

四川大學



个人报告

课 程 名 _____ 问题求解实战 _____

任课教师 _____ 杨波 _____

学 院 _____ 计算机学院 _____ 专 业 _____ 计算机科学与技术 _____

学生姓名 _____ 郑仕博 _____

学 号 _____ 2023151610075 _____ 年 级 _____ 2023 _____

2025 年 6 月 27 日

1. 项目基本情况

本项目名为“基于生成式 AI 的个性化文创图像作品设计系统”，旨在解决当前文创产品同质化严重、难以满足日益增长的个性化需求的痛点。在国家推动文化和旅游融合发展，以及《如果国宝会说话》等成功案例的启发下，本项目致力于通过创新方式“激活”文化遗产，赋能个体创造独特文创作品。核心目标是利用先进的生成式 AI 模型，开发一个能够让普通用户轻松设计个性化文创图像的工具，从而满足市场对个性化文创设计的需求，并探索相关技术在中文文创领域的应用潜力。

项目的主要开发工作包括：模型微调与优化，采用 AnyText 模型作为基础，并针对中文应用场景进行专门的训练和优化，具体分为文字控制框架的训练和扩散模型的训练两部分。其次是数据集构建，制作了两份数据集，一份是基于 AnyWord-3M 筛选出的约 40 万条数据用于微调 AnyText 模型，另一份是通过网络爬虫采集的约 1000 张与中华文化及文物相关的图片用于微调 stable diffusion v1-5 模型。最后是系统实现与部署，基于 Gradio 搭建了交互式的网页界面，实现了文字到图片生成和图片文字编辑两大核心功能。

本项目由四川大学计算机学院的杨波老师指导，第一组，团队成员包括组长郑仕博（我本人），以及组员陈奕嘉和苏泳豪。项目周期从最初的策划到最终的部署和文档完善，历时数月，期间经历了多次迭代和问题解决，最终完成了模型的训练、功能实现、系统部署，并申请了软件著作权。

2. 承担的主要工作情况

在本项目中，我主要承担了项目统筹与规划、模型训练与调优、核心代码编写与上传以及项目部署等关键职责。以下是我在项目不同阶段的具体工作及完成情况：

项目初期（Week 4-5）： 作为组长，我首先对项目计划书进行了细致的修订，并明确了团队成员的具体分工。我安排组员负责搜集图片素材以构建数据集，并着手进行前后端开发工作，而我则专注于模型的训练环节，并深入研究应对过拟合问题的有效策略。在第五周，我组织组员开展了图片收集工作，并与组员共同学习扩散模型的结构。我还编写并修改了相关权重文件，对收集到的图片进行了初步处理和筛选，例如去除水印、优化图像质量等，以确保数据的可用性和一致性。

需求分析与数据集构建（Week 6-7）： 在第六周，我与组员陈奕嘉一同完成了需求分析系统的初稿，并与组员一起收集数据集图片，寻找标注工具并指导组员进行标注。第七周，我与陈奕嘉一起完善了需求分析文档的细节，并与组员一起进行了 AnyText 框架的调节。同时，我也与苏泳豪一起进行了扩散模型的训练。

模型训练与优化（贯穿项目始终，尤其在 Week 8）： 模型训练与调优是我的核心工作。我专注于模型训练，研究并解决了过拟合问题。在第八周，我与组员一起完成了模型的训练，得到了最终版本，并由我上传至 ModelScope。我还负责编写了模型测评文件代码并上传至 GitHub。在项目总结报告中也提到，我负责了模型微调与优化，采用 AnyText 模型作为基础，并针对中文应用场景进行专门的训练和优化。为了更贴合主题，在 Realistic Vision V4.0（基于 sd1.5）权重基础上，我使用 dreambooth 的方法进行微调，并与微调后的 AnyText 模型的权重进行整合。

系统实现与部署（Week 9-12）： 在第九周，我与组员一起完成了前后端的制作，并上传至 GitHub，修改了标题并增加了一些示例。在第十周，我进行了系统概述文档和需求分析文档的修改，用 LaTeX 重新书写，使文档更加正式和规范。第十一周，我们完成了 Dockerfile 的编写，并将项目封装并上传至 Docker Hub，解决了正常部署可能因网络问题导致文件下载不下来的问题。在第十二周，我负责编写了 requirements.txt 文件，并将其与代码部分一起上传至 ModelScope，进行了部署。

