

SUMMARY

- **My North Star: Building Baymax for Mental Health**
Design an emotionally intelligent, speech- and behavior-based digital companion that can sense, respond, and support individuals experiencing mental distress—especially the elderly or underserved populations
- Expertise Field: Signal Processing & Computational Linguistics and Speech
- Research interest: Speech Emotion Recognition (SER), Audio-Visual Emotion Recognition (AVER), multi-modal (speech & text) Automated Deception Detection (ADD), Automatic Speech Recognition (ASR), Speech Large Language Models (LLMs), Text-to-Speech (TTS); Bias and Fairness Evaluation
- First author of top-tier conferences: 1*ACL, 2*SLT, 3*ICASSP; 3*INTERSPEECH
- First author of top-tier journal: 1*IEEE Transactions on Affective Computing
- Co-Lead of top-tier conferences: 1*ICASSP, 2*INTERSPEECH; 3*ASRU
- Ongoing projects: Speech Large Language Models (LLMs) with emotions, fairness and robustness measure of SER and AVER Systems, emotional text-prompt and generated speech alignment evaluations of Text-to-Speech (TTS) models, multilingual SER benchmark
- Selected project: EMO-SUPERB (the first SER benchmark)
- Keywords: affective computing, speech/physiological signal processing, machine/deep learning application, fairness and bias, trustworthy machine

WORK EXPERIENCE

- **Postdoctoral Scholar – Visiting Fellow|University of Southern California** Los Angeles, California, USA
Signal Analysis and Interpretation Laboratory (SAIL Lab) Advisor: Prof. Shri Narayanan Oct. 2025 – Present
 - Research conducted as a recipient of the prestigious Postdoctoral Research Abroad Program fellowship from the National Science and Technology Council (NSTC), Taiwan.
- **Visiting Scholar|National Taiwan University** Taipei City, Taiwan
Speech Processing and Machine Learning Laboratory (SPML Lab) Advisor: Prof. Hung-yi Lee Sep. 2025 – Oct. 2025
 - Working on bias and fairness in speech emotion recognition task.
- **Mandatory Military Service|Ministry of the Interior, Taiwan Government** Taipei City, Taiwan
Department of Social Welfare, Nangang Social Welfare Center, Taipei City Government Sep. 2024 – Sep. 2025
 - Served as a social work assistant supporting frontline welfare services and community outreach for elderly individuals, persons with disabilities, and other vulnerable populations.
 - Completed Emergency Medical Technician-1 (EMT-1) certification and trained as a Disaster Relief Volunteer.
 - Participated in basic military training, including live-fire exercises with Type-65 rifle.
 - Maintained research activity as an **Independent Researcher**, collaborating with students from Prof. Hung-yi Lee’s Lab at National Taiwan University on deep learning and speech-related projects.
 - Served as translator and staff member for the **World Masters Games 2025**, providing multilingual support (Mandarin Chinese, English, Taiwanese Taigi).

