

ПРОГРАММНОЕ СРЕДСТВО РАСПОЗНОВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ И ВИДЕО

Синевич И.В., студент

*Белорусский государственный университет информатики и радиоэлектроники,
Институт информационных технологий,
г. Минск, Республика Беларусь*

Сицко А. Л. – канд. техн. наук, доцент каф. ИСиТ

В данной статье рассмотрена актуальность разработки ПС для распознавания объектов. Рассмотрены основные проблемы, а также предложены пути упрощения их решения. Определены основные концепции создания подобного ПС, а также представлены основные результаты использования такого инструмента.

В современном мире все большую роль играют технологии искусственного интеллекта, особенно в области компьютерного зрения. Компьютерное зрение — это наука, изучающая способы получения, обработки и анализа изображений с целью выделения полезной информации. Одной из важных задач компьютерного зрения является распознавание объектов на изображениях, видео и видео с камеры. Эта задача заключается в том, чтобы определить, какие объекты присутствуют на изображении, где они находятся и как они взаимодействуют друг с другом. Распознавание объектов имеет множество практических применений, таких как безопасность, медицина, робототехника, автоматизация, развлечения и так далее.

Основные задачи компьютерного зрения включают в себя:

Распознавание объектов: системы компьютерного зрения могут быть обучены распознавать и классифицировать объекты на изображениях или видео. Это может включать в себя распознавание лиц, автомобилей, животных, предметов и многого другого.

Сегментация изображений: эта задача заключается в выделении конкретных областей изображения, таких как объекты или фон. Сегментация позволяет более детально анализировать структуру изображения.

Определение движения: компьютерное зрение используется для отслеживания движущихся объектов на видео и анализа их траекторий.

Оценка глубины: системы компьютерного зрения могут стремиться определить глубину объектов на изображении, что полезно, например, для создания эффекта трехмерности.

Распознавание жестов и выражений лица: компьютеры могут быть обучены распознавать жесты рук или выражения лиц, что полезно для взаимодействия с людьми или создания систем управления с использованием жестов.

Распознавание текста: эта задача включает в себя извлечение текста из изображений, что может быть полезно для автоматического распознавания документов или надписей на изображениях.

Для решения этих задач используются различные методы, включая классические компьютерное зрение, глубокое обучение (нейронные сети) и комбинации этих подходов.

Глубокое обучение, особенно с использованием сверточных нейронных сетей, стало ключевым элементом в развитии компьютерного зрения, позволяя системам автоматически извлекать признаки из визуальных данных и обучаться на больших объемах информации.

В современном информационном обществе графические данные становятся неотъемлемой частью цифрового контента, представляя собой важный источник информации. Задачи распознавания объектов на изображениях, в видео и с камер являются ключевыми в области компьютерного зрения и машинного обучения. Эти задачи имеют применение в различных областях, начиная от безопасности и медицины, заканчивая автоматизированным управлением и технологиями виртуальной реальности.

Суть распознавания в том, чтобы анализировать и классифицировать объекты на изображениях и видео по различным критериям.

Для выполнения работы были использованы следующие методы исследования:

аналитический метод – для изучения теоретических основ распознавания образов, нейронных сетей, машинного обучения и компьютерного зрения;

сравнительный метод – для обзора и сравнения существующих методов и программ для распознавания объектов на изображениях и видео;

экспериментальный метод – для разработки и тестирования программы на Python, которая с помощью нейросети распознает объекты на изображениях и видео;

статистический метод – для обработки и анализа данных, полученных в ходе тестирования программы.

Научная значимость работы заключается в том, что она расширяет и углубляет знания в области распознавания образов, нейронных сетей, машинного обучения и компьютерного зрения, а также предлагает новый подход к решению задачи распознавания объектов на изображениях и видео с помощью программы на Python и нейросети.

Практическая значимость работы заключается в том, что разработанная программа на Python, которая с помощью нейросети распознает объекты на изображениях и видео, может быть использована в различных областях и сферах деятельности, где требуется анализ и классификация изображений и видео.

Для распознавания объектов часто используются нейронные сети — математические модели, имитирующие работу биологических нейронов. Нейронные сети способны обучаться на больших объемах данных и адаптироваться к различным условиям. Одним из видов нейронных сетей, которые хорошо подходят для распознавания объектов и которые будут использованы в данном дипломном проекте, являются сверточные нейронные сети, которые в свою очередь состоят из нескольких слоев, каждый из которых применяет некоторую операцию к входным данным, например, свертку, пулинг, активацию, нормализацию и другие [1].

Свертка — это операция, которая применяет маленькую матрицу, называемую фильтром или ядром, к каждой части входного изображения или другого сигнала, чтобы получить выходное изображение или сигнал, называемый картой активации. Свертка позволяет извлекать разные характеристики входного сигнала, такие как ребра, углы, цвета и так далее. Свертка может повторяться несколько раз с разными фильтрами, чтобы получить несколько карт активации на разных уровнях детализации [2].

Пулинг — это операция, которая уменьшает размер карты активации, выбирая из каждой части карты активации одно значение, которое описывает эту часть. Например, можно выбирать максимальное, минимальное или среднее значение. Пулинг позволяет уменьшить количество параметров и вычислений в сети, а также сделать сеть более устойчивой к маленьким изменениям входного сигнала [3].

Для эффективного решения задач распознавания объектов выбор соответствующих инструментов и технологий играет решающую роль. В данном исследовании, для разработки программного средства по распознаванию объектов, был выбран язык программирования Python и ряд современных фреймворков, включая Numpy, SciPy, OpenCV, Pillow, Matplotlib, H5py, Keras и ImageAI. Решение о выборе данных фреймворков обусловлено их высокой эффективностью в работе с графическими данными, а также широкой распространенностью.

Популярность программ распознавания объектов растёт с каждым днём. Уже ни для кого не секрет, что хорошее приложение приносит хорошую прибыль. Оно на порядок выгодней, так как нет

затрат на зарплату для рабочих, которые бы занимались данными вещами. Одним из преимуществ предлагаемой разработки является простой запуск, высокая скорость обработки информации и отсутствие ошибок, которые может допустить сотрудник при работе с большим количеством данных.

Список использованных источников:

- [1] Goodfellow, I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – Cambridge : MIT Press, 2016. – 800 с.
- [2] Оссовский С. Нейронные сети для обработки информации; [пер. с польск. И.Д. Рудинского]. М.: Финансы и статистика, 2004. 344 с.
- [3] Raschka, S. Python Machine Learning / S. Raschka, V. Mirjalili. – Birmingham : Packt Publishing, 2017. – 622 с.