Yi Zhu

Montreal, Canada|+1 514-791-8242| Yi.Zhu@inrs.ca| Personal website: zhu00121.github.io

EDUCATION

Ph.D., Telecommunications **Jan, 2021 - Aug, 2024 (Expected)**
Institut national de la recherche scientifique (INRS) *Montreal, Canada*
Master of Science, Biomedical Engineering (Major), Movement Science (Minor) **July, 2020**
University of Minnesota - Twin Cities *Minneapolis, USA*
Area of Expertise: Foundation models, Speech/Audio/Physiological signal, Cybersecurity, Healthcare

Awards (2023-2024)

- "*Rising Star in Signal Processing"* at ICASSP 2023
- *Philippe-Edwin Bélanger – TD Grant* 2023

RESEARCH EXPERIENCE

Ph.D. Candidate **Feb, 2021 - Current**
Multisensory Signal Analysis and Enhancement Lab (MuSAE) *Montreal, Canada*
Thesis: Development of generalizable, explainable, and privacy-preserving audio applications

- Explore new speech/audio representation learning frameworks applied across various tasks, such as speech synthesis, deepfake detection, health monitoring, speaker verification, etc.
- Build public benchmarks to facilitate the reproduction and evaluation of large foundation models
- 5+ YOE on DSP and ML for speech/audio analysis, sensor signals, and their real-world applications
- Top-tier first-author publications (e.g., *Nature Scientific Data, IEEE-TASLP, IEEE-TIFS, ICASSP, INTERSPEECH*); team-leading in multiple challenges (e.g., ASVspoof, ComParE)

Research Assistant **Jan 2019 – July, 2020**
Human Sensorimotor Control lab *Minneapolis, USA*
Thesis: sEMG-based interface design for individuals with cervical dystonia

- Developed an EMG-based model for cervical movement prediction with an accuracy of 82.5% in predicting 10 orientations of neck movements
- Designed an EMG-controlled wearable device design for individuals with neurological disorders

INDUSTRY EXPERIENCE

Applied Scientist - Audio Model (Internship) **Jan, 2024 - Current**
Reality Defender *Remote*

- Developed self-supervised pretraining methods for generalized audio deepfake detection
- Building tools for interpreting the decision-making of large speech models and improving the generalizability to unseen attacks

Data Scientist (Part-time) **Dec, 2022 – Current**
Nectar *Montreal, Canada*

- Investigated using multi-modal signals (audio, humidity, temp) for context-aware bee monitoring
- Led in bee audio signal processing, feature engineering, data visualization, and pattern recognition
- First-author in two Nature papers (under review) for multi-modal beehive monitoring

R&D engineer (Part-time) **Sept 2018 - May 2019**
University of Minnesota - Twin Cities *Minneapolis, USA*

- Collaborated with a local biomedical company to develop a standing assisted device for the elderly
- Data analysis of body biomechanics using motion capture to understand sit-to-stand ergonomics
- Collaborated with engineers and business analytics in device design and market analysis

SUPERVISING/TEACHING EXPERIENCE

MITACS student supervisor **June-Aug, 2022&2023**
INRS-MITACS *Montreal, Canada*
Summary: Supervising multiple undergraduate students to conduct research in the following fields: (1) visual-audio deepfake detection; (2) acoustic event localization; and (3) pathological sound analysis.

Master's student supervisor **Jan, 2022 – current**
INRS *Montreal, Canada*
Summary: Supervising a master's student on investigating adversarial attacks in SER

Short-term course lecturer **Aug, 2022**
INRS-UPFE (Brazil) *Remote*
Summary: Lectured course "Modulation Spectrum Signal Processing: A Theoretical and Hands-On Course with Applications in Speech, Biomedical, and Cybersecurity Domains" (see Github course link here)

PUBLICATIONS (2023-2024)

Journals:
1. **Y.Zhu**, and T.Falk, "WavTX: a disease-agnostic, generalizable, and privacy-preserving speech health encoder", IEEE Transactions on Audio, Speech, and Language Processing (IEEE-TASLP), *under review*
2. **Y.Zhu** et al., "MSPB: a longitudinal multi-sensor dataset with phenotypic trait measurements from honey bees", Nature Scientific Data, *under review*
3. **Y.Zhu** et al., "On the impact of voice anonymization for speech-based health diagnostics", IEEE Transactions on Information Forensics and Security (IEEE-TIFS), 2024, *published*
4. **Y.Zhu,** and T.Falk, "Spectral-temporal saliency maps and modulation tensorgrams for generalizable COVID-19 detection", Computer Speech & Language, 2023, *published*
5. **Y.Zhu** et al., "Linear prediction and modulation spectrum features for improved COVID-19 detection", IEEE Transactions on Audio, Speech, and Language Processing (IEEE-TASLP), 2023, *published*

Conferences:
1. Y.Zhu et al., "SLIM: a self-supervised training framework to learn style-linguistics mismatch in synthesized audio", Thirty-eighth Annual Conference on Neural Information Processing Systems (Neurips) 2024, *under review*
6. **Y.Zhu**, Saurabh Powar, and T.Falk, "Characterizing the temporal dynamics of universal speech representations for generalizable deepfake detection", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2024, *published*
7. **Y.Zhu** et al., "Early prediction of honeybee hive winter survivability using multi-modal sensor data", IEEE-MetroAgriFor 2023, *published*
8. H.Guimarães, M.Abdollahi, **Y.Zhu**, et al. "Adapting Self-Supervised Features for Background Speech Detection in Beehive Audio Recordings", IEEE-MetroAgriFor 2023, **Best paper award**
9. H.Guimarães, **Y.Zhu,** O.Mengara, A. Avila, T.Falk, "Assessing the Vulnerability of Self-Supervised Speech Representations for Keyword Spotting Under White-Box Adversarial Attacks", SMC 2023, *accepted*
10. **Y.Zhu** et al., "Investigating Biases in Diagnostic Systems Processed with Automated Speech Anonymization Algorithms", ISCA-SPSC 2023, *published*
*Papers before 2023 can be found at my *Google Scholar Profile*

SKILLS

Computer skills: Python, MATLAB, C++
Machine Learning Framework: Pytorch, TensorFlow, SpeechBrain, Numpy, Pandas, Scikit-learn
Languages: English, Mandarin, French