

Yinghao Ma

yinghao.ma@qmul.ac.uk | +44 (0)7741575063 | [nicolaus625.github.io](https://github.com/nicolaus625) | [linkedin.com/in/nicolaus625](https://www.linkedin.com/in/nicolaus625)

EDUCATION

Queen Mary University of London (QMUL)

09/2022 – 09/2026 (expected)

PhD: AI & Music, School of Electronic Engineering and Computer Science. Supervised by Dr Emmanouil Benetos

- Research interests: **Self-supervised learning for music information retrieval and multimodality**
- Selected experience: Co-founder of the [Multimodal Art Projection](#) community ([Hugging-face](#));
Established [Music Audio Representation Benchmark for universal Evaluation](#).

Carnegie Mellon University (CMU)

09/2020 – 08/2022

MS: Music & Technology, School of Music. Supervised by Prof. Richard M. Stern

- Overall GPA: 4.03/4.00 (Top 1 of the grade)
- Awards and Honours: Fellowship for graduate students that covers 50% of tuition fee
National Music Honor Society member. (Theta Xi, Pi Kappa Lambda reward).
- Music Background: Recorded Chinese musical version of Beethoven's serenade for 250th anniversary of his birth during COVID to cheer others up. Released on CMU DL course web
- Selected Modules: Speech Understanding; Convex Optimisation; Introduction to Computer Music

Peking University (PKU)

09/2016 – 07/2020

BS: Mathematics & Applied Mathematics (Probability Theory), School of Mathematical Sciences

- Overall GPA: 3.4/4.0
- Awards and Honours: Outstanding graduates of the School of Mathematical Science at PKU
Preliminary excellence prize for S.-T.Yau College Student Mathematics Contests
- Music Background: One of the student conductors in the orchestra of Chinese Music Institute at PKU
The amateur highest level of Chinese flutes, China Conservatory of Music
- Selected Modules: Advanced Theory of Probability; Statistics; Intro to Stochastic Processes; Topology

RESEARCH EXPERIENCE

Bridging Music & Text with Pre-trained Models for Music Captioning and QA

07/2023 – present

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Developed Music Instruct (MI) query-response dataset based on captions & well-designed prompts to GPT-4.
- Achieved cutting-edge performance in question answering on both MusicQA and Music Instruct datasets.
- Employed instruct fine-tuning techniques on MI to attain state-of-the-art (SOTA) results in captioning.

Instrument Playing Technique (IPT) Detection Using Multi-task Finetuning

06/2023 – 09/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Finetuned a pre-trained model, providing SOTA performance on Guzheng, Guitar, and Chinese flute.
- Further improved results by multi-task training with pitch detection and onset detection.
- Post-processed prediction on frame-level, resulting in significant improvement in event-level metrics.

MARBLE: Music Audio Representation Benchmark for universal Evaluation

01/2023 – 06/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Designing the downstream tasks, datasets, evaluation metrics and state-of-the-art results.
- Implementing the mir_eval metrics with torchmetrics and developing utilisation for sequential tasks.
- Establishing a fair, reproducible and universal music information retrieval benchmark for future work.

Acoustic Music Understanding Model with Large-Scale Self-supervised Training

08/2022 – 05/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Built self-supervised learning systems, acquiring 50k+ downloading of checkpoints on Huggingface.
- Replaced the pseudo-tag from MFCCs to Chroma music features for harmonic information.
- Utilising deep features like Encodec instead of k-means for scaling up models to 1 B parameters.

Learnable Front End for Music, Speech and Audio

09/2021 – 07/2022

Research Assistant, Supervised by Prof. Richard Stern, Carnegie Mellon University

- Constructed temporal modulation learnable front ends, surpassing SOTA methods on music tagging.
- Analysis of the model performance among tags with different genres and instrument tags.

Chinese Flute Playing Technique Classification Based on FCNNs (undergraduate thesis)

02/2020 – 05/2020

Research Assistant, Supervised by Prof. Xiaou Chen, Peking University

- Established music technique detectors based on a series of CNNs with different layers as well as FCNNs.
- Extended models with transpose convolution to support variable length inputs and pixel-level classification.

