

Hongwei Xue

Email: xuehongwei@mail.ustc.edu.cn Tel: (86) 182-2662-5994 Page: hellwayxue.github.io
Address: University of Science and Technology of China, No.96, Jinzhai Road Baohe District, Hefei, Anhui

Education

PhD in School of Information Science and Technology, USTC *Aug 2021- Jul 2024*

Advisers: Jiebo Luo, Houqiang Li.

Master in School of Information Science and Technology, USTC *Sep 2019- Jul 2021*

Advisers: Jiebo Luo, Houqiang Li.

National Scholarship in 2021.

Bachelor of Science in Automation, University of Science and Technology of China *Aug 2015- Jul 2019*

GPA: 3.69/4.3, in the School of Gifted Young.

Scholarship for Outstanding Students in 2016, 2017, 2018. Freshmen Scholarship 2015.

Research & Competition Experiences

Academic Community Service

Reviewer for ACM MM 2021, ICMR 2021, ICME 2022, TMM 2022, CVPR 2023, ICCV 2023

Internship at Microsoft Azure Cognitive Service *Feb 2023-May 2023*

- Study on AI Generated Content, mainly focus on image-to-text generation.

Internship at Shanghai AI Lab *Sep 2022-Feb 2023*

- Studied on Self-Supervised Learning like Masked Image Modeling (MIM) and fundamental computer vision. One paper was accepted by CVPR.

Internship at MSRA *July 2020-Nov 2021*

- The Award of Excellence in Stars of Tomorrow Internship Program.
- Worked on Vision-Language project, including Image/Video-Language Pre-training model and dataset. Proposed HD-VILA, a large-scale and diversified video-text dataset. Proposed CLIP-ViP, a powerful video-language model based on CLIP. Several papers were accepted by ACM MM, NeurIPS, CVPR, ICLR.
- Worked on the Image-to-Video Animation project, which aims to generate a time-lapse video from a single landscape image. One paper was accepted by ACM MM (Oral).
- Contributed to a template recoloring algorithm applied for Office PowerPoint. This algorithm can recolor the template image based on given palettes, which can cut the cost of template design.

ACM Multimedia 2021 Pre-training for Video Understanding Challenge *May 2021-July 2021*

- Proposed Semantic Tag boosted XlanV (ST-XlanV) model to introduce semantic tags into video captioning.
- Reached the 5th place (no extra data) in the challenge and the paper was accepted.

Internship at HUAWEI Cloud *Feb 2019-July 2019*

- Participated in the Video Fingerprint Project. Independently designed and implemented a model for image manipulation classification, which aims to classify the type of editing in video such as horizontal flip, mosaic.
- Integrated the model to a piracy detection system, which is already for commercial use. I also contributed to the implementation of the system, including a video processing pipeline in C++ and some algorithms in Python.

Compete in the RoboGame held by USTC *June 2017-Oct 2017*

- Designed and implemented a robot can draw by a pulley actuated from two poles. Reached the final contest and won the 3rd place. Learned a lot about hardware and actuators.

Leadership & Activities

Poster Sharing in Microsoft Ada Camp

Aug 2021

Shared my Image-to-Video Animation work and MSRA culture in Poster sharing session of Ada Camp 2021 in Microsoft Research of Asia.

Teaching Assistant of Sensor principle and application

Sep 2018-Jan 2019

Basketball Team Leader of the School of Gifted Young

Mar 2016-June 2016

Group Leader of Society Practice

Feb 2016-Mar 2016

Led the team to investigate the red culture in Northeast region and was rated as Excellent.

Group Leader of Freshmen Seminar

Sep 2015-Mar 2016

Led the team to study in the best type of anti-haze masks and gained an 'A' grade.

Skills

Learning distributed deep learning training framework such as DeepSpeed, Horovod.

Frequent user of Linux, Python (PyTorch, Numpy). Fundamental grasp of C/C++, MATLAB.

Publications

- **Hongwei Xue**, Peng Gao, Hongyang Li, et al. Stare at What You See: Masked Image Modeling without Reconstruction. CVPR 2023.
- **Hongwei Xue**, Yuchong Sun, Bei Liu, et al. CLIP-ViP: Adapting Pre-trained Image-Text Model to Video-Language Representation Alignment. ICLR 2023.
- Yuchong Sun, **Hongwei Xue**, Ruihua Song, et al. Long-Form Video-Language Pre-training with Multimodal Temporal Contrastive Learning. NeurIPS 2022.
- **Hongwei Xue**, Tiankai Hang, et al. Advancing High-Resolution Video-Language Representation with Large-Scale Video Transcriptions. CVPR 2022.
- **Hongwei Xue**, Yupan Huang, Bei Liu, et al. Probing Inter-modality: Visual Parsing with Self-Attention for Vision-and-Language Pre-training. NeurIPS 2021.
- Yiqing Huang*, **Hongwei Xue***, et al. Semantic Tag Augmented XlanV Model for Video Captioning. ACM MM 2021.
- Yupan Huang, **Hongwei Xue**, Bei Liu, et al. Unifying Multimodal Transformer for Bi-directional Image and Text Generation. ACM MM 2021.
- **Hongwei Xue**, Bei Liu, Huan Yang, et al. Learning Fine-Grained Motion Embedding for Landscape Animation. ACM MM 2021 (**Oral**).
- **Hongwei Xue**, Haomiao Liu, Jun Li, et al. Sed-Net: Detecting Multi-Type Edits of Images. ICME 2020.