

*For citation:* Feng Junlong, Long Junhao. Design and implementation of AI-assisted painting tools based on MidJourney // Grand Altai Research & Education — Issue 1 (21)'2024 (DOI: 10.25712/ASTU.2410-485X.2024.01) — EDN: <https://elibrary.ru/GHGFMI>

UDK 004.89

## DESIGN AND IMPLEMENTATION OF AI-ASSISTED PAINTING TOOLS BASED ON MIDJOURNEY

*Feng Junlong<sup>1\*</sup>, Long Junhao<sup>2</sup>*

1 Hubei Digital Textile Equipment Key Laboratory, Wuhan Textile University, Wuhan, 430073, China  
E-mail: \*2419785330@qq.com ; 2889373053@qq.com

**Abstract.** With the rapid development of Internet technology, a large number of AI tools have emerged. Among them, the field of painting is the focus of people's attention. Through AI painting tools, people can simplify the painting process, which greatly facilitates people of different levels to create in the field of painting. However, the use of MidJourney is very complicated and consumes users' time. To solve this problem, this paper proposes a new architecture. Based on the Spring Boot architecture, it realizes interaction with MidJourney's official painting tool, combines the database advantages of Redis and MySQL, and introduces the message middleware RabbitMQ to achieve asynchronous communication and task decoupling within the system. The entire system aims to provide users with a stable and efficient platform for obtaining painting resources and promote innovation in the field of painting creation. The SparkDesk interface has been added to the system to realize the universal question and answer function of large models. The CRISPE framework is used to optimize the prompts used in the MidJourney and Spark model interfaces to further optimize the user experience.

**Keywords:** Spring Boot; Redis; Spark; painting field; MidJourney; CRISPE

## 基于 MIDJOURNEY 的 AI 辅助绘画工具设计与实现

*冯俊龙<sup>1\*</sup>, 龙君豪<sup>2</sup>*

1 武汉纺织大学, 湖北省数字化纺织装备重点实验室, 中国, 武汉, 430073  
E-mail: \*2419785330@qq.com ; 2889373053@qq.com

**摘要:** 随着互联网技术的快速发展, 大量的 AI 工具涌现, 其中, 绘画领域是人们重点关注的, 通过 AI 绘画工具, 人们可以简化绘画流程, 大大方便了不同层次的人群在绘画领域上创作. 但 MidJourney 的使用非常复杂, 消耗用户时间. 为解决这一问题, 本文提出了一种新的架构模式. 基于 Spring Boot 架构, 实现了与 MidJourney 官方绘画工具的交互, 结合了 Redis 和 MySQL 的数据库优势, 引入

消息中间件 RabbitMQ, 实现系统内部的异步通信和任务解耦. 整个系统旨在为用户提供一个稳定, 高效的绘画资源获取平台, 推动绘画创作领域的创新发展. 系统中加入了 SparkDesk 接口, 实现了大模型万能问答的功能, 使用 CRISPE 框架优化 MidJourney 和 Spark 模型接口所使用的 Prompt, 进一步优化用户体验.

**关键词:** Spring Boot; Redis; Spark; 绘画领域; MidJourney; CRISPE

## 0引言

人工智能在进入艺术领域之前, 首先应用于人的日常生活中. 人工智能在模仿和效法人的智力的能力上, 得到了很多领域的注意和认可, 逐渐被应用于人们的日常生活和工作中, 作为人们脑力和智力的延伸.

随着技术的飞速发展, 绘画领域方面, 人工智能如雨后春笋, MidJourney 作为其中的佼佼者, 以其出色的创作能力和特有的风格获得了许多赞美. 但是对于非专业人士来说, 其使用门槛仍然非常高, 复杂的注册以及操作流程和特定的命令式交互设计使许多用户难以享受其中的便利. 在这个背景下, 研究一种简化操作流程的系统尤为重要. 本系统旨在提供统一接口, 使用文本输入框的形式降低 MidJourney 的操作难度, 使用户能够更加便利地接触到 AI 绘画的领域, 进一步推动 AI 绘画的发展, 启发用户灵感, 吸引更多的人投入 AI 创作的绘画作品中来.

## 1 AI绘画工具的发展现状

MidJourney 的原理主要基于深度学习技术, 特别是生成对抗网络 (GAN) 和扩散模型的应用. GAN 则是由生成器和判别器两部分组成. 生成器的任务是生成尽可能真实的样本, 而判别器则负责鉴别样本的真实性 [2]. 在这个技术环境下, MidJourney 绘画被广泛应用于休闲. 社交媒体, 建筑设计, 动漫人物原型的研发, 以及美术教学. 系统功能基本模块图如图 1 所示.

