Supplementary Material for Sketch2Model: View-Aware 3D Modeling from Single Free-Hand Sketches

Song-Hai Zhang Yuan-Chen Guo Qing-Wen Gu
BNRist, Department of Computer Science and Technology, Tsinghua University, Beijing
shz@tsinghua.edu.cn, guoyc19@mails.tsinghua.edu.cn, gqw17@mails.tsinghua.edu.cn

1. Network Architecture

We use ImageNet-pretrained ResNet-18 as our image feature extractor. The 512-dim feature vector is mapped to a 512-dim shape code and a 512-dim view code through 2 linear layers with L2-normalization in the end. The two codes are concatenated as the input to the decoder, which consists of multiple linear layers, to get the offsets for each vertex on the template mesh. For the shape discriminator, we tried using latent features in the decoder as input, and also tried directly using the offsets as input. We found that the latter works well while the former could result in the discriminator becoming too strong too quickly, which makes it hard to train the generator and leads to unrealistic results. For the domain discriminator, we concatenate the average pooling and max pooling results of the image feature map as input, like in [2].

2. Viewpoint Estimation

We use Euler angles to represent viewpoints and normalize the angles within $[0, 1]$ for input and prediction. Table.1 shows viewpoint prediction errors on ShapeNet-Synthetic test set, measured in degrees. Azimuth angle can be ambiguous for objects with multiple symmetry planes (bench, cabinet, display, lamp, loudspeaker, table, telephone), which leads to poor numerical results as shown in the table. We think it can be solved by adopting better viewpoint representation, and leave it as a future work.

Category	airplane	bench	cabinet	car	chair	display	lamp	loudspeaker	rifle	sofa	table	telephone	watercraft
elevation	7.0	4.4	2.8	2.2	4.5	4.7	9.9	6.0	8.3	4.0	3.4	6.5	8.4
azimuth	32.8	81.0	78.5	23.4	29.9	78.1	109.8	104.0	37.1	36.7	114.8	89.3	69.8

Table 1. View prediction errors on ShapeNet-Synthetic test set. Errors are measured in degrees.

3. Additional Results

We show some additional results on ShapeNet-Sketch dataset in Fig.2,3,4,5 and on existing free-hand sketch datasets in Fig.6,7,8,9,10,11,12. For free-hand sketch datasets, we also compare with SoftRas [1] to show that our results are more sketch-aligned. Each model is rendered from a standard viewpoint, and projected to the predicted viewpoint for intuitive comparisons.

4. Failure Cases

Figure.1 shows some failure cases of our method. Our method sometimes struggles to generate thin structures, like chair legs. Also, The generated results can be too smooth and lack distinct features, comparing to SoftRas [1]. We find that the problem lies in the shape discriminator. To balance between view-awareness and shape quality, the network tends to generate shapes more similar to the average shape, and ignore specific details. The problems may be solved by carefully choosing viewpoints used in RVR instead of random selection, and by adopting more advanced adversarial training strategy.

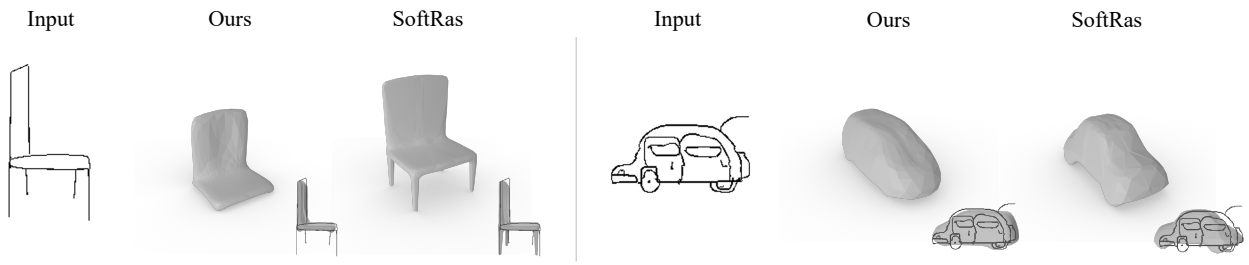

Figure 1. Failure cases. Our method sometimes has difficulty generating thin structures (left), and tends to generate over-smooth shapes (right).

References

[1] Shichen Liu, Tianye Li, Weikai Chen, and Hao Li. Soft rasterizer: A differentiable renderer for image-based 3d reasoning. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 7708–7717, 2019. 1

