

何文彬

硕士 | 男 | 中共党员 | 籍贯: 湖南 | 出生日期: 2000-11
研究方向: 多模态学习、计算机视觉、视觉语言模型、多模态大模型
☎ (+86) 182 2931 7520 ✉ wbhe00@gmail.com
📍 个人主页: <https://bbihh.github.io/>



🎓 教育经历

中央民族大学(985) 民族语言智能分析与安全治理 教育部重点实验室 2022年09月 – 2025年06月

硕士 信息工程学院 软件工程 成绩排名: 2/14

导师: 翁彧 教授 中央民族大学信息工程学院院长

毕业论文题目: 《基于多模态语义适配的跨语言图文问答方法研究》

主要研究多模态相关方向, 包括图文问答、图文检索、文生图/视频等内容, 发表多篇CCF推荐国际会议/期刊论文。参与国家重点研发项目的申报、结项工作骨干成员, 并作为核心开发人员完成多项横向课题项目。

南华大学

2018年09月 – 2022年06月

本科 计算机学院 软件工程(卓越计划班) 成绩排名: 9/442

毕业论文题目: 《基于视觉的智能辅助驾驶系统设计与实现设计》

主要研究计算机视觉方向, 包括目标检测、多目标分类等内容。多次获得奖学金、三好学生, 年级成绩前2%。以队长身份组织参加各种学科竞赛, 获得多项国家级奖项, 主持国家级大学生创新创业项目。

🔬 科研成果

▶ **第一作者论文 3 篇:** 一作论文 3 篇, 均为CCF会议/SCI期刊。

1. (共一作者, IEEE Trans., 中科院二区, CCF-C) Wenbin He*, et al. Zero-Shot Cross-Lingual Knowledge Transfer in VQA via Multimodal Distillation. // IEEE Transactions on Computational Social Systems, TCSS. doi: 10.1109/TCSS.2024.3402270. **已发表**. (视觉问答、迁移学习)
2. (第一作者, CCF-C) Wenbin He, et al. Enhancing Text-Image Person Re-identification via Intra-Class Relevance Learning. // Proceedings of the 24th International Conference on Algorithms and Architectures for Parallel Processing, ICA3PP 2024. **已接收**. (图文检索、对比学习)
3. (学生一作, IEEE Trans., 中科院一区 Top, CCF-B) Wenbin He*, et al. Cross-Lingual Adaptation for Vision-Language Model via Multimodal Semantic Distillation. // IEEE Transactions on Multimedia, TMM. **大修**. (多模态基模型适配、迁移学习)

▶ **共同作者论文 3 篇:** 参与论文包括AIGC、安全方面研究方向。

1. (共同作者) Lidu Lou, Wenbin He, Xiangyun Tang, et al. Enabling Trustable Financing A Verifiable Privacy-preserving Cross-chain Protocol. // 2024 IEEE/CIC International Conference on Communications in China, ICCIC 2024. **已接收**. (区块链相关)
2. (共同作者, CCF-B) Guanfu Chen*, Xiao Liang*, Wenbin He. Enhancing Human-Computer Interaction through Decoupling Motion and Camera Control in Human-Centric Video Generation. // International Journal of Human-Computer Interaction, IJHCI. **已投在审**. (文生视频相关)
3. (共同作者, 中科院一区) Yu Weng, Kunyu Yang, Zheng Liu, Wenbin He, Xiangyun Tang. LGVLM-mIoT: A Lightweight Generative Vision-Language Model for Multilingual IoT Applications. // IEEE Internet of Things Journal, IoTJ. **已投在审**. (多模态应用相关)

▶ **发表专利 1 项:**

1. 翁彧, 邢天娇, 唐湘云, 何文彬, 等. 一种基于中继链的可验证隐私保护跨链系统. 公开号: CN117335955A. 第四作者, 已公开.

