Supplementary Material of Density Map Regression Guided Detection Network for RGB-D Crowd Counting and Localization

Dongze Lian[1*], Jing Li[1*], Jia Zheng[1], Weixin Luo[1,2], Shenghua Gao[1†]
[1] ShanghaiTech University [2] Yoke Intelligence
{liandz, lijing1, zhengjia, luowx, gaoshh}@shanghaitech.edu.cn

In this document, we give more detection results on our dataset and ShanghaiTech Part_B. Further, we show some failure cases and analyze some possible reasons.

1. Results on Our Dataset and ShanghaiTech Part_B

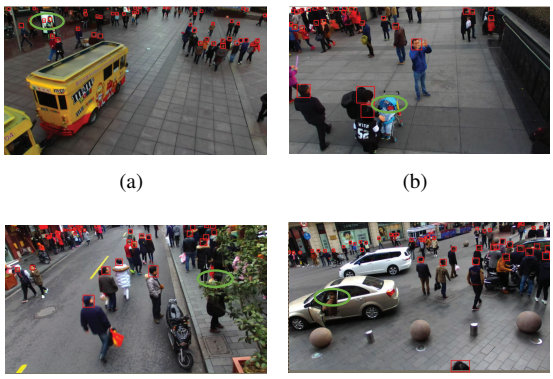
The detection results on our dataset (ShanghaiTechRGBD) and ShanghaiTech Part_B are shown in Figure 1 and 2, respectively.

Figure 1. The detection results on ShanghaiTechRGBD.

Figure 2. The detection results on ShanghaiTech Part_B.

2. Some Failure Cases

Although our RDNet can obtain state-of-the-art performance, there are still some hard cases that cannot be solved. In Figure 3, some heads cannot be detected due to the occlusion (such as (c) (d)) or the high similarity (such as (b)) to background, which are also challenges in regression-based crowd counting. There are still heads in the poster board be detected inaccurately (such as (a)). A possible research direction is that detect head and body together because the human body contains more information for detecting and recognition.



(a) (b) (c) (d)

Figure 3. Some failure cases. The green circles show some false or missing detections.

*Equal contribution.
†Corresponding author.