

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Д.В. Лукашик, Московский технический университет связи и информатики, dan4eg777@gmail.com.

УДК 004.852

Аннотация. В рамках данной статьи рассматривается понятие сегментации изображения, исследуются различные типы задач сегментации, такие как: семантическая сегментация, сегментация экземпляра и паноптическая сегментация. Производится анализ различных сфер применения сегментации изображений, рассматриваются классические методы сегментации, выявляются их преимущества, недостатки и сферы применения. Также исследуются современные методы сегментации, основанные на глубоком обучении.

Ключевые слова: сегментация изображения; сверточные нейронные сети; SegNet; U-Net; DeepLab.

ANALYSIS OF MODERN IMAGE SEGMENTATION METHODS

Denis Lukashik, Moscow Technical University of Communications and Informatics.

Annotation. Within the framework of this work, the concept of image segmentation is considered, various types of segmentation problems are studied, such as: semantic segmentation, instance segmentation and panoptical segmentation. An analysis is made of various areas of application of image segmentation. Next, classical segmentation methods are considered, their advantages, disadvantages and areas of application are identified. After that, modern segmentation methods based on deep learning are explored.

Keywords: image segmentation; convolutional neural networks; SegNet; U-Net; DeepLab.

Введение

В связи с развитием в области искусственного интеллекта и инновациями в глубоком обучении и нейронных сетях, компьютерное зрение получило сильное развитие в последнее время. Сегментация является важной задачей, решение которой люди пытаются найти различными способами, в том числе с использованием нейронных сетей.

В рамках данной статьи проводится исследование задачи сегментации изображений, а именно рассматривается понятие сегментации изображений, видов и сфер ее применения. Автором проводится исследование традиционных методов сегментации и выявляются их преимущества и недостатки, а также осуществляется исследование современных архитектур нейронных сетей для сегментации изображений.

Сегментация изображения

Сегментация изображения – это метод, при котором цифровое изображение разбивается на различные подгруппы, называемые сегментами изображения, что помогает уменьшить сложность изображения и упростить дальнейшую обработку или анализ изображения. Сегментация – это присвоение меток пикселям. Все

элементы изображения или пиксели, принадлежащие к одной и той же категории, имеют общую метку, назначенную им.

Сегментация изображений является расширением классификации изображений, где помимо классификации мы выполняем локализацию. Таким образом, сегментация изображения представляет собой надмножество классификации изображений, при этом модель точно определяет, где находится соответствующий объект, очерчивая его границы [1].

На рис. 1 представлены различные задачи компьютерного зрения.



Рисунок 1

Типы задач сегментации изображений

Задачи сегментации изображения можно разделить на три группы в зависимости от количества и типа информации, которую они передают.

В то время как семантическая сегментация определяет широкую границу объектов, принадлежащих к определенному классу, сегментация экземпляра предоставляет карту сегментов для каждого объекта, который он просматривает на изображении, без какого-либо представления о классе, к которому принадлежит объект [2].

Паноптическая сегментация, безусловно, является наиболее информативной, поскольку она представляет собой сопряжение задач экземплярной и семантической сегментации. Паноптическая сегментация дает нам карты сегментов всех объектов любого конкретного класса, присутствующих на изображении.

Рассмотрим эти задачи более подробно.

1. Семантическая сегментация.

Семантическая сегментация относится к классификации пикселей изображения по семантическим классам. Пиксели, принадлежащие к определенному классу, просто относятся к этому классу без учета другой информации или контекста.

Как и следовало ожидать, применение данного метода является нецелесообразным, когда на изображении тесно сгруппированы несколько экземпляров одного и того же класса. Изображение толпы на улице будет иметь модель семантической сегментации, предсказывающую, что вся область толпы принадлежит к классу «пешеход», таким образом, предоставляя очень мало подробных деталей или информации об изображении.

2. Сегментация экземпляра.

Модели сегментации экземпляров классифицируют пиксели по категориям на основе «экземпляров», а не классов.

