

# Xinhao Mei

☎ +1 4258293065 | @ meixinhao429@gmail.com | 🔗 LinkedIn | 🐙 GitHub | 🌐 WebPage | 📄 Google Scholar

## EDUCATION

---

### University of Surrey

*PhD in Electronic Engineering*  
*Supervisor: Prof. Wenwu Wang*

Guildford, UK

*Jan 2021 – Aug 2024*

### University of Surrey

*Master of Science in Electronic Engineering*  
*Distinction*

Guildford, UK

*Sep 2019 – Sep 2020*

### Southwest Petroleum University

*Bachelor of Engineering in Software Engineering*  
*First Honor*

Chengdu, China

*Sep 2015 – Jun 2019*

## RESEARCH INTERESTS

---

Audio-Language Multimodal Learning, Audio Understanding and Generation

## WORK EXPERIENCE

---

### Meta

*Research Scientist*

Redmond, USA

*Apr 2024 – present*

- Topic: Multi-modality based audio generation

### Meta

*Research Scientist Intern*

Paris, France

*Jul 2023 – Nov 2023*

- Topic: Visual-to-audio generation
- Advisors: Gael Le Lan, Yangyang Shi

### Peking University, Shenzhen Graduate School

*Visiting PhD scholar*

Shenzhen, China

*Oct 2022 – Dec, 2022*

- Topic: Large-scale audio-language multimodal learning
- Supervisor: Prof. Yuexian Zou

## PUBLICATIONS

---

**Xinhao Mei**, Chutong Meng, Haohe Liu, Qiuqiang Kong, Tom Ko, Chengqi Zhao, Mark D Plumbley, Yuexian Zou, and Wenwu Wang. “WavCaps: A ChatGPT-Assisted Weakly-Labelled Audio Captioning Dataset for Audio-Language Multimodal Research,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, 2024.

**Xinhao Mei**, Xubo Liu, Jianyuan Sun, Mark D Plumbley, and Wenwu Wang. “Towards Generating Diverse Audio Captioning via Adversarial Training,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, vol. 32, pp. 3311-3323, 2024.

**Xinhao Mei**, Xubo Liu, Mark D Plumbley, and Wenwu Wang. “Automated Audio Captioning: An Overview of Recent Progress and New Challenges,” *Journal on Audio, Speech, and Music Processing*, 2022,

**Xinhao Mei**, Varun Nagaraja, Gael Le Lan, Zhaoheng Ni, Ernie Chang, Yangyang Shi, and Vikas Chandra. “FoleyGen: Visually Guided Audio Generation,” *arXiv preprint arXiv:2309.10537*, 2023,

Haohe Liu, Zehua Chen, Yi Yuan, **Xinhao Mei**, Xubo Liu, Danilo Mandic, Wenwu Wang, and Mark D Plumbley. “AudioLDM: Text-to-Audio Generation with Latent Diffusion Models,” *International Conference on Machine Learning (ICML)*, PMLR, 2023.

Xubo Liu\*, Qiushi Huang\*, **Xinhao Mei\***, Haohe Liu, Qiuqiang Kong, Jianyuan Sun, Shengchen Li, Tom Ko, Yu Zhang, Lilian H. Tang, Mark D. Plumbley, Volkan Kılıç, and Wenwu Wang. “Visually-Aware Audio Captioning with Adaptive Audio-Visual Attention,” *INTERSPEECH*, ISCA, 2023.

**Xinhao Mei**, Xubo Liu, Jianyuan Sun, Mark D Plumbley, and Wenwu Wang. “On Metric Learning for Audio-Text Cross-Modal Retrieval,” *INTERSPEECH*, ISCA, 2022.

**Xinhao Mei**, Xubo Liu, Jianyuan Sun, Mark D Plumbley, and Wenwu Wang. “Diverse Audio Captioning via Adversarial Training,” *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2022.

**Xinhao Mei**, Xubo Liu, Qiushi Huang, Mark D Plumbley, and Wenwu Wang. “Audio Captioning Transformer,” *Detection and Classification of Acoustic Scenes and Events Workshop (DCASE)*, 2021.

**Xinhao Mei**, Qiushi Huang, Xubo Liu, Gengyun Chen, Jingqian Wu, Yusong Wu, Jinzheng Zhao, Shengchen Li, Tom Ko, H Lilian Tang, Xi Shao, Mark D Plumbley, and Wenwu Wang. “An Encoder-Decoder based Audio Captioning System with Transfer and Reinforcement Learning,” *Detection and Classification of Acoustic Scenes and Events Workshop (DCASE)*, 2021.

## AWARDS & ACHIEVEMENTS

---

**2nd place in Task 6b (Language-Based Audio Retrieval)**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge 2022

**3rd place in Task 6a (Automated Audio Captioning)**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge 2022

**3rd place in Task 6 (Automated Audio Captioning)**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge 2021

**Turing Scheme Mobility Fund (£2000)**, Turing Scheme, UK Government

**University of Surrey-Santander PhD Travel Award (£1200)**, University of Surrey

**Surrey International Masters Scholarship (£5000)**, University of Surrey

## SERVICES

---

**Reviewer:** ICASSP, INTERSPEECH, MLSP, Information Fusion, TALLIP

## SKILLS

---

**Programming:** Python, PyTorch, Numpy

**Languages:** Chinese (Native), English (Fluent)