WENHAO WU

Address: Los Angeles, CA 90015

E-mail: whwu.ucas@gmail.com   **Phone**: (+1) 6506667152

HomePage: https://whwu95.github.io

WORKING EXPERIENCE

Artificial General Intelligence (AGI), Amazon _Seattle, United States_
Applied Scientist. Manager: _Davide Modolo_ _(Incoming) July 2025 -_

Department of Computer Vision Technology (VIS), Baidu _Shenzhen, China_
Senior Researcher. Manager: _Jingdong Wang (IEEE Fellow)_ _July 2020 - July 2023_
Worked on (1) Multimodal LLMs, (2) video-text learning, (3) strong video foundation model, (4) efficient video tuning/inference, and (5) self-supervised video representation learning.

EDUCATION & VISITING

The University of Sydney _Mar 2022 - June 2025_
Doctor of Philosophy in Computer Science _Sydney, Australia_
Supervisor: _Prof. Wanli Ouyang_

The Chinese University of Hong Kong _Oct 2023 - Jan 2024_
Honorary Research Assistant in Multimedia Laboratory (MMLab) _Shatin, Hong Kong_
Supervisor: _Prof. Wanli Ouyang_

University of Chinese Academy of Sciences _Sep 2017 - July 2020_
Master of Science in Engineering in Pattern Recognition & Intelligent System _Beijing, China_
Supervisor: _Prof. Shifeng Chen and Prof. Yu Qiao_

Chinese Academy of Sciences _Oct 2016 - Aug 2017_
Research Assistant in Shenzhen Institute of Advanced Technology (SIAT) _Shenzhen, China_
Supervisor: _Prof. Shifeng Chen and Prof. Yu Qiao_

Central South University _Sep 2013 - June 2017_
Bachelor of Engineering in Automation _Changsha, China_

INDUSTRIAL INTERNSHIP EXPERIENCE

Creative Vision Group, Snap Research _Los Angeles, United States_
Applied Scientist Intern. Mentor: _Sergey Tulyakov, Yanyu Li, Anil Kag_ _Mar 2025 - July 2025_

Amazon Web Services AI _Santa Clara, United States_
Applied Scientist Intern. Mentor: _Shuai Zhang_ _June 2024 - Sep 2024_

Department of Computer Vision Technology (VIS), Baidu _Beijing, China_
Research Intern. Mentor: _Jingdong Wang (IEEE Fellow)_ _Oct 2023 - Mar 2025_

OpenMMLab, SenseTime Research _Shenzhen, China_
Research Intern. Mentor: _Kai Chen_ _Jan 2020 - Feb 2020_

Institute of Deep Learning (IDL), Baidu Research _Shenzhen, China_
Research Intern. Mentor: _Xiao Tan, Dongliang He, Shilei Wen_ _Sep 2018 - Mar 2020_

Video Analysis Group, iQIYI _Beijing, China_
R&D Intern. Mentor: _Yang Yang_ _June 2018 - Sep 2018_

Samsung Advanced Institute of Technology (SAIT), Samsung Research _Beijing, China_
Research Intern. Mentor: _Zhenbo Luo_ _Mar 2018 - June 2018_

PUBLICATIONS

1700+ citations on **Google Scholar** with H-index of 24 (2025.03). Code available on **GitHub**.

(Co-first Author → *, Corresponding Author → ✉, Mentored Interns/Mentees → #)

Journal

1) **Cap4Video++: Enhancing Video Understanding with Auxiliary Captions**
<u>Wenhao Wu</u>, Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, Wanli Ouyang
Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**), 2024

2) **Transferring Vision-Language Models for Visual Recognition: A Classifier Perspective**
<u>Wenhao Wu</u>, Zhun Sun, Yuxin Song, Jingdong Wang, Wanli Ouyang
International Journal of Computer Vision (**IJCV**), 2023

3) **GPT4Ego: Unleashing the Potential of Pre-trained Models for Zero-Shot Egocentric Action Recognition**
#Guangzhao Dai, Xiangbo Shu, <u>Wenhao Wu</u>, Rui Yan, Jiachao Zhang
Transactions on Multimedia (**T-MM**), 2024

4) **Rethinking 3D Cost Aggregation in Stereo Matching**
#Wanshui Gan, <u>Wenhao Wu</u>, Shifeng Chen, Yuxiang Zhao, Pak Kin Wong
Pattern Recognition Letters (**PRL**), 2023

Conference

1) **DistinctAD: Distinctive Audio Description Generation in Contexts**
#Bo Fang, <u>Wenhao Wu</u>, Qiangqiang Wu, Yuxin Song, Antoni B. Chan
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2025

