Segmenting Transparent Objects in the Wild

Enze Xie, Wenjia Wang[2], Wenhai Wang[3], Mingyu Ding[1], Chunhua Shen[4], and Ping Luo[1]

[1] The University of Hong Kong [2] SenseTime Research
[3] Nanjing University [4] The University of Adelaide
{xieenze, mingyuding}@hku.hk, wangwenjia@sensetime.com, wangwenhai362@smail.nju.edu.cn, chunhua.shen@adelaide.edu.au, pluo@cs.hku.hk

1 Detailed annotation information.

Our dataset contains 20 different categories of transparent objects. As shown in Fig. 1(a), we count the number of images for different categories. We see that for things, the "Cup" appears the most frequently while the "Stationery" appears the least. For stuff, the "French Window" appears the most frequently while the "Table" appears the least. Furthermore, our dataset also contains 13 scenarios. As shown in Fig. 1(b), we see that the "Desktop" is the most frequently scene while the "Vechile" is the least scene. In summary, our Trans10K contains abundant category distribution of transparent objects and scenarios, which is not available in TOM-Net [1] and TransCut [2].

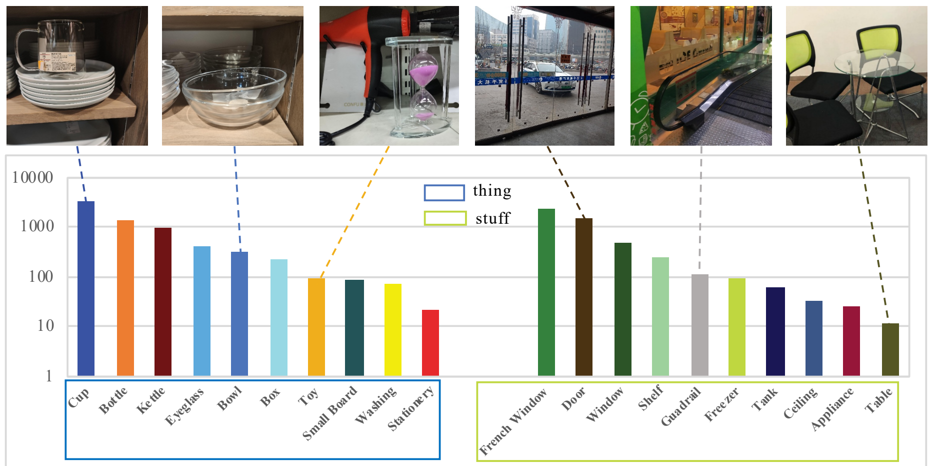
2 More visual results Analysis.

2.1 Failure Samples Analysis

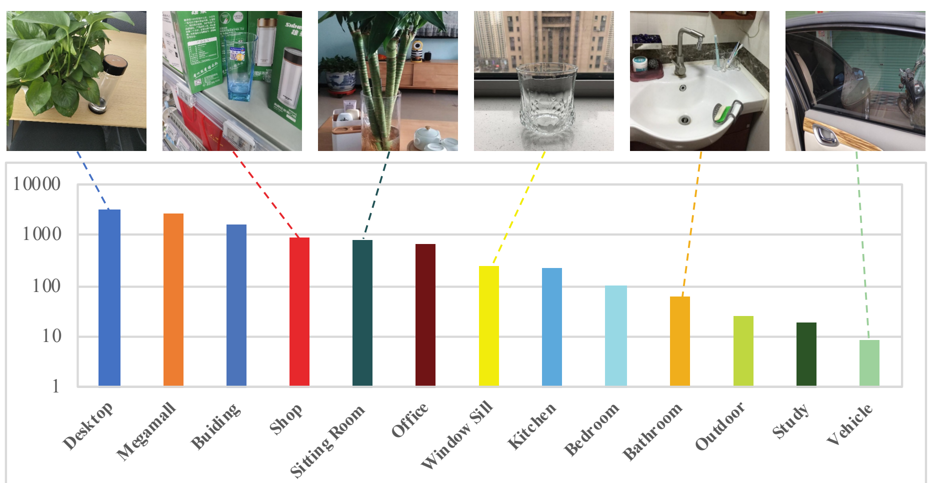
As shown in Fig. 2, our method has some limitations. For example, in Fig. 2 (a), when facing highly-transparency objects, our method will fail to segment in some region. Fig. 2 (b) shows that some objects with strong reflection will also make our method confuse and lead to wrong classification. In Fig. 2 (c), we can find our method does not work well when some objects have overlap and occlusion with transparent objects. Finally, in Fig. 2 (d), when semi-transparent objects are adjacent with transparent objects, our method will also confuse.

