Gene-Ping Yang

geneping2009@gmail.com | homepages.inf.ed.ac.uk/s2064029/ | in geneping-yang | CZyca70AAAAJ

Education

University of Edinburgh — Edinburgh, Scotland
Ph.D. in CSTR, School of Informatics, Supervisor: Prof. Hao Tang — Oct. 2020 - March 2025 (expected)

National Taiwan University — Taipei, Taiwan
M.S. in Computer Science, Supervisor: Prof. Lin-shan Lee and Prof. Hung-yi Lee — Sep. 2017 – June 2019
Thesis: Speech Separation with Time-and-Frequency Cross-Domain Joint Embedding and Clustering
B.S. in Electrical Engineering — Sep. 2013 – June 2017

Research Interests

Self-Supervised Speech Learning: Guide self-supervised speech feature alignment with text modality

Speech Tokenization: Extract discrete speech tokens from pre-trained speech models for speech-to-text applications

Automatic Speech Recognition: Improve implicit alignment between speech and text

Speech Enhancement and Separation: Uncover speech from noisy and reverberant signals

Academic Experience

The Centre for Speech Technology Research (CSTR) — University of Edinburgh
PhD, Supervisor: Prof. Hao Tang — Oct. 2020 - March 2025

- **Self-supervised learning**: Autoregressive predictive coding (Publication 5)
- **Unsupervise speech segmentation / phone-like unit discovery**: Leveraged self-supervised speech features (HuBERT and Wav2vec 2.0) with constrained unsupervised HMMs. (Publication 1)
- **Unsupervise ASR**: Applied optimal transport to align speech embedding with phonetic embedding. (Publication 3)
- **Speech-Text alignment for ASR**: Advanced encoder-decoder framework with supervised attention. (Publication 6)

Speech Processing and Machine Learning Lab — National Taiwan University
Master & undergrad research, Supervisor: Prof. Lin-shan Lee and Prof. Hung-yi Lee — Feb. 2017 - Aug. 2020

- **Modeling for speech separation**: Integrated time-domain and frequency-domain features to design an feature space that facilitates joint feature clustering. (Publication 9)
- **Improved permutation invariant training**: Developed a novel algorithm to improve speaker permutation optimization for speaker-invariant speech separation. (Publication 8)

Work Experience

Microsoft Research — Redmond, WA
Research Intern, Mentor: Sebastian Braun — July 2024 - Sep. 2024

- **Multi-microphone speech enhancement and separation**: designed for distributed asynchronous devices, utilizing novel cross-attention methods that outperform previous microphone aggregation and neural beamforming techniques. (Paper under preparation)

Apple — Cambridge, MA
Research Scientist Intern, Mentor: Zhen Huang, Stefan Braun, Loren Lugosch — July 2023 - Sep. 2023

- **Speech foundation model**: Developed an encoder-decoder pre-training framework using multilingual speech pseudo labels, demonstrating significant improvement in downstream ASR, particularly with large-scale unlabeled data. (Paper under preparation)

Amazon — Sunnyvale, CA
Applied Scientist Intern, Mentor: Yuzong Liu, Yue Gu, Qingming Tang — Sep. 2022 - Dec. 2022

- **Self-supervised model distillation**: Developed a task-agnostic objective for distilling self-supervised models, incorporating redundancy reduction and contrastive learning techniques to minimize data bias. (Publication 4)
- **Quantization-aware training for self-supervised models**: Applied a novel quantization-aware training method to quantize both model weights and activations, preserving optimal model performance. (Publication 2)

Ministry of Science and Technology Taipei, Taiwan

Head Teaching Assistant for Formosa Grand Challenge, Organizer: Prof. Hung-yi Lee July 2017 - Oct. 2017

- **Chinese Question Answering Challenge**: Organized a Chinese question-answering challenge, which included collecting data from TV shows, preparing models with PyTorch, TensorFlow, and Keras, and creating documentation to facilitate understanding and implementation of state-of-the-art question-answering models.