2) **Dense Connector for MLLMs**
#Huanjin Yao*, <u>Wenhao Wu</u>*(✉), Taojiannan Yang, Yuxin Song, Mengxi Zhang, Haocheng Feng, Yifan Sun, Zhiheng Li, Wanli Ouyang, Jingdong Wang
Conference on Neural Information Processing Systems (**NeurIPS**), 2024

3) **Automated Multi-level Preference for MLLMs**
#Mengxi Zhang, <u>Wenhao Wu</u>, Yu Lu, Yuxin Song, Kang Rong, Huanjin Yao, Jianbo Zhang, Fanglong Liu, Yifan Sun, Haocheng Feng, Jingdong Wang
Conference on Neural Information Processing Systems (**NeurIPS**), 2024

4) **DetToolChain: A New Prompting Paradigm to Unleash Detection Ability of MLLM**
Yixuan Wu, Yizhou Wang, Shixiang Tang, <u>Wenhao Wu</u>, Tong He, Wanli Ouyang, Jian Wu, Philip Torr
European Conference on Computer Vision (**ECCV**), 2024

5) **What Can Simple Arithmetic Operations Do for Temporal Modeling?**
<u>Wenhao Wu</u>, Yuxin Song, Zhun Sun, Jingdong Wang, Chang Xu, Wanli Ouyang
IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023

6) **Cap4Video: What Can Auxiliary Captions Do for Text-Video Retrieval?**
<u>Wenhao Wu</u>, Haipeng Luo, Bo Fang, Jingdong Wang, Wanli Ouyang
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2023 **(Highlight, 2.5%)**

7) **Bidirectional Cross-Modal Knowledge Exploration for Video Recognition with Pre-trained Vision-Language Models**
<u>Wenhao Wu</u>, Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, Wanli Ouyang
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2023

8) **Revisiting Classifier: Transferring Vision-Language Models for Video Recognition**
<u>Wenhao Wu</u>, Zhun Sun, Wanli Ouyang
AAAI Conference on Artificial Intelligence (**AAAI**), 2023

9) **DSANet: Dynamic Segment Aggregation Network for Video-Level Representation Learning**
<u>Wenhao Wu</u>, Yuxiang Zhao, Yanwu Xu, Xiao Tan, Dongliang He, Zhikang Zou, Jin Ye, Yingying Li, Mingde Yao, Zichao Dong, Yifeng Shi
ACM International Conference on Multimedia (**ACMMM**), 2021

10) **MVFNet: Multi-View Fusion Network for Efficient Video Recognition**
Wenhao Wu, Dongliang He, Tianwei Lin, Fu Li, Chuang Gan, Errui Ding
AAAI Conference on Artificial Intelligence (**AAAI**), 2021

11) **Dynamic Inference: A New Approach Toward Efficient Video Action Recognition**
Wenhao Wu, Dongliang He, Xiao Tan, Shifeng Chen, Yi Yang, Shilei Wen
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**) Workshop, 2020. **Oral Presentation**

12) **Multi-Agent Reinforcement Learning Based Frame Sampling for Effective Untrimmed Video Recognition**
Wenhao Wu, Dongliang He, Xiao Tan, Shifeng Chen, Shilei Wen
IEEE/CVF International Conference on Computer Vision (**ICCV**), 2019 **(Oral Presentation, 4.3%)**

13) **UATVR: Uncertainty-Adaptive Text-Video Retrieval**
#Bo Fang*, **Wenhao Wu***, Chang Liu, Yu Zhou, Yuxin Song, Weiping Wang, Xiangbo Shu, Xiangyang Ji, Jingdong Wang
IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023

14) **MaMiCo: Macro-to-Micro Semantic Correspondence for Self-supervised Video Representation Learning**
#Bo Fang*, **Wenhao Wu***, Chang Liu, Yu Zhou, Dongliang He, Weiping Wang
ACM International Conference on Multimedia (**ACMMM**), 2022 **(Oral Presentation, 5.0%)**

15) **NSNet: Non-saliency Suppression Sampler for Efficient Video Recognition**
#Boyang Xia*, **Wenhao Wu***(✉), Haoran Wang, Rui Su, Dongliang He, Haosen Yang, Xiaoran Fan, Wanli Ouyang
European Conference on Computer Vision (**ECCV**), 2022

16) **Temporal Saliency Query Network for Efficient Video Recognition**
#Boyang Xia*, #Zhihao Wang*, **Wenhao Wu**(✉), Haoran Wang, Jungong Han
European Conference on Computer Vision (**ECCV**), 2022

17) **Temporal Action Proposal Generation with Background Constraint**
#Haosen Yang*, **Wenhao Wu***, Lining Wang, Sheng Jin, Boyang Xia, Hongxun Yao, Hujie Huang
AAAI Conference on Artificial Intelligence (**AAAI**), 2022

