Fangrui Zhu

Email: zhu.fang@northeastern.edu   Webpage: https://fangruizhu.github.io/

Education	**Doctor of Philosophy, Computer Science** *Northeastern University, 2021.9 – present*

Doctor of Philosophy, Computer Science   *Northeastern University, 2021.9 – present*
- Advisor: Prof. Huaizu Jiang.
- Working on multi-modality models and relationship reasoning.

Master of Science, Statistics   *Fudan University, 2018.9 – 2021.7*
- GPA: 3.58/4.0.
- Advisor: Prof. Yanwei Fu.
- Working on statistical learning and computer vision.
- Thesis: Self-supervised representation learning from videos by exploiting spatio-temporal consistency.

Bachelor of Engineering, Software Engineering   *Tongji University, 2014.9 – 2018.7*
- GPA: 4.68/5.0, Rank: 6/165.
- Advisor: Prof. Lin Zhang.
- Software development and testing, architecture design.
- Thesis: Learning pedestrian attributes for person re-identification.

Research Interests

- Multi-modal LLMs: Developing large language models that process video, text, and 3D data, with a focus on understanding spatial and temporal information from video inputs.
- Visual Grounding & Relationship Understanding: Aligning language with visual content and modeling object relationships to improve scene understanding and reasoning.

Publications and Preprints

[1] **Fangrui Zhu**, Hanhui Wang, Yiming Xie, Jing Gu, Tianye Ding, Jianwei Yang, Huaizu Jiang. Unlocking 3D Spatial Reasoning in LMMs with Perception-Guided Prompting. *Under review.*

[2] **Fangrui Zhu**, Jianwei Yang, Huaizu Jiang. Towards Flexible Visual Relationship Segmentation. The Thirty-Eighth Annual Conference on Neural Information Processing Systems (NeurIPS), 2024. [PDF]

[3] **Fangrui Zhu**, Yiming Xie, Weidi Xie, Huaizu Jiang. Diagnosing human-object interaction detectors. International Journal of Computer Vision (IJCV), 2025. [PDF]

[4] Mu Cai, Reuben Tan, Jianrui Zhang, Bocheng Zou, Kai Zhang, Feng Yao, **Fangrui Zhu**, Jing Gu, Yiwu Zhong, Yuzhang Shang, Yao Dou, Jaden Park, Jianfeng Gao, Yong Jae Lee, Jianwei Yang. Temporalbench: Benchmarking fine-grained temporal understanding for multimodal video models. Workshop on Video-Language Models in NeurIPS, 2024. [PDF]

[5] Alberto M. Ceballos-Arroyo, Hieu T. Nguyen, **Fangrui Zhu**, Shrikanth M. Yadav, Jisoo Kim, Lei Qin, Geoffrey Young, Huaizu Jiang. Vessel-aware Aneurysm Detection Using Multi-scale Deformable 3D Attention, International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2024. [PDF]

[6] Zeyu Han, **Fangrui Zhu**, Qianru Lao, Huaizu Jiang. Zero-shot Referring Expression Comprehension via Structural Similarity Between Images and Captions, IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR), 2024. [PDF]

[7] **Fangrui Zhu**, Yiming Xie, Weidi Xie, Huaizu Jiang. Diagnosing Error in Human-object Interaction Detectors, *to appear* in International Journal of Computer Vision (IJCV). [PDF]

[8] **Fangrui Zhu**, Yi Zhu, Li Zhang, Chongruo Wu, Yanwei Fu, Mu Li. A Unified Efficient Pyramid Transformer for Semantic Segmentation, 2021 Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2667-2677. [PDF]

[9] **Fangrui Zhu**, Li Zhang, Yanwei Fu, Guodong Guo, Weidi Xie. Self-supervised Video Object Segmentation, arXiv 2006.12480. [PDF]

[10] Xuelin Qian, Wenxuan Wang, Li Zhang, **Fangrui Zhu**, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, Xiangyang Xue. Long-Term Cloth-Changing Person Re-identification, 2020 IEEE Asian Conference on Computer Vision (ACCV). [PDF]

[11] Yu Xie, **Fangrui Zhu**, Yanwei Fu. Main-Secondary Network for Defect Segmentation of Textured Surface Images, 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), 3531-3540. [PDF]

[12] Jianhui Chen, **Fangrui Zhu**, James J. Little. A Two-Point Method for PTZ Camera Calibration in Sports, 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 287–295. [PDF]

Experience	**Research Intern**	*Microsoft Research, 2024.6 – 2024.11*

- Advised by Dr. Jianwei Yang.
- Topic: Video Foundation model.
- Deep learning group.

