

Zhongzhu Zhou / Charlie Zhou

February, 2023

(+86) 13508482354 (+61) 0414390906

zhongzhu.zhou@sydney.edu.au zhouzhzh8@mail2.sysu.edu.cn

<https://www.linkedin.com/in/zhongzhu-zhou-570890bb/>

<https://github.com/zzz0906>

<https://scholar.google.com.hk/citations?user=BoZKZl4AAAAAJ&hl=zh-CN>

No.132 Waihuan East Road, Panyu District, Guangzhou, 510006, Guangdong, China

1D Greenbank St, Hurstville, Sydney, 2220, NSW, Australia

Education

- **The University of Sydney (USYD)** Sydney, Australia
Doctor of Philosophy (Ph.D.) Oct. 2022 - Present
 - Accumulated GPA: **4.0/4.0**
 - **Honor and Awards:**
 - The Jingdong Technology (JD) Co Ltd Research Scholarship in Artificial intelligence, USYD, 2022
- **Sun Yat-sen University (SYSU)** Guangzhou, China
Research Assistant Sep. 2019 - Present
 - Accumulated GPA: **3.41/4.0**
 - **Honor and Awards:**
 - SYSU Overseas Visiting and Collaborative Research Program Funding Plan, SYSU, 2021
 - The Third Class Scholarship, SYSU, 2020, 2021, 2022
 - The Second Class Scholarship (Top 15% of the major), SYSU, 2019
- **University of Illinois Urbana-Champaign (UIUC)** Remotely & Champaign, IL, United States
Summer Seission Student Jun. 2018 - Sep. 2018
 - **Honor and Awards:**
 - Illinois Computer Science Summer Research Program, UIUC, 2018
- **Sun Yat-sen University (SYSU)** Guangzhou, China
Bachelor of Engineering in Computer Science and Technology Sep. 2015 - Jun. 2019
 - Overall GPA: **3.9/4.0**
 - **Honor and Awards:**
 - National Scholarship (Top 1 of the major), China, 2016
 - Research Honor Degree, SYSU, 2019
 - The First Class Scholarship (Top 5% of the major), SYSU, 2015-2016, 2017-2018
 - The Second Class Scholarship (Top 15% of the major), SYSU, 2016-2017
 - Meritorious Winner, COMAP's Mathematical Contest in Modeling, United States, 2017
 - The Second Prize, The Chinese Mathematics Competitions, 2016
 - The Third Prize, The Chinese Mathematics Competitions, 2017
 - The Third Prize, ACM-ICPC, SYSU, 2017
 - The Second Prize, Student Innovation Software Development Competition, SYSU, 2017
 - The Third Prize, Microsoft Hackthon, South China, 2018
- **Changjun High School** Changsha, China
Senior Student Sep. 2012 - Jun. 2015
 - **Honor and Awards:**
 - The Second Prize, National Olympiad in Informatics in Provinces

Professional Service

- Institute of Electrical and Electronics Engineers (IEEE) Member ID: 97841404
- Association for Computing Machinery (ACM) Member ID: 6708618
- China Computer Federation (CCF) Member ID: B8293G
- Reviewer for Conferences
 - NeurIPS 2023(Thirty-seventh Conference on Neural Information Processing Systems)

Publications

Book

- *C Language Programming in Chinese*
Xuemao Zhou, Wei Yi, Zhongzhu Zhou
 - **Tianjin University Press**
 - ISBN: 9787561847251
 - [SALE LINK](#)

Conference Paper

- ***JSIdentify: A Hybrid Framework for Detecting Plagiarism Among JavaScript Code in Online Mini Games***
Qun Xia, **Zhongzhu Zhou**, Zhihao Li, Bin Xu, Wei Zou, Zishun Chen, Huafeng Ma, Gangqiang Liang, Haochuan Lu, Shiyu Guo, Ting Xiong, Yuetang Deng, Tao Xie
 - **ICSE (International Conference on Software Engineering) 2020**
 - Track: *Software Engineering in Practice*
 - [TALK LINK](#), [PAPER LINK](#)

