

Научная статья

УДК 004.8

DOI 10.25205/1818-7900-2021-19-3-26-39

Семантическая сегментация изображений в проекте «Duckietown»

Дарья Евгеньевна Шабалина¹
Кристина Сергеевна Ланчуковская²
Татьяна Викторовна Лях³
Константин Владимирович Чайка⁴

¹⁻³ Новосибирский государственный университет
Новосибирск, Россия

⁴ Санкт-Петербургский государственный электротехнический университет
Санкт-Петербург, Россия

⁴ Лаборатория алгоритмов мобильных роботов «JetBrains Research»
Санкт-Петербург, Россия

¹ d.shabalina@g.nsu.ru, <https://orcid.org/0000-0003-2418-6172>

² k.lanchukovskaya@g.nsu.ru, <https://orcid.org/0000-0002-9085-3634>

³ t.liakh@g.nsu.ru, <https://orcid.org/0000-0001-9148-946X>

⁴ pro100kot14@gmail.com, <https://orcid.org/0000-0001-5778-9266>

Аннотация

Статья посвящена сравнительному анализу алгоритмов семантической сегментации и исследованию их применимости на примере проекта «Duckietown». Рассмотрены как классические алгоритмы семантической сегментации, так и алгоритмы, использующие подходы машинного обучения. Исследованы фреймворки машинного обучения с учетом всех ограничений проекта «Duckietown». По результатам исследования для решения задачи сегментации в проекте «Duckietown» были выбраны нейросетевые алгоритмы, основанные на сетях U-Net, SegNet, DeepLab-v3, FC-DenceNet и PSPNet. U-Net и SegNet и протестированы на симуляторе «Duckietown».

Ключевые слова

робототехника, компьютерное зрение, нейронные сети, семантическая сегментация изображений, проект «Duckietown», роботы, искусственный интеллект

Благодарности

Работа выполнена при поддержке Лаборатории алгоритмов мобильных роботов «JetBrains Research»

Для цитирования

Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В. Семантическая сегментация изображений в проекте «Duckietown» // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 26–39. DOI 10.25205/1818-7900-2021-19-3-26-39

© Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 26–39

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 26–39

Semantic Image Segmentation in Duckietown

Daria E. Shabalina¹, Kristina S. Lanchukovskaya²
Tatyana V. Lyakh³, Konstantin V. Chaika⁴

¹⁻³ Novosibirsk State University
Novosibirsk, Russian Federation

⁴ Saint Petersburg Electrotechnical University
St. Petersburg, Russian Federation

⁴ Mobile Robot Algorithms Laboratory “JetBrains Research”
St. Petersburg, Russian Federation

¹ d.shabalina@g.nsu.ru, <https://orcid.org/0000-0003-2418-6172>

² k.lanchukovskaya@g.nsu.ru, <https://orcid.org/0000-0002-9085-3634>

³ t.lyakh@g.nsu.ru, <https://orcid.org/0000-0001-9148-946X>

⁴ pro100kot14@gmail.com, <https://orcid.org/0000-0001-5778-9266>

Abstract

The article is devoted to evaluation of the applicability of existing semantic segmentation algorithms for the “Duckietown” simulator. The article explores classical semantic segmentation algorithms as well as ones based on neural networks. We also examined machine learning frameworks, taking into account all the limitations of the “Duckietown” simulator. According to the research results, we selected neural network algorithms based on U-Net, SegNet, DeepLab-v3, FC-DenceNet and PSPNet networks to solve the segmentation problem in the “Duckietown” project. U-Net and SegNet have been tested on the “Duckietown” simulator.

Keywords

robotics, computer vision, neural networks, semantic image segmentation, Duckietown, duckiebots, artificial intelligence

Acknowledgements

The work was carried out with the support of the Mobile Robot Algorithms Laboratory “JetBrains Research”

For citation

Shabalina D. E., Lanchukovskaya K. S., Liakh T. V., Chaika K. V. Semantic Image Segmentation in Duckietown. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 26–39. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-26-39

Введение

В настоящее время активно развиваются и используются тренажеры, которые имитируют различные роботизированные системы, а также индустриальные технологические системы. Одним из таких тренажеров является «Duckietown»¹. Это открытый исследовательский проект в области автономно управляемых роботов. Он состоит из двух частей – роботов (duckiebots) и города (Duckietown), по которому они передвигаются. В городе есть размеченные дороги, светофоры, дорожные знаки и препятствия. Сами роботы – duckiebots – представляют собой двухколесную платформу, которая управляется двумя моторами постоянного тока, основным вычислительным модулем является одноплатный компьютер Raspberry Pi. На роботах установлены драйвер моторов и одна камера с Fisheye объективом². Несмотря на техническую простоту, это позволяет роботу двигаться согласно разметке, избегать столкновений с препятствиями и другими роботами, определять свое местоположение на карте, самостоятельно передвигаться по городу, взаимодействовать с другими роботами.

