

Peer Review

Review of: "LiLMaps: Learnable Implicit Language Maps"

Motonari Kambara¹

1. Independent researcher

The authors propose a segmentation model for 3D environments. In the context of object manipulation and navigation in household settings, accurately recognizing the positions of furniture and objects in the environment is crucial. In this regard, the task tackled by the authors is considered to have significant research value. Furthermore, the newly proposed update strategy effectively addresses the shortcomings of existing methods, and its effectiveness is demonstrated through experimental results. The experiments are comprehensive, and both quantitative and qualitative evaluations clearly support the validity of the proposed method. In particular, the qualitative results are presented in a clear and comprehensible way.

Questions

- In Table 2, I suggest reordering the methods to improve readability. Specifically, grouping the baseline methods (OpenScene, VLMaps) separately from the proposed methods would make the table easier to follow.
- Regarding the performance gap between LiLMaps_SEM and LiLMaps_LSeg, should this be interpreted as a direct reflection of the performance difference between SEM and LSeg themselves?
- In Table 3, the performance difference between LiLMaps (simple) and LiLMaps (pretrained) appears to be minimal. Similarly, OpenScene (Matterport3D) and OpenScene (nuScenes) also show little difference. Does this imply that the task was relatively easy?
- If nuScenes and ScanNet can be used as training data, it would be beneficial to include evaluation results on those datasets as well.

Declarations

Potential competing interests: No potential competing interests to declare.