现阶段 Stable Diffusion 模型的应用同样广泛. 它可以用于文生图, 图生图等多种任务, Stable Diffusion 的模型处理可以浅分为三种类型: Dream booth 训练, Lora 训练以及运用最广泛的文本逆向即 Textual Inversion 训练. 目前来看, Dream Booth 的配置要求最高, Lora 次之, 而 Textual Inversion 以其配置要求低, 训练模型效率高而收到广泛的喜爱 [3]. Stable Diffusion 是通过 Text Encoder 处理进入 Noise Predictor 进行噪声预测, 最后使用 VAE 进行变形编码, 生成最后的图片, 但缺点是较长的采样时间和较大的模型容量. 由于稳定扩散 (stable diffusion) 模型是开源算法, 并且可以在大部分普通消费级 GPU 上运行, 因此得到了广泛的应用 [4].

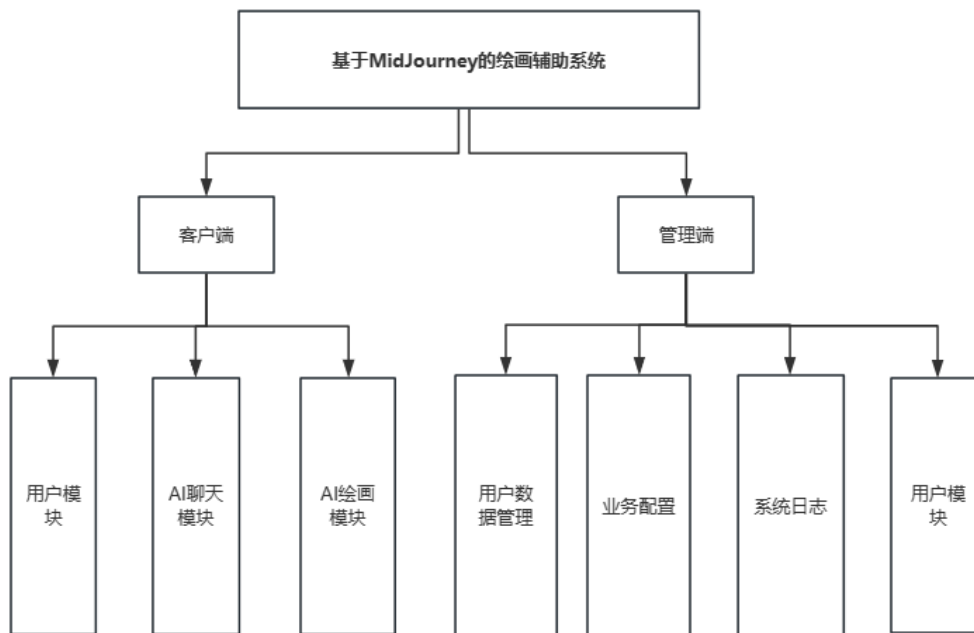


图1 系统功能模块图

Figure 1 System function module diagram

在客户端模块, 用户主要可以进行 AI 聊天与问答和 AI 绘画创作. 其中聊天模块功能参考了市面上大多数的 AI 模型问答窗口, 做到了与市面上问答模型相同的功能. 用户可以新建问答页, 也可继续上一次的问答记录进行操作. AI 绘画模块主要就是实现用户文字生成图片以及图生图功能即 AI 绘画系统辅助用户进行图像创作.

#### (1) SparkDesk 的聊天模块

以下是 SparkDesk 聊天模块各功能图. 如图2 所示:

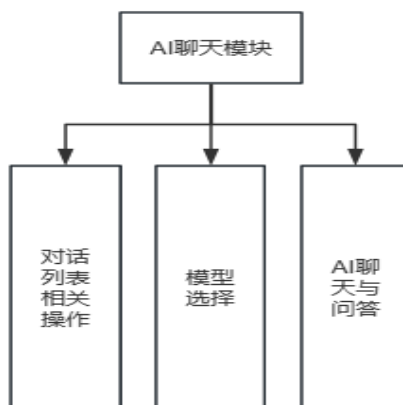


图2 SparkDesk 聊天模块功能图

Figure 2 SparkDesk chat module function diagram

在 SparkDesk 万能问答聊天模块中, 用户点击按钮后可以输入自己的文本对 AI 模型进行操作. 可以询问自己的问题, 也可以让模型给出建议. 在用户聊天时, 左侧栏位会生成当前聊天记录, 用作下次继续聊天的入口. 用户也可以选择