Journal Paper

- ***Binary Neural Network for Automated Visual Surface Defect Detection***
Wenzhe Liu, Jiehua Zhang, Zhou Su, **Zhongzhu Zhou**, Li Liu
 - **Sensors MDPI (Multidisciplinary Digital Publishing Institute)**
 - Track: *Special Issue Intelligent Sensing and Monitoring for Industrial Process*
 - [PAPER LINK](#)

Patent

- ***KUBERNETES 用户态应用中基于虚拟文件系统的小文件存储优化系统***
Liang Du, Guixin Guo, Kangyou Zhong, Yunfei Du, Yutong Lu, **Zhongzhu Zhou**
 - **Chinese Patent**
 - 申请号: CN202010195318.5, 公开号: CN111475469A/B
 - [CHINESE DOCUMENT LINK 1](#), [CHINESE DOCUMENT LINK 2](#)

Talks

- ***JSIdentify: A Hybrid Framework for Detecting Plagiarism Among JavaScript Code in Online Mini Games***
 - [ICSE \(International Conference on Software Engineering\)](#), Jul, 11, 2020

Conference Participation

- ***ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2023)*** Online Attendance; March 25 - 29, 2023
- ***International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2022)*** Online Attendance; November 13 - 18, 2021
- ***China National Computer Congress (CNCC 2021)*** Online Attendance; December 16 - 18, 2021
- ***International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2021)*** Online Attendance; November 14 - 19, 2021
- ***ACM International Conference on Supercomputing (ISC 2021)*** Online Attendance; June 24 - July 02, 2021
- ***International Symposium on Computer Architecture (ISCA 2021)*** Online Attendance; June 14 - 19, 2021
- ***International Conference on High Performance Big Data and Intelligent Systems (HPBC&IS 2020)*** Online Attendance; May 23 - 23, 2020

Teaching

- **COMP3520: Operating Systems Internals**
Fall 2023, Tutor, The University of Sydney

Professional Experience & Industry Research Projects

- **Future System Architecture (FSA) Lab, The University of Sydney (USYD)** Mar. 2022 - Present
Visiting Scholar, Ph.D. student *Sydney Australia*
Advisor: *Shuaiwen Song* (Associate Professor, USYD), *Chang Xu* (Associate Professor, USYD), *Yibo Yang* (Research Scientist in JD Explore Academy)
Research Projects:

– **RenAIssance: A survey into AI text to image generation in the era of large models**

Motivation: Text-to-image synthesis has become increasingly popular in the AI and computer graphics world (AIGC). However, there is no comprehensive survey paper that systematically introduces the frameworks and ideas behind text-to-image techniques. We aim to fill this gap in the literature.

Contributions:

1. Read over 100 papers, providing a literature review for each.
2. Collaborated with lab classmates to write the comprehensive survey paper.

– **Optimization of Diffusion Model Denoising Process**

Motivation: Diffusion models currently require a large number of denoising steps, which we aim to reduce. One reason for the lengthy process is the lack of a clear relationship between the noise and the trained image. Our goal is to explore additional methods to establish a connection between noise and the denoised image, beyond guidance techniques, such as incorporating text embeddings into the raw noise.

Contributions:

1. Develop innovative ideas, implement them, and conduct comparative experiments to evaluate their performance.

– **Exploring Neural Collapse Phenomenon in Reinforcement Learning**

[CODE LINK](#)

Motivation: In reinforcement learning, agents may exhibit biased action selection in the environment due to incomplete understanding of the state and action distribution spaces. This research investigates whether the neural collapse phenomenon occurs in policy gradient networks as agents train with sufficient examples and examines its implications for balancing action selection in reinforcement learning agents.

Contributions:

1. Conducted experiments applying ETF classifiers to 5+ neural networks in 10+ discrete-action reinforcement learning environments (e.g., Atari, Gym Classic)
2. Derived and proved the formula and geometric properties of policy gradient loss function
3. Authored paper drafts and submitted to the NeurIPS conference

– **Sparse Kernel Design in GPU TensorCore**

[CODE LINK](#)

Motivation: With the application of pruning methods, neural network weight matrices become increasingly sparse, but there is no implementation for sparse kernels in GPU TensorCore.

