

ПОЛТЕВ М. В., КУДРЯШОВА О. М.**Система создания изображений из текста с помощью нейросети**

УДК 004.8.032.26, ГРНТИ 28.23.37

*Статья поступила в редакцию 02.07.2025*Система создания изображений из
текста с помощью нейросетиText-to-Image Generation System
Using Neural Networks**М. В. Полтев, О. М. Кудряшова****M. V. Poltev, O. M. Kudryashova**Ухтинский государственный
технический университет, г. Ухта;Ukhta State Technical University,
Ukhta;

В статье представлена работа по проектированию и разработке Системы создания изображений из текста с помощью нейросети. Анализ предметной области показал, что текущий способ создания изображений с помощью нейросети сложен для простого пользователя. Разработка системы создания изображений из текста с помощью нейросети повысит эффективность промптов пользователей и облегчит создание изображений.

The article presents work on the design and development of a Text-to-Image Generation System using neural networks. Analysis of the subject area revealed that the current method of generating images with neural networks is complex for the average user. The development of a text-to-image generation system using neural networks will enhance the effectiveness of user prompts and simplify image creation.

Ключевые слова: ИИ, REACT, JS, нейросеть, генерация, изображения, дизайн

Keywords: AI, REACT, JS, Neural network, image, generation, design

Введение

В 2022 году искусственные нейронные сети (нейросети) начали массово использоваться во всем мире. В этом году появились такие нейросети, как Midjourney и ChatGPT, которые быстро завоевали лидерство в своих областях. Например, ChatGPT достигли 100 миллионов активных пользователей всего через два месяца после запуска. [1]

Нейросети быстро покорили рынок, предоставляя большой спектр решений для широкого круга задач: творчества, работы, бизнеса и т. д. Особенно популярны стали нейросети для генерации изображений из текста. Рядовой пользователь больше не нуждается в услугах дизайнера при стилизации личного блога, начинающий бизнес может сгенерировать уникальный логотип, а в руках

умелого пользователя нейросеть становится мощным инструментом для творчества.

На данный момент нейросети для генерации изображений широко используются во всем мире, но для рядового пользователя это кажется пока сложным.

К этому приводят следующие проблемы:

- Сложность создания промпта. Для получения более ожидаемого результата используют методы управления генерацией, такие как: негативный промпт, указание важности слов, разделение на сегменты и т.д. с помощью соответствующего синтаксиса.
- Проблема поиска дополнительных референсов. На данный момент не на всех сайтах-аналогах есть каталог референсов, показывающий возможности модели нейросети и обладающий возможностью повторить изображение самостоятельно.
- Необходимость сохранять результат локально. Понравившийся результат не сохраняется в системе.
- Сложность решения индивидуальных задач. Создание логотипов требует дополнительный поиск соответствующих стилей и настроек генерации.

Все перечисленные проблемы говорят об актуальности данной темы.

Целью данной работы является облегчение процесса создания изображений с помощью нейросети.

Для выполнения поставленной цели необходимо выполнить следующие задачи:

- Создать инструмент для генерации запроса с помощью нейросети;
- Сохранять результат на сайте для дальнейшего редактирования;
- Интегрировать каталог референсов для генерации изображения;
- Реализовать генерацию логотипов;
- Разработать функциональный редактор результатов.

Предпроектное обследование

Онлайн-генерация изображений, как правило, состоит из нескольких этапов:

- ввод текстового запроса (промпта) или изображения;
- настройку параметров генерации (стиль, разрешение, детализация);
- просмотр и выбор вариантов изображений;
- сохранение, редактирование и экспорт результатов.

Основные подходы к генерации изображений с помощью нейронных сетей можно разделить на два типа:

1. На основе текстового описания. Генерация на основе текстового описания происходит за счет получения на вход нейронной сети словесной характеристики признаков требуемой картинки.