Тренажер отражает особенности массовых индустриальных автопилотируемых систем с ограниченными ресурсами: такие роботы оснащены только камерами, а также имеют сла-

¹ URL: <https://www.duckietown.org/> (дата обращения 31.03.2021).

² Getting the Duckiebot hardware. URL: https://docs.duckietown.org/daffy/opmanual_duckiebot/out/get_db_hw.html (дата обращения 31.03.2021).

бую вычислительную мощность и минимальный набор датчиков. Подобные роботы используются в промышленности, например, в качестве роботов-погрузчиков [1].

На сегодняшний день важной задачей, в том числе в рамках проекта «Duckietown», является семантическая сегментация изображений, которые поступают с камеры робота. Причем данная задача применима как к данным камеры реального мира, так и к симуляционным. Семантическая сегментация [2] – это задача поиска групп пикселей изображения, каждая из которых характеризует один смысловой объект. При семантической сегментации требуется не только обнаружить и классифицировать объекты, но и определить их границы. Она применяется во многих сферах, в том числе и для алгоритмов управления беспилотным транспортом, а также в различных практических сферах, таких как промышленная робототехника, автомобильная промышленность, медицина [3]. В таких системах семантическая сегментация позволяет точнее проанализировать спутниковые изображения и карты, рентгенографические, электронно-микроскопические снимки, изображения, полученные с камер автопилотируемых автомобилей и роботов [4; 5]. Сложность решения этой задачи для тренажера «Duckietown» обусловлена ограничениями, которые накладывает робот: малой вычислительной мощностью, отсутствием дополнительных датчиков, монокулярной камерой.

Статья посвящена сравнительному анализу алгоритмов семантической сегментации и их применимости на примере проекта «Duckietown».

Обзор методов и алгоритмов семантической сегментации изображений

Для решения задач семантической сегментации применяются следующие подходы:

- граничные алгоритмы;
- пороговые алгоритмы;
- алгоритмы наращивания областей;
- алгоритмы сегментации на основе кластеризации;
- нейронные сети.

При анализе подходы оценивались по следующим критериям:

- устойчивость;
- оптимальные условия;
- ограничения и применение в робототехнике;
- ограничения и применение в автономном вождении.

Граничные алгоритмы

Граничные методы [6] основаны на выделении контуров объектов. Работа граничных методов состоит из двух этапов: поиск граничных пикселей с помощью операторов Робертса, Собеля, Лапласа, Кэнни и т. д., и замыкание границ. Основным недостатком граничных методов является сильная зависимость от «шума», что зачастую ведет к пересегментации [7; 8].

Пороговые алгоритмы

Пороговые алгоритмы [9] основаны на разделении гистограммы изображения на различные части по пороговым значениям. Существуют два различных метода: метод с глобальным порогом и метод с адаптивным порогом. Остальные алгоритмы – это производные от этих двух методов. Пороговые алгоритмы используют, когда нужно провести сегментацию простого изображения, области которого значительно различаются по яркости. Но при наличии различных помех – вспышек, теней, неравномерного освещения, что часто встречается в реальной жизни, пороговые методы не дают хорошего результата.

Алгоритмы наращивания областей

Алгоритмы наращивания областей в простейшем варианте основаны на выборе пикселя и изучении смежных с ним пикселей для проверки значений на близость по яркости [10]. Ес-

ли значения яркости близки, то соответствующие пиксели относятся в одну группу. Таким образом, в результате сращивания пикселей формируется область. Существуют и более эффективные варианты, в них в качестве начальной точки используются не отдельные пиксели, а небольшие области изображения. Каждая область проверяется на однородность: если результат проверки неудовлетворительный, то соответствующая область разбивается на более мелкие. Процесс продолжается до тех пор, пока все области не пройдут проверку на однородность. Далее производится формирование областей с помощью наращивания небольших областей, а не отдельных пикселей. Недостатком алгоритма является задача выбора начальной точки для сегментируемой области вручную, поэтому метод не может быть полностью автоматизирован.

Алгоритмы сегментации на основе кластеризации

Кластеризация [11] – задача разбиения множества объектов на непересекающиеся подмножества, называемые кластерами, таким образом, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались. Входными данными для задачи кластеризации являются признаковые описания объектов – наборы характеристик, свойственные каждому объекту. Выходными данными служат кластеры (непересекающиеся множества). Чтобы привести задачу сегментации к задаче кластеризации, нужно отобразить точки изображения в пространство признаков, а также ввести в этом пространстве признаков метрику. К недостаткам кластеризации можно отнести то, что она плохо работает на зашумленных изображениях.

Нейронные сети

Нейронные сети [12] в упрощенном виде можно представлять как способ моделирования в технических системах принципов организации и механизмов функционирования головного мозга человека. В настоящее время наиболее популярным решением проблемы семантической сегментации является использование нейронных сетей. Благодаря им точность семантической сегментации значительно увеличилась. Среди основных характеристик ИНС выделяют обучение на примерах и обобщение полученных знаний, подавление шумов, отказоустойчивость.

