

Н. А. Суранов^{1}, Т. И. Балтыжакова¹, А. Ю. Романчиков¹*

Применение нейронных сетей и методов глубокого обучения для обнаружения объектов на снимках

¹Санкт-Петербургский горный университет императрицы Екатерины II, г. Санкт-Петербург, Российская Федерация

*e-mail: nik.suranov.00@mail.ru

Аннотация. Работа посвящена анализу применения методов глубокого обучения для распознавания объектов на изображениях. Она представляет собой обзор, в котором рассматриваются основные задачи, выделяемые в рамках автоматического анализа изображений, уделяя наибольшее внимание обнаружению объектов и сегментации. Анализируются основные применяемые модели, как в теоретическом, так и в практическом плане. Целью исследования является установление степени проработанности научной проблемы. В публикации помимо преимуществ нейронных сетей идет речь о возникающих проблемах. Результаты показывают, что применение методов глубокого обучения в сфере анализа изображений возрастает.

Ключевые слова: глубокое обучение, нейронные сети, обнаружение объектов, дистанционное зондирование

N. A. Suranov^{1}, T. I. Baltyzhakova¹, A. Yu. Romanchikov¹*

Application of neural networks and deep learning methods for detecting objects in images

¹Saint Petersburg mining university of empress Catherine II, Saint-Petersburg, Russian Federation

*e-mail: nik.suranov.00@mail.ru

Abstract. The paper is dedicated to the analysis of the application of deep learning methods for object detection in images. It presents an overview that takes into account the main tasks that are highlighted in the context of automatic image analysis, with the greatest attention paid to the detection and segmentation of objects. The main models used are analysed from both a theoretical and practical point of view. The aim of the study is to determine the degree of scientific progress. In addition to the advantages of neural networks, the paper discusses the challenges encountered. The results show that the application of deep learning techniques in the field of image analysis is increasing.

Keywords: deep learning, neural networks, object detection, remote sensing

Введение

В настоящее время для изучения земной поверхности широкое применение получили дистанционные методы, которые позволяют получить наиболее актуальную информацию об объектах местности с использованием минимальных трудозатрат со стороны исполнителя. При этом используются как материалы дистанционного зондирования, так и аэрофотосъемка (АФС), преимущественно, с использованием беспилотных воздушных судов. Спутниковые изображения отличаются оперативностью получения данных, но обладают худшим разрешением по сравнению с АФС.

Наиболее современным подходом к обработке изображений является применение нейронных сетей глубокого обучения. В первую очередь при обработке изображений они решают задачу распознавания образов.

В рамках автоматического анализа изображений выделяют следующие задачи:

- классификация – обеспечивает целостное понимание того, что есть на изображении, присваивает метки и классы областям на изображении и предоставляет информацию о классе объекта - имя и прогнозную вероятность (уровень достоверности);

- обнаружение объектов заключается в локализации и идентификации конкретных объектов, представляющих интерес, при которой получают не только класс объекта, но и его местоположение на изображении;

- сегментация – процесс разбиения изображения на значимые области для идентификации объектов, представляющих интерес, в ходе этого процесса информация дополняется формой объекта [1, 2];

Сравнение решаемых задач на примере приведено на рис.1.



Рис. 1 Классификация, обнаружение и сегментация [2]

Применяя модели обнаружения можно построить только ограничивающую рамку, в которой идентифицирован конкретный объект, но ничего нельзя точно сказать о форме объекта. Сегментация изображений дает попиксельную маску для каждого объекта на снимке, что позволяет судить о формах объектов [1].

В данной статье наибольшее внимание уделено методикам обнаружения объектов и сегментации, поскольку задача локализации конкретных объектов часто встречается при решении задач из сферы землеустройства и кадастра с использованием данных дистанционного зондирования [3–5]: идентификация незарегистрированных построек, детектирование распространяющихся сельскохозяйственных вредителей и т.д. Также инструмент распознавания объектов будет полезен при изучении земель лесного фонда [6].

Цель работы – установление степени проработанности научной проблемы, а также определение методов и направлений, обеспечивающих их решение.

В ходе работы поставлены следующие задачи:

- проанализировать существующие модели обнаружения объектов;
- рассмотреть особенности применения этих моделей, а также реализацию в программных продуктах;
- подвести итоги о применимости нейронных сетей для обнаружения объектов.