Correspondence between Speech Melody and Pitch Contour in Sichuan Folk Songs 07/2019 – 09/2019

Research Assistant, Supervised by Prof. Zhiyao Duan, University of Rochester

- Set up a database on Sichuan folk songs with music scores in MusicXML form, and lyrics reading recording.
- Analysed the correspondence between the change of music notes and the fundamental frequency of lyrics.

Automatic Musical Instrument Recognition and Timbre Recognition 02/2019 – 07/2019

Research Assistant, Supervised by Prof. Xiaoou Chen, Peking University

- Built a quartet dataset based on Chinese instrumental music recording in the DCMI dataset.
- Implemented an audio event detection model based on CRNNs on Chinese instrument recognition.

PUBLICATION & CONFERENCE SUBMISSION

- Deng, Z., **Ma, Y.**, Liu, Y. et al. "*MusiLingo: Bridging Music and Text with Pre-trained Language Models for Music Captioning and Query Response.*" [Submitted to The European Chapter of the ACL (EACL) 2024.]
- Li, D., **Ma, Y.**, et al. "*Mertech: Instrument Playing Technique Detection Using Self-supervised Pretrained Model with Multi-task Finetuning.*" [Submitted to International Conference on Acoustics, Speech & Signal Processing 2024.]
- Li, Y., Yuan, R.; Zhang, G., **Ma, Y.**; et al. "*MERT: Acoustic Music Understanding Model with Large-Scale Self-supervised Training.*" [Submitted to International Conference on Learning Representations (ICLR) 2024.]
- Yuan, R., **Ma, Y.**, Li, Y., et al. "*MARBLE: Music Audio Representation Benchmark for universal Evaluation.*" Advances in Neural Information Processing Systems (NeurIPS), 2023.
- **Ma, Y.**, Yuan, R., Li, Y., et al. "*On the Effectiveness of Speech Self-Supervised Learning for Music.*" International Society for Music Information Retrieval (ISMIR), 2023.
- Li, Y., Yuan, R., et al. "*Map-music2vec: A simple and effective baseline for self-supervised music audio representation learning.*" International Society for Music Information Retrieval late-breaking demo, 2022.

WORK EXPERIENCE

Tempo, Beat and Downbeat Detection in Chinese Pop Songs (internship) 06/2020 – 08/2020

Algorithm Engineer, Beijing Deepmusic Technology Co.

- Built beat detection pipelines with LSTMs & TCNs, significantly outperforming librosa & madmom libraries.
- Estimated tempo and beat of Chinese pop songs producing 98% accuracy on the tempo prediction.

Cover Song Detection & Evaluation of Automatic Speech Recognition (internship) 05/2021 – 08/2021

Algorithm Engineer, Tencent Holdings Limited. (Beijing)

- Examined and analysed existing models with learnable frontends on proprietary music datasets.

Teaching Assistant of Digital Signal Processing 02/2022 – 05/2022

Teaching Assistant & Guest Lecturer of Machine Learning for Signal Processing 08/2021 – 12/2021

- Delivered lectures on ICA; designed quizzes and assignments on NMF, SVM, HMM, Compressive sensing etc.

SELECTED ACADEMIC ACHIEVEMENT

Advanced Digital Signal Processing

- 1st rank of the class. Proofread lecture notes' errata, listed in acknowledgements.
- Signal sampling, interpolation, STFT, classical estimation and maximum entropy of PSD function, LPC, adaptive filtering, microphone array and beam forming. MFCC, Introduction to Wavelet Analysis.

Introduction to Deep Learning (A)

Scalability Machine Learning (A+)

Convex Optimisation (A+)

LEADERSHIP

- One of the student conductors at the Chinese Music Institute at Peking University. Guided rehearsals of philharmonic chamber and concert. Organised seminar on music theory and music information retrieval.
- One of the core members and seminar organisers of the Undergraduate Music Composition Association.
- Organised charitable activities. social research and publicity for the disabled and people with rare diseases.
- One of the co-founders of volunteer teaching math to the poor county activities as a member of the Association of Young Volunteers in the School of Mathematical Sciences at Peking University.

SKILLS

- Python (>5k LOC)
- MATLAB, C (>500 LOC)