

УДК: 004.8

DOI: 10.30987/2658-3488-2019-2019-4-38-43

М.С. Любимов, В.И. Лушков, А.А. Азарченков

РАСПОЗНАВАНИЕ ОБЪЕКТОВ ДОРОЖНОЙ ИНФРАСТРУКТУРЫ С ПОМОЩЬЮ ПОЛНОСВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

В данной статье представлен метод распознавания ключевых объектов дорожной инфраструктуры с использованием полносверточной нейронной сети. Результатом работы нейронной сети является сегментированное изображение, где искомые объекты выделены определенными цветами. На этапе постобработки происходит выделение участка дорожного полотна, по которому передвигается машина, а также вычисление параметров ограничивающих прямоугольников, для каждого из объектов. Данный метод позволяет локализовать дорогу, пешеходный переход, автомобили, дорожные знаки, светофоры, пешеходов. Тестирование разработанного алгоритма производилось на макете городской инфраструктуры в масштабе 1:18, где в качестве автомобиля выступал колесный робот.

Ключевые слова: *полносверточная нейронная сеть, локализация объектов, сегментирование изображения, Pyramid Scene Parsing Network, колесный робот.*

M.S. Lubimov, V.I. Lushkov, A.A. Azarchenkov

OBJECT RECOGNITION OF ROAD INFRASTRUCTURE USING A FULLY CONVOLUTIONAL NEURAL NETWORK

This article presents a method for recognizing key objects of the road infrastructure using a fully convolutional neural network. The result of the neural network is a segmented image, where the desired objects are highlighted in certain colors. At the post-processing stage, a section of the roadway along which the car moves is selected, as well as the calculation of the parameters of the bounding rectangles for each of the objects. This method allows you to localize the road, pedestrian crossing, cars, traffic signs, traffic lights, pedestrians. Testing of the developed algorithm was carried out on a model of the urban infrastructure at a scale of 1:18, where a wheeled robot acted as a car.

Keywords: *fully convolutional neural network, object localization, image segmentation, Pyramid Scene Parsing Network, wheeled robot.*

Введение

В настоящее время ведутся активные разработки в области беспилотных транспортных средств. Некоторые компании приближаются к последнему, пятому уровню автономности, при котором становится возможным полностью автономное движение автомобиля без участия водителя. Одной из ключевых задач, без которых невозможно создание беспилотного средства в динамической среде является задача распознавания объектов дорожной сети. Существуют различные подходы для решения этой задачи: от распознавания отдельных элементов, таких как дорожные знаки или светофоры [7], до более комплексных подходов [8], позволяющих детектировать несколько элементов. Распознавание может производиться на основе информации, приходящей от разных сенсоров. Наиболее сенсорами являются камеры, возвращающие изображения и лидары, возвращающие облако точек. Использование камер является более традиционным методом, который имитирует зрение водителя. В области компьютерного зрения распознавание объектов является одной из основных задач, для решения которой используют как стандартные алгоритмы компьютерного зрения [4], так и использования различных архитектур нейронных систем [1]. Метод, описанный в данной статье, основан на использовании полносверточной нейронной сети [6], выполняющей задачу сегментации изображения.

Говоря о разработке беспилотного транспортного средства отдельное внимание следует уделить этапу тестирования алгоритмов. Тестирование на реальных автомобилях может

повлечь как человеческие, так и финансовые жертвы. Поэтому одним из вариантов решения этой задачи может служить использование модели в уменьшенном масштабе, сохраняющей основные свойства реального движения. Еще одно преимущество использования макета заключается в том, что он позволяет приобретать навык людям, желающим развиваться в области беспилотных транспортных средств, без серьезных последствий. Разработанная система распознавание объектов была протестирована на макете, содержащем несколько видов светофоров, дорожные знаки, пешеходов, автомобили, дорожное полотно с разметкой.

1. Распознавание объектов

Основой детектирования объектов является полносверточная нейронная сеть, которая принимает на вход изображение с камеры и возвращает сегментированное изображение (рис. 1). Сегментирование – это процесс, при котором генерируется изображение, содержащее однородные области одинакового цвета, где каждому цвету соответствует определенный объект. Определение объектов на основе сегментированного изображения позволяет решить несколько задач, необходимых для беспилотного движения автомобиля: с одной стороны, она позволяет детектировать объекты дорожной сети, с другой решается задача распознавания дорожного полотна.

