Supplementary Material for the Paper: Knowledge Mining and Transferring for Domain Adaptive Object Detection

Kun Tian[†‡] Chenghao Zhang[†‡] Ying Wang[†] Shiming Xiang[†‡] Chunhong Pan[†]
[†] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
[‡] School of Artificial Intelligence, University of Chinese Academy of Sciences
Email:{kun.tian, chenghao.zhang, ywang, smxiang, chpan}@nlpr.ia.ac.cn

1. Visualization of feature alignment

As stated in Section 3.1 of the original paper, our KTNet can also achieve feature alignment between different domains. In Figure 1, we utilize t-SNE [1] to visualize the distribution of foreground and background features obtained by source-only detector and KTNet in three cross-domain scenarios. It can be observed that there are obvious domain gaps existing in the foreground/background features extracted by source-only model. However, transferring object-related knowledge better aligns the representation of same category in source and target domains. And the KTNet still has a good ability to distinguish foreground and background features, which is attributed to its focus on common knowledge hidden in the same class. Note that when source and target domains are similar, our model also aggregates the background features well. And if there is a significant domain disparity (Sim10k/KITTI→Cityscapes), KTNet does not completely align the background representation, which indicates that our model is able to capture the differences and similarities between domains.

2. More examples for activation maps and testing results

It can be seen from Figure 2 that the detector without adaption pays more attention to the background, which may cause the output of unreasonable false alarms. While KTNet can maintain activation response to foreground objects and suppress the focus on irrelevant background noise.

As displayed in Figure 3, the domain adaptive detector via knowledge transferring can significantly alleviate missing detection, i.e., the number of blue boxes is reduced. Meanwhile, there are more green boxes representing true positives. This intuitively shows that our KTNet does have better testing performance in the target domain. Although there are still some false alarms, they are more concentrated on the foreground objects. We analyze the occurrence of false detection for two reasons. First, detectors predict the same target repeatedly. If intersection over union of multiple detected results and the ground truth are greater than threshold, *e.g.*, 0.5, only the box with the highest confidence will be judged as true positive. Second, the classification and regression accuracy of the model still need to be improved, which is also the direction of our future efforts.

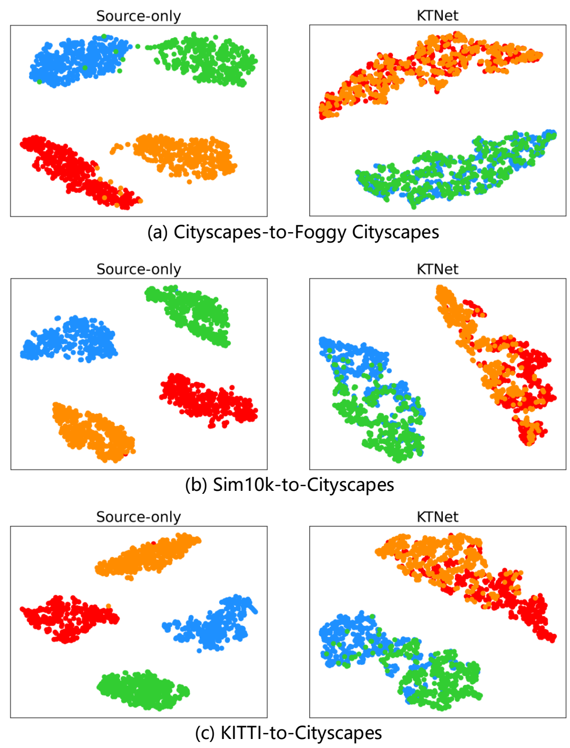

Figure 1. Visualization of foreground and background feature embeddings from three transfer tasks. Red/orange dots are foregrounds of source/target domain. Blue/green dots are backgrounds of source/target domain. For each task setting, the left and right subgraphs are derived from source-only model and our KTNet respectively. Best viewed in color.

