

## Введение

Компьютерное зрение становится все более актуальным в современном мире, так как оно находит применение в различных сферах, включая медицину, транспорт и безопасность. Эта технология значительно упрощает обработку и анализ визуальной информации, что открывает новые возможности для автоматизации и повышения эффективности.

### Актуальность компьютерного зрения

- **Рост объема данных:** С увеличением объемов визуальной информации, необходима автоматизация ее обработки. Компьютерное зрение позволяет эффективно анализировать изображения и видео, извлекая из них полезные данные.
- **Применение в медицине:** В области здравоохранения компьютерное зрение используется для диагностики заболеваний, анализа медицинских изображений и мониторинга состояния пациентов, что способствует более точному и быстрому принятию решений.
- **Безопасность и мониторинг:** В системах видеонаблюдения компьютерное зрение помогает в распознавании лиц, обнаружении подозрительных действий и повышении общей безопасности общественных мест.
- **Автономные транспортные средства:** В автомобильной промышленности технологии компьютерного зрения играют ключевую роль в разработке систем помощи водителю и автономных автомобилей, обеспечивая безопасность на дорогах.
- **Индустрия развлечений:** В игровой индустрии и виртуальной реальности компьютерное зрение улучшает взаимодействие пользователей с контентом, создавая более захватывающий опыт.
- **Анализ поведения пользователей:** В маркетинге и рекламе компьютерное зрение позволяет анализировать поведение клиентов, что

помогает в создании более целевых и эффективных рекламных кампаний.

- **Развитие технологий:** С каждым годом технологии компьютерного зрения становятся более доступными и мощными, что открывает новые горизонты для их применения в различных отраслях.

**Цель проекта:** узнать, что такое компьютерное зрение

**Задача проекта:**

1 Разобраться для чего нужно компьютерное зрение

2 Создать рабочий код

**Гипотеза о будущем компьютерного зрения:**

- **Увеличение применения:** ожидается, что технологии компьютерного зрения будут активно внедряться в новые области, такие как сельское хозяйство, экология и образование, что расширит их влияние на общество.
- **Интеграция с ИИ:** С развитием искусственного интеллекта, компьютерное зрение будет интегрироваться с другими технологиями, что позволит создавать более сложные и умные системы, способные к самообучению и адаптации.
- **Улучшение алгоритмов:** Прогресс в области машинного обучения и глубокого обучения приведет к созданию более точных и быстрых алгоритмов обработки изображений, что повысит качество анализа и снизит затраты на внедрение.
- **Этические и правовые аспекты:** С увеличением использования компьютерного зрения возникнут новые вызовы, связанные с конфиденциальностью и безопасностью данных, что потребует разработки новых этических норм и правовых рамок.
- **Образование и подготовка кадров:** Увеличение спроса на специалистов в области компьютерного зрения будет способствовать

развитию образовательных программ и курсов, что обеспечит подготовку квалифицированных кадров для работы с новыми технологиями.

### ***Практическая значимость искусственного интеллекта (ИИ):***

*Заключается в его способности повышать эффективность, производительность и генерировать инновации в разных отраслях.*

- **Автоматизация.** ИИ может выполнять задачи, которые раньше требовали участия человека, например, автоматическую обработку текста, роботизацию производственных процессов. Это позволяет сэкономить время и деньги для компаний и частных лиц.
- **Персонализация.** ИИ анализирует огромные объёмы данных, чтобы предоставлять персонализированные рекомендации и опыт для отдельных лиц. Это повышает удовлетворённость и лояльность клиентов.
- **Здравоохранение.** ИИ используется для диагностики заболеваний, выявления генетических маркеров и разработки индивидуальных планов лечения. Это приводит к более точным диагнозам и лучшим результатам для пациентов.
- **Эффективность.** ИИ может оптимизировать процессы, в том числе рабочие, делая бизнес более эффективным и продуктивным.
- **Инновации.** ИИ помогает предприятиям и исследователям открыть для себя новые идеи и разработать новые продукты и услуги, которые ранее были невозможны.
- **Транспорт.** ИИ используется для разработки автономных транспортных средств, в которых используются датчики, камеры и другие технологии для навигации по дорогам и обхода препятствий. Это делает их более безопасными и эффективными, чем традиционные транспортные средства.
- **Умные города.** Технологии ИИ содействуют созданию умных городов, где информационные системы контролируют и

оптимизируют инфраструктуру, делая города более устойчивыми и эффективными.

