

Supplementary Material for ADSCNet

Shuai Bai¹, Zhiqun He², Yu Qiao³, Hanzhe Hu⁴, Wei Wu², Junjie Yan²

¹Beijing University of Posts and Telecommunications ²SenseTime Group Limited

³Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences ⁴Peking University

baishuai@bupt.edu.cn {hezhiqun, wuwei, yanjunjie}@sensetime.com

yu.qiao@siat.ac.cn huhz@pku.edu.cn

In this supplementary material, we provide additional visualization results. In section 1, self-corrected density maps with different iteration number are presented. Then some qualitative results on evaluated datasets are shown in section 2.

1. Results with Different Iteration number

As shown in Fig. 1 based on the estimated results, the neighboring dotted annotations are adjusted to close to the response position, and the variances tend to match the response range. However, there are some dotted annotations that are not effectively responding. As the iteration number increases, the variances of these points will continuously increase. The limited variance growth is conducive to the response of the target, and the excessive variance will only introduce the learning of the background information. Therefore, excessive iterations do not bring the improvement of the performance.

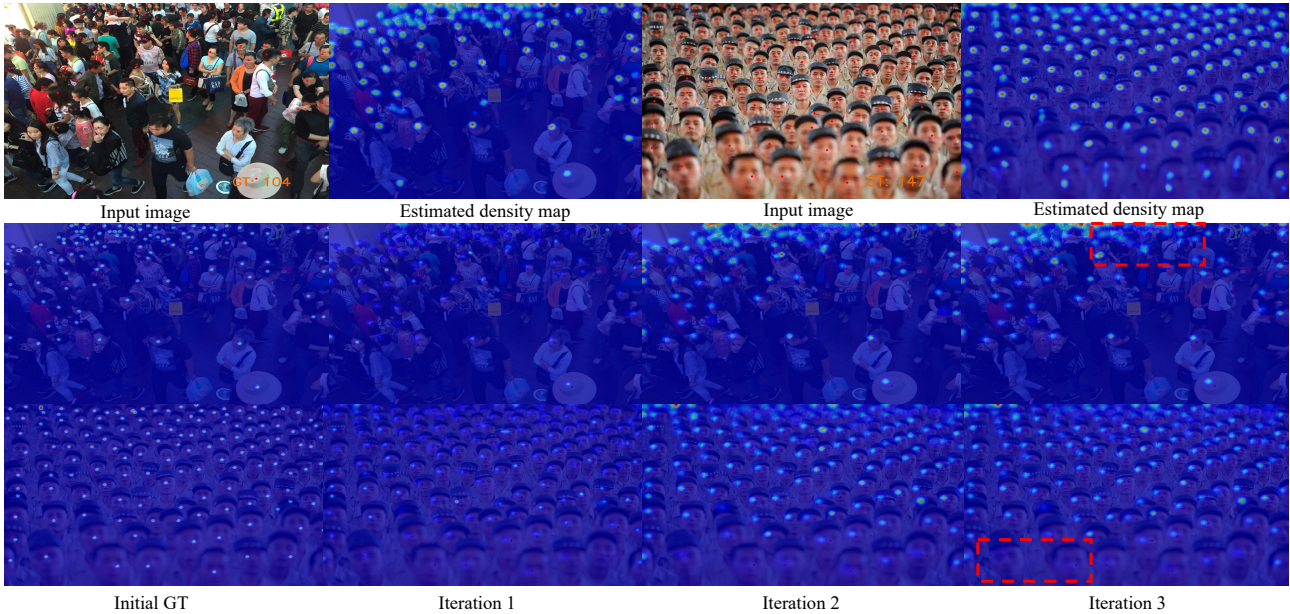


Figure 1. **Results with Different Iteration number.** As the iteration number increases, the positions and variances of most points will match the targets. However, the variances of some points (red boxes) continuously increase, which introduces the background noises.

