

# Yifeng He

---

## CONTACT INFORMATION

**Cell:** (530) 302-6806 | **E-mail:** yfhe@ucdavis.edu | **WWW:** <https://eyh0602.github.io/>

## RESEARCH INTERESTS

Program Analysis (Compilers, Type Theory, and Fuzzing), Code Understanding via Machine Learning & NLP, Programming Languages for Machine Learning & AI, Formal Verifiable Programs, Computer Security.

## EDUCATION

**University of California, Davis**, Davis, California, USA

08/2019 – present

- **Anticipated degree:** B.S. in Computer Science & Applied Mathematics
- **GPA:** 3.879

## RESEARCH

**Computer Security Lab at UC Davis**

10/2022 - present

- Work with one PhD candidate, researchers from Tencent AI Lab, and Prof. Hao Chen on the frontend of programming language embedding.
- Generate fuzzing data (program-level IO pairs) for fine-tuning (Fuzz-tuning) LLM and achieved SOTA on downstream tasks.
- Generate fuzzing data (program-level IO pairs) for pre-training (Fuzz-pretrain) LLM.
- Proposed new tasks for evaluating testcase generation by LLM.
- Instrument open-source projects to generate function-level fuzzing IO-pairs via LLVM Pass.

**Path Academics**

02/2022 - 07/2022

- Conducted research on neural network and its application in AI under the supervision of Prof. Pavlos Protopapas from Harvard
- Attended workshops on gradient descent algorithm, neural network optimizers, regularization of neural network, and other related concepts and architecture
- Analyzed and compared models of deep learning algorithms application,
- Made automatic differentiation to activation functions by hand, visualized receptive fields through max-pooling

## INTERNSHIP

**ByteDance**

04/2021 – 08/2021

*Software Engineering Intern*, Income Platform Team

- Used microservice tech to connect parts of the author income settlement business
- Transformed author-relation data architecture design from relational database (SQL) to graph database (Gremlin) to allow better efficiency for the business model
- Refactored the income calculation control process with visitor design pattern using Python 3

*Software Engineering Intern*, **Xigua Video** Author Experience Team

- Created a data cleaner script with ORM to maintain the size and readability of online data settlement table
- Created the offline flow of Medium Video Encouragement Project for weekly data calculation
- Built the interface for frontend of web and mobile app to display the data visualization

PROJECTS	<b>CourseReco</b>	06/2022 – 09/2022
	<ul style="list-style-type: none"> <li>• Designed the overall system architecture</li> <li>• Led the programming for API server and recommender engine</li> <li>• Negotiated with the third-party provider, SchedGo, for data service</li> <li>• Provided technical leadership to teammates</li> </ul>	
	<b>Music Genre Classifier</b>	05/2022 – 06/2022
	<ul style="list-style-type: none"> <li>• Processed music samples into spectrogram by Short-time Fourier transform</li> <li>• Designed the appropriate model (CNN) to classify spectrograms into category</li> <li>• Analyzed the resulting model and test outputs with saliency maps</li> </ul>	
	<b>ImageOrientation</b>	03/2022 - 04/2022
	<ul style="list-style-type: none"> <li>• Pre-processed image data by rotating them with random generated angles, and assigned these angles as label</li> <li>• Designed the appropriate CNN for regression task, tested and improved the model</li> <li>• Applied Hyper-parameter tuning based on train, validation, and tested results</li> </ul>	
	<b>Dcash-server</b>	05/2021 - 07/2021
	<ul style="list-style-type: none"> <li>• Created a multi-threaded API server using C++ to allow users to create accounts to make deposit and transfer</li> <li>• Used MySQL to store and maintain user data</li> <li>• Made API calls to the Stripe API server to handle credit card information</li> </ul>	
	<b>Genshine Impact Gacha Analyzer</b>	08/2021 - 09/2021
	<ul style="list-style-type: none"> <li>• Designed fetching process of gacha data from MiHoYo and categorized the process</li> <li>• Stored data into local database automatically, wrote into excel for data analysis by option</li> <li>• Generated text or graph visualization report from data analyze results</li> </ul>	
ACTIVITIES	<b>HackerHub Club</b> , UC Davis	07/2020 - present
	<i>Co-founder, President, Technical Officer</i> <ul style="list-style-type: none"> <li>• Design and maintain a course recommendation system, CourseReco, for UC Davis students</li> <li>• Organize and lead the Code Jam Competition on data visualization, AI, augmented reality and virtual reality, and machine learning</li> <li>• Coach in introductory programming workshops, including Assembly, functional programming, recommender system, generative adversarial network, etc.</li> </ul>	
HONORS AND AWARDS	<b>Dean's Honor List</b> , Fall 2019, Spring 2020, Spring 2021, Spring 2022	
PUBLICATIONS	Zhao, J., Rong, Y., Guo, Y., <b>He, Y.</b> , Chen, H. <i>Understanding Programs by Exploiting (Fuzzing) Test Cases</i> , Findings of Association for Computational Linguistics (ACL), 2023.	
	<b>He, Y</b> , <i>Big Data and Deep Learning Techniques Applied in Intelligent Recommender Systems</i> , IEEE 4th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), 2022.	