Yifan Zhang

Phone: +1 919 638-0501 | Email: yifan.zhang.2@vanderbilt.edu | Homepage | LinkedIn | Twitter

Education

Vanderbilt University — *Nashville, TN, USA*
Ph.D. in Computer Science, Specialize in AI for Software Engineering — *Jun. 2022 – Apr. 2026 (Expected)*
- Advisors: Dr. Kevin J. Leach & Dr. Yu Huang

Georgia Institute of Technology — *Atlanta, GA, USA*
M.Sc. in Computer Science, Specialize in Computing Systems — *Aug. 2022 – Jun. 2025*
- Studied Machine Learning & Software Engineering

China University of Petroleum — *Beijing, CN*
M.Eng. in Petroleum Engineering, Specialize in Industrial & System Engineering — *Sep. 2012 – Jun. 2019*
- Double majored in English (TEM-8 holder) & Minor in British Parliamentary Debate

Publication

Book Chapters

- **Yifan Zhang**: Federated Feature Engineering in *Federated Learning: Technology and Practice*. Electronic Industry Press, 2021.

Journals and Conferences

- Yichang He, **Yifan Zhang**, Yunpeng Fan, U-Xuan Tan: *Real-time Vibration Compensation with Long Short-term Memory Recurrent Neural Network and Adaptive Filter*. Under Review by IEEE Transaction on Mechatronics (**IEEE-TMech**), 2023.
- **Yifan Zhang**, Chen Huang, Yueke Zhang, Kevin Cao, Scott Thomas Anderson, Huajie (Jay) Shao, Kevin Leach, Yu Huang: *COMBO: Pre-Training Representations of Binary Code Using Contrastive Learning*. In Preparation for European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE**) [arXiv]
- **Yifan Zhang**, Junwen Yang, Haoyu Dong, Qingchen Wang, Huajie (Jay) Shao, Kevin Leach, Yu Huang: *ASTRO: An AST-Assisted Approach for Generalizable Neural Clone Detection*. Under Review by International Conference on Software Engineering Workshops (**ICSEW**), 2023. [arXiv]
- **Yifan Zhang**: *Leveraging Artificial Intelligence on Binary Code Comprehension*. Automated Software Engineering Doctoral Symposium (**ASE-DS**), 2022. [Paper] [arXiv]
- **Yifan Zhang***, Haoyu Dong*, Nicholas Konz, Hanxue Gu, Maciej Mazurowski: *Lightweight Transformer Backbone for Medical Object Detection*. International Conference on Medical Image Computing and Computer Assisted Intervention Workshops (**MICCAIW**), 2022. [Paper] [arXiv]
- Haoyu Dong, **Yifan Zhang**, Hanxue Gu, Nicholas Konz, Maciej Mazurowski: *SwinSSL: Sliding-Window based Self-Supervised Learning Framework for Anomaly Detection*. Revision from IEEE Transaction on Medical Imaging (**IEEE-TMI**), 2023.
- Haoyu Dong, **Yifan Zhang**, Nicholas Konz, Hanxue Gu, Maciej Mazurowski: *Localized Semi-Supervised Anomaly Detection for Breast Tomosynthesis Lesion Screening*. Under Review by Medical Image Analysis (**MIA**), 2023.
- Chen Huang, **Yifan Zhang**, Kevin Leach, Yu Huang, Wenqiang Lei, Jiancheng Lv: *Cross-Domain Filters for Graph Convolutional Networks by Multiple Kernel Learning*. Under Review by International Joint Conference on Artificial Intelligence (**IJCAI**), 2023.
- Jing Li, Xiangfang Li, Keliu Wu, Dong Feng, Tao Zhang, **Yifan Zhang**: *Thickness and Stability of Water Film Confined inside Nanoslits and Nanocapillaries of Shale and Clay*. International Journal of Coal Geology (**IJCG**), 2017. [Paper]

Preprints

- **Yifan Zhang**, Haoyu Dong, Nicholas Konz, Hanxue Gu, Maciej Mazurowski: *REPLICA: Enhanced Feature Pyramid Network by Local Image Translation and Conjunct Attention for High-Resolution Breast Tumor Detection.* (Preprint) [arXiv]

Academic Experience

Vanderbilt University *Apr. 2022 – Present*
Graduate Research Assistant, Advisors: Dr. Kevin J. Leach & Dr. Yu Huang *Nashville, TN, USA*

- **Binary Code Representation**: Built a contrastive learning method to facilitates representation learning of binary code visualized by distribution analysis, and improves the performance on all three downstream tasks of software engineering and security by 5.45% on average compared to state-of-the-art large-scale language representation models.
- **Domain-Guided Clone Detection**: Designed an AST-based representation for source code that leverages program structure and semantics to improve state-of-the-art code clone detection. Our experimental results show that the model improves state-of-the-art neural clone detection approaches by up to 6.6% in F-1 score.

