Shreya G. Upadhyay

✉ shreya@gapp.nthu.edu.tw | ✉ shreyaupadhyay10@gmail.com | 📱 +886 905763774

in Linkedin | 🎓 GoogleScholar | R⁶ ResearchGate | ⓘD ORCiD | 🌐 HomePage | ⬤ Git

RESEARCH INTEREST

Behavioural Speech Signal Processing, Machine Learning, Speech Emotion Recognition, Affective Computing, Sound Event Detection, Automatic Speech Recognition, Cross-Corpus Transfer Learning

EDUCATION

National Tsing Hua University, Taiwan   2019/09 - 2025/05
PhD in the Department of Electrical Engineering

K.J. Somaiya College of Engineering, Mumbai, India   2016/06 - 2018/10
MTech in the Department of Computer Engineering

Mumbai University, Mumbai, India   2009/06 - 2013/07
BE in the Department of Computer Engineering

WORK EXPERIENCE

Applied Scientist II, Amazon, Canada   2025/06 - Present

RESEARCH EXPERIENCE

Behavioral Information and Interaction Computation Lab (BIIC), NTHU   2019/09 - Present
Speech Emotion Recognition
– Pretrained Layer-Constrained Cross-Lingual SER (INTERSPEECH 24)
– Utilizing Rater-Expanded Label Space for Consensus-based SER (TAFFC 24)
– Affective Priming on Emotional Annotations (ACII 23)
– Phonetic-Anchor based Cross-Lingual SER (ICASSP 23)
– Monolog v/s Conversation Speech Data Analysis (ACII 22)

Trustworthy SER Systems
– Gender Fairness Generalizability in Cross-Corpus Scenario (ICASSP 25)
– Balancing Speaker-Rater Fairness for Gender-Neutral SER (ICASSP 24)

Speech Signal Processing
– Adversarial Attack on One-shot Voice Conversion (INTERSPEECH 24)
– Noise-Robust Bandwidth Expansion with Diffusion Process (INTERSPEECH 24)
– Implementing the Taiwanese Mandarin Forced Aligner
– Utilizing Large Speech Production SSL Models

Sound Event Detection
– Phoneme-Level Acoustic Learning for Rare Sound Event Detection (INTERSPEECH 20)
– Sound Event Localization and Detection

Affective Multimedia
– Integrating Annotator's Subjectivity in Multimedia Emotion Recognition System
– Modelling Multimodal System with Verbal and Non-Verbal Human Sound Events (EUSIPCO 22)

Dataset Collection and Creation
– Taiwanese Mandarin Affective Podcast Dataset Collection (ACII 23)
– Synthetic Scene and Event Background Noise Dataset Creation for ASR

K.J. Somaiya College of Engineering (KJSCE), India 2016/06 - 2018/10
Affective and Behavioural Computing
– Attention-based Modelling for Accent Classification
– Dataset Collection Interface Creation for Judgement Questionnaire
– Early-Later Stage Confidence Behaviour Prediction of Online Platform Learners (ICCUBEA 18)

Cooperated Research & Projects

School of Computer Science, Carnegie Mellon University, USA (Remote) 2024/08 - 2025/05
– Mouth-Gesture Anchoring for Speech Emotion Recognition

University of Texas at Dallas, USA (Remote) 2021/07 - 2024/12
– Affective Mandarin Taiwanese Podcast Database Collection
– Phonetic-based Cross-Lingual Speech Emotion Recognition
– Effect of Affective Priming on Emotional Annotations
– Monolog v/s Conversational Speech on Recognizing Emotion

C-Media Electronics Incorporation (C-Media Inc.), Taiwan 2021/02 - 2021/07
– Implement AI De-Reverberation De-Noise Algorithm Using Deep Noise Suppression

Internships

Amazon, AGI Applied Scientist Intern, USA 2024/09 - 2024/12
– LLM Multi-task Learning
– LoRA Merging/Mixing for efficient
  task scalability