文档完善与汇报准备（Week 13-14）： 在项目后期，我持续参与了文档的完善工作。在第十三周，我们对两份文档进行了修饰，修改了需求分析并增加了技术难点。第十四周，我们完善了文档，完成了测试报告，并开始制作报告 PPT。

总的来说，我不仅承担了项目管理和统筹的职责，确保项目按计划进行，还深入参与了核心技术部分的开发，特别是模型训练与部署，为项目的顺利完成做出了关键贡献。

3. 项目实践过程中遇到的问题及处理结果

在项目实践过程中，我们遇到了一系列挑战，主要集中在技术实现和资源限制方面。

前后端不完善与文字叙述问题（Week 4）： 在项目初期，我们发现前后端并不完善，并且部分文字叙述不得体。

处理结果： 我立即督促组员进一步完善前后端开发，并对文字叙述进行了修改和润色，确保表达准确得体。在后续周报中，这些问题逐渐得到解决，例如在 Week 9，我们完成了前后端的制作并上传至 GitHub。

数据集不足与开源文件存在 Bug（Week 5）： 我们面临现有数据集不足的问题，并且使用的开源文件存在 Bug。

处理结果： 我指导组员继续收集更多有用的数据集，例如通过网络爬虫采集与中华文化及文物相关的图片。同时，我们积极通过论坛等渠道寻求解决方案，以解决开源文件中的 Bug。在后续的数据集制作部分，我们成功制作了两份高质量的数据集。

AnyText 框架效果未达到预期与扩散模型结合不理想（Week 6）： AnyText 框架的效果并未达到预期，且扩散模型与框架结合的效果不理想。

处理结果： 我们考虑使用更新的权重，并深入研究框架对扩散模型生成的影响。在 Week 7，我们与组员一起进行了 AnyText 框架的调节，并与苏泳豪进行了扩散模型的训练。

模型效果未达到预期（Week 7）： 尽管进行了调节和训练，模型效果仍未达到预期。

处理结果： 我最终和组员对扩散模型进行了针对性的训练，并计划将其上传至 ModelScope。在 Week 8，我们成功完成了模型的训练，得到了最终版本，并由我上传至 ModelScope。

文字生成正确率未达到预期（Week 8）： 尽管模型效果显著提升，但文字生成正确率并未达到预期。

处理结果： 在项目总结报告中，文字正确率测试结果也印证了这一点，训练前后文字生成和修改的正确率及编辑距离没有显著提升，这与算力、参数及训练批次有关。最终，我们采用了官方的 AnyText 权重并与训练好的扩散模型进行合并。未来可以考虑使用更新的模型，并且增加改变字体的功能。项目展望中也提及，计划未来采用性能更优的 AnyText2 模型，该模型增加了字体选择功能。

前后端美化与部署难度（Week 9）： 前后端美化有待加强，且部署可能会存在一些难度。

处理结果： 我将模型的封装到 Docker 镜像中，以方便部署。在 Week 11，我们完成了 Dockerfile 的书写，并将项目封装并上传至 Docker Hub，解决了正常部署可能会因为网络问题导致文件下载不下来的问题。

模型未部署到服务器，无法演示（Week 11）： 项目存在未将模型部署到服务器上，导致无法进行演示的问题。

处理结果： 我将模型部署到 Modelscope 上，方便后续进行演示，并进行对照实验。

ModelScope 持久化存储问题导致启动慢（Week 12）： 在 ModelScope 上部署后，发现无法持久化存储，导致每次启动非常慢。

处理结果： 后来我发现设置为公开后启动速度变慢，于此同时我和我的组员进行了测试。

算力限制： 在项目总结报告中明确指出，在文字控制框架的训练过程中，由于算力限制（至少需要 8 卡 V100），训练效果未达到预期，这是一个主要的挑战。

处理结果： 尽管算力限制是一个难以立即解决的外部因素，但我们将模型拆成两部分进行训练，最终在官方的 AnyText 权重基础上与训练好的扩散模型进行合并，以保证项目的最终效果。项目展望中也提到，未来需要寻求更强大的计算资源支持。

