

Решение задач с подписью изображений при помощи SceneXplain Ai

Бокач Никита Александрович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В данной статье исследуется использование модели искусственного интеллекта SceneXplain AI для решения задач с подписью изображений. Сценарий SceneXplain AI основан на архитектуре глубокого обучения и предназначен для автоматического анализа и интерпретации семантики изображений.

Ключевые слова: Ai, искусственный интеллект, описание изображений.

Solving problems with image caption using SceneXplain Ai

Bokach Nikita Aleksandrovich

Sholom-Aleichem Priamursky State University

Student

Abstract

This article explores the use of the SceneXplain AI artificial intelligence model to solve problems with image caption. The SceneXplain AI script is based on a deep learning architecture and is designed for automatic analysis and interpretation of image semantics.

Keywords: Ai, artificial intelligence, image description.

1 Введение

1.1 Актуальность

Новые технологии в области компьютерного зрения и искусственного интеллекта оказывают колоссальное влияние на различные сферы нашей жизни, включая медицину, автомобильную промышленность, маркетинг и многие другие. Одной из самых захватывающих и перспективных областей является компьютерное зрение с использованием искусственного интеллекта для анализа и понимания изображений. В этой статье мы рассмотрим применение метода SceneXplain AI для решения задач с подписью изображений.

1.2 Обзор исследований

Статья Г. В. Прозорова посвящена компьютеризации исследований литологического состава и свойств горных пород по фотографиям керна [1]. Статья И. А. Медведева посвящена такой актуальной проблеме, как совершенствование коммерциализации технологий компьютерного зрения с целью повышения производительности и обеспечения эффективной работы предприятия [2]. В статье К. И. Кий рассмотрел задачу обнаружения сигнальных объектов на изображениях дорожных сцен, и предложены алгоритмы ее решения на основе метода геометризованных гистограмм [3]. В статье А. В. Сириченко изучает вопросы применения технологии компьютерного зрения в

автоматизированных системах управления технологическими процессами в машиностроении [4].

1.3 Цель исследования

Целью исследования является обзор нейросети SceneXplain, которая поможет с описанием изображений.

2 Обзор и работа с нейросетью Rationale

SceneXplain AI – это инновационная технология, основанная на современных методах глубокого обучения. Она позволяет автоматически анализировать и классифицировать изображения, а также присваивать им точные подписи. Благодаря нескольким методам анализа изображений, можно получать как быстрое описание изображений, так и более детальное.

Для начала работы с SceneXplain [5] необходимо зарегистрироваться. Переходим на страницу приложения и нажимаем кнопку авторизоваться. Далее проходим быструю регистрацию с помощью социальных сетей или электронной почты (рис.1).

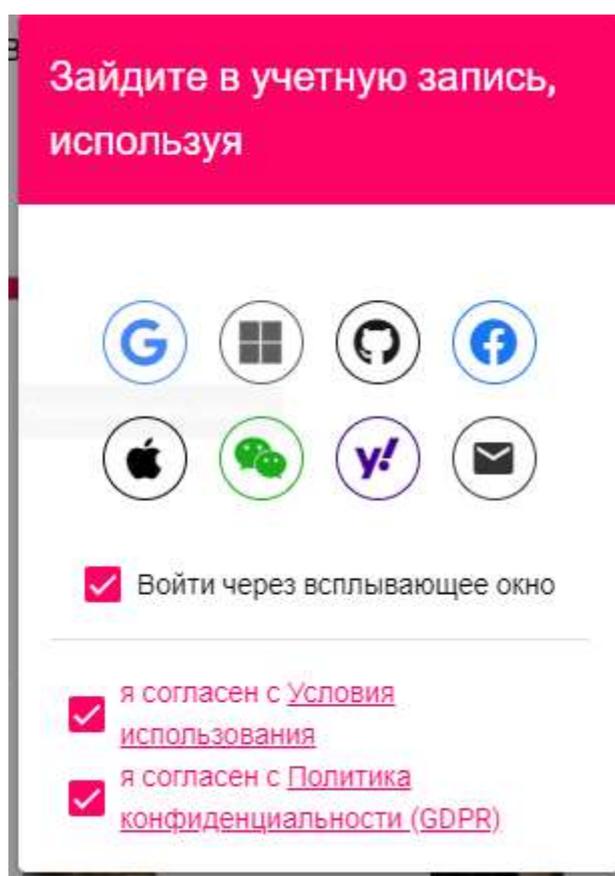


Рисунок 1 – Регистрация в приложении Rationale

В бесплатном тарифном плане приложения доступно 20 запросов, если посещать приложение каждый день, то начисляется 1 запрос. После авторизации открывается главная страница приложения (рис.2).

