

ПЛЕНАРНЫЕ ДОКЛАДЫ

В.А. Головки, А.А. Крошенко
Беларусь, Брест. БрГТУ

НЕЙРОСЕТЕВЫЕ МОДЕЛИ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Задача обнаружения объектов (детекция) на изображениях заключается в выделении отдельных частей изображения, принадлежащих некоторым заранее определенным классам. Модель, осуществляющая подобную операцию, принимает на вход изображение, а на выходе возвращает координаты и размеры прямоугольных областей, включающих искомые объекты, а также вероятность принадлежности заключенного в них объекта заданному классу.

Решение подобной задачи – актуальная тема области компьютерного зрения. Благодаря моделям-детекторам, можно осуществлять анализ фото- и видеоизображений в реальном времени, размещая метки на определенных объектах и осуществляя предопределенные операции обработки.

Задачу обнаружения объекта можно логически разделить на две подзадачи – локализация объекта и его классификация. Существующие в настоящее время подходы к обнаружению объектов на изображениях позволяют объединить эти два разрозненных этапа в одной нейронной сети, которая выполняет обе задачи одновременно и формирует итоговый результат на выходе (рисунок). Это ускоряет анализ и получение результатов, включающих все найденные объекты, без необходимости их последовательной обработки.

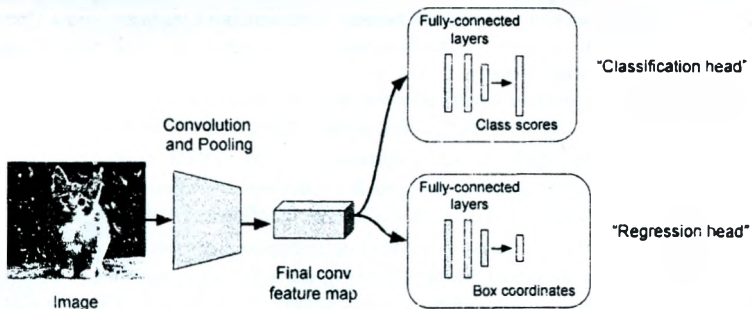


Рисунок – Общий вид нейросетевой модели, применяемой для решения задачи обнаружения объектов на изображении [1]

Существует две категории методов, применяемых для решения задачи обнаружения объектов:

1. Методы с выделением кандидатов (R-CNN, Fast R-CNN, Faster R-CNN).
2. Однопроходные методы (one-look), к которым относятся SSD, YOLO, YOLO9000.

Основное отличие первой категории от второй – то, что для методов первой категории процесс обнаружения объектов делится на два четких этапа: 1) локализация регионов-претендентов; 2) классификация обнаруженных регионов.

При осуществлении локализации может быть выделено большое количество регионов, не все из которых содержат искомые объекты. В этом случае часть областей отсеивается.

R-CNN (Region-based Convolutional Neural Network). R-CNN базируется на идее метода с предварительным выделением претендентов [2]. Вначале осуществляется анализ изображения с локализацией регионов, предположительно являющихся объектами (всего выделяется около 2 000 прямоугольных областей). Затем такие регионы подаются на вход нейронной сети, которая делает вывод о принадлежности объекта определенному классу.

Fast R-CNN. Метод Fast R-CNN [3] представляет собой развитие идей классического метода R-CNN. В этой архитектуре регионы-претенденты проходят через т. н. слой region of interest pooling (ROI), после чего формируется набор карт признаков фиксированного размера, которые затем подаются на полносвязный слой.

Слой ROI pooling фактически действует по принципу подвыборочных (pooling) слоев, отображая проскции интересующих нас областей в области фиксированного (меньшего) размера.

Подобная архитектура позволяет уменьшить количество генерируемых блоков и, как следствие, ускорить работу сети.

Yolo (You-only-look-once). Сеть Yolo [4] представляет собой первую предствительницу группы моделей, позволяющих осуществлять локализацию и распознавание в составе одной единственной модели.

Эта архитектура чаще всего базируется на предобученной сверточной нейронной сети (например, VGG16).