□ 纵向科研项目研究

- ▶ 参与国家重点研发项目 2 项，参与省级重点实验室申报 1 项，省部级重点研发项目 2 项：
1. 支持异构多链互通的跨链系统研发与应用（国家重点研发计划子课题）2020.11 – 2024.6

项目骨干成员，负责课题的组织、管理、开发，重点参与项目整个结项过程。

- 项目介绍：该项目针对目前国内外主流区块链系统的兼容性、互操作性、安全性进行研究，提出了一种创新的可插拔的“跨链即服务”平台，提高供应链金融运转的透明度和效率。

- 参与内容：负责项目中供应链金融场景下的高效跨链方案设计，参与区块链跨链系统的设计与测试，并整理编写项目相关开发文档。相关论文成果已发表至 ICCC 国际会议。
2. 基于新闻报道场景的 AI 辅助写稿机器人系统研发（国家重点研发计划子课题）2021.9 – 2023.10

项目骨干成员，参与项目研究中图文模态对齐相关技术研究。

- 项目介绍：该项目针对新闻事件素材的多文档摘要生成任务，从事件脉络中检测并构建出重要事件脉络，高度凝练素材内容生成事件摘要，实现面向完备数据的事件脉络摘要生成。

- 参与内容：在项目中主要研究多模态对齐技术，针对新闻素材的图文内容做进一步加工，在事件脉络构建中融入图像注释任务以及视觉问答任务辅助上下文内容关联。
3. 海南陵水黎安国际教育创新试验区重点数字实验室（海南省重点实验室申报）2024.3 – 2024.9

项目骨干成员，参与海南重点数字实验室申报工作。主要负责实验室研究内容的数字底座和多模态数字资产生成技术路线调研与搭建。主要工作包括实验室研究内容框架搭建，数字资产流通、加工和生成相关内容的撰写与整合，以及申报文档和PPT的编制与整理。同时辅助管理申报流程进度和文档质量控制，为项目申请提供全面技术支持。
4. 多模态数字资产智能生成关键技术研究（海南省重点研发项目）2024.1 – 2026.12

项目骨干人员，负责其中课题一“跨源数据融合与领域知识图谱构建方法研究”的研究框架和技术路线。主要聚焦于多模态融合技术的研究与应用，负责设计针对多模态信息的知识图谱构建框架，提高实体对齐精度和知识图谱的丰富度。编写详细的研究内容文档，以及绘制系统框图。
5. 规模化开放式图像敏感内容的精细注解关键技术研究（自然科学基金青年基金项目）2024.1 – 2026.12

主要参与人员，研究计算机视觉与多模态交互，聚焦于图像信息脱敏技术。项目旨在解决大规模数据中敏感内容的精细处理，设计人机交互式脱敏框架，涉及跨图像关联和多模态感知。
- ▶ 个人主持科研项目 2 项：
1. 基于迁移学习的多语言视觉问答研究（硕士研究生自主科研项目）2023.6 – 2024.3

该项目旨在解决多语言人工智能系统中的跨语言理解挑战，特别是在视觉问答（VQA）领域。在该项目中设计了一套跨语言多模态知识迁移框架，该框架通过代码混合跨语言掩码建模和多模态知识蒸馏方法，实现了英语视觉问答模型知识向其他语言的高效迁移。相关成果发表在 IEEE 期刊。
2. 基于计算机视觉与行为分析的智能行车安全辅助应用（国家级大学生创新创业项目）2021.6 – 2022.6

模型优化：基于YOLOv5模型，调整CSP结构增加残差注意力机制，提升了模型对不同尺寸车辆与道路交通物体的检测能力，达到78%识别率；数据优化：使用数据增强技术扩充训练数据集，增强了模型在不同光照和天气条件下的鲁棒性；方案优化：在行车场景加入安全场景分类，加入额外轻量模型识别分类道路危险行为，辅助危险道路行为的判断，增强行车辅助应用在复杂场景下的准确性。

