



# Multiple-step Sampling for Dense Object Detection and Counting

Zhaoli Deng, Chenhui Yang

School of Informatics

Xiamen University

Xiamen, China

Email: dzleee@stu.xmu.edu.cn, chyang@xmu.edu.cn

## Abstract

Since the poor performance of Faster R-CNN in dense scenes. We propose a straightforward sampling method, which helps to elevate the ratio of positive samples. By assigning a positive label to more anchors in three steps according to different rules, we increase the number of positive samples and reduce the waste of ground-truth. The experiments prove that the performance of our method reaches nearly the existing state-of-the-art with faster inference speed.

## Multiple-step Sampling

There are there are  $N$  anchors and  $M$  ground-truth boxes in an image and we denote the set of anchors as  $\{a_i\}_{i=1}^N$  and the set of ground-truth boxes as  $\{g_j\}_{j=1}^M$ . We assign anchors with positive labels based on their IoU ratio in the following three step.



In the first step, we associate ground-truth boxes  $g_j$  with anchor  $a_i$ . They are satisfy:  $\text{IoU}(a_i, g_j) \geq 0.7$ .

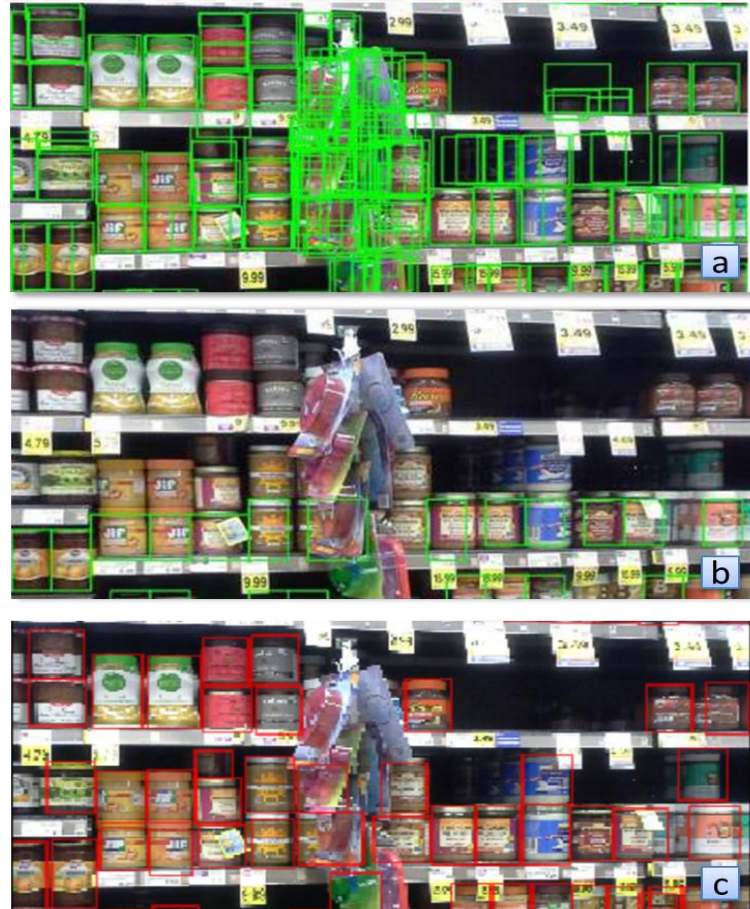
In the second, we select an anchor  $a_i$  without label to match a ground-truth box  $g_j$ .  $\text{IoU}(a_i, g_j) \geq 0.3$ .

In the third step, we select anchor to match the small ground-truth.

DETECTION ON SKU-110K

Method	AP	AP <sup>75</sup>	FPS
RetinaNet [2]	.455	.389	0.58
FPN [4]	.413	.384	0.61
FPN&Soft-IoU [1]	.418	.386	0.60
FPN&EM-Merge [1]	.482	.540	0.24
FPN&Soft-IoU&EM-Merge [1]	<b>.492</b>	<b>.566</b>	0.24
YOLO9000 [8]	.097	.072	5.14
Faster R-CNN [3]	.045	.010	3.15
Faster R-CNN (4,8,16)	.197	.174	3.15
<b>Ours</b>	<b>.418</b>	<b>.427</b>	3.15

Our final model is a Faster R-CNN applied with our proposed method. It gains AP (.418) and AP75 (.427). The detection results in table shows that our proposed sampling method can be applied to Faster R-CNN to improve its performance significantly.



Detection in packed domains. A typical image in SKU-110K, (a) Detection results for the RetinaNet, it returns a lot of overlapping or incorrect boxes. (b) The famous two-stage detector Faster R-CNN detects few objects. (c) Our approach does precise detection.