Предобученная сверточная сеть интегрируется в модель (без своих классифицирующих слоев) и дообучается задачам локализации и распознавания. При этом классы, которые использовались для обучения сверточной нейронной сети, и классы, объекты которых нужно обнаруживать посредством модели, могут не совпадать. Предобученная сеть позволяет начать дообучение Yolo-сети с меньшей ошибкой.

Алгоритм Yolo осуществляет отображение исходного изображения на решетку размера $S \times S$, где S – размерность карты признаков, и для каждой ячейки этой решетки осуществляет прогнозирование параметров B элементарных прямоугольных областей (регионов, боксов), степень доверия к этим областям (confidence) и метки класса C . Таким образом, такие данные могут быть представлены тензором размерности $S \times S \times (B * 5 + C)$.

SSD (Single-shot detector). Модель SSD [5], как и YOLO, принадлежит к категории однопроходных методов, позволяющих решать задачу обнаружения объектов в рамках одной единственной сети.

Отметим основные особенности этой модели.

1. Отличается от других single-shot-детекторов (в частности, от YOLO) тем, что каждый слой модели участвует в формировании информации об объектах и их расположении (при этом учитывается масштаб этих объектов – каждый последующий слой обрабатывает объекты большего размера, чем предыдущий).

2. В качестве базового элемента используется предобученная сеть (VGG или ResNet), которая преобразуется к полностью сверточной НС (FCN).

3. В процессе работы сети используется алгоритм Non-maxima suppression для уменьшения количества боксов.

4. Каждый элемент карты признаков формирует набор т. н. default boxes (или Anchors), отличающихся по масштабу и соотношению сторон.

5. Модель обучается, чтобы для каждого anchor правильно прогнозировать его класс и смещение.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Object Localization and detection [Electronic resource]. – Mode of access: https://leonardoraujosantos.gitbooks.io/artificial-intelligence/content/object_localization_and_detection.html. – Date of access: 02.10.2018.

2. Rich feature hierarchies for accurate object detection and semantic segmentation [Electronic resource] / R. Girshick [et al.] // arXiv. – Mode of access: <https://arxiv.org/pdf/1311.2524v5.pdf>. – Date of access: 02.10.2018.

3. Girshick, R. Fast R-CNN [Electronic resource] / R. Girshick // arXiv. – Mode of access: <https://arxiv.org/pdf/1504.08083.pdf>. – Date of access: 02.10.2018.

4. You Only Look Once: Unified, Real-Time Object Detection [Electronic resource] / J. Redmon [et al.] // arXiv. – 2016. – Mode of access: <https://arxiv.org/pdf/1506.02640.pdf> – Date of access: 02.10.2018.

5. SSD: Single Shot MultiBox Detector / W. Liu [et al.] // arXiv [Electronic resource]. – 2016. – Mode of access: <https://arxiv.org/pdf/1512.02325.pdf>. – Date of access: 02.10.2018.

Т. С. Силюк, О. В. Матысик

Беларусь, Брест, БрГУ имени А.С. Пушкина

ПРОЕКТ МТП «СОЗДАНИЕ ТРАНСГРАНИЧНОЙ ПЛАТФОРМЫ «БИЗНЕСТРАНС»»

Программа трансграничного сотрудничества Польша – Беларусь – Украина на 2007–2013 гг. реализуется в рамках Европейского инструмента соседства и партнерства, который является инициативой Европейской комиссии, с целью развития сотрудничества Европейского союза и партнерских государств для обеспечения интегрированного и устойчивого регионального развития. 1 декабря 2013 г. *партнеры проекта* – Высшая государственная школа имени Папы Яна Павла II в Бялой Подляске и Брестский государственный университет имени А.С. Пушкина – начали реализацию проекта «Создание трансграничной платформы «Бизнесстранс» для поощрения и поддержки сотрудничества между бизнесом и научными учреждениями с целью установления более тесных взаимоотношений.

Территория проекта: с польской стороны – поветы Хайнувка, Семятыче, Бяла Подляска, Влодава, с белорусской стороны – Брестский, Каменецкий, Пружанский, Малоритский районы.

Общий бюджет проекта – 162 982,4 евро. Доля белорусской стороны – 46 471,77 евро (из них 41 824,59 евро – средства ЕС).