

Генерация художественных изображений для книжных иллюстраций по текстовому описанию с использованием нейросети GPT-5.1

Кокухина Софья Александровна

Приамурский государственный университет им.Шолом-Алейхема

Студент

Аннотация

В статье демонстрируется процесс автоматической генерации художественных иллюстраций для книг на основе текстового описания с использованием мультимодальной нейросети GPT-5.1. В работе рассматриваются особенности интерпретации моделью литературных фрагментов, анализируются возможности визуализации художественных образов и оценивается влияние детализации текста на итоговое качество изображения.

Ключевые слова: нейросеть, GPT-5.1, генерация изображений, иллюстрации, искусственный интеллект, художественная визуализация, мультимодели.

Generation of artistic images for book illustrations based on a text description using the GPT-5.1 neural network

Kokukhina Sofya Alexandrovna

Sholom-Aleichem Priamursky State University

Student

Abstract

The article demonstrates the process of automatic generation of artistic illustrations for books based on a text description using a multimodal neural network GPT-5.1. The paper examines the features of the model's interpretation of literary fragments, analyzes the possibilities of visualizing artistic images, and evaluates the impact of text detail on the final image quality.

Keywords: neural network, GPT-5.1, image generation, illustrations, artificial intelligence, artistic visualization, multimodels.

1. Введение

1.1 Актуальность

Книжная иллюстрация является важной частью визуальной культуры и одним из ключевых средств художественной интерпретации литературного произведения. Процесс создания иллюстраций традиционно требует значительных творческих усилий, знания художественных техник и времени.

Появление мультимодальных моделей нового поколения, включая GPT-5.1, предоставляет возможность автоматизировать начальные этапы иллюстрирования. Модель способна анализировать литературные фрагменты,

выделять ключевые образы, персонажей, атмосферу и визуальные детали, преобразуя текст в художественные изображения, которые могут использоваться как черновые эскизы или концепты.

1.2 Обзор исследований

В. И. Курбатова рассматривает использование искусственных нейронных сетей для создания живописных изображений, отмечая, что современные алгоритмы способны генерировать оригинальные картины, комбинируя стили и содержание нескольких исходных изображений. При этом подчёркивается, что художественная ценность результата во многом зависит от постановки задачи и участия человека в выборе исходных данных и интерпретации результата [1]. А. В. Голованёва анализирует применение нейросетей в эскизировании и живописи, показывая, что генеративные модели могут выступать инструментом для быстрого получения вариантов композиции, цветовых решений и стилистических интерпретаций. В работе приводятся примеры использования нейросетей для преобразования набросков и переноса стиля, что делает такие технологии потенциально полезными и для предварительного этапа создания книжных иллюстраций [2]. Т. В. Пирязева рассматривает новые технологии креативного искусственного интеллекта в контексте творческих профессий, указывая на возможности генеративных моделей в поддержке художников, дизайнеров и иллюстраторов. Отмечается, что нейросети могут использоваться не только для автоматического синтеза изображений, но и как инструмент совместного творчества, когда человек задаёт концепцию, а модель помогает быстро получать визуальные варианты [3]. М. В. Кузьменкова исследует влияние искусственных нейронных сетей на художественное творчество, подчёркивая, что генеративные инструменты позволяют выйти за рамки традиционных техник и экспериментировать с новыми формами, цветами и композициями. В работе поднимаются вопросы изменения роли художника в условиях, когда часть визуальной работы выполняется алгоритмами, а также обсуждаются риски и преимущества такого подхода для развития визуальных искусств [4]. Г. В. Антонов рассматривает возможности использования методов машинного обучения и генеративных моделей, таких как Midjourney, для синтеза изображений по текстовому описанию. Показано, что подобные системы позволяют напрямую преобразовывать словесные формулировки в визуальные сцены, что открывает перспективы для автоматизации создания иллюстраций, в том числе для издательской сферы и цифровых учебных материалов [5].

1.3 Цель исследования

Цель исследования — продемонстрировать процесс генерации художественных иллюстраций к литературным произведениям по текстовому описанию с использованием мультимодальной нейросети GPT-5.1 и оценить степень соответствия сгенерированных изображений исходным литературным

фрагментам с точки зрения композиции, стиля и эмоциональной выразительности.

2 Материалы и методы

В исследовании использовалась мультимодальная нейросеть GPT-5.1, обладающая возможностью преобразовывать литературные описания в художественные изображения различной степени детализации и стилизации.

Работа модели рассматривалась как единый процесс интерпретации текста и формирование визуального результата. В ходе исследования использовались фрагменты литературных описаний, содержащие персонажей, атмосферу, природные и архитектурные объекты. Генерация иллюстраций выполнялась по заранее подготовленным текстовым промптам, отражающим структуру сцены, стиль и эмоциональный настрой.

