Supplementary Material for the Paper "3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction"

1 Single-View Real-World Image Reconstruction

In this section, we present more single-view reconstruction results using the PASCAL VOC 2012 dataset and its corresponding 3D models from the PASCAL 3D+ dataset. Please refer to the main paper for more details.

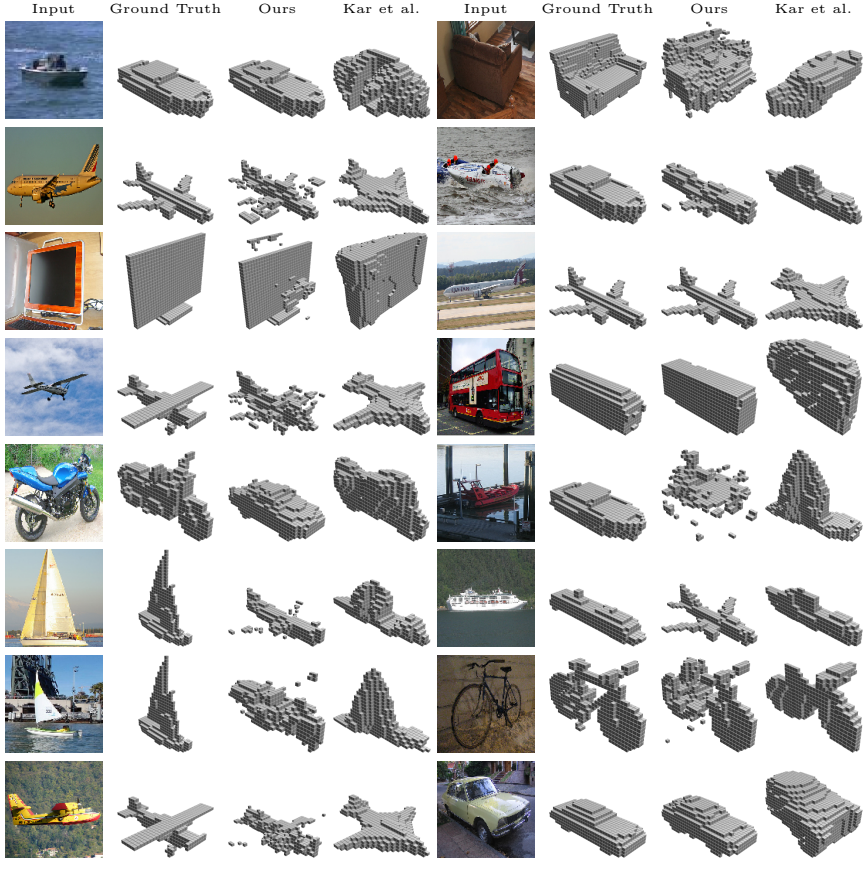

Fig. 1. Further reconstruction samples of PASCAL VOC 2012 dataset.

2 Online Product Dataset Reconstruction

In this section, we present more multi-view reconstruction results using the Online Product dataset. Please refer to the main paper for more details.

