

Human-Computer Interface

Survey Report



实验名称	AI 图像生成平台的用户交互设计对比与优化
小组成员	刘震
	謝文軒
学院	计算机科学与技术学院
专业	软件工程
任课教师	梁爽
日期	2025.3.27

1 引言

随着人工智能技术的快速发展，AI 图像生成工具在创意设计、艺术创作等领域取得了显著突破。MidJourney、DALL-E 和 Stable Diffusion 是目前市场上最受欢迎的三款 AI 图像生成工具。它们利用深度学习和生成对抗网络（GAN）等技术，能够生成高质量的图像，极大地激发了创作者的创意。从用户交互设计方面，他们也有着不同的优缺点，这直接影响了用户的使用体验和创作效率。

本报告旨在对 MidJourney 与 DALL-E、Stable Diffusion 这三款工具的用户交互设计进行比较分析，并讨论以下问题：

- 1. MidJourney 与 DALL-E、Stable Diffusion 的用户交互设计有何异同？
- 2. 这些差异如何影响用户的使用体验？
- 3. 如何优化这些工具的设计，以提升其可用性和功能性？

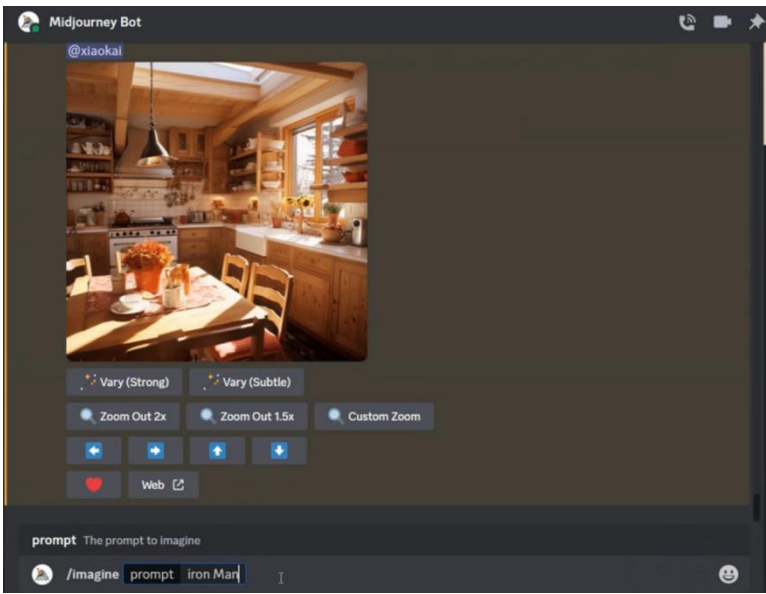
2 针对用户交互技术的分析与比较

2.1 MidJourney 平台

2.1.1 使用方式与用户交互分析

MidJourney 是一款基于人工智能的图像生成工具，用户可以通过在 Discord 服务器或在网页版界面中与 MidJourney Bot 交互，输入文本提示词，即可生成高质量的图像。MidJourney 的用户交互设计注重简洁与直观，主要体现在：

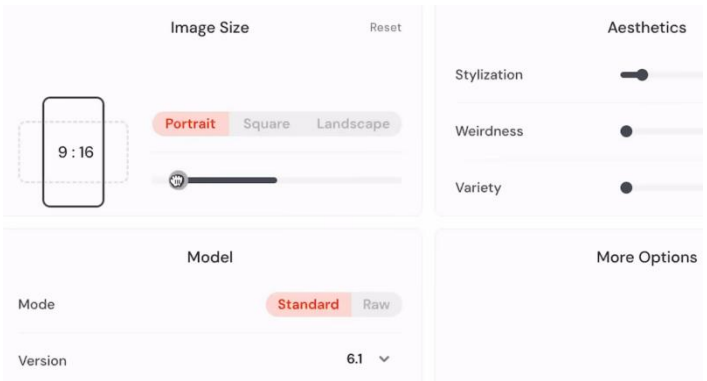
命令式交互:用户通过在 Discord 中输入特定命令（如 /imagine, /ask, /fast）与 MidJourney Bot 进行交互，操作流程清晰明了。



风格化控制:MidJourney 提供丰富的风格化词汇，用户可以在提示词中加入特定的风格描述，以影响生成图像的艺术效果。

参数调整:MidJourney 的参数调整方式非常灵活，允许用户根据自己的创意需求进行个性化定制。用户可以通过简单的文本命令调整图像的风格、质量、比

例等，也可以通过网页版的图形化界面调整，提供了很大的自由度。



2.1.2 优势与劣势

MidJourney 能够生成高质量的艺术性和创意性图像，满足用户在视觉表达上的需求。通过提供多样的风格化选择，用户可以轻松探索不同的艺术风格，提升创作的多样性。此外，MidJourney 通过 Discord 平台的社区互动功能，使用户能够在创作过程中获得反馈、交流创意，从而增强使用体验。