3. Результаты

3.1 Генерация иллюстрации по краткому литературному описанию

На первом этапе проверялась способность модели создавать изображение на основе короткого фрагмента текста, содержащего общую атмосферу сцены.

Промпт:

«Создай художественную иллюстрацию. Описание:

Маленький мальчик стоит на опушке леса и смотрит на рассвет.

Лёгкий туман поднимается над травой, солнце едва появляется над горизонтом.

Стиль: мягкая акварель.»

GPT-5.1 сформировала изображение, в котором были корректно переданы утреннее освещение, приглушённые цвета и акварельная текстура. Несмотря на небольшую детализацию текста, модель отразила ключевые элементы сцены и эмоциональный настрой (Рис. 1).

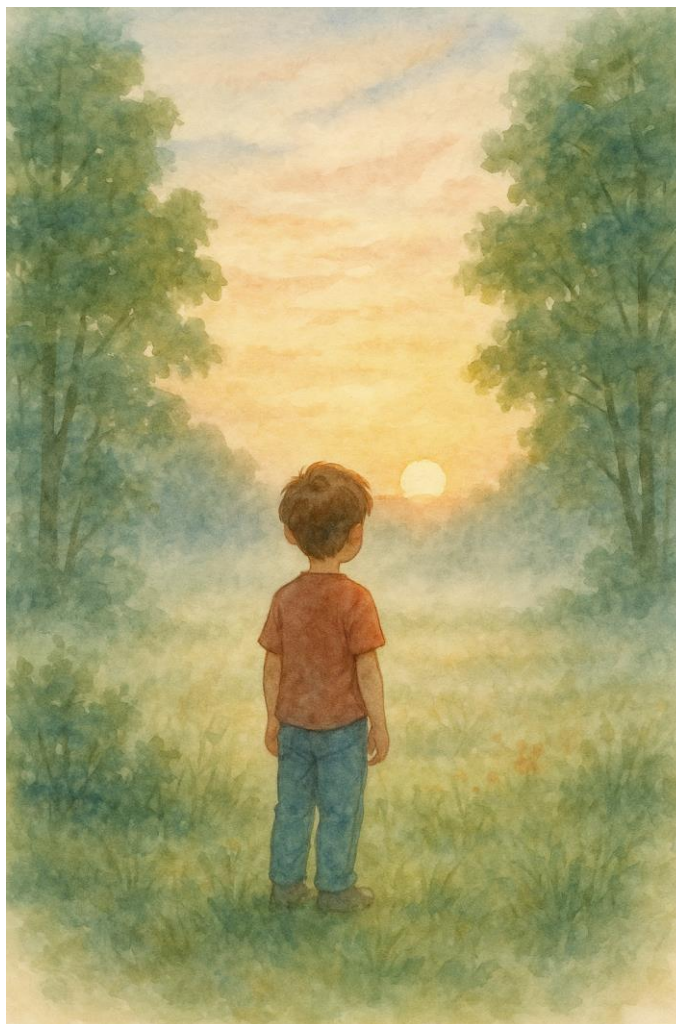


Рисунок 1 — Генерация иллюстрации по краткому описанию

3.2 Генерация иллюстрации по детальному описанию персонажа и окружения

Следующий эксперимент включал использование расширенного текстового описания с указанием деталей внешности персонажа, композиции сцены и художественного стиля.

Промпт:

«Создай художественную иллюстрацию. Описание:

Юная девушка в длинном синем плаще идёт по старинной библиотеке.

Высокие деревянные стеллажи тянутся к потолку, тусклый свет ламп создаёт тёплые блики.

Девушка держит в руках старинную книгу с золотым тиснением.

Стиль: классическая книжная иллюстрация XIX века.»

GPT-5.1 с высокой точностью визуализировала интерьер библиотеки, текстуры дерева и освещение. Книга в руках персонажа соответствовала текстовому описанию, а общий стиль напоминал классические гравюры XIX века (Рис. 2).