Результаты сравнения алгоритмов семантической сегментации приведены в табл. 1.

Выводы по результатам проведенного исследования

Для решения задачи сегментации в Duckietown было решено использовать нейросетевые алгоритмы, так как они отличаются устойчивостью и быстродействием, а также уже использовались в робототехнике и автономном вождении. Для предобработки изображений было предложено использовать классические алгоритмы компьютерного зрения. Были отобраны алгоритмы наращивания областей и алгоритмы кластеризации. Пороговые алгоритмы были отклонены, поскольку успешность применения данного метода зависит от освещения. Граничные алгоритмы не подойдут как самостоятельные алгоритмы для разметки, так как требуется не только выделить границы, но и закрасить их, однако они отлично работают в паре с другими алгоритмами.

Обзор нейронных сетей

Современная тенденция – использование глубоких нейронных сетей для сегментации изображений. В обзоре сравнивались нейронные сети, предназначенные для семантической сегментации изображений, по точности, преимуществам, ограничениям и области применения. Результат сравнения приведен в табл. 2.

Таблица 1

Алгоритмы семантической сегментации

Table 1

Semantic segmentation algorithms

Алгоритмы сегментации изображения	Устойчивость	Дополнительные ограничения	Оптимальные условия	Область применения
Граничные алгоритмы	Любой шум может повлиять на результат сегментации; низкий контраст изображения может влиять на качество обнаружения границ	Выявляет все границы, и бывает сложно найти корреляцию между интересующей областью и полученными границами	Изображения с высоким контрастом между различными областями	Используются в паре с другими алгоритмами сегментации
Пороговые алгоритмы	Успешность применения данного метода зависит от освещения и от правильно выбранного порога	Зависимость результатов сегментации от выбранной величины порога; повышенная чувствительность к шуму и неоднородностям в интенсивности	Простые изображения с высоким контрастом между различными областями	Чаще используется для выбора интересующей области изображения, игнорируя ненужные части; один из этапов в процедуре обработки МРТ-данных; робототехника
Алгоритмы наращивания областей	Наращивание области может быть чувствительно к шуму	Точность и качество сегментации зависят от выбора начальных параметров; необходимость в выборе начальной точки для каждой сегментированной области производится вручную – следовательно, метод не может быть полностью автоматизирован	Если известно расположение стартовых точек, а также грамотно выбран критерий однородности, то результаты качественные	Робототехника
Алгоритмы на основе кластеризации	Могут быть чувствительны к шуму и неоднородностям в интенсивности	Зависит от конкретного алгоритма	Зависит от конкретного алгоритма	Сегментация спутниковых изображений, сегментация цветного изображения рака желудка
Нейронные сети	Обладают повышенной устойчивостью	Зависит от конкретной нейросети	Зависит от конкретной нейросети	Практически во всех видах сегментации

Сравнение нейросетевых алгоритмов

Таблица 2

Comparison of neural network algorithms

Table 2

Сеть для сегментации	Точность, mIoU	Достоинства	Ограничения	Область применения
FCN-8	~ 57	Хорошая производительность; быстрое действие	Необходимость адаптации для сегментации; невысокая точность сегментации	Спутниковые изображения
U-Net	~ 77.5	Хорошие результаты сегментации даже на небольших датасетах; высокая точность; быстрота	Невысокая точность сегментации для некоторых задач	Duckietown; медицина; дорожные сцены
SegNet	~ 56.1	Высокая производительность; высокая скорость работы	Необходимость адаптации для сегментации; низкая точность сегментации	Duckietown; классификация типов дорожных знаков; дорожные сцены
FC-DenceNet	~ 66.9	Хорошая точность	Недостаточная точность сегментации для некоторых задач	Дорожные сцены
RefineNet	~ 73.6	Высокая точность сегментации	Невысокое быстродействие	Применяется везде, в том числе и для сегментации дорог
PSPNet	~ 81.2			Дорожные сцены; медицина
DeepLab-v3	~ 81.3			Дорожные сцены; медицина
LinkNet	~ 76.4	Работает в реальном времени	Необходим мощный процессор	Спутниковые изображения
ESPNet	~ 60.3			Дорожные сцены
ERFNet	~ 68.0			
ICNet	~ 69.5			
ENet	~ 58.3			

1. *FCN* [13] реализует архитектуру «кодер – декодер». В качестве кодера используется CNN без полносвязных слоев, выполняющая операции свертки и пулинга, декодер же использует транспонированные свертки. На входе сети – изображение произвольного размера, на выходе – карта сегментации.