MidJourney 提供了强大的功能，但其操作存在一定的学习曲线。另外，高质量图像的生成需要较高的计算资源，这会影响生成速度，尤其在使用高峰期。同时，MidJourney 对生成内容有一定的审核机制，这可能会限制某些类型图像的创作，影响用户的创作自由度。

2.2 DALL-E 平台

2.2.1 使用方式与用户交互分析

DALL·E 是由 OpenAI 开发的图像生成模型，能够根据用户输入的文本描述生成高质量、细节丰富的图像。用户通过简洁的文本提示，即可生成符合描述的图像，操作过程直观且高效。

其使用方式与界面和普通针对文字的大模型中用户输入文字、模型给出答复的界面类似。



2.2.2 优势与劣势

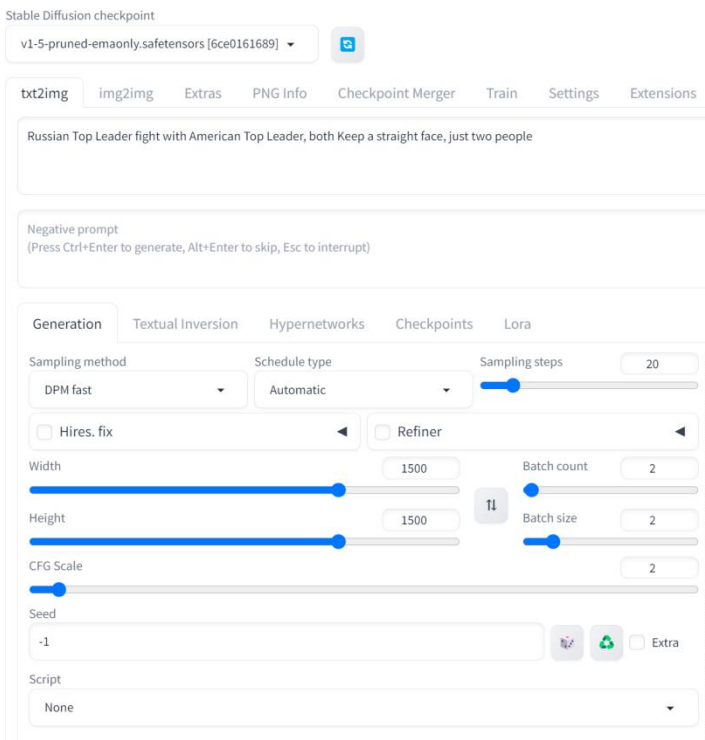
DALL·E 能够根据文本描述生成高质量、细节丰富的图像，满足用户在创意设计和艺术创作中的需求。用户只需输入简单的文本提示，就能生成符合描述的图像，操作流程简单直观，极大地降低了使用门槛。

但是，DALL·E 的图像生成控制精度相对有限，用户对生成结果的影响主要依赖于文本提示的准确性，缺少对图像细节和风格的精细调整功能。

2.3 Stable Diffusion 平台

2.3.1 使用方式与用户交互分析

Stable Diffusion 是一款强大的开源图像生成模型，能够根据文本描述生成高质量的图像。它结合了生成对抗网络（GAN）和自回归模型的技术，支持用户根据输入的文字描述生成图像。用户可以通过简单的文本提示，结合多个可调参数（如迭代步数、采样方法等），生成符合需求的图像。此外，Stable Diffusion 支持在生成过程中进行多样化的自定义设置，使得用户可以精确控制生成图像的风格和质量。



2.3.2 优势与劣势

Stable Diffusion 的优势包括其高质量与多样性，能够生成高质量且多样化的图像，并支持高度自定义，满足不同用户的创作需求。用户可以精确控制图像的细节，如风格、颜色和构图，确保最终创作结果符合预期。

Stable Diffusion 的劣势在于操作复杂性，对于新手用户而言，丰富的参数选

项可能导致学习曲线较为陡峭，且缺乏适当的引导。生成高质量图像时，该平台还需要消耗大量计算资源，特别是在使用高分辨率和多次迭代时，生成速度较慢，可能导致较长的等待时间。