2.2 Visual results on external data.

In this part, we also directly test our TransLab trained on Trans10K dataset to evaluate the generalize ability of Trans10K dataset and the robustness of TransLab. Firstly, we test our method on two prior datasets: TransCut [2] and TOM-Net [1]. As shown in Fig. 3, our method can clearly output very high-quality mask. Moreover, we also test our method on some external data randomly captured by our mobile phones or obtained from Internet videos such as YouTube, TikTok and eBay. We see our TransLab can also successfully segment

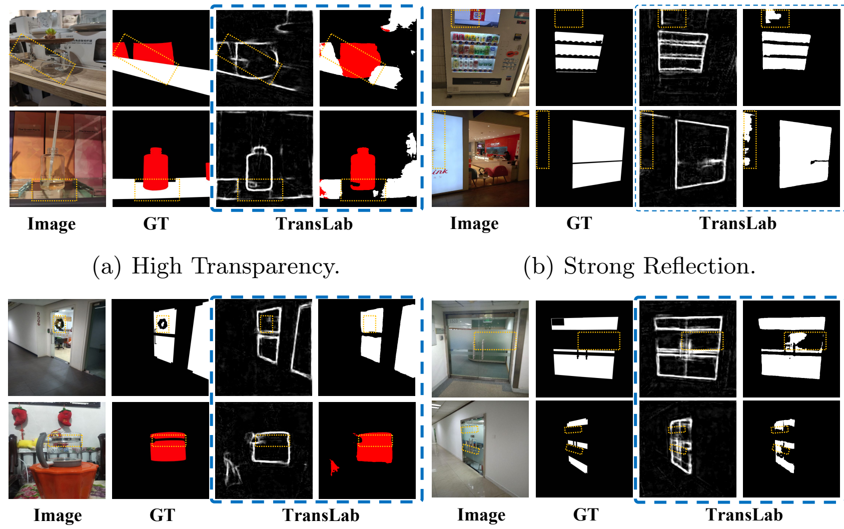


(a) Histogram of image numbers for different transparent object categories. It is split by things and stuff. Event categories are ranked in an descending order based on the image numbers. Example images for specific transparent object classes are shown. Y-axis denotes for image numbers. X-axis denotes for transparent object categories.



(b) Histogram of image numbers for different scene categories. Event categories are ranked in an descending order based on the image numbers. Example images for specific scene classes are shown. Y-axis denotes for image numbers. X-axis denotes for event scene name.

Fig. 1. Statistics of Trans10K dataset.



(a) High Transparency. (b) Strong Reflection.

(c) Overlap & Occlusion. (d) Semi-Transparency.

Fig. 2. Failure cases. Our method fails to segment transparent objects in some complex scenarios.

transparent objects in most cases. In summary, we believe our Trans10K dataset contains high-diversity images which can easily generalize to real scene. Also, our boundary-aware algorithm TransLab is robust enough to segment unseen images.

2.3 More comparison of TransLab with other methods.

In this part, we demonstrate more test examples produced by our TransLab on Trans10K dataset in Fig. 5. From these results, it can be easily observed that with the proposed Boundary Stream and Boundary Attention Module, our method can output high-quality boundary map and better transparent object segmentation mask than other semantic segmentation methods.

