

ДЕТЕКТИРОВАНИЕ И СЕГМЕНТАЦИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Степанов П.А.,

студент 4 курса ВГУ имени П.М. Машерова, г. Витебск, Республика Беларусь

Научный руководитель – Кухарев А.В., канд. физ.-мат. наук

В области компьютерного зрения основными задачами являются классификация, детектирование и сегментация объектов на изображениях. Классификация – это определение класса, к которому относится объект. Детектирование подразумевает нахождение координат прямоугольных рамок, выделяющих на изображении требуемые объекты, а сегментация – это установление, к какому классу относится каждый пиксель изображения.

Наилучших результатов в решении перечисленных задач в настоящее время достигли сверточные нейронные сети. Это разновидность многослойных искусственных нейронных сетей на основе операции свертки, которые были предложены в работе Яна Лекуна в 1998 г. [1]. Благодаря операции свертки, которая работает как фильтр, сверточный слой способен выделять признаки во входном изображении и передавать информацию о них на последующие слои.

Цель настоящей работы – разработать программную реализацию детектора объектов на изображениях на основе сверточной нейронной сети.

Материал и методы. Для сравнительного анализа нейросетевых моделей, пригодных для построения детектора объектов, были выбраны архитектуры YOLO («You Only Look Once»), SSD («Single Shot MultiBox Detector»), а также сверточные нейронные сети на основе регионов: R-CNN, Fast R-CNN, Faster R-CNN и Mask R-CNN [2].

Для программной реализации детектора использовался язык программирования Python и библиотеки глубокого обучения TensorFlow и Keras.

Для обучения нейронной сети использовался автоматически сгенерированный набор изображений, на которых случайным образом размещалось три вида геометрических фигур: квадраты, круги и треугольники. Обучение модели проводилось в два этапа. На первом этапе обучались только последние слои нейронной сети с «заморозкой» низлежащих слоёв. На втором этапе проводилось дообучение всех слоёв сети с уменьшенным в 10 раз коэффициентом обучения.

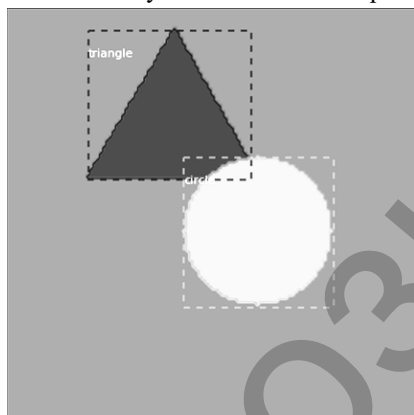


Рисунок 1 – Пример геометрических фигур для детектирования

Результаты и их обсуждение. Из рассмотренных моделей нейронных сетей архитектуры Faster R-CNN и Mask R-CNN обладают наилучшей точностью детектирования объектов на изображениях, однако они уступают моделям YOLO и SSD по времени работы. Основным отличием модели Mask R-CNN от Faster R-CNN является наличие дополнительных слоёв для выполнения сегментации объектов.

На рисунке 1 показан пример детектирования геометрических фигур на изображении с помощью разработанного детектора на основе нейросетевой модели Mask R-CNN. На тестовом изображении все фигуры были обнаружены и безошибочно обведены рамками (штрих-пунктирные линии).

Заключение. Разработана программная реализация детектора объектов на изображениях на основе нейросетевой модели Mask R-CNN. На простейшем датасете из трех классов детектор показывает отличное качество работы как в задаче обнаружения объектов, так и их сегментации. В дальнейшем предполагается

использовать разработанную программную систему для детектирования более сложных объектов, таких как дорожные знаки.

1. LeCun, Y. Gradient-based learning applied to document recognition / Y. LeCun [et al.] // Proc. of the IEEE. – 1998. – Vol. 86. – P. 1–46.
2. He, K. Mask R-CNN / K. He [et al.] // arXiv.org [Electronic resource]. – Mode of access : <https://arxiv.org/abs/1703.06870>. – Date of access : 25.02.2019.