方悦

电话：18938813098 | 邮箱：1206137707@qq.com | 籍贯：广东潮州

个人主页：https://yuefang0211.github.io | 英文：CET6（596）, IELTS（7）

编程：Python，Vibe Coding（Codex,，Claude，Cursor），Harness，Linux



教育背景

东南大学（985） **网络空间安全学院（学硕）** **可信人工智能安全课题组** **2025.09-2028.06**

规格化 GPA：82.7 (前 8%) 荣誉：二等学习优秀奖学金

北京交通大学（211） **电子信息工程学院（本科）** **信息工程** **2021.09-2025.06**

GPA：89.0 (前 8%) 荣誉：二等学习优秀奖学金，校三好学生，二等社会工作奖学金，校优秀团干，校优秀学生干部

实习经历

字节跳动 | 北京 | Seed Model - Robotics | 数据算法实习生 **2026.09-2026.11**

- 多样性造数：设计四阶段标签算子，实现长程交互数据切片，产出技能、物品属性等 10 余类机器人数据字段；
- Physical VLM 造数：为 VLM 提供物理空间与动力学估算、局部意图理解等监督信号，培养模型空间感知与行为决策能力；
- 数据 Infra：参与数据平台算子研发，面向 RUA、VLA 等场景提供多源异构数据入湖、处理与打包等服务。

蔚来 | 北京 | 自动驾驶研发数据算法部门 | 数据算法实习生 **2025.02-2025.06**

- 强化学习：对比验证强化学习相较于 SFT 在红绿灯及交通设施检测、路径规划等驾驶任务中的效果优势；
- 数据构建：围绕模型训练需求构建高质量驾驶数据，补充长尾场景数据、CoT 推理数据及多维度扩展数据；
- VLA 预研：实验验证 RT 系列、π_0系列等在驾驶领域的能力。

酷哇科技 | 上海 | 自动驾驶算法部门 | 数据算法实习生 **2026.04-2026.06**

- Data Harness：构建面向驾驶行为的闭环训练数据优化链路，实现训练数据动态筛选与最优配比，提升模型准确率至 80%；
- World Model 预研：探索使用 V-JEPA2+FM 架构实现驾驶行为预测可行性。

清华大学车辆与运载学院智能出行所 | 北京 | 自动驾驶视觉感知课题 | 联合培养实习生 **2024.05-2024.11**

- 面向自驾在未知场景下的目标检测精度及灾难性遗忘问题，提供一种基于增量式学习的感知模型目标检测方法；
- 项目已转化为专利《一种基于增量式学习的感知模型的目标检测方法及装置》，申请号 CN202411715425。

科研成果

多模态大语言模型幻觉的检测与量化评估 (IMAVIS (中科院二区)，优秀毕业论文) | 第一作者 **2024.09-2025.06**

- VLM 幻觉量化与语义处理：针对 VLM 在复杂场景易产生的空间感知与属性判断等幻觉，对模型输出自然语言进行横纵双向语义处理，提出基于不确定性量化的零参考黑盒评估框架，在复杂开放场景下的实现 VLM 输出置信度的精准量化。

可信多智能体系统通信幻觉缓解 (KAIS (CCF B) 在审) | 第一作者 **2025.09 至今**

- 多智能体信息级联建模：针对多智能体系统显隐式错误信息级联放大问题，建模多机协同场景下信息的时空演化范式；
- 通信拓扑优化：结合强化学习与动态掩码机制构建通信网络，引导智能体在通信过程中解耦错误信息，实现差异化协同。

面向大语言模型多智能体系统的对抗性语义伪装 (JSSOFTWARE (CCF B) 在审) | 第四作者 **2025.10-2026.10**

- 针对多智能体系统关键节点易受攻击问题，设计语义伪装模块与动态诱饵节点，为多机协作场景提供可扩展防御范式。

校园经历

牛津大学奥利尔学院 | 暑期交流 | 交流内容包括 AI 算法、Symbolic AI、AI 安全等，获 A+评分 **2026.07-2026.08**

东南大学艺术团舞蹈团 | 成员 | 参演东南大学跨年演出、毕业专场等 **2025.09 至今**

北京交通大学合唱团 | 声部长 | 一类一等文艺活动优秀奖学金，大艺展银奖 **2021.09-2024.08**

BAS 跨国支教 | 优秀个人，获国际志愿者证书 **2023.01-2023.02**