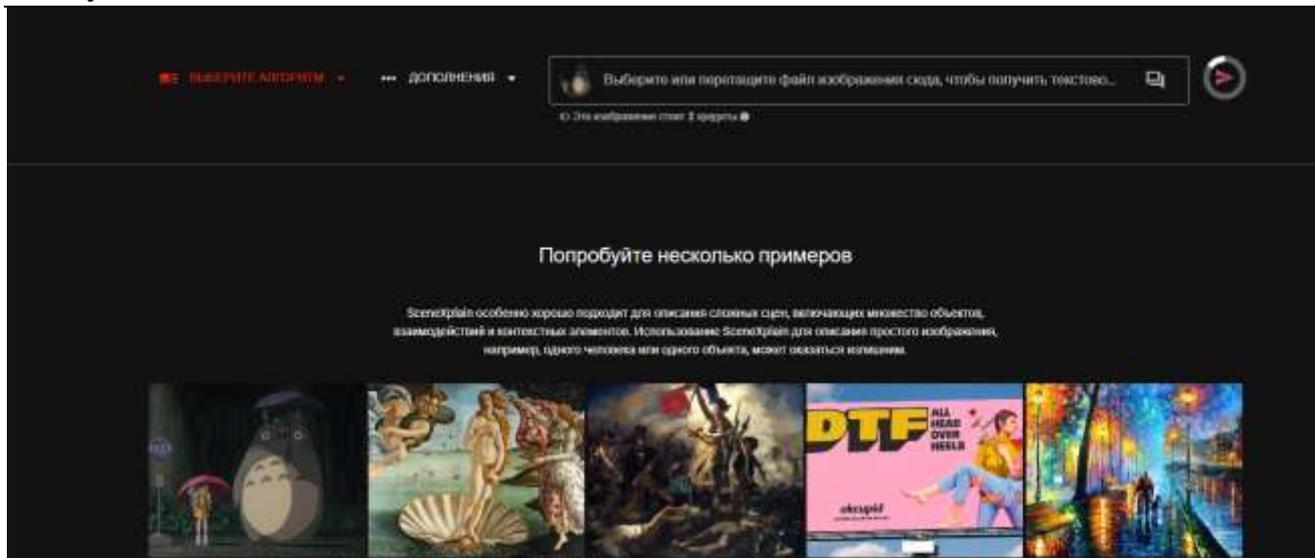


Рисунок 2 – Главная страница приложения

SceneXplain умеет анализировать изображения разными алгоритмами (рис.3).

<p>Dune рекомендуемые</p> <p>Улучшенная по сравнению с «Comet», «Dune» оптимизирована для поп-культуры и распознавания известных личностей. Это также сводит к минимуму проблему галлюцинаций для сложных и простых сцен.</p>	<p>Скорость ●●○○○</p> <p>Подробности ●●●●○</p> <p>Актуальность ●●●●○</p> <p>Фактичность ●●●●○</p>
<p>Ember</p> <p>Алгоритм «Ember» устраняет тенденцию включать краткое описание сцены в финальное предложение, как это видно в алгоритме «Dune», что приводит к более кратким и релевантным описаниям сцены.</p>	<p>Скорость ●●○○○</p> <p>Подробности ●●●●○</p> <p>Актуальность ●●●●○</p> <p>Фактичность ●●●●○</p>
<p>Flash</p> <p>Оптимизированный по скорости, этот алгоритм имеет тенденцию генерировать краткие и менее подробные описания, аналогичные тем, которые выдает BLIP2.</p>	<p>Скорость ●●●●●</p> <p>Подробности ●○○○○</p> <p>Актуальность ●●○○○</p> <p>Фактичность ●●●●○</p>
<p>Comet</p> <p>«Комета» призвана решить проблему галлюцинаций для простых сцен, сохраняя при этом скорость и детализацию «Болта».</p>	<p>Скорость ●●●●○</p> <p>Подробности ●●●●○</p> <p>Актуальность ●○○○○</p> <p>Фактичность ●●●●○</p>
<p>Bolt</p> <p>«Болт» оптимизирован для захвата каждой детали сложной сцены. Это немного медленнее и может вызвать галлюцинации для простых сцен.</p>	<p>Скорость ●○○○○</p> <p>Подробности ●●●●○</p> <p>Актуальность ●●●●○</p> <p>Фактичность ●●○○○</p>
<p>Aqua</p> <p>Aqua – самый ранний алгоритм для SceneXplain. Это быстро и точно для сцен средней сложности.</p>	<p>Скорость ●●○○○</p> <p>Подробности ●●○○○</p> <p>Актуальность ●●●●○</p> <p>Фактичность ●●○○○</p>