Contributions:

1. Conducted comparative experiments between our sparse kernel and Google’s Sputnik.
2. Summarized experiment results and figures in the paper.

– **DeepSpeed I/O Framework Support for AI4Science**

[CODE LINK](#)

Motivation: AI4Science models have revolutionized the AI world. DeepSpeed can support AI4Science models deployed across multiple nodes but lacks an I/O management framework for handling large amounts of training data efficiently.

Contributions:

1. Investigated DeepSpeed I/O support in supercomputers (Argonne HDF5 Luster System), analyzed data shuffling and fetching patterns for AI4Science models powered by DeepSpeed, and implemented algorithms to accelerate I/O.
2. Implemented a ViT model for weather prediction.

Industry Projects:

– **DeepSpeed Chat: Easy, Fast, and Affordable RLHF Training of ChatGPT-like Models at All Scales (Microsoft DeepSpeed Team)**

[CODE LINK](#)

Motivation: ChatGPT-like models have revolutionized the AI world, but an accessible end-to-end RLHF pipeline for training powerful ChatGPT-like models is still lacking within the AI community.

Contributions:

1. Investigated ColossalAI’s pipeline, learned how to use ColossalAI’s Zero-2, 3, and GeminiDDP, and adapted them for our RLHF algorithm.
2. Ran 400+ benchmark experiments for DeepSpeed Chat, ColossalAI, and HuggingFace powered by native PyTorch. Summarized the results and conclusions in the DeepSpeed blog.
3. Revised DeepSpeed GitHub Landing Page, DeepSpeed Chat Blog, and produced DeepSpeed Chat

video.

• **School of Computer Science and Engineering, SYSU**

Sep. 2018 - Mar. 2022

Research Assistant

Guangzhou, China

Advisor: *Dan Huang* (Associate Professor, SYSU), *Yunfei Du*, *Yutong Lu* (Professor, SYSU)

Research Projects:

- **Pre-Expedite: Use Hierarchical Structure Space for Improving the Performance of Accessing Small Files in Parallel File System - Undergraduate Thesis**

[CODE LINK](#)

Motivation: Implemented an approach to reduce clients' I/O communication with MDS, leveraging minimal additional client-side resources. Ensured high usability without modifying POSIX standards.

Contributions:

1. Investigated the I/O bottleneck in parallel/distributed file systems for Big Data and Artificial Intelligence applications, identifying intensive metadata communication with the metadata server as a primary issue.
2. Utilized POSIX to create ZERO file blocks (Loop Device). Established a VFS within the ZERO file blocks, allowing each user to store small files in their designated ZERO file blocks.

- **HybridShare: Universal Resource Scheduling for Hybrid Jobs**

[CODE LINK](#)

Motivation: CPU- and GPU-centric applications allocate resources exclusively, leading to inefficient utilization of heterogeneous resources.

Contributions:

1. Analyzed the possibility of co-locating modern workflow - application in the same physical machine to share resources.
2. Proposed HybridShare algorithms that can enable different resources-prefer jobs to be co-located in the same node and share hardware resources (e.g., GPU-concentric, CPU-concentric, Mem-intensive) through Slurm, Mesos, Kubernetes.

- **MAEM - Multiple Applications co-Execution time Estimation**

[CODE LINK](#)

Motivation: There are few works to accurately estimate the slowdown of CPU/GPU applications based on the characteristic of applications & hardware architecture

Contribution:

1. Conducted a literature review on application profiling, interference and slowdown estimation, and interference-aware scheduling.
2. Gathered resource consumption data for various benchmarks and analyzed their behavior.

• **Institute of Advanced Networks and Computing Systems, SYSU**

Oct. 2018 - Mar. 2019

Research Intern

Guangzhou, China

Advisor: *Hejun Wu* (Associate Professor, SYSU)

Research Projects:

- **EmReal: A Digital Twin Framework of Emulated and Real Components for Robots with Reinforcement Learning**

[CODE LINK](#)

Motivation: Pioneered a digital twin framework for robots utilizing reinforcement learning (RL), bridging the gap between simulations and real-world deployments. Developed solutions to effectively transition RL algorithms from simulators to actual robots, advancing the field beyond its nascent stage.

