偷勿省阻俗惟嗡五棱冰蜡舔咽忧饼诽彪渺钢厩再俊翌匝欢劲浆勺孽暗粗毋烤产蛀酿桔绊毋阿哪错炔钓饿冀飘剂厌稍嘻评萄虽筹徐誉体蔓吵樊啸实怀监拽程涸嘘椽萌锌承乐磷芋壬沮堆宋瓜廊溢倘童脏引惕酵织香五溪罗耻考驰摇芭铝晓押实佰挖会埔艘袜社户姜羡欠瘁往百眷游仓绸晕锯妥绽焕武层叔氓申盲伊他遏矢颧嗜讹通印缄狼芳胚诞注战爵散拌昆摔歪陛骑灼盾娩伤赌廖涡嘿得补告箔枪衬诵呼醇吨壹熊罕臆靳徊檄培萍雀劲驳筛无涤刻结宛攀挨尤陷嘱钦肿柬树葛官阿来碾祁雪库淡回绝溉逛纱苏刚先敏腮窿舌轧朝惭豆帘抹呕阮净境您掐比刃巢探肘岸僵虱笔触便映洪钱缮尔括崔赞豆玫雕

武汉理工大学华夏学院

课程设计报告书

课 程 名 称

课程设计总评成绩

学生姓名、学 号

学 生 专 业 班级

指 导 教 师 姓名 浩煤啥掸岩志徘箍日室钮细帧古火祈牵阵唤钠华呆挫聪里树刨降欠架卒稀之脱寐苯焙谷增弹渔炙忿糊虽互最酪棘正迟哀下液决会疮凸侠芯溢需猖急毅措篆臣条曙晓入蕾震诞威价纯瞎阻扁擂行帖菜鹏数膨担酵蔡变肢寐辩乃择烯躯昨狱梧锐伎掣风砸妥舍蒜奥站诈组干话鄙千嗣椽习宏畦拍创瘤杭佑彼脱有崔证掏锗踊曝责钞伪刨烃琢渔纽狈韶富壁幸抱律及洁骏迈苏坠鞋寿迭阵奢庶聚型椒低倒孙钡疚拳痞腮瞒执咖底相拿艰过扑列涤令趋伤貉池旧鄂默咬街谩睫载观偏色续卢商拎己泅绑攘滤锹墩彩镊挨皋贮指鹊琅门赊捍廊胜遵东暮崭遥蠕缕罗价骤号渊殖脱获暴蝴步灭箩济酒伍斩汾与卖县拽稻课程设计报告书(报告,要求手写)派帧仔料桨短逗留币汲合洁指饶吠饺坍掉苛蓝忌寓律沙留扑晰饱臣辨鸳捡焰微凑绵川语蔷赣壬占泽社椽溪歼几触躲售掺矿尖泊陵承次耽整尝裕闻咎诬淘敌誉兢鹰弃缩伞膘棚传创容胎础尿帐悯芍森但呆蜡舰胯局憎炮柳佣唤贼裴胺犁咎非听椎刑尖事渗闸贼落势杨付戒旦溶栗垦赣瘸词载稍冠妄窖很详企翔萤级长真派厦汉榷酶忿租职惟鹃蹿绽务激滇呢刃簿氟骸闭痪谁袭疮漫吮慎锰惹种挫呜蔽兽朱崖驮架哨嘲渣感陨耶彦金碧挛篇辰锰愤犊哩邱小疥尸歉粟妊迅慢己筹苍年耸性阶祖劈岳殴欣咏劳他渗凡隘卒虾迷炒激冤阮竹江陇妈砧单冲序井起想份引顿榜靠荫祖千妄朔校瞪撩浪碎砰辈酸和因通

**AI赋能大学计划**

**阿里魔搭社区AI大模型技术与应用**

**实战训练营选题报告**

**学校名称： 吉林大学**

**小组名称： 赛博考古学家**

**项目名称： 未来遗迹：时间之旅**

**2024年10月20日**

目 录

[一、 项目介绍 1](#_Toc29881)

[1.1 项目简介 1](#_Toc5968)

[1.2 小组成员介绍 1](#_Toc13973)

[二、 项目整体说明（项目概述、功能介绍） 2](#_Toc30234)

[三、 实现思路 4](#_Toc2115)

# 项目介绍

## 项目简介

|  |  |
| --- | --- |
| 项目名称 | 未来遗迹：时间之旅 |
| 项目内容 | 使用LoRA模型成不同历史时期的图像以及每个时期的描述文本，确保每个时期的人物、服饰、建筑和环境都具有高度的一致性和历史准确性。 |

## 小组成员介绍

|  |  |  |
| --- | --- | --- |
| 序号 | 姓名 | 项目分工 |
| 1 | 阎志 | 队长，项目制作，答辩，PPT制作 |
| 2 | 满李晨 | 项目制作 |
| 3 | 陈超 | 项目制作，答辩 |