- Computational Resource Measure and Development of Atmosphere Detection | RealTek** Hsinchu City, Taiwan
Applied ML/DL and DSP Engineer Intern, CTC, Emerging Tech/Advanced DSP Tech Jun. 2024 – Sep. 2024
 - ✓ Engineered streaming audio emotion recognition feature for monitor devices, utilizing Python and TensorFlow for real-time signal processing, improving device analytics capabilities.
 - ✓ Conducted computational resource estimation for neural network-based KWS and SER systems, leveraging profiling tools like PyTorch Profiler to optimize resource allocation and system efficiency.
 - ✓ Optimized state-of-the-art SER and KWS systems by compressing model sizes (30% reduction) with attention to FLOPs, MACs, and parameter constraints, achieving deployable solutions for low-latency environments.
- Multi-task ASR and SER System | Amazon AGI Alexa** Pittsburgh, Pennsylvania, USA
Applied Scientist Intern, Alexa Speech, Alexa Hybrid Science, Alexa ASR Aug. 2023 – Nov. 2023
 - ✓ Pioneered one of the first unified models for ASR and multi-label SER using a shared-layer architecture, reducing system complexity and computational load by 20%.
 - ✓ Identified transcription accuracy gap in OpenAI's WHISPER model, showing 15% higher error rates on emotional vs. neutral utterances, informing ASR system improvements for emotional speech.
 - ✓ Developed a two-stage fine-tuning approach for the WHISPER model in ASR and SER multitasking, achieving an 8.85% absolute WER reduction by mitigating catastrophic forgetting through customized training schedules.
- Taiwan Subsidiary Startups Operation | JOOLA** Taipei City, Taiwan
Business Advisor Mar. 2023 – July 2023
 - ✓ Develop Pickleball and Table Tennis Marking and Promotion Systems (sales figures increases by more than 300%)
 - ✓ Cooperate with and Sponsor Taiwan National Table Tennis Teams (FISU World University Games: Man Team: Silver Medal; Woman Team: Bronze Medal)
 - ✓ Promote Pickleball sport in Taiwan and research project with NTHU
- Automatic Meeting Summarization System | Industrial Technology Research Institute (ITRI)** Hsinchu, Taiwan
Research Assistant Dec. 2018 – Dec. 2019
 - ✓ Propose a summarization system using a deep learning model (transformer and BERT) from spoken sentences to summarizations under multi-party meeting scenarios
- Research Assistant at BIIC Lab | NTHU** Hsinchu, Taiwan
Graduate Research Assistant Feb. 2016 – Jul. 2024
 - Planned, designed, and collected the first dyadic interaction deception database in Taiwanese Mandarin (DDDM), supporting research in automated deception detection [1].
 - Led the creation of the NNIME database, the first dyadic interaction emotion corpus in Taiwanese Mandarin, enabling research in speech emotion recognition and multimodal interaction [2].
 - Ensured data integrity, annotation consistency, and ethical compliance across multiple datasets and projects, including EMO-SUPERB [3], NNIME [2], and NTUBA [4].
 - Designed and managed outreach efforts, including the development of the NNIME project website and the BIIC Lab Facebook page to promote lab activities and datasets.
 - Mentored junior lab members, guiding Master's and Ph.D. students through thesis research, experimental design, and academic publishing.
 - Co-authored and supported the publication of multiple peer-reviewed conference papers and journal articles in the field of speech and affective computing.
- Discover Modern Polarity Terms Algorithm | ITRI** Hsinchu, Taiwan
Research Intern Jul. 2015 – Sep. 2015
 -

- Developed rule-based algorithms to discover modern polarity terms by natural language processing (NLP) from numerous BBS, Yahoo, and Twitter Mandarin Chinese (Taiwan) online articles

- **Research Assistant at BIIC Lab|NTHU** Hsinchu, Taiwan
Undergraduate Research Assistant Feb. 2014 – Feb. 2016
 - Planned, designed, and collected the first dyadic interaction emotion database in Taiwanese Mandarin (NNIME), supporting research in affective computing and multimodal communication [2].
 - Led participant recruitment and data collection involving 2 years in naturalistic and semi-structured interaction settings.
 - Managed multimodal data recording (audio, video, and facial expressions) and annotation using tools such as Praat, Opensmile.
 - Designed and developed the NNIME project website and BIIC Facebook page to disseminate research, enhance public visibility, and support academic collaboration.

EDUCATION

- **National Tsing-Hua University (NTHU)** Hsinchu City, Taiwan
Ph.D. in the Electrical Engineering Department; GPA: 3.82/4.0 Feb. 2016 – Aug. 2024
 Advisor: Prof. Chi-Chun Lee, Behavioral Informatics & Interaction Computation Laboratory (BIIC lab)
 Research Area: automatic speech emotion recognition, automatic deception detection, group & team dynamics
Thesis Title: Revisiting Modeling and Evaluation Approaches in Speech Emotion Recognition: Considering Subjectivity of Annotators and Ambiguity of Emotions (ACLCLP Doctoral Dissertation Award – Honorable Mention)
- **The University of Texas at Dallas (UTD)** Texas, USA
Visiting Ph.D. Student in the Electrical and Computer Engineering Department Mar. 2021 – Sep. 2022
 Advisor: Prof. Carlos Busso, Multimodal Signal Processing Laboratory (MSP lab)
 Research Area: model subjectivity and ambiguity of emotional annotations in speech emotion recognition
- **NTHU** Hsinchu City, Taiwan
B.S. in the Electrical Engineering Department; GPA: 3.04/4.0 Sep. 2011 – Jan. 2016
 Advisor: Prof. Chi-Chun Lee, Behavioral Informatics & Interaction Computation Laboratory (BIIC lab)
 Independent Study: Recognize the Resonance of Positive Emotion Perception From the Crosstalk Recordings
 Relevant Coursework: Digital Signal Processing Laboratory, Embedded System Laboratory, Operating Systems, Data Structures, Psychology and Modern Life; Social and Cultural Analysis

