

校企项目个人报告

2025 年 5 月 11 日

姓 名：HOOJUN KIM 专 业：信息工程

导 师：梁晓妮 学 院：学生创新中心

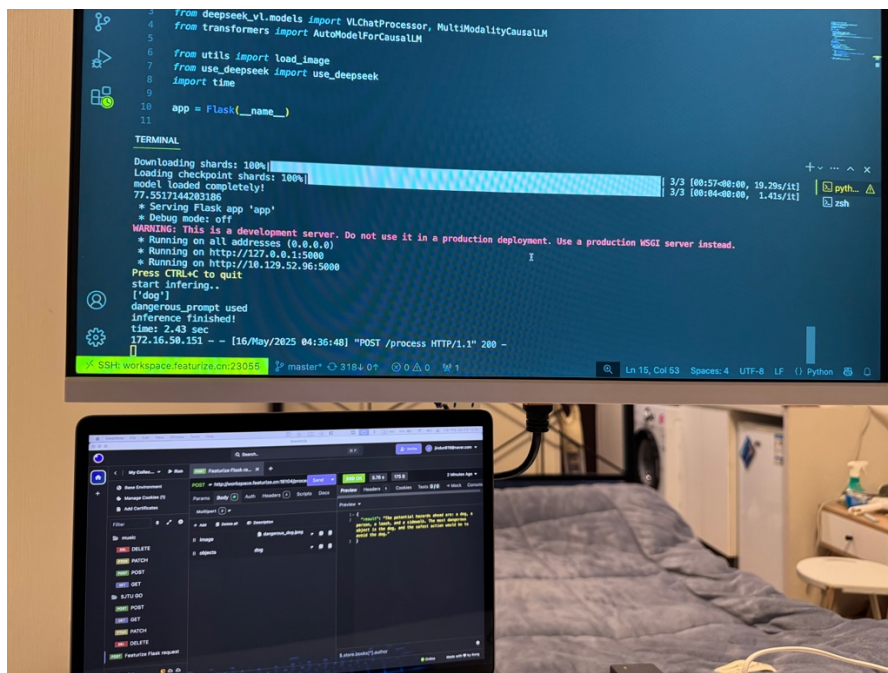
学 号：520030990027

本次项目名为“微型物体识别相机”，旨在为视障人士提供一个便携式、智能化的环境感知工具。通过图像识别、语音反馈等方式，我们希望帮助视障用户准确地理解自己所处的场景与周围的物体，从而提升他们在日常生活中的安全性与独立性。本项目融合了最新的人工智能模型、边缘计算优化技术以及多模态感知系统，是一次理论与实践高度结合的有益尝试，也让我在多个方面获得了显著成长。

一、个人职责与项目贡献

我们团队由三名成员组成，每个人都在不同领域发挥了关键作用。我在团队中主要负责服务器的搭建与后端系统的整体管理。此外，我还积极参与了 AI 模型的部署与优化流程，协助完成了项目结构的初步设计以及后续的技术选型与集成任务。

在服务器建设方面，我经过详细调研与比较，结合我们的项目需求（如推理效率、成本、算力稳定性等），最终选择了 Featurize 平台进行部署。该平台为我们的模型运行提供了可靠的 GPU 支持，并允许我们灵活配置资源。在此基础上，我使用 Flask 框架搭建了后端服务，支持接收来自终端设备的 HTTP 请求，处理图像和文本数据，并调用 DeepSeek-VL 多模态模型进行推理，将结果实时反馈给用户。



【图 1】 服务器响应速度的精确测试

为了提升用户体验，我引入了模型缓存加载机制，通过避免每次请求时重新初始化模型，大幅缩短了响应延迟，提升了整体流畅度。在模型选择方面，我还测试和评估了多个主流视觉-语言模型，如 Qwen2.5-VL-7B-Instruct、LLaVA-3.2-11B-Vision-Instruct 等，从 GPU 占用率、响应时间、准确率等维度进行了详尽的 benchmark，最终协助团队确定了最适合本项目场景的模型组合。

值得一提的是，我在项目初期提出了“模型部署在服务器端、终端设备轻量化”的架构建议，这一构想后来成为了我们系统设计的基础。通过将计算密集型任务集中在云端，我们不仅减轻了设备端的计算压力，还在低功耗的条件下确保了系统的实时性和稳定性。

我还协助进行 YOLO 模型的训练与微调工作，为识别“楼梯”等特殊障碍物训练了特定数据集，提升了系统在实际场景中的适应能力。同时，我也参与了部分 prompt 设计工作，优化了多模态模型对复杂场景的理解与输出能力。

