**软件学院创新项目实训**

**任务书**

**（2024版）**

项目编号 （申请人不必填写）

项目名称 Sapphire：基于SAM的智能辅助数据标注平台

团队负责人 侯玉祺

手机号码 15889766016

山东大学软件学院

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **项目名称** | Sapphire：基于SAM的智能辅助数据标注平台 | | | | | |
| **项目目标** | 基于Segment Anything Model(SAM)的先进图像分割技术，结合众包任务分发概念，搭建AI辅助数据标注平台，填补细分领域的市场空白。 | | | | | |
| **指导教师** | 武蕾 手机号码：15662664499 | | | | | |
| **场地工位** |  | | | | | |
| **团队名称** | 重生之我在山大标数据 | | | | | |
| **参加人员** | **姓名** | **学号** | **专业** | **手机号码** | **电子邮箱** | **项目分工** |
| 侯玉祺 | 202100300341 | 软件工程 | 15889766016 | houyuqi2003@foxmail.com | 组长，负责人 |
| 吴羲勇 | 202100300261 | 软件工程 | 19963645235 | wxyww2049@gmail.com | 代码、文档编写 |
| 朱哲甬 | 202100300082 | 数字媒体技术 | 18967892613 | zhuzy@mail.sdu.edu.cn | 代码、文档编写 |
| 郭苏睿 | 202100300224 | 软件工程 | 13012615613 | 13012615613@163.com | 代码、文档编写 |
| 伍钰棋 | 202102300398 | 数字媒体技术 | 19168852517 | irra.c923@gmail.com | 代码、文档编写 |
| 其他人员 | | | | | |
| **资源地址** | | | | | | |
| 魔搭社区用户手机号码和邮箱：指导教师提供算力资源。  项目CSDN博客地址：https://blog.csdn.net/m0\_59025440/category\_12621020.html  个人CSDN博士地址：  （1）侯玉祺：https://blog.csdn.net/ellyi/category\_12620990.html  （2）郭苏睿：https://blog.csdn.net/qq\_40068267/category\_12620988.html  （3）吴羲勇：https://blog.csdn.net/wxyww2049/category\_12621009.html  （4）伍钰棋：https://blog.csdn.net/xi\_irra/category\_12621018.html  （5）朱哲甬：https://blog.csdn.net/u012185664/category\_12621022.html | | | | | | |
| **项目任务要点（请简明扼要列出关键点，具有可考核、可量化指标）** | | | | | | |
| 1 系统设计 **Sapphire：基于SAM的智能辅助数据标注平台**，基于Segment Anything Model(SAM)的先进图像分割技术，结合众包概念，用冗余任务分发保证标注质量。本平台由标注员提供点或框作为prompt，AI模型基于此自动分割对象，提高了标注效率，填补了该细分领域的市场空白。  主要包括以下六大模块：用户模块、数据集模块、分发模块、AI辅助标注模块、分析统计模块及激励模块。 2 系统功能性需求2.1 角色定义 主要分为两类：   * 管理员：拥有最高权限。 * 普通用户：既可作为众包发布者也可作为标注员。  2.2 基本功能 **管理员端：**   1. 用户管理：权限分配、重置密码、用户数据统计等。 2. 任务管理：管理任务的生命周期，包括任务创建、分发、进度监控等。 3. 数据集管理：创建、维护数据集等。 4. 分析统计：生成各类分析统计报告。 5. 激励机制管理：设置和调整激励机制，如积分规则和排名奖励。   **普通用户端：**   1. 注册与登录 2. 个人资料管理：支持编辑个人信息，如联系方式、密码等。 3. 标注任务管理：领取、标注和提交标注任务。 4. 数据集浏览与下载：用户可以花费积分等，获取平台提供的数据集。 5. 标注工具与界面：使用平台提供的多种标注工具和友好的界面进行数据标注。 6. 激励机制：积分与排名    1. 完成标注任务可获得积分，发布任务需消耗积分    2. 用户可以查看自己的积分，参与平台排名竞争。    3. 标注质量保证：用户需要按照规定的标注准则进行标注，并确保标注结果的准确性和一致性。标注质量呈现在积分中。  3 系统非功能性需求3.1 性能需求 **前端响应时间**：前端界面在与用户的交互中，应保持快速响应，以确保用户操作的流畅性和高效性。前端页面推理蒙版加载时间应控制在**2秒以内**。  **后台任务处理时间**：后台任务自动分发和标注结果比对的处理时间应控制在合理范围内，以确保任务能够及时分发和结果能够快速返回给用户。从用户开始获取任务到得到分发的任务，等待时间不能超过**3s**。在用户标注期间，同时像前端缓冲区预发送任务。从得到用户标注的结果，到完成比对写入数据库，不超过**1s**。 3.2可伸缩性 **并发用户支持**：系统应能够同时支持多个用户进行标注任务，且在高负载情况下能够保持稳定性和性能表现。系统应能够同时支持至少**100**个并发用户进行标注。  **数据集规模支持**：系统应能够处理大规模的数据集，包括大量的标注任务和数据集的存储、检索和管理。系统应能够处理包含至少**10000**个数据样本的数据集，并能完成及时的分发与加工。 3.3 安全性 **数据隐私保护**：系统应采取适当的安全措施来保护用户的数据隐私，防止未经授权的访问、泄露或篡改。用户数据应进行加密存储，并采用访问控制措施限制对敏感数据的访问。  **身份验证和访问控制**：系统应支持身份验证和授权机制，确保只有经过授权的用户才能访问敏感操作和数据。使用用户名和密码进行身份验证，并为不同用户角色分配相应的权限。 3.4 易用性 **用户界面友好性**：前端界面应具备良好的用户体验和易用性，以便用户能够轻松理解和操作系统的功能。使用直观的界面设计、清晰的标注指导和友好的错误提示。  **文档和培训资料**：系统应提供清晰、详尽的用户文档和培训资料，以帮助用户快速上手和充分利用系统功能。 4 实施计划 **Sapphire：基于SAM的智能辅助数据标注平台**项目预计于本学期第五周开始，第十七周结束，共计十二周。  具体的实施计划如下：  第五周（3.25~3.31）：需求分析与产品设计。  第六、七周（4.1~4.14）：核心技术验证。  第八～第十三周（4.15~5.26）：网页端与服务端开发。  第十四～第十五周（5.27~6.9）：测试与部署。  第十六～第十七周（6.10~6.23）：文档编写和项目总结。 | | | | | | |
| 指导教师  （可不填） | （签字） 2024年4月8日 | | | | | |
| 学院负责教师签字 | （签字） 年 月 日 | | | | | |
| 本团队郑重承诺，此申请书内容真实有效。项目成果引用的模型、数据集、论文等成果全部标注引用和来源。    （所有团队成员签字）  2024 年4月 8日 | | | | | | |

注：本文件一式两份，作为项目验收时的完成度证明，请谨慎填写。

1. 如果没有魔搭社区用户，请注明算力资源来源。
2. **必须有项目CSDN博客地址，用于记录项目工作日志；必须有个人CSDN博客地址，记录个人周报。以上内容请可以公开访问，爬虫抓取。**
3. 项目任务要点主要突出可量化，比如“选取了XXX张图片、XXXGB的XX数据，具有XX%的准确率”等描述。
4. 如果有指导教师，必须有指导教师签字。