

*Панина Виктория Сергеевна, студент-магистр, Калужский филиал ФГБОУ
ВО «Московский государственный технический университет имени Н.Э.*

Баумана (национальный исследовательский университет)»

*Амеличев Глеб Эдуардович, аспирант, Калужский филиал ФГБОУ ВО
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»*

*Белов Юрий Сергеевич, научный руководитель, к.ф. -м.н., доцент,
Калужский филиал ФГБОУ ВО «Московский государственный технический
университет имени Н.Э. Баумана (национальный исследовательский
университет)»*

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ MASK R-CNN В ИНТЕЛЛЕКТУАЛЬНЫХ ПАРКОВОЧНЫХ СИСТЕМАХ

Аннотация: Существуют различные методы распознавания свободных парковочных мест. Одним из таких способов является метод, основанный на распознавании изображения, получаемого с камер видеонаблюдения. Этот метод содержит более полную, но менее структурированную информацию о припаркованных автомобилях, в отличие от других методов. Для распознавания нескольких объектов на изображении предпочтительнее использовать более современную архитектуру, основанную на сверточных нейронных сетях (CNN) – Mask R-CNN. В этой статье описывается архитектура сети Mask R-CNN для построения интеллектуальной парковочной системы.

Ключевые слова: Интеллектуальные парковочные системы, сверточные нейронные сети, CNN, компьютерное зрение, сегментация изображений, Mask R-CNN.

Annotation: There are various methods for recognizing free parking spaces.

One of these methods is a method based on image recognition obtained from CCTV cameras. This method contains more complete but less structured information about parked cars than other methods. To recognize multiple objects in an image, it is preferable to use a more modern architecture based on convolutional neural networks (CNN) – Mask R-CNN. This article describes the architecture of the Mask R-CNN network for building an intelligent parking system.

Keywords: Intelligent parking systems, convolutional neural networks, CNN, computer vision, image segmentation, Mask R-CNN.

Введение. Существующие методы распознавание свободных парковочных мест можно разделить на три типа: на основе счетчиков, на основе датчиков и на основе изображений. Методы на основе счетчика ограничены развертыванием на закрытых парковках и работают путем подсчета количества транспортных средств, въезжающих и выезжающих с парковки. Методы на основе датчиков основаны на физических датчиках обнаружения, которые размещаются над или под парковками, но требуют значительных капитальных затрат на большое количество физических датчиков, необходимых для покрытия больших парковок. Методы на основе изображений основаны на системах камер и могут охватывать большие открытые или закрытые парковки, когда есть достаточно высокие и беспрепятственные точки обзора. Методы на основе изображений также содержат более богатую, но менее структурированную информацию, чем методы на основе счетчиков и датчиков; например, можно определить конкретные характеристики транспортного средства с помощью методов на основе изображений, что трудно сделать с помощью методов на основе счетчиков или датчиков [1].

Структура сверточной нейронной сети

Сверточная нейронная сеть (CNN) — это тип искусственной нейронной сети, используемой для распознавания и обработки изображений, которая оптимизирована для обработки данных пикселей [2]. Таким образом, сверточные нейронные сети являются фундаментальными и основными строительными

блоками для задачи компьютерного зрения по сегментации изображений.

Архитектура сверточной нейронной сети состоит из трех основных слоев:

1. Сверточный слой. Этот слой помогает абстрагировать входное изображение в виде карты объектов с помощью фильтров и ядер.
2. Слой объединения. Этот слой помогает уменьшать выборку карт объектов, суммируя присутствие объектов в участках карты объектов.
3. Полносвязный слой. Полностью связанные слои соединяют каждый нейрон в одном слое с каждым нейроном в другом слое.

Объединение слоев CNN позволяет разработанной нейронной сети научиться идентифицировать и распознавать интересующий объект на изображении. Простые сверточные нейронные сети созданы для классификации изображений и обнаружения объектов с помощью одного объекта на изображении. На рисунке 1 показана концепция архитектуры CNN.

