

УДК 004.932.2

ОБНАРУЖЕНИЕ И ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Д.Ю. ТУЧА

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 13 марта 2019*

Аннотация. Предложено программное обеспечение, разработанное на языке Python 3 для обнаружения и детектирования объектов на кадрах видеоизображения на основе алгоритма SSD (от англ. Single Short Derector) и библиотеки TensorFlow. Проведена практическая оценка предложенного программного обеспечения.

Ключевые слова: обнаружение объекта, видеоизображение, машинное обучение, нейронная сеть.

Введение

В последнее время широкое распространение робототехники, систем автоматического управления, видеонаблюдения, дополнительной реальности и других систем, связанных с распознаванием объектов на изображении, вызвало обширные исследования в области компьютерного зрения. Благодаря значительному развитию нейронных сетей, особенно глубокого обучения, системы визуального распознавания достигли значительных результатов. В работе предложена программная реализация метода обнаружения объектов с помощью глубокого обучения на основе языка программирования Python 3.

Машинное обучение

Программа реализована на основе открытой программной библиотеки для машинного обучения TensorFlow. В библиотеке вычисления описываются с использованием графов потока данных (рис. 1).

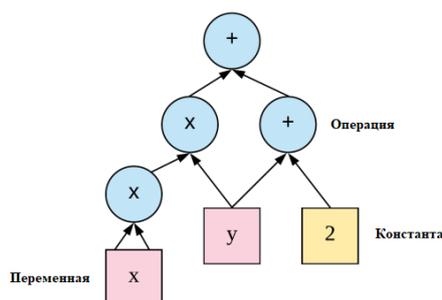


Рис. 1. Граф потока данных

Каждый узел графа представляет собой математическую операцию (например, сложение, деление или умножение), а каждое ребро – многомерный набор данных (тензор), над которым выполняются операции [1]. Тензор представляет собой многомерную матрицу, заполненную числами – компонентами тензора.

Модель глубокого обучения для обнаружения объектов можно описать тремя шагами:

- сбор данных для обучения;
- обучение модели;
- предсказание на новых изображениях.

Применение метода Single Shot Detector

Для обнаружения объектов на видеоизображении используется алгоритм SSD. Обнаружение состоит из извлечения карты признаков и применения свертки для обнаружения объектов. На вход подается изображение размером 300×300 пикселей. К изображению применяются сверточные слои из стандартной модели VGG-16. Далее к каждому выходному слою добавляются специальные сверточные слои, представляющие собой изображение в разных масштабах. Пространственная размерность убывает до тех пор, пока не станет равной единице. Каждый из сверточных слоев позволяет сформировать карту признаков для разных масштабов изображения, например 8×8 (рис. 2) и 4×4 (рис. 3). Далее для карты признаков определяется, какой из ограничивающих прямоугольников в области 3×3 лучше всего совпадает с эталонной разметкой. Все эти карты объединяются в единый выходной слой, в которых может находиться объект. Итоговые области выбираются из прямоугольников с помощью метода подавления немаксимумов [2].

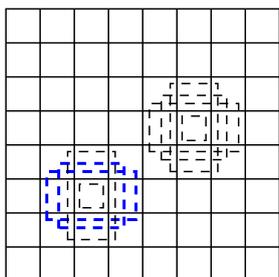


Рис. 2. Карта признаков для масштаба изображения 8×8

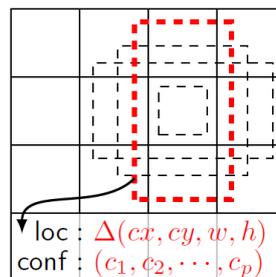


Рис. 3. Карта признаков для масштаба изображения 4×4

Оценка эффективности разработанной программы

Качество обнаружения транспортных средств описывают такие метрики, как полнота R и точность P обнаружения объектов [3]. Полнота R показывает чувствительность алгоритма к ошибкам 2-го рода, то есть, пропускам, и равна отношению количества правильно найденных объектов к общему количеству этих объектов в эталонной разметке:

$$R = \frac{tp}{tp + fn},$$

где tp – истинно-положительные – объекты, которые ожидалось увидеть и которые были получены на выходе, fn – ложноотрицательные объекты, которые ожидалось увидеть, но алгоритм их не определил.