[2] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018. 1

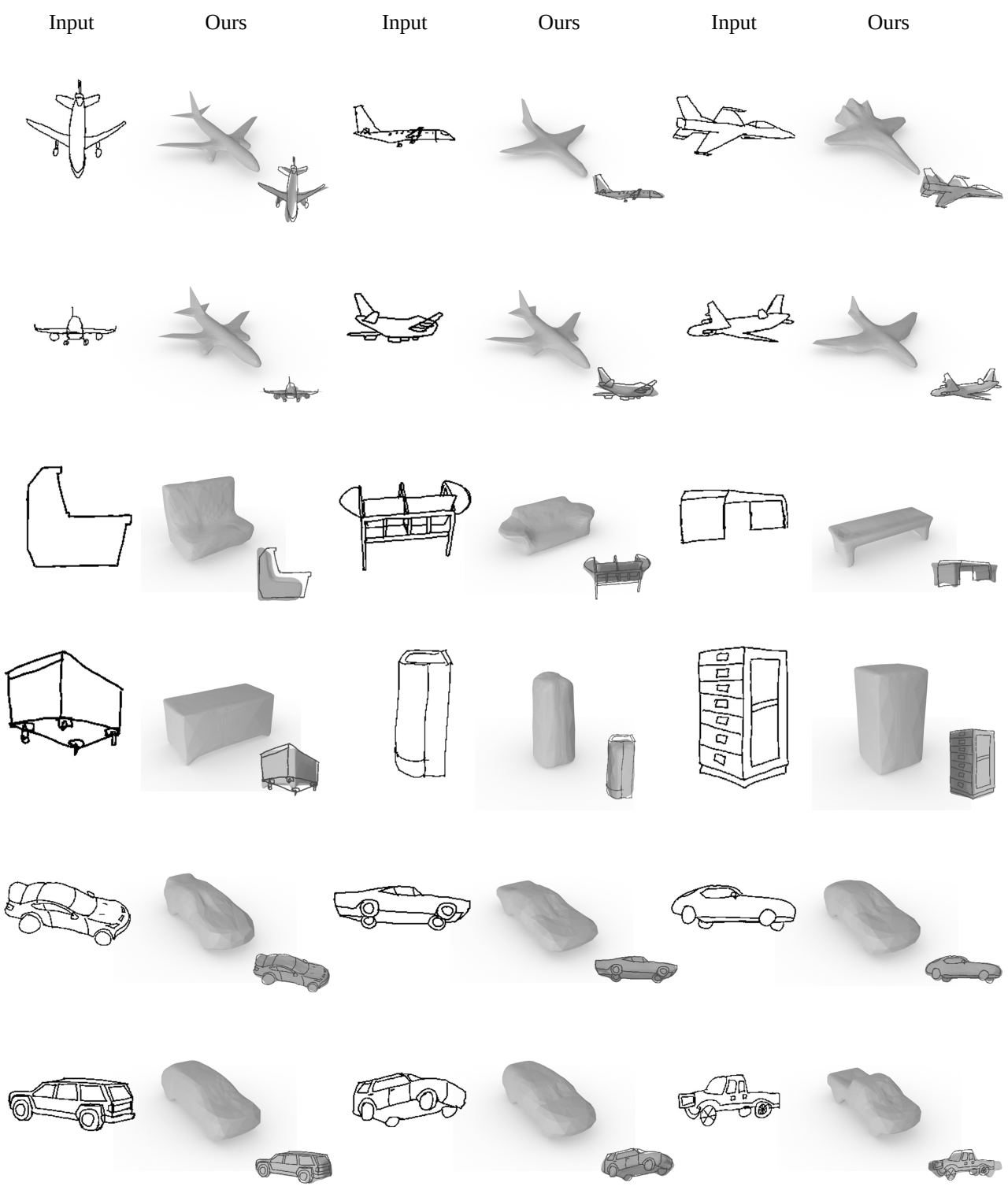

Figure 2. Additional results on ShapeNet-Sketch dataset.

Figure 3. Additional results on ShapeNet-Sketch dataset.

