

РАСПОЗНАВАНИЕ ОБРАЗОВ ДЛЯ ЛЮДЕЙ С НАРУШЕНИЯМИ ЗРЕНИЯ

Богдан Д. И.

Научный руководитель - д.т.н., доц. Машталир С. В.

Харьковский национальный университет радиоэлектроники
(61166, Харьков, просп. Науки, 14, каф. Информатики, тел. (057) 702-14-19)
e-mail: denys.bohdan@nure.ua, тел. (097) 714-83-67.

This work is devoted to applying deep learning for object detection. The task was to create a mobile object detector application for the visually impaired using a convolutional neural network (CNN). For this research, Keras Python library with TensorFlow library was used for object detection. As a result, an API with a convolutional neural network which detects presence of indoors objects on an input image and a mobile application that communicates with the API were created. As a future improvement the trained CNN model can be used directly in the application to support offline mode.

В связи с активным развитием и распространением информационных технологий и использованием камер в мобильных устройствах возрастает необходимость в классификации и нахождении объектов на полученных изображениях. Для человека обычно не является затруднительным распознать объекты, которые он видит. Также человек может определить размеры объектов, их расположение относительно других объектов и окружающего мира и т. д. Но для людей с ограниченными возможностями зрения даже такие базовые задачи являются затруднительными или невыполнимыми. По данным Всемирной организации здравоохранения около 1 млрд. человек имеют различные формы нарушения зрения и от 40 до 65 млн. вовсе не могут видеть. Это все люди, жизнь которых можно сделать проще благодаря последним достижениям в науке и технологии, и в частности - благодаря компьютерному зрению.

В последнее время методы, решающие задачи компьютерного зрения, основаны на нейронных сетях (НС), в частности - сверточных нейронных сетях.

Сверточная нейронная сеть (СНС) - архитектура нейронных сетей, созданная специально для имитации зрительной коры головного мозга. СНС состоит из серии слоев. Входное изображение пропускается через набор сверточных слоев и с помощью полносвязного слоя порождается вывод (класс или вектор вероятностей классов, описывающих изображение).

Сверточный слой представляет собой набор фильтров (матриц), проводящих операцию свертки (поэлементное умножения пикселей изображения на элементы матрицы фильтра). Результатами работы сверточных слоев является набор деталей, от более абстрактных (кривые) к более конкретным (геометрические фигуры).

Наиболее популярными библиотеками для реализации НС являются Keras и TensorFlow. Keras - высокоуровневая библиотека для работы с НС на уровне моделей. Для манипуляции над тензорами и дифференцирования используется библиотека поддержки тензоров TensorFlow.

Для реализации СНС для приложения была выбрана модель MobileNet из библиотеки Keras, специально разработанная компанией Google для использования на устройствах, ограниченных в ресурсах (например, мобильные телефоны).

Этапы работы:

1. Выбор классов изображений. Для данной работы были выбраны следующие классы: дверь, лестница, человек, автомобиль.
2. Подготовка изображений для классификации: разбиваем их на три каталога: изображения для обучения, проверки и тестирования.
3. Создание генератора изображений, используя метод ImageDataGenerator (). Этот генератор произведет аугментацию данных (увеличит количество данных при помощи переноса, вращения, применения фильтров и т. д.). Для каждого каталога необходимо создать свой генератор.
4. Обучение сети при помощи метода fit_generator ().
5. Анализ результатов обучения. При необходимости дообучаем сеть.
6. Создание серверного приложения, которое передает изображение в СНС и возвращает результат классификации.
7. Создание мобильного приложения, которое проводит съемку, отправляет изображения на серверное приложение и отображает результаты классификации.

Недостатками такого подхода является необходимость интернет-подключения у мобильного устройства и использование большого количества интернет трафика для передачи изображений. В дальнейшем планируется избавление от серверного приложения и проведение классификации непосредственно на мобильном устройстве, а также добавление других классов изображений.

Список використаних джерел:

1. Романов А. А. Сверточные нейронные сети [Электронный ресурс]. – 2018. Режим доступа: <https://scientificresearch.ru/images/PDF/2018/21/svertochnye.pdf>, свободный.
2. Samer Hijazi, Rishi Kumar, and Chris Rowen. Using Convolutional Neural Networks for Image Recognition [Электронный ресурс]. – 2018. Режим доступа: https://ip.cadence.com/uploads/901/cnn_wp-pdf, свободный.