Достоинствами данной сети являются хорошая производительность и ее быстрдействие. Но *FCN* имеет некоторые существенные ограничения, а именно: необходимость адаптировать нейронную сеть для сегментации, низкое разрешение карт сегментации на выходе и невысокая точность сегментации по современным меркам. Использовалась для семантической сегментации спутниковых изображений.

2. *U-Net* [14] реализует архитектуру «кодер – декодер», базируясь на *FCN*. В случае *U-Net* кодер, как и у предыдущей сети, состоит из слоев свертки и пулинга, но в отличие от *FCN* у *U-Net* на каждом этапе удваивается количество каналов карт признаков. Декодер же состоит из слоев, повышающих разрешение карт признаков и слоев свертки, которые уменьшают количество каналов признаков.

Достоинствами данной сети являются хорошие результаты сегментации даже на небольших датасетах, высокое качество карт сегментации на выходе, быстрота. Но *U-Net* имеет недостаточную точность сегментации для некоторых задач, например при определении дорожных знаков. Сфера применения *U-Net* достаточно обширная и включает медицину, дорожные сцены. Также эта сеть уже использовалась в *Duckietown*.

3. *SegNet* [15; 16] – *SegNet*, как и *U-Net*, реализует архитектуру «кодер – декодер». Кодер состоит из слоев свертки и пулинга, а декодер – из слоев анпулинга и свертки. В отличие от других сетей, в этой сети применяются более быстрые и эффективные скуп-соединения.

Достоинствами данной сети являются высокая производительность и скорость работы, а также небольшое потребление оперативной памяти компьютера. Но *SegNet* характеризуют некоторые ограничения, например необходимость адаптировать сеть для сегментации, недостаточно хорошее разрешение карт сегментации и низкая точность. Применялась для классификации типов дорожных знаков и для сегментации дороги. Также эта сеть уже использовалась в *Duckietown*.

4. *DeepLab* [17] реализует архитектуру CNN вместе с многомасштабной обработкой изображений для получения карт сегментации. Для увеличения количества захватываемой информации без увеличения числа параметров в *DeepLab* используются разряженные ядра сверток.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но *DeepLab* имеет невысокое быстрдействие. Применялась для сегментации дорог, а также в медицине для сегментирования опухолей головного мозга.

5. *FC-DenseNet*³ расширяет архитектуру *U-net*, реализованную на архитектуре «кодер – декодер». Кодер сети состоит из классификационной сверточной нейронной сети *DenseNet* без полносвязных слоев. В отличие от других сетей, в *FC-DenseNet* реализованы *DenseBlock*, а также используются длинные скуп-соединения между кодером и декодером и короткие – между отдельными этапами кодера.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и хорошая точность сегментации. При этом *FC-DenseNet* имеет недостаточную точность сегментации для некоторых задач. Применяется для сегментации дорожных сцен.

6. *RefineNet*⁴ включает в себя идеи полностью сверточных нейронных сетей и многомасштабной обработки изображений. Кодером является часть классификационной сверточной

³ Review: *FC-DenseNet – One Hundred Layers Tiramisu, Fully Convolutional DenseNet*. URL: <https://towardsdatascience.com/review-fc-densenet-one-hundred-layer-tiramisu-semantic-segmentation-22ee3be434d5> (дата обращения 31.03.2021).

⁴ Review: *RefineNet – Multi-path Refinement Network*. URL: <https://towardsdatascience.com/review-refinenet-multi-path-refinement-network-semantic-segmentation-5763d9da47c1> (дата обращения 31.03.2021).

нейронной сети ResNet. Декодер базируется на специальных блоках, предназначенных для обработки и объединения карт высокого разрешения кодера и карт низкого разрешения из предыдущего блока декодера.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но RefineNet имеет невысокое быстродействие. Применялась везде, в том числе и для сегментации дорожных сцен.

7. *PSPNet* [18] базируется на архитектуре «кодер – декодер» и многомасштабной обработке с использованием пирамиды изображений. Кодер сети основывается на нейронной сети ResNet, в которую дополнительно входят разреженные свертки. Декодер сети использует модуль пулинга пирамиды.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но PSPNet имеет невысокое быстродействие. Применялась в медицинских целях, например для сегментации изображений опухолей головного мозга, а также применялась для сегментации дорог.

8. *ENet* [19] реализует архитектуру «кодер – декодер». В качестве кодера используется сверточная нейронная сеть ResNet.

Достоинством данной сети является возможность работы в «реальном» времени. Главное ограничение ENet – необходимость иметь мощный процессор. Эта сеть разрабатывалась как раз для распознавания городских улиц.

9. *ICNet* [20] – каскадная нейронная сеть, основанная, как и RefineNet, на идеях многомасштабной обработки изображения. Архитектура ICNet базируется на топологии сети PSPNet.

Данная сеть имеет возможность работы в реальном времени. Главное ограничение ICNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

10. *LinkNet* [21; 22] основана на архитектуре «кодер – декодер». Между блоками кодера и декодера сети имеются соединения, которые суммируют соответствующие карты признаков, агрегируя информацию с текущего и предыдущих этапов работы сети.

