



ObjVariantEnsemble: Advancing Point Cloud LLM Evaluation in Challenging Scenes with Subtly Distinguished Objects

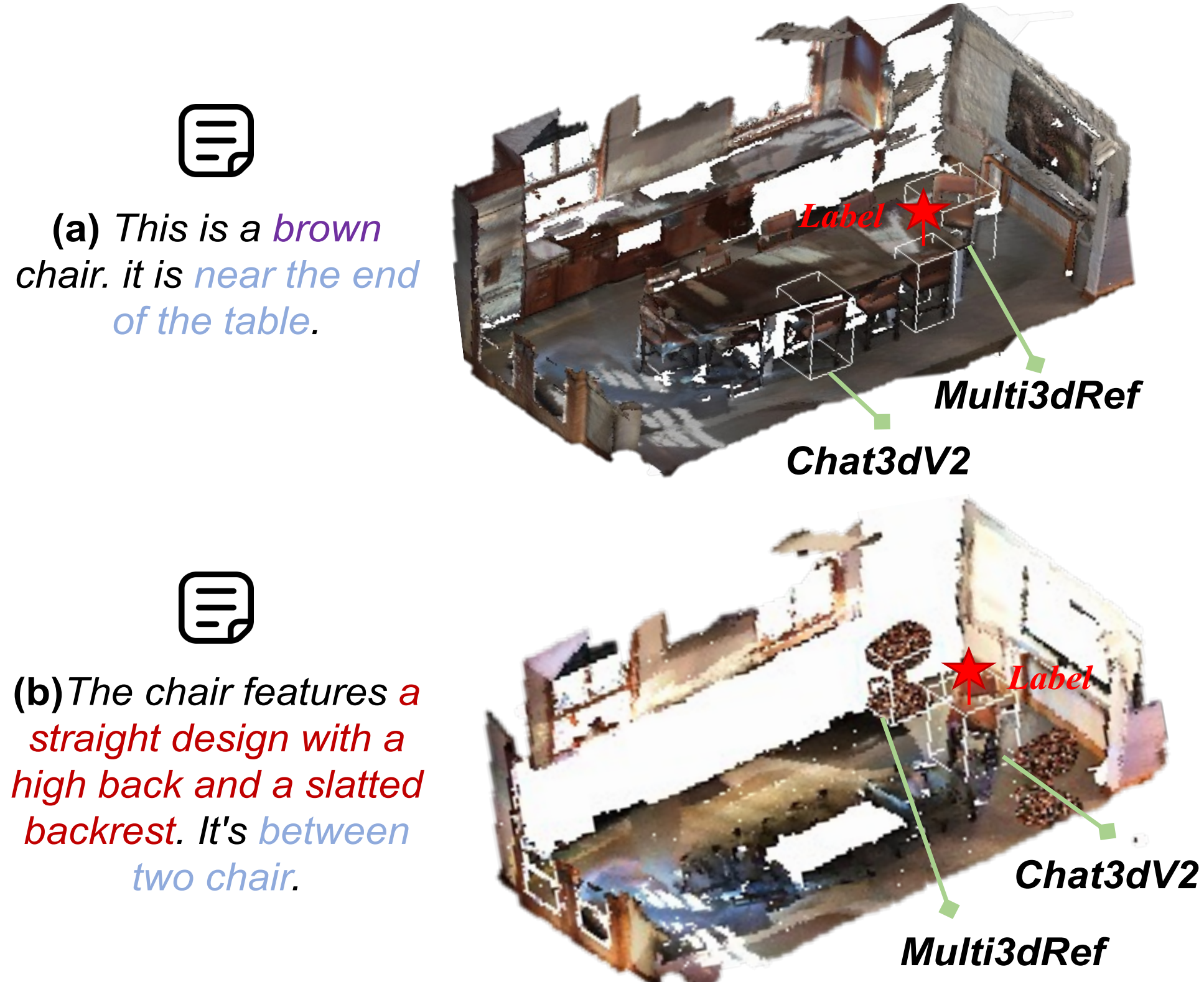


Qihang Cao^{1,2}, Huangxun Chen²

¹ Shanghai Jiao Tong University ² Hong Kong University of Science and Technology (Guangzhou)