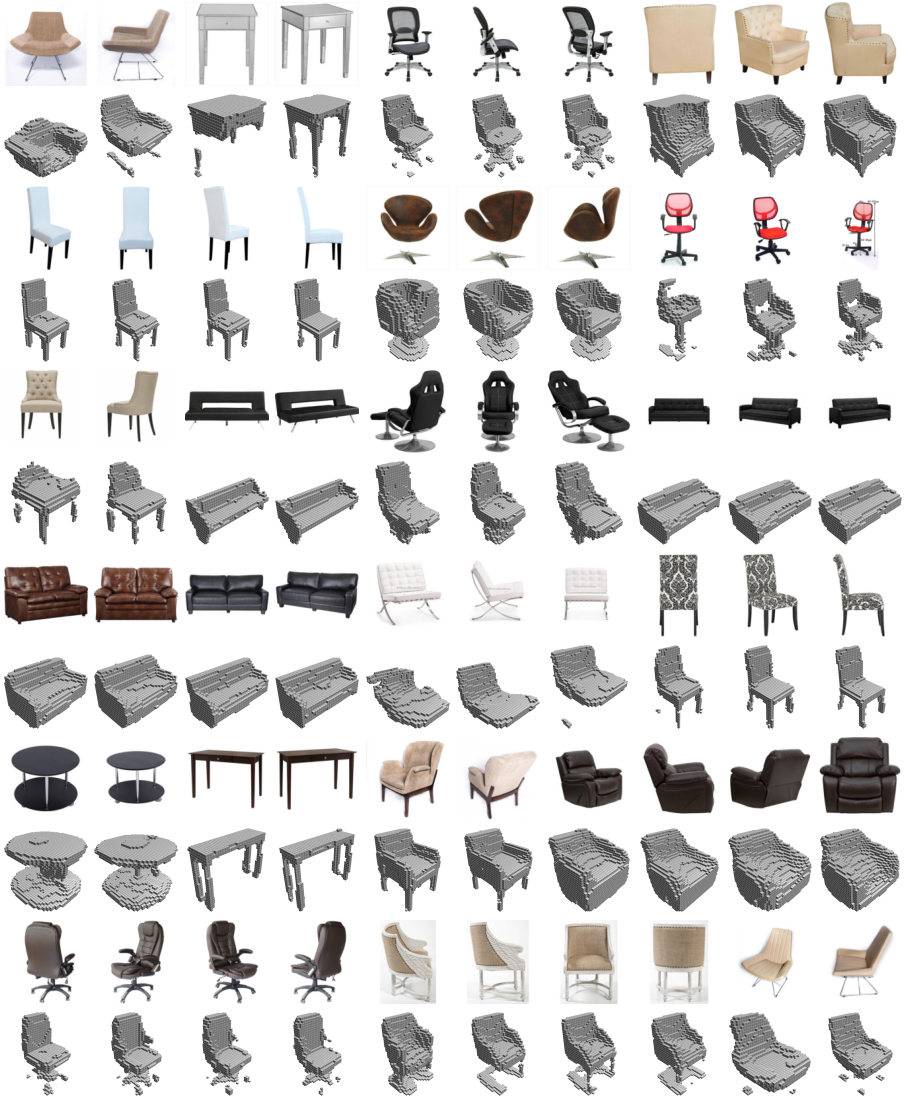

Fig. 2. Reconstructions of the Online Product dataset.

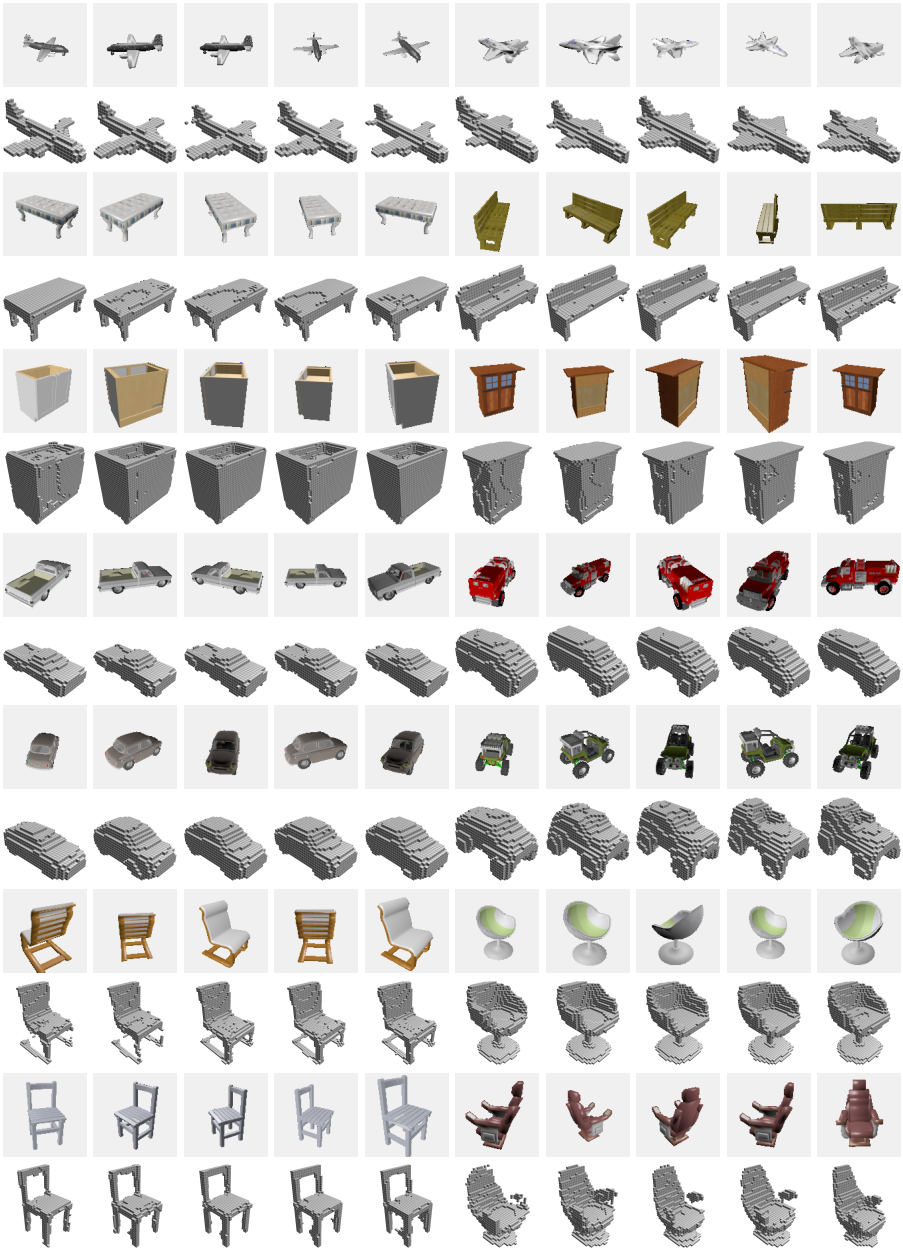

Fig. 3. Reconstructions of the ShapeNet testing set.