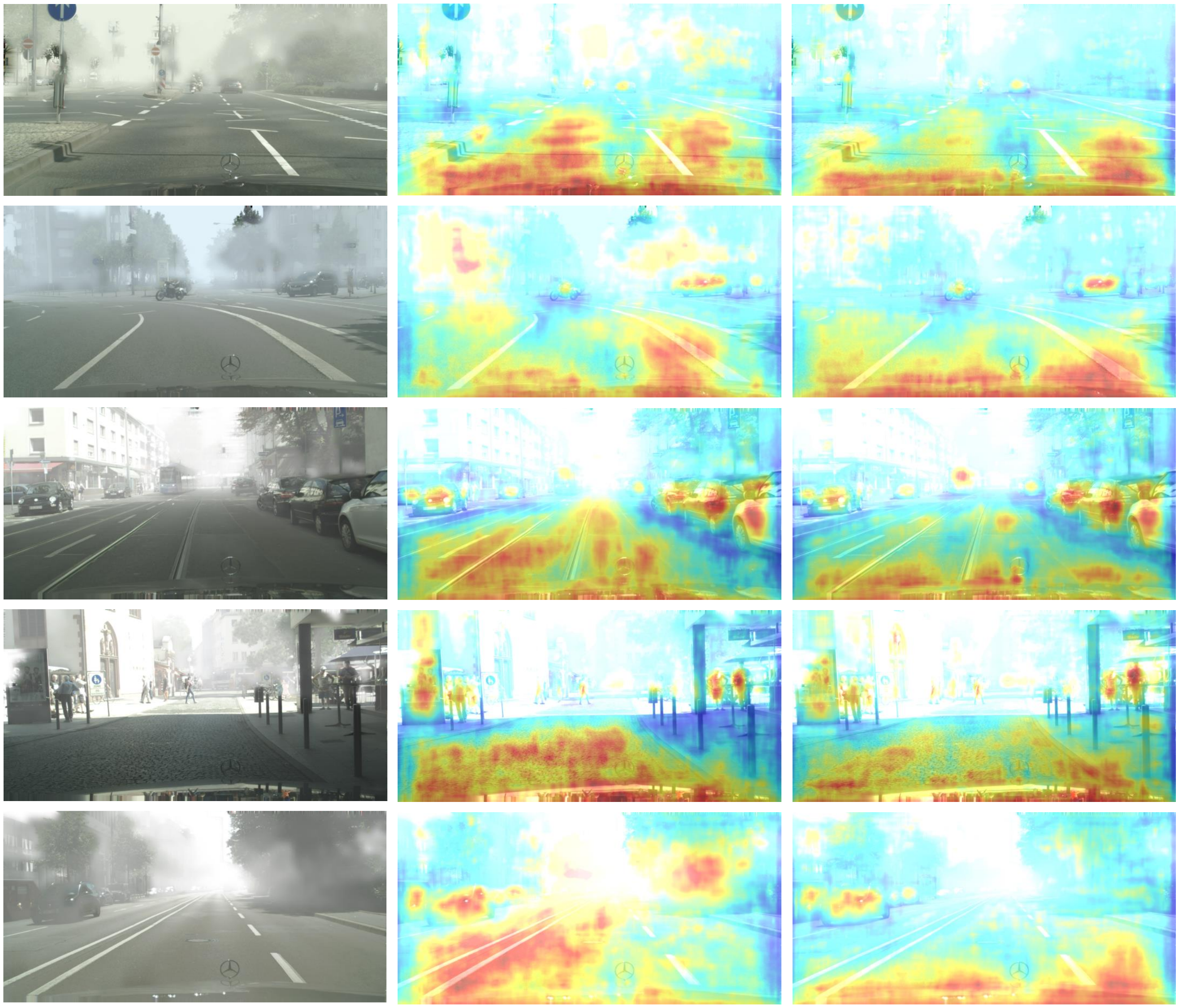

Figure 2. More examples for object activation response. The second and third columns are collected from source-only model and KTNet. The warmer the color, the larger the response value. Please zoom in for a clearer visual effect.

References

[1] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2008.

Figure 3. More examples for testing results in the target domain. The first and second columns correspond to source-only model and KTNet. Green, red and blue boxes are true positives, false alarms and false negatives (missing detection).