
EDUCATION

- Queen Mary University of London, England** **Sep. 2020 - Mar. 2025 (Expected)**
 PhD, UKRI CDT in Artificial Intelligence and Music. Supervisors: Simon Dixon and Sebastian Ewert
- National Chiao Tung University, Taiwan** **Aug. 2013 - Jan. 2016**
 Master of Science, Sound and Music Innovative Technologies Overall GPA: 4.15 / 4.30
 Supervisor: Tai-Shih Chi. Thesis: *Detection of Common Mistakes in Novice Violin Playing* (ISMIR 2015)

RESEARCH INTEREST

Deep Generative Models, Representation Learning, Unsupervised Learning, Controllable Audio Generation

My research focuses on deep generative models and the development of inductive biases through self-supervised or unsupervised techniques for representation learning. My domain of interest is audio, including music, singing voice, and speech. The main goal is to decompose the data into semantically meaningful features and use these features as modular and reusable building blocks to generate novel data in a generalisable and composable manner.

SELECTED PUBLICATIONS

- Y.-J. Luo, K. W. Cheuk, W. S. Choi, W.-H. Liao, K. Toyama, T. Uesaka, K. Saito, C.-H. Lai, Y. Takida, S. Dixon, Y. Mitsufuji. “*Disentangling Multi-Instrument Music Audio for Source-Level Pitch and Timbre Manipulation*,” NeurIPS Workshop 2024.
- Y.-J. Luo, S. Dixon. “*Posterior Variance-Parameterised Gaussian Dropout: Improving Disentangled Sequential Autoencoders for Zero-Shot Voice Conversion*,” ICASSP 2024.
- Y.-J. Luo, S. Ewert, S. Dixon. “*Towards Robust Unsupervised Disentanglement of Sequential Data – A Case Study Using Music Audio*,” IJCAI 2022 (Main Track at 15% acceptance rate).
- Y.-J. Luo, K. W. Cheuk, T. Nakano, M. Goto, D. Herremans. “*Unsupervised Disentanglement of Pitch and Timbre for Isolated Musical Instrument Sounds*,” ISMIR 2020.
- Y.-J. Luo, C.-C. Hsu, K. Agres, D. Herremans. “*Singing Voice Conversion with Disentangled Representations of Singer and Vocal Technique Using Variational Autoencoders*,” ICASSP 2020.
- Y.-J. Luo, K. Agres, D. Herremans. “*Learning Disentangled Representations of Timbre and Pitch for Musical Instrument Sounds Using Gaussian Mixture Variational Autoencoders*,” ISMIR 2019.

RESEARCH EXPERIENCE

- Stability AI, UK** **Sep. 2024 - Present**
 Research Intern, *Semantically Enriched Audio Codecs*
- Music Foundation Model Team, Sony AI, Japan** **Nov. 2023 - Apr. 2024**
 Research Intern, *Source-level Pitch/Timbre Manipulation for Multi-instrument Mixtures*
- Addressed pitch/timbre disentanglement from mixtures of musical instruments, and proposed a diffusion transformer to sample mixtures of novel pitch/timbre combinations (NeurIPS 24 Workshop).
- Centre for Digital Music, Queen Mary University of London, England** **Sep. 2020 - Present**
 Ph.D. Researcher, Supported by Spotify, *Unsupervised and Robust Disentangled Representation Learning (DRL)*
- Proposed a simple Gaussian dropout to enhance multiple unsupervised voice conversion models (ICASSP 24).
 - Devised strategies to more robustly achieve unsupervised DRL of time-invariant and time-varying attributes with real-world music instrument recordings (IJCAI 22), extended for improved audio quality (ICASSP 24).
- Media Interaction Group, National Institute of AIST, Japan** **Mar. 2020 - June 2020**
 Research Intern, *Unsupervised Disentanglement of Pitch and Timbre for Isolated Musical Instrument Sounds*
- Proposed an unsupervised approach to disentangle audio spectrums of musical instruments into discrete pitch and continuous timbre representations (ISMIR 20).
- AMAAI, Singapore University of Technology and Design, Singapore** **Sep. 2018 - July 2020**
 Postgraduate Research, *Disentangled Representation Learning Using Variational Autoencoders (VAEs)*
- Developed Gaussian mixture VAEs to disentangle pitch and timbre of musical instruments for sound synthesis (ISMIR 19) and singer identity and vocal technique for singing voice conversion (ICASSP 20).
- Music and Culture Technology Lab, Academia Sinica, Taiwan** **Apr. 2017 - Aug. 2018**
 Postgraduate Research, *Singing Voice Pitch Correction and Separation by Domain Alignment*
- Developed an autoencoder framework to align data representations across domains (ISMIR/ICASSP 18).

SPECIALISED COURSEWORK & SKILLS

- Coursework: Deep Learning for Audio and Music, Music Informatics, Bayesian Methods for Machine Learning
- Programming: Python, PyTorch, PyTorch Lightning, librosa, scikit-learn