徽标, 公司名称

描述已自动生成

**2023年（第16届）**

**中国大学生计算机设计大赛**

人工智能实践赛作品报告

作品编号：　　　　　　　　　xxxx

作品名称： 智绘连环--基于LLM和Stable Diffusion的“创意”连环画平台

填写日期：　　　　　2024年4月9日

**目 录**

[第1章 作品概述 1](#_Toc308)

[1.1 创意来源 1](#_Toc10301)

[1.2 产生背景 2](#_Toc32054)

[1.3 用户群体 2](#_Toc24699)

[1.4 主要功能与特色 3](#_Toc3789)

[1.5 应用价值 4](#_Toc16230)

[1.6 推广前景 5](#_Toc28445)

[第2章 问题分析 6](#_Toc32481)

[2.1 问题来源 6](#_Toc2186)

[2.2 现有解决方案 6](#_Toc16294)

[2.3 本作品要解决的痛点问题 8](#_Toc24893)

[2.4 解决问题的思路 8](#_Toc7717)

[2.4.1 功能模块 8](#_Toc24336)

[2.4.2 性能需求 9](#_Toc24410)

[2.4.3 测试数据集 9](#_Toc22411)

[第3章 技术方案 11](#_Toc32059)

[3.1 技术路线框架 11](#_Toc13591)

[3.2 文本理解模块 12](#_Toc22444)

[3.2.1 基于vllm框架的LLM模型推理优化策略 13](#_Toc670)

[3.2.2 基于Qwen1.5-14B-Chat-GPTQ-Int4模型的文本理解 14](#_Toc13863)

[3.3 提示词生成模块 15](#_Toc31006)

[3.3.1 本地系统prompt设计 16](#_Toc26480)

[3.3.2 提示词生成流程 17](#_Toc18184)

[3.4 连环画生成模块 19](#_Toc30728)

[3.4.1 基于Stable Diffusion模型的图像生成 19](#_Toc18269)

[3.5 工作流管理模块 20](#_Toc32076)

[3.5.1 基于Litegraph节点流程框架实现Stable Diffusion ComfyUI 20](#_Toc20198)

[3.5.2 统一风格的自动化批量图像生成 21](#_Toc8790)

[第4章 系统实现 22](#_Toc10236)

[4.1 软件设计实现 22](#_Toc19740)

[4.1.1 前端技术 22](#_Toc22966)

[4.1.2 后端技术 24](#_Toc9576)

[4.1.3 接口封装与调用 25](#_Toc32130)

[4.1.4 系统集成 25](#_Toc21652)

[4.2 用户界面 27](#_Toc15977)

[4.2.1 主要使用界面 27](#_Toc19513)

[4.2.2 用户登录界面 29](#_Toc14224)

[4.3 数据来源 29](#_Toc4903)

[4.4 数据训练 30](#_Toc21433)

[4.5 改进过程 32](#_Toc21355)

[4.5.1 选择vllm框架改进模型推理过程 32](#_Toc7631)

[4.5.2 Qwen系列模型的选择 33](#_Toc31016)

[4.5.3 实现节点流界面的框架选择 34](#_Toc12989)

[4.6 系统部署方法 35](#_Toc29068)

[第5章 测试分析 36](#_Toc18215)

[5.1.1 文本理解模块测试分析 36](#_Toc21040)

[5.1.2 提示词生成模块测试分析 37](#_Toc1979)

[5.1.3 连环画生成模块和工作流管理模块的测试分析 38](#_Toc13465)

[第6章 作品总结 39](#_Toc4402)

[6.1 作品特色与创新点 39](#_Toc11967)

[6.1.1 作品特色 39](#_Toc25917)

[6.1.2 作品创新点 40](#_Toc22720)

[6.2 应用推广 40](#_Toc11248)

[6.3 作品展望 41](#_Toc4390)

[参考文献 41](#_Toc19815)

# 作品概述

## 创意来源

智绘连环平台的创意灵感源于三款前沿技术的融合：Stable Diffusion模型、ComfyUI工作流界面以及LLM（Large Language Model）大语言模型。通过这三者的结合构建一套紧密协作且极具创新性的技术栈，实现从文本到生动连环画的转化。

首先，Stable Diffusion模型作为一款文本到图像生成工具，具备将文本提示词转化为视觉图像，或依据文本指令对现有图像进行编辑的强大能力。

ComfyUI则是一个基于节点流程式的Stable Diffusion AI绘图工具，它将Stable Diffusion的复杂流程拆分成直观易懂的节点，从而实现了更加精准的工作流定制和完善的可复现性。通过ComfyUI，用户可以更加自由地定制工作流，实现个性化的图像生成需求。

更为关键的是，智绘连环平台引入了LLM大语言模型。LLM是一种先进的深度学习模型，专为理解和生成人类语言而设计。它能够深入解析文本的含义，并生成与之相关的提示词。这些提示词能够精确引导Stable Diffusion模型生成符合文本描述的图像。

将这三项技术有机整合，智绘连环平台创新性地使用如下工作流程：首先，利用LLM大语言模型对输入文本进行深度解读，提炼出能够精准驱动图像生成的提示词集合。接着，在ComfyUI工作流界面中，依据这些提示词，灵活定制连环画各帧画面的生成流程，细致调整每一幅图像的风格、氛围、角色动作等细节，确保图像序列连贯且富有表现力。实现提供一段文本生成连环画，这种方式让文本与图像得以深度融合，用户的创作自由度大幅提升，使得生成的连环画既忠实于原文本的叙事脉络，又融入了个性化的艺术诠释。

Stable Diffusion模型、ComfyUI工作流界面与LLM大语言模型的三位一体，共同开创了一种崭新的创意实践方式，将一段文本叙述生动转化为连环画，赋予创作者极大的掌控力与表达空间，拓宽了文生图应用的可能性边界。

## 产生背景

近年来，人工智能技术的发展正迎来前所未有的机遇和挑战，深度学习、自然语言处理、计算机视觉、强化学习等人工智能子领域的关键技术取得了突破性进展。算法模型的规模不断扩大，性能持续优化，AI工具的应用日益广泛，深刻影响着社会的各个领域，如大规模语言模型、高保真度的文生图模型等，展现出强大的泛化能力和创造力。与此同时，强大的计算平台和硬件支持，使得人工智能算法能够处理更加复杂的问题和海量的数据，而数据的爆炸式增长为AI技术的发展提供了丰富的资源，为这些复杂模型的训练和应用提供了强大支撑。人工智能技术的快速迭代和成熟，为内容创作领域带来了变革的契机。

此外，政策和资金的支持也为AI技术的发展提供了有力保障。各国政府纷纷出台相关政策，加大对AI产业的扶持力度，鼓励企业和科研机构加强合作，推动AI技术的研发和应用。同时，资本市场也对AI产业投以高度关注，大量的投资涌入AI领域，为AI工具的研发和推广提供了资金保障。

面对激烈的市场竞争和瞬息万变的消费者偏好，企业和个人创作者迫切需要提升内容产出的速度、质量和多样性，以保持竞争优势。AI工具凭借其自动化、智能化的特点，能够显著提升内容创新效率。

在这一背景下，AI工具的应用领域也在不断拓展。无论在什么行业AI工具都发挥着越来越重要的作用。通过合理的运用，它们能够实现业务流程的优化、决策的智能化和用户体验的提升，为社会的发展注入了新的动力。

## 用户群体

智绘连环平台的用户群体广泛，可能的用户群体包括但不限于教育工作者、科普工作者、创意设计师、广告营销人员、文学作家、插画师：

对于教育工作者，平台的连环画生成能力特别适合将复杂知识转化为直观、生动的教学素材，适用于课堂讲解、课后复习或在线教育环境。教育工作者和教师可以利用智绘连环平台将科学概念、历史事件、文学故事等转化为连环画形式，以提高学生的理解和记忆效果。

对于科普工作者，该平台能够助力科普工作者以连环画形式传播科学知识，吸引公众对科学主题的关注。内容创作者可以借助平台创作引人入胜的科普连环画，用于社交媒体分享、科普网站发布或科普读物制作。

对于创意设计师与广告营销人员，智绘连环平台能够帮助设计师和广告营销人员将品牌故事、产品特性或营销信息转化为连环画，提升品牌形象，吸引潜在客户注意力。连环画可以用于社交媒体营销、网站展示、广告海报、产品手册等多种宣传材料中，增加内容的故事性和互动性。

文学作家可以利用平台将小说、剧本或诗歌中的情节、人物和场景转化为连环画，为文学作品提供视觉补充，增强读者的沉浸感。插画师则可以通过平台快速生成概念草图、故事板或初步视觉呈现，作为创作过程中的辅助工具。

这些用户群体受益于平台的文本深度理解、高度自定义工作流、高效可复现的创作体验以及增强的创作自由度，能够在各自领域内实现内容创新与高效产出。

## 主要功能与特色

主要功能：

智绘连环平台创新性地结合Stable Diffusion模型、ComfyUI工作流界面及LLM大语言模型，实现接收用户输入的文本，通过LLM大语言模型理解文本内容，并生成相应的图像生成提示词。然后，利用Stable Diffusion模型把这些提示词转化为连环画中的各张图像。从单一文本输入到一系列连贯、生动的连环画图像的自动转化，将抽象的文字叙述转化为具象的视觉叙事。

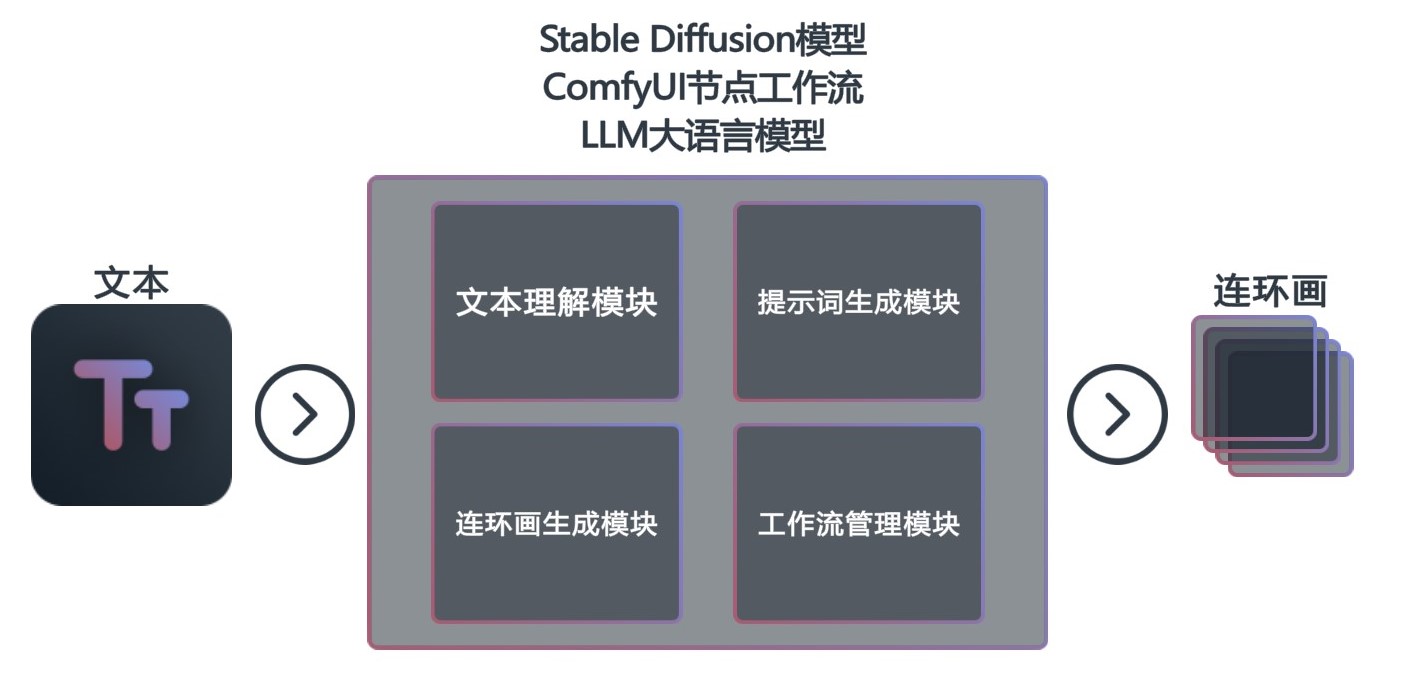


图 1 智绘连环主要功能图

特色：

1. **深度文本理解：**借助LLM大语言模型的深度学习能力，系统能够深入理解输入文本的语义内涵，精准捕捉文本中的关键情节、人物特征、场景描述等元素，为后续图像生成提供丰富的语义指导。
2. **精准提示词生成：**LLM模型将文本解析结果转化为一组针对性强、描述精确的提示词，这些提示词直接对接Stable Diffusion模型，确保图像生成过程严格遵循文本设定，准确再现文本描述的世界。
3. **高度自定义工作流：**通过ComfyUI工作流界面，用户获得对连环画生成过程的精细化控制。该界面将复杂的Stable Diffusion操作流程简化为直观的节点结构，用户可根据需要自由添加、调整节点，定制各帧画面的生成规则，包括但不限于风格设定、氛围渲染、角色动态等，实现个性化连环画创作。
4. **高效可复现的创作体验：**得益于ComfyUI提供的可视化工作流管理，用户可以清晰追踪和修改生成过程，确保创作过程的可追溯性与可复现性，便于反复打磨作品或分享创作经验。
5. **增强用户创作自由度：**智绘连环平台极大地提升了用户在文生图创作中的自主权，使非专业用户也能便捷地运用前沿AI技术，根据个人创意生成高质量、高定制化的连环画作品，拓宽了创意表达的边界。