## Глава 1 Документация

### 1.1 Что такое компьютерное зрение

Компьютерное зрение — это область искусственного интеллекта и компьютерных наук, занимающаяся разработкой алгоритмов и систем, которые позволяют компьютерам и машинам "видеть", интерпретировать и понимать визуальную информацию из окружающего мира. Основная цель компьютерного зрения — автоматизация процессов, связанных с анализом изображений и видео, чтобы извлекать из них полезные данные и принимать решения на основе этой информации.

#### Основные аспекты компьютерного зрения:

1. **Обработка изображений:** это начальный этап, который включает в себя улучшение качества изображений, фильтрацию шумов, изменение размеров и другие операции, которые подготавливают данные для дальнейшего анализа.
2. **Извлечение признаков:** на этом этапе из изображений извлекаются ключевые характеристики (признаки), такие как края, текстуры и формы, которые помогают в идентификации объектов.
3. **Распознавание объектов:** Алгоритмы компьютерного зрения могут идентифицировать и классифицировать объекты на изображениях или в видео, например, распознавание лиц, автомобилей, животных и т.д.
4. **Сегментация:** Этот процесс включает разделение изображения на несколько сегментов или областей, чтобы упростить анализ и выделить интересующие объекты.
5. **Трекинг:** Отслеживание движущихся объектов в видео, что позволяет анализировать их поведение и взаимодействие с окружающей средой.
6. **3D-восстановление:** Создание трехмерных моделей объектов на основе двухмерных изображений, что полезно в таких областях, как архитектура и робототехника.

Компьютерное зрение активно развивается благодаря достижениям в области машинного обучения и глубокого обучения, что позволяет создавать более точные и эффективные системы для обработки визуальной информации.

## 1.2 Как действует компьютерное зрение

Компьютерное зрение работает через несколько ключевых этапов, которые позволяют системам обрабатывать, анализировать и интерпретировать визуальную информацию. Вот основные шаги, которые описывают, как действует компьютерное зрение:

### 1. Сбор данных

**Сенсоры и камеры:** Компьютерное зрение начинается с захвата изображений или видео с помощью различных сенсоров, таких как камеры, сканеры или другие устройства. Эти устройства преобразуют световые волны в цифровые данные.

### 2. Предобработка изображений

**Улучшение качества:** на этом этапе изображения могут быть обработаны для улучшения их качества. Это может включать:

- Устранение шумов (фильтрация).
- Коррекцию яркости и контрастности.
- Изменение размера и ориентации изображений.

### 3. Извлечение признаков

**Анализ структуры:** Изображения анализируются для извлечения ключевых признаков, таких как:

- Контуры и границы объектов.
- Текстуры и цвета.
- Уникальные характеристики, которые помогают в идентификации объектов.

### 4. Сегментация

**Разделение на области:** Изображение делится на несколько сегментов или областей, чтобы выделить интересующие объекты. Это может быть сделано с помощью различных методов, таких как:

- Пороговая сегментация.
- Сегментация на основе кластеризации.
- Сегментация с использованием нейронных сетей.

### 5. Распознавание и классификация

**Идентификация объектов:** на этом этапе система использует алгоритмы машинного обучения или глубокого обучения для распознавания и классификации объектов на изображении. Это может включать:

- Распознавание лиц.
- Идентификация объектов (например, автомобилей, животных и т.д.).
- Классификация сцен (например, город, природа).

## **6. Трекинг (отслеживание)**

**Мониторинг движущихся объектов:** если система работает с видео, она может отслеживать движущиеся объекты в реальном времени, анализируя их перемещение и поведение.

## **7. Интерпретация и принятие решений**

**Анализ данных:** на основе извлеченных данных и распознанных объектов система может принимать решения или генерировать выводы. Например, в системах безопасности это может быть сигнализация о подозрительном поведении.

## **8. Вывод результатов**

**Представление информации:** Результаты анализа могут быть представлены пользователю в виде отчетов, визуализаций или в реальном времени через интерфейсы, такие как экраны или мобильные приложения.

### 1.3 Применение компьютерного зрения

Компьютерное зрение находит широкое применение в различных областях, благодаря своей способности обрабатывать и анализировать визуальную информацию. Вот некоторые из ключевых областей применения компьютерного зрения:

#### 1. Медицина

- **Анализ медицинских изображений:** используется для диагностики заболеваний на основе рентгеновских снимков, МРТ, КТ и УЗИ. Алгоритмы могут выявлять опухоли, аномалии и другие патологии.
- **Мониторинг состояния пациентов:** Системы могут отслеживать изменения в состоянии здоровья пациентов, анализируя изображения и видео.

