

# JIAWEI LIU

jiawei6@illinois.edu ♦ 2nd-year Ph.D. Student ♦ [GitHub](#) ♦ [HomePage](#)

## RESEARCH INTEREST

---

I am interested in *programming systems* and *deep learning*. My research goal is to improve the *programmability*, *performance* and *correctness* of existing systems. Recently I am working on test-case synthesis for DL compiler.

## EDUCATION

---

**University of Illinois at Urbana-Champaign**, IL, US  
Ph.D. in Computer Science; GPA: 4.0/4.0

Aug. 2021 - Present  
Advisor: [Lingming Zhang](#)

**Tongji University**, Shanghai, China  
B.Eng. in Computer Science

Sept. 2017 - Jul. 2021

## PAPERS AND TALKS

---

- [1] NEURI: Diversifying DNN Generation via Inductive Rule Inference.  
[Jiawei Liu](#), Jinjun Peng, Yuyao Wang, Lingming Zhang.  
[pre-print](#)
- [2] **[ASPLOS'23]** NNSMITH: Generating Diverse and Valid Test Cases for Deep Learning Compilers  
[Jiawei Liu](#)<sup>\*</sup>, Jinkun Lin<sup>\*</sup>, Fabian Ruffy, Cheng Tan, Jinyang Li, Aurojit Panda, Lingming Zhang.  
[code](#) ♦ [artifact](#) ♦ [paper](#) ♦ [pre-print](#)
  - Systems & Software Engineering Seminar, UIUC Mar. 2023
- [3] **[OOPSLA'22]** Coverage-Guided Tensor Compiler Fuzzing with Joint IR-Pass Mutation  
[Jiawei Liu](#), Yuxiang Wei, Sen Yang, Yinlin Deng, Lingming Zhang.  
[code](#) ♦ [artifact](#) ♦ [paper](#)
  - SPLASH/OOPSLA Conference Talk Nov 2022
  - SAMPL Lunch Talks, University of Washington May 2022
  - Software Engineering Seminar, UIUC Apr. 2022
- [4] **[ACMMM'21 OSC]** Fast and Flexible Human Pose Estimation with HyperPose  
Yixiao Guo<sup>\*</sup>, [Jiawei Liu](#)<sup>\*</sup>, Guo Li<sup>\*</sup>, Luo Mai, Hao Dong.  
[code](#) ♦ [paper](#)

<sup>\*</sup> *Co-primary.*

## PROFESSIONAL EXPERIENCES

---

- Intern at OctoML** May. - Aug. 2022  
Advised by: [Yuchen Jin](#), [Sunghyun Park](#), [Tianqi Chen](#)  
Designed and implemented a declarative pattern language for simplifying graph optimizations in TVM's Relax IR (PRs: [A](#), [B](#)).  
*Pattern Language*
- Research Assistant at UIUC** Apr. 2021 - Present  
Advised by: [Lingming Zhang](#)  
Doing research in applying PL and formal methods to synthesize Deep-Learning programs for fuzzing and beyond [1]–[3].  
*Fuzzing*
- Intern at DAMO Academy, Alibaba Cloud** Mar. - Aug. 2021  
Advised by: [Yuanwei Fang](#), [Yuan Xie](#) (Univ. of California at Santa Barbara)  
Implemented multi-core graph sampling algorithms in C++ for GNN pre-processing, improving prior implementation by over 20×.  
*GNN Serving*
- Research Assistant at NYU Systems Group** Jul. 2020 - Mar. 2021  
Advised by: [Jinyang Li](#)  
Designed a programming model, *Compare-And-Skip*, for doing video analytic with programmable accuracy-efficiency trade-off.  
*Video Analytics*
- Research Assistant at Peking University** Jan. - Aug. 2020  
Advised by: [Hao Dong](#), [Luo Mai](#) (Univ. of Edinburgh)  
Implemented a CPU-GPU scheduler for pose estimation and post-processing algorithms, with up-to 7.3× speedup over SOTA [4].  
*Optimizing Pose Estim.*
- Intern at ByteDance AI Lab** Feb. - Jul. 2020  
Advised by: [Guanzhe Huang](#), [Chuanxiong Guo](#)  
Developed a model server with dynamic batching and a monitoring sub-system for debugging health/performance of model services.  
*DNN Model Serving*

## NOTABLE ACHIEVEMENTS

---

Conference Presentation Award, UIUC	2023
ASPLOS Student Travel Award	2023
Qidi Innovation Scholarship of Tongji University (Top 1%)	2020
Selected Entrant for 2020 Google Machine Learning Winter Camp	2019
Winner of International Data Science Hackathon (Chinese Region), Covestro [news]	2019
National 2nd Prize and Province-level 1st Prize in RoboMaster, DJI Inc.	2019
National 2nd Prize (0.75~3.84%) and Province-level 1st Prize of Chinese Modeling Contest	2018

## OPEN-SOURCE CONTRIBUTIONS

---

I embrace and grow with the ❤ open-source community.

I lead the **NNSmith** project, a random DNN synthesizer for testing Deep-Learning (DL) systems. I also have been actively contributed to **TVM** for the **dataflow pattern language** of *Relax* IR to simplify graph optimization, as well as bug fixes. Previously, I co-developed **HyperPose** (**1200+ stars**) for efficient pose estimation.

My research facilitates the correctness of DL systems. Since 2021, we found over 210 new bugs for TVM, PyTorch, TensorFlow, TensorRT, and ONNXRuntime. My work have impacted the testing tooling of real-world systems including nvFuser, TensorRT (privately acknowledged by NVIDIA developers) and **TVM Relay**.

## PROFESSIONAL SERVICES

---

Artifact Evaluation Committee:	PLDI'23, OSDI'22, ATC'22.
Reviewer:	AAAI'23@DCAA

## SKILL STACK

---

My general skill set covers program optimization/analysis, fuzzing, visualization and deep learning. Oftentimes, I code with modern C++ for efficiency with a Python front-end for interoperability.

**General:** C++, Python, LLVM, Docker, Git, Grafana, GDB, etc.

**Machine Learning & Systems:** PyTorch, TensorRT, TensorFlow, TVM, ONNX, Pybind11, etc.

**Correctness:** Z3, libFuzzer, Sanitizers, Dafny, Spin, etc.

**Parallel:** C++ Thread Library, CUDA, gRPC, etc.