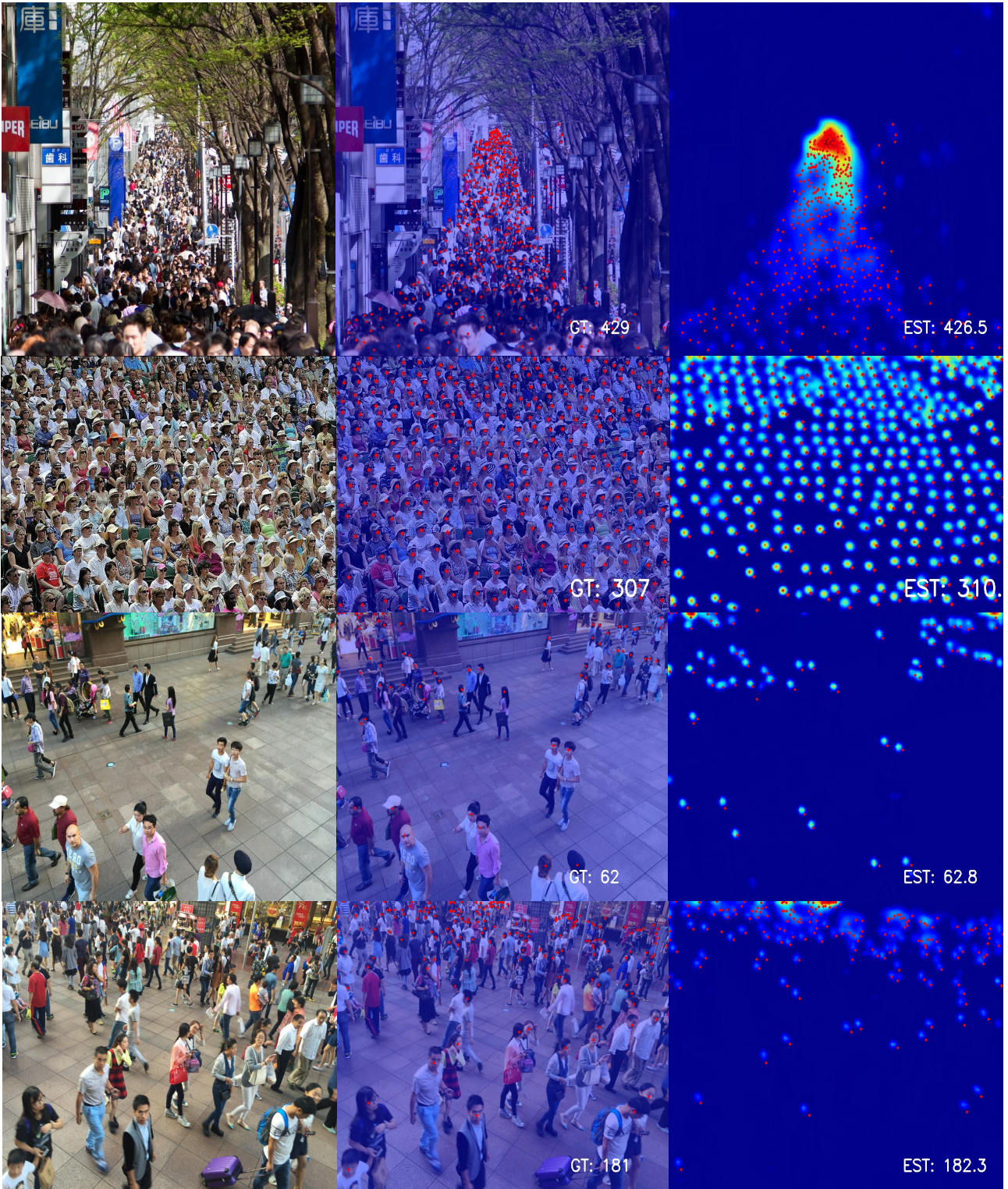
2. Qualitative Results with ADSCNet

We present some qualitative results of ADSCNet on four benchmarks (ShanghaiTech [4], UCF_CC_50 [2], UCF_QNRF [3], TRANCOS [1]) in Fig. 2 to Fig. 5. In the estimated density maps, the response positions are different from the dotted annotations but consistently focus on a certain location of the target, such as the upper left contours of the head for human and

the upper left contours of car for vehicles. Therefore, the network actually counts objects by the contour features. In addition, the results of ADSCNet have uniform response intensity whether in dense-crowd or sparse-crowd regions.