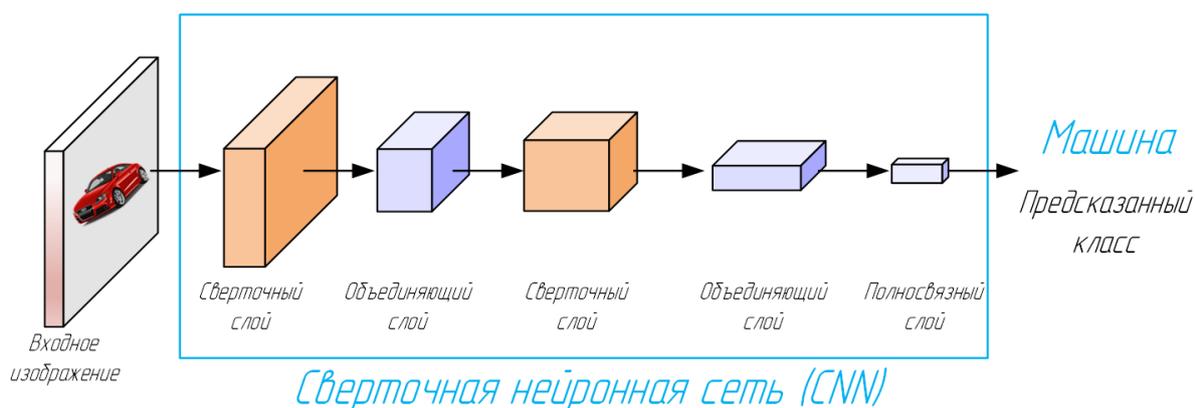


Рисунок 1 – Концепция архитектуры CNN: как работает сверточная нейронная сеть

В более сложной ситуации с несколькими объектами на изображении простая архитектура CNN не является оптимальной. Для таких ситуаций используется Mask R-CNN, представляющая собой более современную архитектуру [3].

Mask R-CNN

Mask R-CNN или Mask RCNN – это сверточная нейронная сеть (CNN), современная с точки зрения сегментации изображений и сегментации экземпляров. Mask R-CNN была разработана на основе Faster R-CNN,

региональной сверточной нейронной сети [4].

Первый шаг к пониманию того, как работает Mask R-CNN, требует понимания концепции сегментации изображения.

Задача компьютерного зрения «Сегментация изображения» — это процесс разделения цифрового изображения на несколько сегментов (наборов пикселей, также известных как объекты изображения). Эта сегментация используется для обнаружения объектов и границ (линий, кривых и т. д.).

Есть 2 основных типа сегментации изображения, которые подпадают под действие Mask R-CNN:

1. Семантическая сегментация
2. Сегментация экземпляра

Семантическая сегментация

Семантическая сегментация классифицирует каждый пиксель по фиксированному набору категорий без дифференциации экземпляров объекта. Другими словами, семантическая сегментация имеет дело с идентификацией/классификацией похожих объектов как единого класса на уровне пикселей.

Сегментация экземпляра

Сегментация экземпляров или распознавание экземпляров занимается правильным обнаружением всех объектов на изображении, а также точным сегментированием каждого экземпляра. Таким образом, это комбинация обнаружения объектов, локализации объектов и классификации объектов. Другими словами, этот тип сегментации идет дальше, чтобы дать четкое различие между каждым объектом, классифицируемым как аналогичные экземпляры.

Общая структурная схема Mask R-CNN

Mask R-CNN была построена с использованием Faster R-CNN. В то время как Faster R-CNN имеет 2 выхода для каждого объекта-кандидата, метку класса и смещение ограничительной рамки, Mask R-CNN — это добавление третьей ветви, которая выводит маску объекта [5]. Дополнительный вывод маски

отличается от вывода класса и блока, требуя извлечения гораздо более тонкой пространственной компоновки объекта.

Mask R-CNN является расширением Faster R-CNN и работает путем добавления ветви для прогнозирования маски объекта (области интереса) параллельно с существующей ветвью для распознавания ограничительной рамки.

Преимущества Маски R-CNN

- Простота. Mask R-CNN прост в обучении.
- Производительность. Mask R-CNN превосходит все существующие записи с одной моделью в каждой задаче.
- Эффективность. этот метод очень эффективен и добавляет лишь небольшие накладные расходы Faster R-CNN.
- Гибкость. Mask R-CNN легко обобщить на другие задачи.

Таким образом, Mask R-CNN — это метод, расширенный из Faster R-CNN путем добавления ветвей для прогнозирования объектов маски параллельно существующей ветви для определения предела площади [6]. Суть в том, чтобы разделить задачу классификации и предсказание маски на уровне пикселей. Общая топология сети Mask R-CNN представлена на рисунке 2.