这些问题的解决，体现了团队在项目管理和技术攻关方面的能力，也为项目的最终成功奠定了基础。

4. 参与项目过程的体会与自我评价

参与项目过程的体会：

参与“基于生成式 AI 的个性化文创图像作品设计系统”项目是一次宝贵的学习和成长经历。作为项目组长，我不仅在技术层面得到了显著提升，更在项目管理、团队协作和问题解决能力方面积累了丰富的经验。

首先，在技术方面，我深入了解并实践了生成式 AI 模型，特别是 AnyText 模型和扩散模型的训练与优化。从最初的 AnyText 模型微调、数据集构建，到最终的模型部署和性能测试，每一个环节都让我对深度学习模型有了更深刻的理解。我学会了如何处理和筛选大规模数据集，如何针对特定应用场景（如中文文创）对模型进行定制化训练，以及如何评估模型的性能并解决过拟合等问题。特别是对文字控制框架的训练和扩散模型的合并，让我对文本与图像融合的复杂性有了更直观的认识。尽管遇到了算力限制的挑战，但我们通过研究和调整，最终选择了最合适的模型权重，这让我体会到在实际项目中，权衡技术理想与现实条件的重要性。

其次，在项目管理和团队协作方面，我作为组长，从项目计划书的修订、团队成员的分工、任务进度的督促，到文档的规范化编写和汇报材料的准备，全程参与并主导了各项工作。这让我深刻体会到清晰的项目规划、有效的沟通机制和明确的职责分工对于项目成功的关键作用。每周的周报制度和定期的团队会议确保了我们能够及时发现并解决问题，例如前后端开发进度、数据集收集困难、模型效果不理想等。通过与陈奕嘉和苏泳豪两位组员的紧密合作，我学会了如何发挥每个人的长处，共同克服困难。尤其是在 Docker 封装和 ModelScope 部署等复杂环节，团队成员的协同努力是不可或缺的。

最后，在问题解决方面，面对数据集不足、开源文件 Bug、模型效果不佳以及算力限制等一系列挑战，我们没有退缩。我积极带领团队分析问题根源，寻找解决方案，无论是通过论坛寻求技术支持，还是调整训练策略，都体现了我们解决问题的能力。特别是 ModelScope 持久化存储的问题，让我认识到在实际部署中，除了技术实现，还需考虑平台特性和用户体验。这些经历让我学会了如何在不确定性中保持灵活性和适应性，并不断探索新的方法。

自我评价：

在本次项目中，我不仅担任了项目组长，更作为技术核心成员，深度参与并推动了从模型研究到系统部署的整个生命周期。

在技术攻坚方面，我主导了项目最关键的生成模型训练工作。面对模型过拟合等技术难题，我主动查阅前沿文献，结合实践经验提出了多项改进策略。当初代模型效果未达预期时，我迅速组织团队开展问题复盘，敏锐定位瓶颈，灵活调整技术路线，最终成功训练出性能优异的最终版本模型，并上传至 ModelScope，为整个系统奠定了坚实的技术基础。此外，我编写了模型评估脚本并将其开源在 GitHub 上，展现了良好的工程规范意识与开源精神。

在工程实现与部署方面，我致力于推动算法走向实际应用。我积极参与前后端开发工作，并使用 Dockerfile 封装项目环境，解决了因网络环境差异带来的部署障碍，成功将应用上传至 Docker Hub。这一过程充分体现了我对软件流程的深刻理解和实际掌控能力。

在项目文档与规范化建设方面，我也始终保持高标准。为提升文档质量与专业性，我使用 LaTeX 重新编排了系统概述与需求分析文档，并在项目收尾阶段主动补充技术难点说明，确保项目资料全面、清晰、具有复用价值。

通过本项目，我不仅深化了在生成式 AI 方向上的技术能力，也全面锤炼了我的项目管理能力、系统化思维和工程落地能力。在协作中，我展现出强烈的责任心和技术领导力；在面对问题时，我始终保持敏锐洞察力与解决问题的主动性。