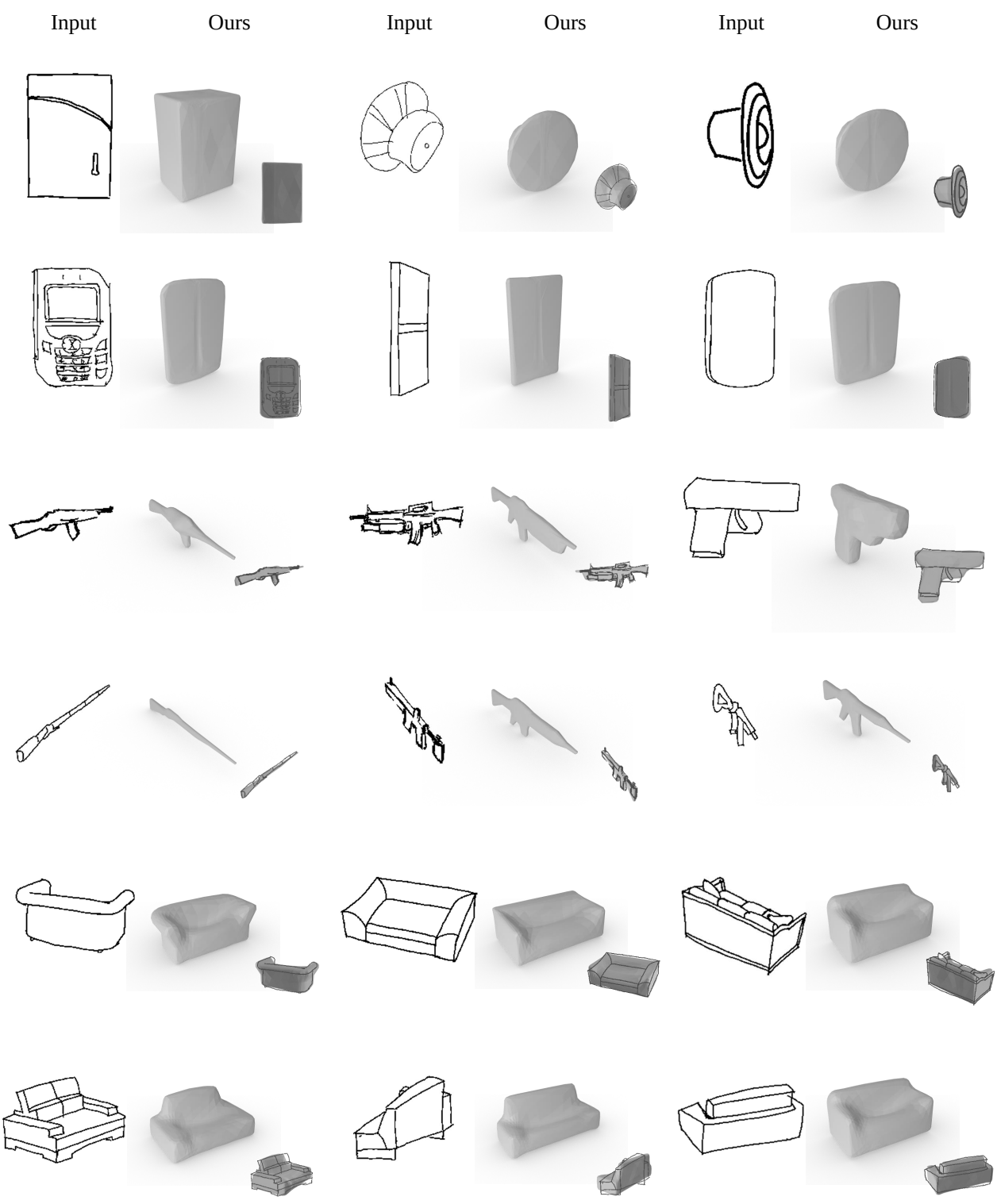

Figure 4. Additional results on ShapeNet-Sketch dataset.

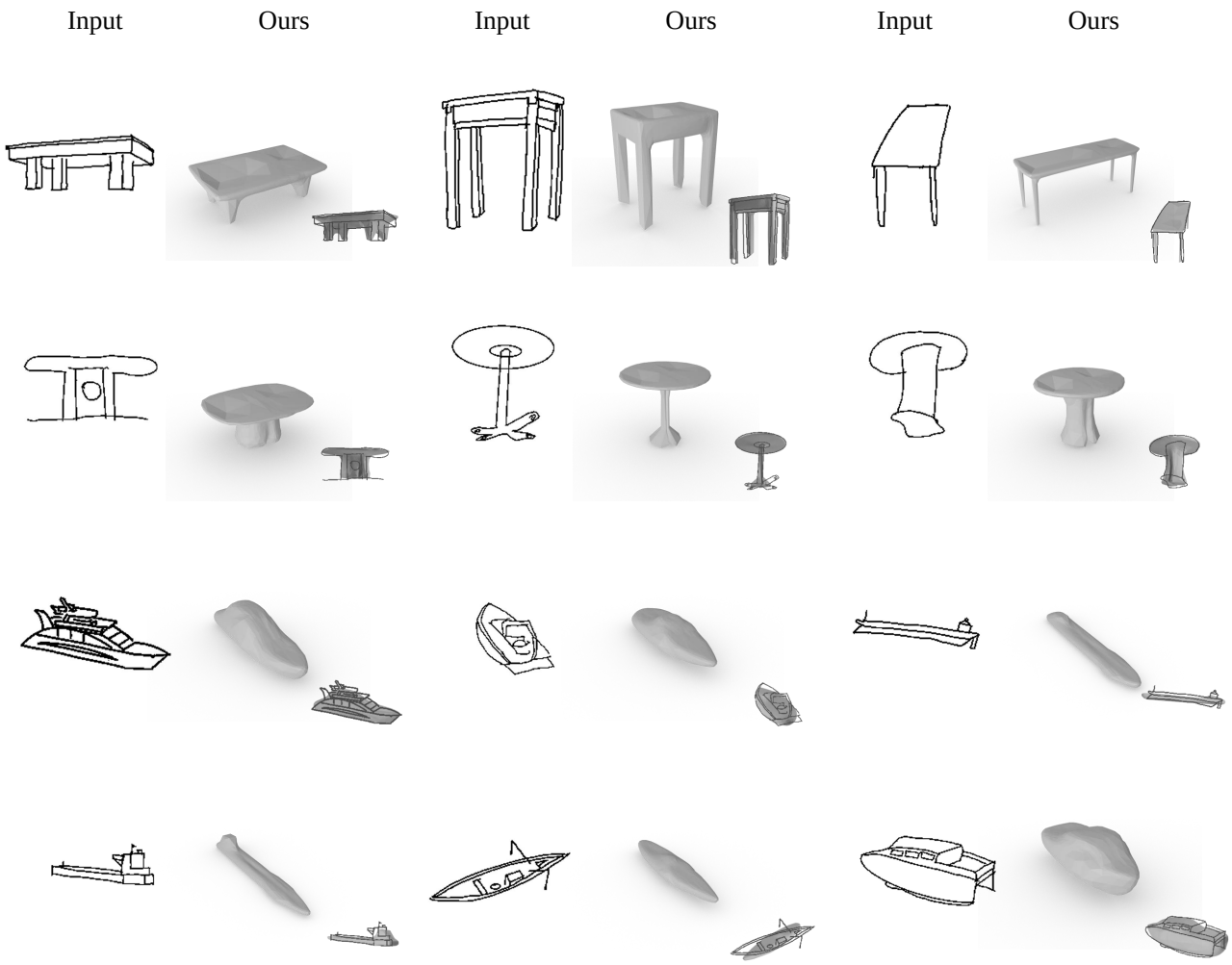

Figure 5. Additional results on ShapeNet-Sketch dataset.

