

Shreya G. Upadhyay

Doctoral Student, EE, National Tsing Hua University, Taiwan

shreyaupadhyay10@gmail.com | shreya@gapp.nthu.edu.tw

+886-905763774

Research Interests

Behavioural Speech Signal Processing, Machine Learning, Affective Multimedia, Sound Event Detection and Scene Analysis, Automatic Speech Recognition, Cross-corpus Transfer Learning

Education

National Tsing Hua University (NTHU), Taiwan

PhD in Department of Electrical Engineering

2019/09 – present

K.J. Somaiya College of Engineering, Mumbai, India

MTech in Department of Computer Engineering

2016/06 – 2018/10

Mumbai University, Mumbai, India

BE in Department of Computer Engineering

2009/06 – 2013/07

Research Experience

Behavioral Information & Interaction Computation Lab (BIIC), NTHU

2019/09 – present

Speech Emotion Recognition

- Monolog v/s conversation speech data analysis and modelling to argue that the conversation samples are more suitable for building emotional corpora
- Utilizing rater-expanded label space to integrate the not only label but also the rater ambiguity in learning to improve consensus-based speech emotion recognition

Cross-Corpus Transfer Learning

- Use phones as an anchor to transfer the emotion from one language data to another language data in a cross-lingual SER setting
- Use an off-the-shelf voice converter to handle complex cross-domain transfer learning by converting source domain data to resemble the target domain

Automatic Speech Recognition

- Code-switching ASR in low-resource languages
- Building an architecture that takes advantage of integrating ASR network representations as additional input when training an acoustic sound event detector

Sound Event Detection and Scene Analysis

- Integrating Phoneme-Level Acoustic Representation for Rare Sound Event Detection
- Performing sound event localization and detection by transfer learning from sound detection model to localization.

Affective Multimedia

- Considering and modelling the annotator's perception and its effect on each other by cross-modal attention for fairness in multimodal multimedia emotion recognition system
- Explicit modelling of verbal and non-verbal human sound events with conventional audio-visual context multimedia emotion recognition model by using Acoustic sound events cues and scene analysis

Dataset Collection and Creation

- Chinese affective podcast emotion AI dataset
- Synthetic synthesized scene and event background noise dataset for ASR with time information

K.J. Somaiya College of Engineering, Mumbai

2016/06 – 2018/10

Affective and Behavioural Computing

- Attention-layer modelling-based Accent classification model for learning different accents of the English language
- Involved in creating dataset collection interface for Judgement questionnaire in self-learning platforms using HTML and CSS
- Early-later stage confidence behaviour analysis of online platform learners using machine learning models by analysing learners' behaviour during the learning and also in task performance

Teaching Assistant Experience

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

Cooperated Projects

UT Dallas MSP-Lab	2021/07 – present
• Chinese podcast database collection in emotion recognition context	
C-Media Electronics Incorporation (C-Media Inc.)	2021/02 – 2021/07
• Implement AI de-reverberation de-noise algorithm based on deep noise suppression	

Awards

Scholarships

- NTHU International Student Scholarship 2022/09 – 2023/08
- NTHU International Student Scholarship 2021/09 – 2022/08
- NTHU International Student Scholarship 2020/09 – 2021/08
- NTHU International Student Scholarship 2019/09 – 2020/08

Honors

- Paper selected for Women in Machine Learning (WiML) workshop at NeurIPS, 2022
- Session Chair at 30th European Signal Processing Conference (EUSIPCO), 2022
- Merry Electroacoustic Paper Award 2020 (Finalists)
- Paper selected for Young Female Researchers in Speech Science and Technology workshop (YFRS) at Interspeech, 2018

Skills

Languages Python, C++, MATLAB, HTML

Libraries & toolkits Pytorch, TensorFlow, Kaldi, Docker, ESPNet

Internship

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

Publications

Journal

- **Upadhyay Shreya G**, Woan-Shiuan Chien, Bo-Hao Su, and Chi-Chun Lee. "Learning with Rater-Expanded Label Space to Improve Consensus-based Speech Emotion Recognition." *IEEE Transactions on Affective Computing*.
- Woan-Shiuan Chien, **Upadhyay Shreya G.**, Wei-Cheng Lin, Carlos Busso, and Chi-Chun Lee. "Monologue versus Conversation: Impact of Emotion Perception and Acoustic Variability on Speech Emotion Recognition." *IEEE Transactions on Affective Computing*.

International Conference

- **Upadhyay Shreya G.***, Woan-Shiuan Chien*, Bo-Hao Su, Carlos Busso, and Chi-Chun Lee. "Towards a Consortium of Naturalistic Affective Speech Corpora through an Intelligently-Controlled Framework." Proc. 11th International conference on affective computing and intelligent interactions (ACII), 2023. (under review)
- 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. (under review)
- **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.
- **Upadhyay, Shreya G.**, Bo-Hao Su, and Chi-Chun Lee. "Improving Induced Valence Recognition by Integrating Acoustic Sound Semantics in Movies." In *Women in Machine Learning WiML workshop (NeurIPS)*, 2022. (Poster)
- 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." *30th 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." Proc. *Interspeech*, 2020: 3102-3106.
- **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.
- **Upadhyay, Shreya G.**, and Kavita M. Kelkar. "Accent Classification using Audio-based Features." In *2018 Young Female Researchers in Speech Science and Technology workshop (Interspeech)*, 2018. (Poster)