Motivation

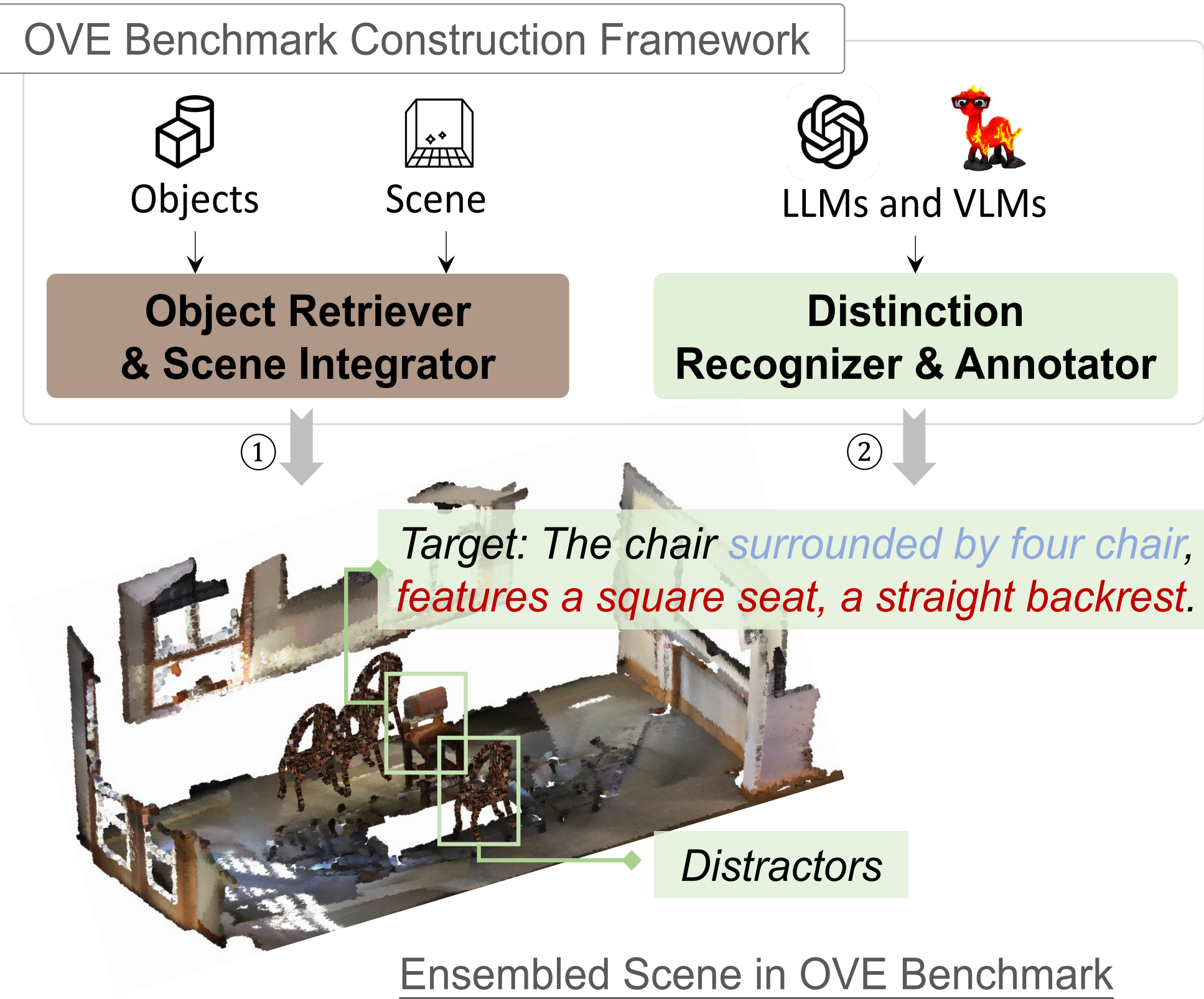
What is challenging scenes with subtly distinguished objects?



