

Aditya Pujari

adityapujari.dev

☎ +1-940-595-3193

✉ adityapujari@my.unt.edu

in [adisp](#)

🌐 [pujariaditya](#)

Denton, Texas 76201

Research Interests

Applied Machine Learning in Speech and Audio:

- **Audio Watermarking and Media Authentication:** dual-watermarking frameworks for integrity verification and source attribution; robust embedding under common distortions; fine-grained temporal tamper localization.
- **Audio Deepfake Detection and Source Tracing:** open-set source attribution of synthetic speech; codec-residual and spectral-phase signatures; spoofing countermeasures.
- **Speech Editing and Synthesis:** LLM-guided diffusion for text-based speech editing; zero-shot text-to-speech; conditioning-space identity redirection for speaker unlearning.

Education

University of North Texas (UNT)

Aug 2023 – Present

Ph.D., Computer Science, Denton, TX

University of North Texas (UNT)

May 2023

M.S., Artificial Intelligence, Denton, TX GPA: 3.93/4.00

Nagesh Karajagi Orchid College of Engineering and Technology

June 2020

B.S., Computer Science, Solapur, India Grade: 9.30/10.00

Academic & Industrial Research Experience

University of North Texas – Visual Computing and Biometric Security Lab (VCBSL) *Oct 2023 – Present*

Research Assistant, Denton, TX

Advisor: Dr. Ajita Rattani

- Co-authored WaveVerify, a FiLM-based audio watermarking framework achieving zero bit-error-rate under common distortions (IJCB 2025), and AudioAuth, a dual-watermarking framework for audio integrity verification and source attribution (IEEE TBIOM 2026).
- Built CoRA, an open-set source-tracing system attributing synthetic speech to its generator from codec-residual and spectral-phase signatures: FPR95 0.50% vs. published 3.36% on MLAAD v5 (65 synthesis models).
- Built wavepainter, a multimodal LLM-guided diffusion model for text-based speech editing: substitution WER 3.48 vs. prior best 4.41, outperforming FluentSpeech and VoiceCraft at speaker similarity 0.943/0.961/0.918.
- Invented cond-ID, a conditioning-space identity-redirection edit that unlearns a speaker from deployed zero-shot TTS without retraining; erasure survives white-box re-finetuning (0.48 robustness on XTTS-v2).

Kanarys

June 2022 – Aug 2022

Data Science Intern, Dallas, TX

- Built sales forecasting models with LSTM, XGBoost, and Facebook Prophet, improving system performance by 27%.

Publications

Journals

- [J1] **Aditya Pujari** and Ajita Rattani. “AudioAuth: A Dual-Watermarking Framework for Robust Audio Integrity and Source Attribution.” *IEEE Transactions on Biometrics, Behavior, and Identity Science*, pp. 1–15, January 2026. [DOI] [code]

Peer-Reviewed Conferences

- [C1] **Aditya Pujari** and Ajita Rattani. “WaveVerify: A Novel Audio Watermarking Framework for Media Authentication and Combatting Deepfakes.” *Proc. IEEE International Joint Conference on Biometrics (IJCB)*, Osaka, Japan, September 8–11, 2025. [arXiv] [code]

Selected Projects

CoRA: Codec-Residual Open-Set Source Tracing for Audio Deepfakes 2026

Python, WavLM-Large, EnCodec, Open-Set Attribution [code]

- Designed an open-set source-tracing system for audio deepfakes: a 288-d embedding fusing frozen WavLM-Large features, a 52-d spectral phase/modulation signature, and EnCodec-24 kHz codec residuals, scored with anchor-margin and relative-Mahalanobis rules to name the synthesis model or report “unknown”.
- Set a new bar on MLAAD v5 (65 known models, model-held-out open set): FPR95 0.50% vs. published 3.36%, OOD-EER 2.45%, 99.86% known-model accuracy; 6.80% EER cross-corpus on held-out models.

Wavepainter: Multimodal LLM-Guided Diffusion for Text-Based Speech Editing 2026

Python, Diffusion Models, Multimodal LLM, Speech Editing [code]

- Built text-based speech editing with a multimodal LLM-guided diffusion model: editing a word re-predicts only that mel-spectrogram span, preserving the rest (PESQ 4.090); one checkpoint handles substitution, insertion, and deletion.
- Substitution WER 3.48 on Ming-Freeform-Audio-Edit vs. prior best 4.41, beating FluentSpeech, Voice-Craft, and Ren et al.; speaker similarity 0.943/0.961/0.918 across tasks.

Cond-id-tts-unlearning: Conditioning-Space Identity Redirection for Speaker Unlearning 2026

Python, XTTS-v2, Tortoise-TTS, IndexTTS-1.5, Projected-Gradient Editing [code] [demo]

- Invented speaker unlearning for deployed zero-shot TTS: a few-step projected-gradient edit on the speaker-conditioning path only, exported as a small adapter patch – no retraining needed for opt-outs.
- One edit ports across XTTS-v2, Tortoise-TTS, and IndexTTS-1.5; VCTK 20-speaker pool (5 splits $\times$ 3 seeds): erasure 0.44–0.71, retain-keep 0.89–0.92, retain-WER shift +0.00 to +0.03.
- Erasure resists white-box re-finetuning (0.48 robustness on XTTS-v2).

AudioAuth: Dual-Watermarking for Audio Integrity and Source Attribution 2026

Python, PyTorch, Audio Watermarking, Frequency-Partitioned Encoding [code]

- Designed frequency-partitioned dual watermarking: a model watermark in even bands for source attribution plus a dynamic data watermark in odd bands for fine-grained temporal tamper localization.
- Built a three-stage training pipeline with inference via `infer.py`, CLI, and Python API; watermarked audio holds 20–22.5 dB SI-SNR, keeping the watermark imperceptible.

WaveVerify: A Novel Audio Watermarking Framework for Media Authentication 2024 – 2025

Python, PyTorch, FiLM, Mixture-of-Experts, Audio Watermarking [code]

- Designed a FiLM-based generator with parallel hierarchical modulation for robust multiband watermark embedding, cutting training time by 80% vs. sequential bottleneck approaches.
- Built a Mixture-of-Experts detector reaching zero bit-error-rate under common distortions and 0.98+ MIoU under severe temporal modifications.
- Created a unified training framework with a dynamic effect scheduler applying multiple distortions in a single backpropagation step; paper accepted at IJCB 2025 (see [C1] above).

Skills

Programming: Python (PyTorch, TensorFlow), JavaScript/TypeScript, $\text{\LaTeX}$

ML & Data: scikit-learn, XGBoost, LightGBM, NumPy, pandas, SciPy, Polars

Cloud & MLOps: AWS (SageMaker, EC2, S3), Google Cloud (Vertex AI), Azure ML, Docker, MLflow, Weights & Biases

LLMs & Foundation Models: fine-tuning, LoRA/QLoRA, PEFT, RAG, RLHF (Llama, Qwen)

Deployment: FastAPI, Streamlit, Gradio, ONNX, TensorRT, vLLM

Certifications

- Oracle Cloud Infrastructure 2024 Generative AI Certified Professional [\[link\]](#)
- Microsoft Certified: Azure Fundamentals [\[link\]](#)
- AWS Certified: Cloud Practitioner [\[link\]](#)
- Machine Learning – Stanford University (Coursera) [\[link\]](#)
- Deep Learning Specialization – DeepLearning.AI (Coursera) [\[link\]](#)
- AWS Certified: Solutions Architect – Associate [\[link\]](#)