

ProjectSketchers:

система распознавания объектов и текста

Состав команды и роль:

Худяков Дмитрий Алексеевич (Распознавание текста и общая координация проекта)

Буров Егор Геннадьевич (Веб-Интерфейс)

Мосейчук Иван Сергеевич (Обработка Yolo)

Крутько Александр Сергеевич (Работа с метаданными)



Добро пожаловать на презентацию, посвященную нашей системе распознавания объектов и текста. Сегодня мы подробно рассмотрим архитектуру, используемые технологии и достигнутые результаты. Актуальность этой системы сложно переоценить, ведь она находит применение в самых разных областях, от автоматизации и повышения безопасности до анализа данных и оптимизации бизнес-процессов. Распознавание объектов позволяет обнаруживать и классифицировать объекты на изображениях, а распознавание текста — извлекать текстовую информацию из изображений и документов.



Сделано командой Мариупольские С

YOLO: Распознавание объектов в реальном времени

Архитектура и преимущества


YOLO (You Only Look Once) — это передовая архитектура нейронной сети, разработанная для быстрого и точного распознавания объектов в реальном времени. Её ключевые преимущества — скорость, точность и простота.

Скорость и точность

YOLO обеспечивает высокую скорость обработки, что критически важно для задач, требующих мгновенной реакции. Несмотря на это, YOLO демонстрирует впечатляющую точность обнаружения объектов, конкурируя с более сложными моделями.

Сравнение с другими моделями

По сравнению с Faster R-CNN и SSD, YOLO выигрывает в скорости обработки, достигая 30 FPS на COCO датасете против 7 FPS у Faster R-CNN и 45 FPS против 22 FPS у SSD.



Yandex Speller: Распознавание и исправление ошибок

1 Интеграция и преимущества

Для распознавания и исправления ошибок в тексте мы интегрировали Yandex Speller. Этот сервис обладает широким языковым охватом и высокой точностью распознавания.

2 Автоматическое исправление ошибок

Speller позволяет автоматически исправлять опечатки и грамматические ошибки в распознанном тексте, адаптируясь к различным стилям и контекстам.

3 API и SDK

Благодаря удобному API, Yandex Speller в нашу систему прошла быстро и эффективно, что позволило нам успешно внедрить исправление грамматических ошибок в тексте.



Tesseract OCR: Распознавание текста

Обзор Tesseract OCR

Tesseract OCR – это технология оптического распознавания символов, обеспечивающая высокую производительность.

Поддержка форматов и адаптация

Система поддерживает различные форматы изображений (JPG, PNG, TIFF, PDF) и адаптируется к различным шрифтам и стилям текста.

Точность и скорость

Tesseract OCR обеспечивает до 98% точности при хорошем качестве изображения, распознавая одну страницу всего за 2-3 секунды.



Архитектура системы: Взаимодействие компонентов

1

YOLO: Обнаружение объектов

YOLO обнаруживает объекты на изображении, такие как номерные знаки или документы.

2

Tesseract OCR: Извлечение текста

Tesseract OCR извлекает текст из обнаруженных объектов, обеспечивая высокую точность распознавания.

3

Yandex Speller: Проверка и исправление

Yandex Speller проверяет и исправляет ошибки в распознанном тексте, повышая общее качество информации.

Реализация и настройка: Практические аспекты

1

Подготовка данных

Первый этап — подготовка датасета, написание программы.

2

Обучение моделей

Далее происходит обучение моделей с использованием Python, TensorFlow и OpenCV.

3

Настройка параметров

Завершающий этап — настройка параметров для оптимизации производительности и масштабируемости системы.



```
from flask import Flask, request, render_template, send_from_direct
import torch
import cv2
import numpy as np
from PIL import Image, ImageDraw, ImageFont, ExifTags
import io
import os
import time
import base64
import pytesseract
import requests
```

Необходимые импорты библиотек для работы программы

Загрузите изображение или выполните поиск по метаданным

Выберите файл

Файл не выбран

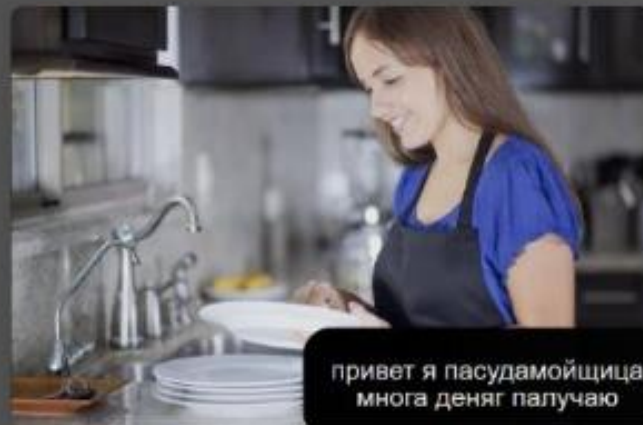
Загрузить изображение

Введите слово для поиска

Поиск по метаданным

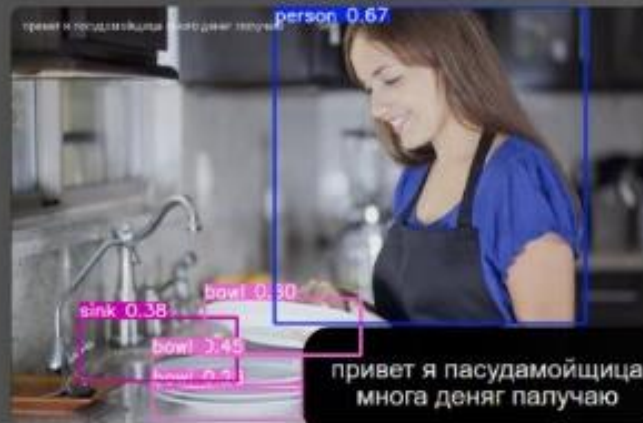
© 2025 Project Sketchers from Mariupol

Исходное изображение



привет я пасудамойщица
много денег палучаю

Исправленное изображение



привет я пасудамойщица
много денег палучаю

Информация о результатах:

Обнаруженные объекты: person (0.67), bowl (0.60), bowl (0.45), sink (0.38), bowl (0.28)

Текст был изменен: 'привет я пасудамойщица много денег палучаю' -> 'привет я посудомойщица много денег получаю'

Загрузить новое изображение



Сортировать ▾



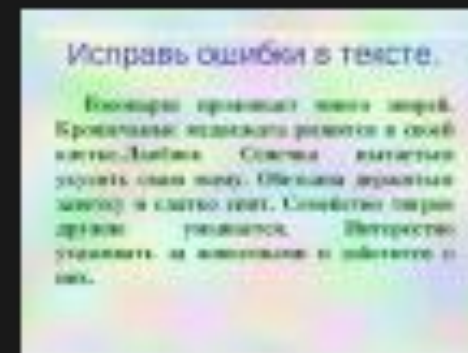
corrected_image
_1742028141.png



corrected_image
_1742028152.png



corrected_image
_1742028213.png



corrected_image
_1742028243.png

Сохраняет обработанные файлы в папку проекта.

Спасибо за
ВНИМАНИЕ

```
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Object Detection & Metadata Search</title>
  <link rel="stylesheet" href="https://fonts.googleapis.com/css2?family=Roboto:wght@400;700;900&display=block">
  <style>
    body {
      font-family: 'Roboto', sans-serif;
      background-color: #f4f4f4;
      margin: 0;
      padding: 0;
      display: flex;
      flex-direction: column;
    }
  </style>
</head>
```