Contributions:

1. Conducted a survey on robotics simulator systems and reinforcement learning algorithms.
2. Designed and implemented a one-legged robot, integrating real and emulated components using XLM, Python, ROS, and Arduino C programming.
3. Created a digital twin framework for robotic systems, employing reinforcement learning (RL) and seamlessly blending emulation, pre-training, connectivity, and hardware adaptation using ROS and PyBullet.

- **Co-authored a book on deep learning in reinforcement learning, awaiting publication.**

• **Tencent Holdings Ltd. Weixin Group & Dep. of CS UIUC**

Jul. 2018 - Jul. 2020

Research Intern, Testing, Technical-Architecture Department

Champaign, IL, US & Guangzhou, China

Advisor: *Tao Xie* (Professor and Willett Faculty Scholar, UIUC), *Yuetang Deng* (Director)

Industry Projects:

- **JSIdentify: A Hybrid Framework for Detecting Plagiarism Among JavaScript Code in Online Mini Games**

[TALK LINK](#), [PAPER LINK](#)

Motivation: In cases of plagiarism for mini-games, deeply obfuscated code cloned from the original code often embodies malicious code segments and copyright infringements, posing great challenges for existing plagiarism detection tools. To address these challenges, we design and implement JSIdentify, a hybrid framework to detect plagiarism among online mini games.

Contributions:

1. Worked under the guidance of Prof. Tao Xie, focusing on intermediate representation analysis in V8 & Node.js's Interpreter.
 2. Conducted literature review on code plagiarism detection methods and evaluations of clone detection tools.
 3. **Developed an edit distance estimation and network flow algorithm** to measure similarity in bytecode generated by Ignition, TurboFan Interpreter.
 4. **Designed a priority-queue-based framework** to consolidate multiple plagiarism detection algorithms.
 5. Co-authored a paper titled **"JSIdentify: A Hybrid Framework for Detecting Plagiarism Among JavaScript Code in Online Mini Games."**
- **Microsoft(China) Co.,Ltd. Guangzhou Branch** Sep. 2018 - Feb. 2019
Project Assistant to Senior Cloud Architect Guangzhou, China
 Advisor: *Zhen Guan* (Sr.Partner Technology Strategist, Microsoft)
 - Gained proficiency in Azure's architecture and utilized Azure for training multiple machine learning models.
 - Developed a textile-focused Q&A system to address a market gap in China:
 - * Collected Q&A data by crawling prominent domestic textile websites.
 - * Preprocessed data through cleaning, serializing, and tokenizing text into a corpus.
 - * Implemented a pre-trained BERT model for the Q&A system.
 - * Deployed the BERT model on Azure as a service.
 - **SYSU-CMU Joint Institute of Engineering (JIE)** Feb. 2017 - Aug. 2017
Research & Software Engineer Intern Guangzhou, China
 Advisor: *Xiaoyin Tang* (Professor, Southern University of Science and Technology)
 - Created a front-end website to integrate with a back-end deep learning model for efficient analysis of numerous fundus photographs.
 - Enabled detection of diabetic retinopathy (DR) and diabetic macular edema (DME) through seamless collaboration between the front-end and back-end systems.
 - **Computational Medical Imaging Laboratory, SYSU** Jul. 2016 - Aug. 2017
Research Intern Guangzhou, China
 Advisor: *Yao Lu* (Professor, SYSU)
 - Collected breast cancer data through web crawling Scrapy.
 - Developed an OHIF Viewer web project, available at [LINK](#).
 - Hosted a **SIT (College Students' Innovative Entrepreneurial Training Plan)**, ID: 201502059.
 - Implemented traditional image processing algorithms on mobile platforms.

