

Shenghua He

✉ shenghh2015@gmail.com 🌐 shenghh2015.github.io

WORK EXPERIENCES

- **Senior Research Scientist** *PAII Inc., Palo Alto, CA* Dec. 2021 - current
RLHF/LLM, 3D computer vision, and deep learning
- **Visiting Ph.D. Student** *University of Illinois Urbana-Champaign (UIUC)* Apr. 2019 - Dec. 2021
Computer vision for biomedical image analysis

EXPERTISE & MENTORSHIP

- **LLMs:** RLHF, Search-based PPO, LLM-judge evaluation, SFT, Megatron-Deepspeed, etc.
- **3D Computer Vision:** Self-calibration, 3D object tracking, 3D geometry, image translation, etc.
- **Programming:** PyTorch, Python, Numpy, Transformers, Opencv-python, etc.
- **Others:** Deep learning (DL), machine learning (ML), linux, docker, etc.
- **Mentees:** Hui Wei

PROJECTS

- ☐ **Reinforcement Learning from Human Feedback (RLHF) for Training LLMs** May 2023 - present
Senior research scientist, NLP team, PAII Inc.
 - **Built and Stabilized a Distributed RLHF Training System based on Megatron-Deepspeed**
 - Our team is the **only one in Ping An Group** (a Global Fortune 500 company and the parent company of PAII Inc.) that developed a system supporting multi-node PPO training (open-source ones only support single-node PPO training).
 - The number of parameters in the SFT, reward, and policy models scales up to 7B.
 - **My contributions:**
 - * Managed data to facilitate efficient collaboration among teammates.
 - * Improved reward modeling: iterative data labeling and effective reward losses.
 - * Stabilized PPO training: learning rate, KL coefficient, and reward/advantage normalization, etc.
 - * Extended the system from a single-node version to a multi-node version.
 - * Evaluated and compared widely-used alignment methods: PPO, DPO and best-of-N sampling.
 - **Innovated RLHF Algorithms: Beam Search-Based PPO**
 - Developed an innovative PPO algorithm that uses a beam search strategy to generate responses for policy updates.
 - Conducted a systematic investigation into the effectiveness of beam search in the PPO training phase and obtained encouraging results.
 - The study will be submitted to ICLR 2025.
 - **Led an LLM-as-a-Judge Evaluation Study**
 - Commercial LLMs, such as GPT-4, are frequently used for evaluating LLM alignment methods (referred to as the LLM-as-a-judge method), but their reliability has not been sufficiently investigated.
 - I led this research project to advance the systematic evaluation of LLM-as-a-judge methods.
 - This work has been submitted to the AI Alignment Track of AAAI 2025 [Arxiv paper].
 - **My contributions:**
 - * Designed the research study plan and defined a set of explainable evaluation metrics (accuracy, position bias, and length bias) with improved theoretical explainability.
 - * Led the development of an evaluation and visualization framework for assessing LLM judges.
 - * Led the systematic evaluation of LLM judges using commercial LLMs and diverse prompt templates on frequently-used LLM alignment datasets.
 - **LLM Application: Multi-modality Medical LLM (MMLLM)**
 - Participated in the development of a multi-modality LLM to help doctors efficiently address patients'

queries through vision and text inputs.

- This system has been deployed in China's largest online healthcare service ("Ping An Good Doctor," serving over 440 million users).

- **My responsibilities:**

- * Provided insights into the unusual results generated by the LLMs.

- * Suggested effective solutions to fine-tune the MMLLM and improve the completion rate of successful conversations with patients.

❑ 3D Computer Vision

Dec. 2021 - May 2023

Senior research scientist, computer vision team, PAII Inc.

- **3D Multi-Object Tracking (MOT) from Surveillance Videos**

- Developed a MOT algorithm (PieTracker) using synthetic video data, which wins [3rd-place in CPVR 2022 MOT challenge workshop].

- Independently developed a **novel self-calibration algorithm** that estimate camera parameters, avoiding inconvenient calibration procedures.

- Developed a **3D tracking algorithm** for MOT and generated promising results on MOT benchmark datasets.

- **A Real-Time iPad Detection and Tracking System on Mobile Devices**

- Independently developed an iPad tracking system based on non-deep learning methods, using homography and intrinsic geometry constraints to **improve the tracking performance significantly**.

- The developed system enables robust and real-time iPad tracking despite irrelevant backgrounds.

