

Highlights

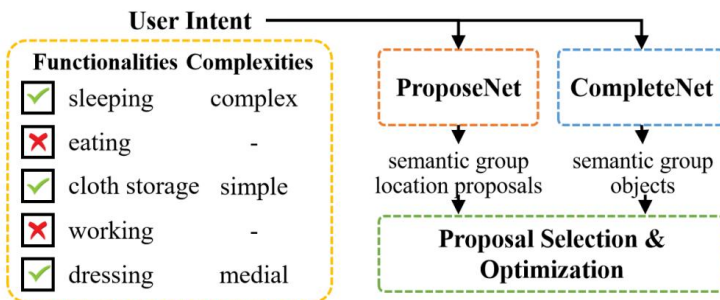
Goal: Building a novel scene synthesis framework for enabling the flexible generation of individualized indoor virtual environments.

Key Contributions:

- Propose a novel flexible propose-and-complete scene generation framework that supports personalized semantic function settings.
- Design a normalizing flow-based ProposeNet for learning the group-level position and rotation distributions.
- Devise a conditional graph variational auto-encoder, CompleteNet, to generate entire semantic groups with user-desired group complexities.

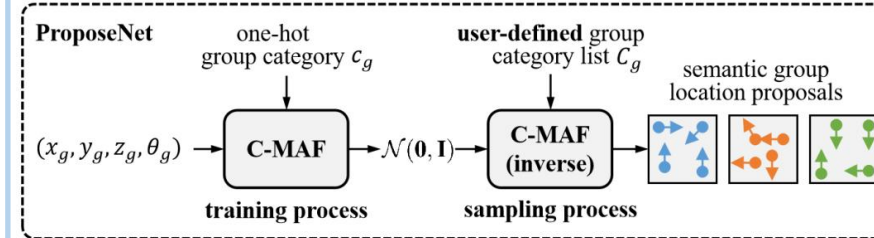
Introduction

Critical Idea: Semantic function is a natural clustering that compactly fills the space delicately. Conversely, we can enrich an empty room by adding various semantic groups with flexible complexities, with a propose-and-complete strategy. Furthermore, we adopt user-specific semantic group categories user-desired object number to guide a conditional-based generative process, in order to meet individualized and flexible generation requirement.



New Method

ProposeNet:

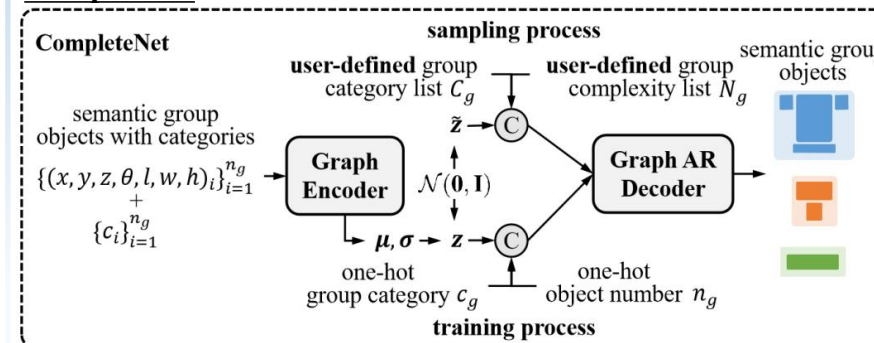


We extend the original MAF model towards a conditional version based on our conditional density $p(\mathbf{x}|c_g)$. The auto-regressive model conditioning on the semantic group category is:

$$p(x_i | \mathbf{x}_{1:i-1}, c_g) = \mathcal{N}(x_i | \mu_i, (\exp \alpha_i)^2)$$

The user selects the room functionalities as the conditioning feature. Our ProposeNet supports sampling the arbitrary combination of group location proposals regardless of the semantic co-occurrence of the training rooms.

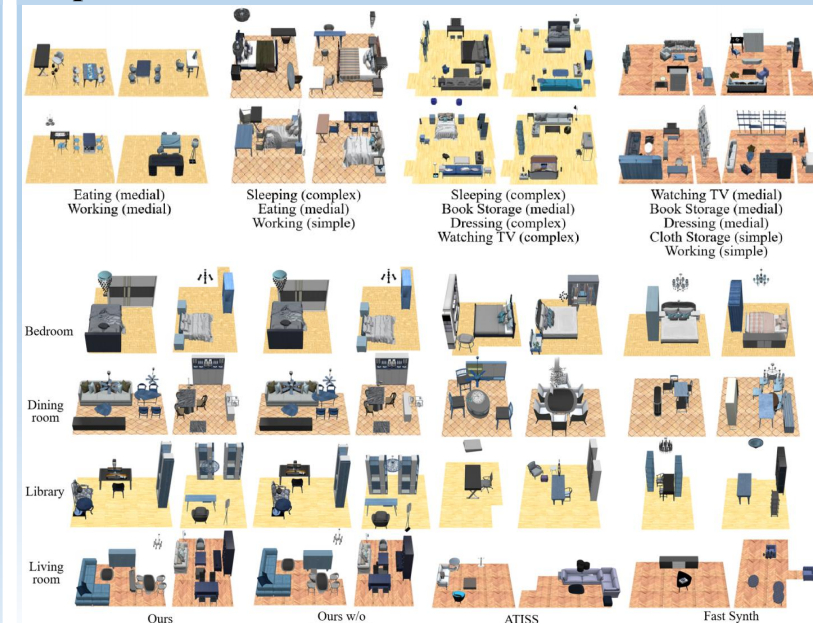
CompeteNet:



CompleteNet has a variational graph auto-encoder structure to generate the group objects, conditioning the group semantics and its target complexity.

We recursively select the most proper location for each group category by scoring the candidates with several common-sense criteria: inter-group collision, group accessibility, free zone area, and compatibility of the floor shape.

Experimental Results:



	FID (↓)			
	Bedroom	Dining	Library	Living
Training Set	56.147	67.318	64.755	49.546
FastSynth	68.871	80.956	76.469	76.629
ATISS	60.304	75.838	69.074	57.580
Ours (w/o opt.)	61.769	75.683	69.022	57.025
Ours ($N_p = 50$)	60.810	74.408	68.340	55.344

Conclusion

- We articulated a novel propose-and-complete framework supporting custom-made scene layout generation with high versatility.
- The key innovation is founded upon a flexible combination of indoor functional semantic groups, with which we propose potential group-level locations by the ProposeNet and complete the detailed intra-group objects by the CompleteNet in a divide-and-conquer fashion.