## 应用价值

本平台通过结合三项前沿技术实现了从文本到连环画的自动化转化，为连环画创作带来了创新性的实现方案。本平台还拥有强大的文本理解能力，大大降低了用户使用门槛，适用于教育、科普、广告、个人创作等多个领域，为用户提供了一种新的、高效的内容表达方式。

1. **教育及科普领域：**在教育教学中，连环画是一种直观且易于理解的教学材料。教师或创作者可以利用本平台将复杂的科学知识、历史故事或文学名著转化为连环画形式，帮助学生更好地理解和记忆。同时，科普工作者也可以利用这一工具制作科普连环画，以生动有趣的方式普及科学知识。
2. **创意设计与广告营销：**设计师和广告商可以利用本平台将品牌故事、产品特点或活动信息转化为连环画，以更具吸引力和故事性的方式呈现给目标受众。这不仅有助于提升品牌形象，还能有效吸引潜在客户的关注。
3. **个人创作与娱乐：**对于普通用户而言，本平台提供了一个发挥创意、表达想法的平台。用户可以将自己的故事、想法或梦境转化为连环画，分享给朋友或社交媒体上的粉丝。此外，连环画还可以作为电子书、漫画或动画的素材，为用户提供更多样化的娱乐体验。
4. **文学与艺术的融合：**本平台为文学家和艺术家提供了一个新的创作工具，使他们能够将文学作品中的情节、人物和场景以连环画的形式呈现，为文学与艺术之间的融合搭建了一座桥梁。这有助于推动文学与艺术的创新与发展，为创作者提供更广阔的创作空间。

## 推广前景

智绘连环--基于LLM和Stable Diffusion的“创意”连环画平台，集成了Stable Diffusion模型、ComfyUI工作流界面和LLM大语言模型三项前沿技术，为用户提供了一个高效、便捷且创新的连环画创作工具。

随着社会对高质量、个性化内容需求的增长，尤其是对于视觉叙事形式的需求，本平台提供的文本到连环画自动化转化服务满足了创作者、教育工作者、广告商等多元用户群体对于快速、精确且富有创意的内容生成工具的需求，具有很大的市场潜力。

其次本平台集成的三项前沿技术——Stable Diffusion模型确保图像生成的高质量与多样性，ComfyUI工作流界面提供用户友好的交互与精细控制，LLM大语言模型确保深度文本理解和精准提示词生成——共同构成了平台的核心竞争优势，使其实现了超越传统创作手段的高效、便捷与创新体验。

同时本平台不仅适用于个人创作者进行艺术表达与娱乐，更能在教育、科普、广告等行业发挥重要作用。无论是作为教学辅助工具提升学生学习兴趣，还是作为营销媒介创新品牌叙事方式，本平台因其具备丰富多样的应用场景，不仅拥有坚实的社会价值基石，更彰显出广泛而深厚的市场接纳潜力。

# 问题分析

## 问题来源

文字叙述虽然能够表达故事情节和人物形象，但往往难以直观呈现具体的场景和细节，而连环画则能够通过图像的方式，将文本信息以直观的形式呈现出来。 而传统的连环画创作过程往往需要创作者花费大量时间和精力进行手绘或设计，效率较低。且传统的绘画创作需要一定的艺术技巧和专业知识，在创作过程还可能受到绘画技巧、材料限制等因素的制约，导致创作自由度受限。

## 现有解决方案

将文本转化成连环画形式表达，目前有以下几种方式：

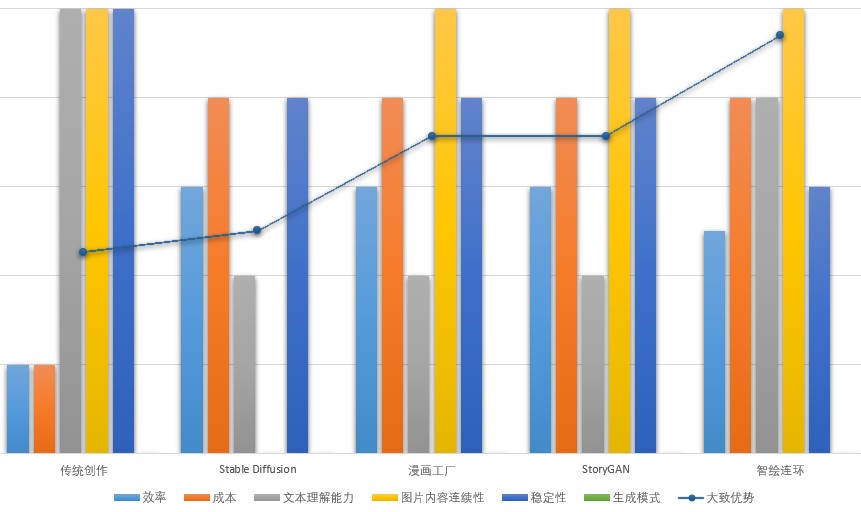
传统的绘画创作者根据文本内容进行手绘或者板绘创作，再基于各种图像编辑工具得到作品。这种方式相对来说效率较低，且成品的质量和风格取决于创作者的创作水平与绘制习惯，成本也较高。

使用目前市面上比较流行的AI工具如Stable Diffusion文生图框架进行图片生成。但是文生图类型的AI工具只能根据提示词生成对应的一张或者几张图片，并不具备生成具有上下文联系关系的连环图的能力

使用专门的将文本生成漫画或连环画的产品和方法，如“漫画工厂”、“StoryGAN”等。这些产品或者方法通常存在生成模式固定、风格固定、文本理解能力不足、使用门槛高等缺点

表 1 文本转化成连环画工具/产品/方法调研

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **工具/产品/方法** | **效率** | **成本** | **文本理解能力** | **图片内容** | **稳定性** | **生成模式** |
| 传统创作 | 低 | 高 | 极高 | 连续 | 无 | 固定 |
| Stable Diffusion | 高 | 低 | 中 | 不连续 | 稳定 | 固定 |
| 漫画工厂 | 高 | 低 | 中 | 连续 | 稳定 | 固定 |
| StoryGAN | 高 | 低 | 中 | 连续 | 稳定 | 固定 |
| 智绘连环 | 中 | 低 | 高 | 连续 | 较稳定 | 可定制 |

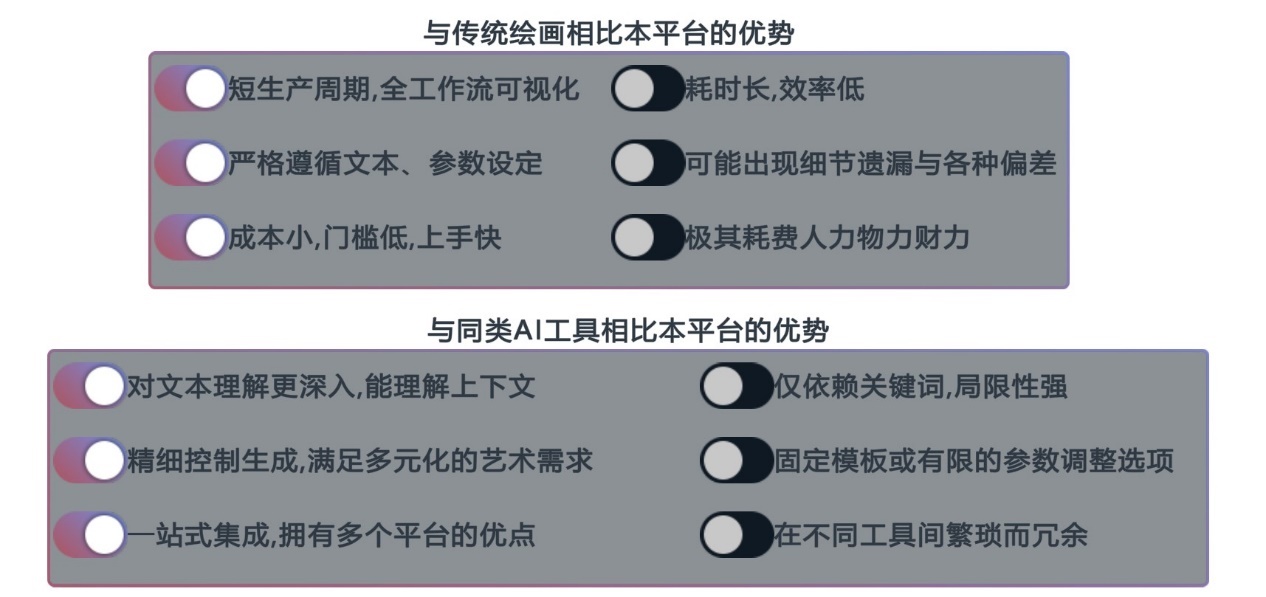


**与传统绘画相比本平台的优势：**

1. **效率与速度：**相较于传统手绘连环画，本平台基于AI技术能够快速将文本输入转化为多帧连贯的图像，极大地缩短了创作周期。传统绘画过程中，从构思、草图、细化到上色等环节往往耗时较长，而该平台能够在短时间内生成高质量的连环画作品，显著提高了创作效率。
2. **一致性与多样性：**平台通过LLM模型精准理解文本并生成精确提示词，再由Stable Diffusion模型生成图像，确保了每一帧图像严格遵循文本设定，减少了人为绘制过程中可能出现的细节遗漏和各种失误问题。同时传统绘制的连环画风格取决于创作者的画风和习惯，而平台在ComfyUI工作流界面中调用Stable Diffusion模型生成图像，可以通过调节各种参数实现多种风格的创作。
3. **节省成本：**传统绘画的成本费用较高，而使用本平台进行的连环画创作成本极低。

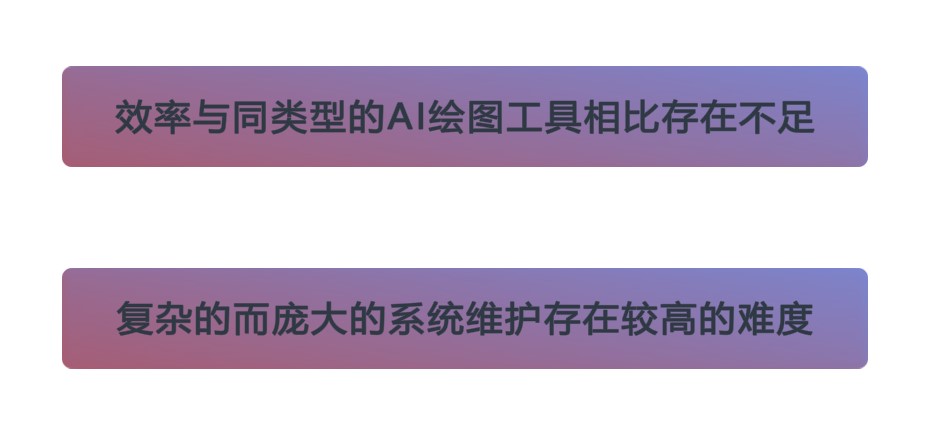
**与同类AI工具相比本平台的优势：**

1. **深度文本理解与精准图像生成：**相较于一些AI绘画工具仅依赖关键词生成图像，本平台结合了LLM大语言模型，能够更深入地理解文本内容，精准捕捉复杂文本中的情感、情境与细节，生成更符合文本描述的图像。这使得连环画作品在情节、人物特征、场景描述等方面更加准确、生动，并保证了上下文的连续性。
2. **高度自定义工作流：**相比于其他AI绘画工具可能提供的固定模板或有限的参数调整选项，该平台的ComfyUI工作流界面允许用户以节点形式精细控制生成过程，包括风格设定、氛围渲染、角色动态等高级属性，提供前所未有的创作自由度，适应多元化的艺术需求。
3. **集成创新与一站式解决方案：**市场上可能存在多种独立的AI绘画工具、文本理解模型和工作流管理软件。而该平台将Stable Diffusion模型、LLM模型与ComfyUI界面无缝集成，为用户提供一站式连环画创作平台，避免了在不同工具间切换的繁琐，提升了整体使用体验与工作效率。

****

**智绘连环平台有待优化之处：**

1. **相较于同类AI绘图工具运行效率的不足：**由于智绘连环平台是基于Stable Diffusion模型、ComfyUI工作流界面以及LLM大语言模型三者的一个结合与创新，虽然平台给用户提供了强大的自定义功能，允许用户高自由度地调整图像的风格和角色等细节，但在实际处理过程牺牲了生成效率。特别是一次性生成多幅图片组合成连环画时，生成时间会比同类AI绘图工具要久，对于原生的其他AI绘画产品来说，本产品的运行效率还有一定的进步空间，后续可就优化算法与模型和升级硬件配置等方面来达到对运行效率的提升。
2. **系统维护的稳定性：**本产品是将多种不同的图像生成技术整合到一个平台中，因此可能会增加系统的复杂性和维护难度。这就意味着要求平台具备强大的技术实力和持续的技术支持，以确保系统的稳定性和性能。

****

图（简单画一张概括优势和缺点的图）

## 本作品要解决的痛点问题

1. 实现对文本的深度理解并转化成Stable Diffusion模型适用的提示词，并保证生成的每一个图像不仅要独立成画，还要与前后图之间保持剧情、角色、视觉风格的高度连贯与一致。
2. 现有的AI绘画工具和传统绘画，用户难以对生成过程进行精细化控制，通过ComfyUI工作流界面将复杂的生成过程分解为可配置的节点结构，让用户能够清晰地看到并干预连环画生成步骤。

