

# Fuzhao XUE

+65 83434586 | xuefuzhao@outlook.com | Singapore

## EDUCATION

<b>Doctor of Philosophy (Ph.D.)</b> National University of Singapore	Jul. 2025 (expected) Singapore
<ul style="list-style-type: none"><li>• Majored in Computer Science</li><li>• President's Graduate Fellowship</li><li>• Published paper at AAAI 2022</li><li>• Published two papers at ICASSP 2022</li><li>• Submitted two papers at IJCAI 2022</li><li>• Submitted two papers at ICML 2022</li></ul>	
<b>Master of Engineering (M.Eng.)</b> Nanyang Technological University	Jul. 2021 Singapore
<ul style="list-style-type: none"><li>• Majored in Computer Science and Engineering</li><li>• Thesis: Refining latent multi-view graph for relation extraction</li><li>• GPA: 5.0/5.0</li><li>• Published paper at AAAI 2021</li><li>• Published paper at ICML 2020</li></ul>	
<b>Bachelor of Engineering (B.Eng.)</b> Harbin Engineering University	Jul. 2020 Harbin, China
<ul style="list-style-type: none"><li>• Majored in Computer Science and Technology</li><li>• 6 times University First-Class Scholarship</li><li>• 4<sup>th</sup> Place of China AI Futurelab Contest</li><li>• 2<sup>nd</sup> Prize of 2017 National Robot and Artificial Intelligence Competition "Human-like Robot Dash Project"</li><li>• National University of Singapore Summer Workshop</li></ul>	
Grade: A	Singapore

## PROFESSIONAL EXPERIENCE

<b>National University of Singapore   Singapore</b> <b>Ph.D. Student</b>	Jul. 2021-Present
<i>In charge of efficient and effective general training framework for both computer vision and natural language processing to scale transformer. Published at AAAI 2022. Submitted at IJCAI 2022, ICML 2022.</i>	
<ul style="list-style-type: none"><li>• 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.</li><li>• Proposed knowledge integration to integrate knowledge from pretrained MoE into dense student for a dense model, which is efficient and hardware-friendly as other dense model but is as effective as MoE counterpart.</li><li>• Proposed sequence parallelism to train transformer with longer sequence from system perspective.</li></ul>	
<b>Nanyang Technological University   Singapore</b> <b>M.Eng. Student</b>	Jul. 2020-Jun. 2021
<i>In charge of design of improving dialogue-level relation extraction by adaptive graph pooling proposed. Published paper at AAAI 2021</i>	
<ul style="list-style-type: none"><li>• Identified indicative words using Dynamic Time Wrapping Pooling in an unsupervised manner with high accuracy.</li><li>• Improved the state-of-the-art of dialogue-level relation extraction by 6% on DialogRE.</li></ul>	
<b>National University of Singapore   Singapore</b> <b>Research Intern</b>	Oct. 2019-Apr. 2020
<i>In charge of design of hybrid speech recognition framework with utilization of Pytorch and Kaldi. Published paper at ICML 2020</i>	
<ul style="list-style-type: none"><li>• Used Pytorch underlying operations to implement acoustic models such as DNN, LSTM, SRU, RRN and RTN.</li><li>• Modeled relational thinking through Deep Graph Process, reducing WER by 10% relatively on Chime2, Chime5 and SWBD.</li></ul>	
<b>National University of Singapore Summer Workshop   Singapore</b>	Jul. 2018-Aug. 2018
<i>Designed fire-fighting trolley with automatic driving, speaker recognition, target detection, image classification, automatic obstacle avoidance, semi-automatic driving with manual driving, voice control, and fully automatic driving. <b>Grade: A</b></i>	
<ul style="list-style-type: none"><li>• Solved problem of image uploading speed of trolley through MQTT protocol being way faster than model recognizing speed, and significantly reduced latency from 9s to 2s.</li><li>• Preprocessed all audio data with sox as well as modeled first 0.25s of each channel voice and uniformed sampling rate, improved speech recognition and voiceprint recognition and audio data quality.</li></ul>	

## SKILLS

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

## **PUBLICATIONS**

---

- **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**
- **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**
- 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.

- **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**
- Shenggui Li\*, **Fuzhao Xue**\*, Yongbin Li, Yang You "Sequence Parallelism: Long Sequence Training from System Perspective" *arXiv preprint* **arXiv: 2105.13120**
- Yuxuan Lou, **Fuzhao Xue**, Zangwei Zheng, Yang You "Cross-token Modeling with Conditional Computation" *arXiv preprint* **arXiv:2109.02008**
- Xiaoxin He, **Fuzhao Xue**, Xiaozhe Ren, Yang You "Large-Scale Deep Learning Optimizations: A Comprehensive Survey" *arXiv preprint* **arXiv:2111.00856**
- 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**