

ABDUL BASIT

Paris, France

basitmal36@gmail.com | +33 7 73 64 29 23 | <https://abdul-basit-ai.github.io/>

Profile

MSc.Graduate student in Machine Vision and Artificial Intelligence at the University of Paris-Saclay. Specializing in large language models (LLMs), Agentic AI, vision-language models (VLMs), video-language understanding, computer vision, deep learning, reinforcement learning, and graph neural networks. Strong interest in research-driven projects and building robust engineering systems that translate advanced AI concepts into practical solutions.

Research Interests: Continual Learning, Vision-Language Models, Large Language Models, Reinforcement Learning, Computer Vision, Agentic Systems, Foundational Models, Alignment.

Education

M.Sc. in Machine Vision and Artificial Intelligence

2024 – 2026

University of Paris-Saclay, France - (Grade M2: 14.6/20 M1: 12/20)

- Relevant coursework: Statistical Machine Learning, Reinforcement Learning, Information Theory, Deep Learning, 3D Data Analysis, Computer Vision, Signal Processing, Virtual and Mixed Reality.

Bachelor of Engineering(EE)

2020 – 2024

Bahria University - (Grade : 3.46/4.00)

- Relevant coursework: Embedded Systems, FPGA Design, Computer Vision, Machine Learning, Signal Processing, Programming, Computer Networks, Communication Systems.

Research Experience

AI Research Engineer Intern

Feb 2026 – Aug 2026

Orange Innovation - (Grade : 15.67/20)

- Modeled complex network topologies and developed performance-optimized Graph Neural Networks (GNNs) for network analysis tasks, including anomaly detection.
- Adapted ANEMONE into an edge-aware GATv2 self-supervised anomaly detection framework.
- Designed a physics-informed synthetic anomaly generator to simulate realistic network degradations.
- Built end-to-end machine learning pipelines with custom training loops and conducted distributed training on high-performance computing clusters using Slurm.
- Compared discrete, temporal and static graph formulations and analyzed FP and FN.

Multimodal Vision-Language Models for Remote Sensing

- Evaluated Qwen2.5-VL-7B-Instruct and InternVL3.5-8B architectures on EarthDial-Mini, a 19,995-example benchmark spanning RGB, synthetic aperture radar (SAR), multispectral optical, and paired cross-modal tasks.
- Conducted test-time ablation studies to measure modality contributions, modality bias, and data bias.
- Analyzed attention distributions to reveal that fine-tuning reinforces text dominance, with models consistently allocating 80-96% of their attention share to text tokens.
- Formulated a structured bias taxonomy connecting pre-training priors, architecture choices, and dataset composition to better understand multimodal failure modes.

Multimodal Video-Language Understanding

- Conducted a detailed study of state-of-the-art techniques for text-guided temporal video retrieval and long-context video understanding.
- Investigated frame sampling strategies including uniform, adaptive, and dynamic sampling for efficient retrieval of informative frames.
- Explored Transformer- and Mamba-based architectures such as STORM, LongVU, VideoCLIP, and related multimodal systems.

Agentic AI Assistant for SME Operations using MCP (Mistral AI Hackathon Project)

- Integrated Facebook with the Mistral Le Chat interface to streamline SME business page management, eliminating the need to juggle multiple applications.

- Enabled end-to-end content workflows from a single unified interface: utilized tool-calling to generate post text and photos, publish directly to Facebook, and instantly retrieve page analytics.

3D Scene Reconstruction from Images

- Surveyed advanced methods for 3D reconstruction from 2D images, including NeRF, Instant-NGP, PixelNeRF, Gaussian Splatting, and Neuralangelo.
- Studied core techniques such as volume rendering, multi-resolution hash encoding, CNN-based feature extraction, and signed distance functions.

Style Transfer and CNN Improvement Using GAN-Based Data Augmentation

- Implemented CNN baselines including ResNet50 and VGG16 for image classification tasks.
- Developed Conditional GANs (C-GANs) with progressive resolution growth to generate high-quality synthetic training data.
- Improved classifier performance using augmented datasets and implemented CycleGAN for unpaired image-to-image style transfer.

AI-Based Plant Monitoring System for Disease Identification

Funded Project

- Built and benchmarked five CNN architectures, including ResNet50, VGG16, VGG19, MobileNetV1, and MobileNetV2, for plant disease classification.
- Deployed optimized models on Raspberry Pi using TensorFlow Lite for real-time edge inference.
- Integrated the vision module with a cable-driven parallel robot and improved robot control from fixed-point navigation to joystick-based operation.

Technical Skills

Programming Languages: Python, C++, C, Verilog/VHDL, HTML/CSS, SQL, Pytorch/DDP, JAX, TensorFlow.

Agentic AI: Reasoning loops, LangChain, LangGraph, prompt engineering, tool calling, MCP, RAG.

Tools and Platforms: Git, Docker, Linux, Kubernetes, CI/CD, FastAPI, MLOPs, Grafana, Slurm

Deep Learning: CNNs, Autoencoders, GANs, Variational Autoencoders, Diffusion Models

Natural Language Processing: RNNs, LSTMs, Transformers, State Space Models

Reinforcement Learning: Value Iteration, Policy Iteration, PPO, DPO, GRPO, Intelligent Agents

Computer Vision: NeRF, Gaussian Splatting, SDF, Meshes, Voxels, Volume Rendering, Ray Tracing

Graph Machine Learning: GCN, GAT, GATv2, Community Detection, Knowledge Graphs, GraphRAG

Explainability: SHAP, LIME, Post-hoc Statistical Feature Attribution

LLM/VLM Concepts: CLIP, SigLIP, RoPE, RMSNorm, KV Cache, Quantization, Tokenization, Adaptation

Languages: English (Fluent), French (Actively Learning)

Selected Implementations

- Implemented Transformer and Vision Transformer from scratch using PyTorch.
- Built a multimodal vision-language model inspired by PaliGemma from scratch.
- Implemented LLaMA-2 architecture from scratch in PyTorch.
- Developed a small language model using reinforcement learning for mathematical reasoning.
- Implemented CLIP from scratch using JAX/Flax.
- Fine-tuned Mistral-7B using QLoRA for instruction-following tasks.

Honours and Awards

- National Grassroots Research Initiative Fund (NGIRI) – Awarded for Bachelor's thesis.
- Government of Pakistan – Awarded a laptop under PMYLS.

References

Kahina Mokrani

Orange Innovation

kahina.mokrani@orange.com

Dr. Hedia Tabia

University of Paris-Saclay

hedi.tabia@univ-evry.fr

Dr. Abdul Attayyab Khan

Bahria University

aakhan.bukc@bahria.edu.pk