Sujia Wang

◇ Phone: (+86) 131-8723-0103 ◇ Email: wsj22@mails.tsinghua.edu.cn ◇ Home ◇ Google Scholar

EDUCATION

Tsinghua University (THU), CHINA Sept. 2022 - Jun. 2025
M.E. in Data Science and Information Technology *GPA: 3.6/4*

- Advised by Prof. **Yansong Tang** and Prof. **Lei Chen**
- Research on Video Understanding and Multimodal Learning
- **Core Courses:** Machine Learning, Learning from Data and Foundations for Big Data Analytics

University of Cambridge (CAMB), UK Nov. 2021 - Mar. 2022
Winter Program *Grand: A*

- Advised by Prof. **Pietro Lio**
- Project on Path planning of unmanned surface vehicles based on deep learning
- **Core Courses:**Machine Learning and Neural Networks

Northeastern University (NEU), CHINA Sept. 2018 - Jun. 2022
B.S. in Mathematics and Applied Mathematics *GPA: 3.3/4*

- Advised by Prof. **Zhiqiong Wang** and Prof. **Xinhui Shao**
- Research on Graph and Information Theory Applications in Biology
- **Core Courses:** Mathematical Analysis, Advanced Algebra, Probability Theory, and Analytic Geometry

PUBLICATIONS AND PATENTS

J = Journel, C = Conference, P = Patent, S = Software Copyright, *: Equal Contribution

Video Understanding

[C1] **Sujia Wang***, Xiangwei Shen*, Yansong Tang, Xin Dong, Wenjia Geng, and Lei Chen. **Localization-Aware Multi-Scale Representation Learning for Repetitive Action Counting**, ***VCIP(Oral)***, 2024.

[C2] Shiyi Zhang, Wenxun Dai, **Sujia Wang**, Xiangwei Shen, Jiwen Lu, Jie Zhou, and Yansong Tang. **LOGO: A Long-Form Video Dataset for Group Action Quality Assessment**, in ***CVPR***, 2023. [Paper] [Code]

[P1] Yansong Tang and **Sujia Wang**. **Action Recognition Method, Device, Computer Equipment, Storage Medium, and Computer Program Product**, Chinese Patent Application Number: 10000533025811, 2024, Under Substantive Examination.

Multimodal Learning

[C3] Wenjia Geng, **Sujia Wang**, Baoliang Tain, Zhang Xuezhong, Wencheng Zhu, Yansong Tang, and Lei Chen. **CoSTL: Comprehensive Spatial-Temporal Representation Learning for Moment Retrieval and Highlight Detection**, in ***CVPR***, 2025, Under Review.

[C4] Wenjia Geng, Yong Liu, Lei Chen, **Sujia Wang**, Jie Zhou, and Yansong Tang. **Learning Multi-Scale Video-Text Correspondence for Weakly Supervised Tasks**, in ***AAAI***, 2023. [Paper]

Graph and Information Theory Application

[J1] **Sujia Wang**, Yunqi Liu, Qixuan Sun and Zhiqiong Wang. **Optimization Method for Weak Association Regulation in Gene Regulatory Networks**, in ***Think Tank Era***, 2022(17): 207-210. [Paper]

[S1] **Sujia Wang**, Zhiqiong Wang, Yunqi Liu, Qixuan Sun, Yunrui Hao, and Renfei Gao. **Optimization System for Gene Regulatory Networks Targeting Key Node Substructures V1.0**, Chinese Software Copyright, Registration Number: 2022SR0278775, 2022.

[S2] **Sujia Wang**, Zhiqiong Wang, Yunqi Liu, Qixuan Sun, Renfei Gao, and Yunrui Hao. **Information Theory-Based Weak Association Regulation Optimization System for Gene Regulatory Networks V1.0**, Chinese Software Copyright, Registration Number: 2022SR0138787, 2022.

[S3] Qixuan Sun, Zhiqiong Wang, Yunrui Hao, Renfei Gao, **Sujia Wang**, and Yunqi Liu. **Optimization System for Gene Regulatory Networks Targeting Compact Structures V1.0**, Chinese Software Copyright, Registration Number: 2022SR0283545, 2022.

