

Обзор обнаружения объектов и его применения в космических исследованиях

Научный руководитель – Богуславский Андрей Александрович

Feng Ji

Студент (магистр)

Lomonosov Moscow State University, Факультет космических исследований, Москва,
Россия

E-mail: 305080234@qq.com

Обнаружение объектов - это компьютерная технология, связанная с компьютерным зрением и обработкой изображений, которая занимается обнаружением экземпляров семантических объектов определенного класса (таких как люди, здания или автомобили) на цифровых изображениях и видео.[1] С развитием компьютерных технологий и широким использованием компьютерного зрения использование обнаружения объектов становится все более популярным. Динамическое обнаружение объектов в режиме реального времени имеет важное прикладное значение в области интеллектуальной транспортировки, интеллектуального мониторинга, обнаружения военных целей и космических исследований.

Методы обнаружения объектов обычно классифицируются как основанные на нейронных сетях или не основанные на нейронных сетях. Для не нейронных методов классическими алгоритмами являются: Haar, SIFT, HOG и т.д. Для нейронных методов классическими алгоритмами являются: Region Proposals, Single Shot MultiBox Detector (SSD), You Only Look Once (YOLO), и т.д. Традиционные методы компьютерного зрения используют хорошо разработанные искусственные характеристики для описания характеристик, в то время как алгоритмы, основанные на нейронных сетях, используют характеристики, изученные самим алгоритмом. Теперь эффект автоматически изучаемых характеристик, основанных на алгоритмах нейронной сети, уже стал лучше, чем у созданных человеком характеристик. Таким образом, основное внимание в этой тезисе уделяется алгоритмам, основанным на нейронных сетях, и обсуждение их приложений в основном касается областей, связанных с космосом.

Нейронная сеть, о которой здесь говорится, в основном относится к CNN (сверточная нейронная сеть). В функциональном анализе свертка - это математический оператор, который генерирует третью функцию через две функции f и g . Но в содержании обработки изображений мы в основном используем дискретную версию свертки [2]. CNN привлекла широкое внимание после 2012 года, когда Алекс Крижевски использовал свою модель CNN - Alexnet, чтобы значительно повысить точность распознавания изображений на конкурсе ImageNet.

Для обнаружения объекта необходимо не только определить классификацию цели на изображении, но и определить положение цели. Чтобы определить местоположение цели, мы можем предложить несколько регионов-кандидатов (region proposal). Существует два подхода к генерации region proposal: алгоритм Sliding window и алгоритм Selective search. Sliding window - это традиционный алгоритм перебора, который использует окна с различными масштабами, скользящие по всему изображению, для генерации большого количества предложений по регионам. Алгоритм sliding window будет генерировать много избыточности и очень неэффективен. Selective search - это алгоритм предложения региона для задач обнаружения объектов, который был предложен в статье IJCV 2012. В настоящее время он в основном заменил традиционный метод sliding window. Selective search группирует пиксели изображения и объединяет похожие кластеры, а затем создает рамки для кластеризованных целей. Таким образом, избыточность намного меньше.

R-CNN (Region Based Convolutional Neural Networks, 2013)[3] - одна из первых попыток применения нейронных сетей для обнаружения объектов. Сначала он использует метод selective search для определения местоположения цели, а затем использует CNN для классификации цели. Однако RCNN нельзя использовать при обработке видео в реальном времени, поскольку для обработки одного изображения требуется около 47 секунд (с GPU Titan X).[4]

Важной причиной того, что RCNN отнимает много времени, является то, что CNN выполняется по каждому region proposal отдельно. Если мы сначала запустим базовый слой свертки для всего изображения, а затем передадим результат в подсеть R-CNN, можно сэкономить много вычислений. Это основная стратегия улучшения Fast R-CNN. Кроме того, в Fast R-CNN есть хороший трюк, когда он классифицирует цель, в дополнение к классификациям целей существует дополнительная классификация под названием "фон". Затем, faster R-CNN (2015) [5] является дальнейшим улучшением, основанным на Fast R-CNN. Он использует сеть RPN (сеть региональных предложений) вместо алгоритма selective search для определения местоположения цели. И скорость обработки становится еще быстрее. Теперь нейронная сеть используется не только для классификации цели, но и для определения положения цели. Наконец, поскольку определение положения и классификация цели выполняются с использованием нейронных сетей, можно ли превратить эти два этапа в один этап, чтобы нейронная сеть могла предсказывать положение и классификацию цели одновременно? Да, более поздние SSD и YOLO являются такими методами, и скорость таких методов была значительно улучшена.

Существует два основных критерия обнаружения объекта: mAP и FPS. Полное название mAP - Mean Average Precision, которая может быть использована для измерения того, точно ли алгоритм распознает цель. Полное название FPS - frame per second, то есть количество кадров, которые могут быть обработаны в секунду, которое может быть использовано для измерения скорости обработки алгоритма. В настоящее время mAP YOLO в наборе данных VOC 2007 достигла 66.4, а скорость обработки достигла 45 FPS (с GPU Titan X). [6] Основная модель обнаружения объектов уже может быть использована для обнаружения целей в видео в реальном времени. Кроме того, разные алгоритмы имеют разные символы (mAP, FPS, потребление памяти и время обучения). Таким образом, мы можем выбирать различные алгоритмы в соответствии с требованиями конкретной задачи или использовать различные алгоритмы в комбинации.

Обнаружение объектов может широко использоваться в области космических исследований. Например, идентифицировать самолет в аэропорту и дополнительно определить скорость самолета; идентифицировать транспортные средства на дороге и дополнительно определить скорость транспортного средства, а затем рассчитать условия дорожного движения; и, конечно, цели, связанные с военными, такие как ракеты или реактивные самолеты, также могут быть обнаружены и отслежены.[7]

Источники и литература

- 1) Object detection / Wikipedia. URL: https://en.wikipedia.org/wiki/Object_detection
- 2) Discrete convolution / Wikipedia. URL: https://en.wikipedia.org/wiki/Convolution#Discrete_convolution
- 3) Region Based Convolutional Neural Networks / Wikipedia. URL: https://en.wikipedia.org/wiki/Region_Based_Convolutional_Neural_Networks
- 4) Ross Girshick, 2014, Rich feature hierarchies for accurate object detection and semantic segmentation Tech report / arXiv> URL: <https://arxiv.org/abs/1312.6229v4>

- 5) Shaoqing Ren, Kaiming He, Ross Girshick, 2015, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / arXiv. URL: <https://arxiv.org/abs/1506.01497>
- 6) Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, 2016, You Only Look Once: Unified, Real-Time Object Detection / arXiv. URL: <https://arxiv.org/pdf/1506.02640.pdf>
- 7) [U+592A] [U+7A7A] [U+76F4] [U+64AD] [U+9AD8] [U+6E05] [U+62CD] [U+6444] [U+6C11] [U+7528] [U+9065] [U+611F] [U+536B] [U+661F] [U+8FFD] [U+8E2A] [U+98DE] [U+884] [U+52A8] [U+6001] [U+76D1] [U+6D4B] [U+65F6] [U+901F] [U+70ED] [U+8FFD] [U+8E2A] [U+5B9] [U+7CBE] [U+786E] [U+5230] [U+6BCF] [U+4E00] [U+4E2A] [U+8F66] [U+8F86] / BiliBili. URL: <https://b23.tv/mSkMKVo>