使用不同的模型，如：星火大模型和百度千帆模型，甚至市面上非常火热的 OpenAI 的模型。

## (2) MidJourney 绘图模块

以下是 MidJourney 绘图模块功能图. 如图3 所示:

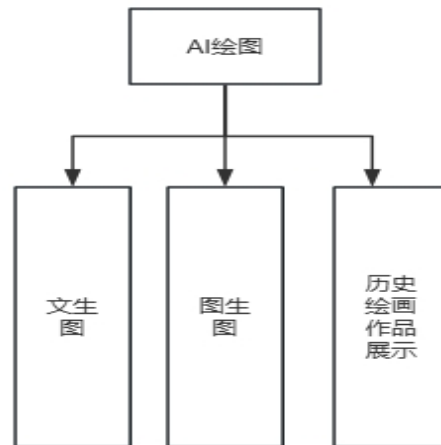


图3 MidJourney 绘图模块功能图

Figure 3. MidJourney drawing module function diagram

## 2 Mid-Prompt 优化-提示词设计

提示词, 是生成图像的短文本提示语, MidJourney 会将提示词, 中提到的单词和短语分解成更小的部分, 叫做 token, 可以将其与其训练的数据进行比较, 然后用于生成图像。

当我们使用中文制作 Prompt 时, 看看效果, 如图4 所示. 可以看到, 用户使用中文红色草莓时, MidJourney 并不能理解我们的提示词和我们的想法, 制作出来的图像往往与初始想法完全不同。

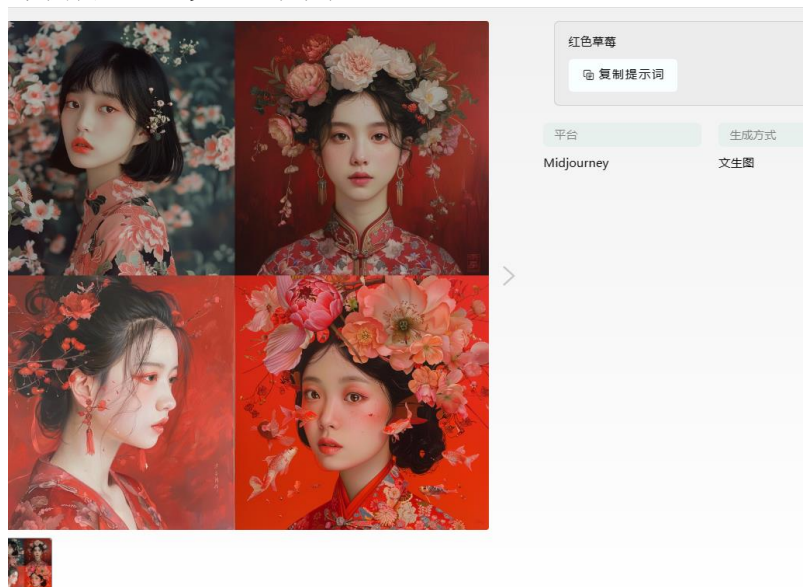


图4 中文Prompt

Figure 4. Chinese Prompt



使用英文制作 Prompt 时, 看看效果, 如图5 所示. 可以看到, 当我们使用英文的时候, 用户可以得到更加优质的图像. 这就是我们设计提词器的初衷.