RESEARCH EXPERIENCE

Comprehensive Human Action Understanding in Videos: From Coarse to Fine Granularity

Temporal-Spatial Action Localization (TAL) Present

- **Led** the formation of a data-labeling team of 10 trained professionals, pre-trained and fine-tuned YOLO-V5 models and X-AnyLabeling tools to accelerate and improve the labeling process, contributing 300 hours of work.
- **Proposed** the Kitchen Monitoring Dataset, which includes over 8,000 multi-person temporal-spatial bounding box labels from real-life scenarios for action recognition and monitoring.

Repetitive Action Counting (RAC) [C1] 2023 - 2024

- **Proposed** the Localization-Aware Multi-Scale Representation Learning (LMRL) framework to improve RAC by addressing noise from action interruptions and inconsistencies.
- **Developed** a Multi-Scale Period-Aware Representation (MPR) to handle diverse action frequencies and a Repetition Foreground Localization (RFL) method to enhance action representation with global semantic information.
- **Achieved** state-of-the-art counting performance on the main benchmarks, one conference paper published on VCIP (2024 oral), supervised by Prof. **Yansong Tang** and Prof. **Lei Chen**.

Action Quality Assessment (AQA) [C2] 2022 - 2023

- **Led** an 8-person team with professional athletes in frame-level labeling of competition videos, utilizing LabelMe and COIN annotation tools, contributing 600 hours of work.
- **Introduced** the LOGO dataset with 200 videos from 26 artistic swimming events and developed the *Group-Aware Attention Module* to enhance AQA representations.
- **Published** one conference paper on CVPR (2023), supervised by Prof. **Yansong Tang** and Prof. **Jiwen Lu**.

Video-Text Multimodal Understanding

Video Highlight and Moment Retrieval (VH and MR) [C3] Present

- **Proposed** the CoSTL framework for video moment retrieval and highlight detection, addressing the challenge of simultaneously capturing fine-grained image-level information and temporal dynamics.
- **Developed** a two-step, text-driven fine-grained image encoder and a multi-scale temporal perception module, improving both spatial and temporal understanding.
- **Achieved** best performance on four main public benchmarks, and one conference paper under the review of CVPR (2025).

Weakly Supervised temporal Article Grounding (WSAG) [C4] 2023 - 2024

- **Proposed** the MVTCL framework for weakly supervised temporal article grounding, addressing the challenge of aligning multi-scale semantic information in both video and text modalities.
- **Developed** a semantic calibration module to align hierarchical textual content with video segments, and introduced a multi-scale contrastive learning module to enhance discriminative representations.
- **Achieved** state-of-the-art through the innovative architecture and supervision design, and one conference paper published on AAAI (2023), supervised by Prof. **Yansong Tang** and Prof. **Jie Zhou**

OxCam Research Programme

Path planning of unmanned surface vehicles based on deep learning 2022

- **Proposed** an unmanned ship IoT model and used U-net to semantically segment water images.
- **Awarded** as Excellent Team and received a Grand A, supervised by Prof. **Pietro Liò** .

Applications of Information Theory and Graph Theory in Biology

Gene Regulatory Network Optimization Technique in Breast Cancer [J1][S1],[S2],[S3] 2021 - 2022

- **Proposed** three new algorithms based on the structural analysis and information theory reducing redundant edges in the network to improve its accuracy.
- **Awarded Provincial Excellent Undergraduate Student Project**, supervised by Prof. **Zhiqiong Wang**.

AWARDS AND HONORS

Academic Contests

- **Second Prize**, Tsinghua-Berkley Shenzhen Institute Student Poster Competition — Aug. 2024
- **Golden Prize**, Global Citizens Open Innovation SDGs Challenge — Dec. 2022
- **Second Prize**, National College Students Mathematical Contest in Modeling in LN — Sept. 2021
- **First Prize**, Northeastern University Mathematical Contest in Modeling — Sept. 2019

Scholarships

- **First Prize Scholarship, Tsinghua University** — 2023 - 2024
- **Second Prize Scholarship, Northeastern University** — 2018 - 2019

SKILLS

Programming Languages: Python, Shell, LaTeX, MATLAB, C++/C, Java/JavaScript
Programming Tools: Git, PyTorch