

Chenyu Wang

✉ wangchenyu0628.wcy@gmail.com ☎ +86-18292065293

🏠 <https://wangchenyu.uno>



RESEARCH PROFILE

I am an M.S. candidate at the University of Chinese Academy of Sciences, based at the Shanghai Institute of Microsystem and Information Technology, and a Multimodal Agent Research Intern at Shanghai Artificial Intelligence Laboratory. My research focuses on **large language models and AI agents for healthcare**. I am the first author of manuscripts on guideline-supervised ophthalmic triage and sensor-domain adaptation, currently under review at **ICLR 2027** and **IEEE Sensors Journal**, respectively. My ongoing work covers retrieval-augmented oncology decision support and agent-tool risk attribution. As the **sole participant and project lead** of `ucas_xlab` at the ECCV 2026 SLoMO Workshop competition, I placed **2nd in AD Special** and **3rd in AD Main**.

EDUCATION BACKGROUND

University of Chinese Academy of Sciences (UCAS) Sep. 2024 – Jun. 2027 (expected)
M.S. in Electronic Information Shanghai, China
Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences
Advisors: Prof. Baoqing Li and Prof. Min Tu

Xi'an University of Technology Sep. 2019 – Jul. 2023
B.Eng. in Electrical Engineering and Automation Xi'an, China
School of Electrical Engineering; Advisor: Assoc. Prof. Kaiyan Wang

INTERNSHIP EXPERIENCE

Shanghai Artificial Intelligence Laboratory Mar. 2026 – Present
Multimodal Agent Research Intern, Frontier Multimodal Research Shanghai, China

Research Highlight 1: Guideline-Supervised Clinical Agents

Guideline-as-Oracle: Ophthalmic Telephone Triage. Mar. 2026 – Aug. 2026.

First-author manuscript under review at ICLR 2027; earlier preprint: arXiv:2608.04772.

- Developed a guideline-to-dialogue pipeline to address the cost of annotating multi-turn clinical conversations, converting a **70-row ophthalmic rule table into 3,000 training dialogues** with rule-derived urgency labels.
- Performed **full-parameter supervised fine-tuning of a 9B model**, using class resampling and nurse-role token supervision. Also investigated Outcome GRPO as an exploratory post-training approach.
- In the prespecified development-reference comparison with the base 9B model, improved agreement from **61.7% to 74.1%** and emergency recall from **9.5% to 69.0%** on **201 internal rule-reference cases**. Additional seed and simulator evaluations showed variable agreement gains. This reference also informed development and model selection, so these results measure internal policy conformance rather than held-out clinical generalization.

Research Highlight 2: Evidence-Grounded Clinical Decision Support

Retrieval-Augmented mCRC MDT Decision Agent.

Shanghai Artificial Intelligence Laboratory. Ongoing. Aug. 2026 – Present.

- Investigating how longitudinal records and retrieved cases support prediction of **five recorded MDT management actions** for metastatic colorectal cancer. Built a single-center data pipeline from **9,779 encounters / 4,983 patient records** spanning 2013–2025; integrity screening retained **8,274 encounters / 4,077 patients**. Restricted inputs to information available before each target meeting.
- Froze a test set of **1,247 encounters from 610 patients**, with no patient overlap across splits. Compared classical machine learning, retrieval, and language-model approaches using a common evaluation protocol.
- Combined **longitudinal history, similar-case retrieval, and XGBoost probability priors** in a patient-time-aligned language-model prompt. Achieved **64.1% accuracy and 0.527 macro-F1**, compared with **57.3% and 0.453** for the strongest classical baseline: gains of **6.8 percentage points in accuracy and 0.074 in macro-F1**.
- For encounters with prior patient history, accuracy increased from **55.4% to 66.1%** relative to the classical model. Current results measure retrospective agreement with recorded MDT actions; prospective clinical benefit remains untested.

DSH: Component-Level Static Auditing and Risk Attribution

Shanghai Artificial Intelligence Laboratory. Ongoing. Aug. 2026 – Present.