References

- [1] Ricardo Guerrero-Gómez-Olmedo, Beatriz Torre-Jiménez, Roberto López-Sastre, Saturnino Maldonado-Bascón, and Daniel Onor-Rubio. Extremely overlapping vehicle counting. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 423–431. Springer, 2015.
- [2] Haroon Idrees, Imran Saleemi, Cody Seibert, and Mubarak Shah. Multi-source multi-scale counting in extremely dense crowd images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2547–2554, 2013.
- [3] Xinlong Wang, Tete Xiao, Yuning Jiang, Shuai Shao, Jian Sun, and Chunhua Shen. Repulsion loss: Detecting pedestrians in a crowd. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7774–7783, 2018.
- [4] Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, and Yi Ma. Single-image crowd counting via multi-column convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 589–597, 2016.

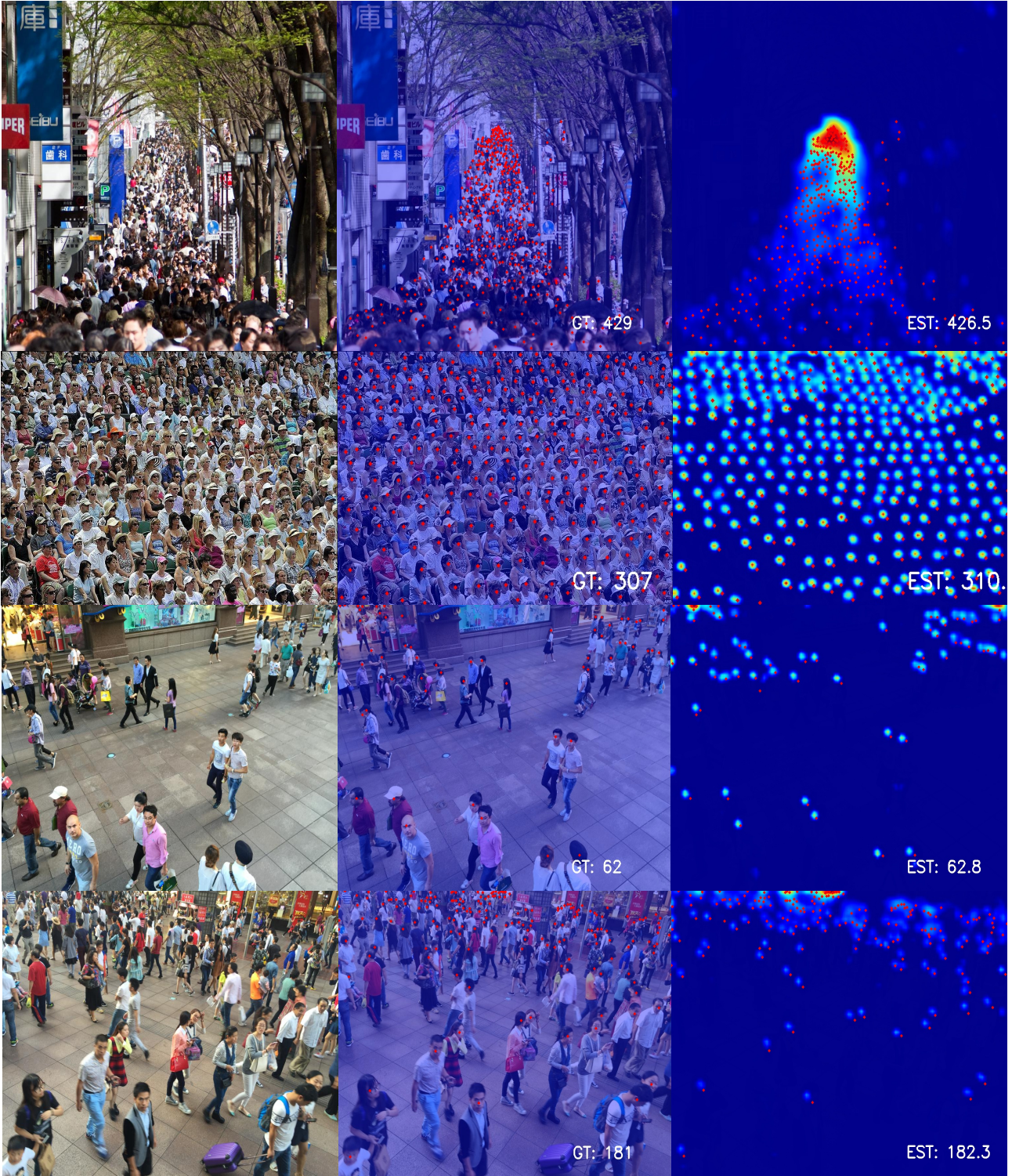


(1) Input images

(2) Annotations

(3) Estimated density maps

Figure 2. **Qualitative results on ShanghaiTech.** The images in first two rows are from ShanghaiTechPartA. The images in last two rows are from ShanghaiTechPartB. The dotted annotations are represented as red points to compare with the estimated density maps.



(1) Input images

(2) Annotations

(3) Estimated density maps

Figure 3. *Qualitative results on UCF_QNRF. The dotted annotations are represented as red points to compare with the estimated density maps.*

