

北 京 邮 电 大 学

本科毕业设计（论文）开题报告

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）基于对话系统的交互式图像编辑的研究与仿真				
	（英文）Research and Simulation of Interactive Image Editing Based on Conversational Systems				
<p>1. 背景和意义</p> <p>随着图像生成技术的不断发展，图像编辑作为其中的关键技术之一，应用广泛，涵盖了媒体娱乐、数字营销和智能医疗等多个领域。然而，传统的图像编辑模型存在着交互性差和生成图像质量受限的问题，迫使我们探索更先进的方法以提高图像生成的质量和用户交互性。通过深度学习和语言大模型的结合，我们有望构建一个创新的交互式图像编辑系统，为图像编辑领域带来新的可能性。</p> <p>2. 研究的基本内容和拟解决的主要问题</p> <p>内容一：熟悉图像编辑技术和生成模型的相关知识</p> <p>在这一部分，基本内容是深入了解图像生成领域的相关知识，包括图像编辑技术、生成模型的原理，以及语言大模型在图像编辑中的应用。这为后续交互式图像编辑系统的构建奠定了理论基础和技术背景。</p> <p>内容二：分析传统图像编辑模型的限制和挑战</p> <p>在这一部分，基本内容是深入研究传统图像编辑模型，识别其存在的问题，特别关注交互性和生成图像质量的限制，为改进图像编辑技术提供方向。</p> <p>内容三：设计和构建交互式图像编辑系统</p> <p>在这一部分，基本内容是明确交互式图像编辑系统的设计目标和实现方式。通过考察的不同参数的多种语言大模型，选择在性能和负载中取得平衡的大语言模型，然后设计系统的核心算法和交互界面。</p> <p>内容四：评估系统性能并进行部署</p> <p>在这一部分，基本内容是从多个纬度评估系统性能，包括生成图像的质量、交互系统对用户请</p>					

求的处理能力和用户体验。通过不断的改进和优化，确保系统能够满足需求，提高系统的可用性和易用性。选择合适的方式部署系统，是系统能够提供服务。

### 3. 研究方法及措施

在解决上述问题的过程中，我们将采用以下研究方法和措施：

（1）深入调研与学习：对深度学习和神经网络的基本原理进行学习，特别关注对抗生成模型和扩散模型等图像生成模型的详细了解。同时，调研语言大模型的最新发展。

（2）总结与优化：分析当前图像编辑模型的创新之处和局限性，提出改进图像编辑模型的方法，以提高其性能。

（3）语言大模型应用：学习如何使用语言大模型，并探索其在图像编辑任务中的应用，为系统集成做好准备。

（4）系统构建与评估：设计并实现一个交互式图像编辑系统，结合深度神经网络的图像生成模型和语言大模型，通过实验评估系统性能。

### 4. 研究工作的步骤与进度

#### 第1阶段（2023.11.20 - 2024.12.03）：背景调研与开题准备

开始调研基于深度神经网络的图像生成模型，了解对抗生成模型（GANs）、扩散模型等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）：图像生成模型深入研究

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）：问题定义与解决思路

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）：实验与性能评估

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。

学习语言大模型的使用方法，为系统集成做准备。

#### 第5阶段（2024.02.26 - 2024.03.10）：语言大模型与系统集成

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）：实验与系统调优

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）：文献更新与中期报告准备

更新文献调研，将最新研究成果与自己的工作相结合，准备中期报告。

第 8 阶段（2024.04.08 - 2024.04.19）：论文撰写与最终总结

完善研究和论文撰写，确保项目完善，准备最终论文和答辩。

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.

允许进入毕业设计（论文）下一阶段：是 <input type="checkbox"/> 否 <input type="checkbox"/>		指导教师	
日期	年    月    日		

注：可根据开题报告的长度加页