Problems of Current benchmark?

- Small scale and insufficient annotations.
- Lack of customizable challenge levels.

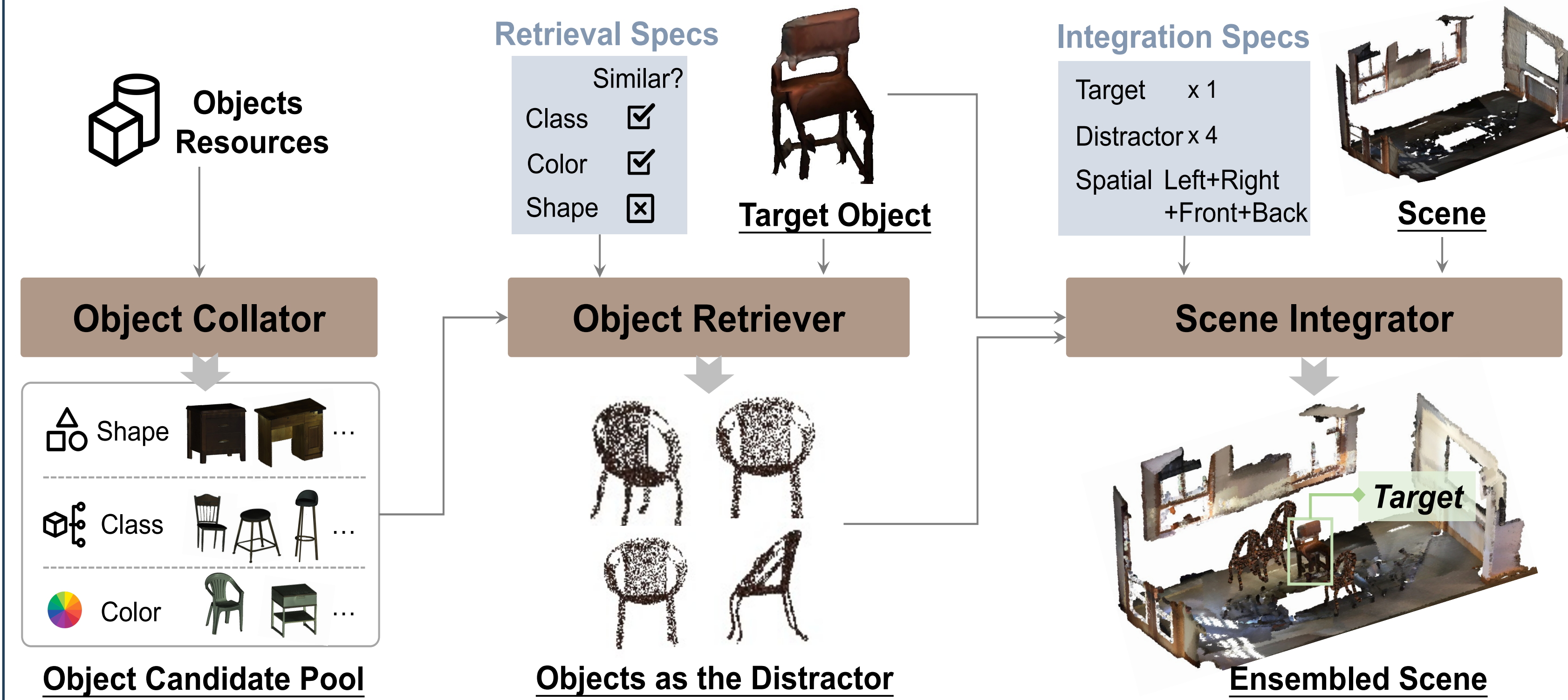
Our Idea



- A systematic and flexible method to introduce subtly distinguished objects adjacent to the target.
- An automated scheme to capture key differences between similar objects.

