

Yusong Wu (吴雨松)

🏠 lukewys.github.io

✉ wuyusongwys@gmail.com

in Yusong Wu | 🌐 lukewys | 🎓 Yusong Wu

EDUCATION

University of Montréal & Mila, Quebec AI Institute

Montréal, Canada

- Ph.D. Candidate in Computer Science - Artificial Intelligence
- M.Sc. in Computer Science - Artificial Intelligence | GPA: 4.3/4.3
- Advisors: Prof. Aaron Courville, Prof. Cheng-Zhi Anna Huang

Sep 2022 - Present

Sep 2020 - Sep 2022

Beijing University of Posts and Telecommunications

Beijing, China

- B.E. in Automation

Sep 2016 - Jun 2020

RESEARCH INTERESTS

I build models that understand and synthesize audio and music from human-centric instructions. My work spans **real-time human-AI music interaction**, **multi-modal audio understanding**, **text-to-music generation**, and **controllable audio synthesis**.

SELECTED OPENSOURCE PROJECTS

- **Contrastive Language-Audio Pretraining (CLAP)** (🌟 1.8K stars): A multi-modal framework that produces both audio and text corresponding representations for broad downstream applications.
- **LAION-Audio-630K** (🌟 712 stars): A large open-source dataset of audio-text pairs used to train CLAP and related models.
- **MIDI-DDSP** (🌟 323 stars): A hierarchical MIDI-to-audio model that infers note, performance, and DDSP synthesis controls for realistic audio, enabling fine control of timbre, vibrato, dynamics, and articulation.
- **MusicLDM** (🌟 178 stars): A text-to-music generation model for investigating the novelty and plagiarism issue on the audio diffusion model via the latent mixup strategy.
- **Chamber Ensemble Generator** (🌟 63 stars): A pipeline that couples Coconet note generation with MIDI-DDSP synthesis to create realistic four-part chorales with rich aligned annotations (MIDI, per-note expression, f_0 , stems, mixes), released as CocoChorales and shown to aid transcription and source separation.

PUBLICATIONS

🎓 Google Scholar (* indicates the equal contribution)

Real-time Human-AI Music Interaction

- [1] **Yusong Wu**, S. Brade, T. Ma, T.-J. Fowler, E. Yang, B. Banar, A. Courville, N. Jaques, and C.-Z. A. Huang, “Generative Adversarial Post-Training Mitigates Reward Hacking in Live Human-AI Music Interaction”, in *In submission to ICLR 2026*.
- [2] **Yusong Wu**, M. Wang, H. Lei, S. Brade, L. Blanchard, S.-L. W. A. Courville, and A. Huang, “Streaming Generation for Music Accompaniment”, in *In submission to ICASSP 2026*.

- [3] **Yusong Wu**, T. Cooijmans, K. Kastner, A. Roberts, I. Simon, A. Scarlatos, C. Donahue, C. Tarakajian, S. Omidshafiei, A. Courville, P. S. Castro, N. Jaques, and C.-Z. A. Huang, “Adaptive Accompaniment with ReaLchords”, in *ICML 2024*.
- [4] A. Scarlatos, **Yusong Wu**, I. Simon, A. Roberts, T. Cooijmans, N. Jaques, C. Tarakajian, and C.-Z. A. Huang, “ReaLJam: Real-Time Human-AI Music Jamming with Reinforcement Learning-Tuned Transformers”, in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA 2025)*.

Multi-modal Audio Understanding

- [5] **Yusong Wu**, C. Tsirigotis, K. Chen, C.-Z. A. Huang, A. Courville, O. Nieto, P. Seetharaman, and J. Salamon, “FLAM: Frame-Wise Language-Audio Modeling”, in *ICML 2025*.
- [8] **Yusong Wu***, K. Chen*, T. Zhang*, Y. Hui*, T. Berg-Kirkpatrick, and S. Dubnov, “Large-scale Contrastive Language-Audio Pretraining with Feature Fusion and Keyword-to-Caption Augmentation”, in *ICASSP 2023*.
- [10] **Yusong Wu**, K. Chen, Z. Wang, X. Zhang, F. Nian, X. Shao, and S. Li, “Audio Captioning Based on Transformer and Pre-Training for 2020 DCASE Audio Captioning Challenge”, in *DCASE 2020 Challenge and 2nd place*.