Алгоритм сегментации экземпляров не знает к какому классу принадлежит классифицируемая область, но может разделить перекрывающиеся или очень похожие области объекта на основе их границ.

Если то же самое изображение толпы, о котором говорилось ранее, передается модели сегментации экземпляра, модель сможет отделить каждого человека от толпы, а также от окружающих объектов, но не сможет предсказать, что каждый регион/объект является экземпляром.

3. Паноптическая сегментация.

Паноптическая сегментация, самая последняя разработанная задача сегментации, может быть выражена как комбинация семантической сегментации и сегментации экземпляров, при которой каждый экземпляр объекта на изображении выделяется и предсказывается идентичность объекта.

Алгоритмы паноптической сегментации находят широкое применение в таких популярных задачах, как самоуправляемые автомобили, где необходимо захватить огромное количество информации о ближайшем окружении с помощью потока изображений.

На рис. 2 представлены задачи сегментации изображений.

Semantic Segmentation vs. Instance Segmentation vs. Panoptic Segmentation

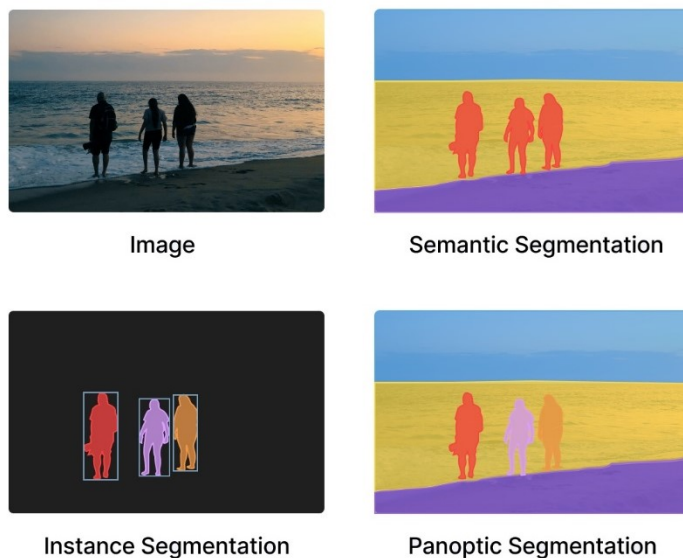


Рисунок 2

Сферы применения сегментации изображений

Сегментация изображения является важной областью компьютерного зрения. Машины должны разделить визуальные данные на сегменты для обработки, специфичной для сегмента.

Таким образом, сегментация изображений находит свое применение в таких важных областях, как робототехника, медицинская визуализация, автономные транспортные средства и интеллектуальная видеоаналитика [3].

Помимо этих приложений, сегментация изображений также используется спутниками на аэрофотоснимках для сегментации дорог, зданий и деревьев.

Ниже приведены несколько наиболее популярных сфер применения сегментации изображений.

1. Медицинская визуализация.

Медицинская визуализация – важная область компьютерного зрения, которая фокусируется на диагностике заболеваний на основе визуальных данных, как в виде простых визуальных данных, так и в виде биомедицинских сканирований.

Сегментация играет важную роль в медицинской визуализации, поскольку она помогает врачам быстро и точно выявлять возможные злокачественные признаки на изображениях.

Используя сегментацию изображений, диагностику заболеваний можно не только ускорить, но и удешевить, что принесет пользу тысячам людей по всему миру.

- Рентгеновская сегментация.
- КТ сегментация органов.
- Сегментация стоматологического экземпляра.
- Сегментация клеток цифровой патологии.
- Аннотация к хирургическому видео.

2. Умные города.

Умные города часто имеют камеры видеонаблюдения для наблюдения в режиме реального времени за пешеходами, дорожным движением и преступностью. Этот мониторинг можно легко автоматизировать с помощью сегментации изображения.

Благодаря мониторингу на основе искусственного интеллекта о преступлениях можно сообщать быстрее, на место дорожно-транспортного происшествия можно немедленно вызвать скорую помощь, а автомобили с превышением скорости легко отследить.

