Fuzhao XUE

+65 83434586 | xuefuzhao@outlook.com | https://xuefuzhao.github.io | Singapore

EDUCATION

Doctor of Philosophy (Ph.D.)   Jul. 2025 (expected)
   National University of Singapore   Singapore
- Majored in Computer Science
- Supervisor: Yang You
- President's Graduate Fellowship
- Published papers at AAAI 2022, ICASSP 2022, CVPR 2022
- Internship at Google Brain and worked with Yi Tay and Mostafa Dehghani

Master of Engineering (M.Eng.)   Jul. 2021
   Nanyang Technological University   Singapore
- Majored in Computer Science and Engineering
- Supervisor: Eng-Siong Chng and Aixin Sun
- Thesis: Refining latent multi-view graph for relation extraction
- GPA: 5.0/5.0
- Published paper at AAAI 2021 and ICML 2020

RESEARCH INTEREST

- **Machine Learning:** Foundation Model and Self-supervised Learning
- **Natural Language Processing:** Language Modeling and Relation Extraction.
- **High Performance Computing:** Model Parallelism and Large-scale Deep Learning System.

PROFESSIONAL EXPERIENCE

Google Brain | Work Remotely from Singapore   *Student Researcher*   Jul. 2022-Present
In charge of adaptive computation with adaptive model depth and elastic input sequence. Submitted part of the work at ICLR 2023.
- Implemented universal transformer and PonderNet on vision transformer based on Scenic and made vision universal transformer open source.
- Proposed AdaTape to achieve adaptive computation with elastic input sequence, which is a novel strategy that enables dynamic computation in neural networks via adaptive tape token.

National University of Singapore | Singapore   *Ph.D. Candidate*   Jul. 2021-Present
In charge of efficient and effective general training framework for both computer vision and natural language processing to scale transformer. Published at AAAI 2022.
- Proposed to go wider instead of deeper, compressing along depth by parameter sharing and scaling along width by mixture-of-experts to construct a parameter-efficient framework.
- Proposed knowledge integration to integrate knowledge from pretrained MoE into dense student model, which is efficient and hardware-friendly as other dense model but is as effective as MoE counterpart.
- Proposed sequence parallelism to train transformer with longer sequence from system perspective.
- Proposed Bamboo idea to tune transformer configuration.

Nanyang Technological University | Singapore   *M.Eng. Student*   Jul. 2020-Jun. 2021
In charge of design of improving dialogue-level relation extraction by adaptive graph pooling proposed. Published paper at AAAI 2021.
- Identified indictive words using Dynamic Time Wrapping Pooling in an unsupervised manner with high accuracy.
- Improved the state-of-the-art of dialogue-level relation extraction by 6% on DialogRE.

National University of Singapore | Singapore   *Research Intern*   Oct. 2019-Apr. 2020
In charge of design of hybrid speech recognition framework with utilization of Pytorch and Kaldi. Published paper at ICML 2020.
- Used Pytorch underlying operations to implement acoustic models such as DNN, LSTM, SRU, RRN and RTN.
- Modeled relational thinking through Deep Graph Process, reducing WER by 10% relatively on Chime2, Chime5 and SWBD.

SKILLS

- **Programming:** Python, C/C++, C#, Matlab
- **Frameworks:** JAX, TensorFlow, PyTorch, Keras, Theano, Scikit-learn, pandas,

SERVICES

- **Conference:** ICML 2022 Reviewer, SIGIR 2022 PC Member, EMNLP 2022 Reviewer

CONFERENCE

- **Fuzhao Xue**, Ziji Shi, Futao Wei, Yong Liu, Yang You "*Go Wider Instead of Deeper*" Published at Association for the Advancement of Artificial Intelligence **(AAAI) 2022**

- **Fuzhao Xue**, Aixin Sun, Hao Zhang, Eng-Siong Chng "*An Embarrassingly Simple Model for Dialogue Relation Extraction*" Published at International Conference on Acoustics, Speech, & Signal Processing **(ICASSP) 2022**

- Andrew Koh, **Fuzhao Xue**, Eng Siong Chng *"Automated Audio Captioning using Transfer Learning and Reconstruction Latent Space Similarity Regularization"* Published at International Conference on Acoustics, Speech, & Signal Processing **(ICASSP) 2022**

- Wangbo Zhao, Kai Wang, Xiangxiang Chu, **Fuzhao Xue**, Xinchao Wang, and Yang You *"Modeling Motion with Multi-Modal Features for Text-Based Video Segmentation"* Published at IEEE Conference on Computer Vision and Pattern Recognition **(CVPR), 2022**

- **Fuzhao Xue**, Aixin Sun, Hao Zhang, Eng-Siong Chng *"GDPNet: Refining Refining Latent Multi-View Graph for Relation Extraction"* Published at Association for the Advancement of Artificial Intelligence **(AAAI) 2021**

- Hengguan Huang, **Fuzhao Xue**, Hao Wang, Ye Wang * "*Deep Graph Random Process for Relational-Thinking-Based Speech Recognition*" Published at International Conference on Machine Learning (**ICML**) **2020**

JOUNRAL

- Xingmei Wang, Jiaxiang Meng, Bin Wen, **Fuzhao Xue** *"RACP: A network with Attention Corrected Prototype for Few-shot Speaker Recognition using Indefinite Distance Metric"* Published at **Neurocomputing**

- Xingmei Wang, **Fuzhao Xue**, Wei Wang, Anhua Liu. *"A network model of speaker identification with new feature extraction methods and asymmetric BLSTM"* Published at **Neurocomputing**

- Xingmei Wang, Anhua Liu, Yu Zhang, **Fuzhao Xue**. *"An Underwater Acoustic Target Recognition: A Combination of Multi-dimensional Fusion Features and Modified Deep Neural Network"* Published at **Remote Sensing**.

PREPRINTS

* denote equal contribution.

- Zangwei Zheng, Pengtai Xu, Xuan Zou, Da Tang, Zhen Li, Chenguang Xi, Peng Wu, Leqi Zou, Yijie Zhu, Ming Chen, Xiangzhuo Ding, **Fuzhao Xue**, Ziheng Qing, Youlong Cheng, Yang You *"CowClip: Reducing CTR Prediction Model Training Time from 12 hours to 10 minutes on 1 GPU"* arXiv preprint **arXiv: 2204.06240**

- **Fuzhao Xue**, Xiaoxin He, Xiaozhe Ren, Yuxuan Lou, Yang You "*One Student Knows All Experts Know: From Sparse to Dense*" *arXiv preprint* **arXiv:2201.10890**

- Xiaoxin He, **Fuzhao Xue**, Xiaozhe Ren, Yang You *"Large-Scale Deep Learning Optimizations: A Comprehensive Survey"* arXiv preprint **arXiv:2111.00856**

- Yuxuan Lou, **Fuzhao Xue**, Zangwei Zheng, Yang You *"Cross-token Modeling with Conditional Computation" arXiv preprint* **arXiv:2109.02008**

- Shenggui Li*, **Fuzhao Xue***, Yongbin Li, Yang You "*Sequence Parallelism: Long Sequence Training from System Perspective*" arXiv preprint **arXiv: 2105.13120**

- Jinjie Ni, Tom Young, Vlad Pandelea, **Fuzhao Xue**, Vinay Adiga, Erik Cambria "*Recent Advances in Deep Learning-based Dialogue Systems" arXiv preprint* **arXiv: 2105.04387**