Рисунок 3 – Алгоритмы анализа изображений

Для тестирования приложения отправим несколько запросов в разных режимах анализа и посмотрим результаты (рис.4-8).

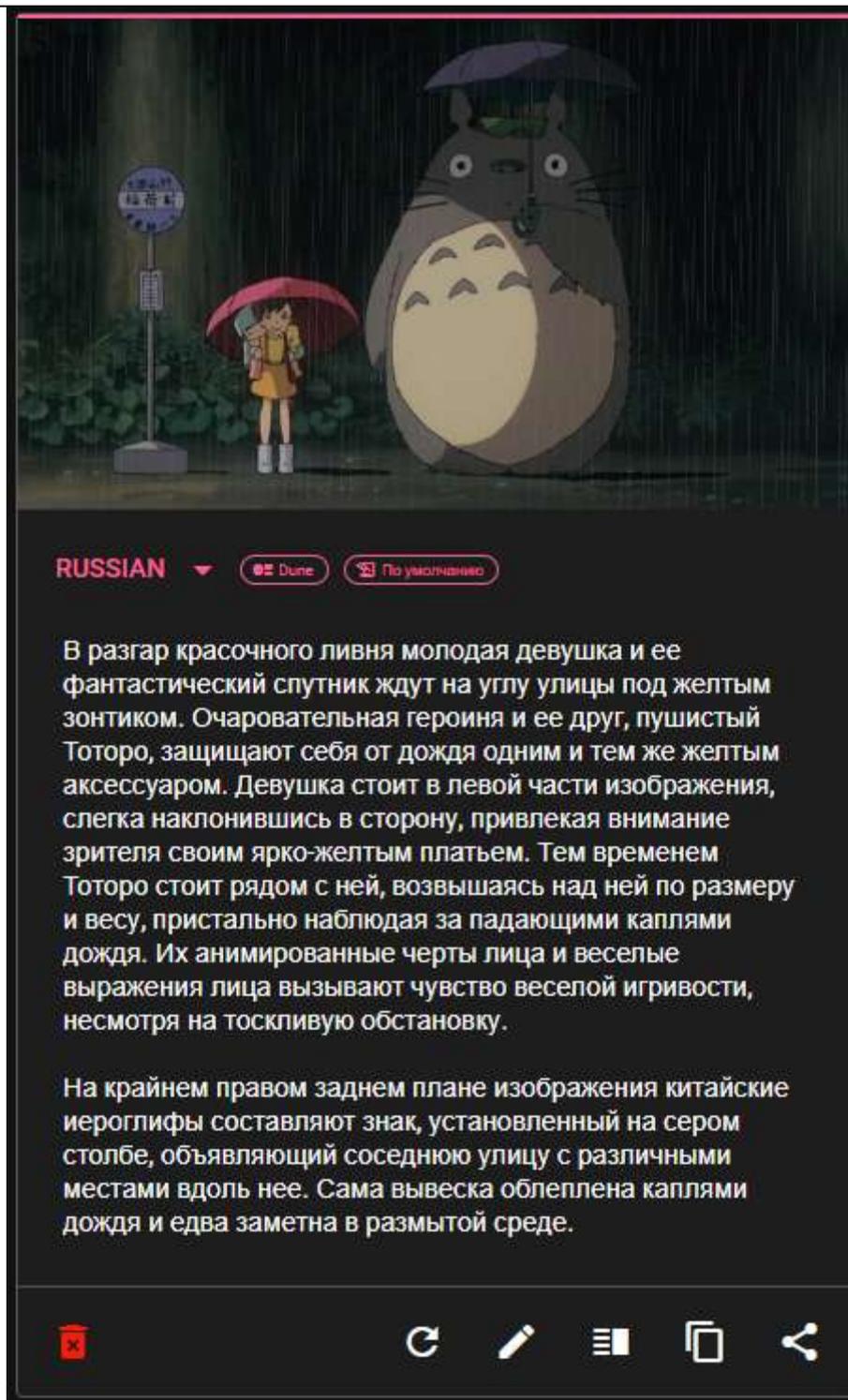


Рисунок 4 – Анализ изображения алгоритмом Dune



Рисунок 5 – Анализ изображения алгоритмом Ember



Рисунок 6 – Анализ изображения алгоритмом Volt

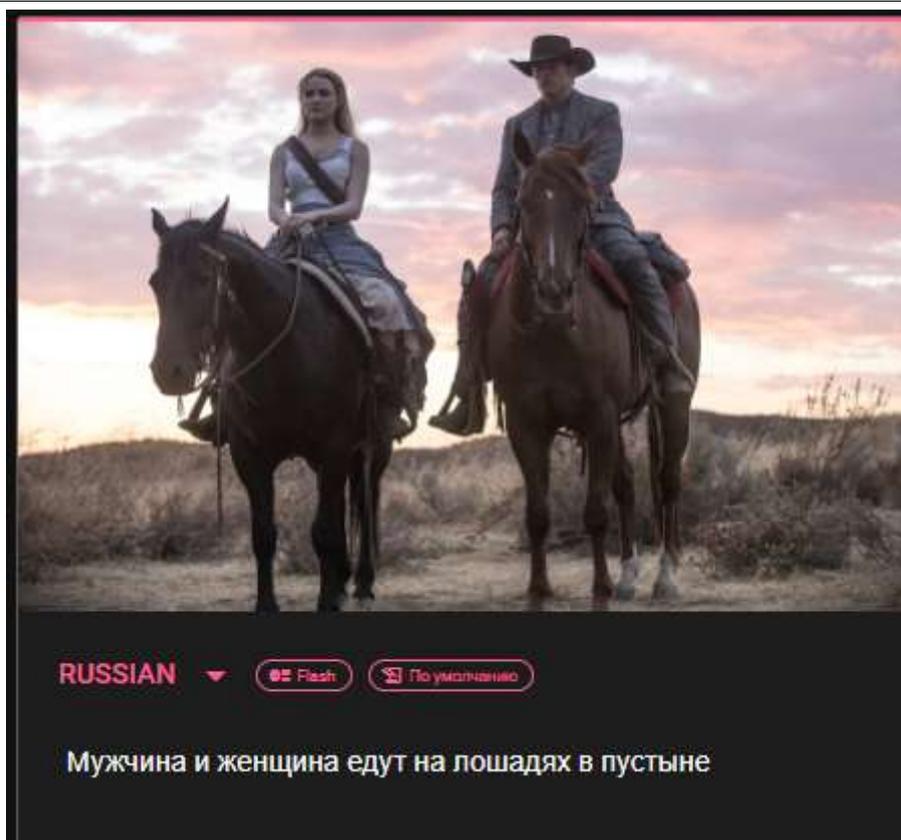


Рисунок 7 – Анализ изображения алгоритмом Flash

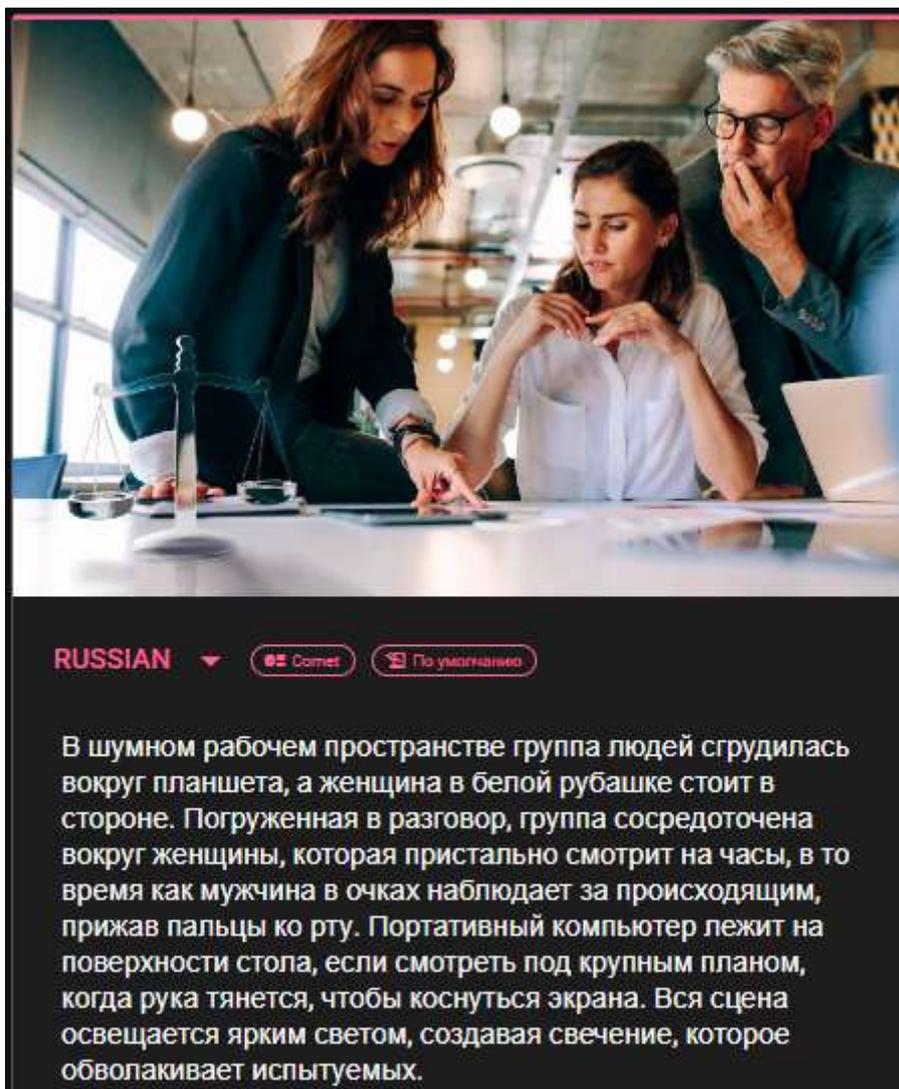


Рисунок 8 – Анализ изображения алгоритмом Comet

В заключение, можно с уверенностью сказать, что использование SceneXplain AI для решения задач с подписью изображений - это инновационный и эффективный подход, который позволяет автоматизировать процесс анализа и обработки визуальной информации. SceneXplain AI обладает не только высокой точностью в определении объектов на изображениях, но и способностью объяснить свои выводы, что придает ему большую прозрачность и доверие со стороны пользователей. Благодаря такому подходу, студенты исследователи, разработчики и специалисты в области компьютерного зрения могут значительно увеличить эффективность своей работы и получить более точные и интерпретируемые результаты.

Библиографический список

1. Прозорова Г. В. Анализ пустотного пространства пород по цифровым фотографиям керна с использованием алгоритмов компьютерного зрения //Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2021. № 4/1.

2. Medvedeva I. A. Automated complex development project based on computer vision technology // Natural and Technical Sciences. 2022. №. 7(170).
3. Кий К. И. Алгоритмы компьютерного зрения для анализа сигнальных объектов в дорожной сцене // Робототехника и техническая кибернетика. 2018. № 1(18).
4. Сириченко А. В. Применение технологии компьютерного зрения для решения задачи контроля качества продукции в машиностроении // Автоматизированное проектирование в машиностроении. 2019. № 7.
5. Система искусственного интеллекта SceneXplain AI URL: <https://scenex.jina.ai/> (дата обращения: 27.06.2023)