## 解决问题的思路

### 功能模块

1. **文本理解模块：**该模块应具备深度文本理解能力，能准确解析用户输入的文本，提取关键情节、人物特征、场景描述等要素。为此，需要集成LLM大语言模型，确保模型具有足够的语义理解深度和广度，能够处理各类复杂文本。
2. **提示词生成模块：**此模块负责将文本理解的结果转化为适合Stable Diffusion模型使用的精确提示词。提示词应充分反映文本内涵，包含必要的细节信息，以便生成连贯、精确的连环画图像。
3. **连环画生成模块：**基于Stable Diffusion模型，该模块应能根据接收到的提示词生成连环画各帧图像，确保图像之间在情节、风格、色彩等方面保持连贯性。
4. **工作流管理模块（ComfyUI）：**提供用户友好的图形化界面，允许用户灵活定制连环画生成流程，包括添加、删除、调整节点以设定各帧图像的生成规则。该模块应具备良好的交互性、易用性，支持实时预览生成效果，并能保存、加载、分享工作流配置。

### 性能需求

1. **文本理解模块：**要求模型对文本的响应速度快、准确率高，能在短时间内生成高质量的提示词集合
2. **提示词生成模块：**要求提示词生成算法高效、稳定，能快速生成与原文本语义高度一致的提示词列表。
3. **连环画生成模块：**保证生成速度、图像质量（分辨率、细节丰富度等）、上下文一致性等，确保生成的连环画视觉效果出色，叙事流畅。
4. **工作流管理模块：**要求界面响应迅速，工作流执行高效，简洁、直观、可控的展示连环画生成流程，便于理解、易于上手，支持大规模连环画项目的管理。

### 测试数据集

* Chinese Squad

1. 数据格式：json
2. 数据来源：https://github.com/junzeng-pluto/ChineseSquad
3. 数据获取方式：下载获取。
4. 数据特点：中文机器阅读理解数据集，通过机器翻译加人工校正的方式从原始Squad转换而来。
5. 数据规模：43498
6. 数据样例：

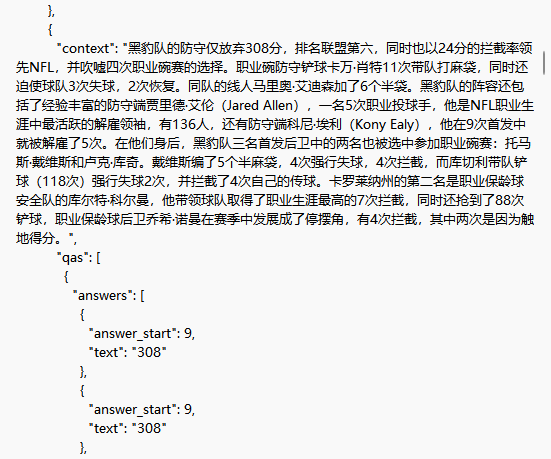


图 2 Chinese Squad数据集样例

1. 数据集用途：作为输入文本，传递给文本理解模块和提示词生成模块，对返回的提示词集合进行评估，根据评估结果进一步完善文本理解模块和提示词生成模块。

* 中文短篇故事数据集

1. 数据格式：txt
2. 数据来源：自制
3. 数据获取方式：原创、收集网络上的短篇故事
4. 数据特点：篇幅较短的有故事性的中文文本
5. 数据规模：100篇
6. 数据样例：

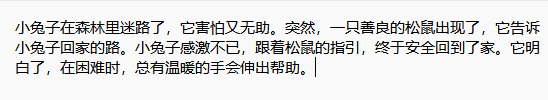


图 3 中文短篇故事数据集样例

1. 数据集用途：用于测试智绘连环平台完整功能，通过提供数据集中的故事文本，对生成的连环画进行运行时间、体验效果、结果质量的评估

# 技术方案

## 技术路线框架

智绘连环平台构建了一套基于文本理解、提示词生成、连环画生成及工作流管理四个关键模块于一体的综合技术框架，以实现从用户提供的文本到连环画图像序列的高效、精准转化，旨在为用户提供高效、个性化且风格统一的连环画创作体验。其中各关键模块间紧密协同，共同驱动连环画创作流程的顺利进行，具体技术路线框架图如下所示。

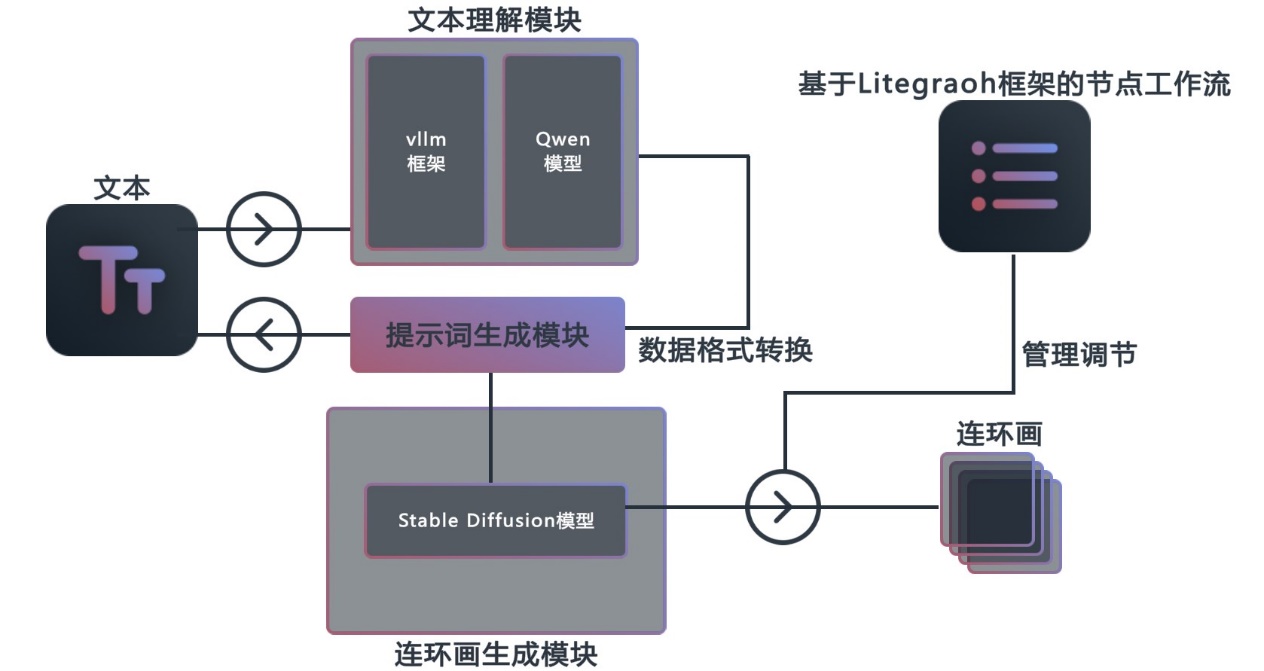


图 4 整体技术路线框架图

技术方案的具体层次结构和逻辑流向如下：

首先，平台的文本理解模块依托vllm框架与Qwen1.5-14B-Chat-GPTQ-Int4模型。vllm框架通过模型剪枝、优化CUDA内核以及连续批处理输入请求等手段，显著提升大型语言模型（Qwen模型）的推理效率。使用的Qwen1.5-14B-Chat-GPTQ-Int4模型基于Transformer架构，运用量化技术有效压缩模型规模、加速推理速度，同时凭借NTK感知插值、动态NTK感知插值、LogN-Scaling以及分层窗口Self-Attention等先进技术，确保在处理长文本时保持高稳定性和准确性。该模块的核心职责是对用户提交的文本进行深度剖析，提炼关键信息，并将其转化为适于连环画制作的结构化数据。

接着，提示词生成模块借助本地内置系统prompt与Qwen模型的定向引导功能。该模块将用户文本与预设的本地内置系统prompt相结合，通过分段处理、要素细化及权重排序等精细化步骤，生成既忠实于用户文本细节与剧情特征，又能精准引导后续图像生成的提示词集合。

随后，连环画生成模块运用Stable Diffusion模型作为核心技术。Stable Diffusion模型基于扩散模型原理，利用文本嵌入对图像生成过程进行精细化引导，确保生成图像在细节、氛围、艺术风格等方面与提示词的语义高度匹配。该模块负责将文本理解模块与提示词生成模块产生的输入转化为实际的连环画图像序列。生成过程涉及文本嵌入、Text Conditioned Latent U-Net去噪处理、潜向量生成及映射回像素空间等多个步骤，确保图像内容与提示词紧密相关且风格统一。

最后，工作流管理模块采用LiteGraph节点流程框架，为用户呈现可视化、可定制、可追溯的连环画生成工作流（ComfyUI）。该模块将连环画生成过程中复杂的Stable Diffusion模型操作流程转化为直观易懂的节点结构，用户可通过图形化界面轻松设计、配置工作流，实现对连环画创作过程的全方位精细化控制。工作流管理模块还确保在多次调用Stable Diffusion模型时，参数设置保持一致，确保连环画整体风格的连贯性。

智绘连环平台通过集成文本理解、提示词生成、连环画生成及工作流管理四大模块，构建了一条从用户文本输入到连环画图像输出的完整技术路径。各模块间紧密协作，充分发挥各自技术优势，共同实现了连环画创作的高效性、个性化与风格一致性。用户通过平台直观的工作流界面，可以轻松设计、控制连环画生成流程，实现从文本到视觉叙事的无缝转化，轻松打造出风格一致、叙事连贯的高质量连环画作品。

## 文本理解模块

智绘连环平台在生成连环画任务中，其主要的输入数据为用户提供的文本，文本理解模块主要职责是对用户输入的文本进行深度解析和理解，准确捕获并提炼出文本的关键信息，为后续的图像生成过程提供详实的语义指导。为此文本理解模块利用集成的LLM（大语言模型）的深度学习能力，能够深入理解输入文本的内在含义，把握文本的主旨、情感色彩、隐含意义等多层次语义信息。

大语言模型 (LLM) 是一类基础模型，经过大量数据训练，使其能够像人类一样理解和生成文本以及其他形式的内容。这种模型有能力从环境中推断，生成连贯且与环境相关的响应，翻译、总结文本，执行各种任务。

智绘连环平台的文本理解模块具体采用了vllm + Qwen的技术方案，vllm 框架和 Qwen 模型在连环画生成平台的文本识别模块中分别扮演推理加速和核心处理的角色，前者优化模型推理效率，后者提供强大的语言理解和生成能力。具体的技术架构图如下所示：

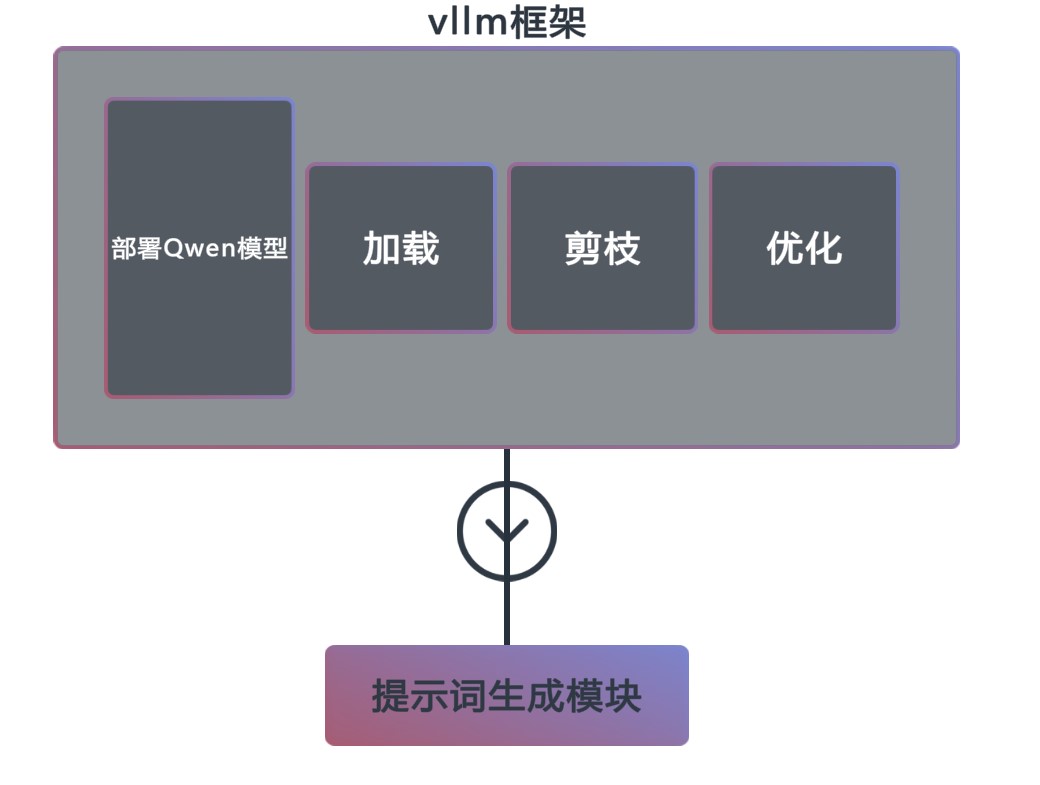


图 5  文本理解模块技术架构图

### 基于vllm框架的LLM模型推理优化策略

其中vllm 是一个基于剪枝技术的大模型推理加速工具，通过去除模型中的冗余参数，可以在保证模型性能的同时显著减少推理时间。vllm由于具备易于使用、具有最先进的服务吞吐量、高效的注意力键值内存管理（通过PagedAttention实现）、连续批处理输入请求、优化的CUDA内核等功能，成为本平台优化模型推理效率的选择。