2. На основе существующих изображений. Генерация на основе существующих изображений происходит за счет получения на вход нейронной

сети существующего изображения, которое необходимо переработать или дополнить. [3]

После изучения предметной области были выделены функциональные требования к системе. Как итог система должна:

- Формировать каталог референсов с запросами;
- Предоставлять возможность генерации изображений из текста;
- Оказать помощь в создании профессионального промпта для получения более ожидаемого результата с помощью текстовой модели нейросети;
- Иметь механизм добавления понравившихся изображений;
- Предоставлять возможность генерации по уже готовому изображению;
- Иметь механизм генерации логотипов с предустановленными администратором настройками.

При поиске программ для автоматизации процесса были рассмотрены следующие решения, такие как:

- FusionBrain

«Fusion Brain» — это платформа, основанная на нейросети Сбербанка Kandinsky. Kandinsky — это российский аналог Midjourney. Решение автоматизирует процессы создания изображений с помощью нейросети. Кроме того, в решении функциональный редактор изображений с помощью инструмента «ластик». [5]

- TurboText

«TurboText» — это интерфейс для одноименной нейросети для генерации изображений. На сайте доступна бесплатная версия, в которой число запросов ограничено. Данная система позволяет генерировать изображения, редактировать полученный результат. [6]

Ни одно из решений не может сохранять результат на сайте для дальнейшего редактирования, кроме того, ни одна из систем не предлагает облегчение решения индивидуальных задач пользователей. Все это говорит о необходимости создания собственного решения.

Проектирование системы

Для разрабатываемой системы была выбрана гибридная архитектурная модель. Компоненты системы развернуты в контейнерах Docker на одном физическом сервере, где контейнер с серверной частью (Node.js) является монолитным ядром, содержащим основную бизнес-логику приложения. Клиентская часть взаимодействует с сервером через Representational State Transfer Application Programming Interface (далее REST API).

Архитектура серверного приложения построена на основе трехзвенного подхода (представление, бизнес-логика, данные) и использует паттерн MVC (Model-View-Controller) с акцентом на разделение ответственности:

Модель (Model): Реализована на серверной части с использованием Node.js с Express. Модель отвечает за обработку бизнес-логики и взаимодействие с базой данных MongoDB. Данные извлекаются по запросу через Mongoose, обеспечивая stateless взаимодействие без хранения состояния на сервере.

Представление (View): Клиентская часть, реализованная с использованием React и React Native с TypeScript, отвечает за отображение данных и взаимодействие с пользователем. Представление формируется динамически на основе данных, полученных от сервера через REST API.

Контроллер (Controller): Логика управления реализована на серверной части через Express маршруты. Контроллер обрабатывает входящие запросы от клиента, координирует взаимодействие между моделью и представлением, а также возвращает данные в формате JavaScript Object Notation (далее JSON). Сервер поддерживает API для асинхронных запросов.

Серверное приложение развернуто на Docker, что обеспечивает масштабируемость и отказоустойчивость. Взаимодействие между клиентом и сервером осуществляется через защищенные протоколы (HTTPS) с аутентификацией JSON Web Token (далее JWT) (Рисунок 1).