ACADEMIC SERVICE

- **Invited Reviewer:**

Journal: IEEE IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP) (2025), IEEE Transactions on Affective Computing (2022-2025), ELSEVIER Information Processing and Management (2025), ELSEVIER Pattern Recognition (2024), ELSEVIER Computer Speech and Language (2022), APSIPA Transactions on Signal and Information Processing (2022); IEEE Transactions on Multimedia (2021)

Conference: ACM Multimedia (2024), IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP) (2022-2023); Conference of the International Speech Communication Association (INTERSPEECH) (2022-2025)
- **Committee: International Speech Communication Association (ISCA) Student Advisory Committee (SAC) (2023-2024), IEEE Spoken Language Technology Workshop (2024)**

Research interest:

Speech Emotion Recognition (SER) and Automatic Deception Detection (ADD)

- **Research in Speech Emotion Recognition (SER) | MSP Lab@UTD | BIIC Lab@NTHU** Jul. 2015 – Aug. 2024
PhD works *Speech*

- **Abstract:** Implantation of algorithms and deep learning approaches to improve performance of SER models based on public benchmark corpus and real-world audio data **considering the ambiguous emotions, subjective perception, imbalanced labels, calibration, and bias.**
- Propose a framework that uses pseudo-labeling with a pre-trained gender detection model and unsupervised clustering on speaker embeddings to mitigate subgroup disparities in multi-label SER. (ISCA INTERSPEECH 2025 [5])
 - ✓ Propose an Implicit Demography Inference (IDI) module that infers latent subgroup information through pseudo-labeling and unsupervised clustering.
 - ✓ Achieve significant improvements in fairness metrics—28.78% using pseudo-labeling and 4.61% with unsupervised clustering—while maintaining high SER accuracy.
 - ✓ Validate the generalizability of the unsupervised approach across gender, race, and age subgroups, providing a scalable solution for mitigating demographic disparities in real-world multi-label SER applications.
- Introduce the first meta-learning framework for listener-personalized categorical multi-label SER, Meta-PerSER (ISCA INTERSPEECH 2025 [6])
 - ✓ Developed Meta-PerSER, the first meta-learning-based framework for listener-personalized multi-label SER.
 - ✓ Introduced Combined-Set Meta-Training (CSMT) and Derivative Annealing (DA) to enhance training stability, reduce computational costs, and improve generalization across listeners.
 - ✓ Demonstrated SOTA performance in both seen and unseen data scenarios on the IEMOCAP dataset.
- Explore the impact of perceptual evaluations elicited by different modalities on performances of SER Systems (IEEE ICASSP 2025 [7]) (IEEE SPS Travel Grant)
 - ✓ Demonstrate that SER works best when trained from voice-only labels instead of using labels from other modalities (video-only or Audio-Visual).
 - ✓ Reveal that SER systems training on the proposed combined labels from all modalities performs best for facial-only and Audio-Visual scenarios.
 - ✓ Indicate that labels elicited by the voice-only stimuli are preferable for speech-only SER applications, drawing a connection to the fact why recent benchmark databases (such as MSP-PODCAST) use audio-only stimuli for labels.
- Utilize two-stage training integrates audio-only, visual-only; audio-visual labels to train AVER models (IEEE Open Journal of Signal Processing 2025, IF=2.9 [8])
 - ✓ Propose an ad-hoc training strategy that mixes unimodal and audio-visual-rated labels for AVER systems to improve recognition rates and model calibration.
 - ✓ Measure AVER model calibration, performance bias, and fairness across different demographics .
 - ✓ Achieve higher emotion recognition rates by using a multi-label formulation that accounts for the co-occurrence of emotions.
- Propose a method that can generalize across various upstream models and datasets in different languages (IEEE ICASSP 2025 [9])
 - ✓ Demonstrate that expressive Speech-to-Speech Translation, like Meta SeamlessExpressive, and large-scale datasets of high-resource language can enhance SER in under-resourced languages.
 - ✓ Introduce a straightforward, effective approach for achieving consistent improvements across various upstream models, languages, and datasets.
 - ✓ Provide guidelines for selecting high-resource datasets to ensure consistent improvements based on our experiments.