首先vllm具备高度专业化的加载预训练模型的能力，且Qwen模型都被vllm所支持，通过其提供的专门的API接口，可将预先训练好的大型语言模型加载到vllm框架中，简化了加载Qwen模型的过程，且确保了模型数据的完整性和准确性。

其次，vllm框架的模型剪枝功能具有高度的专业性和灵活性。剪枝是vllm的核心功能之一，它通过智能地删除模型中的冗余参数，降低模型复杂度，从而实现推理加速。在剪枝过程中，vllm提供了丰富的参数设置选项，根据文本理解模块具体需求和模型特点，可以通过调整这些参数来找到最佳的剪枝策略。

由于经过剪枝优化后的模型在计算过程中所需的内存和计算资源减少，因此推理速度得到了显著提升。同时，vllm还采用了先进的优化算法和技术，如PagedAttention、连续批处理和CUDA核心优化等，进一步提高了Qwen模型推理的效率和准确性。这些专业化的技术使得vllm在处理Qwen模型时能够保持高性能的同时，实现快速推理。

最后，vllm还支持分布式推理，这使得它能够在多台GPU上并行运行模型，进一步提高推理速度。这种分布式推理的支持使得vllm能够应对更大规模的模型和更复杂的推理任务，可以满足智绘连环未来用户不断增长的需求。

### 基于Qwen1.5-14B-Chat-GPTQ-Int4模型的文本理解

文本理解模块具体应用了Qwen1.5-14B-Chat-GPTQ-Int4模型（文中统称Qwen模型）来实现文本理解、语义分析等核心功能，该模型属于Qwen模型系列中的Qwen-Chat（聊天模型，采用人类对齐技术进行微调）模型，Qwen基于数万亿词条的海量数据集上对语言模型进行预训练，然后，使用 SFT 和 RLHF 使 Qwen 进行符合人类偏好对齐，从而得到Qwen-Chat。Qwen 基座模型在众多下游任务中始终表现出卓越的性能，而聊天模型Qwen-Chat拥有先进的工具使用和规划能力，可用于创建agent应用程序，具有很强的竞争力

Qwen 模型采用了改进版的 Transformer架构。具体采用了开源的大型语言模型LLaMA的训练方法。

智绘连环平台所使用的Qwen1.5-14B-Chat-GPTQ-Int4模型是一个量化模型，模型量化技术通过减少模型参数表示的位数来降低模型的计算和存储开销。采用量化模型带来了诸多优势：首先，它显著减小了模型大小，使每个参数占用的存储空间大幅减少，这在移动设备、嵌入式系统或边缘设备上部署模型时尤为关键，因为这些设备的存储资源往往有限。其次，量化技术能加速推理过程，因为使用低位宽的整数或定点数进行计算相较于浮点数更高效，这不仅减少了内存带宽需求，还提高了硬件的并行计算能力，对于实时推理和响应时间敏感的应用至关重要。此外，量化还能降低模型在部署环境中的功耗，计算和存储操作的减少直接减少了能耗，拥有极大的可用性和实用性。

同时，Qwen模型在inference阶段利用了多种技术来实现上下文长度的扩展。首先采用了NTK感知插值（NTK-aware interpolation）技术，这种无需训练的技术可以调整比例参数，有效防止在扩展长度时丢失高频信息。其次使用了动态NTK感知插值（dynamic NTK-aware interpolation），这是NTK感知插值的改进版本，能够以块为单位动态改变比例参数，避免了性能的大幅下降。此外Qwen模型通过LogN-Scaling技术确保了模型在处理长序列时的稳定性和准确性。最后，分层窗口Self-Attention技术也是Qwen模型在处理长上下文时的一个重要策略，通过将注意力限制在一个上下文窗口内，并在不同层采用不同的窗口大小（较低的层使用较短的窗口，而较高的层使用较长的窗口），Qwen模型能够更有效地利用注意力机制，防止模型关注到太远的内容。

智绘连环平台的文本理解模块利用Qwen 模型强大的通用性来理解用户提供的文本指令、故事大纲、对话内容，提取关键信息并在后续转化为连环画制作所需的结构化数据。

## 提示词生成模块

智绘连环平台在处理用户提交的文本以生成连环画的过程中，提示词生成模块采用了独特的本地调试策略来进行提示词生成，确保为Stable Diffusion模型提供最适宜的输入。这一策略的核心是结合用户文本与预设的本地系统prompt，通过定向引导Qwen模型，生成符合连环画生成要求的提示词集合。具体技术架构图如下所示。



图 6 提示词生成模块技术架构图

在人工智能（AI）领域中，"prompt" 是指向模型提供输入以引导其生成特定输出的文本或指令。它是与模型进行交互时用户提供的文本段落，用于描述用户想要从模型获取的信息、回答、文本等内容。prompt 的目的是引导模型产生所需的回应，以便更好地控制生成的输出。

提示词生成模块中将用户提供的文本作为初始prompt，再根据调试好的内置系统prompt，将初始prompt分成几个部分，并与系统prompt结合，再传递给文本理解模块，定向引导文本理解模块中的Qwen模型去理解并返回固定格式的文本集合，最后将数据格式转化并传递给连环画生成模块。

### 本地系统prompt设计

本地内置系统prompt包含以下关键字：

* 指示（Instructions）：

指示是对任务的明确描述，相当于给模型下达了一个命令或请求。它告诉模型应该做什么，是任务执行的基础。

* 例子（Examples）：

例子是给出的一或多个具体示例，用于演示任务的执行方式或所需输出的格式。这种方法在机器学习中被称为示范学习，已被证明对提高输出正确性有帮助。

* 输入（Input）：

输入是任务的具体数据或信息，它是模型需要处理的内容。在Prompt中，输入会被清晰地标识出来，以便模型能够准确地识别和处理。

* 输出（Output）：

输出是模型根据输入和指示生成的结果。在Prompt中，会描述输出的格式，以便后续模块能够自动解析模型的输出结果。

### 提示词生成流程

1. 用户文本与prompt1结合：

用户提交的原始文本首先与本地系统prompt中的prompt1进行合并，形成一个包含了任务指示、示例、输入定义的综合prompt。合并后的prompt被送入文本理解模块，引导Qwen模型对用户文本进行深度解析，并按照指定格式输出四个文本段落，这4个文本段落有着上下文联系，保留着用户提交的文本的全部细节和剧情特征

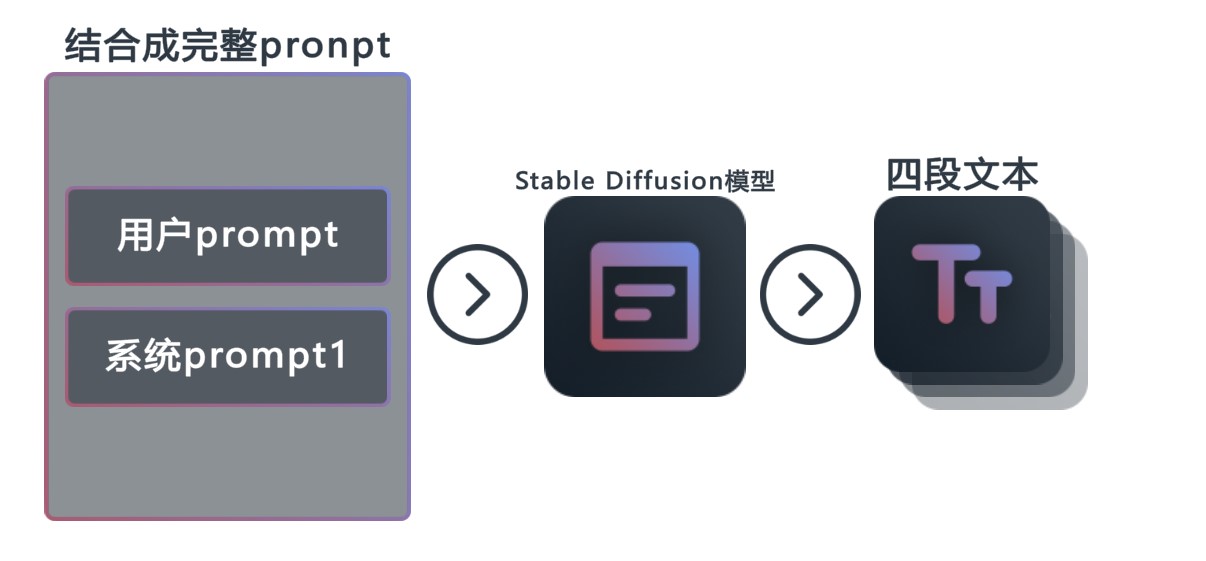


图 7 提示词生成流程1

1. 文本段落与prompt2结合：

根据连环画的4要素（人、事、景、物），每个由Qwen模型生成的文本段落进一步与本地系统prompt中的prompt2结合，形成针对4要素的细化prompt。将4个细化prompt再次被递交给文本理解模块，使用Qwen模型对每个文本段落中的要素进行理解和提炼，返回指定格式的集合。



图 8 提示词生成流程2

1. 数据格式化：

对于每个文本段落的每个要素，收集其对应的关键字列表，按照连环画生成模块中Stable Diffusion模型的标签语法要求，依据关键字的重要性或相关度进行重新排序，形成权重分明的提示词子集。

将四个文本段落的提示词子集整合成一个完整的提示词集合，该集合结构清晰、权重有序，能够有效地引导Stable Diffusion模型生成连环画各帧图像。



图 9 提示词生成流程3

## 连环画生成模块

连环画生成模块是智绘连环平台的核心组件之一，它专门负责将由文本理解模块和提示词生成模块提供的输入数据转化为实际的连环画图像序列。

这一模块是实现从文本叙述到视觉叙事转化的关键环节，具体应用了Stable Diffusion模型进行图像生成，该模型直接利用文本嵌入指导图像生成过程，使得其可以通过精细调整提示词来精确控制生成图像的细节、氛围、艺术风格乃至特定视觉元素，实现对连环画内容的高度定制化。且Stable Diffusion模型通过优化扩散过程，保证了图像生成的稳定性和质量，再结合智绘连环平台的文本理解模块和提示词生成模块，即使面对复杂或抽象的文本描述，也能生成视觉效果出众、艺术表现力强的连环画图像。

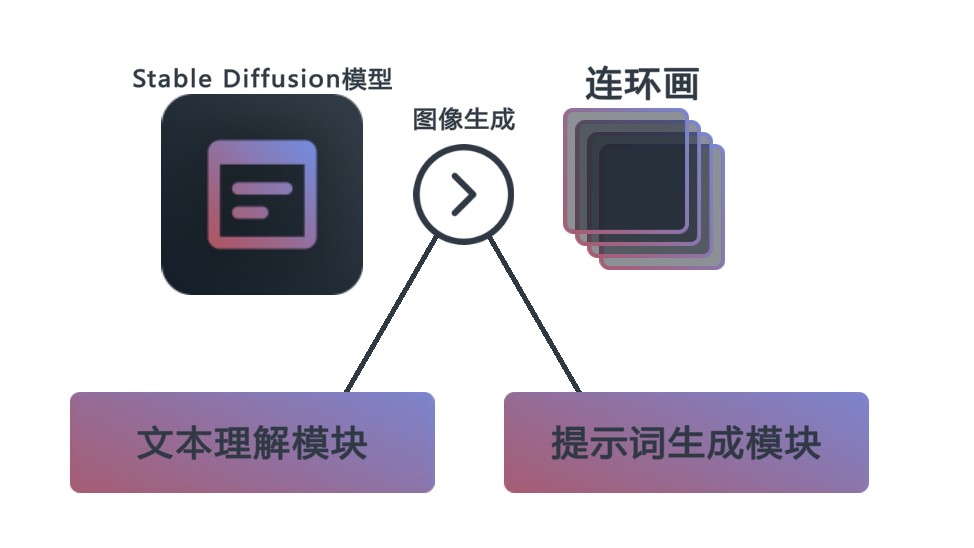


图 ？ 连环画生成模块技术架构图

### 基于Stable Diffusion模型的图像生成

Stable Diffusion模型是一个基于扩散模型的创新迭代版本，扩散模型是一种生成式模型，其灵感来源于物理扩散过程。在扩散模型中，图像生成被视作一个逆向过程，即从完全随机的高斯噪声开始，通过一系列递归步骤逐步还原出清晰的图像。智绘连环平台的连环画生成模块充分利用Stable Diffusion模型强大的“文生图”能力，即根据输入的提示词精确生成与之对应的连环画图像。

对接从文本理解模块和提示词生成模块得到的提示词集合，将每段提示词输入到连环画生成模块中，Stable Diffusion模型对于输入的文字先通过一个预训练的CLIP模型转换为文本嵌入（text embedding），CLIP模型通过联合训练文本编码器和图像编码器，使其在共享的语义空间中对齐，从而能够将任意文本映射到一个与之语义相关的向量表示。这是一种将自然语言文本编码为高维向量的形式，能够捕获文本的语义特征和上下文关联。

得到的文本嵌入与初始图像（通常为随机高斯噪声）一同输入到去噪模块，即Text Conditioned Latent U-Net。U-Net是一种深度卷积神经网络架构，以其独特的编码器-解码器结构和跳层连接著称，其特点在于通过跳层连接（skip connections）将编码器各层的特征直接传递给解码器相应层，从而保留了丰富的空间细节信息。在Stable Diffusion中，模型首先将高斯噪声映射到潜空间，然后在潜空间中执行扩散和去噪步骤，U-Net用于处理潜空间（latent space）中的图像表示，而非直接处理像素级图像。