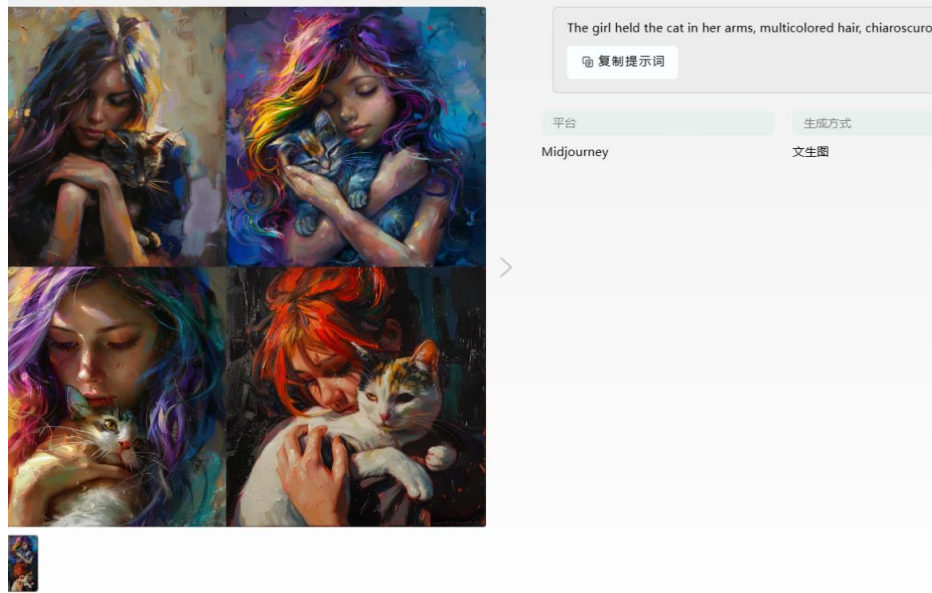


图5 英文 Prompt

Figure 5. Highly thermally conductive, high-strength graphene-based composite films by eliminating nanosheet wrinkles

AI 不太擅长理解人类的语法和句子结构. 单词的选择比较重要, 可以找一些绘画描述的关键词, 而不是自己想当然的描述, 所以 MidJourney 的 Prompt 设计可以参考以下内容: 主题, 媒介, 环境, 照明, 颜色, 情绪, 构图.

当我们使用这种结构的 Prompt 时, 可以产生高质量的图像.

如下图6 所示

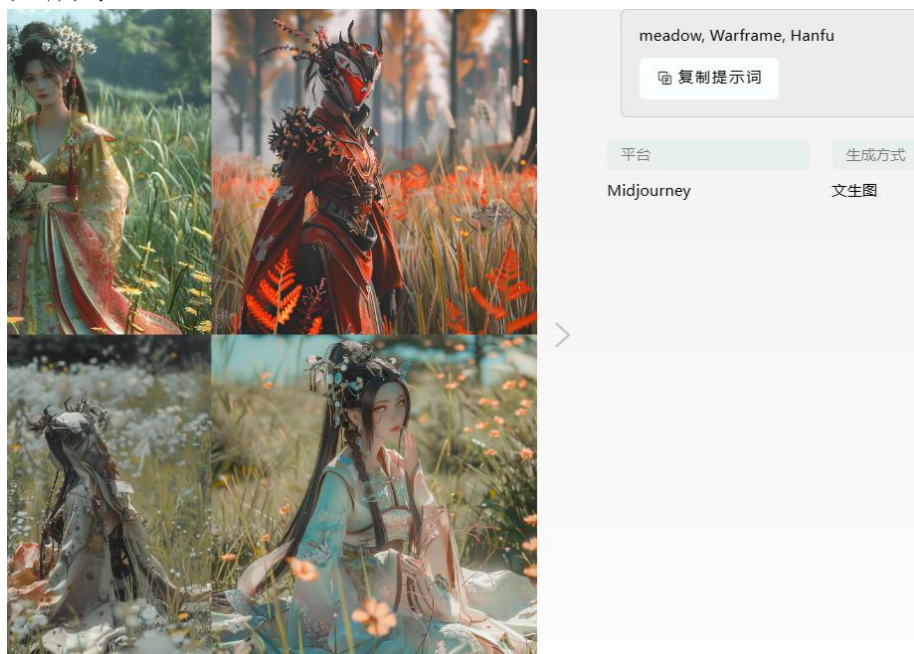


图6 使用结构化 Prompt

Figure 6. Use structured prompts

### 3 Prompt 优化-CRISPE 框架设计

为了更好的使用 Prompt 调试大模型, 我们需要合理使用 Prompt 的优化方式, 这里我们使用的就是 CRISPE 框架, 在这个框架下, 用户可以更有效的与大模型进行互动. 这里我们就得使用 CRISPE 框架, CRISPE 框架即:

**CR:** Capacity and Role (能力与角色). 你希望 AI 扮演怎样的角色.

**I:** Insight (洞察), 提供背景信息和上下文.

**S:** Statement (陈述), 你希望 AI 做什么.

**P:** Personality (个性), 你希望 AI 以什么风格或方式回答你.

**E:** Experiment (实验), 要求 AI 为你提供多个答案.

当我们直接提问 «解释导数概念» 的时候, 大模型给出的答案和教科书给出的解释一致, 不利于初学者掌握知识. 如下图7 所示:

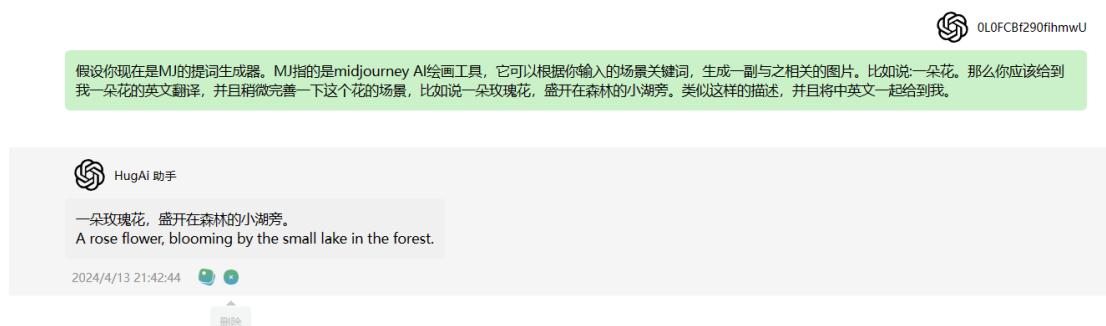


图7 对话训练示例-CRISPE 框架设计

Figure 6. Dialogue training example-CRISPE framework design

### 4 结束语

本系统采用 Java spring boot 技术开发, 使用 MidJourney-proxy 提供的接口方案. 项目思路来源于 GitHub 开源的 chat next web 系列的方案. 系统中使用的 AI 绘画模型方案源自于 MidJourney 提供的 image 操作, 原因在于 MidJourney 提供了 Discord Bot 技术, 使得我们用户可以采用代理的方式通过 Discord 机器人对图像进行操作. 本项目预留了两个接口方案: Stable Diffusion 和 OpenAI 的 AI 绘画接口. 在未来, 相信会有越来越多的人热爱使用人工智能去创作自己的绘画作品, 本系统旨在探索更加便捷的途径去优化 AI 绘画功能, 为此, 将会在本系统中加入 Stable Diffusion 和 OpenAI 的 AI 绘画接口.

### 参考文献

- [1] 王晨薇. 创作论视角下的 «人工智能绘画» 与艺术家绘画比较研究[D]. 南京航空航天大学, 2021. DOI:10.27239/d.cnki.gnhhu.2020.000814.
- [2] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望[J]. 自动化学报, 2017, 43(03):321-332. DOI:10.16383/j.aas.2017.y000003.
- [3] Mokady, Ron, et al. "Null-text Inversion for Editing Real Images using Guided Diffusion Models." ArXiv (2022).
- [4] 马歆嫒. 法理视域下 AI 绘画的未来: 自由与控制的平衡[J]. 智能科学与技术学报, 2023, 5(03):424-430.

- [5] Appleone. (2023-12-05). Prompt 优化 <https://zhuanlan.zhihu.com/p/670552915>.
- [6] 上进 AI 绘画创意基地. 这几个工具太牛了! 一键生成Midjourney提示词! (附 AI 教程). [https://zhuanlan.zhihu.com/p/677432362?utm\\_psn=1762030859241132032](https://zhuanlan.zhihu.com/p/677432362?utm_psn=1762030859241132032).

## References

- [1] Wang Chenwei. Comparative Research on "Artificial Intelligence Painting" and Artist Painting from the Perspective of Creative Theory [D]. Nanjing University, 2021. DOI: 10.27239/D.cnki.gnhhu.2020.000814.
- [2] Wang Kunfeng, Gou Chao, Duan Yanjie, et al. Research progress and prospects of generative adversarial network GAN [J]. Acta Automata, 2017, 43(03): 321-332. DOI: 10.16383/j.aas.2017.y000003.
- [3] Mokady, Ron, et al. "Null-text Inversion for Editing Real Images using Guided Diffusion Models." ArXiv (2022).
- [4] Ma Yuan. The Future of AI Painting from the Perspective of Jurisprudence: The Balance of Freedom and Control [J]. Journal of Intelligent Science and Technology, 2023, 5 (03): 424-430.
- [5] Appleone. (2023-12-05). Prompt optimization <https://zhuanlan.zhihu.com/p/670552915>.
- [6] Shangjin AI painting creative base. These tools are awesome! Generate Midjourney prompt words with one click! (AI tutorial attached). [https://zhuanlan.zhihu.com/p/677432362?utm\\_psn=1762030859241132032](https://zhuanlan.zhihu.com/p/677432362?utm_psn=1762030859241132032).