- Establish the first SER benchmark (IEEE **SLT** 2024 [3])
 - ✓ Code and Website.
 - ✓ Reduce the barriers to reproducibility for SER.
 - ✓ Standardize the train, development, and test splits and setups of four public SER databases that did not provide the partitions.
- Explore an all-inclusive aggregation rule for multi-Label SER (IEEE **SLT** 2024 [10]) (**IEEE SLT Travel Grant**)
 - ✓ Propose an all-inclusive aggregation rule to model the subjective nature of emotion perception for SER.
 - ✓ Show the data and rating loss of the conventional aggregation rules (majority and plurality rules).
 - ✓ Reveal the SER systems trained with the proposed method perform better than the conventional ways.
- Propose a novel label aggregation rule (named **All-Inclusive Rule**) for fully using the labels and data and evaluating the performance of SER systems with the complete test set (IEEE **Transactions on Affective Computing** 2024, IF=11.2, SCI-E [11]) (**Merry Electroacoustics Thesis Silver Award**)
 - ✓ Reveal the performance of SER systems affected by using different aggregation methods for the training set when the models are evaluated on complete and incomplete test data.
 - ✓ Demonstrate the performance of the SER system trained with data selected by the majority or plurality rule on ambiguous emotions.
 - ✓ Show the best label learning strategy for training SER systems when evaluated on the complete test set.
- Propose a new ChatGPT-based prompting strategy that refines existing distributional labels on speech data by incorporating annotators' free-form emotion descriptions. (APSIPA ASC 2024 [12])
 - ✓ Demonstrate how typed descriptions enrich and nuance traditional emotion labels.
 - ✓ Show that the proposed prompt effectively interprets and integrates free-form emotional descriptions.
 - ✓ Integrate ChatGPT-adjusted labels into SER pipelines yields clear improvements in recognition accuracy.
- Introduce a unified evaluation framework, EMO-Codec, to measure how well a variety of neural and legacy speech codecs preserve emotional content—using both objective SER metrics and human listening tests across six emotion datasets. (APSIPA ASC 2024 [13])
 - ✓ EMO-Codec evaluates 14 neural and 3 legacy codecs at multiple bitrates on six diverse emotion corpora, establishing clear emotion-preservation rankings.
 - ✓ The Descript Audio Codec (DAC) family consistently leads in SER accuracy at each bitrate, while AcademiCodec and SpeechTokenizer maintain strong emotional fidelity even under low-bandwidth constraints.
 - ✓ Mixing English and Chinese data for codec training yields only modest improvements for Chinese emotional cues, and negatively valenced emotions (sadness, depression, fear, disgust) suffer the greatest recognition drops after resynthesis.
- Reveal the gender bias in self-surprised learning based SER systems at the large scale (ISCA **INTERSPEECH** 2024 [14])
 - ✓ Analyzes gender bias in SER models trained with Self-Supervised Learning (SSL) at scale.
 - ✓ Reveal that females exhibit slightly higher overall SER performance than males.
 - ✓ Observe that gender-wise emotion distribution differences in training data significantly affect gender bias, while upstream model representation has a limited impact.
- Improve confidence of predictions of SER systems (ISCA **INTERSPEECH** 2023 [15])
 - ✓ Reveal the need for model calibration in SER as a key step to increase the reliability of a system
 - ✓ Demonstrate that integrating a class-balanced objective function during the training process can improve the calibration and performance of a multi-label SER classifier (4.9% absolute improvement in macro-F1 score)
 - ✓ Introduce a multi-label temperature scaling calibration method to improve the confidence of multi-label SER classifiers while preserving the classification performance
- Exploiting co-occurrence frequency of emotions in perceptual evaluations to train a speech emotion classifier (SEC) (ISCA **INTERSPEECH** 2022 [16]) (**ISCA Travel Grant**)