References

1. Chen, G., Han, K., Wong, K.K.: Tom-net: Learning transparent object matting from a single image. In: CVPR. (2018)
2. Xu, Y., Nagahara, H., Shimada, A., Taniguchi, R.: Transcut: Transparent object segmentation from a light-field image. In: ICCV. (2015)

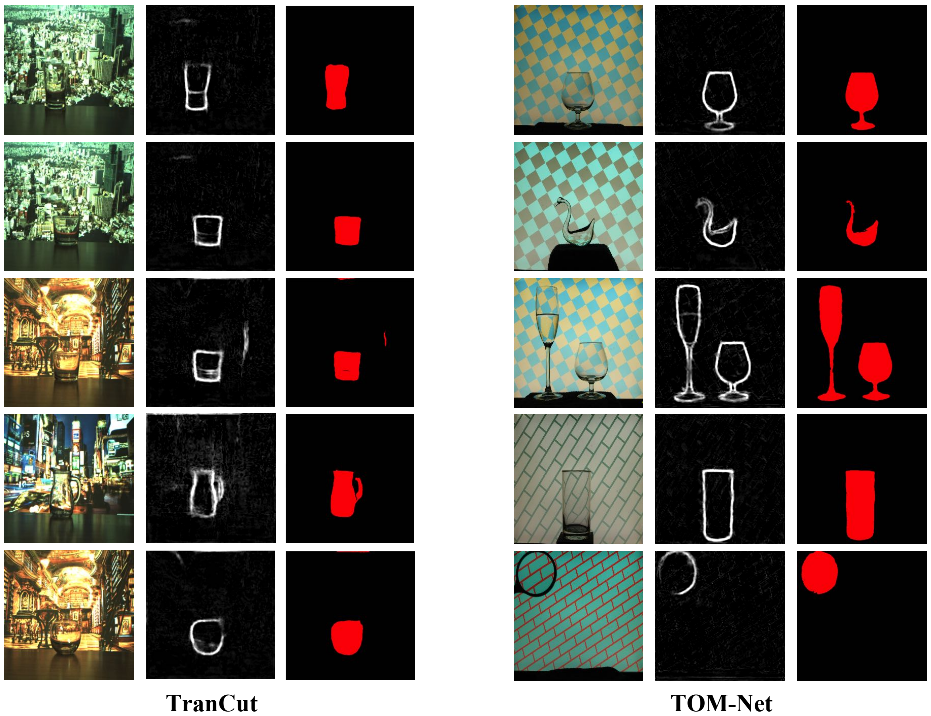


Fig. 3. Some transparent objects segmentation results on two prior datasets: TransCut [2] and TOM-Net [1].

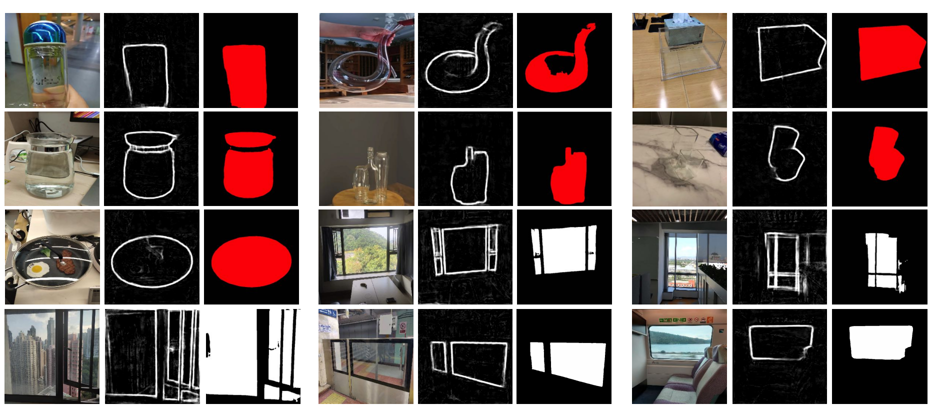


Fig. 4. Some transparent objects segmentation results on challenging images captured by our mobile phones and obtained from Internet such as YouTube, TikTok and eBay.