Microsoft Taipei, Taiwan

Research and Design Intern July 2016 - Oct. 2016

- **Real-time face tracking and emotion recognition**: Developed a real-time face tracking system, integrating a machine learning model for accurate emotion recognition.
- Robot development with embedded systems: Built robots using MediaTek LinkIt Smart 7688 chips, implementing automated instructions through Azure for enhanced functionality.

Publications

1. A Simple HMM with Self-Supervised Representations for Phone Segmentation SLT 2024

Gene-Ping Yang, Hao Tang

2. On-Device Constrained Self-Supervised Learning for Keyword Spotting via Quantization Aware Pre-Training and Fine-tuning ICASSP 2024 Lecture

Gene-Ping Yang, Yue Gu, Sashank Macha, Qingming Tang, Yuzong Liu

3. Towards Matching Phones and Speech Representations ASRU 2023

Gene-Ping Yang, Hao Tang

4. On-device Constrained Self-Supervised Speech Representation Learning for Keyword Spotting via Knowledge Distillation Interspeech 2023 Oral

Gene-Ping Yang, Yue Gu, Qingming Tang, Dongsu Du, Yuzong Liu

5. Autoregressive Predictive Coding: A Comprehensive Study JSTSP 2022

Gene-Ping Yang, Sung-Lin Yeh, Yu-An Chung, James Glass, Hao Tang

6. Supervised Attention In Sequence-to-Sequence Models for Speech Recognition ICASSP 2022 Lecture

Gene-Ping Yang, Hao Tang

7. Stabilizing Label Assignment for Speech Separation by Self-Supervised Pre-Training Interspeech 2021

Sung-Feng Huang, Shun-Po Chuang, Da-Rong Liu, Yi-Chen Chen, Gene-Ping Yang, Hung-yi Lee

8. Interrupted and Cascaded Permutation Invariant Training for Speech Separation ICASSP 2020 Lecture

Gene-Ping Yang, Szu-Lin Wu, Yao-Wen Mao, Hung-yi Lee, Lin-shan Lee

9. Improved Speech Separation with Time-and-Frequency Cross-domain Joint Embedding and Clustering

Gene-Ping Yang, Chao-I Tuan, Hung-Yi Lee, Lin-shan Lee Interspeech 2019 Oral

10. Distributed Asynchronous Device Speech Enhancement via Efficient Temporal Cross Attention

Gene-Ping Yang, Sebastian Braun Under Submission

Teaching

Maching Learning , Lecturer: Hao Tang University of Edinburgh

Lead discussions in five tutorial sessions, covering optimizations, learning, and hands-on implementation.

Applied Deep Learning , Lecturer: Yun-Nung Chen National Taiwan University

Lead coursework on anime face generation based on text descriptions, implementing conditional GANs and testing various GAN objectives such as WGAN, improved WGAN, and ACGAN.

Machine Learning and Having it Deep and Structured, Lecturer: Hung-yi Lee National Taiwan University

Involve in the design of coursework on sequence labeling (phone prediction) using joint training of CNN and RNN, video caption generation with a seq2seq-based model, and automatic game playing through deep reinforcement learning.

Machine Learning , Lecturer: Hung-yi Lee — National Taiwan University

Lead coursework on predicting PM 2.5 levels in the air using a hand-crafted linear regression model with gradient descent, based on historical air quality indicators.

Honors

2022	IEEE Signal Processing Society Travel Grant for ICASSP
2019-2020	Advanced Speech Technologies Scholarship
2019	Foundation For The Advancement of Outstanding Scholarship
2019	International Speech Communication Association (ISCA) Travel Grant
2019	Appier scholarship for Artificial Intelligence and Information Technology
2019	Higher Education Sprout Project

Services

Reviewer	ICLR, WACV, IEEE Signal Processing Letters, ICASSP, Interspeech, ASRU, AAAI, IJCAI