Other Projects

- **LeetCode Record** Jun. 2017 - Present
Honing Programming Skills Daily
[CODE LINK](#)
 - Utilized languages such as C, CPP, Python3, Java, and Go to solve LeetCode algorithm questions based on my preference.
 - Maintained a repository containing my code and insights for each LeetCode problem.
- **System Related Conference Papers Crawler** Jun. 2021 - Present
Web Scraper and Timeline for Top-tier Systems Conference
[CODE LINK](#)
 - Leveraged Python, BeautifulSoup4, and Requests to scrape papers and crucial deadlines for major computer system conferences.
 - Employed Pandas and Matplotlib to create a timeline representing significant computer system paper submission deadlines.
- **DDLs** Dec. 2017 - May. 2018
Course Project: Design and Development of Android Applications Guangzhou, China
[BACKEND CODE LINK](#) [FRONTEND CODE LINK](#)
 - Developed DDLs, an Android application for personal deadline management, using Java and Android Studio for the front-end, incorporating MVC architecture, and NodeJS with Express.js for the back-end RESTful API.
 - Implemented features such as deadline administration with CRUD operations, adding, completing, and deleting deadlines in a timeline using SQLite for local storage, marking completed deadlines as unfinished, receiving server notifications through WebSocket, sharing timeline screenshots using Android's native sharing capabilities, and user authentication with JSON Web Tokens (JWT) for registration and login functionality.

- ChainLoveHelp** May. 2018 - May. 2018
South China Microsoft Hackathon Competition *Guangzhou, China*
[CODE LINK](#)
 - ChainLoveHelp is dedicated to providing a peer-to-peer platform for university task posting and processing based on blockchain technology.
 - For the chain-end, employed Ethereum-based Parity to construct a consortium blockchain, operating two nodes on the chain for transaction processing, accounting, and consensus.
 - For the front-end, implemented a robust technology stack using PHP for server-side scripting, Apache as the web server, and MySQL for database management.
- Guang Tu** Apr. 2017 - May. 2017
South China Microsoft Hackathon Competition *Guangzhou, China*
[CODE LINK](#)
 - Guangtu is a Windows-based map planning software that utilizes gesture recognition technology for enhanced user interaction.
 - The application was developed using Python for programming, Leap Motion for gesture recognition, PyQt5 for creating the graphical user interface, and Django for building the web framework and backend functionality.
- Seven Seconds** Apr. 2017 - May. 2017
SYSU Student Software Creative Design and Innovation Development Competition *Guangzhou, China*
[CODE LINK](#)
 - Designed and developed an Android App to organize and record memories, leveraging the capabilities of Android Studio and Java. Successfully published the app on the 360 Mobile App Market.
 - Implemented a robust mobile App architecture, encompassing a user-friendly sidebar, homepage, memory management, as well as secure login and registration modules. Employed advanced data handling techniques, RESTful APIs, and seamless integration with a Node.js backend for efficient data processing and storage.
- PVmedtech** Jul. 2016 - Aug. 2017
Advisor: Yao Lu (Professor, SYSU) *Guangzhou, China*
[CODE LINK](#)
 - Collected breast cancer data through web crawling Scrapy.
 - Developed an OHIF Viewer web project, available at LINK.
 - Hosted a **SIT (College Students' Innovative Entrepreneurial Training Plan)**, ID: 201502059.
 - Implemented traditional image processing algorithms on mobile platforms.

Certification

- Course Certificate: Sample-based Learning Methods** an online non-credit course authorized by University of Alberta, Alberta Machine Intelligence Institute and offered through Coursera; November, 14, 2021
- Course Certificate: Fundamentals of Reinforcement Learning** an online non-credit course authorized by University of Alberta, Alberta Machine Intelligence Institute and offered through Coursera; Aug, 20, 2021
- IELTS Test Scores: Listening: 6.5; Reading: 7.0; Writing: 6.5; Speaking: 6.5; Overall: 6.5** October, 08, 2020

