Kai Yao

PhD Candidate

Informatics Forum, 10 Crichton Street, EH8 9AB, Edinburgh, UK

kai.yao@ed.ac.uk — kaikaiyao.github.io

Education

- **University of Edinburgh** — Edinburgh, UK
  PhD in Cyber Security, Privacy and Trust — 2023 – Present
  Research Focus: Secure and Trustworthy Machine Learning
  Advisor: Dr. Marc Juarez
  Expected Graduation: 2026

- **Johns Hopkins University** — Baltimore, MD, USA
  MS in Mechanical Engineering — 2020

- **Fudan University** — Shanghai, China
  BS in Theoretical Mechanics — 2017

Research Experience

- **Fingerprinting Generative Models Against Malicious Providers** University of Edinburgh 2024 – 2025
  - Pioneered adversarial model fingerprinting as first work extending generative model fingerprinting to adversarial threat models where providers act maliciously.
  - Achieved near-zero FPR@95%TPR by developing robust fingerprint extraction methods for GANs and diffusion models, resistant to architectural modifications and adversarial attacks.
  - Built comprehensive evaluation framework and empirically validated methods across multiple generative architectures and attack scenarios.
  - Published open-source implementation, advancing AI security and model provenance verification research.

- **Differential Privacy's Disparate Impact in Machine Learning** University of Edinburgh 2023 – 2024
  - Conducted a comprehensive analysis identifying factors that exacerbate fairness disparities in differentially private (DP) machine learning models.
  - Developed a taxonomy categorizing contributing factors across DP mechanisms, model architectures, training data, and data distributions.
  - Performed causal analysis to pinpoint dataset size and group distance to decision boundaries as critical conditions for DP-induced unfairness.
  - Evaluated mitigation strategies, noting limitations such as group label dependencies and computational costs.
  - Identified research gaps in cross-factor interactions, distributional impacts, and conflicting fairness definitions, and proposed future research directions.

- **DL-based 3D Single-Cell Morphology and Size Prediction** Johns Hopkins University 2019 – 2020
  - Developed a high-throughput, label-free AI technique to predict 3D single-cell morphology and size from DIC microscopy images.
  - Designed a microfluidic system employing the fluorescence exclusion method to measure cell morphology.
  - Created image processing algorithms for preprocessing both DIC and FXm images.
  - Implemented a U-Net-based CNN model and optimized its performance via hyperparameter tuning.

- **DL-based Cell Type Classification and Morphological Phenotyping** Johns Hopkins 2018 – 2019
  - Developed a high-throughput, label-free AI method to classify normal versus cancer cells using low-resolution flask images.
  - Created an automated pipeline for screening and preprocessing microscopy images.
  - Designed a CNN-based clustering method to group cells by morphology and analyze tumor cell shapes.
  - Investigated relationships among cell type, density, and morphology to elucidate cancer cell behavior in vitro.

Publications

- **Yao K**, Juarez M. *AuthPrint: Fingerprinting Generative Models Against Malicious Model Providers.* arXiv preprint, 2025.
- **Yao K**, Juarez M. *SoK: What Makes Private Learning Unfair?* Proceedings of the 3rd IEEE Secure and Trustworthy Machine Learning Conference, 2025.
- Rochman ND*, **Yao K***, Gonzalez NA*, Wirtz D, Sun SX. *Single-Cell Volume Measurement Utilizing the Fluorescence Exclusion Method (FXm).* Bio-protocol. 2020 Jun 20;10(12):e3652.
- **Yao K***, Rochman ND*, Sun SX. *CTRL: A Label-Free Artificial Intelligence Method for Dynamic Measurement of Single-Cell Volume.* Journal of Cell Science. 2020 Apr 1;133(7):jcs245050.
- Perez-Gonzalez NA*, Rochman ND*, **Yao K***, Tao J, Le MT, Flanary S, Sablich L, Toler B, Crentsil E, Takaesu F, Lambrus B. *YAP and TAZ Regulate Cell Volume.* Journal of Cell Biology. 2019 Oct 7;218(10):3472–88.
- **Yao K***, Rochman ND*, Sun SX. *Cell Type Classification and Unsupervised Morphological Phenotyping from Low-Resolution Images Using Deep Learning.* Scientific Reports. 2019 Sep 17;9(1):1–3.
- Zhang Q, Meng Z, Zhang Y, **Yao K**, Liu J, Zhang Y, Jing L, Yang X, Paliwal N, Meng H, Wang S. *Phantom-Based Experimental Validation of Fast Virtual Deployment of Self-Expandable Stents for Cerebral Aneurysms.* BioMedical Engineering OnLine. 2016 Dec;15(2):431–7.

*Note: * denotes equal contributions.*

Professional Experience

- **AI Frameworks Engineer, Domain Lead** Intel Corp., Shanghai, China 2021 – 2023
  - Led the development of Neural Coder, an automation tool that optimizes the training and inference throughput of PyTorch and TensorFlow workloads on Intel hardware.
  - Enhanced the Intel Extension for PyTorch by incorporating features that improve computational efficiency on Intel hardware.
  - Developed PyTorch adapter algorithms for Intel Neural Compressor, enabling INT8 quantization for enhanced model throughput.
  - Optimized inference performance for AIGC models (e.g., Stable Diffusion) running on Intel hardware.
  - Designed a comprehensive benchmarking system for Intel's AI software and conducted performance evaluations of AI workloads on both Intel and competitor hardware.
  - Collaborated with corporate partners (e.g., Alibaba, AWS) to integrate Intel AI solutions into their ecosystems.
- **AI Algorithm Engineer** Huawei Technologies Co., Ltd., Shanghai, China 2020 – 2021
  - Developed 5G machine learning algorithms, focusing on MU-MIMO features using MLP, CNN, and RNN architectures.
  - Optimized and compressed models for efficient training and inference.
  - Worked with deployment and validation teams to ensure successful implementation.

Awards, Fellowships, & Grants

- Travel Grant, 3rd IEEE SaTML Conference, IEEE, 2025
- Travel Grant, 2nd IEEE SaTML Conference, IEEE, 2024
- LFCS Travel Fund, PETS Conference, University of Edinburgh, 2023
- School of Informatics PhD Scholarship, University of Edinburgh, 2023
- Division Achievement Award, Fast Stable Diffusion on Intel CPU, Intel AIA, 2023
- Division Recognition Award, Neural Coder Partnering Alibaba Cloud, Intel CESG SW AI, 2023
- Division Achievement Award, Innovation of Neural Coder, Intel AIA, 2022
- Departmental Research Fellowship, Johns Hopkins University, 2017
- Outstanding Graduate of the Year 2017, Fudan University, 2017
- JASSO Full Scholarship for Exchange Students, Japanese Government, 2014

Teaching Experience

- **Privacy and Security with Machine Learning** University of Edinburgh, Edinburgh, UK 2024 – 2025
  Teaching Assistant and Lab Demonstrator
- **Privacy and Security with Machine Learning** University of Edinburgh, Edinburgh, UK 2023 – 2024
  Teaching Assistant and Lab Demonstrator
- **Mathematical Image Analysis** Johns Hopkins University, Baltimore, MD, USA 2019 – 2020
  Teaching Assistant