PhD

SNIPER Efficient Multi-Scale Training ax181213 nips18

2019-12-28 12:38:08 PM

Instead of training on an entire image, extract patches from around GT boxes and train only on these to reduce computation and memory consumption and therefore allow larger batch sizes

Poorly trained RPN used to generate negative samples with the intuition that any patches not generated by it would be easy negatives and therefore can be omitted while training

Image is divided into a regular grid of overlapping patches and greedy selection is used to include patches containing a certain number of GT boxes – Heuristics are used to decide on the smallest and largest boxes included at each scale of the pyramid

References for the claim that object detection and maybe instance segmentation are not very sensitive to faraway context as that might not be covered by the receptive field

Improves on vanilla faster R CNN by about 3 to 4% AP – supposed to be about 5 FPS on GPU which seems to be much slower than the 15 to 20 obtained by the vanilla but not sure what GPU was being used here