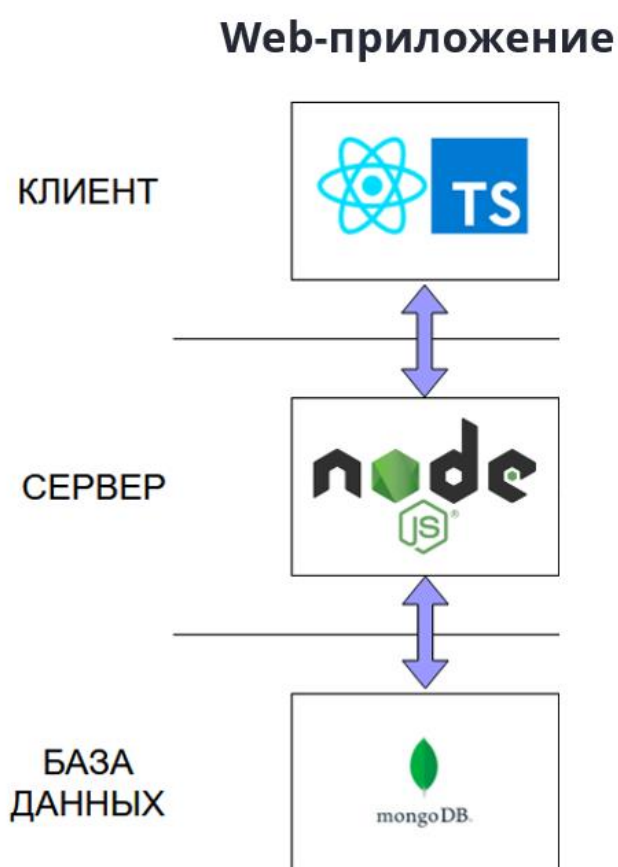


Рисунок 1. Архитектура серверного приложения

Для моделирования бизнес-процесса и проектирования базы данных было выбрано готовое программное решение Visual Paradigm. Visual Paradigm - это инструмент для моделирования и проектирования программного обеспечения, который поддерживает различные методологии разработки, включая UML, DFD, URL и другие. Решение существенно облегчает процесс построения схем и диаграмм, что упрощения процесса разработки.

В качестве среды разработки приложения была выбрана Visual Studio. Visual Studio — это интегрированная среда разработки (IDE), оптимизированная для

создания программного обеспечения для различных платформ. Разработанная Microsoft, она обеспечивает надежную поддержку языков программирования, таких как JavaScript, C#, C++ и Python, среди прочих.

Для описания информационной базы данных было изучено устройство существующих систем [3]. На основе этого, а также выделенных функциональных требований была создана логическая модель базы данных (Рисунок 2).