Методы и материалы

В настоящее время структуры обнаружения объектов можно подразделить на 2 типа [7]:

- двухступенчатые – сначала идентифицируют регионы (области), в которых ожидается обнаружение объектов, а затем обнаруживает объекты только в этих регионах (R-CNN, Fast R-CNN, Faster R-CNN);
- одноступенчатые – используют подход, при котором сеть способна найти все объекты за один проход (Yolo и SSD).

Двухступенчатые структуры имеют несколько лучшую точность, но работают медленнее, в то время как однократные алгоритмы более эффективны и также имеют хорошую точность [7].

Наиболее известными и применимыми алгоритмами глубокого обучения являются: Faster R-CNN, Yolo и SSD. На рис. 2 представлена диаграмма разброса скорости и точности основных методов обнаружения объектов [8].

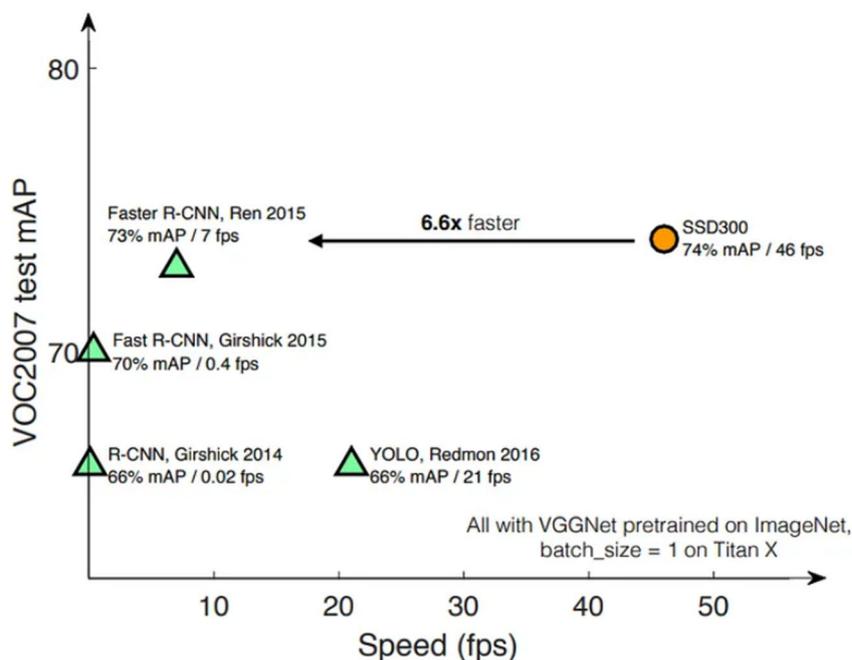


Рис. 2 Сравнение производительности SSD, Faster R-CNN и YOLO [8]

В данном сравнении SSD даже обошел по точности Faster R-CNN.

Таким образом, были рассмотрены основные модели, применяемые для обнаружения объектов на изображениях.

Авторский коллектив из Сингапура выделяет, что начало успеха применения глубоких сверточных нейронных сетей (DCNN) для классификации и обнаружения объектов

ружения объектов приходится на 2012 год (рис. 3). Основными препятствиями на первых этапах внедрения стали: отсутствие крупномасштабных обучающих данных, что приводило к переобучению, ограничения способности вычислительной техники и слабая теоретическая поддержка [9].

DCNN генерируют иерархические представления признаков от сырых необработанных пикселей до семантической высокоуровневой информации, которая автоматически изучается на основании обучающих данных.

