A.7 GENERALIZATION PERFORMANCE FOR LEARNING-BASED MESH RECONSTRUCTION

In the following, we show our model's capabilities to generalize predictions. We trained models on the modalities point clouds (Table 13) and the combination of images and robot data (Table 14) on 13 different objects and evaluated on 4 unseen objects. The evaluation on unseen objects included all sequences.

Table 13: Generalization results for 4 unseen objects for point-cloud-based mesh reconstruction.

Input	$\mathcal{L}_{\mathbf{PFD}} \cdot 10^3 \downarrow$	$\mathcal{L}_{\mathbf{ROI}} \cdot 10^3 \downarrow$	**RPFD** $\downarrow$	$\mathbf{CD}_{\mathbf{UL1}}$**[mm]** $\downarrow$	$J(M_{\mathbf{P}}, M_{\mathbf{GT}}) \uparrow$
Validation set (13 objects)	3.93	3.51	0.698	6.229	0.836
Foam cylinder	4.78	2.38	0.652	8.659	0.794
Plush volleyball	2.78	3.16	0.182	6.260	0.899
Sponge	9.34	3.03	0.603	8.314	0.731
Toilet paper roll	15.43	10.67	0.387	7.546	0.754

Table 14: Generalization results for 4 unseen objects using images and robot data as input.

Input	$\mathcal{L}_{\mathbf{PFD}} \cdot 10^3 \downarrow$	$\mathcal{L}_{\mathbf{ROI}} \cdot 10^3 \downarrow$	**RPFD** $\downarrow$	$\mathbf{CD}_{\mathbf{UL1}}$**[mm]** $\downarrow$	$J(M_{\mathbf{P}}, M_{\mathbf{GT}}) \uparrow$
Validation set (13 objects)	5.07	4.36	0.737	6.840	0.814
Foam cylinder	8.13	5.42	1.206	10.820	0.738
Plush volleyball	8.83	8.19	0.452	9.561	0.829
Sponge	15.90	8.409	1.117	8.410	0.640
Toilet paper roll	37.31	40.45	0.650	9.952	0.667