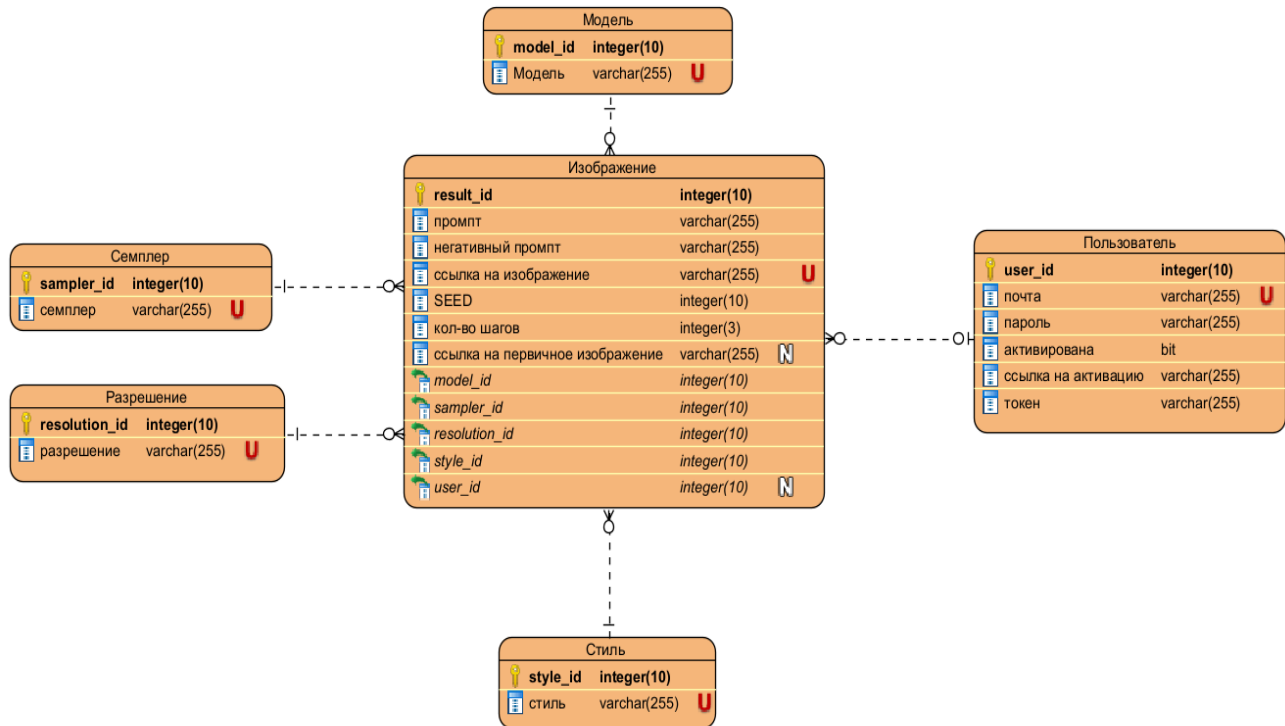


Рисунок 2. Логическая схема базы данных

Реализация и результаты

На данный момент уже реализованы представление системы в виде web-приложения, API для генерации изображений и база данных для хранения данных о пользователях и понравившихся изображениях.

На рисунках представлены интерфейсы системы (Рисунки 3-8).

Navigation

Main Editor

The Main Editor is a powerful tool for creating and editing content with a wide range of features. Perfect for users who need a versatile environment for designing and customizing their projects.

[Go to Main Editor](#)

Logo Editor

The Logo Editor is designed specifically for crafting professional logos with ease. It offers intuitive tools and templates to help you create a unique brand identity.

[Go to Logo Editor](#)

Gallery

Explore and manage your creations in the Gallery. This feature allows you to view, organize, and share your projects in a visually appealing way.

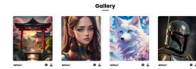
[Go to Gallery](#)

Рисунок 3. Интерфейс страницы навигации

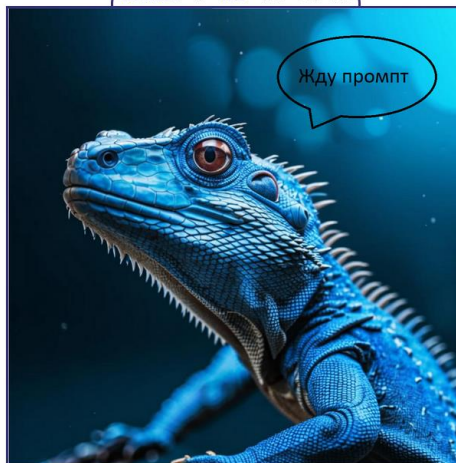
Main Redactor

Prompt:

neon boy

Negative prompt:

Describe negative prompt



Steps: 30    

Model:

Rev Animated v2 

Sampler:

DPM++ 2M 

SEED:

-1 

Рисунок 4. Интерфейс главного редактора



Logo Redactor

Prompt:
logo neon boy

Negative prompt:
Describe negative prompt



Рисунок 5. Интерфейс редактора логотипов



Gallery



REPEAT


[Show detail](#)


REPEAT



Model: revAnimated_v2Rebirth

Prompt: ((best quality)), ((masterpiece)), (detailed), close-up person, long hair, (fantasy art1.3), cute cyborg girl, highly detailed face, beautiful artwork illustration, (portrait composition:1.3), (8k resolution:1.2), (rating_explicit), (score_9, score_8_up, score_7_up, score_6_up, score_5_up, score_4_up, high res, 4k)



REPEAT


[Show detail](#)


REPEAT


[Show detail](#)

