

同行专家业内评价意见书编号： 20250854456

附件1

浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名： 曹世聪

学号： 22260052

申报工程师职称专业类别（领域）： 电子信息

浙江工程师学院（浙江大学工程师学院）制

2025年06月05日

填表说明

一、本报告中相关的技术或数据如涉及知识产权保护、军工项目保密等内容，请作脱密处理。

二、请用宋体小四字号撰写本报告，可另行附页或增加页数，A4纸双面打印。

三、表中所涉及的签名都必须用蓝、黑色墨水笔，亲笔签名或签字章，不可以打印代替。

四、同行专家业内评价意见书编号由工程师学院填写，编号规则为：年份4位+申报工程师职称专业类别(领域)4位+流水号3位，共11位。

一、个人申报

（一）基本情况【围绕《浙江工程师学院（浙江大学工程师学院）工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】

1. 对本专业基础理论知识和专业技术知识掌握情况(不少于200字)

本人在项目实践中掌握了图像处理、通信技术以及人工智能相关的基础理论和应用方法。在参与数字孪生系统（智能算法虚实评测环境）建设过程中，逐步掌握了虚拟建模、多源数据融合、通信同步机制等关键技术，能够较熟练地参与虚实系统的设计与实现。在图像生成方面，近年来重点学习并研究了扩散模型在拟真图像合成中的应用，熟悉其核心原理，如扩散过程建模、神经网络结构设计以及生成质量优化等，能够根据具体需求对算法进行改进与工程落地。日常工作中注重理论联系实际，持续通过项目实践提升自身能力，同时也积极关注行业前沿技术，努力拓展知识广度和深度。

2. 工程实践的经历(不少于200字)

1.
在工程师学院的专业实践中搭建了智能算法虚实评测环境，这是一个类似数字孪生系统的大型平台，包括：a. 将实体车辆接入虚拟系统，对应实的部分；b. 改造开发开源自动驾驶平台，从中采集得到虚拟传感器数据，对应虚的部分；c. 开发虚拟图像数据集的拟真算法，得到大规模拟真图像数据集；d. 实现了多服务器接入的通信架构，支持200个以上节点的实时接入；

2.
在华为公司实习的过程中，设计了LLM多头投机推理方法，主要包括：研究LLM多头投机模型，实现LLM推理加速。提出一个低参数量同时具有高准确性和生成性能的能预测多token投机模型。提出一个两阶段的训练方法：a. 不修改大模型本身参数，仅训练投机模型得到一个具有较好性能的一阶段模型；b. 通过qlora端到端整体训练，得到一个优于Medusa2的二阶段模型

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例（不少于1000字）

一、构建智能算法虚实评测平台：数字孪生系统的工程实现

在工程师学院专业实践阶段，我参与并主导搭建了一个智能算法虚实评测环境。该系统本质上是一个面向智能驾驶算法测试的大型数字孪生平台，涵盖了物理系统接入、虚拟系统建模、图像生成与通信架构构建等多个技术模块。

首先，在“实”的部分，我负责将实体车辆接入平台。通过改造车载传感与通信模块，使其能实时上传车辆状态、传感器数据与控制信号，打通物理系统与数字平台的数据通路。过程中，我深入阅读了CARLA底层代码、车辆内部的CAN协议和相关的通信技术，保证了数据在上传过程中既稳定又具备低延迟特性。

其次，在“虚”的部分，我基于开源自动驾驶仿真平台（CARLA）进行深度改造，使其能生成与实体车辆同步的虚拟场景和传感器数据。我设计了可配置的虚拟摄像头、激光雷达等传感器模块，并建立了道路场景建模机制。在这一过程中，我运用了计算机视觉、图像处理、三维建模等知识，使虚拟数据具有高度还原性，为后续算法验证提供可靠环境。

同时，我为CARLA系统设计了一套适合大规模并发接入的通信系统，这套系统能使得不同服务器能够同步世界中的车辆、人员、天气、地图等，满足了项目对于接入的人员和车辆不少于200个的需求。除此以外，为了适应项目的需要，我还设计了场景中不同主体的攻击和血量等系统，以及攻击的同步方式，将传统的自动驾驶模拟器CARLA改造为了一个适应军事场景的模拟系统。值得一提的是，我为每个实体实现了外部接口，使得里面的每个人可以受到

外部的控制，包括键盘鼠标的控制和大模型的整体调动，实现军事战略动作。

通过上述工作，我在项目中综合运用了通信原理、嵌入式系统、图像处理、系统架构与算法优化等多方面知识，实现了从“虚实融合”建模到“系统部署”落地的完整闭环，提升了我对复杂系统工程的理解与实践能力。

二、在华为实习期间参与多头投机推理模型设计

在华为公司实习期间，我有幸参与大语言模型（LLM）推理加速项目，主要负责多头投机推理机制的研究与实现工作。在当前AI推理成本居高不下的背景下，投机性推理成为提升大模型生成效率的重要技术路径，我所参与的工作正是对该方向的深入探索。

我首先深入研究了现有的多头投机推理架构（如Speculative

Decoding、Medusa等），理解其在token生成预测与模型校验之间的配合机制。在此基础上，我提出了一种轻量化的多头投机推理模型架构，该模型在不显著增加计算资源开销的前提下，实现了对多个token的并行预测，并具有较高的准确性和生成效率。