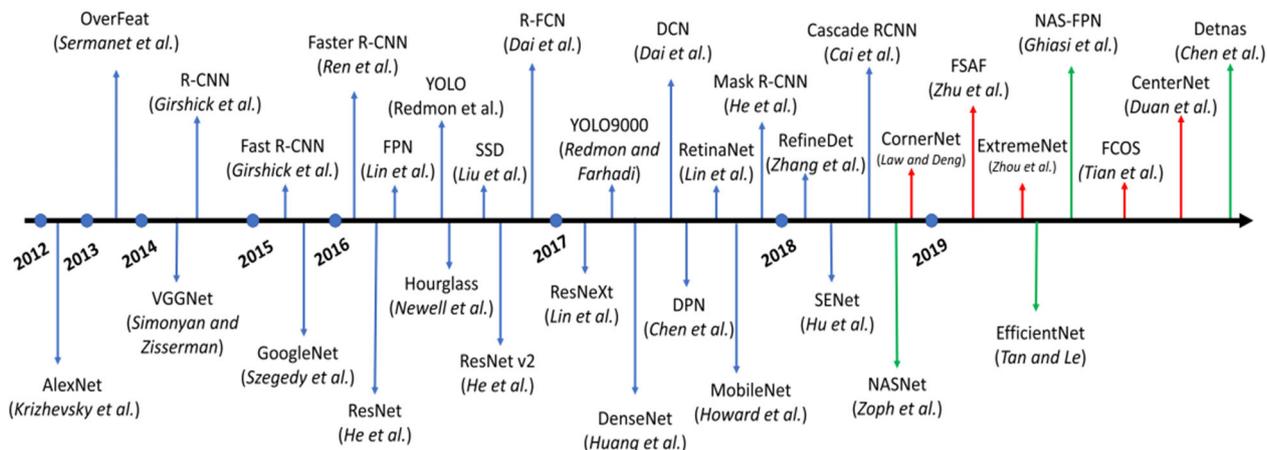


Рис. 3 Основные разработки в области методов обнаружения объектов на основе глубокого обучения с 2012 по 2019 года [9]

Результаты

Отдельной прикладной отраслью применения моделей глубокого обучения для обнаружения объектов является распознавание объектов на аэрофотоснимках и спутниковых снимках. Наиболее часто по ним решаются задачи земельного мониторинга и земельного надзора, а также локализуют объекты, представляющие интерес, для решения разноплановых задач: корабли, транспортные средства, здания, различные природные явления [10, 11].

Применение методов глубокого обучения для распознавания объектов на спутниковом изображении рассмотрено в публикации [12]. В качестве сверточной нейронной сети используется Mask-RCNN, направленная на сегментацию. В исследовании стояла задача распознавания автомобильных дорог, исходный снимок и результаты представлены на рис.4.

Как видно из рисунка сегментированные дороги имели некоторые дефекты распознавания (небольшие разрывы). Это может быть связано с тем, что в исследовании использовались спутниковые снимки WorldView-2 с разрешением 2 м или с недостаточной величиной обучающей выборки.

Одной из доступных программ для использования моделей глубокого обучения в задачах обнаружения объектов является ArcGIS Pro. Перечень задач, качественно решаемых программой, со временем все растет, при этом тех. поддержка остается на высоком уровне. В ArcGIS упрощено обучение самых разных моделей глубокого обучения, что является существенным плюсом для пользователя.

