Особенность второго подхода заключается в выборе альтернатив устному общению, а именно, текст, выводимый на экран.

В целях профилактики и развития таких познавательных процессов учащихся как внимание, мышление и память было разработано специальное программное обеспечение, способное предотвращать ограничение возможностей читать и писать.

Использование информационных технологий в инклюзивном образовании делает образовательный процесс живее и интереснее, поскольку учащийся уже не ущемлен пространственно-временными границами, у него появляется связь с миром. Из чего следует, что для учащихся с ограниченными возможностями здоровья применение информационных технологий в образовательном процессе - прекрасная возможность самореализации.

Применение информационных технологий в инклюзивном образовании способствует облегчению процесса обучения, положительно влияет на интеллектуальное и лингвистическое развитие учащихся с ограниченными возможностями, способствует творческой активности детей, вызывает любопытство и повышает интерес к учебной деятельности. Информационные технологии в инклюзивном образовании гарантируют лишь прогресс при доступе к информации и обеспечивают содействие коммуникации, вследствие чего могут стать фундаментальным дидактическим и коммуникативным инструментом, закладывающим основу для значительного развития индивида, что, в свою очередь, создает условия для участия в жизни общества людей с особыми образовательными потребностями, но ограниченными возможностями.

1. Дозорова Л.В. Основные преимущества использования ИКТ в инклюзивном образовании / Л.В. Дозорова, Е.А. Гурова // Приоритетные направления развития науки и образования: материалы VIII Междунар. науч.-практ. конф. (Чебоксары, 29 янв. 2016 г.) / редкол.: О.Н. Широков [и др.] - Чебоксары: ЦНС «Интерактив плюс», 2016. – С. 117-119.
2. Тринитатская О.Г. Управление развивающей средой в условиях инклюзивного образования: психолого-педагогический ракурс / О.Г. Тринитатская, Н.П. Эпова. – Ростов н/Д, 2015. – 244 с.
3. Юрьева А.А. Применение интерактивных технологий в инклюзивном образовании /А.А.Юрьева// Молодой ученый. - 2018. - №16. - С. 322-323.

Семенова А.Н., Казиева М.Х., Болурова М.И., Салпагарова М.У.
Современные методы локализации объектов на изображении

*Северо-Кавказская государственная академия
(Россия, Черкесск)*

doi: 10.18411/lj-12-2019-28

idsp: ljjournal-12-2019-28

Аннотация

Проблему локализации можно охарактеризовать как проблему разбиения обрабатываемого изображения на изображения объектов, подлежащие дальнейшему признаковому анализу и идентификации автоматической системой распознавания. Из-за недостаточной формализации понятия «изображение» в настоящее время данная проблема относится к числу недостаточно формализованных и обычно рассматривается для изображений из заданной предметной области или решается в контексте распознавания заданных объектов. Тем не менее в случае простейших изображений объектов на однородном фоне она, как правило, не возникает. Поэтому имеет непосредственный смысл разработка алгоритмов, которые позволили бы установить взаимосвязь решения общей проблемы локализации объектов с решением частной задачи их выделения на однородном фоне.

Ключевые слова: классификация изображений, локализация объектов, семантическая сегментация, компьютерное зрение

Abstract

The problem of localization can be described as the problem of dividing the processed image into images of objects that are subject to further feature analysis and identification by an automatic recognition system. Due to insufficient formalization of the concept of “image”, this problem is currently classified as insufficiently formalized and is usually considered for images from a given subject area or is solved in the context of recognition of given objects. Nevertheless, in the case of the simplest images of objects on a uniform background, as a rule, it does not occur. Therefore, it makes direct sense to develop algorithms that would make it possible to establish a relationship between the solution of the general problem of localization of objects and the solution of the particular problem of their allocation against a uniform background.

Keywords: image classification, localization of objects, semantic segmentation, computer vision

Классификация изображений должна распознавать, какие именно объекты находятся на визуальной сцене (рисунок 1 (a)), а обнаружение объектов выявляет, где именно находятся объекты.

Рисунок 1

В этой статье мы сосредоточимся на проблеме того, как именно объекты точно представляются в визуальной сцене с помощью сегментации изображения. Сегментация изображения рассматривается как классификация на уровне пикселей, целью которой является разделение изображения на значимые области путем классификации каждого пикселя в конкретный объект. Технологии глубокого обучения способствовали крупномасштабной контролируемой классификации, переходящей от классификации объектов на уровне изображений к локализации объектов на уровне блоков и далее к пиксельной сегментации объектов. Следовательно, современная сегментация изображения является объектно-ориентированной и может быть разделена на две ветви: (i) семантическая сегментация, которая назначает каждый пиксель изображения классу семантического объекта, как показано на рисунке 2 (c) и (ii) сегментация экземпляра, которая предсказывает разные метки для разных экземпляров объекта в качестве дальнейшего улучшения семантической сегментации, как показано на рисунке 2 (d).

Рисунок 2

В дополнение к классификации и обнаружению, соревнования PASCAL-VOC 2012 и MS-COCO также проводят соревнования по сегментации. Это два важных набора данных для сегментации изображения, по которым оценивается большинство исследовательских работ. Популярные метрики для оценки точности сегментации на уровне пикселей - это точность пикселей (pixel accuracy (PA), т. е. пропорция пикселей, предсказанная правильно), средняя точность пикселей (mPA) для всех классов, среднее пересечение по объединению (mIOU) и взвешенное по частоте пересечение по объединению (FWIOU). Среди них mIOU часто предпочитают из-за своей простоты и представительности.