Duke University *Jul. 2021 – Jun. 2022*
Research Associate, Advisor: Dr. Maciej A. Mazurowski *Durham, NC, USA*

- **Medical Object Detection**: Designed a feature interpolation pipeline for injecting tumors into healthy images as an augmented dataset, and conjuncted a ViT on the outputs of a ResNet as inputs to a FPN in Faster R-CNN for tumor detection. The model mitigates the data-hungry problem of attention and achieves 13.1% improvement in AP50 for detecting tumors.
- **Domain Generalization**: Introduced a method for detecting anomalies in high-resolution medical images by sliding patches, and a domain generalization method by imposing constraints on the feature space and its projection space. Both of the two model achieve state-of-the-art in anomaly detection and domain generalization accuracy.

The University of Hong Kong *Mar. 2021 – Sep. 2021*
Senior Research Assistant, Advisor: Dr. Qingchen Wang *Hong Kong SAR*

- **Financial Decision Making**: Built an entire intelligent debt collection system using data-driven deep reinforcement learning models. The model utilizes transformer as the feature extractor and attaches a offline policy gradient model trained on the embedded sequential-aware hidden features to propose long-term dependent decisions.

Industry Experience

JD.COM *Dec. 2018 – Mar. 2021*
Data Scientist, Supervisors: Dr. Hu Wang & Mr. Chen Huang *Beijing, CN*

- **Action Model**: Built Bi-GRU and DeepFM models on user behavior features to predict the credit use rate and overall profit of every user in cash loan and consumer debt. The model can propose decisions to increase their credit limit for maximizing income, and achieved 21.4% overall profit increase.
- **Credit Score Propagation**: Built a heterogeneous graph on different types of user connections, and applied GraphSage and GCN models to propagate the credit score and improve risk prediction. The model can improve the overall accuracy of the XGB model by 5% in user classification.
- **Privacy-Preserving Collaboration**: Invented one kind of GAN-styled model using differential privacy to improve the efficiency and security of federated learning. Applied for 10 CN patents based on the research outputs, and was listed as 1st or 2nd inventor in 8 of them. One of the patents was awarded as 1st Runner-up in the 3rd JD Discovery Cup Patent Competition.

Academic Services

Program Committee (PC) Member

- 2023 AAAI Workshop on DL-Hardware Co-Design for AI Acceleration (AAAI-DCAA)
- 2023 International Conference on Mining Software Repositories (MSR)
- 2023 AAAI Conference on Artificial Intelligence (AAAI)
- 2022 MICCAI Workshop on Cancer Prevention through early detecTion (MICCAI-CaPTion)

External Conference Reviewer

- 2023 USENIX Security Symposium (USENIX)
- 2022 IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR)

Funds and Awards

Fellowships

- 2022 Research Fellowship from Defense Advanced Research Projects Agency (DARPA) (32500$/year)
- 2021 Research Fellowship from National Institutes of Health (NIH) ($36000/year)
- 2017 Roberto Roca Education Fellowship (Top 10 nationwide)
- 2015 Chinese National Fellowship for Overseas Studies (CA$6000/4 months)
- 2015 & 2016 & 2017 & 2018 First-class Scholarship at China University of Petroleum, Beijing (CUPB)
- 2014 Schlumberger Engineering Fellowship (Top 8 university-wide)
- 2013 Chinese National Scholarship for Outstanding Merits (Top 0.2% nationwide)

Honors and Awards

- 2020 First Runner-up in the 3rd JD.COM Discovery Cup Patent Competition (Top %0.1 company-wide)
- 2020 Silver Medal Award of Distinguished Technical Recruiter at JD.COM (Top %5 company-wide)
- 2020 Bronze Medal Award of Certified Technical Instructor at JD.COM
- 2019 Beijing Outstanding Graduate Award (Top %0.1 nationwide)
- 2015 & 2017 Meritorious Winner of American Mathematical Contest in Modeling (Top %5 worldwide)
- 2015 Dean's List for the Best Undergraduate Students at CUPB (Top %1 university-wide)
- 2013 & 2015 Third Prizes of Chinese National Petroleum Engineering Design Competition (Top 5% nationwide)