在去噪过程中，Text Conditioned Latent U-Net不仅接收当前噪声状态的潜向量，还接受来自CLIP模型的文本嵌入，网络通过联合考虑这两个输入，学习如何在去噪的同时，根据文本提示塑造图像内容，以确保生成的图像与输入提示词语义高度契合。每一去噪步骤都会生成一个新的潜向量，当所有步骤完成后，最后一个潜向量被映射回像素空间，生成最终的一帧连环画图像。

图 ？ stable diffusion模型图像生成原理图

## 工作流管理模块

工作流管理模块在智绘连环平台中用于给用户提供友好的工作流图形化界面（ComfyUI），旨在为用户提供对连环画生成过程的精细化控制和管理。该模块的主要将复杂的Stable Diffusion模型操作流程转化为直观的节点结构。每一个节点代表连环画生成过程中的一个步骤，如风格设定、氛围渲染、角色动态等，并可以设置对应的各种参数用于调整生成的最终图像。

该模块通过提供可视化、可定制、可追溯的连环画生成工作流，极大地提升了用户在连环画创作中的操作便利性、创作自由度和工作效率，使得从文本到连环画的转化过程既高效又个性化，令智绘连环平台可以适用于教育、科普、广告、个人创作等多种应用场景。

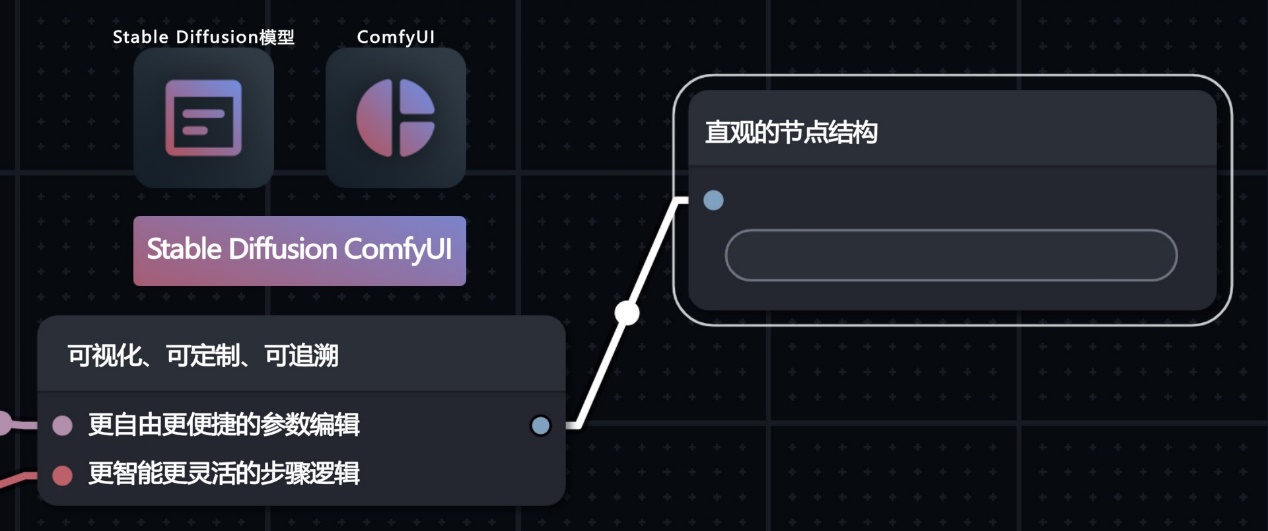


图 ？ 工作流管理模块技术架构图

### 基于Litegraph节点流程框架实现Stable Diffusion ComfyUI

工作流管理模块中使用连环画生成模块的工作流界面（ComfyUI）是基于LiteGraph框架开发的，LiteGraph是一个轻量级、灵活的图形化编程环境，特别适用于构建数据流图和节点式工作流程，可以实现类似虚幻引擎的蓝图编程，包括一个编辑器来构建和测试节点图，支持浏览器和Node.js，可以轻松集成到任何现有的 Web 应用程序中，并且无需编辑器即可运行节点图。

Litegraph框架具备在Canvas2D上渲染的灵活性，优化处理数百个节点的高效性，以及可定制的丰富主题和节点特性。它拥有直观的编辑器，支持实时模式系统，并能在NodeJS中执行。此外，Litegraph.js易于集成到任何JS应用程序中，并支持TypeScript，为智绘连环平台提供了个性化界面创建与图结构管理的强大工具。具体实现流程如下：

1. **节点类型设计：**基于Litegraph框架，根据智绘连环平台的功能模块，设计一系列与之对应的节点类型。例如，创建“文本输入”节点用于接收用户文本，开发“提示词生成”节点对接LLM模型，实现“图像生成”节点与Stable Diffusion模型的接口，以及各种用于风格设定、氛围渲染、角色动态等特定效果的自定义节点。
2. **节点属性与配置：**每个节点定义必要的属性和配置选项，允许用户在ComfyUI界面上直接调整节点参数，如设定生成图像的分辨率、风格权重、角色动作等，这些配置信息能被节点内部逻辑正确解析并应用于连环画生成过程中。
3. **节点间通信机制：**利用LiteGraph的数据线功能，实现节点间的数据传递和交互。定义清晰的数据输入输出接口，确保节点间的连接能够正确反映数据流的流向，如文本数据从“文本输入”节点流向“提示词生成”节点，提示词进一步传递给“图像生成”节点等。
4. **工作流执行与控制：**按照用户定义的节点连接顺序和数据流逻辑执行连环画生成任务。实现暂停、继续、重置工作流，以及单步执行、跳转至指定节点等功能，提供对工作流执行过程的精细控制。
5. **用户界面集成：**将LiteGraph图形界面集成到ComfyUI中，确保界面布局合理、操作流畅。提供必要的辅助工具，如节点搜索、节点库、预设工作流模板、节点属性编辑面板等，使用户的使用体验提升，易于上手。
6. **保存与加载工作流：**实现工作流的序列化与反序列化功能，允许用户保存当前的工作流配置到文件，以及从文件加载已保存的工作流。确保工作流文件格式兼容性强，便于跨平台或跨设备使用。

用户在使用智绘连环平台的连环画生成功能时，可以轻松地通过连接和配置节点来设计连环画生成流程，每个节点代表一个特定的操作或功能，如文本理解、提示词生成、图像渲染等。用户通过连接节点来定义工作流程，直观地看到整个连环画创作的逻辑链条，享受到便捷高效的连环画创作体验。

### 统一风格的自动化批量图像生成

在工作流管理模块中，基于LiteGraph节点流程框架将原本复杂的连环画生成过程转化为一目了然的可视化工作流，并针对Stable Diffusion模型的应用进行了深度集成与优化。使智绘连环平台赋予了用户前所未有的创作自由度与作品一致性控制能力。

首先，连环画生成模块中的Stable Diffusion模型调用、参数设置以及生成进程的控制等关键环节，皆被精心封装为易于操作的节点。用户只需在简洁直观的节点界面中，通过简单的拖曳与连接操作，即可构建出符合自己创作意图的连环画生成流程。无论是模型的选择、艺术风格的设定，还是生成过程的启停，无需深入理解底层技术细节，大大降低了创作的技术门槛。

同时工作流管理模块深度融合了文本理解与提示词生成模块，将该两个模块处理用户输入的文本后得到的提示词集合作为Stable Diffusion模型的输入，引导其依次生成连环画的各帧画面。底层系统会自动多次调用模型，确保一次工作流即能完成整套连环画（默认为4张图片）的绘制。这一设计打破了Stable Diffusion模型单次调用生成单张图片的局限，使得Stable Diffusion模型能够在连环画创作场景下发挥出巨大潜力。

更为重要的是，工作流管理模块确保了在多次调用Stable Diffusion模型的过程中，用户预先设定的各项参数始终保持不变。这意味着无论连环画中的哪一帧画面，模型都将严格遵循用户指定的图像尺寸、色彩风格、细节偏好等要求。这种一致性保障机制确保了连环画每一帧都如同出自同一幅画卷，整体风格高度统一，视觉叙事流畅自然，显著提升了作品的质量。

智绘连环平台借助工作流管理模块，不仅实现了Stable Diffusion模型在连环画创作中的自动化批量图像生成，还确保了生成过程中风格的完美统一。这一设计极大地提升了创作效率，赋予用户更高的创作自由度，同时确保了作品的品质与专业水准。无论是初次尝试连环画创作的新手，还是寻求高效创作工具的专业工作者，都能在智绘连环平台上找到理想的创作环境，轻松打造出风格一致、叙事连贯的高质量连环画作品。

# 系统实现

## 软件设计实现

### 前端技术

前端开发主要采用HTML、CSS、JavaScript（简称JS），构建了用户友好、交互流畅的图形化操作界面。

HTML作为超文本标记语言，用于定义网页的结构和内容，为智绘连环平台构建了清晰的页面布局，确保用户能够直观地访问和操作各项功能。

CSS层叠样式表用于定义网页的样式和外观，赋予了智绘连环平台统一且美观的视觉风格，同时也实现了界面元素的动态响应式设计，确保在不同设备和屏幕尺寸上的良好显示效果。

JavaScript作为客户端脚本语言，负责实现网页的动态行为和用户交互逻辑。在智绘连环平台中，JS主要用于构建基于LiteGraph节点流程框架的连环画生成工作流界面（ComfyUI）。通过JS，智绘连环平台实现了节点的拖放操作、连线逻辑、属性编辑、工作流控制等功能，使得用户能够直观地搭建和调整连环画生成流程。

特别地，智绘连环平台借助LiteGraph框架（一个轻量级、灵活的图形化编程环境），在Canvas2D上高效渲染节点图，支持大规模节点的处理，并提供了丰富的节点主题和特性定制，为用户提供了高度可定制、易于理解的工作流编辑器。此外，通过实时模式系统和良好的浏览器兼容性，确保了用户在Web环境下获得顺畅的图形化编程体验。

具体实现的界面如下：

1. **输入文本生成图像界面：**利用JS实现了文本输入框、生成按钮等组件的交互逻辑。用户输入文本后，可通过JS触发后端请求，将文本数据发送至服务器进行处理，文本数据作为输入进入默认的一套工作流开始生成连环画，待服务器返回生成的图像数据时，再利用JS更新界面，实时展示生成的连环画图像。
2. **展示优质连环画界面：**借助JS实现连环画作品的加载、滑动浏览、查看等交互功能。JS还负责加载服务器保存的优质连环画数据，将其适配到相应的展示模板中，确保作品列表的完整展示。
3. **定制工作流界面：**借助LiteGraph框架，用户可以自定义工作流节点属性、设置节点间依赖关系、添加或删除节点等，定制具有个性的一套来连环发生成工作流。JS实时响应用户的操作，连接相应的后端接口，进行数据传递和接收，动态更新界面展现，确保工作流编辑过程中的即时反馈和精确控制。
4. **用户登录界面：**利用JS实现实时表单验证、记住密码选项、登录按钮状态变化等功能。当用户填写完登录信息并点击登录按钮时，JS发送异步请求至服务器，访问mysql数据库进行身份验证，成功后通过JS更新界面状态，完成用户登录过程。

### 后端技术

后端开发主要使用Python语言，结合一系列关键库和工具，为智绘连环平台提供了强大的文本理解、提示词生成、图像生成以及工作流管理支持。具体包括：

1. **基于vllm框架部署Qwen模型本地服务：**借助于vllm库（一个基于剪枝技术的大模型推理加速工具），实现对Qwen模型的高效加载与推理。vllm通过模型剪枝技术去除冗余参数，显著降低推理时间，并提供高级API接口简化模型加载过程，确保模型数据的完整性和准确性。同时，vllm支持模型分布式推理，能够应对大规模模型和复杂推理任务，为平台的未来发展预留了充足扩展空间。
2. **调用部署在本地的Qwen模型输入输出接口：**Python后端直接与本地部署的Qwen模型进行交互，通过调用其输入接口接收用户提供的文本数据，经由模型处理后，通过输出接口获取深度解析的文本信息，供后续流程使用。
3. **对用户上传的文本数据进行处理：**Python后端负责接收、验证用户上传的文本数据，实现提示词生成模块中prompt集合与文本的结合、最终处理结果的数据格式转换，确保数据满足Qwen模型的输入要求。
4. **调用部署在本地的Stable Diffusion模型各个参数和功能的接口：实现**Python后端与本地部署的Stable Diffusion模型集成，通过调用其接口设置模型参数、触发图像生成任务，并接收生成的连环画图像数据。
5. **开发ComfyUI对应LLM的组件：**通过Python代码实现与前端ComfyUI界面的对接，开发出ComfyUI对应LLM的组件，为前端提供与LLM模型交互的接口，如文本提交、结果接收等，确保前端与后端逻辑的无缝衔接。
6. **调用将Stable Diffusion模型的各个接口与工作流节点对接：**后端通过Python实现对Stable Diffusion模型接口的封装与调度，确保其能够按工作流节点的要求进行操作，如根据节点配置生成特定风格或内容的图像，并将生成结果反馈给前端展示。
7. **使用MySQL数据库进行平台登录验证与数据管理：**智绘连环平台采用MySQL数据库进行用户身份验证、权限管理及各类业务数据存储。实现用户注册、登录、密码找回等身份验证功能，确保用户账户安全。同时，MySQL数据库用于存储用户上传的文本数据、生成的图像结果、工作流状态信息等业务数据，便于数据查询、统计分析及备份恢复，为平台的稳定运行和数据分析提供坚实基础。