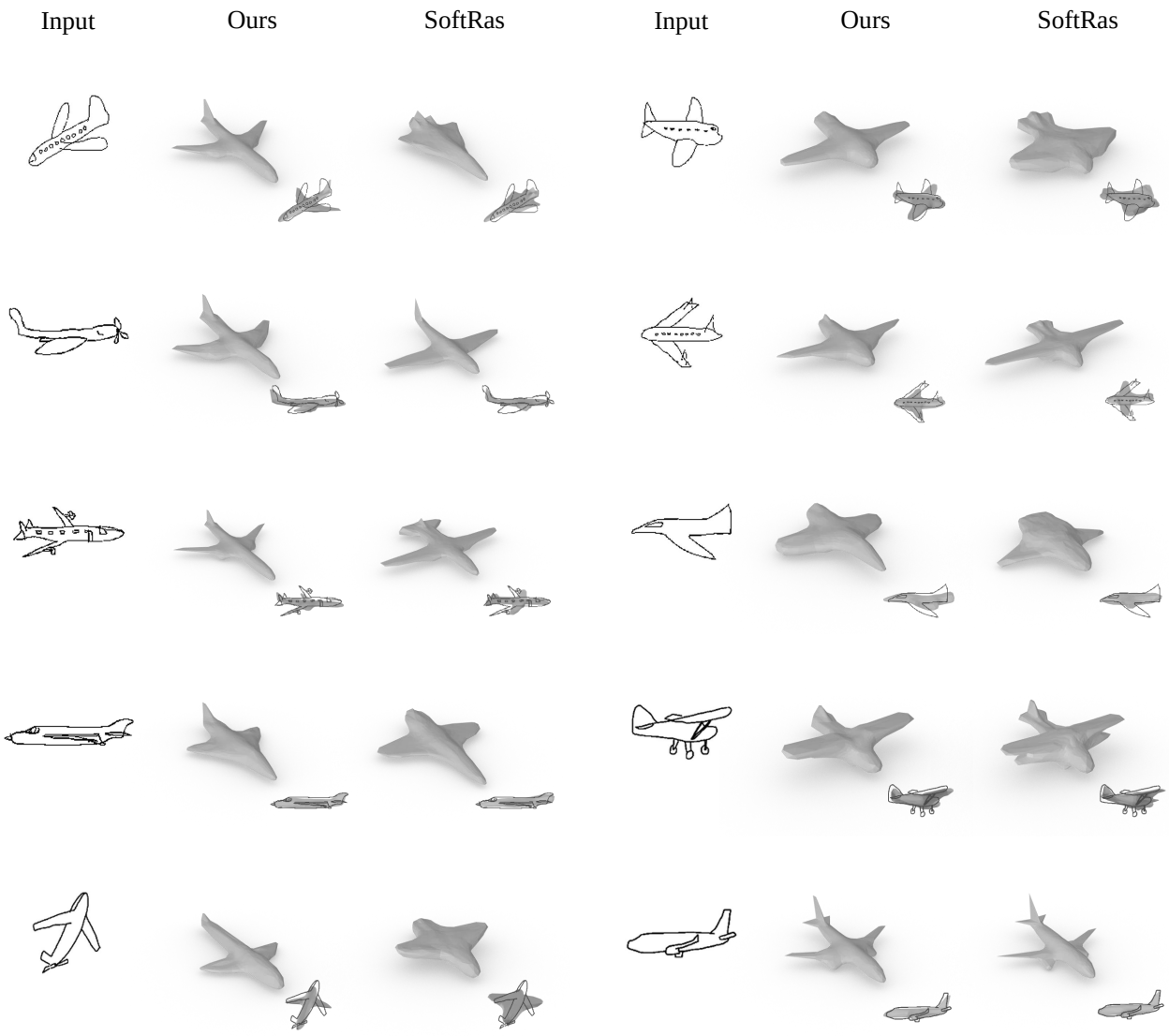

Figure 6. Additional results on free-hand sketch datasets.

