Huabin Liu

🔗 Google Scholar   🔗 Personal Website
📱 18821708622   ✉ huabinliu@sjtu.edu.cn
ℹ 1997/5/29

Curriculum Vitae

Bio. I am a 5th-year PhD Student at Shanghai Jiao Tong University, working with Weiyao Lin on video understanding and analysis. From 12/2024 to 12/2025, I was a visiting PhD student at the University of Amsterdam, advised by Cees Snoek and *Filip Ilievski*.

Research interests. My previous research work covers : **fine-grained image recognition, video understanding, few-shot learning**. Currently, I mainly focus on **multi-modality video understanding and reasoning.**

🎓 Education

Dec 2023 – Dec 2024	**Visiting PhD student**, University of Amsterdam (UvA), Netherlands Video & Image Sense Lab (VIS Lab) Advisor : *Cees Snoek*, *Filip Ilievski*
Sep 2019 – Jun 2025 (expected)	**PhD Candidate** in Department of Electronic Engineering, *Shanghai Jiao Tong University* (SJTU) Institute of Media, Information, and Network (MIN Lab) Advisor : *Weiyao Lin*
Sep 2015 –Jun 2019	**Bachelor of Engineering**, *Xidian University* (XDU), Xi'an Majors in Electromagnetic Fields Micro Wave Technology GPA : 3.72/4.0 Rank : Top 5%

📖 Publications

> **Huabin Liu**, Filip Ilievski, Cees G. M. Snoek,"Commonsense Video Question Answering through Video-Grounded Entailment Tree Reasoning", under review, 2025,
> `video question answering` `explicit reasoning` `explainable AI`

> **Huabin Liu**, Xiao Ma, Cheng Zhong, Yang Zhang, Weiyao Lin, "TimeCraft : Navigate Weakly-Supervised Temporal Grounded Video Question Answering via Bi-directional Reasoning", accepted by **ECCV**, 2024
> `video question answering` `video reasoning` `temporal grounding`

> **Huabin Liu**, Weixian Lv, John See, Weiyao Lin, "Task-adaptive Spatial-Temporal Video Sampler for Few-shot Action Recognition", published in **ACM Multimedia**, 2022.
> `temporal frame selection` `spatial region zoom` `dynamic sampling strategy` `few-shot action recognition`

> **Huabin Liu**, Weiyao Lin, Tieyuan Chen, Yuxi Li, Shuyuan Li, John See, "Few-shot Action Recognition via Intra-and Inter-Video Information Maximization," submitted to International Journal of Computer Vision (**IJCV**), under the 2nd round review
> `video information maximization` `few-shot action recognition`

> **Huabin Liu**, Jianguo Li, Dian Li, John See, Weiyao Lin, "Learning Scale-Consistent Attention Part Network for Fine-grained Image Recognition," published in *IEEE Transactions on Multimedia* (**TMM**), 2021.
> `weakly-supervised` `fine-grained recognition`

> Tieyuan Chen*, **Huabin Liu***(co-first), Cheng Zhong, Yang Zhang, Weiyao Lin, MECD : Unlocking Multi-event Causal Discovery in Video Reasoning, 2024, accepted by **NeurIPS**, 2024
> `video causal discovery` `long-term video understanding`

> Hao Wu*, **Huabin Liu***(co-first), Yu Qiao, Xiao Sun, "DIBS : Enhancing Dense Video Captioning with Unlabeled Videos via Pseudo Boundary Enrichment and Online Refinement", published in **CVPR**, 2024
> `Dense video captioning` `large Language model`

> Shuyuan Li*, **Huabin Liu***(co-first), Rui Qian, Yuxi Li, John See, Mengjuan Fei, Xiaoyuan Yu, Weiyao Lin, "TA2N : Two-Stage Action Alignment Network for Few-shot Action Recognition" published in **AAAI**, 2022
> `spatial-temporal action alignment` `few-shot action recognition`