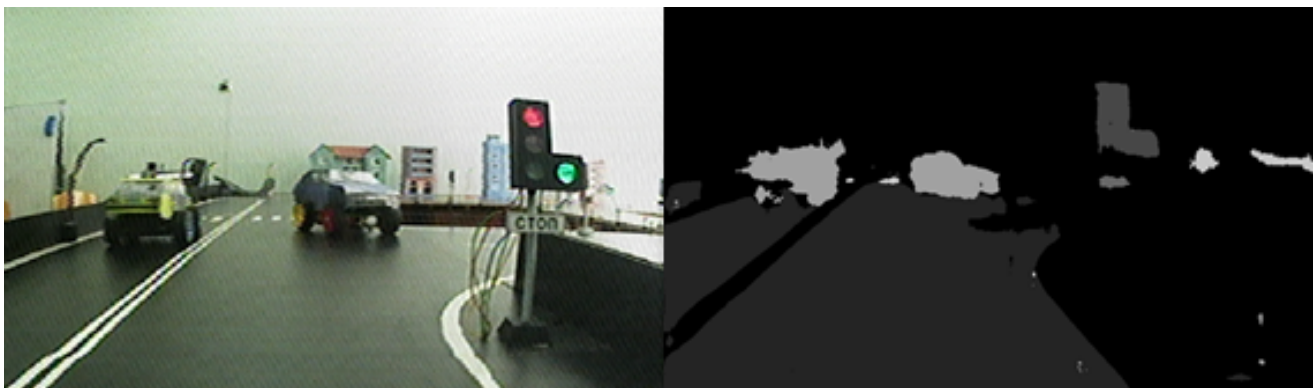


Рис. 1. Пример сегментированного изображения

Используемая нейронная сеть базируется на архитектуре Pyramid Scene Parsing Network [3]. Основными компонентами сети являются предобученная сеть Residual Network (Resnet) [5] и Pyramid Pooling Module.

Модель Resnet является сверточной моделью, отличительной особенностью которой является, наличие остаточной структуры обучения позволяет увеличивать глубину сети без потерь точности. На выходе сверточной сети генерируется карта признаков (feature map), которая поступает на слой подвыборки (pooling layer), данная часть сети называется пирамидальной подвыборкой (Pyramid Pooling Module) потому что состоит из нескольких уровней (рис 2.):

- красный: это самый крупный уровень, который выполняет глобальную подвыборку для каждой карты признаков, в результате чего генерируется одно значение на выходе;
- оранжевый: это второй уровень, который делит карту признаков на 2×2 субрегиона, а затем выполняет среднее объединение для каждого субрегиона;
- синий: это третий уровень, который делит карту объектов на 3×3 субрегиона, а затем выполняет среднее объединение для каждого субрегиона;
- зеленый: это уровень, который делит карту объектов на 6×6 субрегионов, а затем выполняет объединение для каждого субрегиона.

Затем происходит операции увеличения размерности слоев (Upsample), и их объединение.