- ✓Utilize the prior knowledge of co-occurrence of emotions to train a SEC model
- ✓Propose an elegant implementation to incorporate the “penalty loss” in the model, which is flexible and can be easily applied to any emotion classification framework using existing label learning approaches
- ✓Show important relative improvements in macro F1-score for hard-label learning (17.12%), multi-label learning (12.79%), and distribution-label learning (25.8%)
- Maximize utilization of ratings for training SER systems (IEEE **ICASSP** 2022 [17])
 - ✓Gain 6.4% and 16.56% performance gain on 8-class and 16-class emotion classification respectively
 - ✓Utilize annotators’ typed description of emotion perception by NLP techniques
 - ✓Evaluate SER models with all data of emotional database (no data is discarded)
- Investigate impacts of minority views in perceptual evaluations affect confidence on performances of SER Classifiers? (IEEE **ACII** Workshop and Demos 2022 [18])
 - ✓Explore whether removing minority annotations affects the confidence of the SER systems’ predictions
- Predict Inter-annotator Agreements to Improve Calibration and Performance of Speech Emotion Classifiers (SEC) (ISCA **INTERSPEECH** Doctoral Consortium 2022 [19])
 - ✓Investigate whether modeling ambiguity in annotator labels – the disagreements that are traditionally rejected as noise – can improve a SEC
- Build **Personalized** SER systems (ISCA **INTERSPEECH** 2020 [20])
 - ✓Propose per-rater speech emotion perception recognition with soft and hard labels joint training
 - ✓Reduce demands of training data into a half for SER models which achieves competitive results
- Train SER systems with **subjective labels and individual annotators** (IEEE **ICASSP** 2019 [21])
 - ✓Deploy individual SER models corresponding to unique annotators to improve model performance
 - ✓Model the label uncertainty and annotator idiosyncrasy using individual annotators and crowd modeling
- Collect a large-scale Mandarin Chinese interactive multimodal emotion corpus and build Mandarin Chinese SER systems (IEEE **ACII** 2017 [2]) (**AAAC Student Travel Grant**)
 - ✓Release the largest Mandarin Chinese (Taiwan) interactive multimodal emotion corpus
- Amplify a sense of emotion toward drama with a dynamic SER system (ACLCLP **ROCLING** 2017 [22])
 - ✓Modulate the RGB values of frames in video using the predictions of arousal and valence

• **Research in Automated Deception(Trust) Detection (ADD) | BIIC Lab@NTHU (Taiwan)** Jul. 2015 – Dec. 2021
PhD works & Co-work *Speech, NLP*

- **Abstract:** Explore machine/deep learning methods and NLP techniques to develop ADD models based on real-world data
- Improve ADD models by integrating interlocutor’s judgments in conversations (ACL **ACL-IJCNLP** 2021 [23]) (**Merry Electroacoustics Thesis Bronze Award**)
 - ✓Consider human judgments (believe/disbelieve) with three types of speech-language features (acoustic-prosodic, linguistic, and conversational temporal dynamics features) extracted during a conversation for building ADD models
- Propose ADD models using multiple speech and language communicative descriptors in Dialogs (IEEE **APSIPA** TSIP 2021, IF = 3.2, ESCI [24])
 - ✓Model psychologists’ insights and knowledge to improve the performance of detecting deceptions/trusts
- Develop perceived ADD models using multimodal data in dialog games (IEEE **APSIPA** ASC 2020 [25])
 - ✓Incorporate personality attributes as an additional input to the deceptions/trusts recognition network
 - ✓Demonstrate additional evidence indicating that humans are bad at detecting deceptions/trusts
- Design and Propose a Conversational Temporal Dynamics Features (IEEE **APSIPA** ASC 2019 [26]) (**APSIPA ASC Best Regular Paper Award**)
 - ✓Propose a joint learning framework to detect deceptions/trusts by simultaneously considering variations and patterns of the conversation using both interlocutor’s acoustic features and their conversational temporal dynamics

- Daily Deceptive Dialogues Corpus of Mandarin (DDDM) (ISCA **INTERSPEECH** 2019 [1])
 - ✓Release the largest Mandarin Chinese interactive multimodal deception/trust corpus including speech, transcripts, and personality recordings

• **Research in Physiological Signal Processing|BIIC Lab@NTHU**

Hsinchu, Taiwan

Co-works

Jul. 2018 – present

- Develop a rapid and economic in vivo electrocardiogram platform for cardiovascular drug assay and electrophysiology research in adult zebrafish (Nature Scientific Report 2018, IF = 4.857 [27])
 - ✓Design a data analysis system to denoise electrophysiological signals in adult Zebrafish