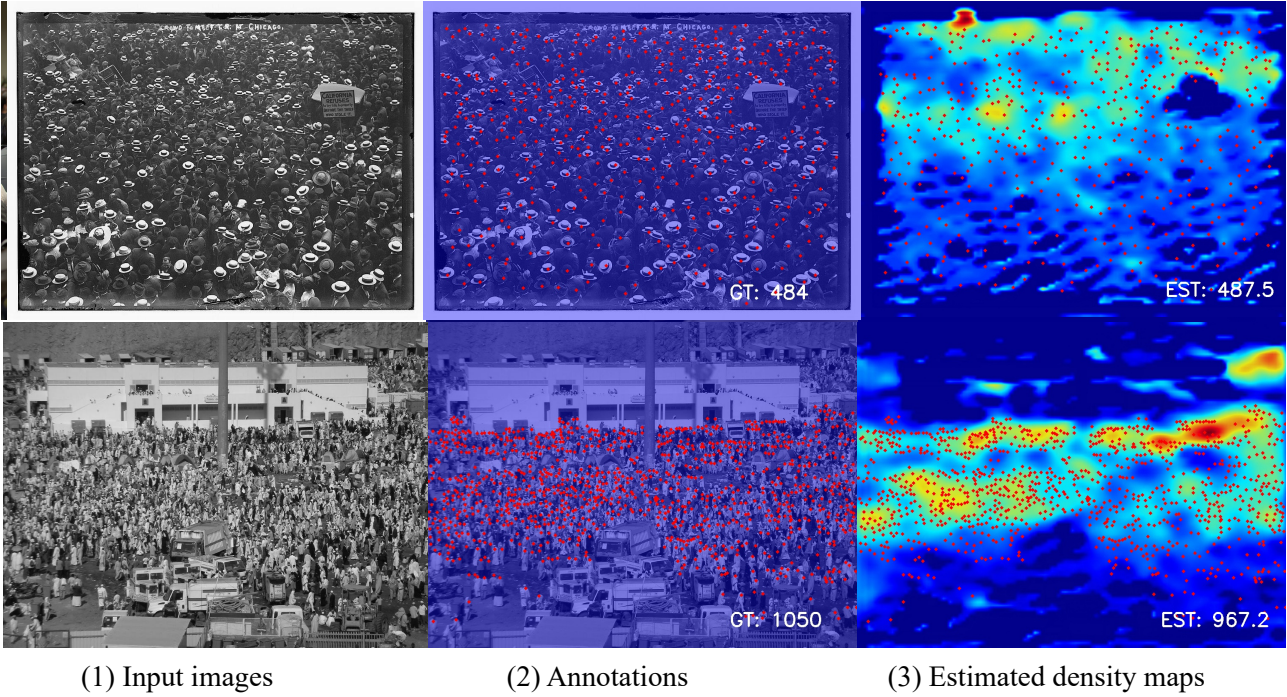


Figure 4. *Qualitative results on UCF_CC_50.* Limited by small dataset size, the density maps are not thoroughly learned. The dotted annotations are represented as red points to