

Обзор статьи: SNIPER: Efficient Multi-Scale Training

Кобелев Максим

14 апреля 2019 г.

SNIPER - эффективный алгоритм задачи распознавания объектов. Реализация алгоритма, описанная в статье основывается на Faster-RCNN с ResNet-101 слоями, и превосходит известные решения как по точности распознавания, так и по времени обучения. Самой главной особенностью алгоритма можно назвать то, что авторы научились не извлекать пирамиду фичей из каждого пикселя картинки, которых во первых может быть много, во вторых - не все из них полезны. Вместо этого они предлагают делить исходную картинку на регионы - Chip'ы низкого разрешения, с помощью Region Proposal Network. Далее внутри них уже применяется стандартная процедура извлечения пирамиды фичей, и работа с ними. Из-за того что Chip'ы имеют низкое разрешение, обучение batch'a может происходить гораздо быстрее, либо мы можем увеличить его размер без потерь по памяти. Результаты также показали, что использовать batch-нормализацию идея очень здравая. В частности потому что низкий размер batch-a позволяет производить обучение на одной GPU, что избавляет разработчика от проведения мероприятий по синхронизации статистик batch-нормализации между кластерами. Ускорение также достигается благодаря умению алгоритма производить адаптивный up/down-sampling, для маленьких и больших изображений соответственно, для того чтобы не производить вычисления над фоновыми пикселями. Finally, имеем, что реализация из статьи может обрабатывать порядка 5 изображений в секунду, и набирает 47.6% точность распознавания объектов ограничивая их обводящими прямоугольниками на датасете COCO.