Shih-Lun Wu

Student, **M.Sc. in Language Technologies (MLT)**, School of Computer Science
Carnegie Mellon University (CMU), Pittsburgh, PA, United States
Email: shihlunw@andrew.cmu.edu | Homepage: https://slseanwu.github.io | Google Scholar | GitHub | LinkedIn

EDUCATION

Master of Science (M.Sc.) | Carnegie Mellon University 08.2022 ~ 08.2024 (expected)
Language Technologies major

- Cumulative QPA -- **4.17/4.33**
- Research areas: **Music & Audio Processing, Multimodal Learning**
- Advisors: Dr. Chris Donahue, Dr. Shinji Watanabe

Bachelor of Science (B.Sc.) | National Taiwan University 09.2017 ~ 06.2021
Computer Science major · **Economics** minor

- Cumulative GPA -- Overall: **4.28/4.30**, Major: **4.28/4.30**, Rank: **1/176**
- Bachelor's thesis: **"Bridging Transformers and Latent Variable Models for User-controllable Conditional Music Generation."** Committee members: Dr. Yi-Hsuan Yang, Dr. Yun-Nung Chen, Dr. Lin-shan Lee [pdf] [defense slides]

HONORS & RECOGNITION

- **Citation count** (Google Scholar)**: 190+** total**, 125+** first-author | **GitHub stars: 400+**

Siebel Scholar, Class of 2024 | The Siebel Foundation 09.2023
- Awarded to 83 graduate students worldwide for outstanding research & leadership ($35K prize money)

Winner (Research Org), Intern Project Showcase | Adobe Inc. 08.2023
- Won with the Music ControlNet and related music generation projects, against 200+ Adobe research interns

1st Prize, Automated Audio Captioning Challenge | DCASE 2023 06.2023
- Leveraged ChatGPT mix-up augmentations and LLM embedding supervision to achieve new SoTA (publication [7])

1st Prize (Ssu-Nien Fu's Award), Best Bachelor's Thesis | National Taiwan University 06.2021
- Awarded to 6 out of 3500+ students in the graduating class for exceptional undergrad research

SELECTED PUBLICATIONS

[8] **Shih-Lun Wu**, Chris Donahue, Shinji Watanabe, Nicholas J. Bryan. "Music ControlNet: Multiple Time-varying Controls for Music Generation." Working manuscript, to be submitted to *IEEE/ACM Transactions on Audio, Speech, & Language Processing (TASLP)*.

[7] **Shih-Lun Wu**, Xuankai Chang, Gordon Wichern, Jee-weon Jung, François Germain, Jonathan Le Roux, and Shinji Watanabe. "Improving Audio Captioning Models with Fine-grained Audio Features, Text Embedding Supervision, and LLM Mix-up Augmentation." Under review at *Int. Conf. on Acoustics, Speech, & Signal Processing (ICASSP)* 2024. [pdf] [DCASE challenge results]

[6] **Shih-Lun Wu**, Yi-Hui Chou, and Liangze Li. "Listener Model for the PhotoBook Referential Game with CLIPScores as Implicit Reference Chain." *Annual Meeting of the Assoc. for Computational Linguistics **(ACL)*** 2023. [pdf] [code]

[5] **Shih-Lun Wu** and Yi-Hsuan Yang. "Compose & Embellish: Well-structured Piano Performance Generation via A Two-Stage Approach." *Int. Conf. on Acoustics, Speech, & Signal Processing **(ICASSP)*** 2023. [pdf] [code]

[4] **Shih-Lun Wu** and Yi-Hsuan Yang. "MuseMorphose: Full-song and Fine-grained Music Style Transfer with One Transformer VAE." *IEEE/ACM Transactions on Audio, Speech, & Language Processing **(TASLP)*** 2023. [pdf] [code] [project website]

[3] Antoine Liutkus, Ondřej Cífka, **Shih-Lun Wu**, Umut Simsekli, Yi-Hsuan Yang, and Gaël Richard. "Relative Positional Encoding for Transformers with Linear Complexity." *International Conference on Machine Learning **(ICML)*** 2021. (**Long talk**, acceptance rate: **3.0%**) [pdf] [code] [presentation video] [project website]

[2] **Shih-Lun Wu** and Yi-Hsuan Yang. "The Jazz Transformer on the Front Line: Exploring the Shortcomings of AI-Composed Music through Quantitative Measures." *International Society for Music Information Retrieval Conference **(ISMIR)*** 2020. [pdf] [code] [poster] [presentation video]

