

УДК 621.391

ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ ПРИ ПОМОЩИ ФРЕЙМВОРКА TENSORFLOW

А.Л. КАЛОША, А.В. САСКЕВИЧ, М.В. СТЕРЖАНОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 01 ноября 2019*

Аннотация. Разработана системы для локализации и классификации объектов на изображениях.

Ключевые слова: TensorFlow, CUDA, нейронная сеть, машинное обучение.

Введение

Человеческий мозг хранит множество загадок. С появлением современной электроники, начались попытки аппаратно-программного воспроизведения его работы. Бурное развитие и применение аппарата искусственных нейронных сетей пришло на последние полвека. На текущий момент задача обнаружения и классификации объектов на изображениях достаточно актуальна. Повсеместно ведется разработка программного обеспечения для автоматического управления транспортными средствами, валидации сканируемых документов и т.д. Другими словами, разрабатываются алгоритмы и программы компьютерного зрения [1].

Основная часть

Правильно обученная нейронная сеть позволяет с высокой точностью найти на изображении объекты и определить их классы.

Нейронная сеть – это громадный распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющих их для последующей обработки [2].

Для решения данной задачи использовалась сверточная нейронная сеть, которая относится к нейронным сетям глубокого обучения.

Под термином глубокое обучение часто подразумевают такие модели, как рекуррентные нейронные сети, сети с долгой краткосрочной памятью, рекуррентные тензорные нейронные сети, сверточные нейронные сети и генеративно-состязательные сети, получившие большую популярность в последние годы [3].

Часто под глубокими нейронными сетями подразумевают нейронные сети с несколькими внутренними слоями, термин «глубокое обучение» в действительности мало чем отличается от искусственных нейронных сетей. Однако, разные архитектуры реализуют уникальные возможности в слоях, что дает им возможность моделировать очень сложные данные [3].

Сверточные нейронные сети (Convolutional Neural Networks, CNN), например, объединяют многослойные перцептроны со сверточным слоем, который итеративно строит карту для извлечения важных признаков, а также имеют этап свертки, уменьшающий размерность признаков, но сохраняющий наиболее информативные их компоненты. Сверточные сети очень эффективны для моделирования изображений и решения задач классификации и обобщения [3].

Для обучения нейронной сети была выбрана библиотека TensorFlow как один из лучших инструментов машинного обучения. TensorFlow – это библиотека программного обеспечения с открытым исходным кодом для численного расчета с использованием графиков потока данных [4].

Для ускорения процесса обучения используется вычислительная мощность видеокарты, а именно технология CUDA. CUDA – это архитектура параллельных вычислений от NVIDIA, позволяющая существенно увеличить вычислительную производительность благодаря использованию GPU (графических процессоров) [5].

В данной работе используется Tensorflow Object Detection API, который упрощает процесс обучения нейронной сети. Например, предоставляет несколько недообученных моделей для обнаружения объектов с разной производительностью и точностью, основанных на датасете COCO (Common Objects in Context). В данной работе используется модель SSD-MobileNet, основанная на алгоритме SSD (Single Shot Detector). Использование данных моделей позволяет сократить время обучения нейронной сети в несколько раз.

Чтобы обучить нейронную сеть новым классам, которые не содержатся в базовой модели, нужно использовать инструмент LabelImg, который упрощает разметку искомым объектам на изображениях. Файл разметки элемента обучающей выборки содержит такие данные как: имя файла изображения, его размер, количество цветов, имя класса, координаты верхнего левого и нижнего правого прямоугольника, содержащего объект (рис. 1).

