Выбор модели обучения

Мой выбор состоял из 3ех моделей машинного обучения подходящих под данную задачу: Faster R-CNN, **SSD**, YOLO. Кратко расскажу о каждой из них.

Faster R-CNN – является прямым продолжением более медленной модели R-CNN (**Regions** with Convolution Neural Networks features), для обнаружения объекта на изображении используется механизм Region Proposal Network (RPN), выделенные RPN регионы передаются на вход обычной нейронной сети классификации. Из-за явных циклов ‘for’ в алгоритме обучения работает сравнительно медленно.

YOLO (You Only Look Once) – как и сказано в названии считывает изображение лишь единожды, данная модель в процессе обучения разбивает картинку на равные ячейки. Каждая ячейка этой сетки предсказывает в ограничивающих прямоугольников с показателем достоверности. Эта уверенность представляет собой просто вероятность обнаружения объекта, умноженную на IoU (intersection Over Union) между предсказанным и наземным полем истинности. Способна работать в реальном времени благодаря ее архитектуре и отсутствию явных циклов.

SSD – ее алгоритм основан на модели YOLO, но ее степень точности немного снижается при идентификации более мелких объектов.

Сравнив эти модели, я выяснил, что модели YOLO, а в частности последняя модель yolov8, является куда более быстрой, и почти не уступающая в точности Faster R-CNN, а в модель SSD не так хорошо подходит под данную задачу, как YOLO.

Yolov8 имеет разные размеры, но мы можем выбрать лишь n – nano, из-за ограниченности времени обучения модели, самую быструю, но не столь точную.