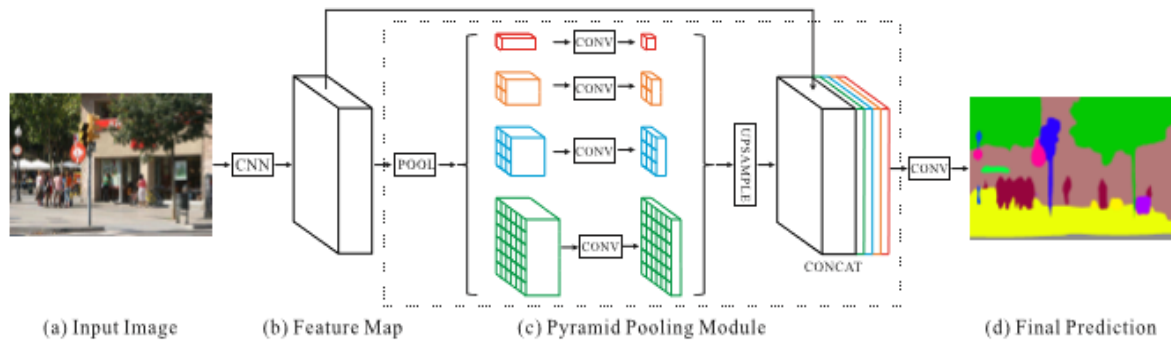


Рис. 2. Pyramid Science Network

2. Обучение

Для обучения нейронной сети был подготовлен собственный набор данных, содержащий макет, по которому передвигался колесный робот. Для каждого изображения вручную выполнялось заполнение областей, принадлежавших разным объектам. Было подготовлено более 1000 кадров, для обучения нейронной сети.

Важным этапом обучения нейронной сети является аугментация данных. Аугментация – это процесс, позволяющий расширить существующий датасет за счет применения методов обработки изображения. Важно, что аугментация данных должна так, чтобы измененные изображения не имели значительного отклонения от потенциально возможных при использовании нейронной сети. В разрабатываемой статье использовались следующие методы:

- изменение яркости изображения;
- изменение экспозиции изображения;
- отражение части изображений по горизонтали;
- добавление шумов на изображение.

Необходимость последнего пункта обоснована тем, что при тестировании алгоритма, на изображении было замечено появление шумов, связанных со способом передачи видеопотока на вычислительный узел. Искусственно создаваемые шумы позволяют уменьшить ошибки при появлении реальных помех

3. Выделение объектов

После генерации сегментированного изображения следует этап поиска объектов на изображении. Для каждого из возможных классов осуществляется поиск по цвету, привязанному к классу. После чего для изображения, содержащего только один цвет выполняется морфологическая операция замыкания, помогающая сделать контуры более однородными (рис. 3.). После поиска областей, происходит выделение их контуров. В том случае, если несколько контуров разрозненны, но расположены близко (рис. 4.) происходит их объединение. Исходя из положения контура, генерируется прямоугольник, ограничивающий объект.

Способ выделения дороги отличался от других объектов. Для дальнейшего поиска дорожной разметки необходимо определение части дороги, по которой может передвигаться автомобиль. В случае, если дорога имеет сплошную разметку, полосы дороги выделяются отдельно. Если разметка прерывистая выделяется все дорожное полотно, включая прерывистую разметку. На этапе выделения находится наибольший контур дороги на изображении, полагая, что машина едет по дороге в момент детекции. Кроме того, дорога,

по которой передвигался колесный робот содержит пешеходные переходы, которые не оказывают влияния на работу дальнейших алгоритмов. В связи с этим было принято решение определять пешеходные переходы нейронной сетью и скрывать их на итоговом сегменте дороги.



Рис. 3. Применение операции замыкания



Рис. 4. Пример необъединенных контуров

4. Эксперименты

Описанный подход использовался для проезда колёсного робота по макету городской инфраструктуры (рис. 5). Макет имеет размер 6x4 и оборудован, дорожным полотном с разметкой, дорожными знаками, другими автомобилями, пешеходами, светофорами, в том числе светофорами с дополнительной секцией. Колесный робот изображён на рисунке 6. Особенность архитектуры заключается в том, что все вычисления происходят на удаленном узле. Благодаря этому появляется возможность работать с модулями, требующими более высокой производительности, что приближает макет к реальным транспортным средствам.