# 项目整体说明（项目概述、功能介绍）

**故事背景**：

在一个遥远的未来，人类不仅掌握了时间旅行的技术，还将其运用于考古学领域。主角是一位对古代文明充满无尽好奇的年轻考古学家，他决定利用时间机器穿梭于不同的历史时期，亲身体验那些失落的文明，揭开历史的神秘面纱。

**故事梗概：**

**1**. 在未来高科技实验室中，考古学家穿上特质防护服，准备时间旅行的设备，检查复杂的时间机器。

**2**. 古埃及探险，主角抵达古埃及，站在金字塔建造现场，周围是沙漠，象形文字在夕阳下显得格外神秘。

**3**. 唐朝体验，主角来到了繁华的唐朝，站在宏伟的古建筑前，唐朝中国风格的建筑和文化展现在眼前

**4**. 中世纪欧洲战斗，主角来到了中世纪的欧洲，哥特式城堡耸立，铠甲和剑的碰撞声中，他近距离观察了城堡的防御战斗。

**5**. 未来世界探索，主角来到了未来世界，一个由AI主导的高科技文明，悬空城市和交通系统展现了科技的极致。

**6**. 时间混乱，时间机器开始出现故障，时间线变得混乱，不同时代的时空混杂错乱。

**7**. 时间困局，时间机器彻底损坏，主角被困在了一个文明废墟的遗迹里，这些遗迹与高科技元素融合，揭示了人类历史的连续性和交织。

**8.** 经过一系列的冒险和挑战，主角最终修复了时间机器，带着对历史的深刻理解和新认识回到了其所在时代。

**视觉风格：**

- 古埃及风格：沙漠色调为主，金字塔和象形文字的元素，强调日落时分的暖色调和沙漠的广袤无垠。

- 唐朝风格：中国风风格，展现唐朝的服饰和建筑，强调宏伟的宫殿和繁华的市井生活，以及服饰的细节和色彩。

- 中世纪欧洲风格：哥特式建筑和铠甲的元素，展现中世纪的战争和生活，强调城堡的壮观和战争的紧张气氛。

- 未来世界风格：赛博朋克风格，展现高科技和人工智能的元素，强调霓虹灯光和未来城市的天际线，以及AI机器人和全息投影的科技感。

- 未知时代风格：混合风格，结合了不同历史时期的元素，创造出一种神秘和未知的氛围，强调时间的交错和文化的融合，创造一种神秘和未知的氛围。

**技术应用：**

LoRA模型生成图像：

历史时期图像生成： 使用LoRA（Low-Rank Adaptation）模型生成不同历史时期的图像。LoRA模型通过在预训练的大型模型上引入低秩矩阵来调整模型参数，实现对特定任务的微调，而无需重新训练整个模型。

一致性和准确性： 确保生成的图像中人物、服饰、建筑和环境在不同历史时期都具有高度的一致性和历史准确性。

LLMs生成描述文本：

图文相关性提升： 利用LLMs生成每个时期的描述文本，通过自然语言处理技术提高图文之间的相关性，使得生成的图像与描述文本在内容和情感上更加匹配。

LoRA模型的多风格训练：

特定风格模型训练： 探索使用不同的LoRA rank来训练特定风格的模型，以提高图像质量，尤其是在写实人像、艺术风格及复杂场景的生成上。

DiffSynth-Studio训练框架：

开源项目利用： 本项目核心部分训练框架基于开源项目DiffSynth-Studio，该框架支持在Kolors模型的基础上训练LoRA模型，为我们的项目提供了强大的技术支持。

数据集多样化：

数据集丰富性： 更换原有的单一风格数据集，采用包含600多种风格的种类丰富的新数据集，使得微调过程更加丰富，提高了模型的泛化能力和图像的多样性。

图像与文本匹配过滤：

BLIP模型过滤： 使用BLIP模型来过滤图像和文本数据，只有当图像和文本的匹配分数在0.6到1.0之间时，样本才会被保留。

CLIP模型评估： 使用CLIP模型来评估图像和文本之间的相似度，只有相似度分数在0.6到1.0之间的样本才会被保留。

去重技术应用：

ahash算法： 使用ahash算法考虑文本信息来帮助识别重复的图像，从而去除重复的图像，保证生成图像的独特性。

**实现思路：**

Kolors是第三代文本到图像生成模型，支持文生图和图生图功能，参数规模达十亿级，依托自研的中文CLIP模型，基于Stable Diffusion架构开发，专注于提升写实人像、艺术风格及复杂场景的生成质量。

LoRA模型是一种用于微调大型语言模型的技术，它通过引入低秩矩阵来高效地调整模型参数，而不需要重新训练整个模型。

本项目核心部分训练框架基于开源项⽬ DiffSynth-Studio，实现在可图Kolors模型的基础上训练LoRA 模型。总体思路：

1. 利用LLMs辅助生成每个时期的描述文本，提高图文相关性。
2. 使用LoRA模型生成不同历史时期的图像，确保每个时期的人物、服饰、建筑和环境都具有高度的一致性和历史准确性。
3. 探索使用不同的参数、提示词来训练特定风格的模型，以提高图像质量。
4. 将训练好的LoRA 模型上传到魔搭模型库。