

# JINGHAO ZHENG

✉ jinghao.zheng@epfl.ch • ☎ (+41) 78 318 24 62

## EDUCATION

<b>École Polytechnique Fédérale de Lausanne (EPFL)</b>	Aug. 2025 – Jan. 2028 (Expected)
M.S. in Computer Science	
<b>Shanghai Jiao Tong University (SJTU)</b>	Sept. 2021 – June 2025
B.E. in Automation	
• Major GPA: 3.82/4.3	• Centesimal grade average: 89.23/100
<b>Related Courses:</b> Calculus II (98), Probability and Statistics (99), Linear Algebra (92), Discrete Mathematics (93), Data Structure (90), Pattern Recognition (96), Artificial Intelligence (90), Robotics (93), Principles of Automatic Control (94)	

## PUBLICATION

D. Nguyen\*, J. Li\*, **J. Zheng\***, B. Mirzasoleiman. Do We Need All the Synthetic Data? Towards Targeted Synthetic Image Augmentation via Diffusion Models. **International Conference on Learning Representations**, 2026. (*Accepted*)  
Z. Huang, X. Cheng, **J. Zheng**, H. Wang, Z. He, T. Li, and X. Huang. Unified Gradient-Based Machine Unlearning with Remain Geometry Enhancement. **Neural Information Processing Systems**, 2024. (*Accepted as a spotlight*)

## RESEARCH & PROJECT EXPERIENCE

<b>Continual Learning for LLMs: A Parameter Sorting Approach to Mitigate Catastrophic Forgetting</b>	
Research Assistant	Nov. 2024 – June 2025

Advisor: **Xiaolin Huang**, Professor, Vice Dean, Department of Automation, SJTU

- Proposed a dynamically block-wise parameter sorting method to mitigate catastrophic forgetting within a continual learning setting for large language models (LLMs).
- Improved overall performance (OP) by 4.88%, backward transfer (BWT) by 6.5%, and training efficiency by 58.37% compared to the state-of-the-art (SOTA) method trained on Qwen2-1.5B.

### Fewer Generated Images for Better Augmentation

Research Assistant	July. 2024 – May. 2025
--------------------	------------------------

Advisor: **Baharan Mirzasoleiman**, Assistant Professor, Computer Science Department, UCLA

- Leveraged the GLIDE, a text-to-image model, with real guidance strategies to diversify the training datasets.
- Conducted experiments on multiple datasets and across multiple architectures using upsampling and our method.
- Improved the classification accuracy by 1% than that without augmentation and training efficiency by 70% than  $2\times$  scale Diffusion-based data augmentation using selectively incorporating generated data.

### Unified Gradient-Based Machine Unlearning (MU) with Remain Geometry Enhancement

Co-author	Mar. 2024 – June. 2024
-----------	------------------------

Advisor: **Xiaolin Huang**, Professor, Vice Dean, Department of Automation, SJTU

- Proposed using KL divergence on the remaining output distribution, instead of Euclidean distance in vanilla methods, as the manifold metric to prevent deviations in the model output on the remaining set, improving MU performance.
- Conducted experiments and parameter tuning to compare the performance of our algorithm with other MU methods in image classification and generation across various datasets and models of different architectures.
- Improved the averaging disparity by 1.8% on average in random subset forgetting on CIFAR-10 in image classification and the FID by 80 on average in class-wise forgetting on ImageNet in image generation.

## HONORS & AWARDS

University-Level Outstanding Graduate of SJTU	2025
3 <sup>rd</sup> Prize, Academic Scholarship of SJTU	2022 & 2023 & 2024
3 <sup>rd</sup> Prize, TI Cup National Undergraduate Electronic Design Contest Shanghai Area	2023

## SKILLS

- Programming: Python, C/C++, ROS/ROS2, Matlab, L<sup>A</sup>T<sub>E</sub>X, Markdown, Stata17
- Techniques: Linux, Windows
- Languages: Chinese (native), English (proficient)
- Leadership Experience: Head of Sports Department, School of Electronic Information and Electrical Engineering (SEIEE) Student Union, Shanghai Jiao Tong University

Dec. 2022 – Dec. 2023