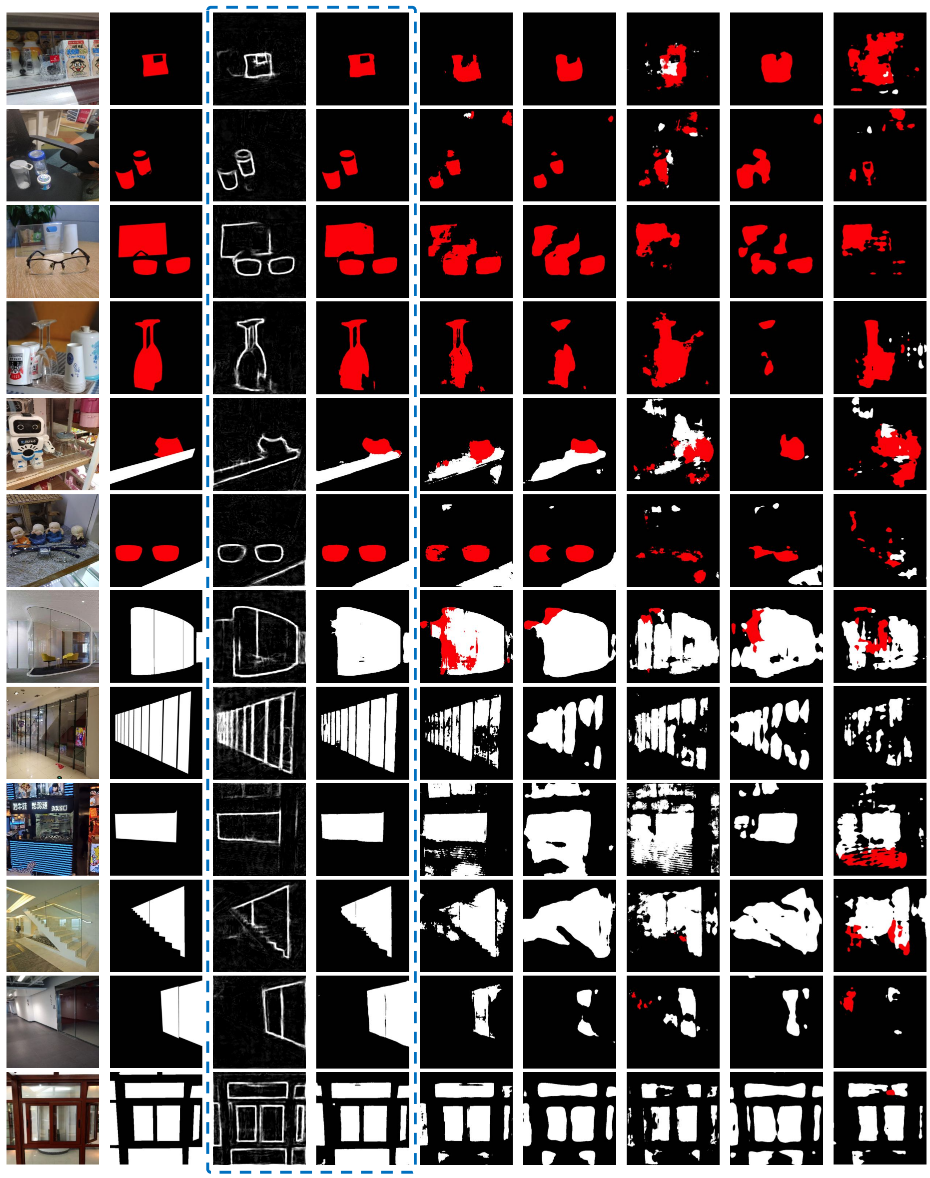


Fig. 5. More visual comparison of TransLab to other semantic segmentation methods. Our TransLab clearly outperforms others thanks to the boundary attention.