Supplementary Material
WEPDTOF: A Dataset and Benchmark Algorithms for In-the-Wild People Detection and Tracking from Overhead Fisheye Cameras

M. Ozan Tezcan[1], Zhihao Duan[2], Mertcan Cokbas[1], Prakash Ishwar[1], and Janusz Konrad[1]

mtezcan@bu.edu, duan90@purdue.edu, [mcokbas, pi, jkonrad]@bu.edu

[1]Boston University, Department of Electrical and Computer Engineering, Boston, MA
[2]Purdue University, Department of Electrical and Computer Engineering, West Lafayette, IN

1. Video-Wise Performance Comparison

In the paper, we provided the cumulative people-detection performance comparison of baseline algorithms and spatio-temporal extensions of RAPiD for all videos in WEPDTOF (Table 3). Here, we break down this comparison into individual videos.

In Table 1 below, we report per-video AP_{50} scores of the 7 algorithms evaluated on WEPDTOF. Although spatio-temporal algorithms outperform the spatial-only algorithms on the majority videos, there is no single best algorithm for *all* of the videos. Even the improved scores of the reported algorithms are not satisfactory for videos with tiny projected bodies at field-of-view periphery (e.g., "Exhibition Setup"), distorted image aspect ratio (e.g., "Street Grocery"), and strong camouflage (e.g., "Printing Store"). We believe the performance for these challenges can be further improved by developing algorithms that address such challenges directly (e.g., via data augmentation to mimic these challenges in the training set).

References

[1] Z. Duan, M. O. Tezcan, H. Nakamura, P. Ishwar, and J. Konrad. RAPiD: Rotation-aware people detection in overhead fisheye images. In *Proc. IEEE Conf. Computer Vision Pattern Recognition Workshops*, pages 636–637, 2020.

[2] S. Li, M. O. Tezcan, P. Ishwar, and J. Konrad. Supervised people counting using an overhead fisheye camera. In *Proc. IEEE Int. Conf. Advanced Video and Signal-Based Surveillance*, pages 1–8, 2019.

[3] M. Tamura, S. Horiguchi, and T. Murakami. Omnidirectional pedestrian detection by rotation invariant training. In *Proc. IEEE Winter Conf. on Appl. of Computer Vision*, pages 1989–1998, 2019.

Table 1: Per-video performance comparison of fisheye people-detection algorithms on WEPDTOF.

Algorithm	Empty Store	Exh. Setup	Conv. Store	Large Office	Ware-house	Exhi-bition	Call Center	Tech Store	Jewel. Store	Street Groc.	Print. Store	Repair Store	IT Office	Kinder-garten
Tamura *et al.* [3]	85.3	33.1	91.3	35.1	70.7	90.3	59.3	69.2	89.9	12.5	20.4	71.2	60.2	49.1
AA [2]	**96.8**	30.5	92.8	40.3	75.0	89.0	64.9	75.5	86.0	26.4	56.0	**86.0**	70.3	66.3
AB [2]	94.6	33.2	92.4	41.0	85.8	89.0	63.4	76.1	92.8	26.2	**59.4**	85.2	**70.7**	67.3
RAPiD [1]	94.5	32.4	97.2	63.4	86.5	92.8	67.8	70.7	78.8	54.2	51.5	76.2	65.3	77.0
RAPiD+REPP	95.3	34.4	**97.7**	66.3	87.6	**94.2**	68.4	71.8	83.0	**58.4**	53.8	77.0	65.8	78.2
RAPiD+FA	91.7	**45.4**	95.9	71.4	**89.0**	91.7	74.5	79.2	**94.1**	57.3	45.1	76.2	65.8	81.4
RAPiD+FGFA	93.5	40.1	96.4	**77.3**	87.5	91.2	**75.4**	**80.2**	93.2	54.0	55.6	78.9	67.4	**81.8**