

КОМПЬЮТЕРНОЕ ЗРЕНИЕ И ЕГО ОСНОВНЫЕ ЗАДАЧИ

Круглая А.А., Харкевич А.П., Щедрова Д.С.

*Белорусский государственный университет информатики и радиоэлектроники¹
г. Минск, Республика Беларусь*

Рыкова О.В.. – канд. физ.-мат. наук

В данной работе мы углубились в такую важную область обучения, как компьютерное зрение. Дано целостное описание появления данной области, её практического применения и дальнейшего развития. Так же нами были рассмотрены две основные задачи компьютерного зрения- Object Detection и Semantic Segmentation. Были изучены основные теоретические и практические методы решения этих задач (5 в первом случае и 3 во втором).

Оказывается, что то, что большинство людей воспринимает как должное — способность видеть и обрабатывать увиденное — чрезвычайно трудно воспроизвести в машинах. Именно на это нацелено компьютерное зрение (CV).

Компьютерное зрение описывает процесс, когда компьютер, использующий алгоритмы искусственного интеллекта, может идентифицировать и обрабатывать изображения (фотографии, видео и т. д.). А затем создавать соответствующие результаты анализа, поскольку компьютер действительно может «понимать» контент. В частности, компьютерное зрение может классифицировать, идентифицировать, проверять и обнаруживать объекты.

Есть три основных компонента компьютерного зрения.

1. Получение изображения: когда цифровая камера захватывает изображение, она создает цифровой файл, состоящий из нулей и единиц.
2. Обработка: алгоритмы используются для определения основных геометрических элементов для построения изображений из двоичных данных.
3. Анализ: на этом последнем этапе компьютерного зрения данные анализируются. Алгоритмы высокого уровня используются для принятия решений на основе изображений.

В последние годы компьютерное зрение стало активно использоваться в промышленности. Более 100 различных компаний начали производить системы машинного зрения. Были разработаны специальные светодиоды для систем машинного зрения, расширялись функции световых сенсоров и архитектуры управления системами компьютерного зрения. Это значительно расширило их функционал при постоянном снижении цен на такие системы.

Развитие Интернета в 90-х годах привело к накоплению огромных массивов оцифрованных изображений, которые стали доступны для анализа в режиме онлайн. Начали активно развиваться программы распознавания печатного и рукописного текста, а также распознавание лиц на фото и видео.

В последние 10-15 лет в связи с развитием облачных технологий, виртуализации вычислений, и таких прикладных областей, как машинное и глубокое обучение, Big Data и пр., развитие компьютерного зрения получило новый импульс для роста.

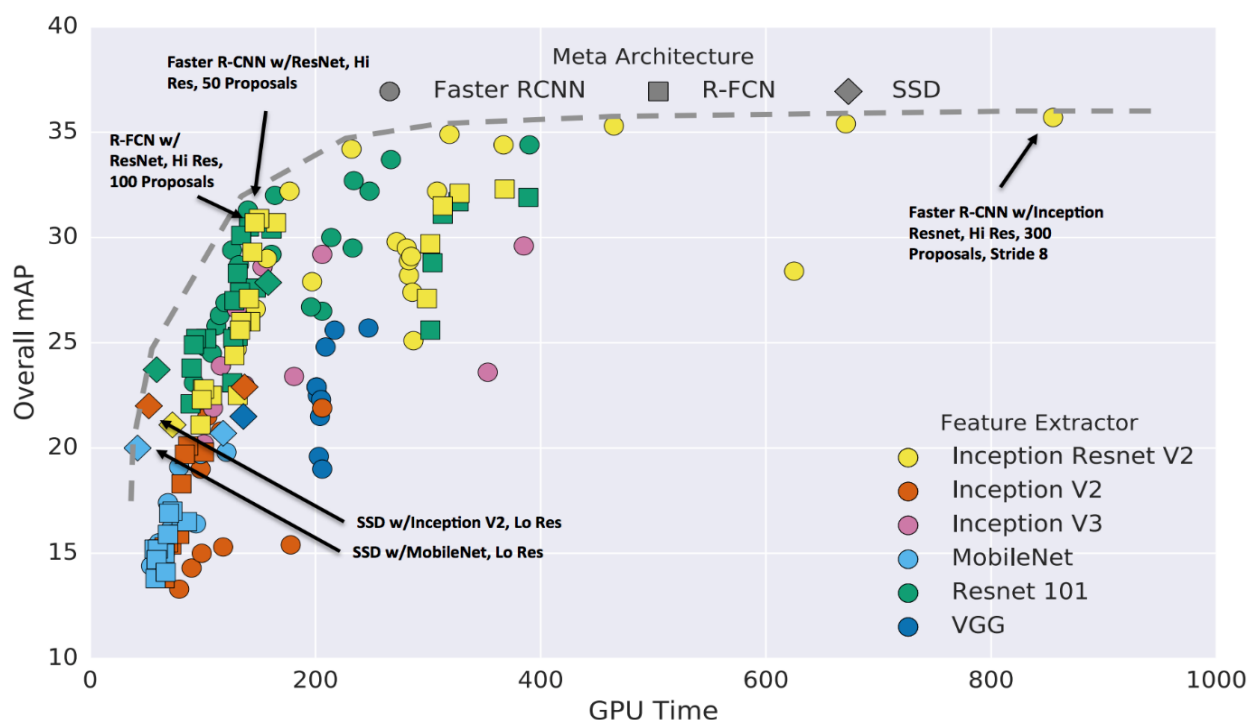
Развитие систем компьютерного зрения пока далеко от реализации всех его возможностей, но они быстро развиваются, и диапазон их применений ширится.

Далее мы проводили сравнительный анализ подходов и сделали следующие выводы:

1. Подход brute force не имеет практического применения, основная сложность в использовании- многократный выбор различных маленьких сегментов на изображении.
2. Подход R-CNN тратит примерно 1 минуту на обработку одного изображения. Основное время тратится на запуск CNN.
3. Подход Fast R-CNN'15 тратит примерно 3 секунду на обработку одного изображение. Основное время он тратит на поиск оптимальных регионов.
4. Подход Faster R-CNN'15 тратит примерно 0.2 секунду на обработку одного изображение. Он показывает наилучшее качество среди данных подходов.
5. Подход YOLO'15 работает со скоростью 45 изображений в секунду, а его быстрая версия работает ещё быстрее- до 155 изображений в секунду! Он работает менее точно, чем CNN подходы, зато сверхбыстро.

Вообще, как правило, ищется баланс между временем работы подхода и его точностью.

Этот баланс хорошо иллюстрирует изображение ниже – на нём можно увидеть как быстро и качественно работает рассмотренные подходы на разных архитектурах.



Для реализации методов Object detection мы выбрали CNN, RCNN и YOLO. Основанием для этого послужило то, что CNN является основой в Object detection, RCNN одна из основных надстроек над CNN. YOLO считается одним из самых прогрессивных методов решения задачи Object detection в наши дни. Данные подходы рассмотрены на конкретных задачах.

Список использованных источников:

1. <https://shalaginov.com/2020/05/16/computer-vision-history/>
2. <https://ru.wikipedia.org>
3. <https://medium.com/profisoft>
4. <https://www.jetinfo.ru/computer-vision-technology-review/>
5. <https://dlcourse.ai/>