Research Intern *Amazon Alexa AI, 2022.6 – 2022.9*
- Advised by Dr. Emre Barut.
- Topic: Mitigating Bias in Image Captioning.

Research Intern *Baidu Research, 2021.7 – 2021.9*
- Research Intern in IDL group.
- Advised by Dr. Guodong Guo.
- Topic: Vision-language representation learning.

Research Intern *Amazon AWS, 2020.6 – 2020.12*
- Research Intern in GluonCV group.
- Advised by Dr. Yi Zhu.
- Topic: Attention-based semantic segmentation.

Research Assistant *Computer Vision Lab at Fudan University, 2018.9 – 2020.5*
- Advisor: Prof. Yanwei Fu.
- Cloth-inconsistent person re-identification (ReID) and occluded ReID.
- Worked closely with Dr. Weidi Xie from the VGG group on self-supervised video object segmentation.

Research Intern *University of British Columbia, 2017.6 – 2017.9*
- Worked as a Mitacs Research Intern.
- Advisor: Prof. James Little and Dr. Jianhui Chen.
- Low-level computer vision based camera calibration.

Selected Awards		
	Second prize, Award on Fudan Excellent Student Scholarship (top 15%)	2020
	Second prize, National Post-Graduate Mathematical Modeling Contest	2019
	Second prize, Award on Fudan Excellent Student Scholarship (top 15%)	2019
	'GouXiong Hui' data mining project funding (¥2,000)	2019
	Shanghai Outstanding Graduate Student Award (top 5%)	2018
	CSC and Mitacs Funding for undergraduate internship ($5,000 CAD)	2017
	'Bo Shi' Scholarship (top 10%)	2017
	Award on Tongji Excellent Social Activity Student Scholarship	2017
	Second prize, 'Challenge Cup' technological innovation competition	2017
	Tongji Outstanding Undergraduate Student (top 5%)	2016

First prize, Award on Tongji Excellent Student Scholarship (¥5,000)(top 5%)
2016

Award on Tongji Excellent Social Activity Student Scholarship   2016

Tongji Outstanding Undergraduate Student (top 5%)   2015

First prize, Award on Tongji Excellent Student Scholarship (¥5,000)(top 5%)
2015

Shanghai Technological Innovation Projects for Undergraduates (funded with
¥10,000)   2015

Teaching Experience

- **Teaching Assistant** for CS7150 Deep Learning (2022 Fall), Northeastern University.
- **Teaching Assistant** for DATA620007 Data Mining (2020 Spring), Fudan University.
- **Teaching Assistant** for DATA130020.01 Principle and Application of Database (2019 Spring), Fudan University.
- **Teaching Assistant** for Deep Learning (2021 Spring), Northeastern University.
- **Teaching Assistant** for Computer Vision (2022 Fall), Northeastern University.

Technical Skills

Programming: Python, C/C++, Matlab, R, JavaScript, HTML, CSS
Tools: PyTorch, Opencv, LaTex, the scientific Python stack (numpy/scipy, scikit-learn, pandas, etc.), NodeJS, Bootstrap, Gurobi, Sketch, Bash, Git
Platform: Linux, Mac OSX, IOS, Amazon Web Services (AWS)
Language: IELTS:7.5; GRE: 323+4.0

Service

Reviewer for ECCV2022, CVPR2023, ICCV2023, CVPR2024, ECCV2024, NeurIPS2024, ICML2025, 3rd Workshop on Advances in Language and Vision Research in ACL2024, TPAMI.