18) **ASCNet: Self-supervised Video Representation Learning with Appearance-Speed Consistency**
#Deng Huang*, **Wenhao Wu***, Weiwen Hu, Xu Liu, Dongliang He, Zhihua Wu, Xiangmiao Wu, Mingkui Tan, Errui Ding
IEEE/CVF International Conference on Computer Vision (**ICCV**), 2021

19) **Good Practices and A Strong Baseline for Traffic Anomaly Detection**
#Yuxiang Zhao*, **Wenhao Wu***, Dongliang He, Xiao Tan, Shifeng Chen, Yi Yang, Shilei Wen
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**) Workshop, 2021. **(Winner Award on CVPR 2021 AI City Challenge)**

20) **CODER: Coupled Diversity-Sensitive Momentum Contrastive Learning for Image-Text Retrieval**
Haoran Wang, Dongliang He, **Wenhao Wu**, Boyang Xia, Min Yang, Fu Li, Yunlong Yu, Zhong Ji, Errui Ding, Jingdong Wang
European Conference on Computer Vision (**ECCV**), 2022

21) **Maximum Spatial Perturbation Consistency for Unpaired Image-to-Image Translation**
Yanwu Xu, Shaoan Xie, **Wenhao Wu**, Kun Zhang, Mingming Gong, Kayhan Batmanghelich
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022

22) **Effective Invertible Arbitrary Image Rescaling**
Zhihong Pan, Baopu Li, Dongliang He, **Wenhao Wu**, Errui Ding
IEEE/CVF Winter Conference on Applications of Computer Vision (**WACV**), 2023

23) **AdaCM: Adaptive ColorMLP for Real-Time Universal Photo-realistic Style Transfer**
Tianwei Lin, Honglin Lin, Fu Li, Dongliang He, **Wenhao Wu**, Meiling Wang, Xin Li, Yong Liu

AAAI Conference on Artificial Intelligence (**AAAI**), 2023

24) **Towards Bidirectional Arbitrary Image Rescaling: Joint Optimization and Cycle Idempotence**
Zhihong Pan, Baopu Li, Dongliang He, Mingde Yao, <u>**Wenhao Wu**</u>, Tianwei Lin, Xin Li, Errui Ding
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022

25) **Coarse to Fine: Domain Adaptive Crowd Counting via Adversarial Scoring Network**
Zhikang Zou, Xiaoye Qu, Pan Zhou, Shuangjie Xu, Xiaoqing Ye, <u>**Wenhao Wu**</u>, Jin Ye
ACM International Conference on Multimedia (**ACMMM**), 2021

26) **Weakly-Supervised Spatio-Temporal Anomaly Detection in Surveillance Video**
Jie Wu, Wei Zhang, Guanbin Li, <u>**Wenhao Wu**</u>, Xiao Tan, Yingying Li, Errui Ding, Liang Lin
International Joint Conference on Artificial Intelligence (**IJCAI**), 2021

27) **Attention-Driven Dynamic Graph Convolutional Network for Multi-Label Image Recognition**
Jin Ye, Junjun He, Xiaojiang Peng, <u>**Wenhao Wu**</u>, Yu Qiao
European Conference on Computer Vision (**ECCV**), 2020

INVENTION PATENTS

- Video recognition method and apparatus, electronic device and storage medium. EP21202755.1
- Method, apparatus, electronic device, readable storage medium and program for classifying video. EP21177154.8
- Method, apparatus, device and storage medium for training video recognition model. US17/983,208
- Method and apparatus for detecting traffic anomaly. US17/963,058
- Video recognition method and apparatus, electronic device and storage medium. US17/411,254
- Method, apparatus, electronic device and readable storage medium for classifying video. US17/338,263
- Robust and efficient blind super-resolution using variational kernel autoencoder. PCT/CN2021/122050
- Method, apparatus, device and storage medium for training video recognition model. PCT/CN2022/075153
- Method and apparatus for detecting traffic anomaly. PCT/CN2022/075071

CONTESTS

· CVPR 2021 AI City Challenge: Traffic Anomaly Detection, ***Winner*** Award	*2021*
· CVPR 2021 NTIRE Challenge: Image Deblurring, ***Runner-Up*** Award	*2021*
· Meritorious Winner in the America Mathematical Contest In Modeling (**MCM**)	*2016*
· 1st **Prize** in National Undergraduate Mechanical Innovation Design Competition	*2016*
· 2nd Prize in National Freescale Cup Intelligent Car Competition (South China Region)	*2015*
· 2nd Prize in Undergraduate Smart Car Racing Competition of Hunan Province	*2015*

