

# Yunbin Tu

Ph.D. Student, University of Chinese Academy of Sciences, Beijing, China

Email: tuyunbin22@mailsucas.ac.cn — Mobile: (+86) 15010751895 — Website: <https://tuyunbin.github.io/>

## RESEARCH INTERESTS

---

- **Vision-and-Language:** video captioning, change captioning, image captioning, multi-modal machine translation.
- **Multimedia content analysis:** multi-modal fusion, text-based information retrieval from video.

## EDUCATION

---

### University of Chinese Academy of Sciences

Ph.D. in Computer Applied Technology  
Advisors: Prof. Li Su and A.P. Liang Li (ICT, CAS)

Beijing, China  
Sep. 2022 — Present

### Kunming University of Science and Technology

M.Eng in Pattern Recognition and Intelligent Systems  
Advisors: Prof. Zhengtao Yu

Kunming, China  
Sep. 2019 — Jun 2022

### Hangzhou Dianzi University

B.Eng in Automation  
Advisors: Prof. Chenggang Yan

Hangzhou, China  
Sep. 2014 — Jun 2018

## INTERN EXPERIENCE

---

### VIPL Research Group, ICT, CAS

*Research Intern with A.P. Liang Li*

Beijing, China  
Nov. 2020 — Aug. 2022

- Worked on video captioning/retrieval and change captioning.

### Tsinghua Shenzhen International Graduate School

*Research Intern with A.P. Xingzheng Wang*

Shenzhen, China  
Mar. 2018 — Jun. 2018

- Worked on video Captioning.

### Prospective Research Lab, ICT, CAS

*Research Intern with Dr. Xishan Zhang*

Beijing, China  
Jun. 2016 — Jun. 2017

- Worked on video captioning.

## Honors & Awards

---

- Merit Student, University of Chinese Academy of Sciences, 2023.
- Excellent Master's Graduation Thesis, Kunming University of Science and Technology, 2022.
- Yunnan Merit Student, Yunnan Province, 2021, 2022.
- National Scholarship for Graduate Students, Ministry of Education of the People's Republic of China, 2021.
- Student Travel Grant, ACM International Conference on Multimedia (ACM MM), 2017.

## Academic Services

---

- PC member of conferences: AAAI 2023/2024, ACL 2023, EMNLP 2023, ACM MM 2023.
- Reviewer of journals: IEEE TCSVT (IF:8.4), IEEE TETCI (IF:5.3), Artificial Intelligence Review (IF:12.0).

## PUBLICATIONS

---

- Citations = 512; h-index = 8; i10-index = 6; as of December 12, 2023.
- Check my google scholar profile for the latest list of publications.

- [1] **Yunbin Tu**, Liang Li, Li Su, Zheng-Jun Zha, Chenggang Yan, Qingming Huang. Self-supervised Cross-view Representation Reconstruction for Change Captioning, **ICCV**, 2023.
- [2] **Yunbin Tu**, Liang Li, Li Su, Junping Du, Ke Lu, Qingming Huang, Viewpoint-Adaptive Representation Disentanglement Network for Change Captioning, IEEE Transactions on Image Processing (**TIP**), 2023 (IF:10.6).

- [3] **Yunbin Tu**, Liang Li, Li Su, Ke Lu, Qingming Huang, Neighborhood Contrastive Transformer for Change Captioning, *IEEE Transactions on Multimedia (TMM)*, 2023 (IF:7.3).
- [4] **Yunbin Tu**, Chang Zhou, Junjun Guo, Huafeng Li, Shengxiang Gao, Zhengtao Yu, Relation-aware attention for video captioning via graph learning, *Pattern Recognition (PR)*, 2023 (IF:8.0).
- [5] Shengbin Yue, **Yunbin Tu**, Liang Li, Ying Yang, Shengxiang Gao, Zhengtao Yu, I3N: Intra- and Inter-representation Interaction Network for Change Captioning, *IEEE Transactions on Multimedia (TMM)*, 2023 (IF:7.3).
- [6] **Yunbin Tu**, Liang Li, Li Su, Shengxiang Gao, Chenggang Yan, Zhengjun Zha, Zhengtao Yu, Qingming Huang, I<sup>2</sup>Transformer: Intra- and Inter-relation Embedding Transformer for TV Show Captioning, *IEEE Transactions on Image Processing (TIP)*, 2022 (IF:10.6).
- [7] Liang Li, Xingyu Gao, Jincan Deng, **Yunbin Tu**, Zhengjun Zha, Qingming Huang, Long Short-Term Relation Transformer with Global Gating for Video Captioning, *IEEE Transactions on Image Processing (TIP)*, 2022 (IF:10.6).
- [8] Gaoxiang Cong, Liang Li, Zhenhuan Liu, **Yunbin Tu**, Weijun Qin, Shenyuan Zhang, Chengang Yan, Wenyu Wang, Bin Jiang: LS-GAN: Iterative Language-based Image Manipulation via Long and Short Term Consistency Reasoning, *ACM MM*, 2022.
- [9] **Yunbin Tu**, Chang Zhou, Junjun Guo, Shengxiang Gao, Zhengtao Yu, Enhancing the Alignment between Target Words and Corresponding Frames for Video Captioning, *Pattern Recognition (PR)*, 2021 (IF:8.0).
- [10] **Yunbin Tu**, Liang Li, Tingting Yao, Jiedong Lou, Shengxiang Gao, Zhengtao Yu, Chenggang Yan, Semantic Relation-aware Difference Representation Learning for Change Captioning, *Findings of ACL (Long)*, 2021.
- [11] **Yunbin Tu**, Liang Li, Chenggang Yan, Shengxiang Gao, Zhengtao Yu, R<sup>3</sup>Net: Relation-embedded Representation Reconstruction Network for Change Captioning, *EMNLP (Long)*, 2021.
- [12] Chengang Yan, **Yunbin Tu**, Xingzheng Wang, Yongbing Zhang, Xinhong Hao, Yongdong Zhang, Qionghai Dai, STAT: spatial-temporal attention mechanism for video captioning, *IEEE Transactions on Multimedia (TMM)*, 2019 (IF:7.3).
- [13] **Yunbin Tu**, Xishan Zhang, Bingtao Liu, Chenggang Yan, Video Description with Spatial-Temporal Attention, *ACM MM*, 2017.