compare with the estimated density maps.

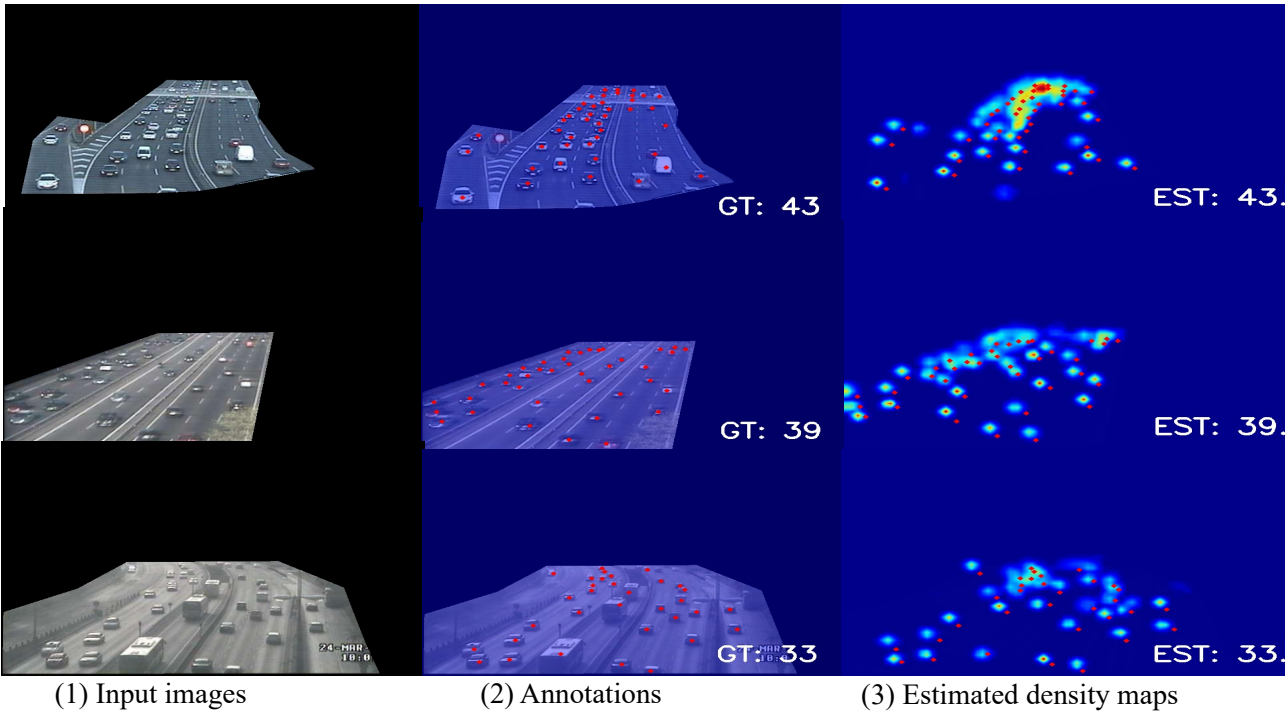


Figure 5. *Qualitative results on TRANCOS.* The dotted annotations are represented as red points to compare with the estimated density maps.