

陈宇航

电话: (+86) 134-2812-2178

邮箱: thisiscyh@gmail.com

政治面貌: 群众

籍贯: 广东省湛江市



教育经历

华中科技大学 计算机科学与技术 大三 2023.9 - 至今

- GPA:** 本科阶段: 4.42/5.0 (Top 5.7%)
- 研究兴趣与技能:** 聚焦世界模型、数据挖掘、多模态智能、模型推理与可解释人工智能等方向; 熟练使用 Python、C/C++ 及 PyTorch, 具备评测基准构建、多模态理解与生成、多模态大模型与智能体系统开发、推理机制分析、代码生成及学术系统原型开发能力
- 主修课程:** 算法分析与设计、线性代数引论、概率论、人工智能原理、机器学习、计算机视觉、计算机组成原理、计算机网络、编译原理
- 荣誉/奖项:** 以共同一作参与的论文发表于 **IEEE 国际计算机视觉与模式识别会议 CVPR 2026**, 以共同二作参与的论文发表于 **知识发现与数据挖掘会议 KDD 2025**, 以第一作者身份发表论文在投于 **CCF A, 学习优秀奖学金, 科技创新奖学金, 2025 全球人工智能算法精英大赛人工智能应用赛道国赛一等奖, 2025 年睿抗机器人开发者大赛智海人工智能算法应用国赛二等奖, 2025 年全国大学生计算机系统能力大赛-数据库管理系统设计赛国赛三等奖**,

科研经历

Research Intern in University of Maryland 2025.3 - 2025.10

- 在 **Prof. Tianyi Zhou** 教授指导下开展研究, 主要关注 **高效推理、可解释推理与以用户为中心的智能系统**。
- 研究推理模型中的思维模式与推理 token 的有效性, 探索如何在保持性能的同时提升推理过程的效率与可解释性。
- 研究面向用户的 AI 系统设计, 探索大语言模型如何为用户提供更多可选方案与启发式辅助。
- 参与统一多模态模型“理解-生成鸿沟”课题, 提出评测框架与量化指标, 并通过系统实验分析该鸿沟背后的机制。

华中科技大学 **ONE Lab** 科研实习生 2024.10 - 2026.2

- 在 **万瑶** 教授指导下开展研究, 主要聚焦 **多模态智能体、多模态评测与学术场景代码生成**。
- 参与 **Paper2Web** 项目研究, 提出基准数据集与自主智能体系统 **PWAgent**, 利用 MCP 驱动工具链将静态 PDF 转化为交互式、布局感知的项目主页。
- 参与 **Judge Anything** 项目研究, 构建统一对称的评测框架, 用于评估 MLLM 作为评审器在 15 类任意模态到任意模态任务中的表现。
- 参与多模态智能体研究, 探索将 MLLM 与多步推理、工具调用和自主工作流结合, 以提升模型在复杂开放式跨模态任务中的能力。

学术成果

Quantifying the Gap between Understanding and Generation within Unified Multimodal Models
CVPR 2026 Findings [PDF] 2025.11-2026.2

- 面向统一多模态模型 (UMM) 在“理解”和“生成”能力之间可能存在割裂的问题, 开展系统性研究, 探索模型跨任务认知一致性的量化方式。
- 参与设计双向评测框架与量化指标, 从理解到生成、生成到理解两个方向刻画模型能力差异, 分析统一多模态模型内部跨模态知识表征的不一致现象。
- 通过多组实证实验研究“理解-生成鸿沟”的表现形式及其潜在成因, 为统一多模态模型的能力评估与机制分析提供了新的视角。

Judge Anything: MLLM as a Judge Across Any Modality
KDD 2025 Datasets and Benchmarks Track [PDF] 2024.11-2025.2

- 面向文本、图像、音频、视频等多种模态生成任务, 研究多模态大模型作为评审器 (MLLM-as-a-Judge) 在跨模态评测场景中的能力边界与适用性。
- 参与构建统一评测框架与基准任务, 覆盖 15 类任意模态到任意模态任务, 系统评估模型在跨模态生成结果判断中的一致性、鲁棒性与泛化能力。
- 参与搭建自动化评测平台, 探索缓解多模态评测过程中偏差与幻觉问题的方法, 为跨模态生成任务建立更统一的评测标准。

PAPER2WEB: LET' S MAKE YOUR PAPER ALIVE!
In submission [PDF] 2025.7-2025.11

- 面向学术论文线上展示形式单一、交互性不足的问题, 研究如何将静态论文 PDF 自动转化为交互式、布局感知的学术项目主页。

- 参与设计基准数据集与自主智能体系统，结合工具调用、页面结构解析与迭代优化机制，实现论文内容抽取、页面布局还原与前端页面自动生成。
- 探索智能体系统在学术场景代码生成中的应用，提升论文内容展示的可访问性、交互性与传播效率。

竞赛经历

- 睿抗机器人开发者大赛 (RAICOM)** 2024.9-2024.12
 - 荣获 2025 年睿抗机器人开发者大赛智海人工智能算法应用国赛二等奖
- 全球校园人工智能算法精英大赛** 2025.9-2025.12
 - 荣获 2025 年第七届全球校园人工智能算法精英大赛算法应用国赛一等奖
- 2025 年全国大学生计算机系统能力大赛** 2025.7-2025.8
 - 2025 年全国大学生计算机系统能力大赛-数据库管理系统设计赛国赛三等奖

项目经历

- Paper2Web**: 开发了 PWebAgent，这是一个利用 MCP 驱动工具和迭代优化机制的自主系统，可将静态 PDF 转换为具有交互性和布局感知能力的学术主页。[[代码仓库](#)]
- Judge Anything**: 提出了 TaskAnything 和 JudgeAnything 两个基准，用于评估多模态大模型作为评审器 (MLLM-as-a-Judge) 在 15 类任意模态到任意模态任务中的表现；开发了 OmniArena 自动化平台，以缓解多模态生成 (MMG) 任务中的评测偏差与幻觉问题。(2025 年 KDD 论文的数据集与评测平台)
- GapEval**: 提出了一个双向评测基准，用于量化统一多模态模型 (UMMs) 在理解与生成之间的认知一致性，并通过知识操控实验揭示了跨模态知识存在割裂的现象。(2026 年 CVPR 论文的数据集)
- RISC-V CPU 设计**: 在 Logisim 上搭建电路，基于 RISC-V ISA 设计 CPU 功能，实现了五级指令流水线、动态预测与重定向机制。
- MiniMind 实现**: 基于 PyTorch 复现 MiniMind，完成模型搭建、训练与推理流程实现，理解了大语言模型在表示学习、自回归生成与训练优化方面的核心机制。

其他经历

- 英语成绩**: 已通过 CET-4、CET-6

个人主页: <https://yuhangchen1.github.io/>,