Yohannes Tadesse **Haile**

Research Associate · AI Engineer

Kigali, Rwanda

(+250) 790-394-898 | yohatad123@gmail.com | yohatad | yohannest

Education

CMU (Carnegie Mellon University Africa) — *Kigali, Rwanda*

M.Sc. in Electrical and Computer Engineering — *Sep. 2022 - May. 2023*

Course work: Cognitive Robotics, Introduction to Deep Learning, Applied Computer Vision, AI system Design

AASTU (Addis Ababa Science and Technology University) — *Addis Ababa, Ethiopia*

M.Sc. in Control and Instrumentation Engineering — *Sep. 2019 - June. 2022*

Course work: Non-linear Control, Optimal Control, Multi-Variable and Robust Control

AASTU (Addis Ababa Science and Technology University) — *Addis Ababa, Ethiopia*

B.Sc. in Electro-Mechanical Engineering — *Sep. 2015 - Jan. 2021*

Skills

Programming C/C++, Python, Scipy, Numpy, LaTeX, Shell/Bash Scripting

Software Engineering Tools Git, Docker, Linux, CI/CD, AWS

Experience

Carnegie Mellon University Africa — *kigali, Rwanda*

Research Associate (Culturally Sensitive Social Robotics for Africa (CSSR4Africa)) Project link — *May 2023 - Present*

- Designed adn Deployed LLM-powered applications using Retrieval-Augmented Generation (RAG) for grounded question answering over domain-specific document collections.
- Built end-to-end RAG pipelines including data ingestion, text chunking, embedding generation, vector search, and prompt orchestration.
- Implemented speech-to-text and text-to-speech pipelines using Whisper ASR and neural TTS (XTTS-v2) to enable natural language interfaces.
- Developed multimodal data processing pipelines combining audio and visual signals for context-aware interaction.
- Optimized inference performance and latency for LLM and speech models across GPU and edge environments.
- Integrated LLM, speech, and perception components into production-ready systems with real-time constraints.

Carnegie Mellon University Language Technologies Institute(LTI) — *Pittsburgh, USA (Remote)*

Teaching Assistant (CMU's flagship deep learning course 11-785 (Prof Bhiksha Raj)) Course link — *May 2023 - Dec 2023*

- Guiding over 400 students and mentoring multiple research projects.
- Covered foundational and advanced topics, including MLPs, CNNs, RNNs, Attention Mechanisms, Transformers, Graph Neural Networks (GNNs), and Hopfield Networks, with practical implementation in PyTorch.
- Conducted recitations on Hyperparameter Tuning Methods, Normalizations, and optimization techniques for improving model performance
- Organized and led hackathons and bootcamps for hands-on coding sessions, focusing on key deep learning components like Dataloaders, MLPs, CNNs, and Neural Networks with Attention.
- Delivered a dedicated recitation on Vision Transformers, including theory, applications, and implementation.

Carnegie Mellon University Africa — *Kigali, Rwanda*

Research Associate (Amharic OCR (Optical Character Recognition) Project) (Part-time) — *June 2025 - Present*

- Developing open-source deep learning models for Amharic OCR to improve text recognition accuracy from scanned documents and images.
- Creating large-scale handwritten and synthetic datasets to augment training data, enhancing model robustness and performance.
- Utilizing OCR tools and computer vision libraries such as Tesseract and OpenCV to optimize text extraction pipelines.
- Development of Amharic OCR challenges in collaboration with AI for Good (ITU), supported by a $10,000 compute grant.

Projects

Training a Speech Recognizer from Synthetic Data GitHub — *Jan 2023 – May 2023*

- Designed a synthetic-data training pipeline for ASR using neural TTS and transformer-based speech models.
- Generated large-scale synthetic speech–text pairs to replace costly human-annotated datasets.
- Trained and evaluated ASR models achieving performance comparable to models trained on natural speech.

EEG Foundation Challenge: From Cross-Task to Cross-Subject EEG Decoding GitHub — *Sep 2025 – Dec 2025*

- Built machine learning models for zero-shot transfer across unseen tasks and subjects using EEG data.
- Developed predictors for mental health–related factors from high-density EEG recordings (128 channels).
- Worked with large-scale EEG datasets (700 GB) from 3,000+ participants across multiple experimental paradigms.
- Benchmarked neural network architectures against demographic and regression-based baselines.
- Placed 67th among 1,200+ participants worldwide in the EEG Foundation Challenge.