Таким образом, использование сегментации изображений и мониторинга на основе ИИ может улучшить образ жизни людей, в части:

- Обнаружения пешеходов.
- Аналитики трафика.
- Распознавания номерного знака.
- Видеонаблюдения.

3. Самоуправляемые автомобили.

Самоуправляемые автомобили – одно из самых больших приложений сегментации изображений, от которого сильно зависит планирование маршрутов и движения.

Семантическая сегментация и сегментация экземпляров помогают этим транспортным средствам идентифицировать дорожные знаки и другие транспортные средства, тем самым обеспечивая беспрепятственную и плавную езду за счет:

- Семантической сегментации дорожного покрытия.
- Сегментации автомобилей и пешеходов.
- Обнаружения объектов в автомобиле (вещи, оставленные пассажирами).
- Обнаружения выбоин.

Традиционные методы сегментации изображений

Первоначально сегментация изображений началась с цифровой обработки изображений в сочетании с алгоритмами оптимизации. Эти примитивные алгоритмы использовали такие методы, как выращивание областей и алгоритм змей [4], где они устанавливали начальные области, а алгоритм сравнивал значения пикселей, чтобы получить представление о карте сегментов.

Эти методы брали локальное представление о функциях изображения и фокусировались на локальных различиях и градиентах в пикселях.

Алгоритмы, которые использовали глобальное представление входного изображения, появились намного позже, когда среди классических методов обработки изображений были предложены такие методы, как адаптивная пороговая обработка, алгоритм Оцу и алгоритмы кластеризации [5].

- Методы на основе порога.

Пороговое определение – один из самых простых методов сегментации изображения, при котором устанавливается пороговое значение для разделения пикселей на два класса. Пиксели, значения которых превышают пороговое значение, устанавливаются равными 1, а пиксели, значения которых меньше порогового значения, устанавливаются равными 0.

Таким образом, изображение преобразуется в двоичную карту, что приводит к процессу, который часто называют бинаризацией. Пороговое значение изображения очень полезно, если разница в значениях пикселей между двумя целевыми классами очень велика, и легко выбрать среднее значение в качестве порога.

Пороговое значение часто используется для бинаризации изображения, чтобы можно было использовать дополнительные алгоритмы, такие как обнаружение контура и идентификация, которые работают только с бинарными изображениями [6].

- Сегментация по регионам.

Алгоритмы сегментации на основе областей работают, ища сходство между соседними пикселями и группируя их в общий класс.

Как правило, процедура сегментации начинается с того, что некоторые пиксели устанавливаются в качестве исходных пикселей, а алгоритм работает, обнаруживая непосредственные границы исходных пикселей и классифицируя их как похожие или несходные.

Затем рассматриваются непосредственные соседи и шаги повторяются до тех пор, пока все изображение не будет сегментировано. Примером подобного алгоритма является популярный алгоритм водораздела для сегментации, который работает, начиная с локальных максимумов карты евклидовых расстояний и растет при условии, что никакие два начальных числа не могут быть классифицированы как принадлежащие к одной и той же области или карте сегмента [7].

- Сегментация краев.

Сегментация краев, также называемая обнаружением краев, представляет собой задачу обнаружения краев на изображениях.

С точки зрения сегментации можно сказать, что обнаружение краев соответствует классификации пикселей изображения, которые являются краевыми пикселями, и соответственно выделению этих краевых пикселей в отдельный класс.

Обнаружение краев обычно выполняется с помощью специальных фильтров, которые дают края изображения после свертки. Эти фильтры рассчитываются с помощью специальных алгоритмов, которые оценивают градиенты изображения в координатах x и y пространственной плоскости.

- Сегментация на основе кластеризации.

Современные процедуры сегментации, которые зависят от методов обработки изображений, обычно используют алгоритмы кластеризации для сегментации.

Алгоритмы кластеризации работают лучше, чем их аналоги, и могут предоставлять достаточно качественные сегменты за небольшой промежуток времени. Популярные алгоритмы, такие как алгоритмы кластеризации K -средних, являются неконтролируемыми алгоритмами, которые работают путем кластеризации пикселей с общими атрибутами вместе как принадлежащих определенному сегменту.