AWARDS AND HONORS

- **IEEE Signal Processing Society (SPS) ICASSP Travel Grants (2025)**
- Merry Electronics Co., Ltd.: Electroacoustics Thesis Award – Silver Award (2025)
- National Science and Technology Council (NSTC) Postdoctoral Research Abroad Program (2025-2026)
- **IEEE Spoken Language Technology Workshop (SLT) Travel Grants (2024)**
- ACLCLP Doctoral Dissertation Award – Honorable Mention (2024)
- The Elite-Well Education Foundation Ph.D. Student Excellence Scholarship Finalist (Top15) (2024)
- **Google East Asia Student Travel Grants (ICASSP 2022[17] & INTERSPEECH 2022[16])**
- NOVATEK MICROELECTRONICS CORP. PhD Excellence Scholarship (2022-2023)
- International Speech Communication Association (ISCA) INTERSPEECH Grants (2022)
- Merry Electronics Co., Ltd.: Electroacoustics Thesis Award – Bronze Award (2021)
- The Rotary Foundation Excellence Scholarship (2020-2021)
- National Science and Technology Council (NSTC) Graduate Students Study Abroad Grant (2020, 2024)
- Foundation for the Advancement of Outstanding Scholarship (FAOS) Outstanding Students Conference Travel Grant (2019, 2022, 2023)
- The Association for Computational Linguistics and Chinese Language Processing (ACLCLP) Outstanding Students Conference Travel Grant (2019, 2022, 2024, 2025)
- Ministry of Science and Technology (MOST) Futuretek Breakthrough Award (2019)
- **APSIPA ASC Best Regular Paper Award (2019)**
- **Association for the Advancement of Affective Computing (AAAC) Affective Computing and Intelligent Interaction (ACII) Student Travel Grant (2017)**
- National Tsing Hua University (NTHU) Dean Ph.D. Student Excellence Scholarship (2016-2020)
- Taiwan Imaging Tek Corporation (TITC) Excellence Scholarship (2015)
- Taiwan Semiconductor Manufacturing Company (TSMC) Excellence Scholarship (2014)

TECHNICAL SKILLS

- **EMO-SUPERB** (SER leaderboard)
- **Database Collection:** Emotion Database (NTHU-NNIME [2] and MSP-Podcast datasets [28]), Deception Database (DDDM) [1], and Group and Team Dynamics Behavior (NTUBA) [4]
- **Multimodality Processing:** **Speech Signal Processing** (Praat, openSMILE, librosa, Fairseq), **Natural Language Processing** (Linguistic Inquiry and Word Count (LIWC), CKIPTagger, part-of-speech, named entity recognition), **Physiological Signal Processing**
- **Machine/Deep Learning:** Supervised learning methods (HuggingFace, scikit-learn, Pytorch, Tensorflow (Keras))
- **Programming:** Python, MATLAB, C++, SQL
- **Other Tools & Familiar OS:** Amazon AWS, Amazon S3, DeepSeed, LaTeX, Git (Github and Gitlab), Linux (Ubuntu and Pop!_OS), Windows

