Bilal Azdad

Final-Year Computer Science Student focusing on AI/ML & MultiMedia

github linkedin bilal.azdad.pro@gmail.com +33 667810509 Toulouse, France

ABOUT ME

Final-year student with a dual focus in AI for multimodal data (NLP, Vision, Speech) and high-performance scientific computing. Experienced in algorithm design, problem formulation, experimental design, and rigorous model evaluation. **Actively seeking a PhD position** to conduct pioneering research and turn fresh insights into transformative AI applications.

EXPERIENCE

Research Intern – Vision–Language Models | **AIRBUS** — March 2025 – Present

Tools: HuggingFace, Pytorch, GCP, LangChain, Sloth, LLM fine-tuning, TensorRT, vLLM — *Toulouse, France*

- Led research on 5+ SOTA Vision-Language Models, leveraging **HuggingFace** on **GCP** for VLM fine-tuning and evaluation. Developed **LangChain**-based multimodal agents for spatio-temporal video analysis, improving insight extraction by 15%.
- Architected and deployed an end-to-end multimodal search platform on GCP, which automated data pipelines for unstructured data (PDFs, images) and enabled cross-modal queries via a unified vector embedding space.
- Presented research findings to internal stakeholders, collaborated with leading external AI experts.

Research Intern – Neural Network Training GUI | **ALTEN** — Oct 2024 – Feb 2025

Tools: PyTorch, Tkinter, OpenCV, Python, Parallel Computing — *Remote*

- Developed a GUI in **Tkinter** to monitor and steer adversarial neural network (GAN) training pipelines written in **PyTorch**.
- Enabled real-time hyper-parameter tuning (learning rate, optimizer choice) of individual GAN components.
- Implemented comparative dashboards to analyse training runs across architectures and optimisers, increasing experiment throughput by 40%, pipelines validated on the **MNIST** dataset.

Research Intern – Autonomous Driving | **IRIT** — June - September 2024

Tools: Kotlin, Android Studio, OpenCV, Yolo, Python, TensorFlow, MiDaS, Quantization — *Toulouse, France*

- Engineered an Android driver-assistance application integrating **YOLO** for object detection and **MiDaS** for depth estimation.
- Architected on-device image processing pipelines for efficient inference of detection and depth models, applying **quantization** to reduce model size by 30% for mobile deployment.
- Boosted overall perception accuracy by 20% by **fine-tuning Yolov8** on public road datasets.

Mathematics Tutor | **Completude** — 2023 - 2025

Courses: Calculus, Linear Algebra & Matrix Theory, Real & Functional Analysis, Geometry & Topology — *Toulouse, France*

- Tutored **undergrad students** in Linear Algebra and Analysis, crafting tailored problem sets and visual aids.

EDUCATION

ENSEEIHT | **Research Master "Performance for Software, Media and Scientific Computing"** | **Toulouse, France** — 2024 - Present

Courses: HPC, Parallel Computing, Optimization, Distributed Systems and Algorithms, Software Engineering...

ENSEEIHT | **MSc Computer Science** | **Toulouse, France** — 2022 - Present

Courses: AI, Machine Learning, Data Analysis, Multimedia Data Processing (Image, Audio, Text), Cloud...

Lycée Vauvenargues | **Preparatory Classes for Engineering Schools** | **Aix-en-Provence, France** — 2020 - 2022

Courses: Probability, Statistics, Linear Algebra, Physics, ...

PROJECTS

End-to-End Medical Vision-Language Model Fine-Tuning | **Personal Project** — March 2025

Tools: PyTorch, Hugging Face (Transformers, PEFT, Datasets), QLoRA, BitsandBytes, Gradio, W&B

- Fine-tuned a **3-billion-parameter Vision Language Model** for the specialized domain of medical pathology VQA, leveraging the 32k-example PathVQA dataset to improve its diagnostic reasoning capabilities.
- Engineered a **memory-efficient training pipeline using QLoRA and PEFT**, successfully fine-tuning the model on a single 8GB consumer GPU. Implemented a robust evaluation framework, benchmarking performance with **accuracy for closed-form answers and BERTScore for open-ended generation**. Deployed a proof-of-concept interactive demo with Gradio.

Multimodal Data Processing | **Team Project** — 2024-2025

Tools: Matlab, OpenGL, OpenCV, C++, Python, TensorFlow

- Background/Foreground Segmentation: K-means clustering, superpixels with SLIC, morphological refinement.
- Source Separation: Decomposition of spectrograms with NMF, separation using Deep Learning (U-net).
- Image & Video Compression: Implemented JPEG and MPEG codecs (DCT, motion-compensated prediction).

SKILLS SUMMARY

Programming Languages:	Python, Java, C/C++, Matlab, Julia, SQL, TypeScript, React, Shell
Frameworks Tools:	PyTorch, TensorFlow, Scikit-learn, OpenCV, HuggingFace, Docker, Git, Linux, MPI, OpenMP, GCP
Soft Skills:	Problem Solving, Team Work, Communication, Adaptability, Reliability, Rigor
Languages:	French (Native), English (C1), Italian (B1), Arabic (A1), German (A1)