Snehit Chunarkar

ML Researcher | Noise Robust Speech Emotion

(+886) 0965742974 Email LinkedIN Website GitHub Hsinchu, Taiwan

Summary

ML Researcher focused on noise-robust speech emotion recognition and multimodal audio-text learning | PhD at NTHU's BIIC Lab under Prof. Chi-Chun Lee (Jeremy).

Research focus: speech/audio representation learning, emotion understanding, and cross-modal fusion (audio–text), with an emphasis on robust modeling and interpretability.

Experience with top-tier models: WavLM, Wav2Vec2, CLAP, Whisper, RoBERTa. Globally #2 in ICASSP 2026 SP Grand Challenge (XACLE) with 92% better prediction; led a noise-robust speech emotion recognition system with 32% improvement.

Education

PhD, Electrical Engineering (2027 expected)
National Tsing Hua University (NTHU), Hsinchu, Taiwan

M.Tech., Instrumentation and Signal Processing
Indian Institute of Technology (IIT), Roorkee, India

B.E., Instrumentation Engineering
Government College of Engineering, Chandrapur, India

Research Advisor

Prof. Chi-Chun Lee (Jeremy)
Professor & Associate Chair, NTHU; Research Fellow, Academia Sinica; Director, NVIDIA–NTHU Joint Innovation Center

Publications

[Accepted] Reasoning Driven Captions to Assist Noise Robust Speech Emotion Recognition
IEEE ICASSP 2026 (Barcelona, Spain)
Snehit B. Chunarkar, Chi-Chun Lee

[Accepted] Cross-Modal Semantic Alignment Via Ensemble Audio-Text Features for XACLE Challenge
IEEE ICASSP Workshop Proceedings 2026 (Barcelona, Spain)
Snehit B. Chunarkar, Krishnagiri Hamza, Chi-Chun Lee

STELIN-US: A Spatio-Temporally Linked Neighborhood Urban Sound Database
Detection and Classification of Acoustic Scenes and Events 2023 Workshop (DCASE2023) (Tampere, Finland)
S. Chunarkar, B. Su, C. Lee

Mixed Language Separation Using Deep Neural Network
ICEECCOT 2021, (Mysuru, India)
S. Chunarkar, S. Chiluveru, M. Tripathy

Efficient Hardware Implementation of Nonlinear Activation Function For Inference Model
International SoC Design Conference (ISOCC) 2024, (Sapporo, Japan)
S. Chunarkar, S. Chiluveru

Projects

Reasoning-driven-SER: (Speech Emotion Recognition in Noisy Environment)

ICASSP 2026 | Most Competitive year with only 40% acceptance rate

Integrating Reasoning-enriched caption and Audio using Cross-Attention to enhance Speech Emotion Recognition (SER) in Noisy Environment.

Improved SER by 32% over the conventional Audio-Only approach at SNR: -10dB.

Tech stack: Python, PyTorch, Hugging Face Transformers, Zenodo

---

EnsembleSVR-XACLE: (Audio-Text Alignment Prediction)

ICASSP 2026 SP Grand Challenge: XACLE | Rank #2 (Globally)

Engineered end-to-end pipeline to predict audio-text alignment scores in multimodal datasets

Ranked #2 on official leaderboard among 40+ submissions

Tech stack: Python, PyTorch, Hugging Face Transformers, scikit-learn

---

STeLiN-US Database (Dataset synthesis for real-world SED)

DCASE Publication

Synthesized a dataset to mimic real-world acoustic environments (overlapping events, spatial distance scaling, temporal correlation) for training sound event detection systems.

Tech stack: Python, PyTorch, scikit-learn, pydub.

Pawpularity Contest: (Cuteness Score Prediction for Pet's Image)

NTHU Machine Learning Course Project (11230EE 655000) | Rank #2 in class

Built ensemble regression model (Gradient Boosting, SVR, Neural Networks) to predict pet adoption likelihood from image and metadata features

Ranked #2 among 16 course teams; with state-of-the-art performance on Pawpularity dataset

Tech stack: Python, PyTorch, timm, scikit-learn, SciPy, cuML (GPU-accelerated)

Skills

Programming Languages: Python (95%), Bash (50%), Markdown (85%)

Machine Learning & Deep Learning: PyTorch, NumPy, Pandas, Matplotlib, scikit-learn, Hugging Face Transformers

Specialized Tools: Audio processing (librosa, pydub), Signal processing (SciPy), GPU computing (CUDA, cuML), Dataset platforms (Kaggle, Zenodo)

Development Environments: VS Code, Spyder, Jupyter Notebook, MATLAB, MobaXterm

Cloud & HPC: Google Colab, TWCC (Taiwan Computing Cloud)

Additional Information

Languages: English (fluent), Hindi (native), Marathi (native)

Active Engagement: Open-source GitHub, Zenodo contributions