

YUCHONG SUN

(+86) 157 3550 0163 ycsun@ruc.edu.cn ycsun1972.github.io

EDUCATION

Renmin University of China
Ph.D. Student in Artificial Intelligence
Adviser: Ruihua Song

Beijing, China
Sep. 2020 - Present

South China University of Technology
B.S. in Information Engineering
GPA: 3.82/4

Guangzhou, China
Sep. 2016 - Jun. 2020

RESEARCH INTEREST

Interested in multi-modal natural language understanding and expression, for the purpose of giving machines the ability to understand and express like humans. Recently I am focusing on vision and language pre-training and downstream applications, *e.g.*, *multi-modal dialog*, *image caption* and *video understanding*.

RESEARCH EXPERIENCE

Internship at **Microsoft Research Asia (MSRA)**

Aug. 2021 - Aug. 2022

- The Award of Excellence in Stars of Tomorrow Internship Program.
- Worked on Video-Language Project, including collecting HD-VILA-100M, a large-scale and diversified video-text dataset, and HD-VILA, LF-VILA, CLIP-ViP, several video-language pre-training models.
- Three papers were accepted by CVPR, NeurIPS, ICLR, respectively.

Internship at **Beijing Academy of Artificial Intelligence (BAAI)**

Dec. 2020 - Jul. 2021

- **WenLan** project of large-scale multi-modal pre-training.

PUBLICATIONS

* means equal contribution.

1. Hongwei Xue*, **Yuchong Sun***, Bei Liu, Jianlong Fu, Ruihua Song, Houqiang Li, Jiebo Luo. CLIP-ViP: Adapting Pre-trained Image-Text Model to Video-Language Representation Alignment. ICLR 2023.
2. Xu Gu*, **Yuchong Sun***, Feiyue Ni, Shizhe Chen, Ruihua Song, Boyuan Li, Xiang Cao. Translating Text Synopses to Video Storyboards. arXiv 2022.
3. **Yuchong Sun**, Hongwei Xue, Ruihua Song, Bei Liu, Huan Yang, Jianlong Fu. Long-Form Video-Language Pre-training with Multimodal Temporal Contrastive Learning. NeurIPS 2022.
4. Hongwei Xue*, Tiankai Hang*, Yanhong Zeng*, **Yuchong Sun***, Bei Liu, Huan Yang, Jianlong Fu, Baining Guo. Advancing High-Resolution Video-Language Representation with Large-Scale Video Transcriptions. CVPR 2022.
5. **Yuchong Sun** Xiwei Cheng, Ruihua Song, Wanxiang Che, Zhiwu Lu, Jirong Wen. Difference between Multi-modal vs. Text Pre-trained Models in Embedding Text. NLPCC 2022.

SKILLS

Programming Languages and Frameworks

Python, C++, MATLAB, PyTorch, TensorFlow

Distributed Deep Learning Training Framework

DeepSpeed, Horovod, Accelerate

AWARDS

The Award of Excellence in Stars of Tomorrow Internship Program, MSRA	<i>2022</i>
First-class Scholarship, Renmin University of China	<i>2022</i>
Merit Student, South China University of Technology	<i>2017</i>

ACADEMIC SERVICE

Journal and Conference Reviewer: TMM, ICCV 2023, CVPR 2023