#### 2. Автономные транспортные средства

- **Навигация и управление:** Компьютерное зрение позволяет автомобилям распознавать дорожные знаки, разметку, пешеходов и другие транспортные средства, что критически важно для безопасного вождения.
- **Системы помощи водителю:** Технологии, такие как автоматическое экстренное торможение и адаптивный круиз-контроль, используют компьютерное зрение для повышения безопасности.

#### 3. Безопасность и видеонаблюдение

- **Распознавание лиц:** используется в системах безопасности для идентификации людей и контроля доступа.
- **Анализ поведения:** Системы могут отслеживать и анализировать поведение людей в общественных местах для выявления подозрительных действий.

#### 4. Промышленность и автоматизация

- **Контроль качества:** Компьютерное зрение используется для автоматического контроля качества продукции на производственных линиях, выявляя дефекты и несоответствия.
- **Робототехника:** Роботы используют компьютерное зрение для навигации, манипуляции объектами и выполнения задач в сложных условиях.

## 5. Розничная торговля

- **Анализ поведения покупателей:** Системы могут отслеживать, как клиенты перемещаются по магазину, что помогает в оптимизации размещения товаров и улучшении клиентского опыта.
- **Автоматизация кассовых операций:** Использование технологий распознавания для автоматического считывания товаров и упрощения процесса оплаты.

## 6. Индустрия развлечений

- **Виртуальная и дополненная реальность:** Компьютерное зрение используется для отслеживания движений пользователей и создания интерактивного контента.
- **Игры:** В играх технологии компьютерного зрения могут использоваться для распознавания жестов и движений игроков.

## 7. Сельское хозяйство

- **Мониторинг состояния растений:** Использование дронов и камер для анализа состояния посевов, выявления заболеваний и оптимизации процессов орошения.
- **Автоматизация сбора урожая:** Роботы, оснащенные системами компьютерного зрения, могут автоматически собирать урожай, распознавая зрелые плоды.

## 8. Научные исследования

- **Анализ изображений в биологии и экологии:** Компьютерное зрение используется для анализа данных, полученных из микроскопов, камер и других источников, что помогает в исследованиях в области биологии и экологии.

## **9. Образование**

- **Интерактивные обучающие системы:** Использование компьютерного зрения для создания интерактивных учебных материалов и систем, которые могут адаптироваться к потребностям учащихся.

## **10. Спорт**

- **Анализ производительности:** Использование видеоанализа для оценки техники спортсменов и улучшения их результатов.
- **Судейство:** Технологии компьютерного зрения могут использоваться для автоматического определения нарушений и других событий во время соревнований.

## Глава 2. Практика

### 2.1 Процесс создания кода

**Первым шагом** скачиваем все нужные нам библиотеки и данные (в моем случае Yolo4 его datasets и веса и конфигурацию)

**cv2:** библиотека OpenCV, используемая для обработки изображений и видео.

**numpy:** библиотека для работы с многомерными массивами и матрицами, а также для выполнения математических операций.

```
Pip install numpy
```

```
Pip install OpenCV-python
```

**Второй шаг** импортируем библиотеки

```
Import numpy
```

```
Import cv2
```

**Третий шаг:** Загрузка классов объектов

```
with open("coco.names", "r") as f:
```

```
    classes = [line.strip() for line in f.readlines()]
```

**Четвертый шаг:** Загрузка модели YOLO

```
net = cv2.dnn.readNet("yolov4.weights", "yolov4.cfg")
```

**cv2.dnn.readNet:** загружает предобученную модель YOLO, используя веса (**yolov4.weights**) и конфигурационный файл (**yolov4.cfg**).