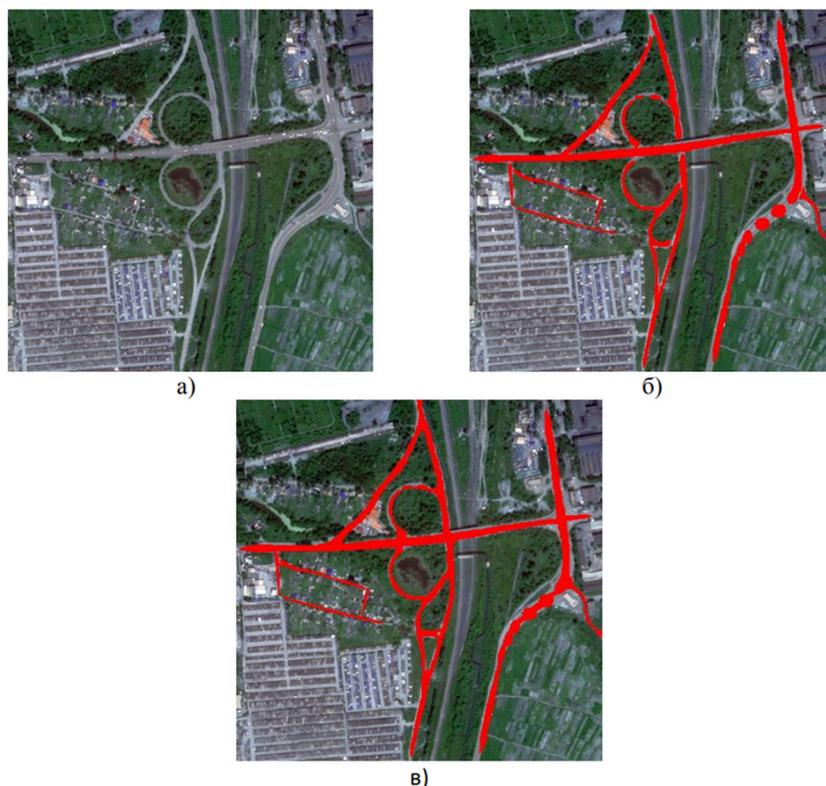


Рис. 4 Изображения дорожной сети. а – исходный снимок, б – результат работы Mask-RCNN, в – результат постобработки [12]

Для обнаружения объектов ArcGIS включается в себя несколько основных моделей обнаружения объектов, таких как SingleShotDetector (SSD) , RetinaNet , YOLOv3 и FasterRCNN [13]. Модели могут обнаруживать объекты даже на видео.

Помимо обнаружения объектов ArcGIS Pro также может решать задачи сегментации с помощью уже упоминавшейся выше модели – MaskRCNN. На рис. 5 приведен пример извлечение контуров зданий по снимку.



Рис. 5 Результат обнаружения контуров зданий [13]

Еще одним средством применяющим глубокое обучение для обнаружения объектов или сегментации можно назвать Mapflow AI от компании GeoAlert [14]. Это плагин на базе QGIS – самой популярной ГИС с открытым исходным кодом.

Ранее плагин был способен распознавать только здания, строительные площадки, сельскохозяйственные поля, леса и дороги. Но с апреля 2023 года появились открытая модель, которая может идентифицировать любые объекты на любом изображении – Segment anything with Mapflow [15]. На рис.6 представлен результат распознавания при 14-ти кратном уровне масштабирования. В зависимости от входного разрешения могут генерировать разные объекты.



Рис. 6 Результат распознавания землепользований, лесов, полей и водоемов [15]

По представленным примерам можно судить о широком распространении методов глубокого обучения в сфере обработки аэрофотоснимков и спутниковых изображений, а также об их внедрении в различные программы.

Обсуждение

Несмотря на все представленные выше преимущества методов глубокого обучения, у них есть ограничения, ввиду чего они не всегда применимы:

- размер объекта: Некоторые объекты на снимках занимают небольшую площадь. Универсальные модели хорошо работают с объектами среднего и большого размера, занимающими большое число пикселей, а при уменьшении масштаба их производительность может резко снизиться [16];

- некоторые модели чувствительны к тому, что объекты могут появляться в сильно разных масштабах на изображениях;

- сильная зависимость от количества и качества входных данных;

- вес данных и требовательность к вычислительной мощности устройств.

– вопреки представленным выше проблемам, алгоритмы глубокого обучения остаются по-прежнему востребованными в разных областях науки, в том числе в сфере распознавания объектов на снимках.

Заключение

В публикации приведен краткий обзор на наиболее применяемые модели нейронных сетей, используемых для распознавания объектов на изображениях. Приведены примеры их применения, а также сравнение производительности. Помимо этого рассмотрены некоторые программные продукты, позволяющие проводить распознавание объектов на снимках.

На сегодняшний день алгоритмы глубокого обучения достигли значимого прогресса в качестве инструмента обнаружения объектов. Результаты анализа доказывают, что их внедрение расширяется, а точность возрастает. Эффективность их применения зависит от характера решаемой проблемы.