### 接口封装与调用

为了保证各模块间通信的规范性、安全性与易维护性，智绘连环平台采用OpenAPI（开放API规范）对前后端接口进行统一定义与封装。OpenAPI规范为平台的接口设计提供了清晰的结构化描述，包括接口路径、请求方法、参数定义、响应格式等信息，使得开发者可以方便地查阅、理解和调用接口。通过遵循OpenAPI标准，智绘连环平台确保了接口的标准化、文档化和自动化测试，有效促进了前后端开发团队的协作效率，也简化了未来可能的第三方集成工作。

### 系统集成

智绘连环平台的架构设计充分考虑了系统的功能需求、性能要求与用户体验，选择了HTML/CSS/JS与Python作为前后端开发语言，依托LiteGraph、vllm等关键库与工具，构建了高效、易用且可扩展的技术栈。通过OpenAPI规范对各模块接口进行封装，确保了系统内部及对外部服务的高效、稳定通信。这样的架构设计为平台提供了坚实的技术基础，有力支撑了连环画生成过程中的文本理解、提示词生成、图像生成及工作流管理等核心功能的实现。

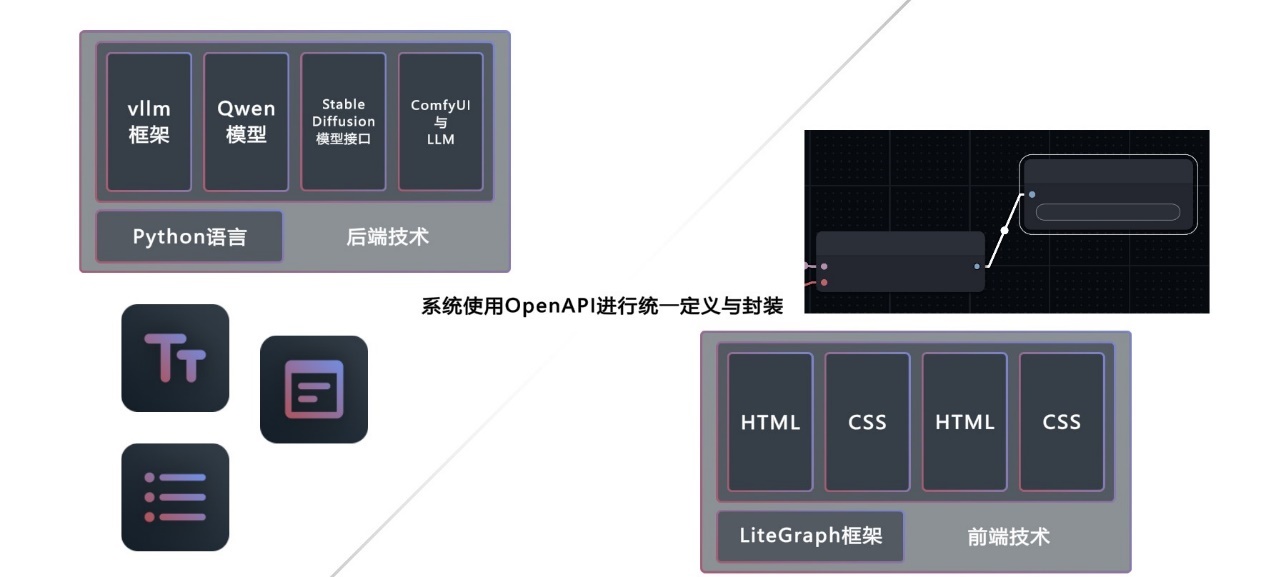


图 ？ 系统集成介绍图

作为一个集成化、智能化的连环画创作系统，智绘连环平台高效运作得益于各模块间的紧密协作与无缝集成。通过接口调用、事件驱动等机制，文本理解模块、提示词生成模块、连环画生成模块以及工作流管理模块相互配合，共同构建了一个从文本输入到连环画输出的完整创作流程。以下详述各模块间的系统集成方式：

1. 文本理解与提示词生成模块的协同：

文本理解模块通过对用户提供的文本进行深度解析，提取关键信息并转化为结构化数据。其输出结果作为提示词生成模块的输入，触发后续的提示词生成流程：

接口调用：文本理解模块通过API接口向提示词生成模块提供已解析的文本信息，包括文本摘要、关键情节、人物关系等结构化数据。

事件驱动：当文本理解模块完成对用户文本的处理并生成结果时，触发一个“文本解析完成”事件，通知提示词生成模块可以开始进行下一步操作。

1. 提示词生成与连环画生成模块的交互：

提示词生成模块基于文本理解模块提供的结构化数据，结合本地系统prompt，生成连环画制作所需的提示词集合。这一集合随后被传递给连环画生成模块，引导模型生成连环画图像：

接口调用：提示词生成模块通过API接口向连环画生成模块提供生成的提示词集合，每个提示词对应连环画的一帧图像。

事件驱动：当提示词生成模块完成提示词集合的生成时，触发“提示词生成完毕”事件，促使连环画生成模块启动图像生成任务。

1. 各核心模块与工作流管理模块的联动：

文本理解模块、提示词生成模块、连环画生成模块与工作流管理模块完整集成，对于一整个创作连环画的工作流程，用户可以通过工作流管理模块去动态调整、使用其它所有核心模块的功能。从文本理解模块接受文本，到提示词生成模块输出对应提示词集合，再到连环画生成模块将提示词转化为连环画图像，其工作流程均受工作流管理模块调度和控制。工作流管理模块根据用户在ComfyUI界面中搭建的节点流程，通过接口调用和事件机制与各模块交互：

接口调用：工作流管理模块通过调用各个模块的API，发送对应指令、传递用户设定的参数（如风格、分辨率等）以及接收各个返回值。

1. 全局事件与状态同步：

在整个系统中，还存在全局事件与状态同步机制，确保各模块对系统状态有统一的认知，并能及时响应变化：

状态查询接口：各模块可通过查询接口获取其他模块的当前状态，如文本理解模块的解析进度、连环画生成模块的生成进度等，以便进行协同决策。

广播事件：系统中发生重要状态变更时（如用户暂停或恢复工作流、更改生成参数等），通过事件广播机制通知所有相关模块，确保各模块及时做出相应的调整。

通过以上接口调用、事件驱动以及状态同步机制，智绘连环平台的各模块得以高效协同，形成了一个高度集成的连环画创作系统。这种集成方式不仅确保了系统内部数据流的顺畅流转，还为用户提供了一致且实时的创作环境，极大地提升了连环画创作的效率与体验。无论是文本解析、提示词生成、图像生成，还是工作流管理，各个环节紧密衔接、有机融合，共同构成了一个强大而易用的连环画创作平台。

## 用户界面

### 主要使用界面

智绘连环平台采用HTML、CSS、JavaScript构建用户友好、交互流畅的图形化操作界面，遵循一致性、简洁性、反馈性、容错性等用户体验原则，确保用户在Web环境下高效进行连环画创作。其主要使用界面的内容具体包括：

1. **输入文本生成图像板块：**界面包含文本输入框、生成按钮、简单参数设置（可选生成单图或4格连环画）、预览区域。用户输入文本并选择生成单图或4格连环画后，点击生成按钮，前端通过异步请求将文本发送到后端。后端通过文本处理和提示词生成，再传递给连环画生成模块，生成图像后返回给前端。前端接收数据后实时更新预览区域，展示生成的连环画。具体界面详情如下图所示。



图 10 输入文本生成图像界面

1. **展示优质连环画板块：**界面包括优质的连环画图像列表，主要用于展示一些优质连环画作品让用户观赏。用户每次进入界面，JS从保存在服务器本地的连环画自动加载到该区域，呈现在作品列表中供用户查看观赏。具体界面详情如下图所示。



图 11 展示优质连环画板块界面

1. **定制工作流界面：**在主要使用界面，用户如果想定制自己的连环画生成工作流进行精细化生成，可以通过点击固定的跳转组件，转入连环画创作的节点工作流界面。具体界面详情如下图所示。

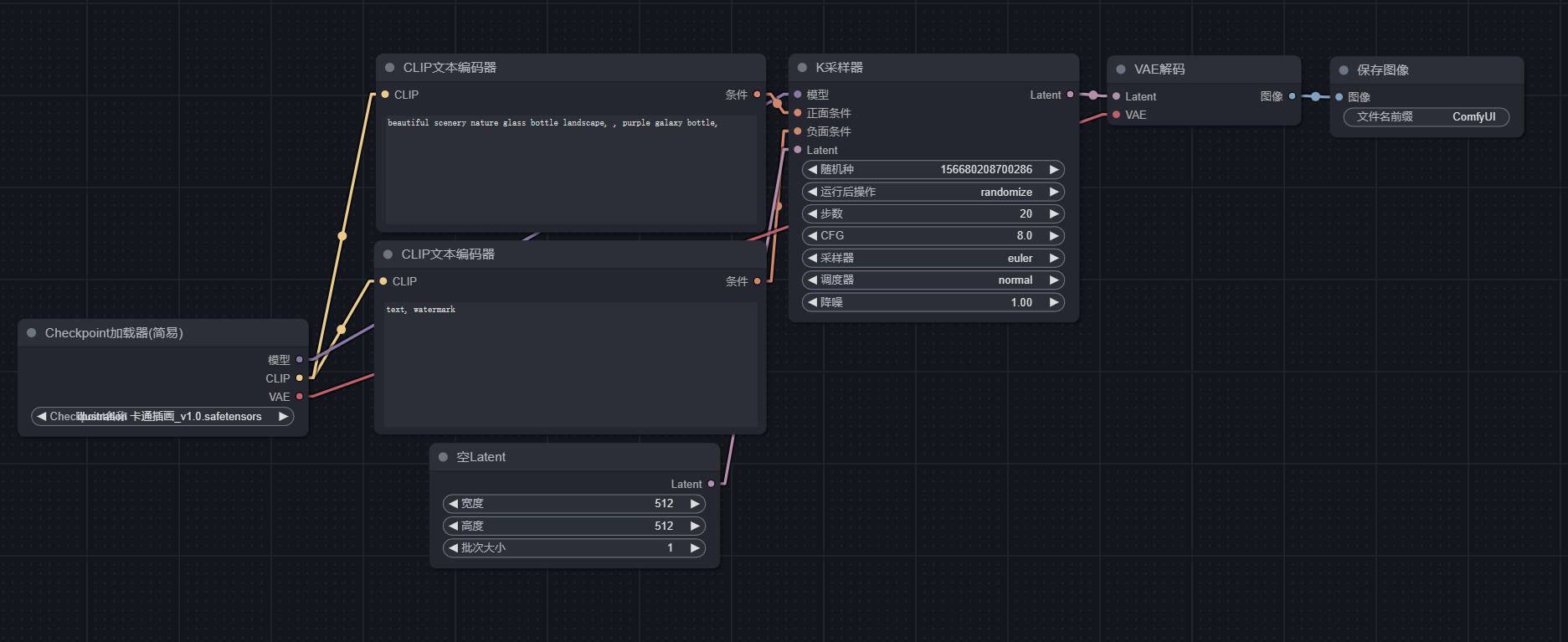


图 12  节点工作流界面

用户可以在节点工作流界面的工作区通过JS驱动的拖放操作添加、移动、连接节点。首先在节点库中选取所需节点类型，点击添加到工作区。然后在属性面板中设置节点的参数，调整节点间依赖关系。最后通过控制面板启动、暂停、保存、加载工作流进行连环画的创作。实时响应用户操作，通过接口与后端通信，更新节点状态与工作流执行情况。具体界面详情如下图所示。



图 13  工作流跳转界面

### 用户登录界面

智绘连环平台的登录界面是用户访问平台和使用平台功能的入口，具体组件包括接收用户输入的邮箱、密码的表单，以及注册和登录两个按钮。

主要功能是进行用户身份验证，用户通过输入预先注册的邮箱和密码，平台通过后端与MySQL数据库进行比对，确认用户提供的凭据与系统中存储的用户信息是否匹配。这一过程确保只有经过授权的用户才能访问个人账户、使用平台服务，防止未授权访问和数据泄露，保障用户数据的安全性和隐私权。

登录界面是实施权限管理的第一道关口。在未来智绘连环可能进行的商业化扩展，对于不同类型的用户（如普通用户、管理员、VIP用户等）可能享有不同的平台权限和功能访问范围。通过登录验证，平台可以根据用户角色分配相应的功能模块、资源配额、特权服务等，确保用户在各自的权限范围内操作，维持平台秩序与公平性。

## 数据来源

智绘连环平台的主要输入为一段具有故事性的文本，然后基于该输入生成连环画，为了系统性地测试智绘连环平台的整体功能和综合能力，项目团队自制数据集“中文短篇故事数据集”，该数据集专为满足智绘连环平台的输入需求而设计，确保测试过程中的数据质量、适用性及合法性，具体属性如下：

1. **数据格式：**所有收录的故事文本均统一采用纯文本（TXT）格式，这种标准化格式简洁明了，易于被智绘连环平台的各个功能模块直接读取和解析，有利于进行无缝对接和高效测试。
2. **数据来源：**数据集的全部内容由项目团队自主研发，确保了数据的针对性和可控性。此数据集不仅能够精准契合智绘连环平台的功能需求，还便于在测试过程中灵活调整数据内容以适应各种特定测试场景。
3. 数据获取方式：

