

Table 1: Composition of our proposed GSD-S dataset. We collect glass images from four existing RGB image datasets with semantic annotations. Note that these datasets initially lack refined annotations of ground truth glass surface masks. When constructing our dataset, we re-labeled the GT masks for the glass surfaces.

Dataset	Whole	Train	Test
SUN RGB-D [38]	1,203	920	283
2D-3D-Semantics [39]	600	488	112
Matterport3D [2]	1,206	992	213
COCO-Stuff [37]	1,511	1,511	N/A
Total	4,519	3,911	608

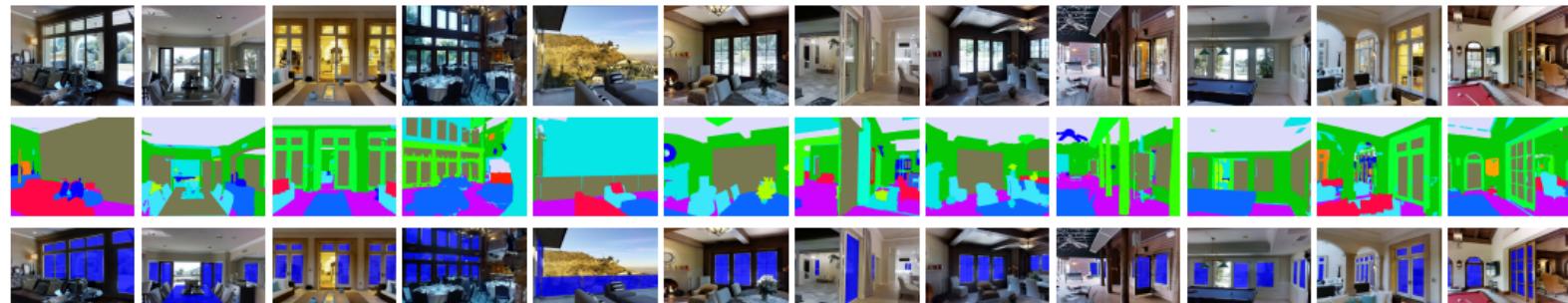


Figure 2: Preview of the GSD-S Dataset. Triplets of (RGB images, semantic maps, overlaid GT mask of glass surfaces).