Достоинством данной сети является возможность работы в реальном времени. Главное ограничение LinkNet – необходимость иметь мощный процессор. Использовалась для семантической сегментации спутниковых изображений

11. *ESPNet* [23], как и сеть PSPNet, реализует архитектуру, включающую в себя идеи как архитектуры «кодер – декодер», так и многомасштабной обработки изображения в виде построения пирамиды карт признаков.

Данная сеть имеет возможность работы в реальном времени. Главное ограничение ESPNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

12. *ERFNet* [24] объединяет идеи нескольких высокопроизводительных глубоких нейронных сетей, как Inception, Xception и ResNet, а также архитектуры «кодер – декодер» для решения задачи семантической сегментации на устройствах с ограниченными вычислительными ресурсами.

Достоинством данной сети является возможность работы в реальном времени. Главное ограничение ESPNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

Исходя из данных табл. 2 и накладываемых ограничений (область применения, приемлемое отношение скорости работы к точности карт сегментации, малая мощность процессора Raspberry Pi), можно сделать следующие выводы:

- FCN-8, LinkNet не подходят, так как область применения не включает ничего, связанного с дорожным движением;
- RefineNet не подходит из-за невысокого быстродействия и относительно не самой высокой точности сегментации по отношению к сетям с таким же ограничением;
- для ENet, ICNet, ESPNet, ERFNet нужен мощный процессор.

В итоге самыми подходящими стали сети U-Net, SegNet, DeepLab-v3, FC-DenceNet [25] и PSPNet [26], которые подходят по всем критериям.

Применение нейронных сетей на примере Duckietown

Ранее уже предпринимались попытки использовать нейронные сети для семантической сегментации в тренажере «Duckietown», но в обоих случаях сегментация захватывала только дорожную разметку, игнорируя дорожные знаки, других duckiebots и “жителей” города.

- U-Net

В работе «Semantic segmentation transfer in Duckietown»⁵ обучающие данные были получены из симулятора Duckietown. Датасет собран вручную. Для данной задачи в нейронной сети U-Net уменьшено количество нейронных слоев. Затем модель была обучена на датасете, размеченном с помощью Photoshop.

В качестве оптимизатора использовался стохастический градиентный спуск (SGD) с импульсом (momentum) 0,9, скоростью обучения 0,01 и уменьшением веса 10⁻⁴.

На смоделированных изображениях точность проверки пикселей к концу обучения составила 98,42 %. Тестирование модели проходило как на смоделированных, так и на реальных изображениях.

- SegNet

В работе [28] задача сегментации изображений на тренажере решалась в 2018 г. с помощью нейронной сети SegNet и GAN⁶.

Обзор фреймворков

Для семантической сегментации с помощью нейросетевых алгоритмов были исследованы библиотеки машинного обучения (табл. 3). Во всех представленных фреймворках используется язык программирования Python. Каждая библиотека является opensource проектом.

Таблица 3

Сравнение фреймворков для нейросетевых алгоритмов

Table 3

Comparison of frameworks for neural network algorithms

Фреймворк	Требования	Область применения
TensorFlow	ОС: Linux, Mac OS X, Microsoft Windows, iOS, Android	Для проектов машинного обучения. Для проектов нейронных сетей. В автоматизированном программном обеспечении для создания титров, таких как DeepDream
Keras	ОС: Linux, Mac OS X, Microsoft Windows	Для исследования глубокого обучения. Популярный среди стартапов, разрабатывающих продукты, основанные на глубоком обучении
PyTorch	ОС: Linux Mac OS X Microsoft Windows, iOS, Android	Используется для решения различных задач: компьютерное зрение, обработка естественного языка

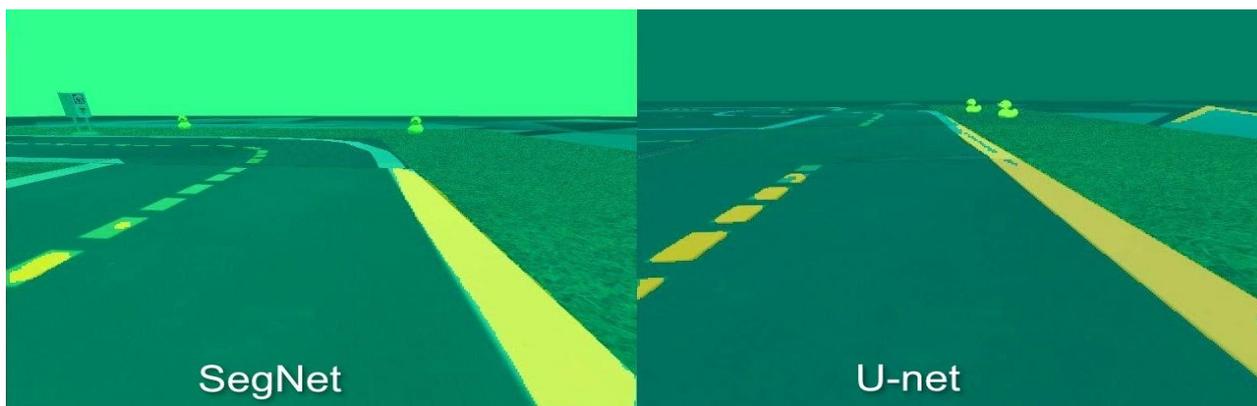
⁵ URL: <https://github.com/duckietown/segmentation-transfer> (дата обращения 31.03.2021).