REFERENCES

- [1] Chih-Hsiang Huang, **Huang-Cheng Chou**, Yi-Tong Wu, Chi-Chun Lee, and Yi-Wen Liu. “Acoustic Indicators of Deception in Mandarin Daily Conversations Recorded from an Interactive Game”. In: *Proc. Interspeech 2019*. 2019, pp. 1731–1735. DOI: 10.21437/Interspeech.2019-2216. URL: <http://dx.doi.org/10.21437/Interspeech.2019-2216>.
- [2] **Huang-Cheng Chou**, Wei-Cheng Lin, Lien-Chiang Chang, Chyi-Chang Li, Hsi-Pin Ma, and Chi-Chun Lee. “NNIME: The NTHU-NTUA Chinese interactive multimodal emotion corpus”. In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. 2017, pp. 292–298. DOI: 10.1109/ACII.2017.8273615.
- [3] **Huang-Cheng Chou**, Haibin Wu, Kai-Wei Chang, Lucas Goncalves, Jiawei Du, Jyh-Shing Roger Jang, Chi-Chun Lee, and Hung-Yi Lee. “Open-Emotion: A Reproducible EMO-Superb For Speech Emotion Recognition Systems”. In: *2024 IEEE Spoken Language Technology Workshop (SLT)*. 2024, pp. 510–517. DOI: 10.1109/SLT61566.2024.10832296.
- [4] Woan-Shiuan Chien, **Huang-Cheng Chou**, and Chi-Chun Lee. “Belongingness and Satisfaction Recognition from Physiological Synchrony with A Group-Modulated Attentive BLSTM under Small-group Conversation”. In: *Companion Publication of the 2021 International Conference on Multimodal Interaction*. 2021. DOI: 10.1145/3461615.3485410.
- [5] Yi-Cheng Lin, **Huang-Cheng Chou**, and Hung-Yi Lee. “Mitigating Subgroup Disparities in Multi-Label Speech Emotion Recognition: A Pseudo-Labeling and Unsupervised Learning Approach”. In: *Proc. Interspeech*. 2025.
- [6] Liang-Yeh Shen, Shi-Xin Fang, Yi-Cheng Lin, **Huang-Cheng Chou**, and Hung-Yi Lee. “Meta-PerSER: Few-Shot Listener Personalized Speech Emotion Recognition via Meta-learning”. In: *Proc. Interspeech*. 2025.
- [7] **Huang-Cheng Chou**, Haibin Wu, and Chi-Chun Lee. “Stimulus Modality Matters: Impact of Perceptual Evaluations from Different Modalities on Speech Emotion Recognition System Performance”. In: *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2025, pp. 1–5. DOI: 10.1109/ICASSP49660.2025.10890558.
- [8] Lucas Goncalves, **Huang-Cheng Chou**, Ali N. Salman, Chi-Chun Lee, and Carlos Busso. “Jointly Learning From Unimodal and Multimodal-Rated Labels in Audio-Visual Emotion Recognition”. In: *IEEE Open Journal of Signal Processing* 6 (2025), pp. 165–174. DOI: 10.1109/OJSP.2025.3530274.
- [9] Hsi-Che Lin, Yi-Cheng Lin, **Huang-Cheng Chou**, and Hung-yi Lee. “Improving Speech Emotion Recognition in Under-Resourced Languages via Speech-to-Speech Translation with Bootstrapping Data Selection”. In: *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2025, pp. 1–5. DOI: 10.1109/ICASSP49660.2025.10887615.
- [10] **Huang-Cheng Chou**, Lucas Goncalves, Haibin Wu, Hung-Yi Lee, and Chi-Chun Lee. “Embracing Ambiguity And Subjectivity Using The All-Inclusive Aggregation Rule For Evaluating Multi-Label Speech Emotion Recognition Systems”. In: *2024 IEEE Spoken Language Technology Workshop (SLT)*. 2024, pp. 502–509. DOI: 10.1109/SLT61566.2024.10832302.
- [11] **Huang-Cheng Chou**, Lucas Goncalves, Seong-Gyun Leem, Ali N. Salman, Chi-Chun Lee, and Carlos Busso. “Minority Views Matter: Evaluating Speech Emotion Classifiers with Human Subjective Annotations by an All-Inclusive Aggregation Rule”. In: *IEEE Transactions on Affective Computing* (2024), pp. 1–15. DOI: 10.1109/TAFFC.2024.3411290.
- [12] **Huang-Cheng Chou**, Haibin Wu, Kai-Wei Chang, Lucas Goncalves, Jiawei Du, Jyh-Shing Roger Jang, Chi-Chun Lee, and Hung-Yi Lee. “Empower Typed Descriptions by Large Language Models for Speech Emotion Recognition”. In: *2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2024, pp. 1–6. DOI: 10.1109/APSIPAASC63619.2025.10848758.
- [13] Wenze Ren, Yi-Cheng Lin, **Huang-Cheng Chou**, Haibin Wu, Yi-Chiao Wu, Chi-Chun Lee, Hung-Yi Lee, Hsin-Min Wang and Tsao Yu. “EMO-Codec: An In-Depth Look at Emotion Preservation Capacity of Legacy and Neural Codec Models with Subjective and Objective Evaluations”. In: *2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2024, pp. 1–6. DOI: 10.1109/APSIPAASC63619.2025.10849259.