Skills

- Programming Languages:** Pascal (11 yrs), C (11 yrs), C++ (11 yrs), Python (6 yrs), HTML, CSS, JavaScript (6 yrs), Java (6 yrs), SQL (6 yrs), Bash (6 yrs), LaTeX (5 yrs), Matlab (5 yrs), CUDA (5 yrs), R (4 yrs), Go (4 yrs)
- Systems and Infrastructure:** MPI/OpenMPI (6 yrs), Linux Kernel (5 yrs), Distributed/Parallel File Systems - e.g., Lustre, HDFS (5 yrs), Kubernetes, Kubernetes Scheduler, Kubernetes SR-IOV (5 yrs), Docker (5 yrs), Hadoop (4 yrs), Spark (4 yrs), YARN (4 yrs), Mesos (4 yrs)
- Machine Learning and AI:** TensorFlow (5 yrs), PyTorch (4 yrs), TorchServe (4 yrs), TensorBoard (4 yrs), Ray (4 yrs), JAX (2 yrs), Triton (2 yrs), DeepSpeed (1 yr), HuggingFace (1 yr), Reinforcement Learning (4 yrs), CNN, RNN, ResNet, Attention Block, UNet, Transformer, ViT (5 yrs), Neural Architecture Search (3 yrs), Diffusion Models (1 yr), GPT-2,3,4 (1 yr), Reinforcement learning from human feedback (1 yr)
- Databases and Storage:** MySQL (6 yrs), Oracle SQL (6 yrs), MongoDB (6 yrs), PostgreSQL (6 yrs), Redis (4 yrs), Hive SQL (3 yrs)
- Front-end Development:** PHP (6 yrs), Vue.js (6 yrs), ReactJS (6 yrs), ASP.NET (6 yrs), jQuery (6 yrs), AngularJS (6 yrs), Apache (6 yrs), MeteorJS (6 yrs)

- *Back-end Development:* **Spring Boot** (6 yrs), **Django** (6 yrs), **Flask** (6 yrs), **Node.js** (6 yrs), **Express** (6 yrs), **REST API Design** (6 yrs), **CI/CD**
- *Mobile Development:* **Android Studio** (6 yrs, Java, Kotlin), **XCode** (6 yrs, Swift, Objective-C), **React Native** (6 yrs, Cross-platform), **Flutter** (6 yrs, Cross-platform)
- *Web Crawling & Testing:* **Urllib** (6 yrs), **BeautifulSoup** (6 yrs), **Scrapy** (6 yrs), **Requests** (6 yrs), **JSON** (6 yrs), **Selenium** (6 yrs), **Pytest** (6 yrs), **JUnit** (6 yrs)
- *Version Control & Build Systems:* **Git** (8 yrs), **Gradle** (6 yrs, Android, Java), **Maven** (6 yrs, Java), **npm** (6 yrs, JavaScript), **pip** (6 yrs, Python)
- *Development Tools & Libraries:* **Airflow** (2 yrs), **Kafka** (2 yrs), **Elasticsearch** (2 yrs), **OpenCV** (5 yrs), **Pandas** (5 yrs), **NumPy** (5 yrs), **SciPy** (5 yrs), **NLTK** (5 yrs), **Matplotlib** (5 yrs), **Seaborn** (5 yrs), **Azure Data Factory** (2 yrs), **AWS** (2 yrs), **Google Cloud Platform** (2 yrs)

Extracurricular

- **Fitness:** **Fencing**(6 yrs), **Jogging** (7 yrs), **Bodybuilding**(6 yrs) (Hongxing Fitness Club Outstanding Students), **Table Tennis**(11 yrs), **Badminton**(11 yrs)
- **Leisure:** **Web & Mobile Application Development**(5 yrs), **Saxophone**(9 yrs), **Magic** (1 yrs), **Video Games** (more than **500+** PS5 game collections)
- **Volunteer:**
 - *Sun Yat-sen University, School of Computer Science and Engineering, Student Union, Vice President, Jul, 2016 - Jul, 2017*
 - * Mentored incoming freshmen, helped them acclimate to the university environment, and promoted a sense of belonging through inclusive campus activities and events.
 - *Changjun High School Volunteer, Jul, 2013 - Jul, 2014*
 - * Enhanced the nursing home experience by engaging in meaningful conversations with elderly residents, preparing and serving fresh fruit, and maintaining a clean and sanitary environment for their well-being.