SELECTED HONORS AND AWARDS

Academic (Undergraduate/Master's/PhD) Honors

- OpenAI Researcher Access Grant ($5000)   *2024*

- **Chinese Government Award for Outstanding Self-financed Students Abroad** (The highest award granted by the Chinese government to Chinese students overseas)   *2024*

- **11th Baidu Fellowship** (only 10 postgraduate student worldwide, 200,000 RMB)   *2023*

- Faculty of Engineering Research Scholarship at USYD (**Full PhD Scholarship**)   *2022-2026*

- Outstanding Student of University of Chinese Academy of Sciences   *2020*

- UCAS Postgraduate Scholarship (**Full Master Scholarship**)   *2017-2020*

- **Graduate with Honour** of Central South University (Top 5%)   *2017*

- National Encouragement Scholarship (Top 5%) *2016*
- Scholarship for Academic Excellence of Central South University *2014,2015,2016*

Professional Honors

- Australian Good Design Award *2022*
- Baidu AIG-TPG Technical Innovation Team Award *2021 H1*
- Baidu TPG-VIS Spotlight Project Award *2021 Q1*
- Baidu AIG-TPG Best Innovation Team Award *2020*
- Baidu Best Newcomer Award *2020*
- Baidu Best Intern Award *2019*

ACADEMIC SERVICES

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)
- IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**)
- IEEE Transaction on Image Processing (**TIP**)
- IEEE Transaction on Circuits and Systems for Video Technology (**TCSVT**)
- IEEE Transactions on Multimedia (**TMM**)
- Computer Vision and Image Understanding (**CVIU**)
- IEEE Transactions on Biomedical Engineering (**TBME**)
- Knowledge-Based Systems (**KBS**)
- International Journal of Multimedia Information Retrieval (**IJMIR**)
- IEEE Intelligent Transportation Systems Transactions (**TITS**)
- Internet of Things (**IOT**)
- Transactions on Multimedia Computing Communications and Applications (**TOMM**)

Program Committee Member/Conference Reviewer

- International Conference on Machine Learning (**ICML**), 2025
- Conference on Neural Information Processing Systems (**NeurIPS**), 2024
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2021, 2022, 2023, 2024, 2025
- IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023, 2025
- European Conference on Computer Vision (**ECCV**), 2022, 2024
- International Joint Conference on Artificial Intelligence (**IJCAI**), 2021
- AAAI Conference on Artificial Intelligence (**AAAI**), 2021, 2022, 2023
- ACM International Conference on Multimedia (**ACMMM**), 2023, 2024
- Winter Conference on Applications of Computer Vision (**WACV**), 2022
- International Conference on Digital Image Computing: Techniques and Applications (**DICTA**), 2022

Member of IEEE, ACM, AAAI and CVF

Off-Campus Mentor of Tsinghua University, 2023-2025

MENTORING EXPERIENCE

($\rightarrow$: Post-mentorship status)

Mentored Interns at Baidu

- Bo Fang (Master, CAS $\rightarrow$ PhD, CityU)   *2021-2023, 2024-Now*
  MaMiCo (ACMMM'23 Oral), UATVR (ICCV'23), DistinctAD (CVPR'25)

- Huanjin Yao (Master, Tsinghua University $\rightarrow$ PhD, NTU)   *2023-Now*
  Side4Video, Dense Connector (NeurIPS'24), Mulberry

- Mengxi Zhang (Master, Tianjin University $\rightarrow$ Bytedance AI Lab)   *2024*
  AMP (NeurIPS'24)

- Boyang Xia (Master, CAS $\rightarrow$ Kuaishou)   *2021-2022*
  NSNet (ECCV'22), TSQNet (ECCV'22)

- Haosen Yang (Master, HIT $\rightarrow$ PhD, University of Surrey)   *2021*
  BCNet (AAAI'22)

- Deng Huang (Master, SCUT $\rightarrow$ AutoX)   *2021*
  ASCNet (ICCV'21)

- Yuguo Wang (Bachelor, UESTC $\rightarrow$ Master, Duke University)   *2022*

- Hengyuan Zhao (Bachelor, NUPT $\rightarrow$ PhD, NUS)   *2021*

- Yuxiang Zhao (Master, CAS $\rightarrow$ PhD, Peking University)   *2021*
  DSANet (ACMMM'21), CVPR'21 AICity Winner

University Mentees

- Guangzhao Dai (PhD, NJUST)   *2023*
  GPT4Ego (TMM'24)

- Haipeng Luo (Master, CAS $\rightarrow$ PhD, Tsinghua University)   *2023*
  Cap4Video (CVPR'23)

- Zhihao Wang (Master, CAS $\rightarrow$ MiniMax)   *2022*
  TSQNet (ECCV'22)

CODE SKILLS

Python, Linux, Matlab, C/C++ and deep learning frameworks such as PyTorch, TensorFlow, Caffe.