

To Prospective Members

Last update: March, 2026

If you are looking for a position in our lab, please read the following instructions carefully.

Position Types / Openings

- **PhD** - PhD student (no positions anymore for 2025 fall)
- **Visiting** - visiting student (if you are currently a student in a university and want to visit our lab temporarily, like 3 - 6 months, this is probably what you want)
- **RA** - research assistant (this is a full-time job for at least 1 year)
- **Postdoc** - (this is also a full-time job for at least 2 years)

In general, the bar for prospective PhD students > RA/Postdoc > visiting.

What kind of students are we looking for? The lab mission, as detailed in a set of internal slides, is as follows.

Our Mission & Core Values



- **Mission.** The mission of the lab is to cultivate scientific research talents with integrity, independence, and international competitiveness (“正直、独立、具有国际竞争力的科研力量”).

Ph.D. students are expected to demonstrate this kind of research capability; in particular, we value independence and responsibility. Because of this, we require prospective Ph.D. students **to have on-site collaboration with us if they apply for the Ph.D. position in our lab**. If you are looking for a Ph.D. student position for *Year N, Fall*, we suggest you contact us before the start of Year N-1 Summer (typically around the NeurIPS deadline, *i.e.*, May of *Year N-1*), so that we can arrange a summer internship for you. The internship is paper-oriented (typically targeting ICLR or CVPR). We shall provide compensation for your on-site work.

Self-Introduction

My name is **Huan Wang** (王欢, he/him/his). I am a **Tenure-Track Assistant Professor** at [Westlake University](#) (Hangzhou, China), **principal investigator (PI)** of **ENCODE Lab** (Efficient Neural Computing and Design Lab, “高效智能计算实验室” in Chinese). I earned my Ph.D. (2024) at Northeastern University (Boston, USA). Before that, I spent seven wonderful years at Zhejiang University (Hangzhou, China) earning my bachelor's and master's degrees in the College of ISEE (浙大信电学院). I mainly work on problems in AI (Efficient AI) and computing, with a focus on multimodal and generative AI now. Please visit my [webpage](#) for more information.

Lab Introduction

- **General background.** I am with the Department of Artificial Intelligence in the School of Engineering. We are dedicated to advancing influential innovations in the theories and applications of various AI subfields, aiming to cultivate the next generation of global leaders through inspiring research projects that motivate their professional development.

- **Research directions of the lab.** I am interested in a variety of topics in AI and computing. At this point, the lab is focused on **Efficient AI, multimodal AI, and generative AI**, with particular interest in the following topics:
 - “hot topics”:
 - LLMs/MLLMs, RL training;
 - test-time scaling, reasoning, CoT;
 - generative modelling (diffusion models, flow-based models);
 - agents / agentic learning;
 - token compression, KV cache compression;
 - (mainly focus on inference, not training)
 - classical model compression and acceleration algorithms:
 - quantization, network pruning, distillation, low-rank decomposition, efficient architecture design, or search;
 - speculative decoding.
 - MLSys, AI4Sys and Sys4AI:
 - LLM serving, AI infra;
 - CUDA / Triton / parallel computing / HPC;
 - code generation, kernel generation, AI for SWE;
 - Mobile and edge AI:
 - mobile deployment (iOS and Android deployment);
 - edge computing, embedded systems.
 - Interdisciplinary AI: AI for science, AI for X, AI for social science, AI ethics.
- **Advising style and lab culture.**
 - I am pretty **hands-on**. I’ll regularly meet with students, discuss, help write the paper, code (if needed), etc., especially for junior students.
 - Meanwhile, you will have **sufficient freedom** to develop your ideas and capabilities.
 - We value an **equal, chill, and creative** working environment. Essentially, we are colleagues at Westlake, working together to solve challenging problems.
- **What we can provide.**
 - **Research-oriented projects** with competitive compensation and a comfortable working environment.
 - We encourage students to do research-oriented **internships**. We will do our best to help find opportunities.
 - We encourage students to **attend academic conferences** (either domestic or international) and provide reimbursement support.
 - We have **sufficient GPUs** (A100 clusters of Westlake, lab-owned A6000, 4090, etc.). Also, we have collaborations *with* industry, which also offers computing support.
 - **For visiting students:** We prefer **on-site** visits (there are very few positions for online visits for exceptional students). Competitive compensation will be provided to help you rent a room near campus and cover most of your daily expenses (such as on-campus dining).

Qualifications / What Kind of Students We Prefer

1. **Strong motivation.** Motivation outweighs everything else.
2. The prospective students are encouraged to be a “三好学生” (“Merit Students”), which is due to the beloved Dr. [Jian Sun](#), when he once gave a talk at Zhejiang University:
 - **Good at math** (“数学好”) - Familiar with the fundamentals of ML (such as linear algebra, calculus, probability theory, statistics, etc.)
 - **Good at coding** (“编程好”) - Strong coding ability with Python, C++ (CUDA programming). Proficient with deep learning platforms like PyTorch / TensorFlow. Familiar with typical deep learning backbones like resnets, transformers.
 - **With a good attitude** (“态度好”) - Good attitude, to me, mostly means
 - **You know why you are here** - Whether it is for dreams, grand research vision, or potential good job opportunities on the market, you *really* think it over and be committed to it.
 - You are **strongly motivated** to explore, hoping to contribute something valuable to the community. As a result, **never give up** when research setbacks happen to you (as they are quite often), and **always stay positive**¹.
 - You act on the **responsibility** to yourself and your colleagues.
3. (Top-tier) Publications are preferred but not required.

How to Apply?

Please follow the instructions exactly - very important.

- 1. Prepare the interview slides, including
 - your basic background: bio/education, GPA, ranking, awards, competitions, etc.;
 - your research experience (not too many) - select the most important one for details, and briefly introduce the rest;
 - your future plan in our lab: what kind of topics you are interested in.
- 2. Fill in [this form](#) about some important questions.
- 3. Send an email² to encodelab@westlake.edu.cn, cc wanghuan@westlake.edu.cn, attaching
 - **your CV** (including, if any, your personal webpage, GitHub link, Google Scholar link, etc.),
 - Pay attention to your **email subject**:
 - For Chinese: 应聘<岗位>-<姓名>-<学校>
 - For others: **Apply for <position>-<your_name>-<your_university>**

Fyi, the position should be one of the 4 available types: PhD/Visiting/RA/Postdoc.
- Note, when you send the email, it means you are ready for the interview. **The interview can come immediately once we receive your email** (this is why you need to prepare your slides in advance).

¹ Okay, being negative for a while is actually [not that bad](#) :-)

² Due to workload, please understand that I may not be able to reply to every email.

- There will be at least two rounds of online interviews:
 - Two rounds: One hosted by the students, the other by the PI (Huan Wang).
 - Either can come first.
 - Duration: short (**10-15mins**) or long (**30-45 mins**). It depends. So, no need to prepare too many slides. Be concentrated.
 - We use **Tencent Meeting** for the interview. We will provide the Tencent Meeting link in your email. Please install it first, and you are encouraged to open the camera during the interview. Check your OS permission ect. in advance to make sure you can open the camera successfully.
 - Both the PI's and the students' opinions will be considered when deciding whether to issue the offer.