Kuan-Yu Chen (Daniel)

Ph.D. Student | Speech, Audio, and Trustworthy AI
de666vv@gmail.com | danielchen1128.github.io | GitHub | Google Scholar

Profile

Ph.D. student and AI researcher developing generative speech and audio systems for real-world deployment. Research spans robust speech editing, instruction-guided TTS, speech fairness, audio security, and music information retrieval. Collaborates with Academia Sinica Bio-ASP Lab and Inventec AI Research Center.

Technical Focus

Machine learning: deep learning for speech and audio, generative modeling, zero-shot learning, model evaluation, bias analysis.
Signal processing: acoustic analysis, audio feature engineering, music information retrieval, vibration and sensor signals.
Programming and tools: Python, MATLAB, Simulink, Arduino, SolidWorks.

Research and Industry Experience

Ph.D. Student, Bio-ASP Lab, Academia Sinica Feb. 2026 – Present
Conduct research on ECG watermarking for biomedical signals, alongside collaborative work on trustworthy speech and audio AI, including fairness and robustness in speech technologies.

AI Researcher Intern, AI Research Center, Inventec Sep. 2024 – Present
Conduct research on robust speech generation, audio security, and instruction-guided text-to-speech. Develop and evaluate deep-learning methods for real-world speech and audio applications.

Audio Engineer Intern, Inventec Jul. 2024 – Aug. 2024
Conducted acoustic and signal analysis for industrial applications, including investigation of gear-rattling behavior and its engineering implications.

Machine Learning Engineer Intern, Leaderg Jan. 2024 – Jun. 2024
Supported AI model development, algorithm research, data labeling and cleaning, signal processing, and AI application software development.

Selected Research Projects

Robust zero-shot speech editing Paper | Code
Developed a background-noise-aware speech editing approach using in-context enhancement for insertion and replacement under real-world acoustic conditions.

Instruction TTS fairness and evaluation Paper | Code
Analyzed how social-status, career, and persona cues combine to produce gender bias in instruction TTS; co-authored a broad survey of fairness and bias in speech AI.

Audio watermark security Paper | Code
Investigated watermarking as a mechanism for stealthy backdoors in audio models and its implications for model and data security.

Music information retrieval and generation Paper
Built learning-based query-by-humming systems and developed diffusion-based guitar tone morphing methods in latent audio spaces.

Education

National Taiwan University, Taipei, Taiwan — Sep. 2024 – Expected Jun. 2029
Ph.D. in Communications Engineering, Direct Ph.D. Program
Advisors: Yu Tsao and Jian-Jiun Ding

National Taiwan University, Taipei, Taiwan — Sep. 2019 – Jun. 2024
B.S. in Biomechatronics Engineering

Selected Publications

The Binding Effect: Analyzing How Multi-Dimensional Cues Form Gender Bias in Instruction TTS. **Kuan-Yu Chen**, Yi-Cheng Lin, Po-Chung Hsieh, Huang-Cheng Chou, Chih-Fan Hsu, Jeng-Lin Li, Hung-yi Lee, Jian-Jiun Ding. *Interspeech*, 2026.

Bloodroot: When Watermarking Turns Poisonous for Stealthy Backdoor. **Kuan-Yu Chen**, Yi-Cheng Lin, Jeng-Lin Li, Jian-Jiun Ding. *ICASSP*, 2026.

SeamlessEdit: Background Noise Aware Zero-Shot Speech Editing with in-Context Enhancement. **Kuan-Yu Chen**, Jeng-Lin Li, De-Yan Lu, Jian-Jiun Ding. *EUSIPCO*, 2026.

Toward Fair Speech Technologies: A Comprehensive Survey of Bias and Fairness in Speech AI. Yi-Cheng Lin, Yun-Shao Tsai*, **Kuan-Yu Chen***, Hsiao-Ying Huang*, Huang-Cheng Chou*, Shrikanth Narayanan, Yu Tsao, Jian-Jiun Ding, Hung-yi Lee. *arXiv:2605.01597*, 2026.

Guitar Tone Morphing by Diffusion-based Model. **Kuan-Yu Chen**, Kuan-Lin Chen, Yu-Chieh Yu, Jian-Jiun Ding. *APSIPA ASC*, 2025.

Improved Query by Humming Using Conformer-based Network with Harmonic-Aware Mechanism. Yu-Hsuan Kung, **Kuan-Yu Chen**, Jian-Jiun Ding. *ISCAS*, 2026.

Additional Publications

Findings of the First Teaching Monster Challenge: A Benchmark of Pedagogical Content Knowledge in AI Agents. Yi-Cheng Lin, Yu-Kai Guo, Szu-Chi Chen, Bo-Han Feng, Yun-Man Hsu, Hsiang Hsieh, Yu-Jung Lin, Yue-Ling Wu, Jia-Kai Dong, An-Yu Cheng, Yu-Han Huang, Lok-Lam Ieong, **Kuan-Yu Chen**, Ming-Douo Tchouang, Shao-Hua Sun, Che Lin, Jian-Jiun Ding, Hung-yi Lee. *arXiv:2608.08852*, 2026.

Speaker Identity in Non-Verbal Vocalizations: Conditional Distillation and Mixture of Experts Approach. Tzu-Chieh Wei, Yi-Cheng Lin, Huang-Cheng Chou, **Kuan-Yu Chen**, Hsin-Yen Sung, Shrikanth Narayanan, Hung-yi Lee. *Interspeech*, 2026.

Do You Hear What I Mean? Quantifying the Instruction-Perception Gap in Instruction-Guided Expressive Text-to-Speech Systems. Yi-Cheng Lin, Huang-Cheng Chou, Tzu-Chieh Wei, **Kuan-Yu Chen**, Hung-yi Lee. *ICASSP*, 2026.

Creativity in LLM-based Multi-Agent Systems: A Survey. Yi-Cheng Lin*, Kang-Chieh Chen*, Zhe-Yan Li*, Tzu-Heng Wu*, Tzu-Hsuan Wu*, **Kuan-Yu Chen***, Hung-yi Lee, Yun-Nung Chen. *EMNLP*, 2025.

Chromagram Features Analysis for Learning-Based Query by Humming Systems. **Kuan-Yu Chen**, Jian-Jiun Ding. *ICEIC*, 2025.

Mixed Music Instrument Classification Based on Instantaneous Frequency Analysis and Time-Variant Spectrum Information. **Kuan-Yu Chen**, Jian-Jiun Ding, Yuan-Kai Lee. *ICGSP*, 2024.

* Equal contribution. Full publication list: Google Scholar.