Size Wu

S-Lab for Advanced Intelligence
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798

Cellphone: +65 8048 1678
Email: size001@e.ntu.edu.sg
GitHub: https://github.com/wusize

EDUCATION

Nanyang Technological University (NTU) *Singapore, Singapore*
Ph.D. student in School of Computer Science and Engineering Jan 2022—Present
University of Science and Technology of China (USTC) *Hefei, China*
Bachelor of Electronic and Information Engineering (AI) Aug 2017—Jun 2021
GPA: 3.8/4.3 (Ranking Top 10%)

RESEARCH INTERESTS

Open-vocabulary Visual Recognition; Object Detection; Vision-language Models; Multi-view Human Pose Estimation

PAPERS & PATENTS

[1] **Size Wu**, Wenwei Zhang, Sheng Jin, Wentao Liu, and Chen Change Loy. Aligning bag of regions for open-vocabulary object detection. In *CVPR*, 2023.

[2] **Size Wu**, Sheng Jin, Wentao Liu, Lei Bai, Chen Qian, Dong Liu, and Wanli Ouyang. Graph-based 3d multi-person pose estimation using multi-view images. In *ICCV*, 2021.

[3] Jiahang Wang, **Size Wu**, Sheng Jin, Wentao Liu, and Chen Qian. Object detection method, device, electronic equipment and storage medium, 2022. Chinese Invention Patent, CN115131705A.

[4] **Size Wu**, Sheng Jin, Wentao Liu, and Chen Qian. 3d keypoint detection method, device, electronic equipment and storage medium, 2021. Chinese Invention Patent, CN113610967A.

[5] **Size Wu**, Sheng Jin, Wentao Liu, and Chen Qian. 3d human pose adjustment method, device, electronic equipment and storage medium, 2021. Chinese Invention Patent, CN113610966A.

RESEARCH EXPERIENCE

Ph.D. Student **NTU**, Jan 2022—Present

- **Augmentation for Open-Vocabulary Object Detection**: Propose an augmentation on image-text pairs for open-vocabulary object detection (OVD), which improves several main-stream OVD methods on multiple benchmarks. **Submitted to ICCV 2023.**

- **Aligning Bag of Regions for Open-vocabulary Object detection**: Propose to distill knowledge from pretrained vision-language models (VLMs) on a bag of regions for open-vocabulary object detection (OVD). The method effectively exploits the VLMs' ability to represent co-existing and contextually related concepts. It achieves state-of-the-art performance on multiple OVD benchmarks. **CVPR 2023. Code**.

Research Intern **SenseTime, Beijing**, Oct 2020—Dec 2021

- **Multi-view 3D Human Pose Estimation**: Propose three task-specific graph neural networks (GNNs) for multiview Human Pose Estimation. The GNNs efficiently match human centers, locate human locations and refine human pose estimations. The proposed method achieves state-of-the-art performance on CMU Panoptic and Shelf datasets with significantly lower computation complexity. **ICCV 2021. Code**.

- **Multi-view Multi-person Tracking**: Give solution to the Multi-camera Multiple People Tracking (MMP-Tracking) Challenge by aggregating geometric and appearance features. The method is applied in industrial projects. Achieve the 3rd place in the top-down view track without expensive training in a challenge in the ICCV 2021 Multi-camera Multiple People Tracking Workshop.

National Innovation Training Program **USTC**, April 2019—Dec 2020

- **Image Colorization**: Use deep learning approach to automatically colorize gray-scale images. The techniques exploited in this project mainly include feature extraction/aggregation, self- attention mechanism, adversarial training, etc. **Code**.

OPEN-SOURCE PROJECT

MMPose Oct 2020—Present
I am one of the developers of MMPose, an open-source project in GitHub for 2d and 3d human pose estimation. And I am the core developer of the multi-view 3d pose estimation in MMPose.

CORE QUALIFICATIONS

Programming: Python, C/C++, MATLAB
English: TOEFL 106/120,   **GRE 337/340+4.0/5.0**

Awards and Honors

- Scholarship For Outstanding students, USTC. 2018—2020.
- National Encouragement Scholarship, USTC. 2018, 2019.
- Talent Program in Information Science and Technology. USTC. 2017—2021.
- Security AI Challenger: Ninth Place Award (out of 36489 teams), Alibaba Cloud Tianchi. 2021
- AI Innovation and Application Competition: Second Place Award, CAICT. 2021
- Multi-camera Multiple People Tracking Challenge: Third place in Top-down View Track, ICCV2021.