Table 1: We evaluate reconstruction accuracy for "*challenge*" tasks that come from concepts / categories not present in the target training set. For both layout and 3D CSG, we observe that our joint paradigm that integrates an edit network with *one-shot* models outperforms the alternative of using only *one-shot* models.

	Layout *cIoU* ⇑	3D CSG *IoU* ⇑
OS Only	75.8	60.8
OS + Edit (Ours)	**87.6**	**70.9**

Method	Chamfer Distance ⇓
Ours (default)	**0.111**
No FT	0.321
No one-shot FT	0.230
No edit FT	0.123
No edit PT	0.145

Figure 1: We extend the ablation study from our main submission and evaluate how a subset of these conditions performs on the *2D CSG* domain.

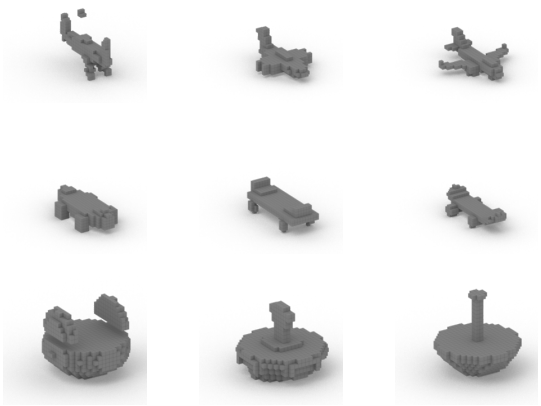

Figure 2: Qualitative reconstructions of "challenge" tasks for the 3D CSG domain.

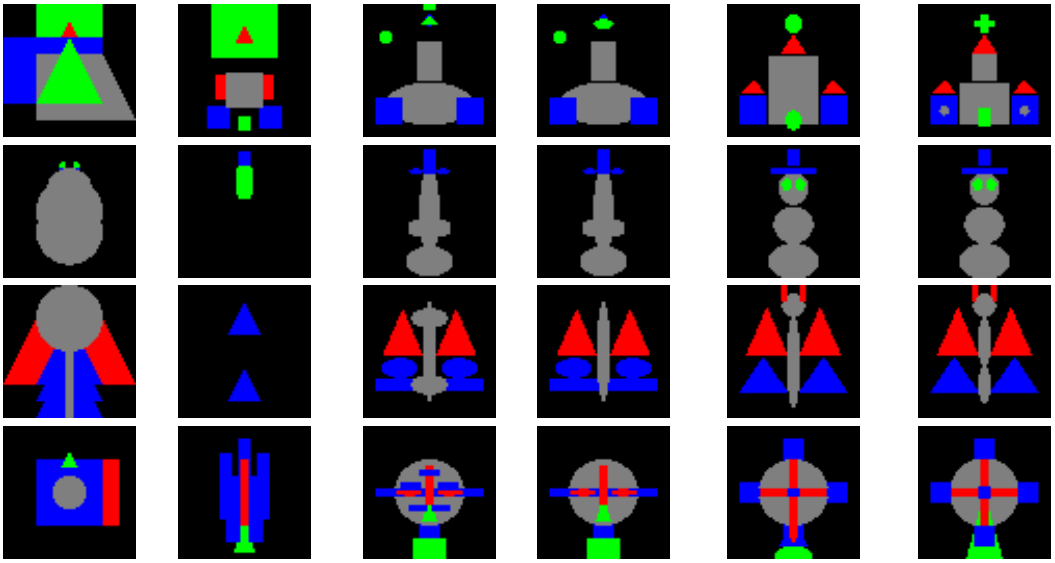

Figure 3: Qualitative reconstructions of "challenge" tasks for the layout domain. We compare against GPT-4V in a zero-shot setting (column 1), when an in-content example (ICE) is provided in the prompt (column 2), and when the one-shot model's predicted program is provided as input (column 3). Our approach (column 5) finds more accurate reconstructions of these out-of-distribution targets (column 6) compared with using only the one-shot network (column 4).