WENXIN JIANG

github.com/wenxin-jiang wenxin-jiang.github.io jiang784@purdue.edu
West Lafayette, Indiana (765)-409-1715 wenxin-jiang

My research interest lie in *Software Engineering* and *Trustworthy AI*, especially the reusability and reliability of deep learning pre-trained models. I am also interested in studying the reproducibility of computer vision models.

EDUCATION

Ph.D., Electrical and Computer Engineering, Purdue University, IN, USA *Aug 2020 - present*
GPA: **3.7/4.0**, Supervised by Prof. **James C. Davis**

B.S., Applied Physics, Southeast University, Jiangsu, China *Aug 2016 - Jun 2020*
- **Study Abroad Program**, *GPA:* **3.8/4.0**, Engineering Physics, UC Santa Barbara, CA, USA *Mar 2019 - Jun 2019*

RESEARCH & WORK EXPERIENCE

Software engineering support for pre-trained deep learning models, *Research Assistant* *Feb 2022 - present*
-**Lead a team** of students, conducted an empirical study on pre-trained deep learning model supply chain.
- Measure the trustworthiness of deep learning model registries through **model audit**, **risk analysis**, and **interviews**.

TensorFlow Model Garden Team (Google x Purdue), *ML engineer*, *team leader* *Sep 2021 - present*
- **Lead a team** of students reproducing a recent computer vision model, *i.e.,* Yolox.
- Contribute to development and testing of model architecture and component integration.

A Machine Shop Dataset for Computer Vision, *Research Assistant* *May 2022 - Aug 2022*
-**Lead a team** of students, **collected video data** of safe and unsafe behavior for low power computer vision models.

Empirical Study on Computer Vision Reengineering, *Research Assistant* *Jan 2021 - Mar 2022*
- **Lead a team** of students, collect open-source bug reports, and analyze the dataset.
- Conduct a **case study** on a reengineering team cooperating with Google, and summarize the practices and challenges.

PUBLICATIONS

1. **Jiang**, Synovic, Sethi, Indarapu, Hyatt, Schorlemmer, Thiruvathukal, and Davis. *An Empirical Study of Artifacts and Security Risks in the Pre-trained Model Supply Chain*. Proceedings of the 1st ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (**SCORED'22**).

2. Montes, Peerapatanapokin, Schultz, Guo, **Jiang**, and Davis. *Discrepancies among Pre-trained Deep Neural Networks: A New Threat to Model Zoo Reliability*. Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering — Ideas, Visions, and Reflections track (**ESEC/FSE-IVR'22**).

3. Synovic, Hyatt, Sethi, Thota, Shilpika, Miller, **Jiang**, Amobi, Pinderski, Laufer, Hayward, Kingensmith, Davis, and Thiruvathukal. *Snapshot Metrics Are Not Enough: Analyzing Software Repositories with Longitudinal Metrics*. Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering — Demonstrations track (**ASE-Tool Demonstrations'22**).

4. Veselsky, West, Ahlgren, Thiruvathukal, Klingensmith, Goel, **Jiang**, Davis, Lee, and Kim. *Establishing trust in vehicle-to-vehicle coordination: a sensor fusion approach*. roceedings of the 2nd Workshop on Data-Driven and Intelligent Cyber-Physical Systems for Smart Cities (**DI-CPS'22**).

5. Banna, Chinnakotla, Yan, Vegesana, Vivek, Krishnappa, **Jiang**, Lu, Thiruvathukal, and Davis. *An Experience Report on Machine Learning Reproducibility: Guidance for Practitioners and TensorFlow Model Garden Contributors*. arXiv. 2021.

SKILLS AND INTERESTS

Technical Skills: Python (*Advanced*), Java (*Intermediate*), C/C++ (*Basic*), Unix, Git
Libraries: TensorFlow, Pytorch, Numpy, Pandas, MatPlotlib, OpenCV
Languages: Mandarin (Native), English
Personal Interests: Photography, Cooking, Guitar