⁶ Duckietown Simulator. URL: https://docs.duckietown.org/daffey/AIDO/draft/dt_simulator.html (дата обращения 31.03.2021).

В итоге были выбраны библиотеки машинного обучения Keras и PyTorch, так как порог вхождения у них намного ниже, чем у TensorFlow, а функции они выполняют те же самые.

Тестирование нейросетей SegNet и U-Net на изображениях с тренажера Duckietown

В эксперименте были получены изображения с тренажера Duckietown⁷ и сформирован датасет для обучения SegNet и U-Net. Целью эксперимента было сегментировать исходное изображение нейросетью на 2 класса – дорогу и фон. При схожих затратах по производительности, нейронная сеть U-net показала лучший результат:



Заключение

Проведен анализ применимости алгоритмов семантической сегментации в тренажере Duckietown, который имитирует автопилотируемых роботов в условиях городской среды. Были проанализированы подходы, основанные на классических алгоритмах компьютерного зрения, и подходы, основанные на нейросетевых алгоритмах. Установлено, что методы, основанные на классических алгоритмах компьютерного зрения, уступают методам, основанных на нейронных сетях, по устойчивости, точности сегментации и скорости работы. Было предложено использовать классические алгоритмы компьютерного зрения для разметки изображений и подготовки датасетов.

Для сегментации изображений на duckiebots было предложено использовать нейросети. Проведен анализ нейросетей, использующихся для сегментации, с учетом особенностей тренажера Duckietown (малая вычислительная мощность, отсутствие дополнительных датчиков, монокулярная камера).

Таким образом, для сегментации изображений были выбраны нейросетевые алгоритмы, основанные на сетях U-Net, SegNet, DeepLab-v3, FC-DenceNet и PSPNet. Для предобработки изображений выбраны классические алгоритмы компьютерного зрения, такие как алгоритмы наращивания областей и алгоритмы кластеризации. Для апробации на тренажере Duckietown были выбраны библиотеки машинного обучения Keras и PyTorch.

Лучший результата в ходе апробации получен с использованием U-net.

В дальнейшем планируется апробация DeepLab-v3, FC-DenceNet и PSPNet, а также подготовка разработанного решения сегментации изображений к развертыванию на реальном роботе.

⁷ Duckietown Simulator. URL: https://docs.duckietown.org/daffy/AIDO/draft/dt_simulator.html (дата обращения 31.03.2021).