在实际测试阶段，我们虽然还未能将系统最终部署到真正的边缘设备上，但成功地在本地计算机环境下完成了完整的功能验证。我们将 RealSense 深度相机连接至本地计算机，成功识别环境中的物体，并将数据发送至我们搭建的服务器，再通过语音合成（TTS）模块将信息以语音形式播报出来。这一模拟测试不仅验证了系统端到端的功能实现，也在项目汇报时得到了展示，获得了老师们的认可。

**【图 2】 首次完整链路测试成功的瞬间**

二、技术能力的提升与系统性学习

本项目让我深入理解了 AI 视觉系统的构建流程，尤其是涉及视觉语言模型（Vision-Language Models）与目标检测模型（如 YOLOv5）的协同工作机制。我不仅系统性地

学习了模型架构、训练流程及各类参数的作用，还掌握了如何根据实际需求调节这些参数，以实现性能最优化。

在模型优化方面，我学会了如何使用 mAP50、Recall、Precision 等指标评估模型输出质量，并通过实验不断调整超参数，实现精度和效率之间的平衡。尤其在 YOLO 模型的迁移学习过程中，我深刻体会到数据集选择、预处理、增强策略对模型最终表现的影响。



【图 3】 训练专属识别模型的背后

此外，我掌握了在 Linux 服务器环境下，部署 AI 模型并与终端设备进行 HTTP 通信的完整流程。通过实际搭建与调试，我提升了自己的后端开发能力，也学会了如何使用如 Insomnia 这类工具进行接口测试和问题排查，保证系统的可靠性与稳定性。

同时，我也掌握了如何处理高并发请求和大规模模型推理任务，例如在服务器压力增大时如何优化资源调度，保障服务不中断，为将来从事 AI 工程类工作打下了坚实基础。

通过这一系列的技术积累，我不仅了解了前沿的视觉-语言模型技术，还锻炼了如何在复杂系统中整合多种 AI 组件，真正做到“理论联系实际”，提升了我的系统工程能力。

三、团队合作与项目管理实践

我们在团队协作方面做了充分的准备与规划。在项目初期，我们通过多轮讨论明确了目标与分工，并制定了详细的时间线。我们每周定期召开线上会议，及时同步各自的

进度和遇到的问题；每位成员也会将工作成果记录在文档中，以便其他成员查阅与反馈。我们还使用 GitHub 进行代码管理，确保版本控制规范、更新透明，提升了整体协作效率。

我认为这次项目在团队协作上最大的优势在于信息透明与交流顺畅，我们在项目进行过程中不断调整分工、互相支持。在任务执行过程中，我也时常帮助其他成员解决他们遇到的技术问题，共同推动项目进展。

当然，也存在一些可以改进的地方。例如在项目初期，由于任务分配标准不够明确，出现了某些工作重复、资源浪费的情况。这提醒我们，项目启动阶段应花更多时间明确角色与边界，确保每个成员专注于最适合的工作方向。

四、创新实践与问题解决能力的提升

本项目的核心创意与技术架构均起源于我个人在前期的构思与提案中。我认为将多模态模型部署在云端而不是终端，是当前技术条件下更合理、更高效的方式。因此，我主动进行技术可行性分析，并通过实际部署验证这一方案的优势，最终得到了团队的支持。

在整个开发过程中，我始终保持“主动探索、不断验证”的态度，对技术细节保持好奇心和思辨能力。从模型对比、性能测试、系统集成，到部署调优，我几乎每一个环节都亲身参与，收获了宝贵的项目经验。

同时，这也是一次挑战传统辅助技术边界的创新实践。我们希望通过将最先进的 AI 模型应用于现实生活中真正有需求的人群，推动人工智能的“落地”，真正服务于社会。

五、未来展望与个人感悟

虽然目前我们尚未完成将整套系统嵌入实际的微型边缘设备，但通过本地环境的模拟测试，我们已成功验证了整体方案的可行性。接下来，我计划继续推进项目，将 AI 推理能力整合进轻量化的便携设备中，打造真正可用的“微型物体识别相机”。

未来的功能规划包括：开发多种场景下的安全预警系统，如识别交通信号灯、即将驶来的汽车或斑马线，帮助用户安全过马路；设计智能导航功能，协助视障人士在复杂环境中自主移动；同时，还希望开发“寻找物品”功能，帮助用户快速定位日常生活中的小物件。

我认为，在当前 AI 高速发展的时代，亲自参与并体验基于人工智能的产品研发，无论未来是否从事相关职业，都是极其宝贵的经历。这不仅增强了我的技术实力，更拓宽了我的视野，为未来职业发展积累了重要的资本。

这次项目让我体会到，技术创新不仅仅是追求理论上的突破，更重要的是结合实际需求，改善人们的生活质量。作为开发者，我们肩负着用技术赋能弱势群体的责任和使命，这也激励我在未来继续投身于有社会意义的科技工作中。



【图4】 从用户视角验证系统可行性