原创故事：团队成员依据平台需求专门撰写故事，确保内容新颖独特，富有创意，且与平台目标用户群体的偏好相吻合。原创作品有助于考察智绘连环平台对多样文风、个性化叙事手法以及新颖情节设定的适应能力。

网络采集：严格遵循相关版权法律法规，从公开授权或无版权限制的网络资源中精心挑选并整理出符合测试要求的短篇故事。此举确保了数据使用的合法合规性，同时引入了更广泛的故事风格和主题，增强了数据集的多样性和普适性。

1. 数据特点：

篇幅较短：每篇故事都精炼紧凑，平均长度约为150字，既适合快速阅读理解，又有利于智绘连环平台高效处理和转化为连环画形式。

有故事性：所有文本均包含明确的情节主线、角色设定、冲突发展与解决等元素，构成完整的故事结构，以激发连环画创作的丰富性和连贯性。

中文文本：所有故事均以中文撰写，语言规范，表述清晰，贴合中国用户的阅读习惯和智绘连环平台的中文处理能力。

1. **数据规模：**该数据集共包含100篇短篇故事，数量充足，足以涵盖多种题材、风格和情节类型，为全面、深入测试智绘连环平台的功能提供了丰富的素材基础。

## 数据训练

为了全面、系统地评估智绘连环平台的性能与效果，特制定以下训练测试方案：

1. 训练目标
2. **功能完整性：**验证智绘连环平台能否准确识别并处理数据集中各类故事文本的关键要素（情节主线、角色设定、冲突发展与解决），将其成功转化为连环画形式。
3. **艺术表现力：**评估平台生成的连环画在视觉艺术、叙事逻辑、情感传达等方面的表现，是否能忠实再现原文本的故事氛围与情感基调。
4. **适应性与稳定性：**测试平台对不同题材、风格、篇幅的故事文本的适应能力，以及在大量连续测试中的稳定性。
5. **用户体验：**通过模拟用户视角，评价平台界面友好性、操作便捷性以及生成连环画的速度、质量对用户满意度的影响。
6. 训练方法与步骤
7. 单元测试

选取数据集中各类型故事样本，输入选定的故事文本至智绘连环平台，观察平台处理过程，记录生成连环画所需时间，然后审查生成的连环画，对照原文本，检查关键情节、角色、冲突及其解决是否准确呈现。

评估指标如下：

1. 文本理解准确率：统计平台正确文本内容所占的比例。
2. 生成时间：记录单个故事转化成连环画的平均耗时。
3. 艺术评分：依据预设的艺术评价标准（如画面构图、色彩搭配、动态表现等）对连环画进行主观评分。
4. 系统测试

随机抽取数据集中的部分或全部故事，批量输入故事文本至平台进行批量测试，监控平台在处理大量输入时的运行状态，记录所有可能出现的异常情况，最后对生成的连环画进行抽样审查，评估整体质量。

评估指标如下：

1. 系统稳定性：记录处理过程中出现错误、延迟或其他异常的次数与比例。
2. 连环画质量一致性：比较批量生成连环画的整体艺术表现力与单元测试结果的差异。
3. 数据分析
4. **数据整理：**汇总单元测试、系统测试与用户体验测试的各项量化指标及主观评价。
5. **性能分析：**对比测试结果与预期目标，分析智绘连环平台在功能完整性、艺术表现力、适应性与稳定性、用户体验方面的实际表现，识别优势与不足。
6. **结果解读：**结合具体测试案例，深入解读数据背后的原因。
7. **建议提出：**根据测试结果与分析，为智绘连环平台的优化升级提供针对性改进建议，包括技术改进、界面调整等。

通过以上训练测试，在开发过程中定期对智绘连环平台的综合性能进行全面、深入的评估，为其持续优化提供坚实的数据支持。

## 改进过程

### 选择vllm框架改进模型推理过程

vllm 是一个领先的 LLM 推理与服务库，它不仅高速，还非常易于使用。其具备业界领先的服务吞吐量、利用 PagedAttention 技术实现高效管理注意力机制中的键和值、连续批处理接入的请求、通过充分优化的 CUDA 核心确保速度与效率、针对分布式推理提供张量并行支持等多种优势，理论上使用vllm框架去部署和加载需要的LLM大语言模型会得到更好的反馈和使用过程。具体的使用vllm与不使用vllm进行llm模型推理测试过程如下：

1. **实验设计：**选取一定数量的代表性文本输入样本（为了更好反映LLM大语言模型推理时间，均使用长度1500字的文本内容），分别使用原模型（Qwen模型）与vllm框架改进后的模型（vllm + Qwen模型）进行推理，对比模型推理回答的响应时间。
2. **数据收集：**记录每个样本的推理开始时间、结束时间。同时，通过模拟用户行为，记录页面加载时间、交互响应时间等用户体验指标。
3. **数据处理：**计算平均推理时间、平均页面加载时间、平均交互响应时间等关键指标，用于后续分析
4. 测试结果：

采用VLLM框架后，平均推理时间从原模型的3.2秒降至改进后的1.1秒，缩短了66%。这意味着用户等待连环画生成的时间显著减少，提高了创作效率。

平均加载时间从15秒降低到12秒，降幅约12.5%。更快的加载速度有助于提升用户首次访问时的满意度。

平均交互响应时间从500毫秒减少到419毫秒，提高了约15%。这可以使连环画创作整体操作更加流畅，增强了创作过程的沉浸感。

基于改进比较测试，可以断定引入vllm框架对智绘连环平台的文本理解模块中Qwen模型进行文本推理的过程能够产生显著积极影响，可以有效实现在实际体验中智绘连环平台综合性能、质量和用户体验的多维度提升，为用户带来了更为卓越的连环画创作体验。

### Qwen系列模型的选择

对于智绘连环平台的文本理解模块中核心的LLM大语言模型的选择，关乎着整个平台最终的水平高低。在众多的LLM大语言模型中，智绘连环平台最终选择了使用Qwen系列模型中的Qwen1.5-14B-Chat-GPTQ-Int4模型。

选择Qwen系列模型，是因为其拥有卓越的性能表现、灵活的适用性、高效的计算资源利用以及持续的技术创新，是自然语言处理方面的优秀选择。在文本分类、情感分析还是问答系统等方面，该系列模型都能准确高效地处理复杂的语言现象，满足不同领域和场景的需求。同时，通过采用先进的模型压缩和量化技术，Qwen系列模型在保证性能的同时，降低了部署和运行的门槛，使得大规模数据处理和实时应用变得更为便捷。此外，Qwen系列模型背后的专业团队持续进行技术创新，不断优化和升级模型性能与功能，确保该系列模型始终走在行业前列。

其中Qwen1.5-14B-Chat-GPTQ-Int4模型是Qwen系列中的一款重要模型，拥有140亿参数量，其大规模的参数容量赋予了模型强大的语言建模能力和深度语义理解能力。与其他模型相比，它在多个方面展现了显著的优势。在模拟基准测试中，该模型在SQuAD 2.0阅读理解任务、GLUE综合自然语言理解评测均高于Qwen系列其它模型。

同时Qwen1.5-14B-Chat-GPTQ-Int4模型采用GPTQ量化技术，实现了模型的高效压缩和加速推理。在相同硬件条件、同等资源限制下，该模型能够处理更大批量的任务或更复杂的文本输入，且对硬件设备的要求有所降低。尚未采用类似GPTQ的量化技术或采用不同的量化方案的其它模型，其资源效率不及Qwen1.5-14B-Chat-GPTQ-Int4模型。所以选择应用Qwen1.5-14B-Chat-GPTQ-Int4模型在长文本的理解和实时应用中具有显著优势。

还有Qwen1.5-14B-Chat-GPTQ-Int4模型特别标注为“Chat”版本，表明该模型在对话交互场景下进行了针对性优化。模拟对话评估显示，该模型在多轮对话连续性理解任务上的平均准确率比非“Chat”版本更高，能有效捕捉对话历史信息，生成连贯、相关性强的回复。这对于智绘连环平台的工作流管理模块进行批量操作时运行结果更准确。

### 实现节点流界面的框架选择

Litegraph框架强大的功能集和易用性使其成为构建类似虚幻引擎蓝图编程的理想之选，其具体具备了以下优势：

1. **跨平台兼容与无缝集成：** Litegraph支持浏览器环境和Node.js环境，能够灵活适应多种应用场景。其轻量级的设计（仅需单个文件，无外部依赖）使得集成至任何现有的Web应用程序中变得极其简便，无论是前端用户界面还是后端服务，均能实现平滑对接。
2. **高性能可视化与交互：**Litegraph在Canvas2D上高效渲染节点图，支持放大、缩小、平移操作，便于用户清晰观察和精细调整复杂的流程图结构。其还具备WebGLTexture内部使用的能力，确保即使在大规模节点图中也能保持流畅的视觉效果。此外，Litegraph提供搜索框、键盘快捷键、多选、上下文菜单等丰富的编辑器功能，极大地提升了用户在设计和调试节点图时的工作效率。
3. **大规模节点图支持与优化：**专为处理包含数百个节点的大型节点图而优化，无论是编辑阶段还是运行时，都能保持出色的性能表现。这种高并发处理能力确保智绘连环平台在面对复杂工作流或大规模创作项目时仍能保持稳定、快速的响应。
4. **高度可定制化：**Litegraph允许用户自定义主题，调整颜色、形状、背景等视觉元素，以契合智绘连环平台的整体风格和品牌一致性。每个节点的动作、绘图和事件均可设置回调函数，实现个性化定制。节点本身也支持高度定制，包括颜色、形状、插槽布局（垂直或水平）、内置小部件以及自定义渲染逻辑，从而满足平台多样化的创作需求和用户偏好。

选择Litegraph作为智绘连环平台节点流界面的实现框架，不仅因其具备优秀的可视化编辑、高性能处理、深度定制化和实时交互能力，而且其跨平台兼容性、无缝集成特性，均与平台的技术栈和未来发展需求高度契合，从而为用户提供高效、灵活、直观的节点流界面设计与操作体验。

## 系统部署方法

1. 系统需求
2. 硬件需求：

服务器硬件：CPU推荐英特尔 Xeon W系列或AMD Ryzen Threadripper PRO系列、内存64G、GPU推荐RTX4090。

1. 网络环境：需要网络。
2. 环境准备
3. Web服务器与相关软件安装：需要安装Visual Studio Code和MySQL。
4. 编程语言环境与依赖：使用的编程语言为Python，依赖库包括Transformer、pytorch、vllm等。
5. 智绘连环平台网站部署

部署步骤：

1. 对前端和后端的相关代码进行编译
2. 将编译后的文件上传至服务器，并在指定目录下解压缩。
3. 部署Qwen1.5-14B-Chat-GPTQ-Int4模型和Stable Diffusion模型到本地。
4. 服务启动与验证
5. 网站启动与访问测试：登录服务器，打开终端，导航至后端服务代码所在目录运行启动脚本。打开浏览器，输入平台的公网域名或IP地址以及相应的端口，观察页面是否正常加载。
6. 功能完整性验证：注册或登录平台账号，创建一个新的连环画项目，在文本输入框内键入一段描述故事场景的文字，如“在一个阳光明媚的早晨，小狐狸在森林里遇到了正在采蘑菇的小兔子。”确认文本保存成功，无丢失或乱码现象，选择已输入文本的场景，触发连环画生成过程，观察后台任务状态，确认任务开始并正在执行，待任务完成后检查生成的连环画图像是否符合文本描述，画面清晰、内容连贯，无明显逻辑错误或视觉瑕疵。

# 测试分析

### 文本理解模块测试分析

* 测试数据集

本次测试采用了中文机器阅读理解数据集——Chinese Squad。该数据集包含大量的篇章（context）及其对应的问题（question）与答案（answer），旨在衡量模型在复杂语境下理解和抽取关键信息的能力。选择此数据集作为测试基准，可以确保文本理解模块在面对多样化、结构化文本时的表现具有较高的信度和效度。

* 测试过程

1. 数据预处理：从Chinese Squad数据集中提取出每个样本的“context”字段，作为待处理的原始文本。
2. 文本处理：将提取出的每个文本与具有“将文本分成4部分故事情节”的prompt相结合，作为输入传入文本理解模块。此prompt指导Qwen模型将长篇文本切分为四个逻辑连贯、内容独立的段落，模拟连环画的四格叙事结构。
3. 输出评估：对Qwen模型理解返回的每个文本的四个分段进行评估，主要关注以下几个方面：

处理时长，记录模型处理每个文本所需的时间，以衡量其处理效率。

分段效果，评价各段落是否主题明确、内容完整，且彼此间逻辑衔接紧密，无明显信息断裂或冗余。

上下文联系，分析各段落是否在保持相对独立的同时，共同构成了对原文本核心内容的整体呈现，无重要信息遗漏或矛盾。

* 测试结果

每个文本的处理时长：统计测试结果数据，Qwen模型处理每个文本的平均时长为2秒，表现出高效的文本理解与分段能力，能够在短时间内完成大量文本的处理任务，满足实时应用的需求。

分段效果：经评估，Qwen模型对文本的分段效果表现优异。各段落主题鲜明，内容精炼且自成体系，既体现了原文本的关键信息点，又避免了无关细节的干扰。段落间的过渡自然，形成了连贯的故事情节链，有效适应了连环画的叙事要求。

上下文联系：在保持各段落独立性的前提下，Qwen模型成功地保持了各段落与原文本整体内容的强关联。各段落共同构建了对原文本核心主题的立体展现，无明显的信息遗漏或逻辑冲突，确保了生成的连环画能够忠实反映原文本的主旨和情感。