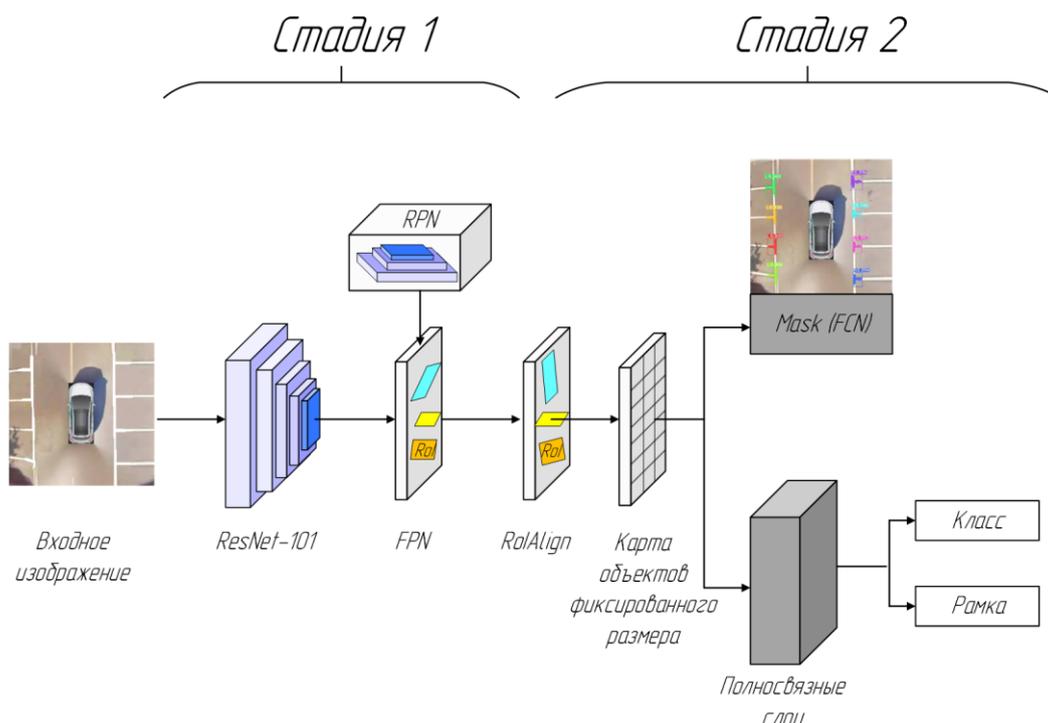


Рисунок 2 – Общая структурная схема Mask R-CNN

На основе фреймворка Faster R-CNN была добавлена третья ветвь для прогнозирования масок объектов параллельно с существующими ветвями для классификации и локализации. Маска ответвления — это полностью подключенная небольшая сеть, которая применяется к каждой области интереса, прогнозируя сегментацию маски от пикселя к пикселю. Дополнительная маска позволяет разумно выбирать пиксели в каждом сегменте объекта и извлекать каждый объект отдельно без какого-либо фона (что невозможно при семантической сегментации). Результаты прогнозирования R-CNN для набора данных Common Objects in Context (COCO) позволяют детально идентифицировать границы объектов на уровне пикселей. Метод Mask R-CNN — это простой метод, который добавляет лишь небольшие накладные расходы, чтобы заставить его работать со скоростью пять кадров в секунду. Этот метод успешно превосходит все записи с одной моделью в каждой заданной задаче. Результатом, полученным методом Mask R-CNN, является квадратный предел и маска сегментации в каждом объекте, сталкиваемом с сегментацией экземпляра [7].

Поскольку сегментация на уровне пикселей требует более точного выравнивания, чем граничные квадраты, Mask R-CNN улучшает слой пула RoI (RoIAlign), чтобы его можно было лучше и точнее сопоставить с областью исходного изображения. Модель R-CNN обрабатывает входное изображение, предоставленное сети, путем выполнения выборочного поиска. Затем он использует область результатов выборочного поиска для извлечения признаков и классификации с использованием уже обученной CNN.

Выводы

Mask R-CNN — это модель, созданная для обнаружения объектов по всему изображению с вычислительной точки зрения, без использования традиционного подхода скользящего окна, используемого в большинстве методов обнаружения объектов. Она работает довольно быстро по сравнению с традиционными

методами скользящего окна. Когда используется современный графический процессор, Mask-RCNN всегда может распознавать объекты в видео высокого разрешения со скоростью несколько кадров в секунду. В дополнение к преимуществам Mask R-CNN предлагает достаточно информации о каждом обнаруженном объекте. В основном алгоритмы обнаружения объектов возвращают только ограничительную рамку каждого объекта, однако Mask R-CNN не предлагает только размещение каждого объекта, он одновременно предлагает объект, связанный (или маскирующий) вокруг идентифицированных объектов [8].

Библиографический список:

1. Панина В.С., Амеличев Г.Э., Белов Ю.С. Интеллектуальная парковочная система как часть интеллектуальной транспортной системы // E-Scio. – 2022. – № 1(64). – С. 445-452.

2. Панина В.С., Амеличев Г.Э., Белов Ю.С. Интеллектуальная парковочная система на основе сверточных нейронных сетей // Научное обозрение. Технические науки. – 2022. – № 1. – С. 29-33; URL: <https://science-engineering.ru/ru/article/view?id=1382> (дата обращения: 31.05.2022).

3. Majgaonkar S., Waghela K., Shah R., Pavate A. and Mishra P. Developing a Parking Monitoring System using Mask-RCNN // SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology. 12, SUP 1 (Jun. 2020), pp. 386-389.

4. Naufal A.R., Faticah C., & Suciati N. Preprocessed Mask RCNN for Parking Space Detection in Smart Parking Systems // International Journal of Intelligent Engineering and Systems, 2020, 13, pp. 255-265.

5. Jiang S., Jiang H., Ma S., & Jiang Z. Detection of Parking Slots Based on Mask R-CNN // Applied Sciences, 2020, 10, 4295.

6. Nyambal J. and Klein R., Automated parking space detection using convolutional neural networks // Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), 2017, pp. 1-6.

7. Patel R. and Meduri P., Car detection based algorithm for automatic parking

space detection // 19th IEEE International Conference on Machine Learning and Applications (ICMLA), 2020, pp. 1418-1423.

8. He K., Gkioxari G., Dollár P. and Girshick R., Mask R-CNN // IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980-2988.