Peiwen (Sam) Sun

Beijing University of Posts and Telecommunications, Beijing, China

Email: sunpeiwen@bupt.edu.cn

Personal Website: https://peiwensun2000.github.io/

EDUCATION

- **Beijing University of Posts and Telecommunications**　　　　　　　　*Sept. 2022 - Present*
  MPhil Artificial Intelligence, and Certificate in Computer Science, Expected June 2025
  - **Overall GPA**: 3.88/4.00 (Top 15%)

- **Beijing University of Posts and Telecommunications**　　　　　　　　*Sept. 2018 - Jun. 2022*
  BSc with Honours, First Class Telecommunication Engineering
  - **Overall GPA**: 3.76/4.00, (Top 5%)

RESEARCH

- Papers of **First or equal contributor**: 1×**ACM MM**'2024 Oral (CCF-A/CORE-A*), 2×**ECCV**'2024 (CCF-B/CORE-A*), 1×**Interspeech**'2023 (CCF-C/CORE-A), 1×**IJCNN**'2024 (CCF-C/CORE-B), 2×**Arxiv**, 1×National **Patent**

- Papers of **collaborator**: 1×**ECCV**'2024 (CCF-B/CORE-A*), 1×**ACM MM**'2024 (CCF-A/CORE-A*), 1×**MICAD**'2024 (Springer),

- Ongoing Papers: ICLR'2025 (-/CORE-A*)

- Reviewer: ICLR('2024,'2025), CVPR'2024, ECCV'2024

Interest:

- Audio-visual scenery understanding, MLLM hallucination.

- Matric learning including speaker verification, face recognition, and multimodal person identification.

- Segmentation with modality guidance, open-world, medical diagnosis.

PAPERS

Unveiling and Mitigating Bias in Audio Visual Segmentation
Advisor: Di Hu, Renmin University of China.　　　　　　　　*Sept. 2023 - March 2024*

- In this work, it is recognized that previous AVS methods is heavily influenced inherent preferences and distributions. We propose a series of methods to unveil and mitigating bias caused by real-world inherent preferences and distributions in Audio Visual Segmentation.

- As a **mutual contributor**, preprint (here) is published in ACM MM'2024 Oral.

Ref-AVS: Refer and Segment Objects in Audio-Visual Scenes
Advisor: Di Hu, Renmin University of China.　　　　　　　　*Sept. 2023 - March 2024*

- In this work, it is recognized that previous AVS methods show a heavy reliance on detrimental segmentation preferences related to audible objects, rather than precise audio guidance. We argue that the primary reason is that audio lacks robust semantics compared to vision, especially in multi-source sounding scenes, resulting in weak audio guidance over the visual space.

- As a **mutual contributor**, preprint (here) is published in ECCV'2024.

Can Textual Semantics Mitigate Sounding Object Segmentation Preference?
Advisor: Di Hu, Renmin University of China.　　　　　　　　*Sept. 2023 - March 2024*

- We introduce a novel task called Reference Audio-Visual Segmentation (Ref-AVS), which seeks to segment objects within the visual domain based on expressions containing multimodal cues.

- As a **mutual contributor**, preprint (here) is published in ECCV'2024.

A Method of Audio-Visual Person Verification by Mining Connections between Time Series
Advisor: Shanshan Zhang, Tencent Inc.   *Mar. 2022 - Aug. 2022*

- A novel audio-visual strategy in-person verification that considers connections between time series from a generative perspective. Generative methods including *GAN*, *Contrastive Learning* and *Diffusion* are used here in an attempt to see the relationship between keyframes of different modalities on a time series as a style transformation.

- As the **main contributor**, the paper (here) is published in Interspeech'2023.

Other Papers

- As a **mutual contributor**, a paper about few-shot learning accepted by IJCNN'2024 (Best Paper Candidate).

- As the **second contributor**, a paper about medical diagnosis using segmentation is published in MICAD'2023 (here).

WORKING EXPERIENCES

Research Intern in Tencent Research Institute   *Feb. 2022 - July. 2022*

- Audio-visual person verification in the wild. The paper is mentioned above.

- Supervised and unsupervised speaker verification.

Engineering Intern in Megvii Research Institute   *Sept. 2021 - Jan. 2022*

- Face recognition on complicated ethnicity and facial expression situations.

- Engineering projects, including quantization and chip-level applications.

HONOURS

Outstanding Graduates Prize and **Undergraduate College Prize**

- Outstanding Graduates Prize of 2022 by Beijing University of Posts and Telecommunications.

Dean's List

- Dean's list for Fall 2019, Fall 2020, and Fall 2021 during BSc. Dean's list for Fall 2022 and Fall 2023 during MPhil.

Outstanding Student Cadres Price

- Granted in Fall 2019, Fall 2020 and Fall 2021.

SKILLS

Python (Proficient)   LaTeX(Proficient)
C (Proficient)   Java (Intermediate)

LANGUAGES

English: TOEFL 102/120 (expired)

GRADUATE LEVEL COURSES

Matrix Theory and Methods	92/100	Machine Learning	97/100
Information Search and Artificial Intelligence	93/100	Computer Vision	93/100
Image and Video Processing	93/100	Advanced Mathmatics I & II	90/100
Engineering Mathmatics	96/100	Linear Algebra	97/100
Data Structure	94/100	Java Programming	92/100