[1] **Shih-Lun Wu***, Ching-Yuan Bai*, Kai-Chieh Chang, Yi-Ting Shieh, Chao Huang, Chung-Wei Lin, Eunsuk Kang and Qi Zhu. "Efficient System Verification with Multiple Weakly-hard Constraints for Runtime Monitoring." *International Conference on Runtime Verification **(RV)*** 2020. (*: equal contribution) [pdf] [publisher page]

Shih-Lun Wu | M.Sc. student, Carnegie Mellon University | shihlunw@andrew.cmu.edu

RESEARCH EXPERIENCE

Research Scientist/Engineer Intern | Adobe Research 05.2023 ~

Audio AI Lab. *Supervisors: Dr. Nick Bryan, Dr. Gautham Mysore*

- Invented Music ControlNet, enabling precise melody, dynamics, rhythm controls for diffusion music generators
- Demonstrated compositionality of proposed controls, and out-of-domain generalizability to user-specified controls
- Outperformed Meta AI's MusicGen by 49% on melody control, despite using 35x fewer params & 11x less data

Research Engineer | Taiwan AI Labs 08.2021 ~ 03.2022

Research Intern | Taiwan AI Labs 07.2020 ~ 07.2021

AI Music Team. *Supervisor: Dr. Yi-Hsuan Yang*

- Designed mechanisms to exert time-varying control on Transformers for sequence generation (publication [4])
- Bridged Transformers, the mechanism above, and Variational Autoencoders for fine-grained style transfer of long musical pieces, allowing users to harness harmonic & rhythmic intensities down to the bar level (publication [4])
- Made a 3-stage model to generate well-structured music with recurring & developing content (some results in [5])

Undergraduate Research Assistant | Academia Sinica 02.2020 ~ 06.2021

Music and AI Lab, Research Center for IT Innovation. *Advisor: Dr. Yi-Hsuan Yang*

- Collaborated with researchers @ INRIA / Télécom Paris on positional encodings for O(n) Transformers (publ. [3])
- Developed a set of widely-used quantitative metrics to assess the quality of machine-generated music (publ. [2])
- Improved musical structure of Transformer-generated pieces by inserting structure-related tokens (publ. [2])

Undergraduate Research Assistant | National Taiwan University 02.2019 ~ 06.2020

Cyber-Physical Systems Lab, Dept. of CSIE. *Advisor: Dr. Chung-Wei Lin*

- Formulated the formal verification problem under multiple weakly-hard constraints on environmental faults
- Discovered and proved the mathematical properties between pairs of weakly-hard constraints
- Devised a lowest-cost-first heuristic using the properties, accelerating verification algorithm by up to 12x (publ. [1])

OTHER WORK EXPERIENCE

Software Engineering Intern | Asus Inc. 07.2019 ~ 08.2019

Cloud Infrastructure Team, Asus Intelligent Cloud Services (AICS) Center

- Developed a Kubernetes + Python (Flask) template for launching containerized, cloud-based ML solutions
- Integrated Azure Key Vault, Mutual TLS auth & Azure App Insights to the template to streamline model deployment

OTHER SELECTED PROJECTS

MuseOptimus: Interactive AI Composition Webapp | *React · Flask · PyTorch* 01.2021

- Realized an immersive user interface for the music generation model developed by me @ Taiwan AI Labs
- Implemented interactive features, including dynamic note display, song rating, and song recommendation
- Scored the highest among 100+ final projects in NTU's Web Programming course (by Prof. Ric Huang) [slides]

EXTRACURRICULAR ACTIVITIES & SERVICE

Pianist, Violist, & Director of General Affairs 09.2018 ~ 06.2021

Symphony Orchestra, National Taiwan University

- Participated actively in concerts [playlist] and handled procurement, musical scores, and transportation affairs

Peer Reviewer

- Conferences: **ICMLA** (2020), **ISMIR** (2021, 2022, 2023), **ICASSP** (2024)
- Journals: **TISMIR** (2021), **ACM Computing Surveys** (2023)

Teaching Assistant

- **Algorithms Design and Analysis** (NTU, Fall 2019)

SKILLS & QUALIFICATIONS

- Programming Languages & Infrastructure: *Python · C/C++ · JavaScript · ReactJS · LaTeX · Linux · Kubernetes*
- Machine Learning Frameworks*: PyTorch · Keras · Tensorflow · HuggingFace · PyTorch Lightning*
- Selected Coursework: **Straight A+'s** in the following courses
  -- CS fundamentals: DS & Algo, Algo Design & Analysis, Formal Language & Automata, Linear Algebra
  -- ML-/DL-related: ML Techniques, Advanced NLP, Speech Recognition & Understanding, Multimodal ML