

Robust Small-scale Pedestrian Detection with Cued Recall via Memory Learning

– *Supplementary Material* –

Jung Uk Kim* Sungjune Park* Yong Man Ro[†]
Image and Video Systems Lab, KAIST, South Korea
{jukim0701, sungjune-p, ymro}@kaist.ac.kr

1. Pedestrian Scale Distribution in CVC-14

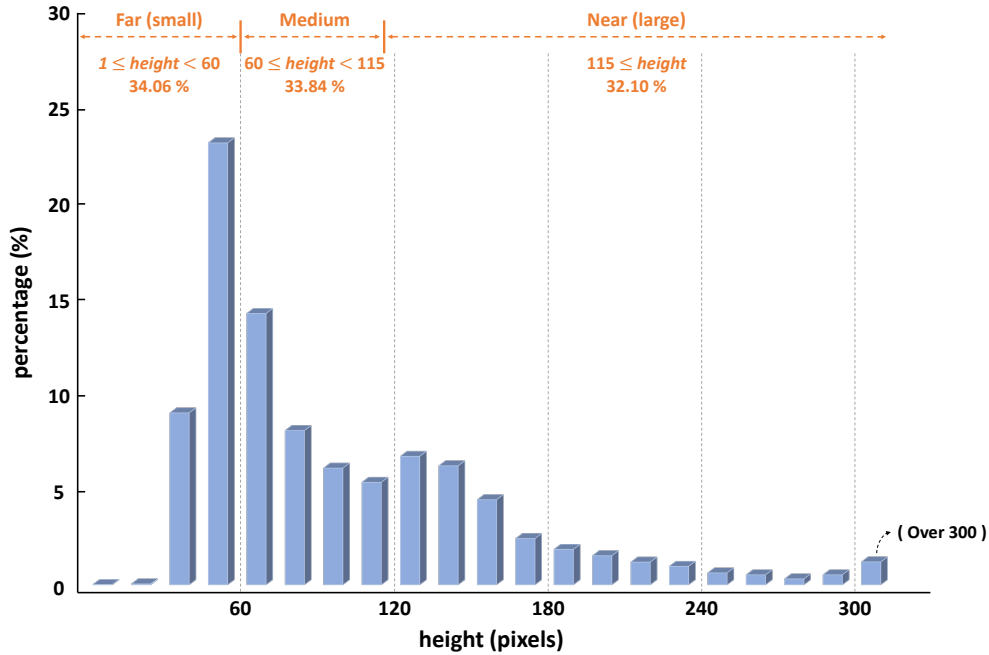


Figure 1. The distribution of pedestrians in CVC-14 [1] with respect to the height scale. We divide the pedestrians into three categories, ‘Far’ (small), ‘Medium’, and ‘Near’ (large), as their heights. Note that the rightmost bar, notated by ‘Over 300’, indicates the pedestrians whose heights are larger than 300.

For evaluation of the proposed framework on CVC-14 [1], it is required to split the distribution of pedestrians into three categories, ‘Far’ (small), ‘Medium’, and ‘Near’ (large). Therefore, we analyze the scale distribution of pedestrians in CVC-14 as their heights, similar with [2]. It is shown in Figure 1. Compared to the KAIST dataset [2], the overall scales of pedestrians in CVC-14 are relatively large. Therefore, we set the height thresholds as 60 and 115 pixels for the proportion of each category to be as evenly as possible (i.e. 34.06%, 33.84%, and 32.10%). With these height thresholds, we measure the detection performances according to each scale as described in Table 4 of the main paper.

2. More RoI Feature Visualization

Due to the proposed LPR Memory, the pedestrian detection framework could recall the visual appearance of large-scale pedestrian, which seems to be captured easily, from the input small-scale pedestrian feature. Figure 5 of the main paper

*Both authors have contributed equally to this work.

[†]Corresponding author

compares the original small-scale pedestrian RoI feature \mathbf{F}_S with the refined small-scale pedestrian RoI feature \mathbf{F}_S^R by RoI feature visualization, following [3]. For the more RoI feature visualization, we provide the feature transition process via the proposed LPR Memory in the “02339_supp_More_Visualization_Results_Video.mp4” video file.

References

- [1] A. González, Z. Fang, Y. Socarras, J. Serrat, D. Vázquez, J. Xu, and A. M. López. Pedestrian detection at day/night time with visible and fir cameras: A comparison. *Sensors*, 2016. 1
- [2] Soonmin Hwang, Jaesik Park, Namil Kim, Yookyung Choi, and In So Kweon. Multispectral pedestrian detection: Benchmark dataset and baseline. In *CVPR*, 2015. 1
- [3] Chang-Dong Xu, Xing-Ran Zhao, Xin Jin, and Xiu-Shen Wei. Exploring categorical regularization for domain adaptive object detection. In *CVPR*, 2020. 2