**Пятый шаг:** Загрузка объектов и проверка загрузки объекта и его размера

```
image_path = "Car1.jpg"
```

```
frame = cv2.imread(image_path)
```

```
if frame is None:
```

```
    print("Не удалось загрузить изображение. Проверьте путь к файлу.")
```

else:

```
new_size = (800, 600) # Новый размер (ширина, высота)
```

```
frame = cv2.resize(frame, new_size)
```

```
height, width, _ = frame.shape
```

**Шестой шаг:** Подготовка кадров для Yolo 4

```
blob = cv2.dnn.blobFromImage(frame, 0.00392, (416, 416), (0, 0, 0), True,  
crop=False)
```

```
net.setInput(blob)
```

- **cv2.dnn.blobFromImage**: преобразует изображение в формат, подходящий для обработки нейронной сетью. Здесь:
- **0.00392**: масштабирование (1/255, чтобы привести значения пикселей к диапазону [0, 1]).
- **(416, 416)**: размер, к которому будет изменено изображение.
- **(0, 0, 0)**: значения среднего цвета (в данном случае не используются).
- **True**: указывает, что изображение будет транспонировано.
- **crop=False**: не обрезать изображение.
- **net.setInput(blob)**: устанавливает подготовленный blob в качестве входных данных для сети.

**Седьмой шаг:** Получение выходных слоев

```
layer_names = net.getLayerNames()
```

```
output_layers = [layer_names[i - 1] for i in net.getUnconnectedOutLayers()]
```

- **net.getLayerNames()**: получает имена всех слоев в модели.
- **net.getUnconnectedOutLayers()**: получает индексы выходных слоев, которые не подключены к другим слоям.
- Мы создаем список **output\_layers**, который содержит имена выходных слоев.

**Восьмой шаг:** строим прямой проход через сеть

```
outs = net.forward(output_layers)
```

- **net.forward(output\_layers):** выполняет прямой проход через сеть и получает выходные данные из выходных слоев.

**Девятый шаг:** Обработка выходных данных

```
boxes = []
```

```
confidences = []
```

```
class_ids = []
```

- Инициализируем списки для хранения ограничивающих рамок (**boxes**), уверенности (**confidences**) и идентификаторов классов (**class\_ids**).

**Десятый шаг:** строим цикл по выходным данным

```
for out in outs:
```

```
    for detection in out:
```

```
        scores = detection[5:]
```

```
        class_id = np.argmax(scores)
```

```
        confidence = scores[class_id]
```

```
        if confidence > 0.5: # Порог уверенности
```

```
            center_x = int(detection[0] * width)
```

```
            center_y = int(detection[1] * height)
```

```
            w = int(detection[2] * width)
```

```
            h = int(detection[3] * height)
```

```
            # Прямоугольник
```

```
            x = int(center_x - w / 2)
```

```
            y = int(center_y - h / 2)
```

```
boxes.append([x, y, w, h])
```

```
confidences.append(float(confidence))
```

```
class_ids.append(class_id)
```

```
import cv2
import numpy as np

# Загрузка классов
with open("coco.names", "r") as f:
    classes = [line.strip() for line in f.readlines()]

# Загрузка модели YOLO
net = cv2.dnn.readNet("yolov4.weights", "yolov4.cfg")

# Укажите путь к вашему изображению
image_path = "1500452276_preview_banana.jpg"
frame = cv2.imread(image_path)

if frame is None:
    print("Не удалось загрузить изображение. Проверьте путь к файлу.")
else:
    # Изменение размера изображения
    new_size = (800, 600) # Новый размер (ширина, высота)
    frame = cv2.resize(frame, new_size)

    height, width, _ = frame.shape

    # Подготовка кадра для YOLO
    blob = cv2.dnn.blobFromImage(frame, 0.00392, (416, 416), (0, 0, 0), True,
crop=False)
    net.setInput(blob)

    # Получение выходных слоев
    layer_names = net.getLayerNames()
    output_layers = [layer_names[i - 1] for i in
net.getUnconnectedOutLayers()]

    # Прямой проход через сеть
    outs = net.forward(output_layers)

    # Обработка выходных данных
    boxes = []
    confidences = []
    class_ids = []

    for out in outs:
        for detection in out:
            scores = detection[5:]
            class_id = np.argmax(scores)
            confidence = scores[class_id]
            if confidence > 0.5: # Порог уверенности
                center_x = int(detection[0] * width)
                center_y = int(detection[1] * height)
                w = int(detection[2] * width)
                h = int(detection[3] * height)

                # Прямоугольник
                x = int(center_x - w / 2)
                y = int(center_y - h / 2)

                boxes.append([x, y, w, h])
                confidences.append(float(confidence))
```

```

class_ids.append(class_id)

# Нанесение ограничивающих рамок
indexes = cv2.dnn.NMSBoxes (boxes, confidences, 0.5, 0.4)

for i in indexes.flatten(): # Используем flatten для получения
одномерного массива
    x, y, w, h = boxes[i]
    label = str(classes[class_ids[i]])
    confidence = confidences[i]
    cv2.rectangle(frame, (x, y), (x + w, y + h), (0, 255, 0), 2) #
Рисуем прямоугольник
    cv2.putText (frame, f"{label} {confidence:.2f}", (x, y - 5),
cv2.FONT_HERSHEY_SIMPLEX, 0.5, (0, 255, 0), 2) # Добавляем текст

# Отображение результата
cv2.imshow("Image", frame)
cv2.waitKey(0)
cv2.destroyAllWindows()

