

## Engineering and Professional Experience

---

- |  |  |                                |
|--|--|--------------------------------|
| <b>Research Intern</b>   | <b>ISIR - Sorbonne University</b>            | <b>April 2025 – Sept. 2025</b> |
| <ul style="list-style-type: none"><li>• Designed and developed end-to-end distributed pipelines for fine-tuning of large-scale vision-language models (VLMs) and LLMs (Qwen, Llama, Mistral) using Python, PyTorch, and Hugging Face.</li><li>• Designed, implemented, and evaluated new reinforcement learning algorithms enhancing performance for language and agentic tasks.</li><li>• Built full ML system lifecycle end-to-end: data collection, preprocessing, model training, and inference for VLMs using Python, PyTorch, HF Transformers, slurm</li><li>• Collected and curated 10k+ vision-language-action training examples from an agentic environment.</li></ul>  |  |                                |
| <b>Research Engineer</b>   | <b>L2TI - Sorbonne Paris-Nord University</b> | <b>2023 – 2024</b>             |
| <ul style="list-style-type: none"><li>• Designed, developed, and trained large computer vision models on large scale datasets including COCO, and satellite imageries using <b>Python, Pytorch, Transformers</b> resulting to increased accuracy</li><li>• Developed automated data pipelines processing 5k+ images and 100k annotations, focusing on robustness and fault tolerance using <b>Python, Pytorch, HF Datasets</b> resulted to a research grade dataset</li><li>• Implemented comprehensive testing framework for model architectures ensuring consistency and reliability</li><li>• Led code reviews and pull request management for team of 3-4 research interns, ensuring code quality and maintainability of shared CV model library, and good usage of GPUs</li></ul> |  |                                |
| <b>Research Intern</b>   | <b>L2TI - Sorbonne Paris-Nord University</b> | <b>2023</b>                    |
| <ul style="list-style-type: none"><li>• Fine-tuned and benchmarked Vision transformer model for object recognition task, using <b>Python, and Pytorch</b></li><li>• Designed transformer-based architectures: Implemented Vision Transformer models and multimodal fusion for object recognition in multimodal satellite images.</li></ul>   |  |                                |
| <b>Data Analyst/ML Engineer</b>  | <b>United Nations Environment Programme</b>  | <b>2018 - 2021</b>             |
| <ul style="list-style-type: none"><li>• Developed and deployed classic ML models for geospatial change detection, including temporal modeling for multi-frame analysis</li><li>• Built <b>backend APIs and services</b> in Python FastAPI, including <b>unit tests</b> and data validation logic</li><li>• Created, cleaned, and managed datasets with robust data pipelines using python</li><li>• Developed APIs, and backend with unit tests using python</li></ul>   |  |                                |

## Education

---

- **Master-2 MVA (Math, Vision, Apprentissage)**, ENS Paris-Saclay, Paris, France . **2024–2025**
- **Master Informatics**, University of Toulouse-Capitole, Toulouse, France. **2021–2023**
- **Courseworks:** LLM, Reinforcement Learning, Robotics, Computation statistics, Computer Vision

## Open source, projects, and Publications

---

- **Top 3 in Gadium & Pyannotate iterate hackathon (Dec. 2025)** Implemented a personal voice assistant that manages, and updates tasks, and notes just over a phone call and speech.
- **Contribution to Hugging Face diffusers: (Nov. 2025)** Efficient implementation of an inference mechanism in Hugging Face diffusers, with corresponding benchmarking and test scripts.
- **Realtime meeting copilot:** A web application integrating LLM, Speech models, and RAG for recording, storing, and prompting meeting minutes. Tech: Python, Transformers, Langchain, Typescript
- **Bahaduri, B., Ming, Z., Feng, F., & Mokraoui, A.** Multimodal Transformer Using Cross-Channel attention for Object Detection in Remote Sensing Images. In ICIP (oral), 2024.
- **Bahaduri, B., Talaoubrid, H., Feng, F., Ming, Z., & Mokraoui, A.** Indirect Attention: Turning Context Misalignment into a Feature. Arxiv-preprint, 2025.
- Bahadur, B., et al. CADOT: Cityscape Aerial Image Dataset For Object Detection. ICIP, 2025.