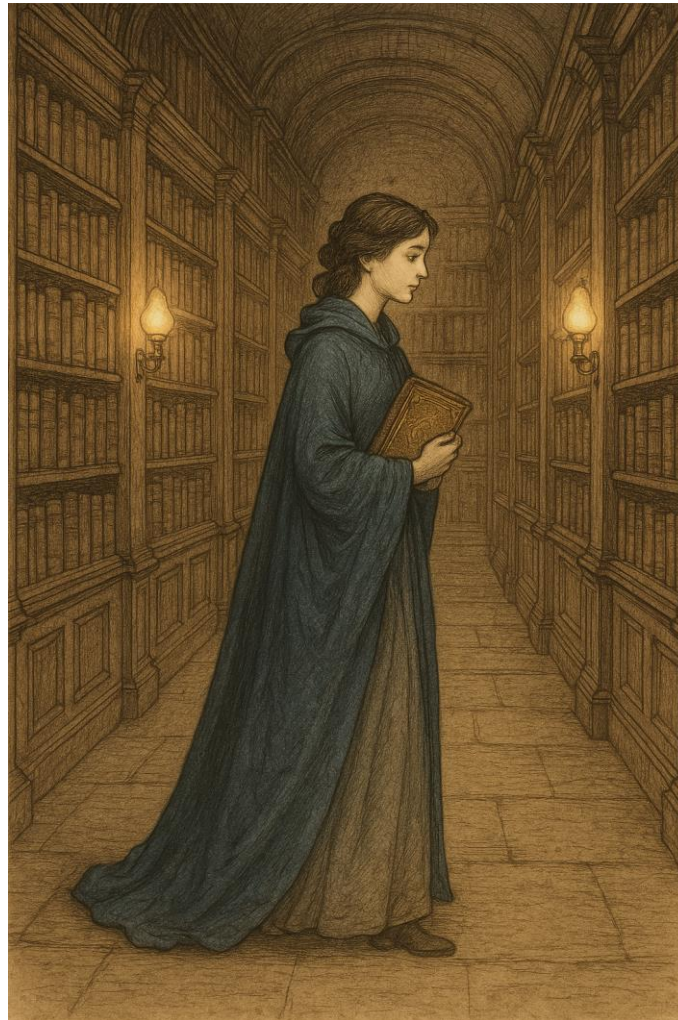


Рисунок 2 — Генерация иллюстрации с детализированным описанием

3.3 Генерация художественной сцены с заданием композиции и атмосферы

На заключительном этапе исследовалась способность модели передавать сложные эмоциональные и композиционные элементы.

Промпт:

«Создай иллюстрацию для фэнтезийной книги. Описание:

На вершине скалы стоит рыцарь в старых доспехах.

Перед ним — огромное тёмное море, над которым сгущаются грозовые облака.

Позади рыцаря развивается красный плащ, ветер поднимает сухие листья.

Стиль: тёмное фэнтези, умеренная детализация.»

Модель корректно интерпретировала композицию, передала драматическую атмосферу и цветовую гамму жанра. Иллюстрация включала описанные детали — развивающийся плащ, пасмурное небо, морской горизонт (Рис. 3).



Рисунок 3 — Генерация фэнтезийной иллюстрации

3.4 Сравнение результатов

Проведённый анализ показывает, что модель способна создавать художественные изображения, соответствующие как лаконичным, так и детализированным текстовым описаниям. Уровень соответствия зависит от полноты исходного текста: чем подробнее описание, тем точнее композиция, стиль и образность итоговой иллюстрации.

4 Выводы

Исследование подтвердило, что GPT-5.1 обладает высокой способностью интерпретировать литературные текстовые описания и преобразовывать их в художественные изображения. Модель корректно передаёт атмосферу произведения, композицию сцены и ключевые художественные элементы.

Использование GPT-5.1 позволяет:

- ускорить процесс создания эскизов для книжных иллюстраций;
- генерировать варианты визуализации по одному и тому же описанию;
- применять разные художественные стили без изменения исходного текста;
- снизить затраты на предварительные этапы художественного

проектирования.

Ограничением является отсутствие гарантированной точности соответствия специфике авторского стиля и конструктивным особенностям изображённых объектов, что требует последующей работы художника.

Библиографический список

1. Курбатова В. И. и др. Разработка проектов для Международного конкурса «Искусственный интеллект в создании картин» // Современные информационные технологии в образовании, науке и промышленности. Искусственный интеллект в создании картин. Издательство «Экон-Информ», 2021. С. 202–205.
2. Голованева А. В., Алибекова М. И. Нейромода: использование нейросетей в эскизировании и создании модных изделий // Инновации и технологии к развитию теории современной моды, «Мода (Материалы. Одежда. Дизайн. Аксессуары)», посвящённая Ф. М. Пармону. Российский государственный университет имени А. Н. Косыгина (Технологии. Дизайн. Искусство), 2022. С. 318–321.
3. Пирязева Т. В., Соколов И. В. Новые технологии креативного искусственного интеллекта для специалистов творческих профессий // Современные информационные технологии в образовании, науке и промышленности. М.: Издательство «Экон-Информ», 2023. С. 64–69.
4. Кузьменкова М. В. Влияние искусственных нейронных сетей на современное искусство // XVI Машеровские чтения: материалы международной научно-практической конференции студентов, аспирантов и молодых ученых, Витебск, 21 октября 2022 года. Т. 1. Витебск: Витебский государственный университет им. П. М. Машерова, 2022. С. 353–356.
5. Антонов Г. В., Иванов С. И. Возможности использования машинного обучения в области искусства // Технологии и инновации: сборник научных статей научно-педагогических работников, аспирантов и обучающихся / ФГБОУ ВО Великолукская ГСХА. Великие Луки: Великолукская государственная сельскохозяйственная академия, 2022. С. 8–10.