

Yicheng Rui

Ph.D. Candidate in Astronomy, Tsung-Dao Lee Institute, Shanghai Jiao Tong University

✉ ruiyicheng@sjtu.edu.cn Google Scholar  github.com/ruiyicheng X/Twitter

Education

2022–present	Ph.D. Candidate in Astronomy , Tsung-Dao Lee Institute, Shanghai Jiao Tong University Advisor: Prof. Fabo Feng GPA: 4.0/4.0 Expected Ph.D. conferral: June 2027
2018–2022	B.Eng. in Physics , College of Sciences, Shanghai University Advisor: Prof. Xianhui Ge GPA: 3.93/4.0

Research Experience

AI Agent evaluation using scientific problems July 2026–present

- Contributed two astronomy tasks to Terminal-Bench Science: [Variable star vetting](#) and [RV+Astrometry orbital parameter estimation](#). Both tasks were accepted for v0.1.
- Served as a reviewer for Terminal-Bench Science.

Foundation Models and Representation Learning for Light Curves Nov. 2025–present

- Developing foundation-model-style representation learning methods for astronomical light curves, using JEPA-inspired objectives to learn transferable embeddings from time-domain observations.
- Pre-training models with distributed data-parallel workflows on an 8-card V100 machine using the LEAVES dataset.
- Investigating how learned astronomical representations can support downstream classification, anomaly detection, photometric zero-point drift detection, and interpretable scientific decision-making.
- Investigating the potential application of this model on the time series of other disciplines, e.g. finance.

AI Agent for Variable Star Vetting Mar. 2026–present

- Developing a vision-language-model-based AI agent for variable-star vetting, aiming for zero-shot transfer across instruments and human-expert-level diagnostic accuracy.
- Designed an agentic light-curve diagnosis workflow that automatically inspects data schema, identifies time/flux/error/band columns, supports multi-band inputs, and retrieves source context from VSX, LAMOST, and Gaia catalogs.
- Integrated VLM-based morphology assessment with GLS/BLS period searches, phase-folded analysis, recursive zoom-in diagnostics, and automated scientific reporting of classifications, confidence levels, evidence, and follow-up priorities.
- Project page: github.com/ruiyicheng/agentic-light-curve-diagnosis

LenghuSky-8: Cloud segmentation and weather Nowcasting dataset Oct. 2025–Feb. 2026

- Built an eight-year all-sky camera dataset from Lenghu, Qinghai, designed for cloud segmentation, cloud-motion nowcasting, astronomical site monitoring, and autonomous telescope scheduling.
- Performed absolute astrometric and alt-az calibration for all-sky images, enabling physically meaningful spatial representations of cloud and sky structures.
- Constructed star-aware cloud masks and benchmarked segmentation and nowcasting models, including DINO-style representation learning and ConvLSTM-based temporal modeling.

- Designed the dataset and benchmarks to support reusable AI tools for astronomy and model generalization across weather conditions, seasons, observing nights, and sky regions.
- Project page: github.com/ruiyicheng/LenghuSky-8

Scientific yield and Software system of Tianyu project

Jun. 2023–present

Tianyu is a 1-m robotic optical survey telescope located in Lenghu, Qinghai, China, designed for exoplanets, variable stars, transients, and other time-domain phenomena.

Role: full-stack developer for observation control, scheduling, data processing, and survey-analysis systems.

- Estimated exoplanet and transient yields for the Tianyu telescope.
- Designed and implemented scalable software architecture for Tianyu data processing, including modular photometry components, workflow orchestration, and database-backed management of high-throughput time-domain observations.
- Built a MySQL + RabbitMQ based distributed DAG system for large-scale, fault-tolerant image calibration, alignment, flux extraction, and light-curve production.
- Developed a relative-photometry pipeline tested on real telescope data. The extracted light curves approach the theoretical noise limit and support minute-cadence observations for transiting exoplanets and variable stars.
- Contributed to survey strategy and scientific-yield estimation for Tianyu’s exoplanet, transient, and variable-star programs.
- Project page: tianyu.sjtu.edu.cn
- Media Report: [Global Times](#); [China News](#)

DeepAP: Deep Learning for Fast Aperture Photometry

Nov. 2024–May 2025

- Proposed and contributed to DeepAP, a two-stage deep-learning framework for aperture photometry, combining a Vision Transformer for feasibility assessment and a ResNet for aperture-size prediction.
- Reframed a classical astronomical measurement problem as a fast learned inference problem, enabling approximately $59,000\times$ speed-up over exhaustive aperture-size enumeration.
- Demonstrated how learned models can accelerate scientific pipelines while preserving measurement quality, making them suitable for large-scale time-domain surveys.
- Software: github.com/ruiyicheng/DeepAP

AstroSkyFlow: Astronomical Image Simulation

Nov. 2024–Jun. 2026

- Leading the development of an end-to-end sky-image simulator for transient events, combining physical event modeling, instrument effects, background sky, and observational noise.
- Building large-scale synthetic datasets for transient-detection and representation-learning tasks, with the goal of improving data efficiency and robustness for rare astronomical events.