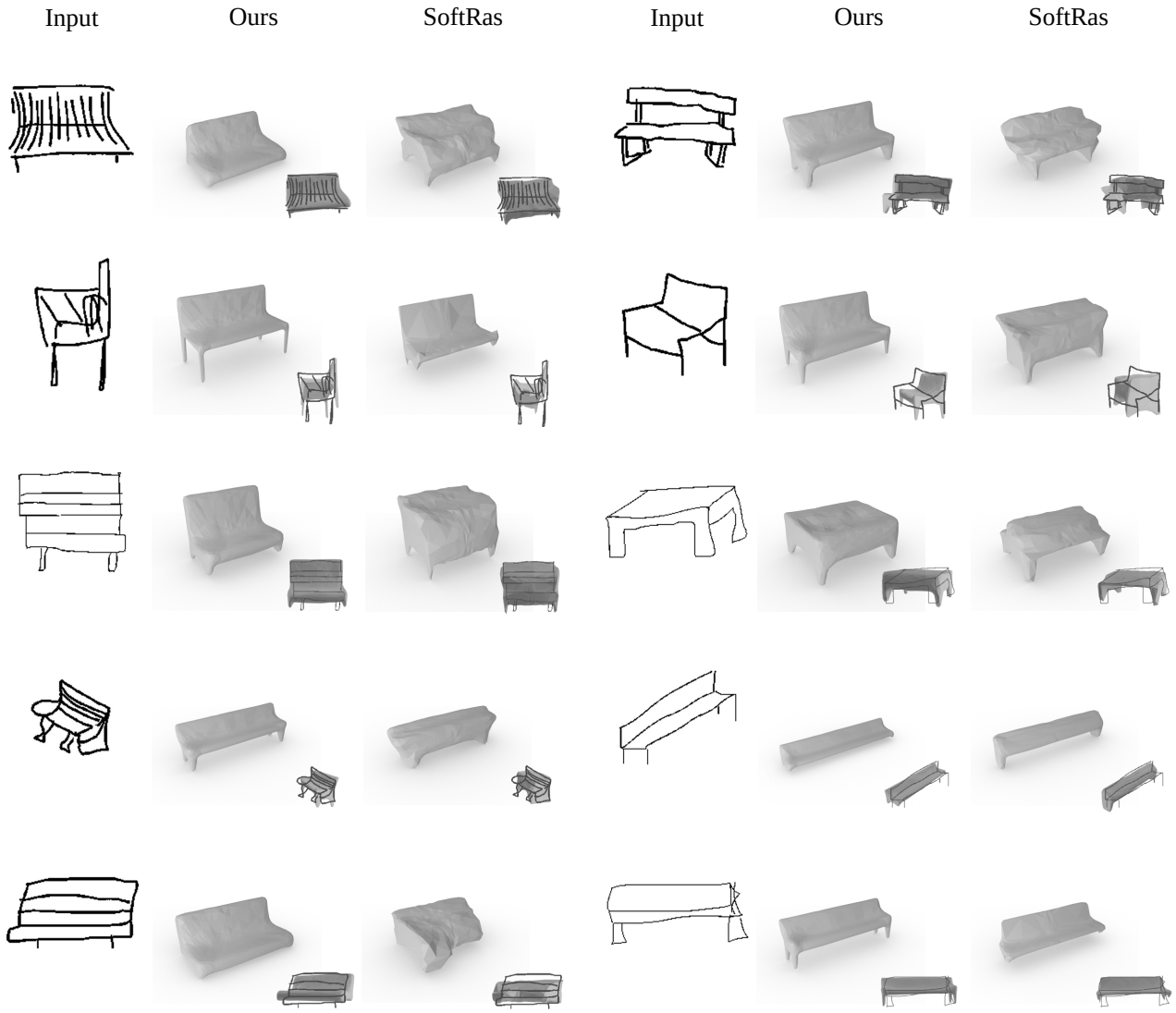

Figure 7. Additional results on free-hand sketch datasets.

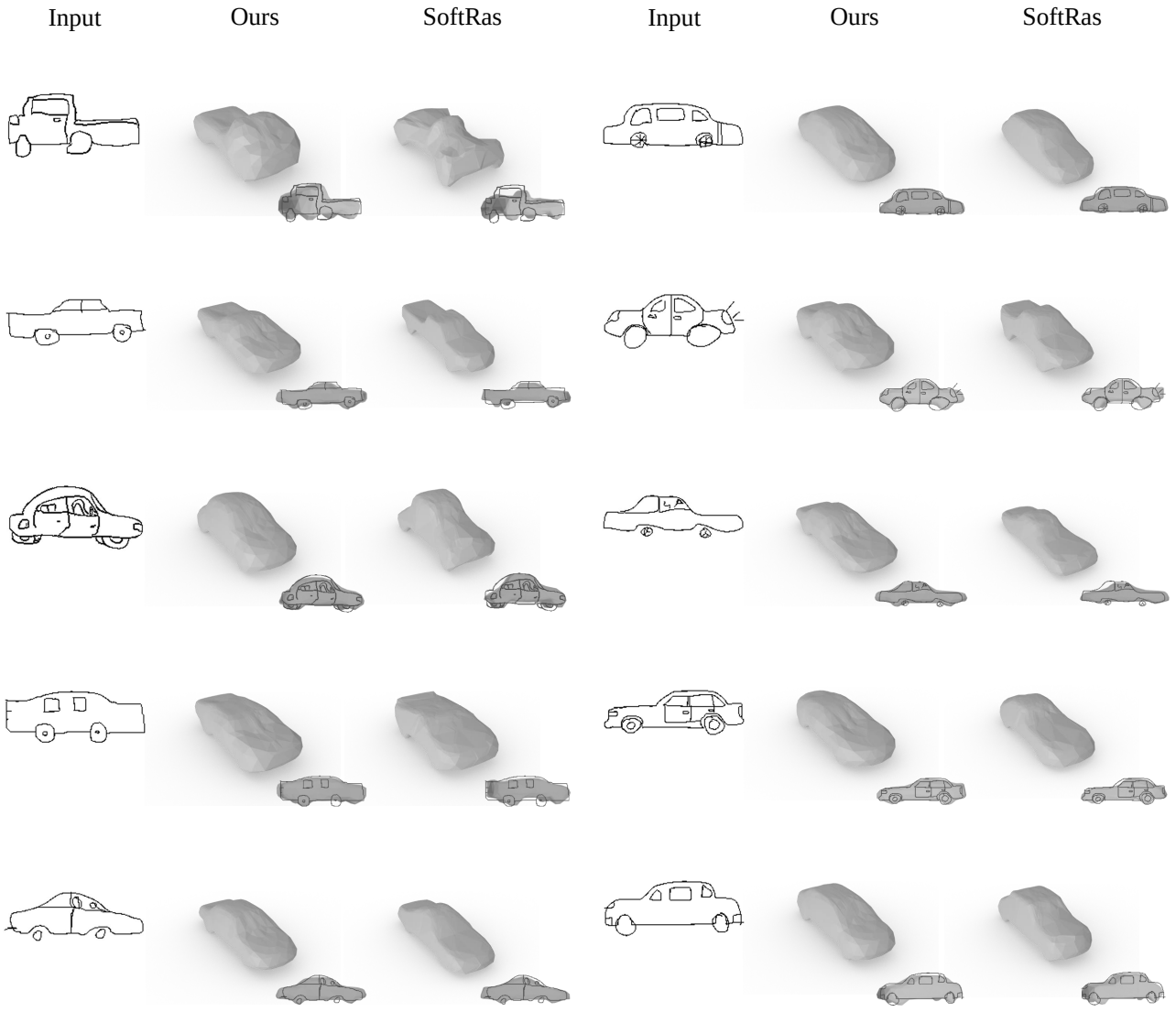

Figure 8. Additional results on free-hand sketch datasets.