▣ 横向项目经历

- 华为鲲鹏众智项目-OpenStack云管平台性能优化项目

核心开发人员

2022年09月 – 2024年03月
- ▶ 自动化测试脚本设计：使用Intel 6448、海光7485和鲲鹏7260服务器，搭建OpenStack环境，设计测试用例，使用Python编写自动化测试脚本获取基线数据，需求实现率100%，遗留缺陷密度为零；
- ▶ 云管平台性能优化：负责对OpenStack的大规模虚拟机并发创建、启动、关闭及删除功能进行深度优化，用C++对关键组件进行重新设计编码，显著提升其在鲲鹏服务器的欧拉操作系统上的性能；分析该平台批量管理虚拟机的性能瓶颈，采用缓存调度器，缓存主机信息，总体优化调度时间提升23%。

- ▶ 数据库插件开发：TimescaleDB是一个专为时序数据设计的开源数据库插件，建立在PostgreSQL的基础上，项目主要负责将Timescale-DB数据库扩展功能适配到openGauss数据库管理系统上，以解决大规模时序数据处理问题，开发出同样功能的数据库插件；
- ▶ 数据库内核优化：修改openGauss内核代码以提供计划树钩子函数，使openGauss能够精确地按照TimescaleDB的计划树生成数据块，从而提升数据库在处理海量数据插入和查询时的性能。通过改变计划树，数据插入速度提升2.6倍，数据查询速度提升16倍。

- ▶ 古建筑数字化重建：针对建筑数字化要求，使用点云扫描技术建立建筑数字内容，调整数据采集的精度要求、数据库建模方法等，确保数据的完整性和准确性；消除扫描点云数据中存在的噪声点，提升数据质量。

🏠 科研工作

- Conference: 5th EAI International Conference on Data and Information in Online Environments (EAI DIONE 2024) Workshops Chair
- Conference: 8th EAI International Conference on Machine Learning and Intelligent Communications (EAI MLICOM 2023) TPC Member, Reviewer

🏆 获奖情况

- ▶ 学科竞赛类：获得国家级奖项 3 项，省级奖项 5 项，校级奖项若干。
 - (国家级三等奖) 中国大学生服务外包创新创业大赛全国总决赛三等奖
队长身份，比赛题目：智能网联汽车辅助驾驶安全信息检测系统
 - (国家级三等奖) 蓝桥杯程序设计大赛C/C++赛道全国总决赛三等奖
 - (国家级三等奖) 全国大学生软件测试大赛全国总决赛三等奖
 - (省级一等奖) 湖南省第17届计算机程序设计大赛-应用开发类一等奖
队长身份，比赛题目：车行眼-基于目标检测与行为分析的驾驶安全辅助系统
 - (省级三等奖) 湖南省第16届大学生计算机程序设计大赛-ACM竞赛三等奖
 - (省级三等奖) 中国大学生计算机设计大赛中南地区赛三等奖
队长身份，比赛题目：行车之眼-基于计算机视觉的智能驾驶辅助系统
- ▶ 奖学金类：研究生期间获特等奖学金、三等奖学金等；本科期间获得三好学生、二等奖学金等。

🔧 个人技能

- 🔧 编程能力：主要使用编程语言Python、C/C++，熟练使用PyTorch深度学习框架，熟练使用Latex写作。
- 🔧 绘图技能：主要使用PPT绘图，熟练使用Visio、Matlab绘图。
- 🔧 外语水平：通过英语六级，具有良好的英文听说读写能力。
- 🔧 其它方面：羽毛球、健身、游泳初学者；团队建设，人际交往。

🇨🇳 自我评价

- 🔧 专业背景：具备扎实的计算机知识与较强的学习能力，关注最新技术发展，不断提高专业水平。
- 🔧 综合素质：具备较好的沟通协调和团队协作能力，能吃苦耐劳，注重个人品德修养和职业道德规范。
- 🔧 研究兴趣：计算机视觉CVer -> 多模态MMer，当前研究兴趣领域：多模态融合与对齐、视觉语言模型、多模态大模型、多模态理解与推理。
- 🔧 科研愿景：人生难免迷茫，但我对人工智能的热情始终如一。我渴望亲眼见证并参与人工智能的演进，哪怕只是推动了一小步。即使个人的力量有限，但也希望能为这个世界带来一些真正有意义的改变。