity via 3D mesh rendering for synthetic training data generation, multi-class YOLOv8 detection, and MegaPose for pose estimation.
- Developed and implemented high-performance computation kernels, reducing system latency in data transformations and computer vision tasks, improving processing speeds and system efficiency by 2x.
- Built a lightweight multitask CNN for gender (binary) and age-group (4-class) classification on a custom edge device, combining multiple Asian face datasets to overcome data scarcity and achieving 94% accuracy.
- Designed a prototype solution for fault diagnosis from machinery sensor signal data, achieving 98% accuracy for fault classification utilizing Fourier and wavelet-based noise reduction and LSTM and dilated convolution layers.

Qualitas Technologies

Bangalore, India

Machine Learning Engineer

Sep. 2020 - Nov. 2021

- Developed and implemented Faster RCNN and SSD-based object detection models, achieving more than 95% mAP score for chassis number detection.
- Implemented DeepLabV3 segmentation for manufacturing defect detection and optimized inference using TensorRT, OpenVINO, mixed precision, and INT8 quantization, achieving ~35% latency reduction with <1% accuracy drop.
- Reduced manual image annotation effort by ~90% by designing and building an active learning pipeline (Faster R-CNN pseudo-labeling + confidence filtering) for manufacturing inspection data, cutting labeling time from ~83 hours to ~8 hours per 10K-image project.

- Worked on PCB component detection and segmentation for recycling and low-light image enhancement as the first author, text removal from scene images and PCB component classification as the second author, global Fourier self-attention for multimodal downstream tasks, and Nuclei instance segmentation as the third author.
- Defined research problem, identified past research works, implemented custom modules to mitigate their drawbacks, and ran the experiments on a FICS PCB dataset for the PCB project and on the LOLv1 dataset for the low-light project.
- Performed experiments on previous methods and contributed to the data processing and post-processing of the proposed method for the multimodal project and the nuclei instance segmentation project.
- Contributed to writing different paper modules and preparing possible reviewer questions to strengthen the paper.

Education

LJIET (LJ Institute of Engineering and Technology)	Aug. 2016 - May 2020
B.E. in Computer Engineering	CGPA: 9.16/10
Relevant coursework: Advanced Engineering Mathematics, Python Programming, Data Mining, Machine Learning, Artificial Intelligence	

Publications

LIVENet: A novel network for real-world low-light image denoising and enhancement	2024
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	
Textual Alchemy: CoFormer for Scene Text Understanding	2024
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	
PCBSegClassNet — A light-weight network for segmentation and classification of PCB component	2023
Elsevier Expert Systems with Applications	
TPFNet: A Novel Text Inpainting Transformer For Text Removal	2023
International Conference on Document Analysis and Recognition (ICDAR)	
GAFNet: A Global Fourier Self Attention Based Novel Network for multi-modal downstream Tasks	2023
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	
DInPNet: A Novel Model for Printed Circuit Board (PCB) Components Classification	2023
International Symposium on Quality Electronic Design (ISQED)	
ACLNet: an attention and clustering-based cloud segmentation network	2022
Taylor & Francis Remote Sensing Letters	
WaferSegClassNet - Network for classification and segmentation of semiconductor wafer defects	2022
Elsevier Computers in Industry	
FEEDNet: Network for Nuclei Instance Segmentation for Histopathological Diagnosis	2022
IOP Physics in Medicine & Biology	

Certificates

TensorFlow Developer Certificate	2022
Microsoft Certified Azure Data Scientist Associate (DP100)	2021
Microsoft Certified Azure AI Fundamentals (AI900)	2020
Google Cloud Certified Cloud Digital Leader	2024
Google Cloud Certified Machine Learning Engineer	2024

Honors and Awards

Mukhyamantri Yuva Swavalamban Yojana Scholarship - ₹36,500 (50% of tuition fees)	2016-2020
Secured place in top 5 in TinyML contest at International Conference on Computer-Aided Design (ICCAD)	2022

Student Life & Volunteerism

IIT Bombay	
Social Media Marketer & College Ambassador	2019
• Served as a social media marketer and college ambassador for the IIT Bombay Tech Fest at my university. • Promoted the tech-fest on various social media platforms and within the university community, actively encouraging student participation and fostering engagement.	
LJIET	
Volunteer	2018
• Volunteered at a state-level technical coding competition that attracted 300 participants. • Assisted participants by explaining tasks and providing necessary resources, contributing to smooth event execution while supporting and guiding students.	