

# Rui Huang

<https://ruihuangai.github.io/>



UESTC | “Everest Project” Computer Top Talent Experimental Class

## Information

Age: 20      Class of 2023      Wechat: huangrui\_dby      Email: huang\_rui@std.uestc.edu.cn  
paulafixamiworali@gmail.com

## Education

UESTC - “Everest Project” Computer Top Talent Experimental Class      2023.09 – 2027.06

- Grade: 97.75      Calculus: 97      Linear Algebra: 98      GPA: 3.99/4      Ranking: 1/334

## Selected Honors and Awards

National Scholarship of 2024 & 2025      Top 6 in College      SenseTime Scholarship      30 Recipients Nationwide

National College Students’ Career Planning Contest      First Prize      Ganen Modern Science Fellowship      Top 10 in School

## Selected Publications

2023.12 – Now

### Post-Training for Reasoning in MLLM/LLMs

- P1: Mastering Physics Olympiads with Reinforcement Learning**

Tech Report Contributor

- P1-VL: Bridging Visual Perception and Scientific Reasoning in Physics Olympiads**

Tech Report Contributor

A series of models that achieved **gold medal-level performance** in the Physics Olympiad (IPhO), **beating all the models**, including Gemini 2.5 Pro, GPT-5 and Grok 4. Reported by [量子位].

### Generative AI and Multimodal Modeling

- Diffusion Dataset Condensation: Training Your Diffusion Model Faster with Less Data**

CVPR 2026 First Author

Proposed D<sup>2</sup>C: Diffusion Dataset Condensation for diffusion models, enabling 100× faster training with 0.8%–4% data; **trained on hundreds of A800/H100 GPUs. Received strong all-positive reviews(scores 4/6, 4/6, 5/6).**

- Beyond Fixed Formulas: Data-Driven Linear Predictor for Efficient Diffusion Models**

CVPR 2026 Co-First Author

Proposed L<sup>2</sup>P, one of the first learnable linear predictors for DiTs, achieving 4.15×–7.18× acceleration on FLUX.1-dev, Qwen-Image, and Hunyuan-Video. **Received strong all-positive reviews(scores 4/6, 5/6, 5/6).**

- Can We Generate Images with CoT? Let’s Verify and Reinforce Image Generation Step by Step**

IJCV (Under Review) Core Contributor GitHub 800+stars, 80+cite

Proposed CoT-Image with step-wise reasoning and novel reward models (PARM/PARM++), improving autoregressive image generation by 24% via test-time verification and preference alignment.

- UniVA: Universal Video Agents towards Next-Generation Video Intelligence**

Tech Report Core Contributor

UniVA **unifies understanding/segmentation/editing/generation** into traceable multi-step video workflows via Plan-Act agents, multi-level memory, and MCP-based tool servers, plus UniVA-Bench. Reported by AK and [新智元].

- Wavelet-Assisted Multi-Frequency Attention Network for Pansharping**

AAAI 2025 (Oral) Co-First Author

Proposed WFANet for image fusion, combining wavelet transformation with attention, achieving SOTA on benchmark.

**AI for Smart Power Systems**

- **Complementary Online Learning Network for Probabilistic Load Forecasting Against Extreme Weather**  
**IEEE Trans. on Industrial Informatics**(SCI Q1, IF 10.2) **First Author**

Proposed the Complementary Online Learning Network (COLNet) with a Weather-aware gating mechanism for high-precision probabilistic and point forecasting under extreme weather.

- **A Novel Spatiotemporal Pyramidal Graph Modeling Approach for Short-Term Residential Load Forecasting**  
**IEEE Trans. on Industrial Informatics**(SCI Q1, IF 10.2) **Core Contributor**

Proposed a low-complexity spatiotemporal pyramidal attention transformer (PATNet) that builds temporal and spatial pyramidal graphs to capture multi-resolution dependencies for accurate multi-step residential load forecasting under volatile residential behaviors.

- **Causal Mechanism-Enabled Zero-Label Learning for Power Generation Forecasting of Newly-Built PV Sites**  
**IEEE Trans. on Sustainable Energy** (SCI Q1, IF 7.9) **Core Contributor**

Developed a causality-enabled domain adaptation framework (CEDAN) that learns domain-invariant PV generation mechanisms to achieve zero-label deterministic and probabilistic power forecasting for newly built PV sites.

**Research Trajectory**

2023.12 – Now

**Stage1: AI for Smart Power Systems 23.12-24.10**

IEEE TII First Author ×1, IEEE Trans. ×4, Student Second Author

**Stage2: Visual Generation / Multimodal / LLM 24.7-Now**

- |  |   |   |
|--|---|---|
| ① <b>Make an Image Clearer</b><br>AAAI 25 Oral CCC 25 Oral CVPR 26 | ② <b>Pretrain a DiT to Generate Images</b><br>CVPR 26 | ③ <b>How to Reason While Generating Images?</b><br>IJCV Under Review<br>Github 800+stars, 80+Cite |
| ⑦ <b>Agentic/LLM reasoning, RL</b><br>Tech Report, <b>Now</b>      | ⑥ <b>World Model</b><br><b>Now</b>                    | ④ <b>Create a Unified Video Agent System</b><br>ICLR 26 Under Review                              |
|  |   | ⑤ <b>How to Make Video and Image Inference Faster?</b><br>CVPR 26                                 |

**Stage3: World Model / Embodied AI / Cognitive Reasoning/Agentic RL /Future**

⑧、⑨、⑩ .....

**Internships / Visiting****Shanghai AI Lab****Research Assistant**

Supervised by Ganqu Cui and Ning Ding, mainly researches LLM SFT and RL.

**CUHK MM Lab****Research Assistant**

Conducted CoT-Image research under Renrui Zhang; studied R1 like video generation .

**HKUST(Guangzhou)****Research Assistant**

Completed D<sup>2</sup>C (CVPR 2026 submission) as first author; focused on diffusion models and data condensation.

**University of Cambridge****Visiting Student**

Participated in AI track; conducted a load forecasting project and received Excellent Student Award.