- [14] Yi-Cheng Lin, Haibin Wu, **Huang-Cheng Chou**, Chi-Chun Lee, and Hung-Yi Lee. “Emo-bias: A Large Scale Evaluation of Social Bias on Speech Emotion Recognition”. In: *Proc. Interspeech*. 2024, pp. 4633–4637. DOI: 10.21437/Interspeech.2024-1073.
- [15] **Huang-Cheng Chou**, Lucas Goncalves, Seong-Gyun Leem, Chi-Chun Lee, and Carlos Busso. “The Importance of Calibration: Rethinking Confidence and Performance of Speech Multi-label Emotion Classifiers”. In: *Proc. Interspeech*. 2023, pp. 641–645. DOI: 10.21437/Interspeech.2023-1113.
- [16] **Huang-Cheng Chou**, Chi-Chun Lee, and Carlos Busso. “Exploiting Co-occurrence Frequency of Emotions in Perceptual Evaluations To Train A Speech Emotion Classifier”. In: *Proc. Interspeech 2022*. 2022. URL: https://www.isca-speech.org/archive/interspeech_2022/chou22_interspeech.html.
- [17] **Huang-Cheng Chou**, Wei-Cheng Lin, Chi-Chun Lee, and Carlos Busso. “Exploiting Annotators’ Typed Description of Emotion Perception to Maximize Utilization of Ratings for Speech Emotion Recognition”. In: *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2022. URL: <https://ieeexplore.ieee.org/abstract/document/9746990>.
- [18] **Huang-Cheng Chou**. “Do Minority Views in Perceptual Evaluations Affect Confidence of Speech Emotion Classifiers?” In: *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE. 2022, pp. 1–5. DOI: 10.1109/ACIIW57231.2022.10085994.
- [19] **Huang-Cheng Chou**. “Predicting Inter-annotator Agreements to Improve Calibration and Performance of Speech Emotion Classifiers”. In: *8th Doctoral Consortium Proc. Interspeech 2022*. 2022. URL: http://www.isca-students.org/sacweb/images/files/resources/8thDC/3_8DC_paper_1859.pdf.
- [20] **Huang-Cheng Chou** and Chi-Chun Lee. “Learning to Recognize Per-Rater’s Emotion Perception Using Co-Rater Training Strategy with Soft and Hard Labels”. In: *Proc. Interspeech 2020*. 2020, pp. 4108–4112. DOI: 10.21437/Interspeech.2020-1714. URL: <http://dx.doi.org/10.21437/Interspeech.2020-1714>.
- [21] **Huang-Cheng Chou** and Chi-Chun Lee. “Every Rating Matters: Joint Learning of Subjective Labels and Individual Annotators for Speech Emotion Classification”. In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 5886–5890. DOI: 10.1109/ICASSP2019.8682170.
- [22] **Huang-Cheng Chou**, Chun-Min Chang, Yu-Shuo Liu, Shiuan-Kai Kao, and Chi-Chun Lee. “Amplifying a Sense of Emotion toward Drama-Long Short-Term Memory Recurrent Neural Network for Dynamic Emotion Recognition”. In: *Proceedings of the 29th Conference on Computational Linguistics and Speech Processing (ROCLING 2017)*. 2017, pp. 136–147. URL: <https://www.aclweb.org/anthology/O17-1013.pdf>.
- [23] **Huang-Cheng Chou**, Woan-Shiuan Chien, Da-Cheng Juan , and Chi-Chun Lee. ““Does it Matter When I Think You Are Lying?” Improving Deception Detection by Integrating Interlocutor’s Judgements in Conversations”. In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Online: Association for Computational Linguistics, Aug. 2021, pp. 1846–1860. DOI: 10.18653/v1/2021.findings-acl.162. URL: <https://aclanthology.org/2021.findings-acl.162>.
- [24] **Huang-Cheng Chou**, Yi-Wen Liu, and Chi-Chun Lee. “Automatic Deception Detection using Multiple Speech and Language Communicative Descriptors in Dialogs”. In: *APSIPA Transactions on Signal and Information Processing 10 (2021)*, e5. DOI: 10.1017/ATSIP2021.6.
- [25] **Huang-Cheng Chou** and Chi-Chun Lee. ““Your Behavior Makes Me Think It Is a Lie”: Recognizing Perceived Deception using Multimodal Data in Dialog Games”. In: *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2020, pp. 393–402. URL: <https://ieeexplore.ieee.org/abstract/document/9306359>.
- [26] **Huang-Cheng Chou**, Yi-Wen Liu, and Chi-Chun Lee. “JOINT LEARNING OF CONVERSATIONAL TEMPORAL DYNAMICS AND ACOUSTIC FEATURES FOR SPEECH DECEPTION DETECTION IN DIALOG GAMES”. In: *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2019, pp. 1044–1050. DOI: 10.1109/APSIPAASC47483.2019.9023050.

- [27] Min-Hsuan Lin and **Huang-Cheng Chou** and Yu-Fu Chen, Wangta Liu, Chi-Chun Lee, Lawrence Yu-Min Liu, and Yung-Jen, Chuang. “Development of a rapid and economic in vivo electrocardiogram platform for cardiovascular drug assay and electrophysiology research in adult zebrafish”. In: *Scientific Reports* 8, 15986 (Oct. 2018), p. 15986. DOI: 10.1038/s41598-018-33577-7.
- [28] Carlos Busso, Reza Lotfian, Kusha Sridhar, Wei-Cheng Lin, Srinivas Parthasarathy, Abinay Naini, Jeri Luckenbaugh, Luz Martinez-Lucas, Ali N. Salman, Lucas Goncalves, and **Huang-Cheng Chou**. “The MSP-Podcast 2.0”. In: *IEEE Transactions on Affective Computing (submitted)* (2025).