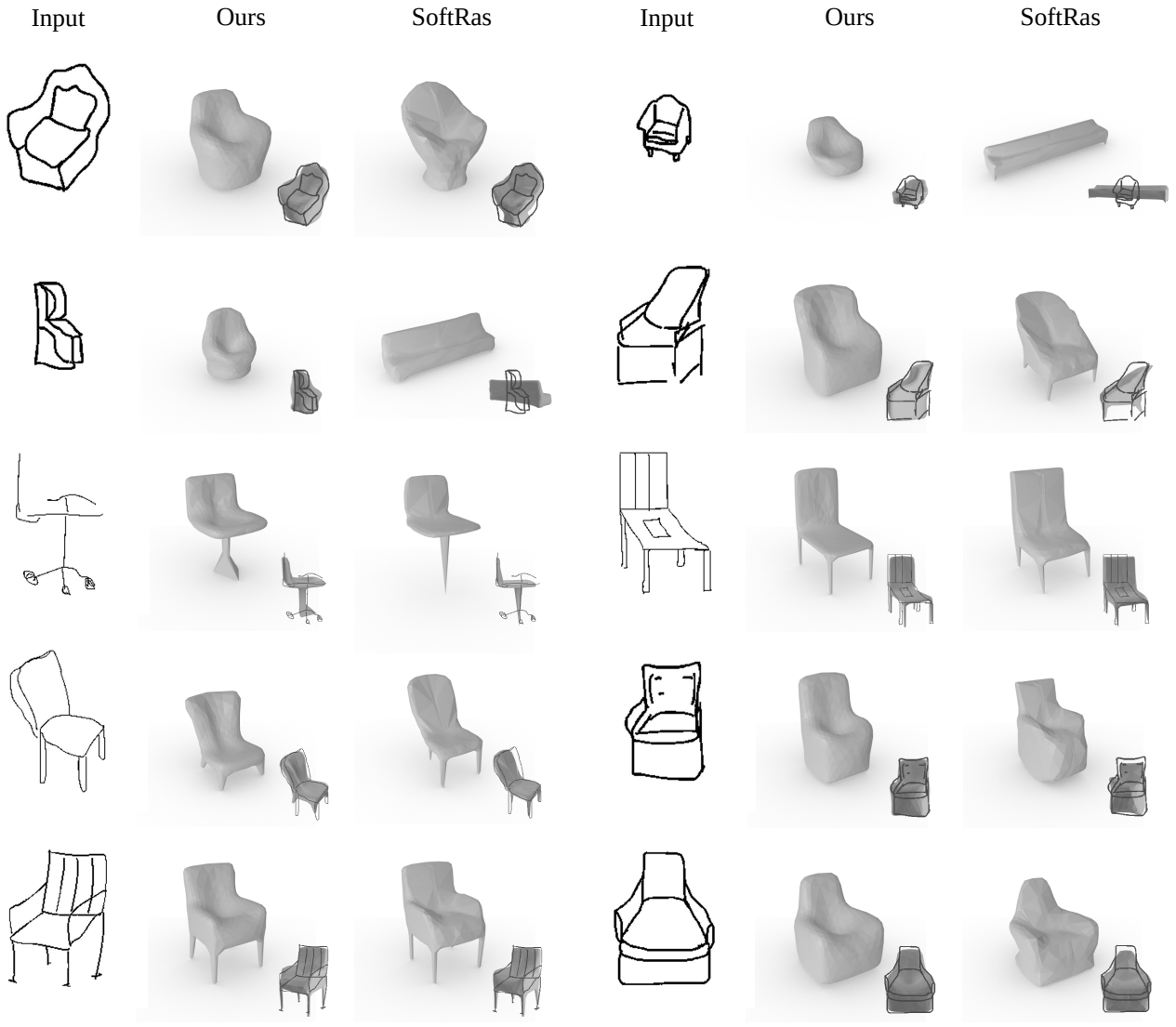

Figure 9. Additional results on free-hand sketch datasets.

