

CHATREARRANGE: LEARNING TEXT-GUIDED 3D SCENE REARRANGEMENT

Anonymous authors

Paper under double-blind review

1 CODE AND DATA SAMPLE

For reproducibility, the implementation and data sample are provided in the anonymous repository. Please refer to `Code/README.md` for detailed setup, data usage, and model training and testing instructions.

2 QUALITATIVE RESULTS

More qualitative results are presented in Fig. 1 (in the next page). Each row represents a single task with the text input in the first column. The furniture-related terms are highlighted in blue, and orientation-related descriptions are colored in red. Our model’s predicted layout aligns the best with the text input and the ground truth. However, the baselines often fail to preserve all the described spatial relationships and latent features contained in the furniture labels. These results suggest our model has promising performance in the text-guided 3D scene rearrangement task.



Figure 1: Qualitative comparison between LLM baselines and our model. Each row represents a single task. The first column shows the input text prompt, with furniture-related terms highlighted in blue and direction-related terms colored in red. The ground truth is shown in the last column. Our model captures both the latent and explicit spatial relationships described in the text, generating layouts that most closely align with the input descriptions. In contrast, the baseline models often fail to preserve all key spatial relationships and semantic coherence.