Рисунок 6. Интерфейс галереи референсов

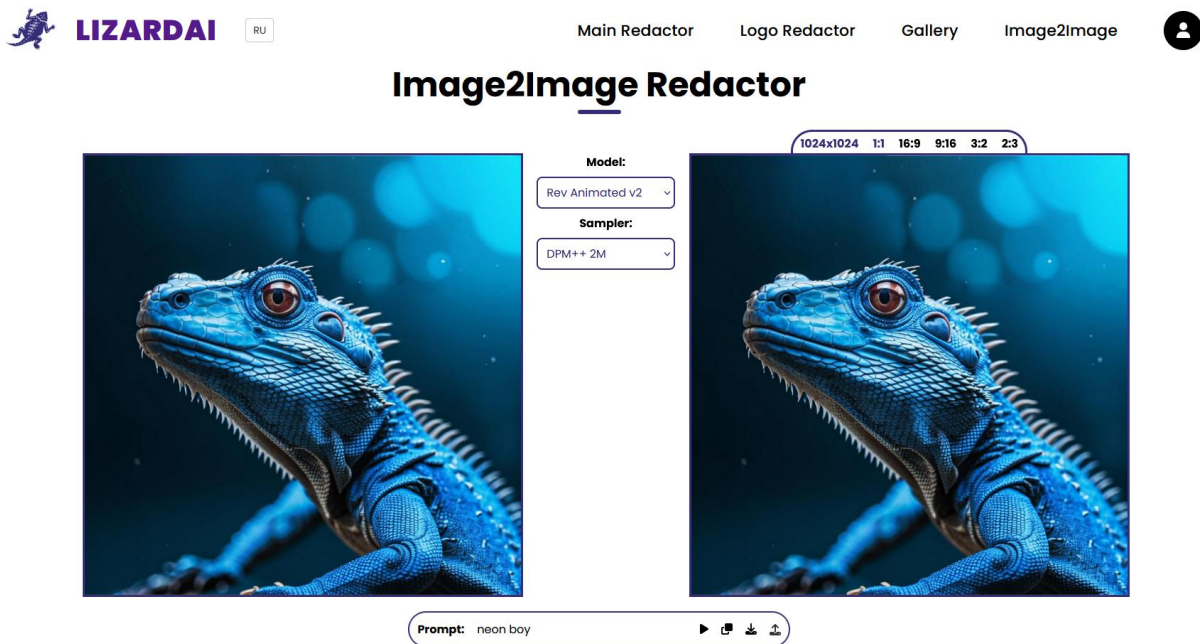


Рисунок 7. Интерфейс редактора Image2Image

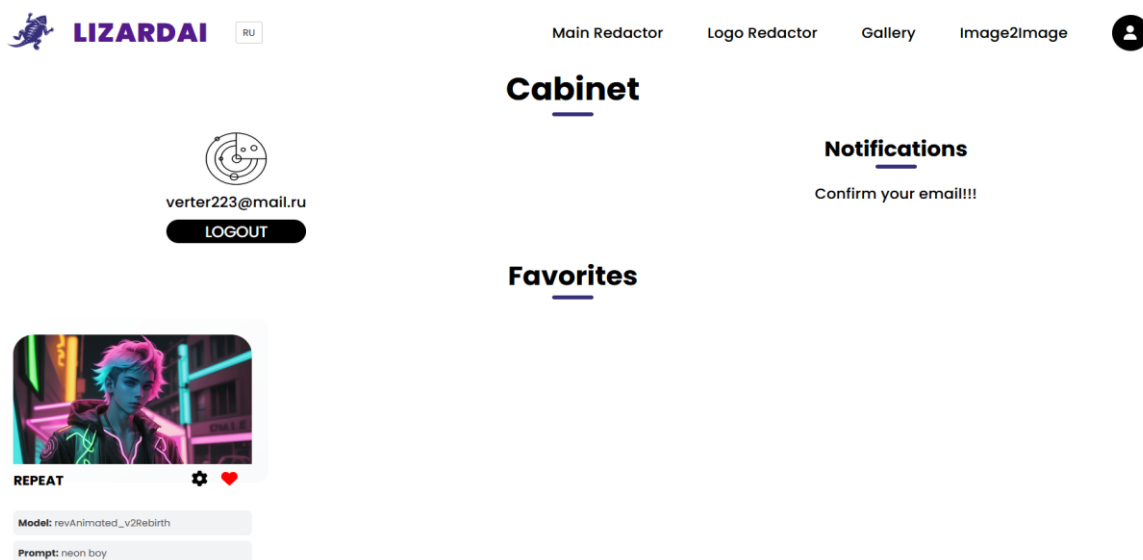


Рисунок 8. Интерфейс профиля пользователя

Заключение

В результате работы разработана и реализована информационная система для генерации изображений с помощью нейросетей, устраняющая ключевые проблемы пользователей: сложность создания промптов, отсутствие референсов и необходимость локального сохранения результатов.

Архитектура на основе Node.js, React и MongoDB обеспечивает генерацию изображений, создание логотипов и работу веб-редактора.

Цель работы достигнута: созданное решение снижает порог входа для пользователей и упрощает работу с нейросетевыми моделями. Перспективы развития включают интеграцию новых моделей и расширение функционала редактора.

Список использованных источников и литературы:

1. Петерс С. В. Нейросети для генерации изображений: области применения и юридические проблемы эксплуатации [Электронный ресурс]//CyberLeninka.– Режим доступа: <https://cyberleninka.ru/article/n/neyroseti-dlya-generatsii-izobrazheniy-oblasti-primeneniya-i-yuridicheskie-problemy-ekspluatatsii/viewer> (дата обращения: 25.05.2025).
2. Самарина А. Е. Нейросети для генерации изображений: педагогический потенциал в высшем образовании [Электронный ресурс]//CyberLeninka. – Режим доступа: <https://cyberleninka.ru/article/n/neyroseti-dlya-generatsii-izobrazheniy-pedagogicheskiy-potentsial-v-vyshhem-obrazovanii> (дата обращения: 25.05.2025).