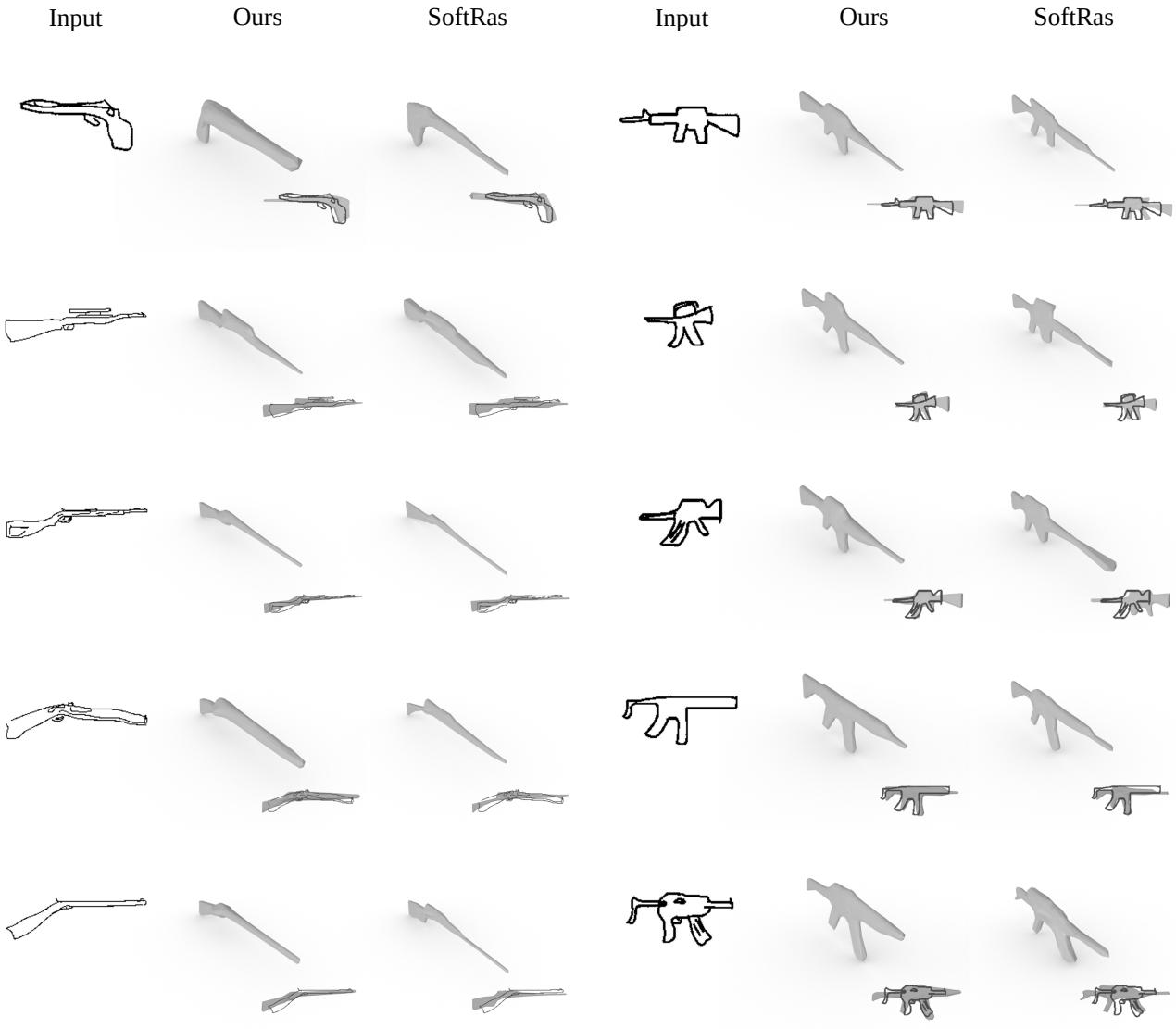

Figure 10. Additional results on free-hand sketch datasets.