```

## 2.2 Ошибки, возникшие в ходе создания кода

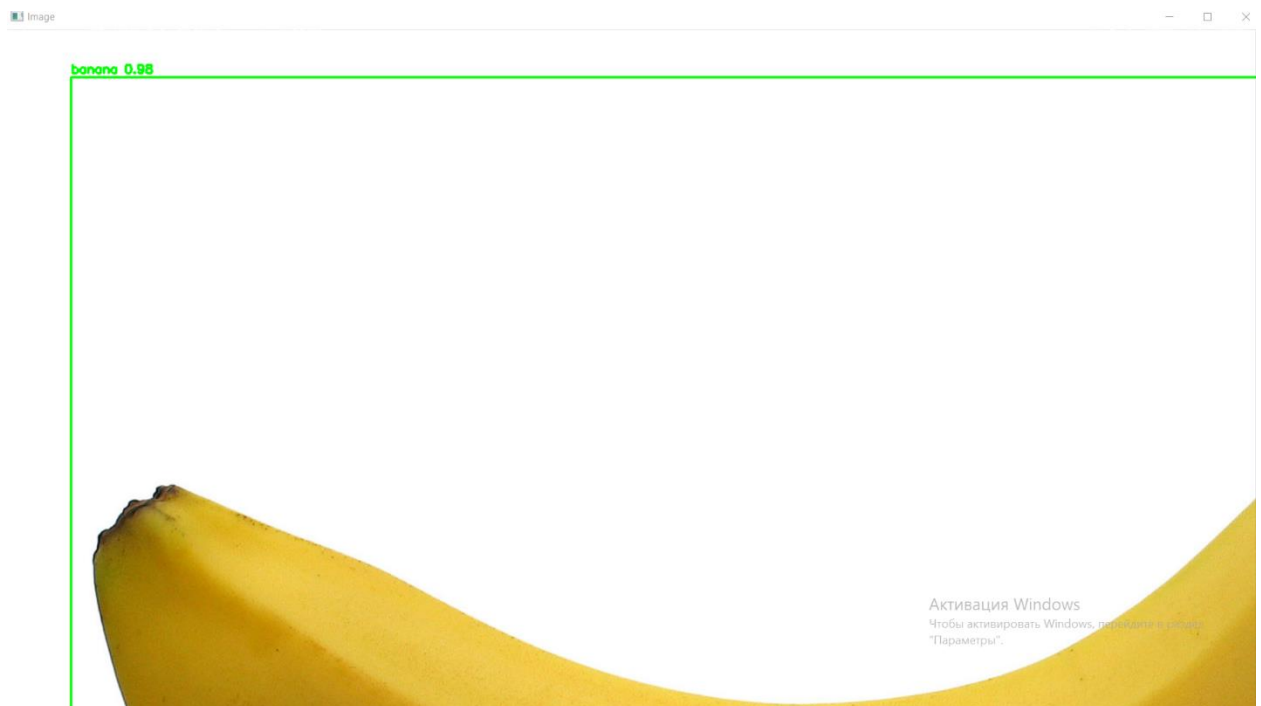
Самая глобальная проблема была в том, что Python не хотел читать библиотеки.

Я решил действовать радикально и переустановил Python и с чистого листа пришлось заново начинать

Вторая проблема была в создании путь к камере на Windows, я хотел, чтобы код работал в режиме прямой трансляции, но из-за защиты Windows и его неадаптивность под процесс создания кода я не смог сделать прямую трансляцию

Третья проблемы возникла в самом конце, она возникла из-за того, что многие файлы не хотят переходить в нужны размер