Music Generation

- [6] **Yusong Wu**, E. Manilow, Y. Deng, R. Swavely, K. Kastner, T. Cooijmans, A. Courville, C.-Z. A. Huang, and J. Engel, “MIDI-DDSP: Detailed Control of Musical Performance via Hierarchical Modeling”, in *ICLR 2022 (Oral and top 5% and tied for first at score release)*.
- [7] K. Chen, **Yusong Wu**, H. Liu, M. Nezhurina, T. Berg-Kirkpatrick, and S. Dubnov, “MusicLDM: Enhancing Novelty in Text-to-Music Generation Using Beat-Synchronous Mixup Strategies”, in *ICASSP 2024*.
- [9] **Yusong Wu**, J. Gardner, E. Manilow, I. Simon, C. Hawthorne, and J. Engel, “The Chamber Ensemble Generator: Limitless High-Quality MIR Data via Generative Modeling”, in *arXiv preprint arXiv:2209.14458*.
- [11] **Yusong Wu**, S. Li, C. Yu, H. Lu, C. Weng, and D. Yu, “Peking Opera Synthesis via Duration Informed Attention Network”, in *INTERSPEECH 2020*.
- [12] L. Zhang, C. Yu, H. Lu, C. Weng, **Yusong Wu**, X. Xie, Z. Li, and D. Yu, “DurIAN-SC: Duration Informed Attention Network based Singing Voice Conversion System”, in *INTERSPEECH 2020*.

Other

- [13] **Yusong Wu** and S. Li, “Guqin Dataset: A Symbolic Music Dataset of Chinese Guqin Collection”, in *China Conference on Sound and Music Technology (CSMT 2019)*.
- [14] **Yusong Wu** and S. Li, “Distinguishing Chinese Guqin and Western Baroque Pieces Based on Statistical Model Analysis of Melodies”, in *International Symposium on Computer Music Multidisciplinary Research (CMMR 2019)*.
- [15] **Yusong Wu**, K. Kastner, T. Cooijmans, C.-Z. A. Huang, and A. Courville, “Datasets That Are Not: Evolving Novelty Through Sparsity and Iterated Learning”, in *Workshop on Machine Learning for Creativity and Design and NeurIPS 2022*.

INDUSTRY INTERN EXPERIENCE

Adobe Research

Research Intern by Justin Salamon, Prem Seetharaman, and Oriol Nieto

Remote

July 2024 - Dec 2024

- Design open-vocabulary sound event detection via frame-wise language-audio contrastive training.

Google DeepMind

Student Researcher by Cheng-Zhi Anna Huang, Pablo Samuel Castro, and Natasha Jaques

Montréal, Canada

Aug 2023 - May 2024

- Build real-time human-AI music interaction models for melody-to-chord accompaniment.

Tencent AI Lab

Research Intern by Chengzhu Yu and Heng Lu

Beijing, China

Aug 2019 - May 2020

- Build singing synthesis systems for Peking Opera and singing voice transfer.

ACADEMIC SERVICE

Invited Talks & Presentations

· Guest lecture on ReaLchords & ReaLJam at Prof. Julian McAuley’s course, UCSD	May 2025
· Invited talk on ReaLchords at Musical AI Agent meeting hosted by SFU	Apr 2025
· Invited talk on neural audio codec at Prof. Dina Katabi’s lab meeting, MIT	Dec 2024
· Poster presentation on ReaLchords at NEMISIG 2024	Mar 2024
· ISMIR 2022 tutorial T3(M): Designing Controllable Synthesis Systems for Musical Signals	2022
· Talk on MIDI-DDSP at CMU music reading group & Bryan Pardo’s lab, Northwestern University	2022
· Talk on MIDI-DDSP at MLNLP community	2022

Community Contribution

· Co-organized Music+AI Reading Group	2022 - 2023
---------------------------------------	-------------

Reviewing

· Neurips 2023, SIGGRAPH 2023, ICML 2025, ISMIR 2025, ICLR 2025 & 2026
--

SELECTED AWARDS

• 3rd Place , AI Song Contest 2022	2022
• Outstanding Paper , 1st CtrlGen Workshop at NeurIPS	2021
• Reproducible System Award & 2nd Place , DCASE Challenge 2020 Task 6	2020
• 2nd Prize Academic Scholarship (Top 15%)	2019
• Gold Prize , Beijing University Orchestra Performance	2018

SKILLS

- **Language:** Native in Chinese; Fluent in English
- **Programming:** Python, JavaScript, HTML
- **Frameworks:** PyTorch, JAX, TensorFlow
- **Audio Processing in Python:** Librosa, Torchaudio, Audiotools
- **Hobbies:** Snowboarding, Hiking, Badminton, Harmonica, General aviation flying

MUSIC EXPERIENCE

- Over 10 years of percussion experience in orchestra, wind symphony, and marching band; proficient in **timpani**.
- Performed with school symphony orchestra in Musikverein Golden Hall, 2007. Won gold medal for Summa Cum Laude International Youth Music Festival.
- Performed with Chinese pop singer Jie Zhang on the TV show “Singer” in 2016.