Industrial Technology Research Institute (ITRI), Taiwan 2022/06 - 2022/12
– Golf Swing Action Detection and Confidence Prediction
– Event Detection in Baseball Matches Using Audio-Only Features

Research Innovation Incubation Design Labs (Riidl), India 2018/07 - 2019/02
– Live Video Object Detection
– Speech Acoustic Analysis

Teaching Assistance

Introduction to Digital Signal Processing, NTHU	2022/02 - 2022/07
Introduction to Machine Learning, NTHU	2021/09 - 2022/01
Machine Learning, KJSCE	2017/02 - 2017/06
Design and Analysis of Algorithm, KJSCE	2016/08 - 2017/01

Awards & Honors

ICASSP Travel Grant, 2025
National Science and Technology Council (NSTC) Travel Grant, 2023 - 2025
NTHU International Student Scholarship, 2019 - 2024
CTCI Foundation Research Scholarship Award, 2023
ICASSP PROGRESS Award, 2023
Google Travel Grant, 2023 - 2024
Poster in Women in Machine Learning (WiML) Workshop, NeurIPS, 2022
ACII Student Travel Grant, 2022
Session Chair, 30th European Signal Processing Conference (EUSIPCO), 2022
Merry Electroacoustic Paper Award (Finalists), 2020
Poster in YFRS Workshop, Interspeech, 2018

Skills

Corpora Collection	BIIC-Podcast (Affective Corpus)
Speech Signal Processing Tools	openSMILE, librosa, Praat, Fairseq
Speech Production Models	Wav2vec2.0, Hubert, WavLM, Whisper
Computer Vision Tools	OpenCV, OpenFace
ASR Tools	Kaldi, Forced Aligners (MFA)
Programming	Python, C++, MATLAB, HTML
Machine Learning Libraries	Pytorch, TensorFlow, Docker, scikit-learn
Other Tools & Familiar OS	LaTex, Git (Github and Gitlab), Linux, Windows

Reviewer Experience

Interspeech

International Conference on Affective Computing and Intelligent Interactions (ACII)

ACM International Conference on Multimodal Interaction (ICMI)

IEEE Transactions on Affective Computing

IEEE Access

APSIPA Transactions on Signal and Information Processing

Research Funding Acknowledgment

Publications

Journal Paper

– **Upadhyay Shreya G.**, Luz Martinez-Lucas, William Katz, Carlos Busso, and Chi-Chun Lee. "Phonetically-Anchored Domain Adaptation for Cross-Lingual Speech Emotion Recognition." IEEE Transactions on Affective Computing. 2025. 🔗

– Luz Martinez-Lucas, Ali Salman, Seong-Gyun Leem, **Upadhyay Shreya G.**, Chi-Chun Lee, and Carlos Busso. "Affective Priming in Emotional Annotations and its Effect on Speech Emotion Recognition." IEEE Transactions on Affective Computing, 2025.

– **Upadhyay Shreya G.**, Woan-Shiuan Chien, Bo-Hao Su, and Chi-Chun Lee. "Learning With Rater-Expanded Label Space to Improve Speech Emotion Recognition." IEEE Transactions on Affective Computing, 2024. 🔗

– Woan-Shiuan Chien, **Upadhyay Shreya G.**, Wei-Cheng Lin, Carlos Busso, and Chi-Chun Lee. "Differential Impacts of Monologue and Conversation on Speech Emotion Recognition." IEEE Transactions on Affective Computing, 2024. 🔗

Conference Paper

– **Upadhyay Shreya G.**, Ali N. Salman, Chi-Chun Lee, and Carlos Busso. "Mouth Articulation-Based Anchoring for Improved Cross-Corpus Speech Emotion Recognition". IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2025. 🔗

– **Upadhyay Shreya G.**, Woan-Shiuan Chien, Chi-Chun Lee. "Is It Still Fair? Investigating Gender Fairness in Cross-Corpus Speech Emotion Recognition". IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2025. 🔗

– Bo-Hao Su, **Upadhyay Shreya G.**, and Chi-Chun Lee. "Toward Zero-Shot Speech Emotion Recognition by Utilizing LLMs in the Absence of Target Data" IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2025. 🔗