我采用两阶段训练策略进行模型构建：第一阶段在保持原始大模型参数冻结的前提下，仅训练一个辅助投机预测模型，快速收敛至具备良好初步性能的投机模块；第二阶段通过QLoRA微调手段，在保留主模型稳定性的前提下，实现了端到端的训练与精度提升。最终模型在准确率与吞吐量方面均优于当前业内常用的Medusa方案。

在该项目中，我不仅应用了自然语言处理、深度学习、分布式推理架构等理论知识，还与团队成员协同完成了系统调试、推理优化与性能评估，进一步提升了工程实践能力与团队协作能力。该项目成果已在公司内部模型加速平台中部署试运行，为后续大模型应用提供了有力支持。

(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩： 89 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1.2 年（要求1年及以上） 考核成绩： 83 分
本人承诺	
个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！	
申报人签名：曹世聪	

22260057

二、日常表现考核评价及申报材料审核公示结果

日常表现 考核评价	非定向生由德育导师考核评价、定向生由所在工作单位考核评价 ☒ 优秀 ☐ 良好 ☐ 合格 ☐ 不合格 德育导师/定向生所在工作单位分管领导签字（公章）  年 月 日
申报材料 审核公示	根据评审条件，工程师学院已对申报人员进行材料审核（学位课程成绩、专业实践训练时间及考核、学位论文、代表作等情况），并将符合要求的申报材料在学院网站公示不少于5个工作日，具体公示结果如下： ☐ 通过 ☐ 不通过（具体原因： 工程师学院教学管理办公室审核签字（公章）： 年 月 日

浙江大学研究生院
攻读硕士学位研究生成绩表

学号：22260052		姓名：曹世聪		性别：男		学院：工程师学院			专业：电子信息			学制：2.5年			
毕业时最低应获：26.0学分				已获得：29.0学分				入学年月：2022-09			毕业年月：				
学位证书号：						毕业证书号：						授予学位：			
学习时间		课程名称		备注	学分	成绩	课程性质	学习时间		课程名称		备注	学分	成绩	课程性质
2022-2023学年秋季学期		工程技术创新前沿			1.5	87	专业学位课	2022-2023学年冬季学期		产业技术发展前沿			1.5	90	专业学位课
2022-2023学年秋季学期		运动素质课			1.0	通过	公共学位课	2022-2023学年春季学期		自然辩证法概论			1.0	84	公共学位课
2022-2023学年秋季学期		数值计算方法			2.0	96	专业选修课	2022-2023学年春季学期		电子与通信工程领域前沿讲座			2.0	94	专业选修课
2022-2023学年冬季学期		车辆控制理论与技术			3.0	91	专业学位课	2022-2023学年春夏学期		工程伦理			2.0	95	公共学位课
2022-2023学年秋冬学期		研究生论文写作指导			1.0	84	专业学位课	2022-2023学年春夏学期		高阶工程认知实践			3.0	89	专业学位课
2022-2023学年秋冬学期		科技创新案例探讨与实战			2.0	88	专业选修课	2023-2024学年夏季学期		研究生英语应用能力提升			2.0	免修	公共学位课
2022-2023学年冬季学期		新时代中国特色社会主义思想理论与实践			2.0	94	公共学位课			硕士生读书报告			2.0	通过	
2022-2023学年冬季学期		车辆信息传感与通信技术			3.0	94	专业学位课								

说明：1. 研究生课程按三种方法计分：百分制，两级制（通过、不通过），五级制（优、良、中、及格、不及格）。

2. 备注中“*”表示重修课程。

学院成绩校核章：
成绩校核人：张梦依
打印日期：2025-06-03



国家知识产权局

310013

浙江省杭州市西湖区古墩路 701 号紫金广场 B 座 1103 室 杭州求是
专利事务所有限公司
邱启旺(0571-87911726-808)

发文日:

2024 年 09 月 29 日



申请号: 202411370318.9

发文序号: 2024092901097860

专利申请受理通知书

根据专利法第 28 条及其实施细则第 43 条、第 44 条的规定, 申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下:

申请号: 2024113703189

申请日: 2024 年 09 月 29 日

申请人: 浙江大学

发明人: 曹世聪, 杜歆

发明创造名称: 一种用于自动驾驶场景的拟真图像生成方法

经核实, 国家知识产权局确认收到文件如下:

权利要求书 1 份 2 页, 权利要求项数: 6 项

说明书 1 份 7 页

说明书附图 1 份 2 页

说明书摘要 1 份 1 页

专利代理委托书 1 份 2 页

发明专利请求书 1 份 4 页

实质审查请求书 文件份数: 1 份

申请方案卷号: 邱-241-303-静

提示:

1. 申请人收到专利申请受理通知书之后, 认为其记载的内容与申请人所提交的相应内容不一致时, 可以向国家知识产权局请求更正。

2. 申请人收到专利申请受理通知书之后, 再向国家知识产权局办理各种手续时, 均应当准确、清晰地写明申请号。

审查员: 自动受理

联系电话: 010-62356655

审查部门: 初审及流程管理部



200101
2023.03

纸件申请, 回函请寄: 100088 北京市海淀区蓟门桥西土城路 6 号 国家知识产权局专利局受理处收
电子申请, 应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外, 以纸件等其他形式提交的文件视为未提交。