Wei ZENG

(+65) · 9379-9504 ◇ w.zeng@u.nus.edu ◇ Homepage ◇ LinkedIn

EDUCATION

National University of Singapore Aug. 2023 – May. 2027 (Expected)
- **Ph.D.** *in Computer Science* *Overall GPA: 4.85/5.00*
  - *Advisor: Prof. Ye WANG, Sound and Music Computing (SMC) Lab*

National University of Singapore Aug. 2021 – Jun. 2023
- **M.Eng.** *in Mechanical Engineering* *Overall GPA: 4.63/5.00*
  - *Advisor: Prof. Chee Kong CHUI*

Shanghai Jiao Tong University Sep. 2016 – Jun. 2020
- **B.Eng.** *in Energy & Power Engineering (Major)* *Overall GPA: 3.45/4.00 (84.7)*
- **B.A.** *in Musicology (Minor)* *Overall GPA: 3.92/4.00 (92.3)*

PUBLICATIONS

[1] **W. Zeng**, J. Zhao, and Y. Wang, "Bridging Piano Transcription and Rendering via Disentangled Score Content and Style," *The Fourteenth International Conference on Learning Representations* **(ICLR)** 2026, accepted. arXiv:2509.23878

[2] X. He, **W. Zeng**, and Y. Wang, "Peransformer: Improving Low-informed Expressive Performance Rendering with Score-aware Discriminator," *Proc. APSIPA Annual Summit and Conference* **(APSIPA ASC)** 2025, pp. 2353–2358. DOI: 10.1109/APSIPAASC65261.2025.11249357

[3] **W. Zeng**, X. He, and Y. Wang, "End-to-End Real-World Polyphonic Piano Audio-to-Score Transcription with Hierarchical Decoding," *Proc. Int. Joint Conf. on Artificial Intelligence* **(IJCAI)** 2024, pp. 7788–7795. DOI: 10.24963/ijcai.2024/862

[4] X. Gu, L. Ou, **W. Zeng**, J. Zhang, N. Wong, and Y. Wang, "Automatic Lyric Transcription and Automatic Music Transcription from Multimodal Singing," *ACM Trans. Multim. Comput. Commun. Appl.* **(TOMM)** 20(7): 209:1–209:29, 2024. DOI: 10.1145/3651310

[5] X. Gu, **W. Zeng**, and Y. Wang, "Elucidate Gender Fairness in Singing Voice Transcription," *Proc. ACM Int. Conf. on Multimedia* **(MM)** 2023, pp. 8760–8769. DOI: 10.1145/3581783.3612272

PREPRINTS

[1] J. Zhao, **W. Zeng**, T. Lyu, and Y. Wang, "CoMelSinger: Discrete Token-Based Zero-Shot Singing Synthesis With Structured Melody Control and Guidance," 2025. arXiv:2509.19883

RESEARCH EXPERIENCE

Zero-Shot Singing Voice Synthesis with Melody Control Jan. 2025 – Sep. 2025
Advisor: Prof. Ye Wang, SMC Lab, National University of Singapore

- Proposed *CoMelSinger*, a discrete token-based zero-shot singing voice synthesis framework built upon a non-autoregressive MaskGCT architecture, enabling structured and disentangled melody control from lyrics and pitch tokens
- Designed a coarse-to-fine contrastive learning strategy at both sequence and frame levels to suppress prosody leakage from acoustic prompts and improve melody–timbre disentanglement
- Developed a lightweight encoder-only Singing Voice Transcription (SVT) module to provide frame-level pitch supervision, enhancing pitch accuracy and rhythmic stability under zero-shot settings

Joint Piano Transcription and Performance Rendering Sep. 2024 – May 2025
Advisor: Prof. Ye Wang, SMC Lab, National University of Singapore

- Proposed a unified transformer-based sequence-to-sequence framework that jointly models Automatic Piano Transcription (APT) and Expressive Performance Rendering (EPR) by disentangling note-level score content and global performance style representations
- Formulated EPR as an alignment-free Seq2Seq task, enabling scalable training using only sequence-aligned data without requiring fine-grained note-level alignment
- Designed a diffusion-based Performance Style Recommendation (PSR) module to generate stylistically appropriate performance embeddings directly from score content, supporting controllable rendering and style transfer

Hierarchical Audio-to-Score Transcription for Piano Aug. 2023 – Feb. 2024
Advisor: Prof. Ye Wang, SMC Lab, National University of Singapore

- Proposed an end-to-end Seq2Seq piano audio-to-score transcription framework with a hierarchical decoder to jointly predict bar-level musical structure (key and time signatures) and note-level score content
- Designed a multi-task learning architecture with separate bar-level and staff-wise note-level decoders, enabling transcription of polyphonic piano music with unconstrained voices
- Developed a two-stage training scheme leveraging expressive performance rendering (EPR) for pre-training on synthetic audio and fine-tuning on real human performance recordings to bridge the synthetic–real domain gap

INTERNSHIP EXPERIENCE

Shanghai Artificial Intelligence Research Institute Nov. 2020 – Apr. 2021
Mentor: Zikai Wang, NLP Intern

- Created database and designed corresponding data tables using MySQL; merged and inserted about 150 million JSON lines from two open source thesis paper databases MAG and AMiner
- Fine-tuned a BERT-based model with supervised learning using paper abstracts and their category labels; generated global semantic embeddings for each paper abstract and stored them in a MySQL database, then built a FAISS-based semantic retrieval system enabling natural language queries to retrieve the most relevant academic papers
- Demonstrated the effectiveness of semantic embedding–based retrieval for large-scale academic literature, enabling efficient topic-level exploration across heterogeneous paper corpora

MISCELLANEOUS

Languages	Mandarin (first language), English (TOEFL 102, GRE 324)	
Skills	Python, MATLAB, C++	
Service	Member of Student Area Search Committee, SoC	AY2025
	Reviewer for TASLP, ICLR, etc.	
Teaching	IT5005 Artificial Intelligence	2024 – 2026
	CS5647 Sound and Music Computing	Sem1 AY2025/26
Awards	SoC Honour List of Student Tutors	AY2024/25
	Teaching Fellowship of SoC	2025
	Research Achievement Award of SoC	AY2023/24
	President's Graduate Fellowship	2023 – 2027
	Academic Excellence Scholarship of SJTU	2016 – 2020