

# Xintong Wang

AI RESEARCH ENGINEER · SPEECH AND SINGING VOICE

Singapore

☎ 65-9120-7974 | ✉ xintongwang9709@gmail.com | 🏠 <https://oshindow.github.io/>

“Be the change that you want to see in the world.”

## Summary

---

Research Assistant in School of Computing at National University of Singapore. Machine Learning Engineer in xiaobing.ai (formerly Microsoft Xiaoice team). 3+ years of experience specializing in AI speech and singing voice, large-scale software development. Always keep a sharp mind in both academic and industry. Super excited about making an impactful product with new possibilities.

## Work Experience

---

### Sound and Music Computing Lab

Singapore, Singapore

RESEARCH ASSISTANT · SCHOOL OF COMPUTING · NATIONAL UNIVERSITY OF SINGAPORE

Oct. 2023 - Present

- Leading the team to develop and deploy a Mandarin Chinese speech assessment system.
- Collaborated with the medical school of NUS to facilitate the learning of Chinese for Singaporean medical students.
- Deeply optimized ASR performance for second-language mispronunciation detection and diagnosis.

### xiaobing.ai

Beijing, China

MACHINE LEARNING ENGINEER

Jul. 2022 - Oct. 2023

- Investigated few-shot singing voice conversion model based on self-supervised learning model
- Developed a multi-language acoustic model via Generative Adversarial Network, which is a single model for three languages (Mandarin/English/Japanese), transforming a monolingual singer into a multilingual singer
- Introduced energy parameter into acoustic model and develop an energy parameter prediction model.
- Developed and robusted a iSTFT-based vocoder via Generative Adversarial Network, which is predicting phase using STFT spectrogram, generating 48k high-fidelity audio high consistency with the ground truth.
- Investigated diffusion-based acoustic models, such as DiffSinger.
- Deployed a CPU-based gRPC speech recognition service using the Wenet framework, with the Wenet pre-trained model for Chinese trained on 10,000 hours.
- Developed a phoneme correction module based on carefully annotated data to correct the position of voiced consonants after automatic alignment by MFA. Further develop phoneme alignment models based on HMM-GMM and CTC, using Kaldi, k2, Funasr toolkits.

### xiaobing.ai

Beijing, China

SPEECH RECOGNITION INTERN

May 2021 - Jul. 2022

- Build, prune and merge N-gram Language Models from large scale corpus of several domains.
- Compile Language Model, Pronunciation Lexicon and Acoustic Model into an HCLG.fst, and decode with Kaldi to obtain a better CER on benchmarks.
- Given a MIDI library and a specific chords progression, develop a program that automatically selects segments from the library that match the given chords progression. Create a new musical score using these selected segments, and use X-Studio 1.8 to synthesize a mashup song.

## Publications

---

### CrossSinger: A Cross-Lingual Multi-Singer High-Fidelity Singing Voice Synthesizer Trained on Monolingual Singers

XINTONG WANG, CHANG ZENG, JUN CHEN, CHUNHUI WANG

ASRU2023

### A 2D Convolutional Gating Mechanism for Mandarin Streaming Speech Recognition

XINTONG WANG, CHUANGANG ZHAO

Information, 12.4 (2021): 165.

## Education

---

### Beijing Forestry University

Beijing, China

B.S. IN MATHEMATICS AND APPLIED MATHEMATICS

2018 - 2022

- Attended a project that developing an ASR system for telecommunications fraud.

# Honors and Awards

---

## Honorable Mention

COMAP MATHEMATICAL CONTEST IN MATHEMATICAL CONTEST IN MODELING

2020

## Outstanding Student Scholarship

COLLAGE OF SCIENCE · BEIJING FORESTRY UNIVERSITY

2019 Fall - 2020 Fall

## Academic Excellence Scholarship

COLLAGE OF SCIENCE · BEIJING FORESTRY UNIVERSITY

2019 Fall - 2020 Fall