3 Network Input Gate Analysis

In this section, we visualize the network input gate activations to observe how the input gates react as new viewpoints become available. We used the [Res3D-GRU-3] network and tested on ShapeNet images taken sequentially from the front viewpoint to the side viewpoint for this controlled experiment. We set the viewpoints to be the same for all experiments to make the analysis easier, and the activations are shown in Fig. 4. First, we fed in the images sequentially (Fig. 4 top row) and selected a specific channel of the input gate activation to analyze, giving us a $4 \times 4 \times 4$ tensor. To visualize this tensor, we extracted four 4×4 grids, with the ith grid corresponding to the ith slice of the $4 \times 4 \times 4$ tensor in the 3rd dimension. Lastly, the grids are concatenated to create a 16×4 grid in a manner such that the top grid corresponds to activations at the top of the object and the bottom grid corresponds to activations at the bottom of the object.

For this particular example, we show the input gate activations from channel #48. The input gates open up strongly at first to make adjustments to the predictions but soon close down as the predictions get more accurate. However, when a part of the prediction mismatches the observation, the corresponding input gates open up to update the prediction.

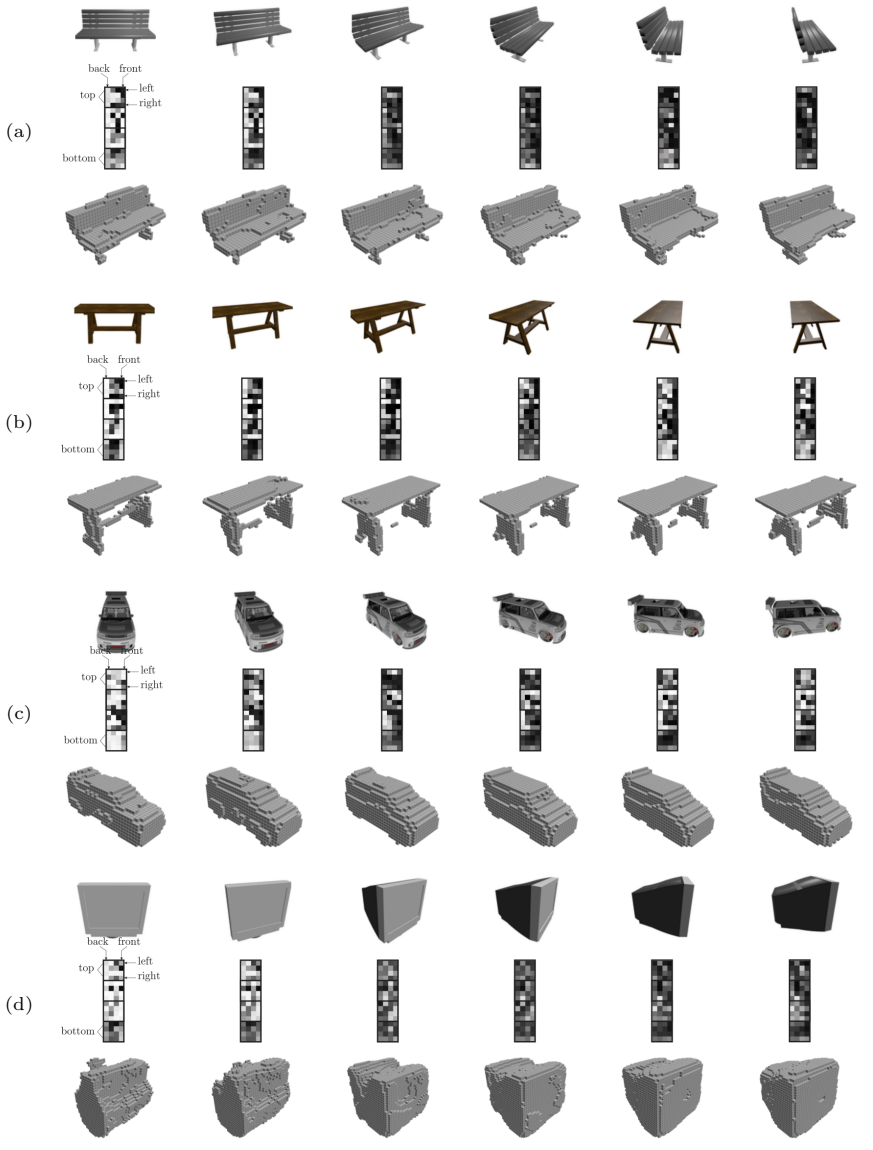

Fig. 4. Input gate channel #48 activations as the network observes more images. Top row: input images x_t, Center row: Input gate activations corresponding to the input image. White indicates strong activation (gate open), and black indicates no activation (gate closed). Bottom row: corresponding reconstructions.