

Hongwei Xue

Email: gh051120@mail.ustc.edu.cn Tel: (86) 182-2662-5994

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**, ranking: **71/375 in the School of Gifted Young**.

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

Research & Competition Experiences

Academic Community Services

Reviewer in ACM Multimedia 2021 Industrial Track, ICMR 2021, ICME 2022.

Internship in MSRA

July 2020-Nov 2021

- Participated in the Video-and-language Pretraining Project. Designed a pre-training framework to learn spatiotemporal information from high-resolution videos. The paper is submitted to CVPR 2022.
- Proposed a model first adopts self-attention to learn visual features for vision-language pre-training and designed a probing method to reveal the advantage of this model. Wrote a paper accepted by NeurIPS 2021.
- Collaborated in presenting a unified image-and-text generative framework based on a Transformer model. The paper was accepted by ACM MM 2021.
- Participated in the Image-to-Video Animation project, which aims to generate a time-lapse video from a single landscape image. Existing works suffer from inaccurate motion generation. Tackled this problem by proposing a fine-grained motion embedding method. Wrote a paper accepted by ACM MM 2021 as 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 in the challenge and the paper was accepted.

Internship in 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 Ada Camp

Aug 2021

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

Teaching Assistant, 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

- Xue H, Tiankai H, et al. Advancing High-Resolution Video-Language Representation with Large-Scale Video Transcriptions. CVPR 2022.
- Xue H, Huang Y, Liu B, et al. Probing Inter-modality: Visual Parsing with Self-Attention for Vision-and-Language Pre-training. Advances in Neural Information Processing Systems, 2021, 34.
- Huang Y*, Xue H*, Chen J, et al. Semantic Tag Augmented XlanV Model for Video Captioning. Proceedings of the 29th ACM International Conference on Multimedia. 2021: 4818-4822. **(Equal contribution)**
- Huang Y, Xue H, Liu B, et al. Unifying Multimodal Transformer for Bi-directional Image and Text Generation. Proceedings of the 29th ACM International Conference on Multimedia. 2021: 1138-1147.
- Xue H, Liu B, Yang H, et al. Learning Fine-Grained Motion Embedding for Landscape Animation. Proceedings of the 29th ACM International Conference on Multimedia. 2021: 291-299. **(Oral)**
- Xue H, Liu H, Li J, et al. Sed-Net: Detecting Multi-Type Edits of Images. 2020 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2020: 1-6.