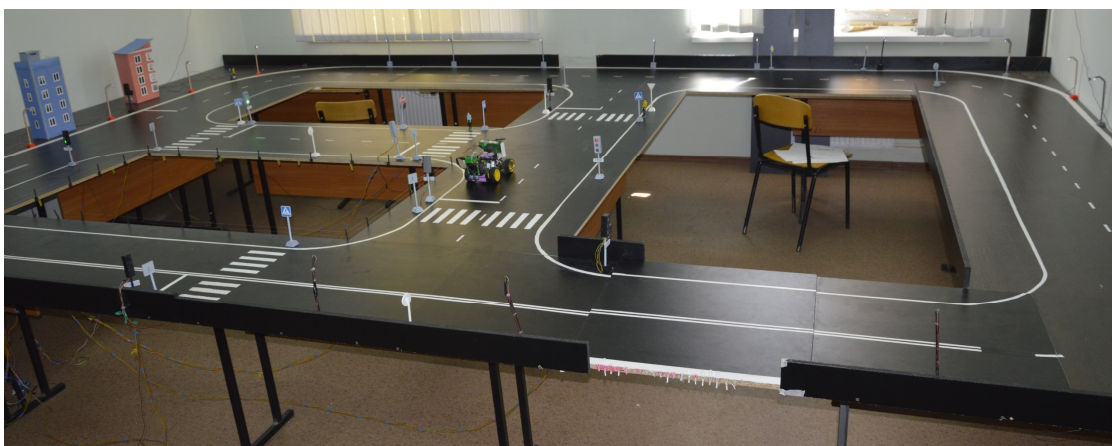


Рис. 5. Макет дорожной инфраструктуры

Запуск нейронной сети проводился на видеокарте NVIDIA GTX 1080. Время генерации одного сегментированного изображения составляет в среднем 0.063 сек, что обеспечивает скорость, достаточную для передвижения по макету. Для тестирования полученная нейронная сеть и подсистема выделения объектов были интегрированы в основной проект, написанный с использованием фреймворка ROS. Первый узел подписывается на

изображение с камеры, генерирует сегментированное изображение и публикует его. Второй узел подписывается на обычное и сегментированное изображение, выделяет объекты и публикует следующие сообщения: дорожное полотно, дорожный знак, который отправляется на классификацию, светофор, который отправляется на модуль определения сигнала, сообщение, содержащее пешеходов и автомобиль, которое отправляется на модуль слияния объектов. Найденные объекты добавляются в базу, после чего пересылаются в узел генерации управляющего сигнала, в котором происходит анализ данных и создание управляющих команд для колёсного робота.

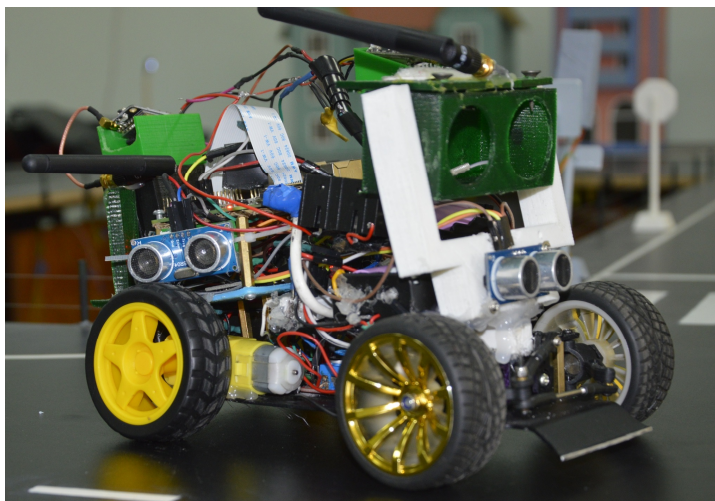


Рис. 6. Колёсный робот

Выводы

В рамках работы была получена полносверточная нейронная сеть, локализующая основные объекты дорожной инфраструктуры, такие как дорожное полотно, пешеходные переходы, автомобили, светофоры, дорожные знаки. На основе данной нейронной и модуля визуализации были написаны два узла ROS, интегрированные в основной проект. Результаты были протестированы на макете, имитирующим городскую среду.

Список литературы:

1. А. Буйвал, М. Любимов, А. Габдуллин «Обнаружение и локализация дорожных знаков на основе данных камеры и лидара» // 11-я Международная конференция по машинному зрению (ICMV 2018) - Мюнхен, Германия 2018.
2. Кордтс М., Омран М., Рамос С., Рехфельд Т., Энцивайлер М., Бененсон Р., Франке У., Рот С., Шиле Б. «Набор данных городских пейзажей для понимания семантической городской сцены» / M. Cordts. 2016г.
3. Хэншуан З., Цзяньпин С., Сяоцзюань В., Сяоган В., Цзя Дж. «Сеть пирамидального разбора сцены» / З. Хэншуан. CVPR 2017 г.
4. С. Хоссейн, З. Хайдер «Обнаружение и распознавание дорожных знаков для автомобильного транспорта», Международный журнал компьютерных приложений (2015).
5. Кайминг Х., Сяньюй З., Шаоцин Р., Цзянь С. «Глубокое остаточное обучение для распознавания изображений» / Х. Кайминг, З. Сяньюй, Р. Шаоцин, С. Цзянь – Microsoft Research 2015 г.