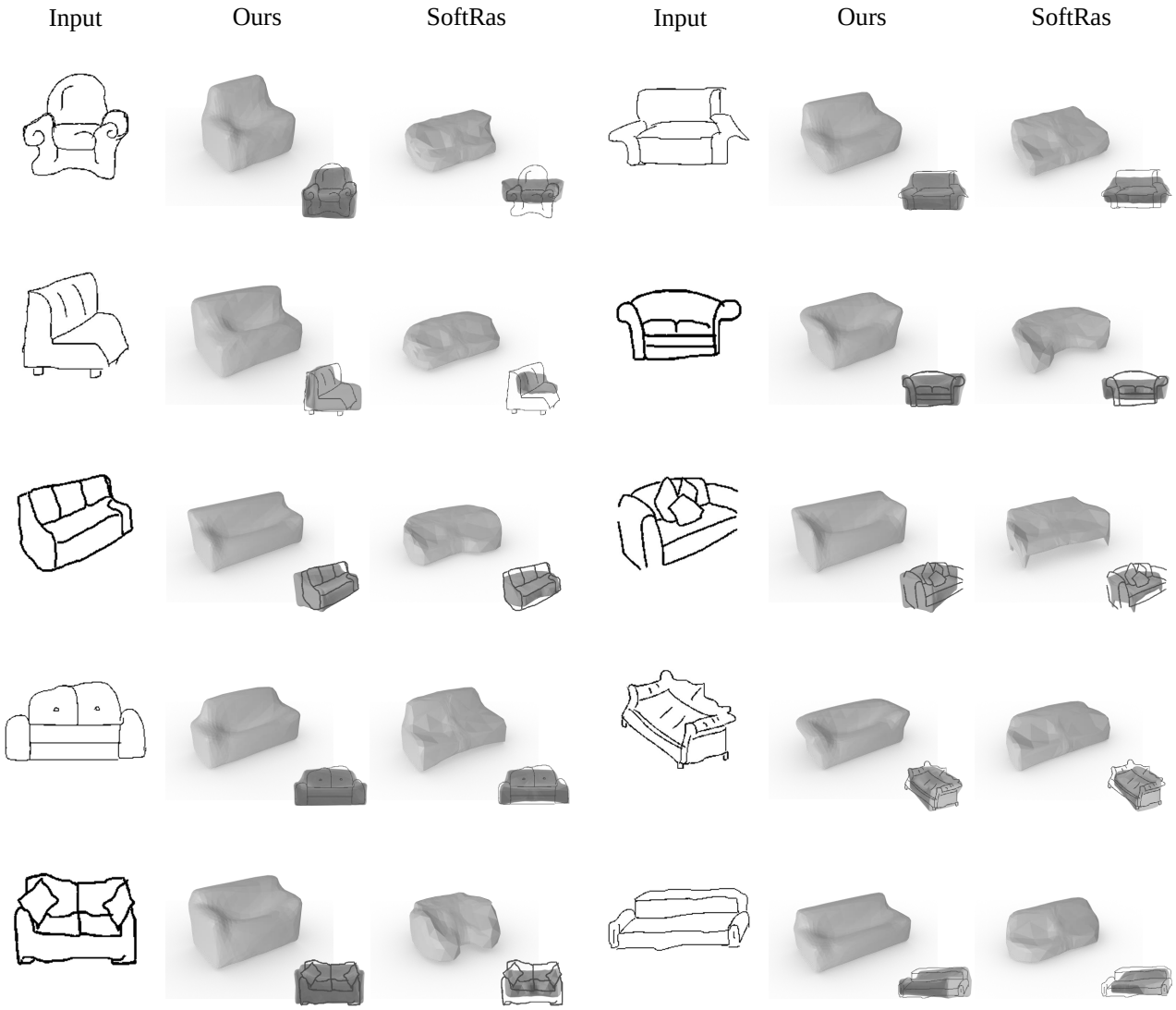

Figure 11. Additional results on free-hand sketch datasets.

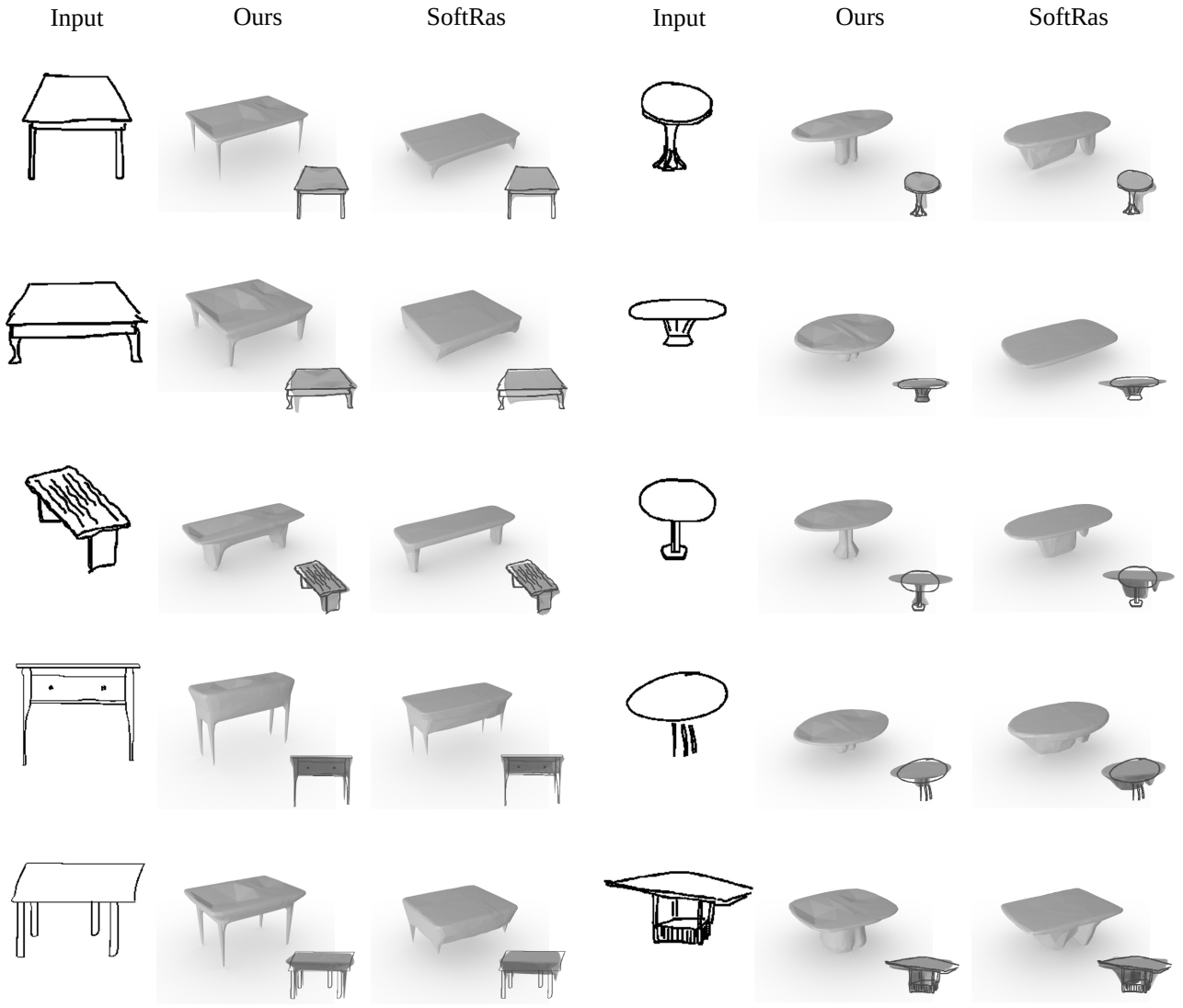

Figure 12. Additional results on free-hand sketch datasets.