Xiaohui Zhang

E-mail: xiaohuizhang.bjtu@outlook.com| 21120320@bjtu.edu.cn | Tel: (+86) 18610987253
Personal Website: https://cecile-hi.github.io/
No.3 Shang Yuan Village, Haidian District, Beijing, China

RESEARCH INTERESTS

- Continual Learning in Machine Learning Theory
- Multimodal Learning
- Speech and Audio Signal Processing
- AI for Health

EDUCATIONAL BACKGROUND

- **Beijing Jiaotong University (BJTU, "211")** GPA:87/100 Master's degree Sep.2021 - Present
  M.Eng in Computer Information and Technology (Supervised by: Prof. Man-gui Liang)
  Relevant Coursework: C programming, advanced programming, Python programming, microcomputer principles, etc.
- **Beijing Jiaotong University (BJTU, "211")** GPA:3.81/4.0 Bachelor's degree Sep.2017 - Jun.2021
  Bachelor Thesis: "Using Deep Learning for Multi-modal Emotion Classification" (Supervised by: Prof. MingQiang Zhu)
  Relevant Coursework: Probability Theory, Calculus, Linear Algebra, Numerical Computation, Stochastic Process, Matrix Analysis
  Machine learning: Machine Learning, Pattern Recognition, Data Mining, etc.

INSTITUTIONS& ASSOCIATIONS

- International Speech Communication Association (General Membership) Apr.2023 - Present
- Chinese Association for Artificial Intelligence - Metaverse Committee (General Membership) Dec.2022 - Present

PUBLICATIONS AND WORKING PAPER

- **Published:**

1. **Zhang, X**., Yi, J., Tao, J., Wang, C., Zhang, C. "Do You Remember? Overcoming catastrophic forgetting with Regularized Adaptive Weight Modification". Fortieth International Conference on Machine Learning, ICML 2023. [Paper][Code][Slides].
2. **Zhang, X**., Yi, J., Tao, J. et al., "What to Remember: Self-Adaptive Continual Learning with Radian Weight Modification". The 38th Annual AAAI Conference on Artificial Intelligence, AAAI 2024. [Paper][Code]
3. **Xiaohui Zhang**, Jaehong Yoon, Mohit Bansal, Huaxiu Yao, Multimodal Representation Learning by Alternating Unimodal Adaptation. The IEEE / CVF Computer Vision and Pattern Recognition Conference, CVPR 2024 (Accepted) [Paper][Code]
4. **Zhang, X**., Yi, J., Tao, J. et al. "Adaptive Fake Audio Detection with Low-Rank Model Squeezing". International Joint Conferences on Artificial Intelligence, IJCAI 2023 Workshop on Deepfake Audio Detection and Analysis. [Paper].
5. **Zhang, X**., Liang, M., Tian, Z., Yi, J. and Tao, J. "TST: Time-Sparse Transducer for Automatic Speech Recognition". International Conference on Artificial Intelligence 2023, CICAI 2023 [Paper].
6. **Zhang, X**., Yi, J., Tao, J. and Zhu, J. "The Elastic Orthogonal Weight Modification for Synthetic Audio Detection in Continual Learning". Journal of Computer Research and Development (Accepted).
7. Wang, C., He, J., Yi, J., Tao, J., Zhang, C. **Zhang, X**. "Multi-Scale Permutation Entropy For Audio Deepfake Detection". 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2024 (Accepted)
8. Wang, C., Yi, J., **Zhang, X**. et al. "Low-rank Adaptation Method for Wav2vec2-based Fake Audio Detection". IJCAI 2023 Workshop on Deepfake Audio Detection and Analysis.[Paper].
9. Yi, J., Tao, J., Fu, R., Yan, X., Wang, C., Wang, T., Zhang, C., **Zhang, X**. et al. "ADD 2023: the Second Audio Deepfake Detection Challenge". IJCAI 2023 Workshop on Deepfake Audio Detection and Analysis. [Paper].
10. Zhang Z, **Zhang X**, Guo M, et al. "A Multilingual Framework Based on Pre-training Model for Speech Emotion Recognition". 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2021: 750-755. [Paper].
11. Lin S, **Zhang X**, Guo M, et el. Research on audio-video based emotion recognition methods [J]. Signal processing, 2021, 37(10): 1889-1898.

- **Submitted:**

1. **Zhang, X**., Yi, J., Tao, J., Wang, C., "Robust continuous learning benchmark for multimodal deepfake detection". IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

2. Yi, J., Tao, J., Wang, C., **Zhang, X**. and Zhao, Y. "Deepfake Audio Detection: A Survey". IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

RESEARCH EXPERIENCE

1. **Multimodal Emotion Recognition Project** 2020 - 2021

SATLab Laboratory, Department of Electronics, Tsinghua University

- Research Objective: Develop a vehicular multimodal emotion analysis system that includes speech, image, and EEG.
- Independently undertook the development of the mid-term speech emotion recognition system; participated in the construction of the multi-modal dataset and the deployment and operation of the fusion recognition model.
- Primarily researched and implemented modality fusion and participated in the planning and collection of the early-stage dataset. Proposed a framework for emotion recognition from speech, image, and EEG under three fusion strategies and published 2 papers.

2. **The competition for Alzheimer's disease detection at NCMMSC 2021** 2020 - 2021

SATLab Laboratory, Department of Electronics, Tsinghua University

- Research Objective: Provide a baseline system for the Alzheimer's disease Detection competition at the National Conference on Man-Machine Speech Communication (NCMMSC 2021)
- Undertook the design, development, and release of the Alzheimer's disease detection system.
- Proposed an Alzheimer's speech detection system based on the HGFM model.

3. **Forged Audio and Video Detection with Continual Learning Project** 2023 - Present

State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences

- Research Objective: Design a robust, scalable system for the generation and identification of forged audio and video.
- Independently undertook the research and development of the incremental training of the audio identification model and passed the mid-term acceptance successfully.
- Integrated different incremental learning methods according to the characteristics of the identification model, and compared them with the direct fine-tuning method, highlighting the good identification effect of the continual learning method when facing new forgery algorithms.

4. **DADA 2023 Workshop at IJCAI 2023. Jointly held by Tsinghua University, Institute of Automation of the Chinese Academy of Sciences, and CUHK (Shenzhen)** 2023 - Present

State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences

- Research Objective: Provide a baseline system for the forged segment localization problem in Track 2 of the DADA 2023 competition.
- Prepared the dataset for Track 2, researched, and provided a baseline system for locating forged segments in forged speech, and provided a test metric calculation code and script.
- We specifically optimized the structure of the LCNN to successfully solve the problem of segment localization.

AWARDS/HONORS

1. First Class Master's Academic Scholarship at BJTU Jan. 2024
2. Third Class Master's Academic Scholarship at BJTU Jan. 2023
3. First Class Master's Academic Scholarship at BJTU Nov. 2021
4. Second Class Academic Excellence Scholarship at BJTU Dec. 2019
5. Outstanding Student Award at BJTU Dec. 2019
6. Second Class Academic Excellence Scholarship at BJTU Dec. 2018
7. Outstanding Student Award at BJTU Dec. 2018
8. Second Prize, 10th National College Mathematics Competition, Popularization Committee of Chinese Mathematical Society Nov. 2018
9. Second Prize, 29th Beijing College Student Mathematics Competition, Beijing Mathematical Society (BMS) Nov. 2018

SKILLS

- Continuous Learning Method, Apparatus, and Electronic Device for Generating Speech Discrimination Models