References:

1. A. Buyval, M. Lyubimov, A. Gabdullin «Road sign detection and localization based on camera and lidar data» //The 11th International Conference on Machine Vision (ICMV 2018) — Munich, Germany 2018г.
2. Cordts M., Omran M., Ramos S.,Rehfeld T.,Enzweiler M., Benenson R., Franke U., Roth S., Schiele B. «The Cityscapes Dataset for Semantic Urban Scene Understanding» / M. Cordts. 2016г.
3. Hengshuang Z., Jianping S., Xiaojuan Q., Xiaogang W., Jiaya J. «Pyramid Scene Parsing Network» / Z. Hengshuang. CVPR 2017г.
4. Md. S. Hossain, Z. Hyder «Traffic Road Sign Detection and Recognition for Automotive Vehicles», International Journal of Computer Applications (2015г.)
5. Kaiming H, Xiangyu Z,Shaoqing R,Jian S «Deep Residual Learning for Image Recognition» / H. Kaiming, Z. Xiangyu, R. Shaoqing, S Jian – Microsoft Research 2015 г.

6. Дж. Лонг, Э. Шелхамер, Т. Даррелл «Полносверточные сети для семантической сегментации» // CVPR 2015 г.
7. А. Могельмосе, М. М. Триведи, Т. Б. Меслунд, «Обнаружение и анализ дорожных знаков на основе видения для интеллектуальных систем помощи водителю: перспективы и исследование», «Операции IEEE в интеллектуальных транспортных системах» (2012 г.)
8. А. Д. Пон, О. Андриенко, А. Хараке, С. Л. Уослендер «Иерархическая глубокая архитектура и метод выбора мини-партии для совместного обнаружения дорожных знаков и света» 15-я конференция IEEE по компьютерному и робототехническому видению (2018 г.)
6. J. Long, E. Shelhamer, T. Darrell «Fully Convolutional Networks for Semantic Segmentation» // CVPR 2015г.
7. A. Møgelmoose, M. M. Trivedi, T. B. Moeslund, «Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Surve.» IEEE Transactions on Intelligent Transportation Systems, (2012).
8. A. D. Pon, O. Andrienko, A. Harakeh, S. L. Waslander «A Hierarchical Deep Architecture and Mini-Batch Selection Method For Joint Traffic Sign and Light Detection» IEEE 15th Conference on Computer and Robot Vision (2018).

Статья поступила в редколлегию 18.09.19.

Рецензент: д.т.н., профессор,

Брянский государственный технический университет

Лозбинев Ф.Ю.

Статья принята к публикации 30.09.19.

Сведения об авторах:

Азарченков Андрей Анатольевич

Кандидат технических наук, доцент,
ФГБОУ ВО «Брянский государственный технический университет»
Тел.: +7 (910) 230-90-35
E-mail: azarchenkovaa@gmail.com

Любимов Максим Сергеевич

Студент ФГБОУ ВО «Брянский государственный технический университет»
Тел.: +7 (930) 720-90-00,
E-mail: max32@inbox.ru

Лущков Владислав Игоревич

Студент ФГБОУ ВО «Брянский государственный технический университет»
Тел.: +7 (952) 967-55-47,
E-mail: lyshkoff2017@yandex.ru

Information about authors:

Azarchenkov Andrey Anatolevich

Ph.D, Associate Professor of IT, FSBEI HE «Bryansk State Technical University»,
Tel.: +7 (910) 230-90-35
E-mail: azarchenkovaa@gmail.com

Lyubimov Maxim Sergeevich

Student FSBEI HE «Bryansk State Technical University»
Tel.: +7 (930) 720-90-00,
E-mail: max32@inbox.ru

Lushkov Vladislav Igorevich

Student FSBEI HE «Bryansk State Technical University»,
Tel.: +7 (952) 967-55-47,
E-mail: lyshkoff2017@yandex.ru