Кластеризация K -средних, в частности, учитывает все пиксели и группирует их в k классы. В отличие от методов наращивания регионов, методы на основе кластеризации не требуют исходной точки для начала сегментации [8].

Методы, основанные на глубоком обучении

В настоящее время наилучшие результаты в области сегментации изображений показывают методы, основанные на машинном обучении.

В общем случае модели сегментации предоставляют карты сегментов в качестве выходных данных, соответствующие входным данным.

Эти карты сегментов часто являются n -канальными, где n – это количество классов, которые модель должна сегментировать. Каждый из этих n -каналов является бинарным по своей природе, при этом местоположения объектов «заполнены» единицами, а пустые области состоят из нулей (рис. 3) [9]. На рис. 3 показана сегментация изображения.

Нейронные сети, выполняющие сегментацию, обычно используют структуру кодер-декодер, в которой за кодером следует узкое место, а декодер или уровни повышения дискретизации следуют непосредственно после узкого места (рис. 4).

Архитектуры кодер-декодер для семантической сегментации стали популярными с появлением такой архитектуры, как *SegNet* в 2015 г. [10].

SegNet предлагает использовать комбинацию блоков свертки и понижающей дискретизации, чтобы сформировать представление входных данных. Затем декодер реконструирует входную информацию, чтобы сформировать карту сегментов, выделяющую регионы на входе и группирующие их по классам.

Наконец, декодер имеет сигмовидную функцию активации в конце, которая сжимает выходные значения в диапазоне (0,1). На рис. 4 показана архитектура кодер-декодер.

SegNet сопровождался выпуском еще одной независимой работы по сегментации в то же время, *U-Net* [11], которая впервые использовала пропуски соединений в глубоком обучении как решение проблемы потери информации, наблюдаемой в слоях понижения дискретизации типичных сети кодер-декодер.

Пропускные соединения (*Skip Connections*) – это соединения, которые идут от кодера напрямую к декодеру, минуя узкое место [12].