3. Ильинская Е. В. Применение генеративно-сопоставительных нейросетей для генерации изображений [Электронный ресурс]//CyberLeninka. – Режим доступа: <https://cyberleninka.ru/article/n/primenenie-generativno-sostyazatelnyh-neyrosetey-dlya-generatsii-izobrazheniy/viewer> (дата обращения: 25.05.2025).
4. Сайт "Fusion Brain" [Электронный ресурс]. – Режим доступа: <https://fusionbrain.ai/> (дата обращения: 02.05.2025).
5. Сайт "TurboText" [Электронный ресурс]. – Режим доступа: <https://www.turbotext.ru/> (дата обращения: 02.05.2025).

List of references

1. Peters S. V. Neural networks for image generation: areas of application and legal problems of operation [Electronic resource]//CyberLeninka.– Access mode: <https://cyberleninka.ru/article/n/neyroseti-dlya-generatsii-izobrazheniy-oblasti-primeneniya-i-yuridicheskie-problemy-ekspluatatsii/viewer> (date of access: 05/25/2025).
2. Samarina A. E. Neural networks for image generation: pedagogical potential in higher education [Electronic resource]//CyberLeninka. – Access mode: <https://cyberleninka.ru/article/n/neyroseti-dlya-generatsii-izobrazheniy-pedagogicheskiy-potentsial-v-vyshhem-obrazovanii> (date of access: 05.25.2025).
3. Ilyinskaya E. V. Application of generative adversarial neural networks for image generation [Electronic resource]//CyberLeninka. – Access mode: <https://cyberleninka.ru/article/n/primenenie-generativno-sostyazatelnyh-neyrosetey-dlya-generatsii-izobrazheniy/viewer> (access date: 05/25/2025).
4. Website "Fusion Brain" [Electronic resource]. – Access mode: <https://fusionbrain.ai/> (access date: 05/02/2025).
5. Website "TurboText" [Electronic resource]. – Access mode: <https://www.turbotext.ru/> (access date: 05/02/2025).