Ранние методы сегментации, основанные на глубоком обучении, обычно применяют сверточные нейронные сети (CNN) в качестве дескриптора функции для каждого пикселя, который описывается окружающим патчем. Эта основанная на CNN структура является проблематичной по эффективности и недостаточно точной для извлечения избыточных признаков. Полностью сверточная сеть (FCN) [81] является предвестником, который успешно реализует плотные по пикселям предсказания для семантической сегментации в сквозной структуре CNN. Он заменяет полностью связанные слои известных архитектур классификации (например, VGG [22], GoogleNet [23] и т. д.). Потери при прогнозировании затем измеряются по пиксельным потерям между увеличенной тепловой картой, используя деконволюцию, и маркированным изображением оригинального размера. FCN демонстрирует значительное улучшение точности пикселей по сравнению с традиционными методами сегментации в PASCAL-VOC 2012. Однако базовая структура FCN не в состоянии охватить большое количество функций и не учитывает пространственную согласованность между пикселями, что затрудняет ее применение для определенных проблем и сценариев. В любом случае, успех архитектуры FCN делает ее популярной, и за ней последовали многие последующие работы по сегментации.

Обычно использование классификационных моделей без полностью связанных слоев в качестве магистральной сети для создания карт характеристик низкого разрешения называется *кодером*, в то время как симметричное отображение из изображения с низким разрешением в пиксельный результат классификации называется *декодером*. С известной магистральной сетью в качестве кодера альтернативные работы по сегментации на основе CNN обычно являются вариантами в реализации декодера. Например, стадия декодирования SegNet использует индексы максимального пула из соответствующих карт характеристик в своем кодере для повышения дискретизации. Результирующие карты затем сворачиваются с набором фильтров для генерации восстановленных карт объектов для плотного прогнозирования. В другой типичной структуре кодера-декодера U-Net, предназначенной для сегментации биомедицинских изображений, уровень повышающей дискретизации напрямую объединяется с набором обрезанных дубликатов соответствующих карт характеристик в кодере для повышения разрешения во время процесса декодера.

Сегментация изображения представляет собой сложную проблему, которая требует как хорошей точности на уровне пикселей, которая зависит от мелкозернистых локальных особенностей, так и точности классификации, для которой глобальный контекст изображения имеет решающее значение для устранения локальных неоднозначностей. Однако стратегия объединения в классических архитектурах CNN является дефектом для потери подробной информации, когда выполняются этапы с несколькими пулами.

Несколько недавно предложенных методов, таких как RefineNet и PSPNet, попытайтесь избежать или восстановить потерю понижающей дискретизации в кодере, совмещая низкоуровневые и высокоуровневые функции. RefineNet разрабатывает модуль декодера, используя как остаточные соединения ближнего, так и дальнего радиуса действия для сбора богатой контекстной информации. Он достиг современного

уровня производительности на 7 общедоступных наборах данных. В PSPNet предлагается модуль пирамидального пула для агрегирования различной региональной контекстной информации для использования возможностей глобальной контекстной информации.

Вдохновленный пространственным пирамидным пулом, DeepLab V2 исследует модуль пространственного пирамидного пула (ASPP), объединяя расширенную свертку с различными частотами выборки и пространственное пирамидальное пул для сбора многоуровневой контекстной информации. В последней версии DeepLab V3+, модуль Xception вводится в кодер вместе с улучшенным модулем ASPP. Он получает самые современные характеристики Pascal-VOC 2012, как показано на рисунке 3.

Рисунок 3

Вышеупомянутые сквозные архитектуры в основном сосредоточены на семантической сегментации. Для сравнения, большинство работ по сегментации экземпляров следуют конвейеру, в котором сегментация предшествует распознаванию. Полностью сверточная сегментация экземпляров (FCIS) объединяет систему предложения сегментов и обнаружение объектов для прогнозирования классов объектов, блоков и масок одновременно. Однако этот метод не так точен, как Mask-RCNN в проблеме сегментации экземпляра MS-COCO. Mask-RCNN расширяет Faster-RCNN, добавляя ветвь предсказания маски в дополнение к регрессии ограничивающего прямоугольника и распознаванию классов. Самая последняя сеть агрегации путей (PANet) улучшает иерархию объектов путем увеличения пути снизу вверх. С небольшими затратами вычислительных ресурсов он занимает первое место в задаче MS-COCO задачи сегментации экземпляров.

Хотя пиксельная сегментация изображения быстро прогрессирует с превосходной точностью, она все еще далека от практического использования, такого как семантическая сегментация видео, из-за высокой сложности плотного прогнозирования. Поэтому разработка высокоэффективной инфраструктуры сегментации изображений является одной из главных задач в сообществе компьютерного зрения.

1. Харинов М.В. Устойчивая сегментация изображения // Вестн. Бурят. гос. ун-та. - 2012. - Вып. 9. - С. 64-69.
2. Бугаев А.С., Хельвас А. В. Поисковые исследования и разработка методов и средств анализа и автоматического распознавания потоковой информации в глобальных информационных системах. Шифр «Лацкан» // Отчет по НИР. - М.: Изд-во МФТИ, 2001. - Т. 1. - 140 с.
3. Харинов М.В. Устойчивая сегментация цифрового изображения // Компьютерная графика и зрение (Графикон'2012): материалы XXII Междунар. конф. по компьютерной графике и зрению. М.: Изд-во МГУ им. М. В. Ломоносова, 2012. С. 208-213.
4. Айвазян С.А. и др. Прикладная статистика: классификация и снижение размерности. - М.: Финансы и статистика, 1989. - 607 с.