Список литературы

1. **Sharma Y.** Adoption of next generation robotics: A case study on Amazon. Pune, 2017, 15 p.
2. **Csurka G., Larlus D., Perronnin F.** What is a good evaluation measure for semantic segmentation. Meylan, 2013, 11 p.
3. **Зуенко Е. А., Шулунова А. А.** Автоматическое выделение боковых желудочков головного мозга на МРТ изображениях // Вестник НГУ. Серия: Информационные технологии. 2017. Т. 15, № 4. С. 22–31. DOI 10.25205/1818-7900-2017-15-4-22-31
4. **Мусатян С. А., Ломакин А. В., Саргасов С. Ю., Попыванов Л. К., Монахов И. Б., Чижова А. С.** Способы сегментации медицинских изображений // Тр. Ин-та системного программирования РАН. М., 2018. DOI 10.15514/ISPRAS-2018-30(4)-12
5. **Горбачева В. А., Криворотова И. А., Маркелов А. О., Котлярова Е. В.** Семантическая сегментация спутниковых снимков аэропортов с помощью свёрточных нейронных сетей. Мю, 2019. 10 с.
6. **Ozmen A., Akman E. T.** Edge detection using steerable filters and CNN. In: 11th European Signal Processing Conference. Toulouse, 2002, pp. 1–4.
7. **Pestunov I., Rylov S., Berikov V.** Hierarchical Ensemble Clustering Algorithm for Multispectral Image Segmentation. In: Proceedings 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-2014). Koblenz, Uni. of Koblenz-Landau Press, 2015, pp. 123–127.
8. **Рылов С. А., Мельников П. В., Пестунов И. А.** Спектрально-текстурная классификация гиперспектральных изображений высокого пространственного разрешения // Интерэкспо ГЕО-Сибирь. 2016. Т. 4, № 1. С. 78–84.
9. **Чичварин Н. В.** Пороговые методы. М., 2016.
10. **Луценко М. Т., Ульянычев Н. В., Семичевская Н. П.** Автоматизированная система распознавания объектов на микроскопических изображениях биологических образцов. Благовещенск, 1999.
11. **Пестунов И. А., Синявский Ю. Н.** Алгоритмы кластеризации в задачах сегментации спутниковых изображений // Вестник КемГУ. 2012. Т. 52, № 4/2. С. 110–125.
12. **Li B., Shi Y., Qi Z., Chen Z.** A Survey on Semantic Segmentation. Beijing, 2018. DOI 10.1109/ICDMW.2018.00176
13. **Noh H., Hong S., Han B.** Learning deconvolution network for semantic segmentation. Pohang, 2015.
14. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional Networks for Biomedical Image Segmentation. Berlin, 2015.
15. **Badrinarayanan V., Kendall A., Cipolla R.** SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 12, pp. 2481–2495. DOI 10.1109/TPAMI.2016.2644615
16. **Емельянов А. В.** Анализ методов семантической сегментации изображений на основе нейронных сетей. М., 2019.
17. **Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. L.** DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, vol. 40, no. 4, pp. 834–848. DOI 10.1109/TPAMI.2017.2699184
18. **Zhou J., Hao M., Zhang D., Zou P., Zhang W.** Fusion PSPnet Image Segmentation Based Method for Multi-Focus Image Fusion. *IEEE Photonics Journal*, 2019, vol. 11, no. 6, pp. 1–12. DOI 10.1109/JPHOT.2019.2950949
19. **Paszke A.** ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. Ithaca, 2016, 10 p.
20. **Zhao H., Qi X., Shen X., Shi J., Jia J.** ICNet for Real-Time Semantic Segmentation on High-Resolution Images. Hong Kong, 2018, 16 p.

21. **Chaurasia A., Culurciello E.** LinkNet: Exploiting encoder representations for efficient semantic segmentation. *EEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4. DOI 10.1109/VCIP.2017.8305148
22. **Емельянов А. В., Галиахметов Д. Г.** Анализ методов семантической сегментации изображений в реальном времени на основе нейронных сетей. М., 2020.
23. **Mehta S., Rastegari M., Caspi A., Shapiro L., Hajishirzi H.** ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation. Seattle, 2018, 29 p.
24. **Romera E., Alvarez J. M., Bergasa L. M., Arroyo R.** ERFNet: Efficient Residual Factorized ConvNet for Real-time Semantic Segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 19, no. 1, pp. 263–272. DOI 10.1109/TITS.2017.2750080
25. **Brahimi S., Aoun N. B., Chokri B. A., Benoit A., Lambert P.** Multiscale Fully Convolutional DenseNet for Semantic Segmentation. Pilsen, 2018.
26. **Li X.** Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. Philadelphia, 2020, 16 p.
27. **Tóth Z.** Deep Learning-based Semantic Segmentation in Simulation and Real-World for Autonomous Vehicles. URL: <https://smartlabai.medium.com/deep-learning-based-semantic-segmentation-in-simulation-and-real-world-for-autonomous-vehicles-e7fe25cef816> (accessed 31.03.2021).
28. **Айрапетов А. Э., Коваленко А. А.** Исследование генеративно-состязательной сети // Политехнический молодежный журнал МГТУ им. Н. Э. Баумана. 2018. № 10. DOI 10.18698/2541-8009-2018-10-380

References

1. **Sharma Y.** Adoption of next generation robotics: A case study on Amazon. Pune, 2017, 15 p.
2. **Csurka G., Larlus D., Perronnin F.** What is a good evaluation measure for semantic segmentation. Meylan, 2013, 11 p.
3. **Zuenko E. A., Shulunova A. A.** Automated Segmentation of the Lateral Ventricles from MRI Image. *Vestnik NSU. Series: Information Technologies*, 2017, vol. 15, no. 4, pp. 22–31. (in Russ.) DOI 10.25205/1818-7900-2017-15-4-22-31
4. **Musatyan S. A., Lomakin A. V., Sartasov S. Yu., Popyvanov L. K., Monakhov I. B., Chizhova A. S.** Methods for segmentation of medical images. In: Proceedings of the Institute for System Programming of the Russian Academy of Sciences. Moscow, 2018. (in Russ.) DOI 10.15514/ISPRAS-2018-30(4)-12
5. **Gorbacheva V. A., Krivorotova I. A., Markelov A. O., Kotlyarova E. V.** Semantic segmentation of satellite images of airports using convolutional neural networks. Moscow, 2019. 10 p. (in Russ.)
6. **Ozmen A., Akman E. T.** Edge detection using steerable filters and CNN. In: 11th European Signal Processing Conference. Toulouse, 2002, pp. 1–4.
7. **Pestunov I., Rylov S., Berikov V.** Hierarchical Ensemble Clustering Algorithm for Multispectral Image Segmentation. In: Proceedings 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-2014). Koblenz, Uni. of Koblenz-Landau Press, 2015, pp. 123–127.
8. **Rylov S. A., Melnikov P. V., Pestunov I. A.** Spectral and texture classification of high spatial resolution hyperspectral images. *Interexpo GEO-Siberia*, 2016, vol. 4, no. 1, pp. 78–84. (in Russ.)
9. **Chichvarin N. V.** Threshold methods. Moscow, 2016. (in Russ.)
10. **Lutsenko M. T., Ulyanychev N. V., Semichevskaya N. P.** Automatic system for object recognition in microscopic images of biological samples. Blagoveshchensk, 1999. (in Russ.)
11. **Pestunov I. A., Sinyavsky Yu. N.** Clustering algorithms in satellite images segmentation tasks. *Vestnik KSU*, 2012, vol. 52, no. 4/2, pp. 110–125. (in Russ.)

