YUSONG WU (吴雨松)

University of Montréal, Montréal, Canada

+1-438-941-7566 | wuyusongwys@gmail.com | https://lukewys.github.io/ | Google Scholar

EDUCATION

Beijing University of Posts and Telecommunications Beijing, China

BE in Automation 09/2016 - 06/2020

University of Montréal & Quebec Artificial Intelligence Institute (MILA) Montréal, Canada

Advisor: Prof. Aaron Courville, Prof. Chengzhi Anna Huang

Ph.D. in Computer Science – Artificial Intelligence 09/2022 - now

MSc in Computer Science – Artificial Intelligence | GPA: 4.3/4.3 09/2020 - 09/2022

SELECTED PUBLICATIONS

- **Yusong Wu**, Christos Tsirigotis, Ke Chen, Cheng-Zhi Anna Huang, Aaron Courville, Oriol Nieto, Prem Seetharaman, Justin Salamon: *FLAM: Frame-Wise Language-Audio Modeling*. **Accepted by ICML 2025**
- **Yusong Wu**, Tim Cooijmans, Kyle Kastner, Adam Roberts, Ian Simon, Alexander Scarlatos, Chris Donahue, Cassie Tarakajian, Shayegan Omidshafiei, Aaron Courville, Pablo Samuel Castro, Natasha Jaques, Cheng-Zhi Anna Huang: *Adaptive Accompaniment with ReaLchords*. **ICML 2024**
- Ke Chen, **Yusong Wu**, Haohe Liu, Marianna Nezhurina, Taylor Berg-Kirkpatrick, Shlomo Dubnov: *MusicLDM: Enhancing Novelty in Text-to-Music Generation Using Beat-Synchronous Mixup Strategies.* **ICASSP 2024**.
- **Yusong Wu***, Ke Chen*, Tianyu Zhang, Yuchen Hui, Taylor Berg-Kirkpatrick, Shlomo Dubnov: *Large-scale Contrastive Language-Audio Pretraining with Feature Fusion and Keyword-to-Caption Augmentation*. **ICASSP 2023**
- **Yusong Wu**, Josh Gardner, Ethan Manilow, Ian Simon, Curtis Hawthorne, Jesse Engel: *The Chamber Ensemble Generator: Limitless High-Quality MIR Data via Generative Modeling*. *arXiv preprint arXiv:2209.14458*.
- **Yusong Wu**, Ethan Manilow, Yi Deng, Rigel Swavely, Kyle Kastner, Tim Cooijmans, Aaron Courville, Cheng-Zhi Anna Huang, Jesse Engel: *MIDI-DDSP: detailed control of musical performance via hierarchical modeling.* **ICLR 2022 Oral (top 5%)**
- **Yusong Wu**, Kun Chen, Ziyue Wang, Xuan Zhang, Fudong Nian, Xi Shao, Shengchen Li: *Audio Captioning Based on Transformer and Pre-Training for 2020 DCASE Audio Captioning Challenge.* Technical Report, **DCASE2020 Challenge (2nd place)**
- **Yusong Wu**, Shengchen Li, Chenzhu Yu, Heng Lu, Chao Weng, Dong Yu: *Peking Opera Synthesis via Duration Informed Attention Network.* **INTERSPEECH 2020**
- Liqiang Zhang, Chengzhu Yu, Heng Lu, Chao Weng, **Yusong Wu**, Xiang Xie, Zijin Li, Dong Yu: *DurIAN-SC: Duration Informed Attention Network based Singing Voice Conversion System*. **INTERSPEECH 2020**

SELECTED RESEARCH EXPERIENCE

Interactive Real-time Music Accompaniment Models 07/2022-now

Ph.D. Research Topic

Ongoing Projects:

- **MARLChords**: co-adaptive real-time melody-to-chord accompaniment agent trained with multi-agent RL.
- **Online Stem Gen**: supervised real-time mixture-to-stem music accompaniment in audio domain.

