ZHUOYUAN CHEN

Research Scientist · Frontier Foundation Models · Pretraining, Post-Training & RL
chenzhuoyuan07@gmail.com · 847-868-4043 · levintsky.github.io · Google Scholar (10,000+ citations)

Research Profile

Research scientist with 15+ years of experience building frontier foundation models and large-scale learning systems across xAI, Google, Apple, Meta FAIR, and Baidu. Worked across the full model lifecycle—data and scaling ladders, pretraining, mid-training, SFT, reinforcement learning, evaluation, and production deployment—with contributions to Grok, Gemini, and Apple Vision Pro. Brings particular depth in multimodal learning, generative modeling, and large-scale RL. Published at ICML, NeurIPS, CVPR, ICCV, and ECCV, with 10,000+ citations.

Experience

xAI | Palo Alto, CA 2025–2026

Member of Technical Staff, Omni / Foundation Models

- Contributed to successive Grok model generations across scaling ladders, pretraining, mid-training, post-training, and evaluation, working at the intersection of model training, large-scale data, and multimodal capabilities.
- Built web-scale data and supervision pipelines that transformed heterogeneous video, speech, and language signals into hundreds of billions of high-quality training tokens.
- Designed and analyzed data mixtures, training ablations, and evaluation suites to improve spatial-temporal reasoning without regressing general model capabilities.

Google | Sunnyvale, CA 2024–2025

Staff Research Engineer and Tech Lead, Vertex AI / Gemini

- Developed and validated high-value training recipes for Gemini through billion-scale data construction, end-to-end training, and controlled ablations; successful improvements were integrated into Gemini core pipelines.
- Built LoRA, SFT, RL, and release-evaluation workflows with Gemini Core teams; contributed to Gemini 1.5, 2.0, and 2.5.

Apple | Cupertino, CA 2020–2024

Staff Research Scientist and Tech Lead, Applied Foundation Models

- Led voxel-based 3D perception deployed in Apple Vision Pro; co-developed ARKitScenes (NeurIPS 2021) and GAUDI (NeurIPS 2022).

Uber Advanced Technologies Group | San Francisco, CA 2019–2020

Senior Research Scientist, Autonomous Driving Research

- Developed shape-aware adversarial 3D point clouds to expose long-tail safety failures in autonomous-driving perception and improve robustness (ShapeAdv, ECCV 2020).

Meta AI Research (FAIR) | Menlo Park, CA 2017–2019

Research Engineer

- Built large-scale reinforcement-learning systems for ELF OpenGo, an open reproduction and analysis of AlphaZero spanning distributed training, inference, and evaluation (ICML 2019 Oral).
- Led autoregressive 3D generative modeling with an order-aware VoxelCNN, an early application of autoregressive generation to 3D worlds (first author, ICCV 2019).

Experience (continued)

Baidu Research, Institute of Deep Learning | Sunnyvale, CA 2014–2017

Senior Research Scientist

- Developed deep representation learning for stereo perception in autonomous driving (first author, ICCV 2015) and explored reinforcement learning for embodied agents.
- Contributed CUDA and systems work to PaddlePaddle, gaining early experience in industrial deep-learning infrastructure, numerical performance, and scalable training.

Adobe Research | San Jose, CA 2012–2013

Research Intern, Imagination Lab

- Developed motion estimation and video-understanding methods resulting in CVPR 2013 and ECCV 2016 publications and multiple U.S. patents.

Selected Publications & Model Releases

- Grok 4.5 (xAI Model Release, 2026) Grok 4.3 (xAI Model Release, 2026) Grok 4.20 System Card (xAI Technical Report, 2026)
- Gemini 3 (Google Model Release, 2025) Gemini 2.5. Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next-Generation Agentic Capabilities, 2025.
- Gemini 2.0 (Google Model Release, 2024) Gemini 1.5. Unlocking Multimodal Understanding Across Millions of Tokens of Context, 2024.
- GAUDI. A Neural Architect for Immersive 3D Scene Generation, NeurIPS 2022. ARKitScenes. A Diverse Real-World Dataset for 3D Indoor Scene Understanding, NeurIPS 2021.
- ShapeAdv. Generating Shape-Aware Adversarial 3D Point Clouds, ECCV 2020. Order-Aware Generative Modeling Using the 3D-Craft Dataset. ICCV 2019.
- ELF OpenGo. An Analysis and Open Reimplementation of AlphaZero, ICML 2019 Oral. Large Displacement Optical Flow from Nearest Neighbor Fields. CVPR 2013.

Education

Northwestern University — Ph.D., Electrical Engineering and Computer Science 2014
Tsinghua University — M.S., Computer Science (Outstanding Thesis Award) 2010
Tsinghua University — B.S., Mathematics and Physics, Academic Talent Program 2007