12. **Li B., Shi Y., Qi Z., Chen Z.** A Survey on Semantic Segmentation. Beijing, 2018. DOI 10.1109/ICDMW.2018.00176
13. **Noh H., Hong S., Han B.** Learning deconvolution network for semantic segmentation. Pohang, 2015.
14. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional Networks for Biomedical Image Segmentation. Berlin, 2015.
15. **Badrinarayanan V., Kendall A., Cipolla R.** SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 12, pp. 2481–2495. DOI 10.1109/TPAMI.2016.2644615
16. **Emelyanov A. V.** Analysis of Semantic Image Segmentation Methods Based on Neural Networks. Moscow, 2019. (in Russ.)
17. **Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. L.** DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, vol. 40, no. 4, pp. 834–848. DOI 10.1109/TPAMI.2017.2699184
18. **Zhou J., Hao M., Zhang D., Zou P., Zhang W.** Fusion PSPnet Image Segmentation Based Method for Multi-Focus Image Fusion. *IEEE Photonics Journal*, 2019, vol. 11, no. 6, pp. 1–12. DOI 10.1109/JPHOT.2019.2950949
19. **Paszke A.** ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. Ithaca, 2016, 10 p.
20. **Zhao H., Qi X., Shen X., Shi J., Jia J.** ICNet for Real-Time Semantic Segmentation on High-Resolution Images. Hong Kong, 2018, 16 p.
21. **Chaurasia A., Culurciello E.** LinkNet: Exploiting encoder representations for efficient semantic segmentation. *EEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4. DOI 10.1109/VCIP.2017.8305148
22. **Emelyanov A. V., Galiahmetov D. G.** Analysis of Semantic Image Segmentation Methods in Real Time Based on Neural Networks. Moscow, 2020. (in Russ.)
23. **Mehta S., Rastegari M., Caspi A., Shapiro L., Hajishirzi H.** ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation. Seattle, 2018, 29 p.
24. **Romera E., Alvarez J. M., Bergasa L. M., Arroyo R.** ERFNet: Efficient Residual Factorized ConvNet for Real-time Semantic Segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 19, no. 1, pp. 263–272. DOI 10.1109/TITS.2017.2750080
25. **Brahimi S., Aoun N. B., Chokri B. A., Benoit A., Lambert P.** Multiscale Fully Convolutional DenseNet for Semantic Segmentation. Pilsen, 2018.
26. **Li X.** Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. Philadelphia, 2020, 16 p.
27. **Tóth Z.** Deep Learning-based Semantic Segmentation in Simulation and Real-World for Autonomous Vehicles. URL: <https://smartlabai.medium.com/deep-learning-based-semantic-segmentation-in-simulation-and-real-world-for-autonomous-vehicles-e7fe25cef816> (accessed 31.03.2021).
28. **Ajrapetov A. E., Kovalenko A. A.** Research of the generative adversarial network. *Polytechnic Youth Journal of the Moscow State Technical University. N. E. Bauman*, 2018, no. 10. (in Russ.) DOI 10.18698/2541-8009-2018-10-380

Информация об авторах

Дарья Евгеньевна Шабалина, студент бакалавриата, 3 курс
Кристина Сергеевна Ланчуковская, студент бакалавриата, 3 курс
Татьяна Викторовна Лях, кандидат технических наук
Константин Владимирович Чайка, аспирант 3 года обучения

Information about the Authors

Daria E. Shabalina, Bachelor Student
Kristina S. Lanchukovskaya, Bachelor Student
Tatyana V. Lyakh, Candidate of Sciences (Engineering)
Konstantin V. Chaika, Post-Graduate Student

*Статья поступила в редакцию 20.06.2021;
одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021
The article was submitted 20.06.2021;
approved after reviewing 01.08.2021; accepted for publication 01.08.2021*