Ahmed Adel Attia, Ph.D. Researcher

469-596-4371 - ahmadadelattia@gmail.com - http://www.ahmadadelattia.com ahmedadelattia
ahmedadelattia rQ7y81gAAAAJ aadel4

EDUCATION

University Of Maryland — Ph.D. in Electrical and Computer Engineering

2025, Maryland, USA

Conducted research in deep learning-based natural language processing and speech, with courses in these areas as well as unsupervised learning and speech and audio processing

Alexandria University — B.Sc. in Communications And Electronics Engineering

2015, Alexandria, Egypt

GPA: 3.82, Top fifth student in a class of 300.

SKILLS

Programming: Proficient in Python3, including expertise in TensorFlow, TensorFlow Dataset, TensorFlow Profiler, and PyTorch. I am also experienced in C, C++, Matlab, LLVM, Verilog, and VHDL.

Research: Demonstrable research experience in Deep Learning applications including Automatic Speech Recognition (ASR), Speech Processing and Modeling, Real-time Deep Learning Models, Self-Supervised Learning, and Natural Language Processing.

EXPERIENCE

University Of Maryland - Graduate Research Assistant, Lab Manager

January 2022 - PRESENT, Maryland, USA

- Adapting and improving ASR and diarization models for children's speech
- Researching to develop Deep Learning and Machine Learning algorithms for acoustic and articulatory speech data to understand speech production better.
- Authored and published a conference paper (arXiv: 2210.15195) to IEEE ICASSP 2023 within the first 10 months of the position.
- Several other publications in top Speech and Signal Processing conferences

Omnispeech, LLC - Deep Learning Consultant (Remote)

June 2021 - January 2022 Maryland, USA

- Worked on developing lightweight real-time speech enhancement deep learning models.
- Successfully scaled-down large Speech Enhancement GAN models from 37 million parameters to less than 1 Million parameters, maintaining good clarity and noise cancellation.
- Developed an efficient data pipeline using TensorFlow Dataset API and TensorFlow profiler for more than a terabyte of audio data achieving optimal performance and ~ 100% GPU utilization.

University Of Arizona - Deep Learning Research Intern

July 2019 - September 2019, Arizona, USA

- Conducted research on Generative Adversarial Networks (GANs) under the supervision of Prof. Ravi Tandon.
- Achieved over 98% accuracy using Mutual Information Neural Estimators, and unsupervised Outlier Detection systems

PROFESSIONAL ASSOCIATIONS

Affiliation with Stanford University - Sponsored by Professor Dora Demszky, Stanford University

October 2023 - PRESENT

- Affiliated with Stanford University through sponsorship by Professor Dorottya Demszky.

- Involvement in a collaborative grant project focused on the development of automated feedback solutions for teachers in classroom settings.
- Concurrently, served as a Graduate Research Assistant at the University of Maryland.

PUBLICATIONS

Kid-Whisper: Towards Bridging the Performance Gap in Automatic Speech Recognition for Children VS. Adults - arXiv:2309.07927

Under Review: ICASSP 2024

- Achieved SOTA performance on children's speech datasets.
- Improved the WER of Whisper Medium on the popular children's dataset MyST from 13.23% to 8.85%, and CSLU Kids from 31.85% to 16.53%.
- Outlined current challenges in children's ASR, like the adaptation of language models to the variability in children's speech.

Improving Speech Inversion Through Self-Supervised Embeddings and Enhanced Tract Variables - arXiv:2309.09220

Under Review: ICASSP 2024

- Studied the effect of input and output representations on the performance of speech inversion systems.
- Improved the input representation of the acoustic signal using HuBERT self-supervised embeddings.
- Refined previously proposed geometric transformation of articulatory data.
- Improved the performance of the speech inversion system from 74.52% to 81.41%

Audio Data Augmentation for Acoustic to Articulatory Speech Inversion using Bidirectional Gated RNNs - arXiv:2205.13086

EUSIPCO 2023

- Implemented a Bidirectional Gated Recurrent Neural Network (BiGRNN) as the speech inversion system and utilized three data augmentation methods to improve the performance of the system on both clean and noisy speech data.
- Achieved a 5% relative improvement in correlation for clean speech data compared to the baseline noise-robust system.
- Demonstrated an additional 6% improvement in average correlation by adapting the pre-trained model to unseen speakers in the test set.

Enhancing Speech Articulation Analysis Using A Geometric Transformation Of The X-ray Microbeam Dataset - arXiv:2305.10775

Interspeech 2023

- Developed a novel geometric transformation that improves the accuracy of speech articulation measurements in the X-ray Microbeam Dataset (XRMB).
- Proposed extending the palatal trace to include the soft palate and anterior pharyngeal wall, enhancing the calculation of tongue body constriction, particularly in back vowels.
- Addressed limitations of previous models by incorporating additional anatomical landmarks, resulting in a more accurate representation of speech production.

Masked Autoencoders Are Articulatory Learners - arXiv:2210.15195

ICASSP 2023

- Developed a novel masked autoencoder model for reconstructing mistracked speech recordings in the X-ray Microbeam (XRMB) dataset.
- Achieved state-of-the-art performance in reconstructing articulatory parameters, with a 96.4% success rate in reconstructing concurrently mistracked features.
- The main novel contribution of my paper was reconstructing recordings where multiple articulatory features were concurrently mistracked.