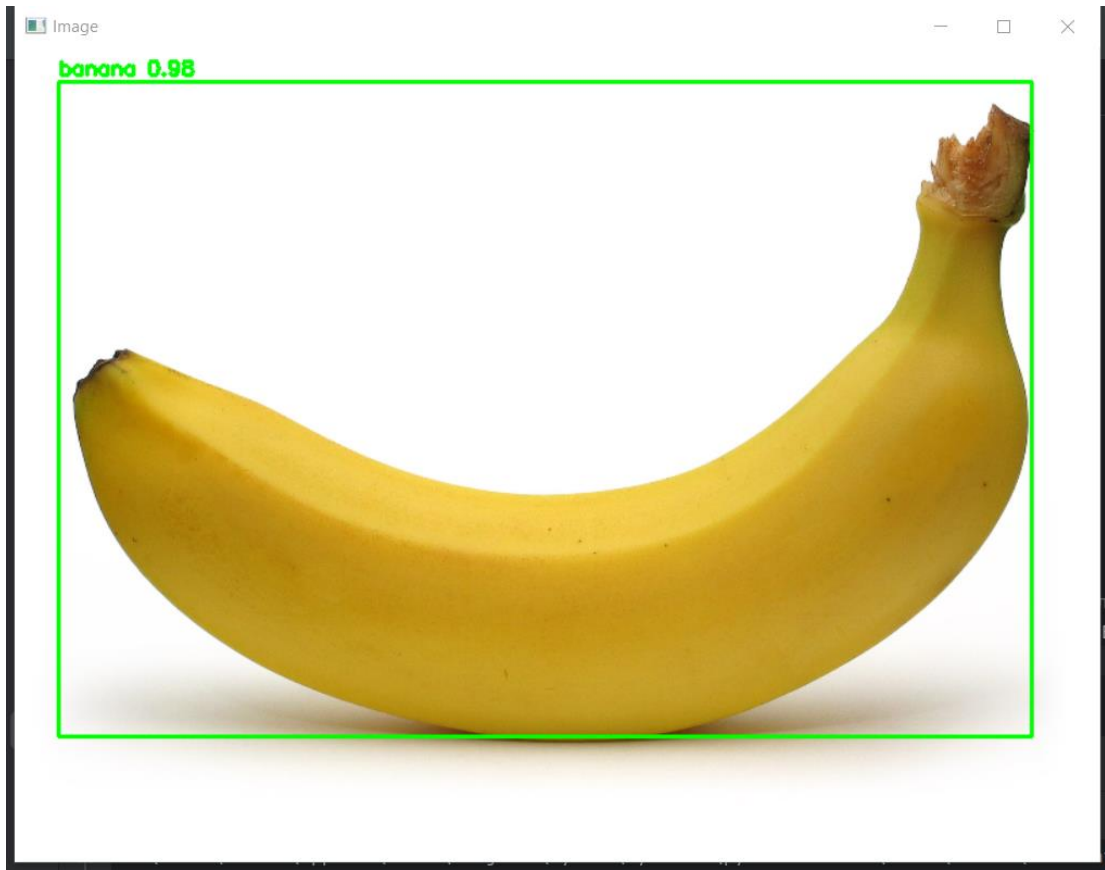
Изображение1.



Но я исправил это, задав новый размер определяемому файлу  
`new_size = (800, 600) # Новый размер (ширина, высота)`  
`frame = cv2.resize(frame, new_size)`

## 2.3 Примеры работы компьютерного зрения

Изображение 1. Попытка 2.



Изображение 2. Попытка 1



## Заключение

Компьютерное зрение — это мощная и быстро развивающаяся область, которая находит применение в самых различных сферах, от медицины и безопасности до автономных транспортных средств и развлечений. Используя алгоритмы и технологии, такие как глубокое обучение и обработка изображений, компьютерное зрение позволяет системам "видеть" и интерпретировать визуальную информацию, что открывает новые возможности для автоматизации и повышения эффективности.

В данном контексте мы рассмотрели, как можно уменьшить размер изображения при его обработке, используя библиотеки, такие как OpenCV. Уменьшение размера изображения не только помогает оптимизировать использование ресурсов, но и улучшает производительность алгоритмов, особенно в задачах, связанных с реальным временем, таких как распознавание объектов.

Применение методов изменения размера и сжатия изображений позволяет эффективно управлять данными, что особенно важно в условиях ограниченных вычислительных ресурсов или при работе с большими объемами визуальной информации. В результате, компьютерное зрение становится неотъемлемой частью современных технологий, способствуя инновациям и улучшая качество жизни в различных аспектах.

С учетом постоянного развития технологий и алгоритмов, можно ожидать, что в будущем компьютерное зрение будет продолжать расширять свои возможности и находить новые применения, что сделает его еще более важным инструментом в различных отраслях.