TENG WANG

☎ (+86)18814115767 ✉ ttengwang@gmail.com github.com/ttengwang ttengwang.com

EDUCATION

The University of Hong Kong, PhD Candidate Sept. 2020 – Jun. 2024
Computer Science, supervised by Prof. Ping Luo and Prof. Feng Zheng *Hong Kong*

Sun Yat-sen University, Master of Engineering (GPA: 3.93/4.00) Sept. 2017 – Jun. 2020
Pattern Recognition and Intelligent Systems, supervised by Prof. Huicheng Zheng *Guangzhou, China*

Sun Yat-sen University, Bachelor of Engineering (GPA: 3.70/4.00) Aug. 2013 – Jun. 2017
Telecommunication Engineering *Guangzhou, China*

INTERNSHIP

Video Captioning & Chinese Image Captioning May. 2019 – Aug. 2019
Research Intern at Tencent AI Lab, with Dr. Jingwen Wang & Dr. Lin Ma

RESEARCH INTEREST

Vision-Language Pretraining & Adaptation: pretraining a model by mining multimodal associations from large-scale unlabeled image-text data, and efficiently adapting the pretrained models for diverse downstream applications.

Fine-Grained Multimodal Models: exploiting fine-grained, region-level alignments between visual and textual data in complex scenes, *e.g.*, diverse captioning with multimodal controls, dense captioning & localization in videos.

PUBLICATIONS

(***** equal contribution)

[12] Baoshuo Kan*, **Teng Wang***, Wenpeng Lu, Xiantong Zhen, Weili Guan and Feng Zheng, "Knowledge-Aware Prompt Tuning for Generalizable Vision-Language Models," International Conference on Computer Vision (***ICCV***), 2023.

[11] Junjie Fei*, **Teng Wang***, Jinrui Zhang, Zhenyu He, Chengjie Wang and Feng Zheng, "Transferable Decoding with Visual Entities for Zero-Shot Image Captioning," International Conference on Computer Vision (***ICCV***), 2023.

[10] Dong Lu, Zhiqiang Wang, **Teng Wang**, Weili Guan, Hongchang Gao and Feng Zheng, "Cross-Modal Guided Attack: Boosting Adversarial Transferability of Vision-Language Pre-training Models," International Conference on Computer Vision (***ICCV***), 2023.

[9] **Teng Wang**, Yixiao Ge, Feng Zheng, Ying Shan, Xiaohu Qie and Ping Luo, "Free Language Modeling for Efficient Vision-Language Pretraining," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2023.

[8] Tiantian Geng, **Teng Wang**, Jinming Duan, Runmin Cong and Feng Zheng, "Dense-Localizing Audio-Visual Events in Untrimmed Videos: A Large-Scale Benchmark and Baseline," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2023.

[7] Chengyue Wu, **Teng Wang**, Yixiao Ge, Zeyu Lu, Ruisong Zhou, Ping Luo, Ying Shan, "π-Tuning: Transferring Multimodal Foundation Models with Optimal Multi-task Interpolation," in *Proceedings of International Conference on Machine Learning (**ICML**)*, 2023.

[6] **Teng Wang**, Wenhao Jiang, Zhichao Lu, Feng Zheng, Ran Cheng, Chengguo Yin and Ping Luo, "VLMixer: Unpaired Vision-Language Pre-training via Cross-Modal CutMix," in *Proceedings of International Conference on Machine Learning (**ICML**)*, 2022.

[5] Zhu Liu, **Teng Wang**, Jinrui Zhang, Feng Zheng, Wenhao Jiang and Ke Lu, "Show, Tell and Rephrase: Diverse Video Captioning via Two-stage Progressive Training," *IEEE Transactions on Multimedia (**TMM**)*, 2022.

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

[3] **Teng Wang**, Ruimao Zhang, Zhichao Lu, Feng Zheng, Ran Cheng and Ping Luo, "End-to-End Dense Video Captioning with Parallel Decoding," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (**ICCV**)*, 2021.

[2] **Teng Wang**, Huicheng Zheng, Mingjing Yu, Qian Tian and Haifeng Hu, "Event-centric Hierarchical Representation for Dense Video Captioning," in *IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**)*, vol. 31, no. 5, pp. 1890-1900, 2020.

[1] **Teng Wang**, Chen He and Haifeng Hu, "Image Caption with Endogenous-Exogenous Attention," in *Neural Processing Letters (**NPL**)*, vol. 50, no. 1, pp 431-443, 2019.

PREPRINTS

(***** equal contribution)

Teng Wang*, Jinrui Zhang*, Junjie Fei* et al., "Caption Anything: Interactive Image Description with Multimodal Controls," in arXiv:2305.02677, 2023.

Teng Wang*, Jinrui Zhang*, Feng Zheng, Wenhao Jiang, Ran Cheng and Ping Luo, "Learning Grounded Vision-Language Representation for Versatile Understanding in Untrimmed Videos," in arXiv:2303.06378, 2023.

COMPETITIONS & AWARDS

· 1st Place in Generic Event Boundary Captioning Track of LOVEU Challenge at CVPR 2023
· 1st Place in Make-up Dense Video Captioning Track of PIC Challenge at ACMMM 2022
· 1st Place in Make-up Temporal Video Grounding Track of PIC Challenge at ACMMM 2022
· 2nd Place in Generic Event Boundary Captioning Track of LOVEU Challenge at CVPR 2022
· 2nd Place in Event Dense-Captioning Track of ActivityNet Challenge at CVPR 2020-2022
· 3rd Place in TinyAction Challenge at CVPR 2021
· 2nd Place in SYSU Data Mining Competition, Sun Yat-sen University, 2015
· Postgraduate Scholarship, The University of Hong Kong, 2020-2024
· Academic Excellence Scholarships, Sun Yat-sen University, 2017-2020
· Academic Excellence Scholarships, Sun Yat-sen University, 2014-2016

ACADEMIC SERVICE

Journal Reviewer: International Journal of Computer Vision (IJCV), IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), IEEE Transactions on Artificial Intelligence (TAI), Complex & Intelligent Systems
Conference Reviewer: ICML, NeurIPS, CVPR, ICCV, ECCV

TECHNICAL SKILLS

Programming Languages	Python, MATLAB, C++, LaTex, R
Deep Learning Libraries	PyTorch, TensorFlow