

# Jing Tan

---

CONTACT INFORMATION	Room 703, Ho Sin-Hang Engineering Building, The Chinese University of Hong Kong, Shatin N.T. Hong Kong	(+852) 9360-4025 jingtan@ie.cuhk.edu.hk
RESEARCH INTERESTS	<b>Computer Vision:</b> Video analysis and understanding, Action and event detection <b>Machine Learning:</b> Representation learning, Deep learning	
EDUCATION	<b>Department of Information Engineering, The Chinese University of Hong Kong</b> Hong Kong SAR Ph.D. Student in <b>MMLab</b> Aug. 2023 – Present Supervisor: <b>Prof. Dahua Lin</b> <b>Department of Computer Science and Technology, Nanjing University</b> Nanjing, China Master in <b>MCG Lab</b> Sept. 2020 – June 2023 Supervisor: <b>Prof. Limin Wang</b> <b>Kuang Yaming Honors School, Nanjing University</b> Nanjing, China B.Sc., Major in Computer Science (GPA 4.44/5, Rank 3/29). Sept. 2016 – June 2020 <b>International Study Program in Letters and Science, UC Berkeley</b> Berkeley, USA Obtained <b>A</b> in COMPSCI 188 and STATC 100. Aug. 2018 – Dec. 2018	
PUBLICATION	<b>Jing Tan</b> , Yuhong Wang, Gangshan Wu, Limin Wang, Temporal Perceiver: A General Architecture for Generic Boundary Detection, in <i>IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)</i> , 2023. <b>Jing Tan</b> , Xiaotong Zhao, Xintian Shi, Bin Kang, Limin Wang, PointTAD: Multi-label Temporal Action Detection with Learnable Query Points, in <i>Advances in Neural Information Processing Systems 35 (NeurIPS)</i> , 2022. <b>Jing Tan*</b> , Jiaqi Tang*, Limin Wang, Gangshan Wu, Relaxed Transformer Decoders for Direct Action Proposal Generation, in <i>IEEE International Conference on Computer Vision (ICCV)</i> , 2021.	
HONORS AND AWARDS (SELECTED)	<ul style="list-style-type: none"><li>• <b>Outstanding Master Thesis Award</b> (3/226) Nanjing University 2023</li><li>• <b>Outstanding Graduate of Nanjing University</b> (20%) Nanjing University 2023</li><li>• <b>National Scholarship</b> (6/226) Ministry of Education 2022</li><li>• <b>3rd Place, GEBD-LOVEU Challenge @ CVPR'22</b> Workshop Committee 2022</li><li>• <b>Outstanding Postgraduate Student Award</b> Nanjing University 2021</li><li>• <b>1st Prize, Scholarship for Postgraduate Students</b> (20%) Nanjing University 2020 &amp; 2021</li><li>• <b>2nd Prize, Academic Outstanding Scholarship for Elite Class</b> Nanjing University 2017 &amp; 2018</li></ul>	
RESEARCH EXPERIENCE	<b>EgoTAD: Exploring Spatial Semantics for Egocentric Temporal Action Detection</b> Advisor: <b>Prof. Limin Wang</b> Dec. 2022 - Mar. 2023 <ul style="list-style-type: none"><li>• As the first author, proposed a temporal action detection (TAD) framework especially for egocentric videos. The work emphasizes the importance of spatial modeling over sophisticated temporal modeling in egocentric scenario and leverages spatial object semantics with learned key object tokens for accurate egocentric TAD.</li><li>• Extensive experiments on Epic Kitchens 100 and Ego4D Memory Queries benchmarks show that EgoTAD achieves superior performance compared to previous methods.</li></ul> <b>PointTAD: Multi-label Temporal Action Detection with Learnable Query Points.</b> Advisor: <b>Prof. Limin Wang</b> Dec. 2021 - Jun. 2022 <ul style="list-style-type: none"><li>• We study the complex multi-label TAD that requires fine-grained class discrimination and dedicated module to precisely localize co-occurring action instances. As the first author, I presented a sparse, query-based temporal action detector with learnable query points to tackle the concurrent instances and their richer relations in multi-label TAD setup.</li><li>• PointTAD streamlines an end-to-end learnable framework taking RGB input only. Extensive experiments on MultiTHUMOS and Charades show that PointTAD surpasses all previous methods by a large margin under detection-mAP and achieves promising results under segmentation-mAP. The code is available at <a href="https://github.com/MCG-NJU/PointTAD">https://github.com/MCG-NJU/PointTAD</a>.</li></ul>	