Real-time Melody-to-Chord Accompaniment with Reinforcement Learning

Student Researcher at Google Magenta team, Deepmind 08/2023-05/2024

- Focus on building generative models that support real-time music interaction with users.
- Leverage RL and Multi-agent RL to train generative models that anticipate and adapt to real-time human inputs, and effectively recover from error.
- ReaLchords: real-time melody-to-chord accompaniment model via RL finetuning.
- ReaLJam: real-time interactive system with delay-tolerance model inference and anticipatory model output visualization.

Open-Vocabulary Audio Event Localization 07/2024-12/2024

Internship project at Adobe

- Build FLAM, a frame-wise audio-language contrastive model that localizes the occurrence of acoustic events described in any text, in addition to traditional ranking and retrieval of audio samples.
- Synthesized a large-scale (1M samples) dataset for open-vocabulary audio event localization via automated data augmentation of diverse text-labeled acoustic events.
- Developed a logit adjustment technique for debiasing frame-level contrastive training, effectively addressing label imbalance and spurious correlations, resulting in significantly improved event localization accuracy.

Audio-Language Multimodal Learning 03/2022-05/2023

Collaborate with LAION.ai, and work as project lead

- Collect text-audio data, train multi-model understanding models between language and audio, and train text-to-audio synthesis models.
- As the initial work, build an audio-language contrastive learning model. Collected the biggest audio-text dataset at the time containing 630k pairs. Opensourced Model GitHub Repo & Dataset GitHub Repo with 2k+ stars combined.

Hierarchical Music Generation with Detailed Control 09/2020 – 07/2022

Collaborate with advisors and members of Google Magenta team

- Propose MIDI-DDSP, a hierarchical music generation model with explicit and interpretable representation for controlling musical performance and synthesis.
- MIDI-DDSP can reconstruct high-fidelity audio, accurately predict performance attributes for a note sequence, independently manipulate the attributes of a given performance, and as a complete system, generate realistic audio from a novel note sequence.
- GitHub repo with 300+ stars

2nd Place in DCASE 2020 Challenge Task 6: Automatic Audio Captioning 03/2020 - 07/2020

Singing Synthesis System 08/2019 – 05/2020

Research Intern, Tencent AI Lab.

MUSIC EXPERIENCE

- Over 10 years of percussion experience in orchestra, wind symphony, and marching band. Proficient in Timpani. Started playing percussion at age 6.
- Played with famous Chinese pop singer Jie Zhang in 2016 on the show "Singer".
- Participant in AI Song Contest 2022, won 3rd place.

COMMUNITY CONTRIBUTION AND ACADEMIC ACTIVITY

- Guest lecture on ReaLchords & ReaLJam at courses by Prof. Julian McAuley & Shlomo Dubnov 05/2025
- Invited talk on ReaLchords at Musical AI agent meeting held by SFU 04/2025
- Invited talk at Prof. Dina Katabi's lab meeting on neural audio codec, MIT 12/2024
- Poster presentation on ReaLchords at NEMISIG 2024 03/2024
- Co-organize Music+AI Reading Group: YouTube Channel 2022-2023
- ISMIR 2022 tutorial - T3(M): Designing Controllable Synthesis System for Musical Signals 2022
- Talk on MIDI-DDSP at CMU music reading group & Bryan Pardo's lab at North Western University 2022

ACADEMIC REVIEW EXPERIENCE

- **Neurips 2023, ICLR 2025, SIGGRAPH 2023, ICML 2025, ISMIR 2025**

SELECTED AWARD

- 3rd place in AI Song Contest 2022 2022
- Outstanding Paper – 1st CtrlGen Workshop at NeurIPS 2021 2021
- Reproducible System Award & 2nd Place on automated audio captioning – DCASE Challenge 2020
- 2rd Prize of Academic Scholarship (Top 15%) 2019
- Gold Price in Beijing University Orchestra Performance 2018