– **Upadhyay Shreya G.**, Carlos Busso, and Chi-Chun Lee. "A Layer-Anchoring Strategy for Enhancing Cross-Lingual Speech Emotion Recognition." Interspeech, 2024. 🔗

– Ching-Yu Yang, **Upadhyay Shreya G.**, Ya-Tse Wu, Bo-Hao Su and Chi-Chun Lee. "RW-VoiceShield: Raw Waveform-based Adversarial Attack on One-shot Voice Conversion." Interspeech, 2024. 🔗

- Yin-Tse Lin, **Upadhyay Shreya G.**, Bo-Hao Su and Chi-Chun Lee. "SWiBE: A Parameterized Stochastic Diffusion Process for Noise-Robust Bandwidth Expansion." Interspeech, 2024. 🔗
- Woan-Shiuan Chien, **Upadhyay Shreya G.**, and Chi-Chun Lee. "Balancing Speaker-Rater Fairness for Gender-Neutral Speech Emotion Recognition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024. 🔗
- **Upadhyay Shreya G.**, Woan-Shiuan Chien, Bo-Hao Su, Lucas Goncalves, Ya-Tse Wu, Ali N. Salman, Carlos Busso, and Chi-Chun Lee. "An Intelligent Infrastructure Toward Large Scale Naturalistic Affective Speech Corpora Collection." Proc. 11th International conference on affective computing and intelligent interactions (ACII), 2023. 🔗
- Luz Martinez-Lucas, Ali Salman, Seong-Gyun Leem, **Upadhyay Shreya G.**, Chi-Chun Lee, and Carlos Busso. "Analyzing the Effect of Affective Priming on Emotional Annotations." Proc. 11th International conference on affective computing and intelligent interactions (ACII), 2023. 🔗
- **Upadhyay Shreya G.**, Luz Martinez-Lucas, Bo-Hao Su, Wei-Cheng Lin, Woan-Shiuan Chien, Ya-Tse Wu, William Katz, Carlos Busso, and Chi-Chun Lee. "Phonetic Anchor-Based Transfer Learning to Facilitate Unsupervised Cross-Lingual Speech Emotion Recognition." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023. 🔗
- Woan-Shiuan Chien, **Upadhyay Shreya G.**, Wei-Cheng Lin, Ya-Tse Wu, Bo-Hao Su, Carlos Busso, and Chi-Chun Lee. "Monologue versus Conversation: Differences in Emotion Perception and Acoustic Expressivity." Proc. 10th International conference on affective computing and intelligent interactions (ACII), 2022. 🔗
- **Upadhyay Shreya G.**, Bo-Hao Su, and Chi-Chun Lee. "Improving Induced Valence Recognition by Integrating Acoustic Sound Semantics in Movies." European Signal Processing Conference (EUSIPCO), 2022. 🔗
- **Upadhyay Shreya G.**, Bo-Hao Su, and Chi-Chun Lee. "Attentive Convolutional Recurrent Neural Network Using Phoneme-Level Acoustic Representation for Rare Sound Event Detection." Interspeech, 2020. 🔗
- **Upadhyay, Shreya G.**, and Kavita M. Kelkar. "Predicting Learner's Confidence from their Behaviour Using a Judgement Questionnaire." Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). IEEE, 2018. 🔗

Accepted Abstract

- **Upadhyay Shreya G.**. "Articulatory Gesture-Constrained Transfer Learning for Cross-Corpus Speech Emotion Recognition." ISCA-SAC Doctoral Consortium, Interspeech, 2024.
- **Upadhyay Shreya G.**, Bo-Hao Su, and Chi-Chun Lee. "Improving Induced Valence Recognition by Integrating Acoustic Sound Semantics in Movies." Women in Machine Learning (WiML) workshop, NeurIPS, 2022.
- **Upadhyay, Shreya G.**, and Kavita M. Kelkar. "Accent Classification using Audio-based Features." 2018 Young Female Researchers in Speech Science and Technology workshop, Interspeech, 2018.