## Temporal Perceiver: A General Architecture for Arbitrary Boundary Detection.

Advisor: Prof. Limin Wang

July 2021 - Nov. 2021

- As the first author, presented a general and efficient architecture for generic boundary detection, namely Temporal Perceiver, which detects temporal boundaries that divide the input videos into taxonomy-free and semantic coherent units and generalizes well to arbitrary generic boundaries.
- Temporal Perceiver achieves state-of-the-art performances on shot-level, event-level and scene-level generic boundary benchmarks, with 6× faster inference speed and 160× less FLOPs than previous SOTA. The paper is available at <https://arxiv.org/abs/2203.00307>.

## Relaxed Transformer Decoders for Direct Action Proposal Generation

Advisor: Prof. Limin Wang

Sept. 2020 - Mar. 2021

- As the first author, presented a simple and efficient framework (RTD-Net) based on Transformer-alike architecture for direct and sparse action proposal generation, which locates action instances in untrimmed videos with more complete and accurate proposals.
- RTD-Net tackles the essential difference between time and space, with several important improvements on Transformers: boundary-attentive modules for discriminative feature encoding, relaxed matching criteria for optimal training and three-branch head design for accurate and reliable proposal generation.
- RTD-Net outperforms the existing state-of-the-art methods on THUMOS14 dataset and achieved comparable performance on ActivityNetv1.3. The code is available at <https://github.com/MCG-NJU/RTD-Action>.

CONTEST  
EXPERIENCE

## Generic Event Boundary Detection (GEBD) Challenge.

April 2022 - June 2022

LOVEU Workshop & International Challenge @ CVPR'22.

- We won the third place in the challenge, the submission report is available at <https://arxiv.org/abs/2206.15268>.
- Contribution: re-designed the boundary-attentive modules for discriminative representations and adopted the Transformer decoders to achieve accurate proposals without dense duplicates.

INTERNSHIP

## Video Algorithm Intern

Beijing, China

Platform and Content Group, Tencent

Dec. 2021 - June 2023

- Our group provides online and automatic checking services for newly uploaded video content based on deep neural networks.
- Build a sparse, query-based temporal action detector that is able to precisely localize concurrent instances in multi-label videos and egocentric videos.
- This temporal action detector is end-to-end trainable with feature extractor and requires only RGB frame input, which is easy for online deployment and automatic video processing.

TEACHING  
EXPERIENCE

## Digital Logic Design and Computer Organization

Nanjing University

Teaching assistant

Mar. 2021 - July 2021

COURSE  
PROJECTS  
(SELECTED)

## NYC Taxi Rides Analysis.

Berkeley, USA

STATC100: Principles and Techniques of Data Science.

Sept. 2018 - Dec. 2018

- Predicted accidents and popular pick-up locations in New York City with data capturing, cleaning and visualization, EDA and feature engineering on NYC taxi rides dataset.

## Advanced HTTP server.

Hong Kong SAR

Competition held during short visit to HKUST

Feb. 2018

- Developed an advanced HTTP server that supports media display and multi-thread feature.
- Won the 2nd place in final evaluation.

SKILLS

- Programming: Python, PyTorch,  $\LaTeX$ , C++, C,
- Languages: Mandarin, English (TOFEL 113, R 30, L 29, S 26, W 28)