Точность P показывает чувствительность алгоритма к ошибкам 1-го рода, то есть ложным срабатываниям, и равна отношению количества правильно найденных объектов к общему количеству найденных алгоритмом прямоугольников:

$$P = \frac{tp}{tp + fp},$$

где fp – ложноположительные объекты, которых на выходе не должно быть, но алгоритм их ошибочно вернул на выходе.

Оценка эффективности разработанной программы проводилась на кадрах изображений, полученных с автомобильного видеорежистратора (рис. 4).

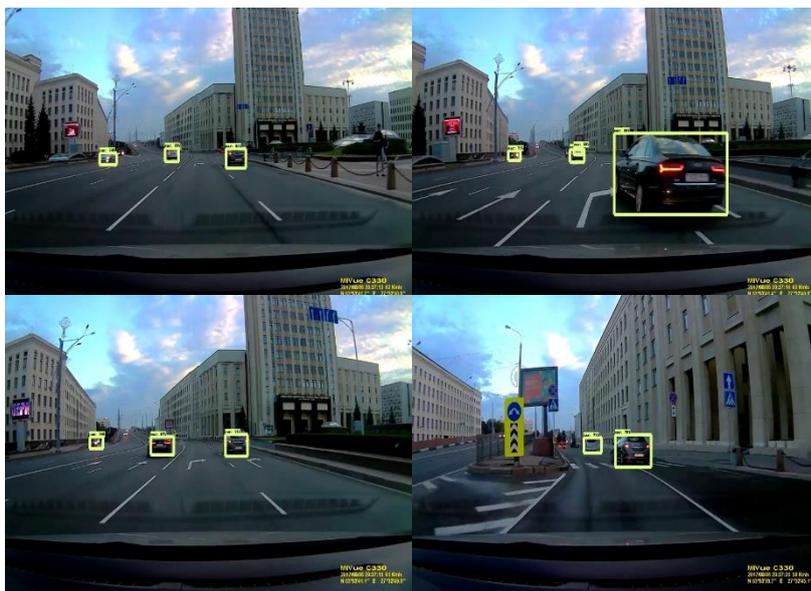


Рис. 4. Примеры обнаружения транспортных средств

В первом тесте использовалась уже обученная нейронная сеть на выборке `ssd_mobilenet_v1_coco`. Второе тестирование проводилось на дополнительно обученной выборке из первого теста. Полученные результаты представлены в таблице.

Результаты обнаружения транспортных средств на готовой выборке и обученной выборке

Номер теста	Полнота	Точность
Тест 1	0,464	0,890
Тест 2	0,993	0,989

Заключение

Разработано программное обеспечение на языке Python 3, которое позволяет обнаружить и детектировать объекты на видеоизображении. На основе результатов его испытаний установлено, что чем больше различных изображений объектов используется при обучении нейронной сети, тем выше результаты качества распознавания. Имеется возможность улучшить производительность работы программного обеспечения путем перехода на графические процессоры с использованием технологии nVidia CUDA.

OBJECTS DETECTION AND RECOGNITION ON VIDEOIMAGES USING MACHINE LEARNING

D.Yu. TUCHA

Abstract. Software for object detection and recognition on video images using algorithm SSD (Single Shot Detector) and open source software library TensorFlow is proposed.

Keywords: detection, videoimage, machine learning, neural network.

Список литературы

1. Ramsundar B., Zadeh R.B. // O' Reilly Media Publishers. 2018. Vol. 2. P. 19–42.
2. SSD: Single Shot MultiBox Detector. [Электронный ресурс]. URL: <https://arxiv.org> (дата обращения: 15.03.2019).
3. Everingham M. [et. al.] // Int. J. of Computer Vision. 2010. Vol. 88. P. 303–338.