> Weiyao Lin, **Huabin Liu**, Shizhan Liu, Yuxi Li, Ning Xu, Hongkai Xiong, Guo-Jun Qi, Nicu Sebe "HiEve : A large-scale benchmark for human-centric video analysis in complex events" accepted by *International Journal of Computer Vision (***IJCV***)*, 2023
> `human-centric video dataset` `large-scale benchmark` `co-learning`

> Tianyao He, **Huabin Liu**, Zeling Ni, Yuxi Li, Xiao Ma, Cheng Zhong, Yang Zhang, Yingxue Wang, Weiyao Lin "Achieving Procedure-Aware Instructional Video Correlation Learning under Weak Supervision from a Collaborative Perspective" published in International Journal of Computer Vision (**IJCV**), 2024
> `video correlation learning` `procedure learning` `metric learning`

> Tianyao He, **Huabin Liu**, Yuxi Li, Xiao Ma, Cheng Zhong, Yang Zhang, Weiyao Lin "Collaborative Weakly Supervised Video Correlation Learning for Procedure-Aware Instructional Video Analysis" published in **AAAI**, 2024

instructional video | step learning | procedure verification

> Jiahao Fan, **Huabin Liu**, Wenjie Yang, John See, Weiyao Lin, "Speed Up Object Detection on Gigapixel-Level Images With Patch Arrangement" published in **CVPR**, 2022
  reinforcement learning | object detection | super high-resolution

> Rui Qian, Yuxi Li, **Huabin Liu**, John See, Shuangrui Ding, Xian Liu, Dian Li, Weiyao Lin, "Enhancing Self-supervised Video Representation Learning via Multi-level Feature Optimization" published in **ICCV**, 2021
  contrastive learning | video representation

</> Projects & Experiences

Dec 2023 – current | **Visiting PhD student, University of Amsterdam, Amsterdam, Netherlands**
> Exploring video-language models in video causal reasoning and question-answering
  video reasoning | video question answering | video-language model

Feb 2023 – current | **Computer vision research intern, Shanghai AI Laboratory, Shanghai**
> May-Nov 2023, devised a weakly-supervised framework for dense video captioning using large-scale unlabeled video data. The paper of this work has been accepted by CVPR 2024.
> Nov-now, exploring the LLM-driven long-term video understanding.
  video understanding | Large language model

Jan 2020 – Jun 2020 | **Organization of HiEve dataset and corresponding challenge 2020, SJTU, ACM MM 2020**
> Participant in the data collection, official website design, evaluation server, and organization of challenge. Get the Best Grand Challenge organization award in ACM MM.
> Responsible for the paper on the HiEve dataset; proposed two novel enhanced baselines for pose estimation and action recognition on HiEve. The paper has been published in IJCV.
> ⚙ : http://humaninevents.org/
  Human-centric large-scale video dataset | challenge | novel enhanced baselines | HiEve paper

Jan 2019 – Sep 2019 | **Computer vision research intern, Aibee Inc., Beijing**
> Jan-Apr 2019, working on an interpretable fine-grained image recognition method. Its performances achieve SOTA at that time. This work was also nominated as an outstanding dissertation for bachelor's degree at Xidian University.
> May-Sep 2019, devised an attention-based model for fine-grained recognition. The paper has been published in IEEE Transactions on Multimedia, 2021.
  fine-grained image | discriminative part localization | weakly supervised | multi-scale part selection | attention

🔨 Honors & Awards

2024	Founded by China Scholarship Council (CSC) to visit the University of Amsterdam as a researcher for 1 year
Fall 2022	CETC scholarship for Ph.D. students (¥20000).
Fall 2020	Selected as a member of Wu Wenjun AI honorary doctor class in Shanghai Jiao Tong University (<1%).
Spring 2019	Outstanding undergraduate student of Xidian University (XDU).
2015 - 2018	First Prize Scholarship of Xidian University (<4%), 3 times
2015 - 2018	Outstanding Student of Xidian University (<3%), 2 times
Fall 2017	First prize of National college students Mathematical Contest in Shaanxi (<3%).

💡 Self-evaluation

> Work seriously and tidily, with independent thinking and analytical ability.
> A quick learner in new research fields.
> Good communication and cooperation ability.