An overview of Semantic Image Segmentation

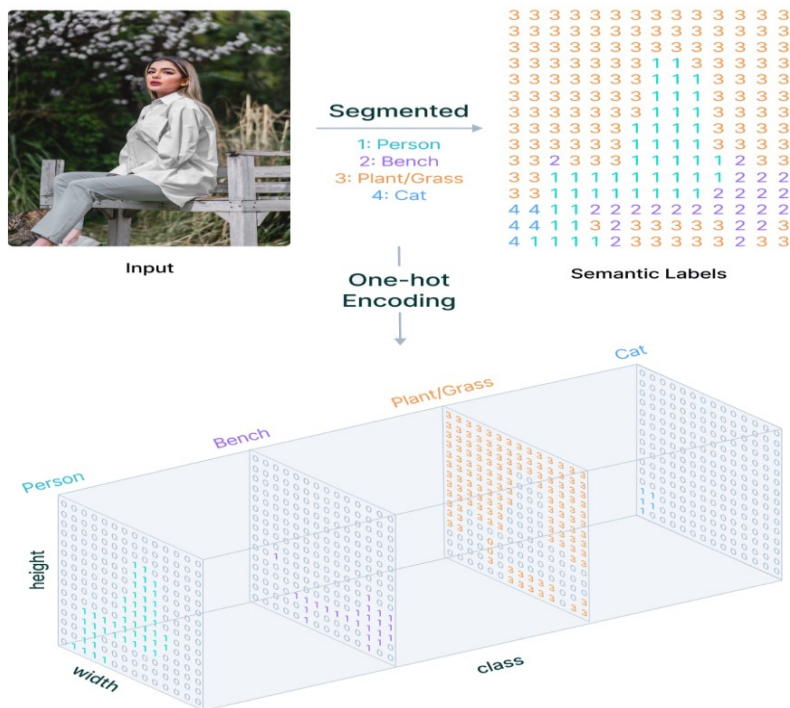


Рисунок 3

Convolutional encoder-decoder

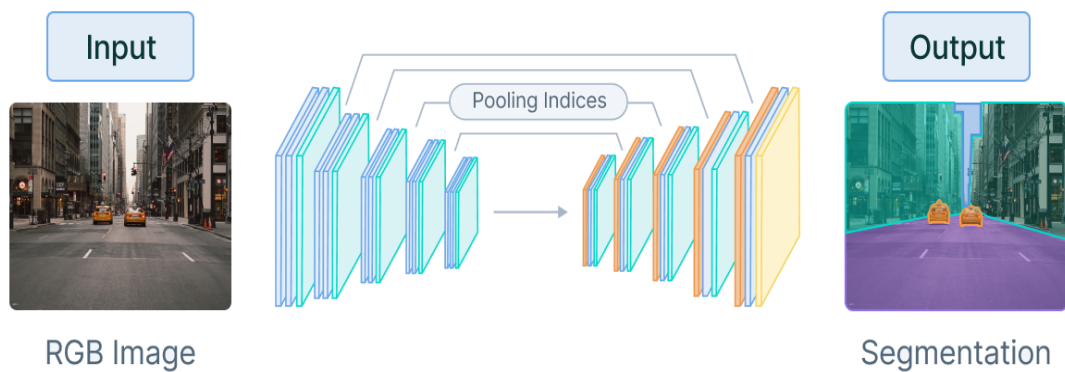


Рисунок 4

Другими словами, карты признаков на различных уровнях закодированных представлений фиксируются и объединяются в карты признаков в декодере. Это помогает уменьшить потерю данных за счет агрессивного объединения и

понижения частоты дискретизации, как это делается в блоках кодировщика архитектуры кодер-декодер.

Skip Connections пользуется большим успехом, особенно в области медицинской визуализации, поскольку *U-Net* предоставила самые современные результаты в области сегментации клеток для диагностики заболеваний (рис. 5). На рис. 5 изображена архитектура *U-Net*.

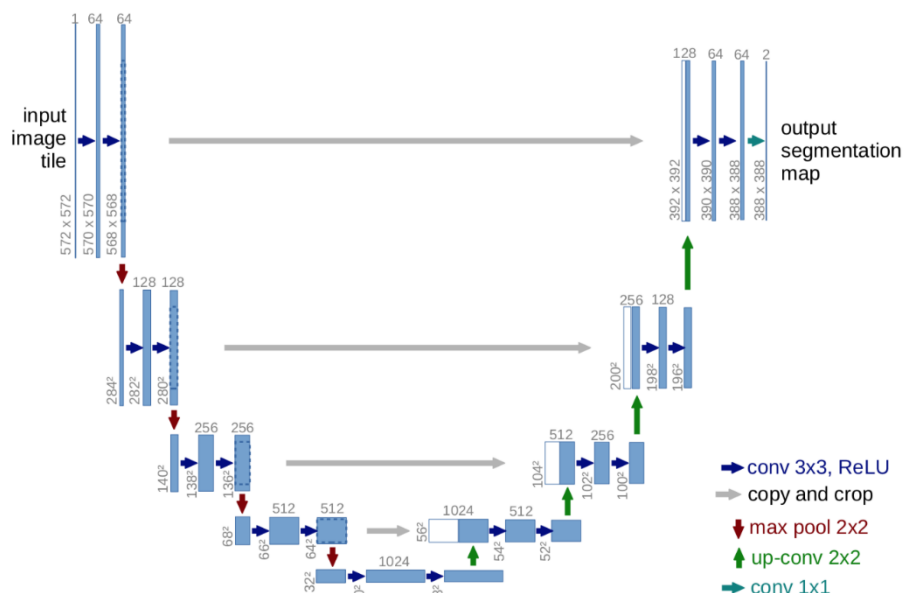


Рисунок 5

Вслед за *U-Net* вехой появилась архитектура *DeepLab*, предоставившая самые современные результаты по семантической сегментации. [13]

DeepLab использует сложные свертки, заменив простые операции объединения и предотвратив значительную потерю информации при даунсемплинге. Кроме того, реализовано многомасштабное извлечение признаков с помощью *Atrous Spatial Pyramid Pooling*, что помогает обнаруживать сегмент независимо от его размера [14].