3 调研结果与比较分析

平台	优点	缺点
MidJourney	1. 高艺术性和创意性 2. 丰富的风格选择，允许个性化表达。 3. 社区支持强，交流创意便捷。	1. 命令行和参数调整较为复杂，学习曲线较陡。 2. 高计算资源消耗，生成速度慢。
DALL·E	1. 简单直观，用户输入文本即可生成图像。 2. 高质量的图像生成，细节丰富。	1. 控制精度有限，无法细致调整图像细节和风格。 2. 高计算资源消耗，生成速度慢。
Stable Diffusion	1. 高度自定义，允许精细控制细节。 2. 丰富的参数设置，适合个性化创作。 3. 开源，灵活性强。	1. 操作复杂，学习曲线较陡，尤其是新手用户。 2. 高计算资源消耗，尤其在高分辨率生成时较慢。

4 针对不同平台的改进建议

4.1 MidJourney

增强图形化界面（GUI）交互：目前依赖于命令行输入和参数设置，可能导致新用户在使用过程中感到困惑。建议开发一个简化的图形化界面，提供一系列模板和预设选项，减少复杂的命令输入，提升交互流畅度。

引入智能提示系统：为了减少学习曲线，可以通过智能提示功能，自动为用户提
供常用参数设置或推荐的风格，帮助用户快速获得理想的图像生成效果。

4.2 DALL·E

改进用户输入反馈：目前用户输入文本后等待生成结果，可能存在一定的时间延迟。可以优化反馈机制，让用户在等待期间得到实时的生成进度或提示，让整体体验更具互动性。

增强输入辅助工具：提供更多的自动化输入建议，特别是在文本描述上。可以引入“关键词推荐”功能，帮助用户更精确地表达想法，从而提升图像生成的质量与精度。

4.3 Stable Diffusion

简化参数调节界面：由于参数设置复杂，且功能强大，许多新用户可能感到困扰。可以优化用户界面，将常用的参数设置进行分类和快捷化，允许用户通过滑动条、下拉菜单等方式快速调整图像生成的各个细节。

增加实时反馈机制：在用户调整参数时，提供即时的预览效果，减少生成过程中的不确定性，提升用户对结果的可控性。

4.4 总结

在对 MidJourney、DALL·E 和 Stable Diffusion 这三款 AI 图像生成工具的用户交互设计进行对比分析后，我们可以看到，虽然它们都在创意设计领域具有强大的潜力，但每个平台的用户交互体验各具特点。MidJourney 以其强大的艺术风格转换功能和社区支持赢得用户喜爱，但操作的复杂性和学习曲线较高，使得部分新用户面临一定的挑战。DALL·E 的文本生成方式简单直观，适合快速创作，但它缺乏对生成细节的精细控制，用户对图像的个性化调节受限。Stable Diffusion 虽然在功能和灵活性上具有极大优势，能够精准控制图像的生成，但其过于复杂的参数设置和较高的计算资源需求使得新手用户在使用上感到困惑和困难。总体而言，这三款工具在创意表现力和生成效果上都具有显著优势，但在用户交互设计上尚有改进空间。

随着 AI 图像生成技术的不断进步，用户交互设计的优化将成为提升用户体验的关键因素。我认为，未来的改进应聚焦于简化平台操作、提升实时反馈效果，并引入智能化辅助功能，如自动推荐、智能提示和快速预览等。这将帮助用户更轻松地驾驭平台功能，减少学习成本，增强创作的效率与灵活性。同时，为了让这些工具更具普适性，平台应增强对细节调整的精确控制，并优化计算资源的使用，降低等待时间，提升图像生成的速度和质量。通过这些改进，AI 图像生成工具将能够更好地服务于创作者，推动创意产业的进一步发展，尤其在艺术创作、广告设计、影视制作等领域，赋予用户更加高效、流畅的创作体验。

5 参考文献

1. "Huge "foundation models" are turbo-charging AI progress". The Economist. Archived from the original on November 17, 2022. Retrieved June 26, 2022.
2. Hertzmann, Aaron (June 10, 2022). "Give this AI a few words of description and it produces a stunning image – but is it art?". The Conversation. Archived from the original on June 10, 2022. Retrieved June 26, 2022.
3. David, Emilia (20 September 2023). "OpenAI releases third version of DALL-E". The Verge. Archived from the original on 20 September 2023. Retrieved 21 September 2023.
4. "OpenAI Platform". platform.openai.com. Archived from the original on 20 March 2023. Retrieved 10 November 2023.
5. Borji, A. (2022). Generated Faces in the Wild: Quantitative Comparison of Stable Diffusion, Midjourney and DALL-E 2. *arXiv preprint arXiv:2210.00586*.
6. Digital Art Face-Off: Stable Diffusion, DALL-E, and Midjourney. Retrieved from <https://hub.baai.ac.cn/view/21656>