- Developing dependency-closure attribution and four-level evidence scoring across **12 security domains** for agent software components. Coverage includes **207 repositories, 12,830 components, and 11,124 entry points**. The method distinguishes shared-dependency findings from risks specific to individual components.
- Combined AST parsing with intra-procedural taint analysis to reduce items requiring manual review from **4,794 to 153**; the remaining 153 represent **1.4% of all 11,124 entry points**. Consolidated findings into **24 code defects**. Two clean-environment reruns produced byte-identical results with fixed source and parser versions.

■ RESEARCH EXPERIENCE

Shanghai Institute of Microsystem and Information Technology, CAS

Graduate researcher; advisors: Prof. Baoqing Li and Prof. Min Tu

UnifiedGas: Domain Adaptation for Sensor Drift.

First-author manuscript under review at *IEEE Sensors Journal*.

Jul. 2025 – Mar. 2026

Shanghai, China

- Developed a single-stage unsupervised adaptation framework for temporal drift and cross-device distribution shift, combining hierarchical MK-MMD, CORAL, dual-level center regularization, and multi-head classification; also studied a relative-response front end and the conditional adversarial variant UnifiedGas-C.
- On UCI Drift, obtained **81.10% / 84.36%** mean accuracy for fixed-source / sequential transfer using **retrospective target-label-based best-epoch selection**.
- On the **full-cycle Twin Arrays track**, fixed-source / adjacent-board accuracy was **100.00% / 100.00%** for UnifiedGas-C and **100.00% / 99.94%** for CDAN under the same frontend-matched protocol.

■ PUBLICATIONS AND PREPRINTS

- [1] **Chenyu Wang**, et al. Guideline-as-Oracle: Rule-Derived Labels Teach a Triage Agent When to Escalate. *ICLR 2027*, under review, first author. Earlier preprint: arXiv:2608.04772.
- [2] **Chenyu Wang**, et al. UnifiedGas: End-to-End Unsupervised Domain Adaptation for Drift-Robust Gas Classification. *IEEE Sensors Journal*, under review, first author.

■ FUTURE RESEARCH INTERESTS

I am interested in applying large language models and AI agents to medicine and exploring how they can reshape the way AI is used in healthcare. Building on my experience in medical dialogue modeling and clinical decision support, I hope to develop intelligent systems that support more interactive, adaptive, and reliable clinical workflows.

■ COMPETITIONS

ECCV 2026 SLoMO Workshop. Sole participant and project lead (registered as ucas_xlab). Independently developed systems for audio description (AD) and movie question answering (MovieQA), covering temporal context, visual evidence, and candidate selection. Official final rankings.

- **AD Special: 2nd place, Private AD Score 49.80.** Developed a zero-shot Qwen3.5-35B-A3B pipeline combining dense frame sampling, temporal context, and time-budget-aware descriptions.
- **AD Main: 3rd place, Private AD Score 53.47.** Designed a description-refinement pipeline using baseline anchors, visual-evidence checks, and blind dual-judge selection.
- **MovieQA Main: 4th place, Private Accuracy 64.40%.** Built a movie question-answering pipeline combining multi-scale temporal sampling, subtitle/ASR alignment, cross-clip evidence retrieval, and model consensus.
- **MovieQA Special: 5th place, Private Accuracy 51.02%.** Combined 8/64/128-frame views from a frozen Qwen3-VL-8B model with weighted Jaccard consensus.

■ AWARDS & ACHIEVEMENTS

- | | |
|--|------------------|
| • Merit Student, University of Chinese Academy of Sciences | 2026 |
| • First-Class Scholarship, Xi'an University of Technology | 2020, 2021, 2022 |
| • Third Prize for Innovation Achievement, Xi'an University of Technology | 2022 |
| • Outstanding Student Cadre, Xi'an University of Technology | 2020, 2021, 2022 |

■ SKILLS

Programming and modeling: Python, PyTorch, scikit-learn, XGBoost; full-parameter SFT, GRPO exploration, retrieval-augmented generation, multimodal reasoning, and unsupervised domain adaptation.

Research engineering: clinical data curation, patient/time-aware evaluation, source ablations, reproducible pipelines, model calibration, AST parsing, and static taint analysis.