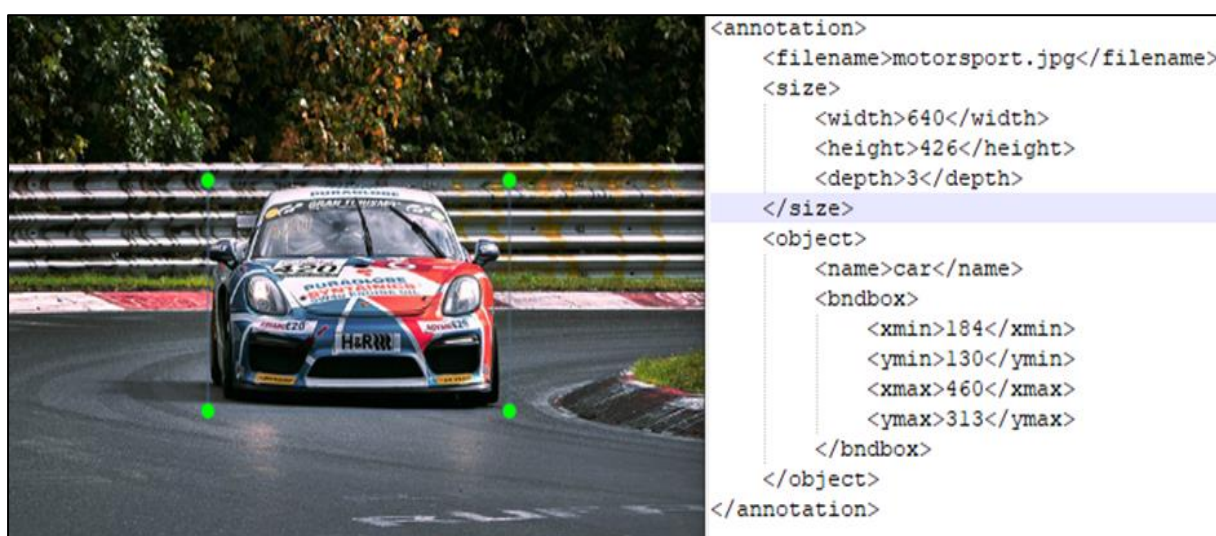


Рис. 1. Пример эталонной разметки изображений обучающей и тестовой выборки

Чтение данных является узким местом обучения. Частично эту проблему решает новый формат данных TFRecord, который предназначен для хранения последовательности двоичных записей. Данный формат позволяет легко комбинировать несколько наборов данных и интегрируется с функциями импорта и предварительной обработки данных.

Для обучения нейронной сети должны использоваться как позитивные, так и негативные примеры областей на изображении. Как правило, негативных примеров требуется в несколько раз больше. При обучении нейронной сети с использованием Tensorflow API можно не указывать негативные примеры. В таком случае все области на изображении, не являющимися позитивными, по умолчанию считаются негативными. Данная возможность позволяет сократить время на подготовку данных для обучения.

Реализация метода SSD осуществлена на языке программирования Python 3 с использованием библиотек глубокого обучения TensorFlow и Keras, а также технологии NVIDIA CUDA.

Обучение и тестирование модели SSD было произведено на компьютере со следующими характеристиками: процессор Intel Core i5-8600; оперативная память 16,0 ГБ; видеокарта Nvidia GeForce GTX 2060, частота графического процессора 1365 МГц, 6144 Мб видеопамати GDDR6, частота видеопамати 3500 МГц, число универсальных процессоров 1920.

Результат работы приложения представлен на рис. 2.

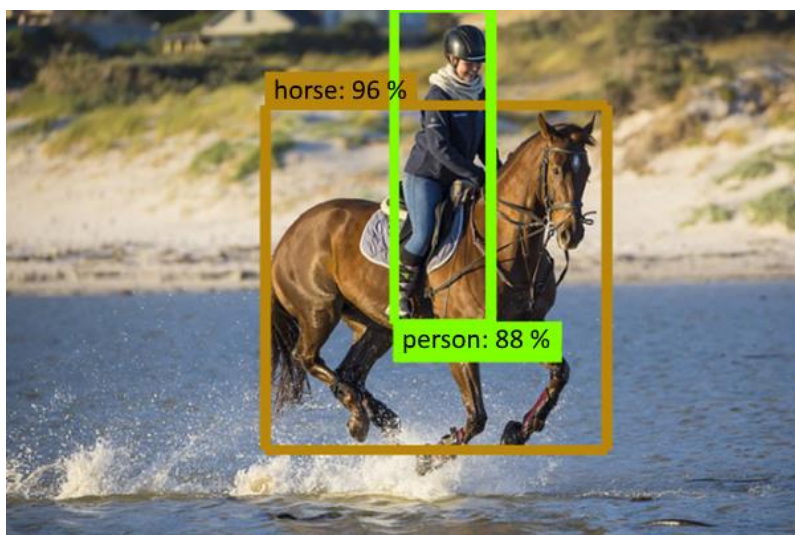


Рис. 2. Результат работы приложения

Заключение

Обучение сети производилось на более чем 3000 изображений, что занимает от 3 до 6 часов, в зависимости от глубины обучения и точности результата. В результате система способна локализовать и классифицировать объекты на изображении. Средняя точность классификации при оптимальных условиях составляет 80 %. Для увеличения точности классификации можно использовать другой тип модели или увеличить выборку для обучения.

OBJECTS DETECTION AND RECOGNITION ON IMAGES USING TENSORFLOW

A.L. KALOSHA, A.V. SASKEVICH, M.V. STERJANOV

Abstract. Software for object detection and recognition on images using TensorFlow was developed.

Keywords: TensorFlow, CUDA, neural network, machine learning.

Список литературы

1. Ситкин И.В. // Научное сообщество студентов XXI столетия. Технические науки: сб. ст. по мат. LXIV междунар. студ. науч. – практ. конф. № 4(63). [Электронный ресурс] URL: [https://sibac.info/archive/technic/4\(63\).pdf](https://sibac.info/archive/technic/4(63).pdf). (дата доступа: 30.10.2019).
2. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006.
3. Бенгфорт Б., Билбро Р., Охеда Т. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. СПб.: Питер, 2019.
4. Library for numerical computation using data flow graphs [Электронный ресурс] URL: <https://www.tensorflow.org>. (дата доступа: 30.10.2019).
5. Параллельные вычисления CUDA [Электронный ресурс] URL: <http://www.nvidia.ru/object/cuda-parallel-computingru.html>. (дата доступа: 30.10.2019).