**北 京 邮 电 大 学**

**本科毕业设计（论文）任务书**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 学院 | 人工智能学院 | | | 专业 | | 智能科学与技术 | |
| 学生姓名 | 罗彬慈 | | 学号 | 2020212053 | | 班级 | 2020219107 |
| 指导教师姓名 | 李佩佩 | | 所在单位 | 人工智能学院 | | 职称 | 副教授 |
| 设计(论文)题目 | （中文）基于LLM的交互式多模态图像编辑系统的设计与搭建 | | | | | | |
| （英文）Design and Construction of Interactive Multimodal Image Editing System Based on LLM | | | | | | |
| 题目类型 | 工程实践类□ 研究设计类☑ 理论分析类□ 文献综述类□ 其他□ | | | | | | |
| 题目来源 | 题目是否来源于科研项目 是☑ 否□ | | | | | | |
| 科研项目名称： | | | | | | |
| 科研项目负责人： | | | | | | |
| 主要内容：  一：熟悉图像编辑技术和生成模型的相关知识；  支撑指标点：☑1.6 □2.1 □2.3 □3.1 □3.2 ☑4.1 □5.3 □10.1 □11.2  二：分析传统图像编辑模型的限制和挑战；  支撑指标点：□1.6 ☑2.1 ☑2.3 □3.1 □3.2 □4.1 □5.3 ☑10.1 □11.2  三：设计和构建交互式图像编辑系统；  支撑指标点：□1.6 □2.1 □2.3 ☑3.1 ☑3.2 □4.1 □5.3 □10.1 ☑11.2  四：评估系统性能并进行质量控制；  支撑指标点：□1.6 □2.1 □2.3 □3.1 □3.2 □4.1 ☑5.3 □10.1 □11.2 | | | | | | | |
| 主要(技术)要求：  内容一：熟悉图像编辑技术和生成模型的相关知识；  1.6 掌握图像生成与语言大模型基础知识及原理，能够将其和计算机知识与原理、数学与工程方法以及计算求解能力用于分析和解决复杂工程问题，并能够对解决方案进行比较和综合。  4.1 能够采用科学方法，通过文献研究和应用案例分析等方法，调研和分析领域图像生成与语言大模复杂工程问题的解决方案。  内容二：分析传统图像编辑模型的限制和挑战；  2.1 针对交互式图像编辑领域的复杂工程问题进行问题识别，分析其功能需求与非功能需求，识别其面临的各种制约条件，对任务目标给出需求描述。  2.3 针对已建立的交互式图像编辑领域的复杂工程问题的抽象模型，论证模型的合理性；并通过文献研究，针对改进的可能性进行分析，确定解决方案，获得有效结论。  10.1 能够以撰写报告、设计文稿、口头陈述等方式，针对交互式图像编辑领域复杂工程问题，与业界同行及社会公众进行有效的沟通和交流。  内容三：设计和构建交互式图像编辑系统；  3.1 了解交互式图像编辑系统开发的一般流程，掌握交互式图像编辑系统开发及工程化的基本方法和技术。  3.2 能够针对特定需求，对交互式图像编辑问题进行分解和细化，具有设计/开发功能模块及智能系统的能力。  11.2 能够在多学科环境下，在设计开发交互式图像编辑系统解决方案的过程中，运用工程项目管理与经济决策方法。  内容四：评估系统性能并进行质量控制；  5.3 能够针对交互式图像编辑系统中的具体问题，开发满足特定需求的现代工具，进行仿真和测试，并能够分析其局限性。 | | | | | | | |
| 主要参考文献：  [1] ACHIAM, Josh, et al. Gpt-4 technical report. arXiv preprint arXiv:2303.08774, 2023.  [2] FLORIDI, Luciano; CHIRIATTI, Massimo. GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 2020, 30: 681-694.  [3] ROMBACH, Robin, et al. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022. p. 10684-10695.  [4] HU, Edward J., et al. Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685, 2021.  [5] ZHANG, Lvmin; RAO, Anyi; AGRAWALA, Maneesh. Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. p. 3836-3847.  [6] VAN HUYNH, Nguyen, et al. DeepFake: Deep dueling-based deception strategy to defeat reactive jammers. IEEE Transactions on Wireless Communications, 2021, 20.10: 6898-6914. | | | | | | | |
| 进度安排：  第1阶段（2023.11.20 - 2024.12.03）:  开始调研基于深度神经网络的图像生成模型，包括对抗生成模型（GANs）、扩散模型等，以了解它们的基本原理和应用领域。  准备开题报告，明确研究目标和方法。  第2阶段（2023.12.04 - 2023.12.17）:  深入研究复现的图像生成模型，理解其优点和限制，探索其在图像编辑任务中的潜在应用。  第3阶段（2023.12.18 - 2023.12.31）:  针对图像编辑任务，提出创新性的问题，明确解决思路。  第4阶段（2024.01.01 - 2024.01.14）:  进行图像编辑模型的实验，收集和分析实验结果，识别性能瓶颈和问题。  研究已有文献和代码，选择合适的图像生成模型进行复现，确保对模型的理解和实现代码能力。  根据图像编辑任务的需求，设计和实施性能提升的方案，可能包括模型改进、数据增强等。  同时开始调研语言大模型，了解其发展现状和应用领域，确保掌握如何使用语言大模型。  第5阶段（2024.02.26 - 2024.03.10）:  结合语言大模型的研究，探讨如何将语言大模型与图像编辑模型相结合，构建交互式图像编辑系统。  第6阶段（2024.03.11 - 2024.03.24）:  进行实验，评估交互式图像编辑系统的性能，进行调整和改进。  第7阶段（2024.03.25 - 2024.04.07）:  文献更新和总结，更新文献调研，将最新研究成果与自己的工作相结合，确保研究与学术前沿保持同步，准备中期报告。  第8阶段（2024.04.08 - 2024.04.19）:  完善研究和论文撰写，完善能量模块的功能，并开始论文撰写，完成研究项目，准备最终论文和答辩 | | | | | | | |
| 指导教师签字 | |  | | 日期 | 2023年 11月 15日 | | |