测试结果样例：

* 分析与结论

本理解模块中的Qwen模型在中文机器阅读理解数据集Chinese Squad上的测试表现优异，证明其具备高效、精准地将长篇文本转化为适合连环画呈现的文本理解能力。

### 提示词生成模块测试分析

* 测试数据集

本次测试的数据集源自对文本理解模块的测试分析，通过收集文本理解模块在处理Chinese Squad数据集时对每个文本所生成的四个分段，并这些分段内容经过逻辑梳理和结构优化，形成了连环画所需的独立故事情节，作为提示词生成模块的输入数据集。此数据集确保了测试的针对性，即考察模块能否准确提取关键信息并转化为适用于Stable Diffusion模型的提示词。

* 测试过程

1. 数据输入：将上述测试数据集中的每个文本分段作为独立输入，送入提示词生成模块。
2. 提示词生成：模块针对每个文本分段生成一组提示词，包含描述性关键词及权重分配。
3. 格式验证：对比每组生成的提示词集合与Stable Diffusion模型所适用的提示词格式要求，确认是否在结构、语法、关键词使用等方面严格遵循模型规范。
4. 内容评估：对每组提示词的描述内容进行深入分析，评估其是否准确捕捉到了对应文本分段的核心信息，无明显遗漏，能否有效引导Stable Diffusion模型生成与原文相符的视觉内容。

* 测试结果

统计测试结果数据，提示词生成模块在生成每个提示词集合时所需的时间为2秒，体现出高效的处理速度，有利于实现大批量文本到连环画的快速转化。

且生成的提示词集合均被验证为符合Stable Diffusion模型提示词标准。这表明模块在生成过程中严格遵守了目标模型的输入规范，确保生成的提示词可以直接用于驱动模型进行图像生成，无需额外的格式调整。

观察每次返回的提示词内容，生成的提示词集合在描述内容上与原文相比没有明显遗漏。这意味着模块成功提炼出了文本分段的核心情节、角色、氛围等要素，为后续的图像生成提供了详实、准确的信息基础。

测试结果样例：

* 分析与结论

提示词生成模块在本次测试中表现优秀，能够高效、准确地将文本理解模块处理后的文本分段转化为符合Stable Diffusion模型要求的提示词集合。

### 连环画生成模块和工作流管理模块的测试分析

* 测试过程

本次测试聚焦于连环画生成模块的效率、质量及其与工作流管理模块的集成情况，旨在验证以下几个核心方面：

1. **连环画生成模块效率：**衡量模块在接收到提示词集合后生成单张连环画图像所需的时间。
2. **连环画生成模块质量：**评估生成的连环画图像是否准确、生动地再现了对应的文本提示词内容，包括场景、角色、动作、情感等要素。
3. **工作流管理模块集成：**检查工作流管理模块是否能无缝对接连环画生成模块，通过设定和调度节点工作流，实现对大批量连环画生成任务的高效管理和参数控制。
4. **批量生成一致性：**测试工作流管理模块在批量生成连环画时，是否能确保所有图像在指定风格和固定参数下保持高度一致性。

* 测试结果

统计测试数据，连环画生成模块在接收到提示词集合后生成单张图像的平均时间为10秒，显示出模块具备高效处理能力，能够在短时间内完成大量连环画的生成任务。

且通过工作流管理模块批量生成的一系列连环画图像，经审查其风格呈现出高度一致的特点，表明模块在执行批量任务时能够精确保持预设的风格参数，确保整套连环画作品视觉风格的统一性。

在测试过程中，通过对工作流管理模块中的节点进行参数调整，连环画生成模块相应地产生了预期的变化，证明节点工作流对连环画生成模块的各项参数控制有效。这体现了工作流管理模块对连环画生成过程的精细调控能力，可以根据需求灵活调整生成结果。

测试结果样例：

* 分析与结论

连环画生成模块与工作流管理模块紧密结合，不仅确保了连环画生成的高效性与高质量，而且通过灵活的工作流配置实现了风格一致且参数可控的大批量连环画生成。

# 作品总结

## 作品特色与创新点

### 作品特色

1. **深度文本理解与精准图像生成：**智绘连环平台依托LLM大语言模型，实现了对输入文本的深度理解，精准捕捉文本中的情感、情境与细节。结合Stable Diffusion模型，平台能生成高度符合文本设定的连环画图像，确保情节连贯、人物特征鲜明、场景描述生动，极大提升了文生图作品的艺术表现力与叙事准确性。
2. **高度自定义工作流：**ComfyUI工作流界面为用户提供了前所未有的创作自由度，允许用户通过添加、调整节点来精细控制连环画生成过程，包括风格设定、氛围渲染、角色动态等高级属性。这种灵活性使平台能够满足多元化艺术需求，适应不同用户群体的个性化创作意图。
3. **集成创新与一站式解决方案：**智绘连环平台整合了前沿的LLM模型、Stable Diffusion模型与创新的工作流管理系统，形成了一站式连环画创作平台。这种集成设计简化了用户操作流程，提升了整体使用体验与工作效率，降低了技术门槛，使非专业人士也能便捷地创作高质量、高定制化的连环画作品。

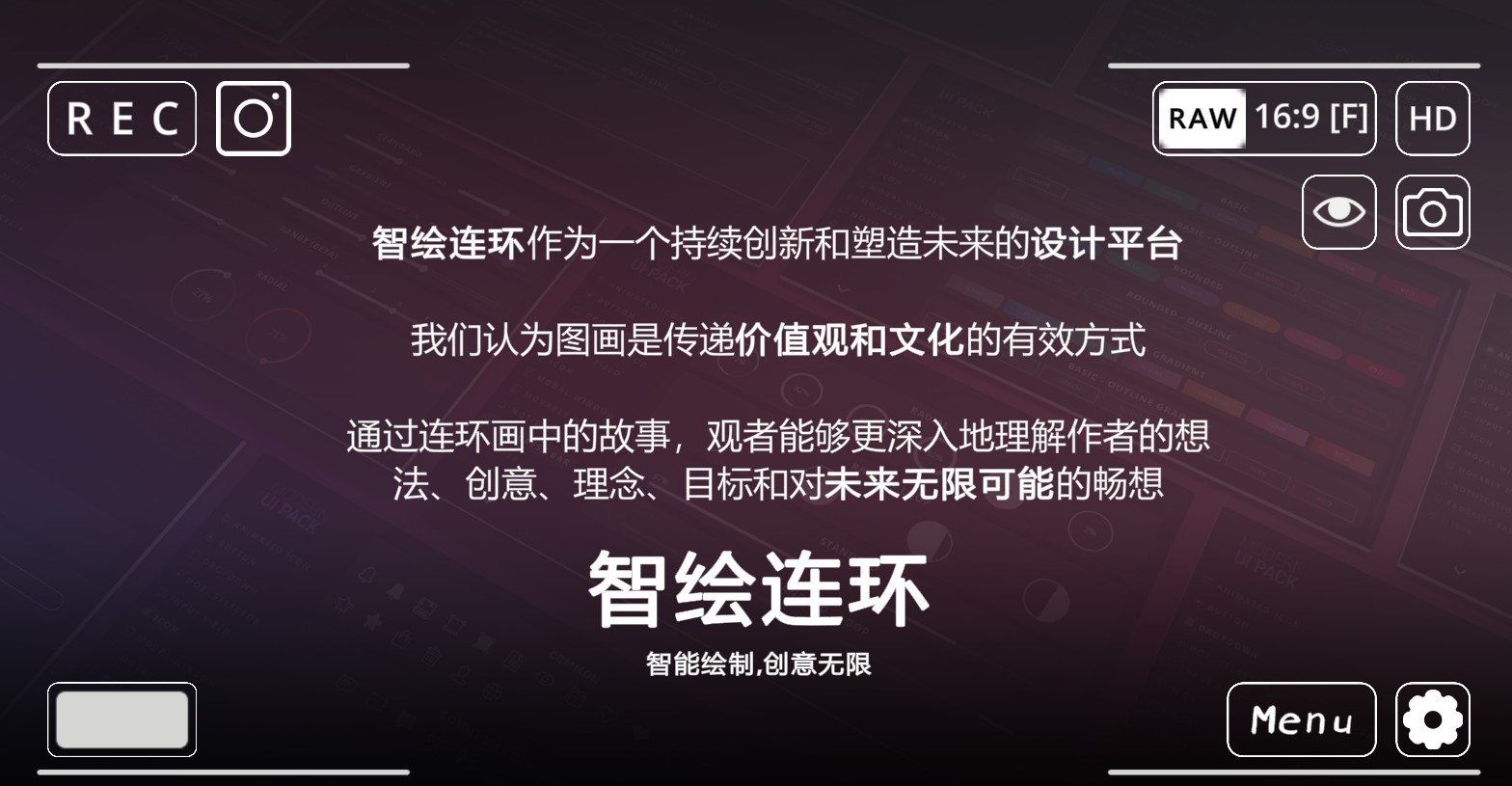
### 作品创新点

1. **连环画自动化生成技术：**平台创新性地结合LLM模型与Stable Diffusion模型，不仅克服了单一技术在内容创作上的局限性，更在文本到图像的转化过程中实现了质的飞跃，为创作者提供了前所未有的创作体验。实现从单一文本输入到一系列连贯、生动的连环画图像的自动转化，首创了基于上下文联系的AI连环图生成模式。
2. **智能化工作流管理系统：**ComfyUI工作流界面创新性地将复杂的Stable Diffusion操作流程简化为直观的节点结构，用户可以清晰追踪和修改生成过程，确保创作过程的可追溯性与可复现性，这是现有AI绘画工具所缺乏的重要功能。
3. **跨领域的应用价值：**智绘连环平台不仅革新了连环画创作方式，还在教育、科普、广告、个人创作、文学与艺术融合等领域展现出广泛应用潜力，为用户提供了一种新的、高效的内容表达方式，拓宽了创意表达的边界。
4. 可复用的解决方案：对生成的作品可以重复生成重复修改，而当前的市场普遍只能支持独立生成，而不能对已生成的作品进行调整，使得入手门槛与成品品质、工作效率有着卓越提升

## 应用推广

1. **教育及科普领域：**可以与教育机构、出版社合作，开发基于智绘连环平台的教材插图、在线课程资源包，助力教师和科普工作者将抽象知识转化为生动的连环画内容，提升教学效果与科普传播影响力。
2. **创意设计与广告营销：**可以与设计公司、广告代理商建立合作关系，将智绘连环平台纳入其服务项目，为品牌客户提供连环画形式的故事化营销内容，增强品牌形象，吸引消费者注意力。
3. **个人创作与娱乐：**通过社交媒体、线上社区、创意共享平台等渠道推广智绘连环平台，鼓励用户分享使用平台创作的作品，举办线上连环画创作比赛，激发大众创作热情，打造用户社群。
4. **文学与艺术融合：**可以与文学网站、出版社、艺术机构合作，推出将文学名著转化成连环画的活动项目，鼓励作家、艺术家使用智绘连环平台将文学作品转化为连环画，推动文学与视觉艺术的跨界交流与创新。

## 作品展望



图（看看能不能设计一个广告图）

随着人工智能技术的不断发展，智绘连环平台将致力于实现技术的持续升级与优化，密切关注并跟进最新的研究进展，适时更新LLM大语言模型和Stable Diffusion模型，以提升对文本内容的深度理解能力和图像生成的精度与质量。同时，ComfyUI工作流界面也将得到持续优化，增加更多高级编辑功能，为用户提供更加流畅、高效的创作体验。

跨平台与移动应用的开发将是智绘连环平台未来的重要方向。为了满足用户在不同场景下的创作需求，平台将推出桌面端和移动端应用，实现跨平台创作。用户无论是在办公室、家中还是户外，都能随时随地进行连环画创作，充分释放创作灵感。

社交互动与协作功能的引入将为智绘连环平台增添更多活力。平台将允许用户分享自己的作品，与其他用户进行评论互动，甚至支持多人协作共同创作连环画项目。这将有助于形成一个线上创作社区，汇聚更多创意与灵感，推动连环画艺术的繁荣与发展。

在商业模式探索方面，智绘连环平台将考虑推出付费高级功能、订阅服务以及版权交易市场等，为平台的可持续发展提供经济支持。这些商业模式将基于用户需求和市场反馈进行精心设计，旨在为用户提供更多价值的同时，也回馈创作者，激励他们产出更多优质内容。

此外，智绘连环平台还将密切关注AI生成内容的相关法律法规，并积极参与制定行业标准。在尊重知识产权、保护用户隐私的前提下，平台将确保健康、合规的发展，为用户提供一个安全、可靠的创作环境。

# 参考文献

[1]何文睿,高丹阳,周羿旭,等.基于扩散模型的多模态引导图像合成系统[J].北京信息科技大学学报(自然科学版),2023,38(06):80-87.

[2]刘泽润,尹宇飞,薛文灏,等.基于扩散模型的条件引导图像生成综述[J].浙江大学学报(理学版),2023,50(06):651-667.

[3]曹磊,俞剑红.AIGC技术在电影数字化创作与制作平台的创新应用[J].北京电影学院学报,2023,(11):80-91.

[4]陈嗣荣,冯敬益.提示词模板在人工智能对话中的实践研究[J].电脑知识与技术,2023,19(34):13-15+22.

[5]陈英,马洪涛.AIGC在艺术设计专业领域的神助攻——以Stable Diffusion为例[J].服装设计师,2024,(01):73-84.