Чтобы восстановить информацию о границах, одну из наиболее важных частей семантической сегментации, а также сегментации экземпляров, используется полностью связанные условные случайные поля (*CRF*).

Сочетание высокой точности локализации *CRF* и способности распознавания сверточных нейронных сетей (*CNN*) привело к тому, что *DeepLab* предоставляет высокоточные карты сегментов, значительно превосходящие по точности такие методы, как *FCN* [15] и *SegNet*.

Литература

1. URL https://neerc.ifmo.ru/wiki/index.php?title=Сегментация_изображений (дата обращения – март 2022 г.).
2. URL <https://habr.com/ru/company/samsung/blog/508342/> (дата обращения – март 2022 г.).
3. URL <https://vc.ru/ml/166105-kompyuternoe-zrenie-zadachi-oblasti-primeneniya-perspektivy> (дата обращения – март 2022 г.).
4. Исрафилов Х.С., Исследование методов бинаризации изображений, 2017. – С. 8.

5. URL <http://home.iitj.ac.in/~manpreet.bedi/btp/documents/sar.pdf> (дата обращения – март 2022 г.).
6. URL [https://ru.bmstu.wiki/Пороговые_методы_\(Сегментация\)](https://ru.bmstu.wiki/Пороговые_методы_(Сегментация)) (дата обращения – март 2022 г.).
7. URL [https://ru.bmstu.wiki/Метод_водоразделов_\(Сегментация\)](https://ru.bmstu.wiki/Метод_водоразделов_(Сегментация)) (дата обращения – март 2022 г.).
8. URL <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1> (дата обращения – март 2022 г.).
9. URL <https://www.v7labs.com/blog/image-segmentation-guide> (дата обращения – март 2022 г.).
10. Vijay Badrinarayanan, Alex Kendall, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, 2016. – С 14.
11. Olaf Ronneberger, Philipp Fischer, U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015. – С. 8.
12. URL <https://ichi.pro/ru/obzor-u-net-resnet-vaznost-soedinenij-s-dlinnym-i-korotkim-propuskom-segmentacia-biomedicinskih-izobrazenij-225122864725827> (дата обращения – март 2022 г.).
13. Liang-Chieh Chen, George Papandreou, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, 2017. – С. 14.
14. XuhangLiana, YanweiPang, Cascaded hierarchical atrous spatial pyramid pooling module for semantic segmentation, 2020.
15. Jonathan Long, Evan Shelhamer, Fully Convolutional Networks for Semantic, 2015. – С. 10.

АНАЛИЗ РАЗЛИЧНЫХ ИНСТРУМЕНТОВ УПРАВЛЕНИЯ И МОНИТОРИНГА ОБЛАЧНОЙ ИНФРАСТРУКТУРОЙ

*А.И. Воронцов, Московский финансово-промышленный университет «Синергия»,
160297@mail.ru;*

*А.А. Бусенков, Национальный исследовательский университет «Высшая школа
экономики», albus.rank@yandex.ru;*

*О.Д. Куприков, Московский технический университет связи и информатики,
kod808@yandex.ru.*

УДК 004.77:658

Аннотация. Поскольку в последние годы облачные вычисления набирают обороты, с их развитием растет потребность в инструментах управления и цифрового мониторинга облачной инфраструктурой. Эти инструменты помогают эффективно управлять частными, общедоступными и гибридными облачными средами. Программные средства помогают облегчить доступ к программным и аппаратным ресурсам, вычислительной мощности и доступу к хранилищам. Инструменты управления и мониторинга позволяют оперировать мультиоблачной платформой из одного места. Большинство организаций выбирают платформу управления облаками для оптимизации использования облаков и управления общими затратами на них. Все доступные инструменты облачного мониторинга помогают в сборе данных, но используют разные подходы, предоставляя разные