- **3D Mesh-Guided Image Wrapping for Pixel-to-Pixel Image Translation**

- Developed a shape-to-image translation approach by using 3D mesh correspondence to wrap the shape and learning to map the wrapped shape to the target image.

- The approach generated images of improved qualities compared to conventional end-to-end shape-to-image models.

❑ Computer Vision in Biomedical Image Analysis

Aug. 2016 - Dec. 2021

Ph.D. student in WashU and visiting Ph.D. student in UIUC

- **Object Counting and Object Segmentation**

- Proposed and developed SOTA deep learning models for reliable cell/nuclei counting and segmentation in microscopy images.

- The work were published on prestigious journals, such as *Nature Photonics*, *Nature Communications*, *Medical Image Analysis* etc. and presented in top-tier conferences, such as *SPIE in Medical Imaging*.

EDUCATION

Washington University in St. Louis, St. Louis, MO

Aug. 2016 - Dec. 2021

Ph.D. in Computer Science

Advisors: Prof. Mark A. Anastasio, Prof. Hua Li

Beijing University of Posts and Telecommunications, Beijing, China

Sep. 2012 - Mar. 2015

M.E. in Electronics and Communications Engineering

Advisor: Xiangming Wen

AWARDS

- Honors (top 15-20%), Department of Computer Science, WashU, 2020 - 2021
- National Scholarship (top 0.2%), Ministry of Education, China, 2013 - 2014

PUBLICATIONS

Summary: 12 **journal papers**, 13 **conference papers**, 1 **US patent application**, 10 **abstracts**. The selected publications are below; a full list is shown in [Google Scholar](#).

Pre-prints (* indicates equal contribution)

1. Hui Wei*, **Shenghua He***, Tian Xia, Andy Wong, Jingyang Lin, and Mei Han, "Systematic Evaluation of LLM-as-a-Judge in LLM Alignment Tasks: Explainable Metrics and Diverse Prompt Templates", Submitted to the AI alignment track of *AAAI 2025*.

Journals

1. Xi Chen, Mikhail E. Kandel, **Shenghua He**, Chenfei Hu, Young Jae Lee, Kathryn Sullivan, Gregory Tracy, Hee Jung Chung, Hyun Joon Kong, Mark Anastasio, and Gabriel Popescu, "Artificial confocal microscopy for deep label-free imaging", *Nature Photonics*, 2023. (**impact factor 34.9**)
2. Chenfei Hu*, **Shenghua He***, Young Jae Lee, Yuchen He, Edward Minjae Kong, Hua Li, Mark A. Anastasio, and Gabriel Popescu. "Live-dead assay on unlabeled cells using phase imaging with computational specificity", *Nature Communication*, 2022. (**impact factor 14.7**)
3. **Shenghua He**, Kyaw Thu Minn, Lilianna Solnica-Krezel, Mark A. Anastasio, and Hua Li. "Deeply-supervised density regression for automatic cell counting in microscopy images", *Medical Image Analysis*, 2021. (**impact factor 10.7**).
4. Yuchen He*, **Shenghua He***, Mikhail E Kandel*, Young Jae Lee, Chenfei Hu, Nahil Sobh, Mark A Anastasio, and Gabriel Popescu, "Cell cycle stage classification using phase imaging with computational specificity", *ACS Photonics*, 2022. (**impact factor 6.5**)

Conferences

1. Youbao Tang, Ning Zhang, Yirui Wang, **Shenghua He**, Mei Han, Jing Xiao, and Ruei-Sung Lin, "Accurate and Robust Lesion RECIST Diameter Prediction and Segmentation with Transformers", In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2022.
2. **Shenghua He**, Weimin Zhou, Hua Li, and Mark A. Anastasio, "Learning numerical observers using unsupervised domain adaptation," *Proc. of SPIE Medical Imaging*, 2020.

US Patents

1. Popescu, Gabriel, Mark A. Anastasio, Chenfei Hu, Shenghua He, and Yuchen He, "Complex System for Contextual Spectrum Mask Generation Based on Quantitative Imaging", U.S. Patent Application 17/826,392, filed December 1, 2022.

PROFESSIONAL SERVICES

- **Conference Reviewer:** AAAI, SIGKDD, MICCAI
- **Journal Reviewer:** IEEE Transactions on Neural Networks and Learning Systems, Medical Image Analysis, Pattern Recognition, IEEE Access, Artificial Intelligence in Medicine, IEEE Intelligent Systems, Neurocomputing, etc.