

EDUCATION

Peking University

Expected 07/2026

B.S. in Physics (School of Physics)

- **GPA:** 83.7/100
- **Relevant Fields:** Reinforcement Learning, Computer Vision, AI for Science.

ACADEMIC VISITS

Shenzhen International Quantum Academy

01/2025 – 02/2025

AWARDS

Alishan Scholarship

8/261 recipients

Outstanding Research Award

30/261 recipients

National Endeavor Scholarship

PUBLICATIONS

Let's Verify and Reinforce Image Generation Step by Step

CVPR 2025

Renrui Zhang*, Chengzhuo Tong*, **Zhizheng Zhao***, Ziyu Guo*, Huaidong Zhang, Manyuan Zhang, Peng Gao, Hongsheng Li

(* Co-first authorship)

RESEARCH EXPERIENCE

Reward and Policy Distribution Optimization in GRPO

03/2025 – 07/2025

University of Illinois Urbana-Champaign (UIUC) | Advisor: Prof. Minjia Zhang

Conducted a comprehensive analysis of Group-Based Reward Policy Optimization (GRPO), structuring the investigation into two primary technical thrusts:

- **Part 1: Enhancing Advantage Granularity** (Addressing coarse sample-level feedback)
 - *Token-Level Critic*: Designed a critic model to learn token “vitality” ($b_\phi(t)$), attempting to assign fine-grained credit to critical tokens within a sample.
 - *Probability Compensation*: Developed a mechanism to amplify gradients for “rare gems” (high-reward but low-probability samples) to accelerate learning from sparse signals.
- **Part 2: Addressing Distribution Mismatch** (Solving the “All-Fail” group phenomenon)
 - *Resource Re-allocation*: Experimented with culling “solved” prompts to concentrate computational budget on hard instances (Concentration of Force).
 - *Negative Advantage*: Proposed explicitly suppressing consistent failure regions by assigning negative rewards to zero-gradient groups (validating NGRPO concepts).
 - *Dense Reward Shaping*: Explored replacing binary rewards with Ground Truth Likelihood to guide reasoning chains in mathematical tasks.

Chain-of-Thought Reasoning for Advanced Image Generation

09/2024 – 01/2025

CUHK & ByteDance | Advisor: Dr. Renrui Zhang

- Applied **Chain-of-Thought (CoT)** strategies to autoregressive text-to-image generation, integrating Direct Preference Optimization (DPO) to enhance reasoning capabilities.
- Independently proposed and developed the **Potential Assessment Reward Model (PARM)**. This model adaptively evaluates intermediate generation steps to select high-potential reasoning paths, balancing test-time compute efficiency with performance.

- Enhanced the *Show-o* model, achieving a **+24% improvement on GenEval** and surpassing Stable Diffusion 3 by **+15%** in benchmark performance.

Improving Crystal Structure Prediction via Niggli Reduction

07/2025 – 09/2025

The Chinese University of Hong Kong (CUHK) | Advisor: Prof. Shengchao Liu

- Identified a fundamental theoretical limitation in the state-of-the-art **DiffCSP** framework: reliance on $E(3)$ equivariance fails to account for **lattice basis transformation invariance**.
- Proposed a principled framework incorporating **Niggli reduction** to align predictions with canonical lattice representations, effectively closing the theoretical gap.
- Designed a **differentiable Proxy Loss** to enable effective training on these canonical representations, significantly improving the physical validity of predicted crystal structures.

Resistive Plate Counter Development

03/2024 – 07/2024

Peking University | Advisor: Prof. Qite Li

- Assembled and commissioned a Resistive Plate Counter (RPC) detector for high-energy physics experiments.
- Developed signal processing pipelines and feature extraction algorithms for precise Particle Identification (PID).

SELECTED COURSE PROJECTS

Neural Network Solver for Electric Fields

12/2024

Course: Physics and AI | Advisor: Prof. Yanqing Ma

- Implemented a **Physics-Informed Neural Network (PINN)** to solve complex electric field distributions.
- Designed custom loss functions integrating boundary conditions and PDE residuals to approximate physical solutions.

SKILLS

<i>Languages</i>	Python, LaTeX, CERN ROOT
<i>Tools</i>	PyTorch, Git, Linux Environments