TAO SHI

Tel: (+86) 159-8931-6450 | Email: shitao21@mails.tsinghua.edu.cn
Homepage: https://taoshi1998.github.io.

EDUCATION

Tsinghua-UC Berkeley Shenzhen Institute (TBSI), Tsinghua University — Shenzhen, China
Master of Science in Data Science and Information Technology — *Sep 2021 - Present*
Cumulative GPA: 3.94/4.00

China University of Mining and Technology (CUMT) — Xuzhou, China
Bachelor of Engineering in Computer Science and Technology — *Sep 2016 - Jun 2020*
Cumulative GPA: 89.73/100, Ranking: 8/243

Australian National University — Canberra, Australia
Exchange Program at the School of Computing — *Feb 2018 - Jul 2018*
Cumulative GPA: 6.75/7.00

RESEARCH INTERESTS

Natural Language Processing, Multimodal Learning, Machine Learning

PUBLICATIONS

* indicates equal contribution

Tao Shi and Shao-Lun Huang. "MultiEMO: An Attention-Based Correlation-Aware Multimodal Fusion Framework for Emotion Recognition in Conversations". *ACL 2023*. [PDF] [Code]

Tao Shi*, Xiao Liang*, Yaoyuan Liang, Xinyi Tong, and Shao-Lun Huang. "SSLCL: An Efficient Model-Agnostic Supervised Contrastive Learning Framework for Emotion Recognition in Conversations". *Under Review by AAAI 2024*. [ArXiv Preprint] [Code]

Xiao Liang*, **Tao Shi***, Yaoyuan Liang, Te Tao, and Shao-Lun Huang. "Exploring Iterative Refinement with Diffusion Models for Video Grounding". *Under Review by AAAI 2024*.

RESEARCH EXPERIENCE

Exploring Iterative Refinement with Diffusion Models for Video Grounding — May 2023 - Aug 2023
Advisor: Prof. Shao-Lun Huang — *Tsinghua University*

- We innovatively formulated video grounding as a conditional generative task using diffusion models, which enabled iterative refinements of predicted spans through the reversed denoising diffusion process.
- A video-centered multimodal encoder was designed to facilitate the interaction between video and sentence features, and a specialized span refining decoder was introduced to effectively generate target spans.

SSLCL: An Efficient Model-Agnostic Supervised Contrastive Learning Framework for Emotion Recognition in Conversations — Feb 2023 - Aug 2023
Advisor: Prof. Shao-Lun Huang — *Tsinghua University*

- Through a novel utilization of label representations, we effectively addressed the constraints posed by large batch sizes and incompatibility with most existing ERC architectures encountered in current supervised contrastive learning (SCL)-based methods.
- We were the first in the SCL community to leverage Soft-HGR maximum correlation as a measure of similarity.
- We innovatively leveraged multimodal information as data augmentation to enhance model performances.

MultiEMO: An Attention-Based Correlation-Aware Multimodal Fusion Framework for Emotion Recognition in Conversations Aug 2022 - Jan 2023

Advisor: Prof. Shao-Lun Huang *Tsinghua University*

- We proposed a novel visual feature extraction network named VisExtNet, which effectively captured visual cues of interlocutors without modeling redundant scene information.
- We designed a multimodal fusion model called MultiAttn based on bidirectional multi-head cross-attention layers, which successfully modeled the complicated correlations across textual, audio and visual modalities.
- A sample-weighted focal contrastive (SWFC) loss was introduced to address the difficulty of classifying minority and semantically similar emotion classes.

AWARDS AND HONORS

Runner-Up of the 6th TBSI Retreat Poster Competition [Poster], *Tsinghua University* (top 1%) *Jul 2023*
Runner-Up of the 4th Tsinghua SDG Open Hack Competition [Slides], *Tsinghua University* (top 5%) *Nov 2022*
Outstanding Graduate of the Class of 2020, *CUMT* (top 3%) *Jun 2020*
Outstanding Graduate Thesis and Dissertation Award, *CUMT* (top 1%) *Jun 2020*
First Prize in 2019 National English Competition for College Students, *Ministry of Education* (top 1%) *May 2019*
Outstanding Undergraduate International Exchange Scholarship, *China Scholarship Council* (top 3%) *Nov 2017*

WORK EXPERIENCE

Teaching Assistant of Seminar in Data Science and Information Technology, *Tsinghua University* *Spring 2023*

ACADEMIC SERVICE

Conference Reviewer at *EMNLP 2023* *Aug 2023 - Sep 2023*

ENGLISH PROFICIENCY

TOEFL: 110, GRE: 333 + 4.0