JOON KIM

joonkim1@berkeley.edu • **Linkedin** • **GitHub**

Education

University of California, Berkeley *Aug 2021 – May 2026 (Expected)*

B.S. Electrical Engineering & Computer Science, GPA: 4.00/4.00

- **Relevant Coursework:** CS 61B(Data Structures: A+), CS 70(Discrete Math & Probability: A+), EECS 16A/B(Circuits & Control: A+), CS 61A(Python: A+), E 7(MATLAB & Numerical Analysis: A+), DATA 8(Pandas: A), CS 170(Algorithms), CS 188(Artificial Intelligence)

Experience

Berkeley Artificial Intelligence Research - CHEN Lab Berkeley, CA

Undergraduate Researcher *Jul. 2024 - Current*

- Designing zero-shot LLM pseudo-label pipeline to improve semi-supervised learning accuracy; advised by Professor Irene Chen
- Took charge of image experiments; investigated LLM agents for image labeling such as CLIP, and showed results on CIFAR-100
- Currently working on RadQA dataset; implemented FixMatch on a non-inference task for baseline comparison

JLK Group Seoul, South Korea

Research Intern, First Author *Feb. 2024 - May. 2024*

- Developed Federated Learning models reaching near identical performance to commercially deployed U-Net models using Python
- Collaborated with four M.D. professionals to investigate the use of Federated Learning in medicine; advised by Dr. Wi-Sun Ryu

Keimyung University Daegu, South Korea

Independent Researcher, First Author *Feb. 2023 - Jul. 2024*

- Proposed a randomized masking algorithm as an obfuscation technique against Deep Leakage in image-based Federated Learning
- Designed experiments to compare performance-privacy trade-offs amongst SOTA defense algorithms; advised by Prof. Sejin Park

Impact AI Seoul, South Korea

Data Engineering Intern *Jul. 2022 - Aug. 2022*

- Developed a data preprocessing pipeline to pattern-match raw datasets of various formats from multiple companies using Python
- Researched and presented eight AI-based B2B SaaS business case studies, showcasing their strengths, weaknesses, and outlooks
- Contributed in designing SQL-like UI/UX features for the main page of web and native applications deployed to client companies

Studio.geo @ UC Berkeley Berkeley, CA

Undergraduate Researcher *Feb. 2022 - May. 2022*

- Experimented Progressive-GAN on the Savio cluster to generate artificial maps using Python; advised by Prof. Clancy Wilmott
- Pictures of 4x4 grid of generated maps of 256x256 pixels trained on real colored maps included in Prof. WIlmott's book proposal

Independent Biomedical Research Seoul, South Korea

Independent Researcher *Jan. 2020 - Jun. 2020*

- Proposed a microarray analysis model for screening early schizophrenia with RNA genetic samples; overcame the lack of public RNA data with oversampling techniques and elected a Deep Neural Network model for inference; advised by Ph.D. Taehyun Kim
- Verbally presented research findings at the 2020 Society of Interdisciplinary Business Research Conference as a representative

Selected Publications

- **In-Silo Federated Learning vs. Centralized Learning for Segmenting Acute and Chronic Ischemic Brain Lesions** *(Medrxiv)*; J. Kim, H. Lee, W. Ryu, et al.; Comparative analysis of Federated and Centralized Learning on brain MRI images. Showed Federated Learning is as effective as Centralized learning on real-life non-i.i.d. brain lesion datasets of ~10,000 patients over 9 institutions
- **Random Gradient Masking as a Defensive Measure to Deep Leakage in Federated Learning** *(Arxiv)*; J. Kim, S. Park; Compared the efficacy of randomly masking gradients from Federated Learning submissions against other defenses against Deep Leakage from Gradients such as Pruning, Compression, and Noising on Convolution Neural Networks; Submitted to journal Applied Intelligence

Skills

- Python, PyTorch, Tensorflow, Docker, C, RISC-V, MATLAB, Pandas, Rust, Go, Java, Javascript, Flutter, Firebase