

Yunlong TANG

Email: yunlong.tang@rochester.edu | Mobile: (+1) 585-616-0074 | Website: yunlong10.github.io

EDUCATION

University of Rochester

Aug. 2023 - Jun. 2028 (Expected)

Ph.D. Student in Computer Science, advised by [Prof. Chenliang Xu](#)

Rochester, NY, US

Southern University of Science and Technology (SUSTech)

Aug. 2019 - Jun. 2023

B.Eng. in Intelligence Science and Technology (GPA: 3.66/4.00), advised by [Prof. Feng Zheng](#)

Shenzhen, CN

RESEARCH INTERESTS

- *Multimodal Learning, Video Understanding, AI-Agents*

RESEARCH EXPERIENCES

University of Rochester

Aug. 2023 - Present

Graduate Research Assistant, supervised by [Prof. Chenliang Xu](#)

Rochester, NY, US

- Leading a survey on large language models for video understanding (Vid-LLMs). The continuously updated survey paper [3] covers core aspects of Vid-LLMs, including their evolution, modeling techniques, capabilities, and applications, alongside an evaluation of the tasks, datasets, and benchmarks essential for their advancement.
- Working on a project involving audio-visual video understanding. Proposed pseudo long-form audio-visual dataset and pseudo temporal boundary alignment to improve the fine-grained video understanding capacity of audio-visual LLMs.
- Collaborated on a project focused on unifying egocentric video understanding tasks, contributed to ideation, developing instruction-tuning datasets, conducting experiments, and manuscript preparation. The findings have been submitted for consideration at the CVPR 2024.

SUSTech VIP Lab

Aug. 2022 - Jul. 2023

Undergraduate Student Researcher, supervised by [Prof. Feng Zheng](#)

Shenzhen, CN

- Collaborated on Caption-Anything project, contributed to the segmentation module for supporting interactive visual prompts, and involved in the technical report [5] writing.
- Proposed LaunchpadGPT, which aims to utilize language model to generate music visualization in the form of Launchpad displaying video. Results [2] accepted to International Computer Music Conference (ICMC), 2023.
- Participated in the Generic Event Boundary Captioning competition at CVPR 2023 Long-form Video Understanding Workshop, proposed and developed the LLMVA-GEBC model [4] that won the championship.

Tencent

Sept. 2021 - Aug. 2022

ML Research Intern, supervised by Ms. Qin Lin and [Dr. Wenhao Jiang](#)

Shenzhen, CN

- Proposed and developed multi-modal segment assemblage network (M-SAN) and importance-coherence reward for training. The method improves efficiency and accuracy when compared to current automatic advertisement video editing techniques. Results [1] accepted to ACCV 2022.
- Deployed the model in Tencent servers online to perform efficient and accurate ad video editing, and filed the patent “An Approach for Automatic Ad Video Editing”.

PUBLICATIONS

(* equal contribution)

[1] **Yunlong Tang**, Siting Xu, Teng Wang, Qin Lin, Qinglin Lu, Feng Zheng, “Multi-modal Segment Assemblage Network for Ad Video Editing with Importance-Coherence Reward”, in *Proceedings of 16th Asian Conference on Computer Vision (ACCV)*, 2022.

[2] Siting Xu*, **Yunlong Tang***, Feng Zheng, “LaunchpadGPT: Language Model as Music Visualization Designer on Launchpad”, in *Proceedings of International Computer Music Conference (ICMC)*, 2023.

PREPRINTS

(* equal contribution)

[3] **Yunlong Tang***, Jing Bi*, Siting Xu*, Luchuan Song, Susan Liang, Teng Wang, Daoan Zhang, Jie An, Jingyang Lin, Rongyi Zhu, Ali Vosoughi, Chao Huang, Zeliang Zhang, Feng Zheng, Jianguo Zhang, Ping Luo, Jiebo Luo, Chenliang Xu, “Video Understanding with Large Language Models: A Survey”, in *arXiv:2312.17432*, 2023.

[4] **Yunlong Tang**, Jinrui Zhang, Xiangchen Wang, Teng Wang, Feng Zheng, “LLMVA-GEBC: Large Language Model with Video Adapter for Generic Event Boundary Captioning”, in *arXiv:2306.10354*, 2023.

[5] Teng Wang*, Jinrui Zhang*, Junjie Fei*, Hao Zheng, **Yunlong Tang**, Zhe Li, Mingqi Gao, Shanshan Zhao, “Caption Anything: Interactive Image Description with Multimodal Controls”, in *arXiv:2305.02677*, 2023.

TEACHING EXPERIENCE

SUSTech

Sept. 2022 - Jun. 2023

Teaching Assistant for SUSTech CS308 Computer Vision

Shenzhen, CN

HONORS & AWARDS

- The First Place in Generic Event Boundary Captioning Track of [LOVEU Challenge](#) at CVPR 2023.
- Excellent Graduate for Exceptional Performance, SUSTech, 2023.
- Excellent Undergraduate Thesis, the Department of Computer Science and Engineering, SUSTech, 2023.
- The First Class of Merit Student Scholarship for Exceptional Performance, SUSTech, 2021-2022.
- Research Innovation Award, Shude College, SUSTech, 2020-2021.

ACADEMIC SERVICE

- **Journal Reviewer:** IEEE Transactions on Multimedia (TMM)

SKILLS LIST

- Programming Languages: Python, C++, Java, JavaScript, MATLAB
- Libraries/Tools: PyTorch, HuggingFace, OpenCV, FFmpeg, LangChain
- Language: Chinese (native), English (fluent)