Lin Li (李淋)

PhD student in Machine Learning, Department of Informatics, King's College London, London, UK
Phone: +86 189 0022 0595 | Email: lin.3.li@kcl.ac.uk | Page: treelli.github.io | Git: github.com/TreeLLi

RESEARCH INTEREST

- Trustworthy ML: robustness, AI safety, AI Alignment
- Data-centric ML: generative models for data augmentation, large-scale data profiling, data pruning
- AI + applications: healthcare, robotics

EDUCATION

M.Phil/PhD, Department of Informatics, King's College London, London, UK Oct. 2019 – Aug. 2024

- Supervisors: Dr. Michael Spratling (primary) and Dr. Dimitrios Letsios
- Thesis: *Towards Robust Visual Classification through Adversarial Training*
- Thesis examiners: Prof. George D. Magoulas (Birkbeck, University of London) and Dr. Adel Bibi (Oxford)

MSc, Department of Computing, Imperial College London, London, UK Oct. 2017 – Sep. 2018

- Advisor: Prof. Wayne Luk
- Grade: Overall Distinction (Exam + Thesis)
- Thesis: *Understanding Deep CNNs via Interpretable Individual Units*

BBM, Department of Finance, Xiamen University, Xiamen, China Sep. 2013 – June 2017

- Advisor: Prof. Zheng Qiao
- Thesis: *Quantitatively Measuring Investor's Sentiment via Search Index*

PROFESSIONAL EXPERIENCE

Associate Member, Sea AI Lab, Singapore Dec. 2023 – present

- advised by: Dr. Tianyu Pang and Dr. Chao Du
- project: LLM Safety

Research Intern, Robotics X Lab, Tencent, Shenzhen, China Dec. 2021 – Oct. 2022

- advised by: Dr. Lipeng Chen
- project: Advancing Robots with Greater Dynamic Dexterity: A Large-Scale Multi-View and Multi-Modal Dataset of Human-Human Throw&Catch of Arbitrary Objects

Teaching Assistant, Department of Informatics, King's College London, London, UK Jan. 2021 – Dec. 2021

- courses: Machine Learning and Pattern Recognition, Introduction to Artificial Intelligence

PUBLICATIONS

1. **Lin Li**, Yifei Wang, Chawin Sitawarin, Michael Spratling, OODRobustBench: a benchmarking and large-scale analysis of adversarial robustness under distribution shift, International Conference on Machine Learning (**ICML**) 2024 and ICLR workshop Data-centric Machine Learning Research, 2024
2. **Lin Li***, Haoyan Guan*, Jianing Qiu, Michael Spratling, One Prompt Word is Enough to Boost Adversarial Robustness for Pre-trained Vision-Language Models, IEEE/CVF Computer Vision and Pattern Recognition Conference (**CVPR**), 2024
3. **Lin Li**, Michael Spratling, Data augmentation alone can improve adversarial training, International Conference on Learning Representations (**ICLR**), 2023
4. **Lin Li**, Michael Spratling, Understanding and combating robust overfitting via input loss landscape analysis and regularization, Pattern Recognition (**PR**), 2023
5. Jianing Qiu, **Lin Li**, Jiankai Sun, and Jiachuan Peng, Peilun Shi, Ruiyang Zhang, Yinzhao Dong, Kyle Lam, Frank P.-W. Lo, Bo Xiao, Wu Yuan, Dong Xu, Benny Lo, Large AI Models in Health Informatics: Applications, Challenges, and the Future, IEEE Journal of Biomedical and Health Informatics (**JBHI**), 2023
6. **Lin Li**, Michael Spratling, Improved Adversarial Training Through Adaptive Instance-wise Loss Smoothing, in submission to IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), 2023
7. **Lin Li**, Jianing Qiu, Michael Spratling, AROID: Improving Adversarial Robustness through Online Instance-wise

Data Augmentation, in submission (major revision) to the International Journal of Computer Vision (IJCV), 2023

8. Jianing Qiu, Jian Wu, Hao Wei, and Peilun Shi, Minqing Zhang, Yunyun Sun, **Lin Li**, Hanruo Liu, Hongyi Liu, Simeng Hou, Yuyang Zhao, Xuehui Shi, Junfang Xian, Xiaoxia Qu, Sirui Zhu, Lijie Pan, Xiaoniao Chen, Xiaojia Zhang, Shuai Jiang, Kebing Wang, Chenlong Yang, Mingqiang Chen, Sujie Fan, Jianhua Hu, Aiguo Lv, Hui Miao, Li Guo, Shujun Zhang, Cheng Pei, Xiaojuan Fan, Jianqin Lei, Ting Wei, Junguo Duan, Chun Liu, Xiaobo Xia, Siqi Xiong, Junhong Li, Benny Lo, Yih Chung Tham, Tien Yin Wong, Ningli Wang, Wu Yuan, VisionFM: a Multi-Modal Multi-Task Vision Foundation Model for Generalist Ophthalmic Artificial Intelligence, in submission (major revision) to the New England Journal of Medicine Artificial Intelligence (NEJM AI), 2023

9. Lipeng Chen*, Jianing Qiu*, **Lin Li***, Xi Luo, Guoyi Chi, Yu Zheng, Advancing Robots with Greater Dynamic Dexterity: A Large-Scale Multi-View and Multi-Modal Dataset of Human-Human Throw&Catch of Arbitrary Objects, in submission (major revision) to the International Journal of Robotics Research (IJRR), 2023

PROJECTS

Detecting objects for hotel rooms, Microsoft, London, UK 2018

- co-supervised by Dr. Anandha Gopalan, Mr. Lee Stott
- Blog (Microsoft), Git, Report, Presentation, Opensource Contributions, Demo

HONORS & AWARDS & GRANT

PGR Research Support, King's College London 2023
King's-China Scholarship, King's College London and China Scholarship Council (CSC) 2019
1st Class (Xiangyu) University Scholarship, Xiamen University 2016
Excellent Academic Performance Scholarship, Xiamen University 2015
3rd prize winner, Jinyuan Creativity and Startup Contest, Xiamen City 2016
3rd prize winner, ChinaNet Dream Accelerator Programming Contest, China 2015

TALKS & PRESENTATIONS

Prompting VLMs for adversarial robustness, AI Time and VALSE 2024
Data augmentation can improve adversarial training, AI Time Youth PhD Talk 2023
Data augmentation for adversarial robustness, ADA talk, King's College London 2023
Defending DNNs against adversarial examples, Departmental Research Showcase, King's College London 2023

ACADEMIC SERVICE

- **Reviewer (conference)**: NeurIPS, ICML, ICLR
- **reviewer (journal)**: IEEE T-Dependable and Secure Computing, IEEE Journal of Biomedical and Health Informatics
- **program committee**: ICRA 2024 workshop WIHR

SKILLS

- **Programming languages**: Python, C++, C, Java, Objective-C
- **Machine learning**: CNN, ViT, adversarial attack and defense, AutoML, diffusion models, large multimodal models
- **Machine learning frameworks**: PyTorch, TensorFlow

REFEREES

Dr. Michael Spratling
Reader, Department of Informatics, King's College London, London, UK
Phone: +44 020 7848 2027, Email: michael.spratling@kcl.ac.uk

Dr. Benny Lo
Reader, the Hamlyn Centre & the Department of Surgery and Cancer, Imperial College London, London, UK
Phone: +44 (0)20 7594 0806, Email: benny.lo@imperial.ac.uk

Dr. Lipeng Chen
Senior Research Scientist, Robotics X Lab, Tencent, Shenzhen, China
Phone: +86 18267157219, Email: lipengchen@tencent.com