Глубокое обучение – это быстроразвивающаяся отрасль, в которой постоянно появляются инновации и новые модели. В целом, можно ожидать, что область глубокого обучения будет продолжать развиваться и прогрессировать, что приведет к дальнейшему улучшению и расширению возможностей автоматического обнаружения объектов на изображениях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Segmentation vs Detection vs Classification in Computer Vision: A Comparative Analysis [Электронный ресурс]. – Режим доступа: <https://www.picsellia.com/post/segmentation-vs-detection-vs-classification-in-computer-vision-a-comparative-analysis#:~:text=Detection%20vs%20Classification%3A%20Differentiating%20Factors,labels%20to%20images%20or%20regions> (дата обращения: 28.10.2023).
2. How to Detect Objects in Images Using the YOLOv8 Neural Network [Электронный ресурс]. – Режим доступа: <https://www.freecodecamp.org/news/how-to-detect-objects-in-images-using-yolov8/> (дата обращения: 28.10.2023).
3. Осипов, А. Г., Дмитриев В. В., Ковязин В.Ф. Методика оценки и картографирования природно-аграрного потенциала ландшафтов // Геодезия и картография. – 2021. – Т. 82, № 9. – С. 11-20.
4. Скачкова, М. Е., Дьячкова И. С., Кремчеев Э.А. Оценка степени градостроительных преобразований урбанизированных территорий с использованием ГИС-технологий // Природообустройство. – 2018. – № 4. – С. 59-65.
5. Быкова, Е. Н., Баникевич Т. Д., Рагузин И.И. Современные особенности кадастровой оценки земель сельскохозяйственного назначения // Инженерный вестник Дона. – 2022. – № 6(90). – С. 1-14.
6. Ковязин В. Ф., Лепихина О. Ю., Демидова П. М., Колесник О. А. Причины слабой изученности земель лесного фонда России // Геодезия, картография, геоинформатика и кадастры. Производство и образование : Сборник материалов IV всероссийской научно-практической конференции. – Санкт-Петербург: Издательство «Политехника», 2021. – С. 478-482.
7. How single-shot detector (SSD) works? [Электронный ресурс]. – Режим доступа: <https://developers.arcgis.com/python/guide/how-ssd-works/> (дата обращения: 29.10.2023).
8. Faster R-CNN vs YOLO vs SSD – Object Detection Algorithms [Электронный ресурс]. – Режим доступа: <https://medium.com/ibm-data-ai/faster-r-cnn-vs-yolo-vs-ssd-object-detection-algorithms-18badb0e02dc> (дата обращения: 29.10.2023).

9. Xiongwei Wu, Doyen S., Steven C.H. Hoi. Recent advances in deep learning for object detection // *Neurocomputing*. – 2020. – Vol. 396. – P. 39–64.
10. Скрипачев В. О., Гуйда М. В., Гуйда Н. В., Жуков А. О. Исследование сверточных нейронных сетей для обнаружения объектов на аэрокосмических снимках // *International Journal of Open Information Technologies*. – 2022. – Т. 10, № 7. – С. 54-64.
11. Скачкова, М. Е., Гурьева О. С. Мониторинг состояния зеленых насаждений Санкт-Петербурга по материалам дистанционного зондирования // *Экология и промышленность России*. – 2023. – Т. 27, № 5. – С. 51-57.
12. Тусикова, А. А., Вихтенко Э. М. О распознавании автомобильных дорог на спутниковых снимках с использованием сверточных сетей mask-rcnn // *Информационные технологии и высокопроизводительные вычисления : материалы V Международной научно-практической конференции*. – Хабаровск: Тихоокеанский государственный университет, 2019. – С. 308-314.
13. Deep learning models in arcgis.learn [Электронный ресурс]. – Режим доступа: <https://www.esri.com/arcgis-blog/products/api-python/analytics/deep-learning-models-in-arcgis-learn/> (дата обращения: 29.10.2023).
14. Geoalert Аналитическая и картографическая платформа [Электронный ресурс]. – Режим доступа: <https://geoalert.io/solutions/power> (дата обращения: 29.10.2023).
15. Segment Anything with Mapflow [Электронный ресурс]. – Режим доступа: <https://medium.com/geoalert-platform-urban-monitoring/segment-anything-with-mapflow-b0d3d3b79c1> (дата обращения: 29.10.2023).
16. Yi W., Syed Muhammad A. B., Mahrukh Kh., Qudrat U., Rui W., Yilin S., Zhe G., Yilong N. Remote sensing image super-resolution and object detection: Benchmark and state of the art // *Expert Systems with Applications*. – 2022. – Vol. 197. – 116793.

© Н. А. Суранов, Т. И. Балтыжакова, А. Ю. Романчиков, 2024