Mitigating Astrometric Bias in Barycentric Correction

Dec. 2021–Jun. 2023

- Developed a gradient-based method to correct biases in stellar astrometric parameters for high-precision barycentric correction of radial-velocity time series.
- Demonstrated the importance of modeling astrometric systematics for ultra-precise radial-velocity searches for Earth-twin exoplanets.

A Universal Brown Dwarf Desert Formed Between Planets and Stars Nov. 2023–Feb. 2026

- Proved that a discrete grid-based Bernoulli occurrence model converges to an inhomogeneous Poisson point-process likelihood in the infinitesimal-bin limit.
- Derived the corresponding occurrence-rate likelihood, including the exponential contribution from empty parameter-space regions and the local density terms from detected objects.
- Used the expectation–maximization (EM) algorithm to initialize the parameters of the brown-dwarf subpopulations for the subsequent population analysis.

Thermodynamic Properties of 2D JT Gravity and the SYK Model Feb. 2020–Jun. 2021

- Numerically computed complex-time Green functions and studied thermodynamic properties of 2D Jackiw–Teitelboim gravity and the Sachdev–Ye–Kitaev model.
- Estimated critical exponents and developed experience in numerical analysis, scientific computing, and large-scale simulations.

Industry Experience

Quantitative Research Intern, Berkeley Brothers Feb. 2026–Apr. 2026

- Modeled financial time series using machine-learning techniques, with emphasis on statistical signal extraction, robust validation, and scalable analysis workflows. (Evaluated the transferability of a light-curve foundation model to financial time-series forecasting, identifying limitations in my given task.)

Technical Skills

Machine learning	Distributed data-parallel training, Vision Transformers, ResNet, DINO-style representation learning, ConvLSTM, CLIP/JEPA-style learning, vision-language models, diffusion models, deep reinforcement learning, Bayesian statistics, Language models
Foundation models	Representation learning, transfer learning, self-supervised learning, model evaluation, generalization across instruments/datasets, scientific data curation, synthetic data generation
Astronomical data analysis	Time-series analysis, relative photometry, aperture photometry, astrometry, light-curve extraction, survey scheduling, telescope data processing, camera characterization
Software and systems	Python, C/C++, Shell, MATLAB, Mathematica, Linux, Git, Docker, Slurm, SQL/MySQL, RabbitMQ, distributed workflows, data pipelines, high-throughput computing
Instrumentation	Optical camera testing, collimator-based measurement, integrating sphere calibration, optical power measurement, observatory control systems

Presentations

2026	Shanghai University Seminar	<i>AI for Accelerating Discovery in Astronomy</i>
2026	SAM Seminar	<i>The application of AI Agents in Astronomy</i>
2025	Tianyu Workshop II	<i>Lessons Learned from Developing the Tianyu Pipeline</i>
2025	Stellar Physics Guest Lecture	<i>History, Naming, and Classification of Variable Stars</i>
2025	SAM Seminar	<i>Scientific Research with SAM 1-meter Telescope</i>
2024	Qinghai University Seminar	<i>Exoplanet Detection Methods</i>
2023	Tianyu Workshop I	<i>Photometry Accuracy and Survey Strategy of Tianyu</i>
2023	EPRV5	<i>A Systematic Reduction of Earth Motion from Radial Velocity</i>

Teaching

2020	Analytical Mechanics in Two Hours	Lecturer	Delivered an online lecture introducing the fundamentals of analytical mechanics.
2020–2021	Teach You to Play with AI	Lecturer	Taught a two-semester course for local high school students on the fundamentals of Python programming and AI.
2022–2023	Classical Mechanics	TA	Served as a teaching assistant for an undergraduate classical mechanics course.
2023	Physical Experiments	Lecturer	Guided undergraduate students in experimental physics coursework.
2023	Gaia Astrometric Error Analysis	Co-supervisor	Co-supervised undergraduate students on a research project.
2024–2026	From Earth to Extraterrestrial Civilization	TA	Designed hands-on problems and supported course instruction.
2025	Stellar Physics	Guest Lecturer	Delivered a guest lecture on variable stars.
2025	Flat-field Optimization	Co-supervisor	Co-supervised an undergraduate thesis.
2025	Deep Learning for Aperture Photometry	Co-supervisor	Co-supervised a master's thesis.

Selected Honors and Awards

2026	National Scholarship
2024	Guo Xie Birong Scholarship, Tsung-Dao Lee Institute
2024	Tsung-Dao Lee Ph.D. Student Scholarship
2023	Outstanding Teaching Assistant, Shanghai Jiao Tong University
2020	National Scholarship
2019, 2021	National First Prize, China Undergraduate Mathematical Contest in Modeling
2019	National Champion, China Robot Competition
2019	National Champion, RoboCup@Home
2018	World Champion, RoboCup Junior Maze

Language

TOEFL iBT: 104 Reading 30, Listening 26, Speaking 24, Writing 24.