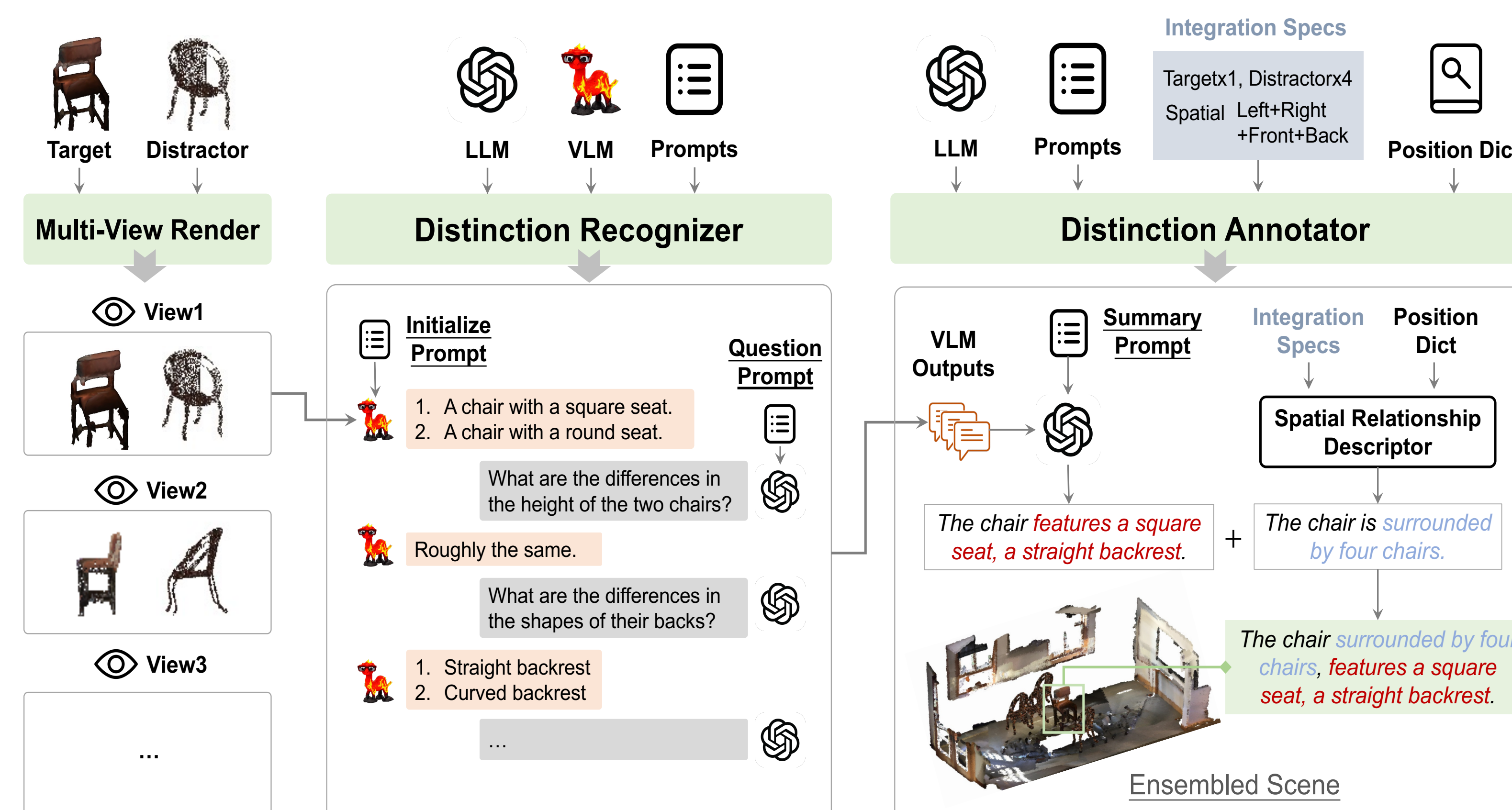
Key Design

1. Scene Data Generation



- Sorts available object-level 3D resources by their features (Class, Color, Shape).
- Assemble multiple objects with varying levels of similarity into a specific scene background to construct a new scene.

2. Capturing Annotations with Key Distinguishing Information



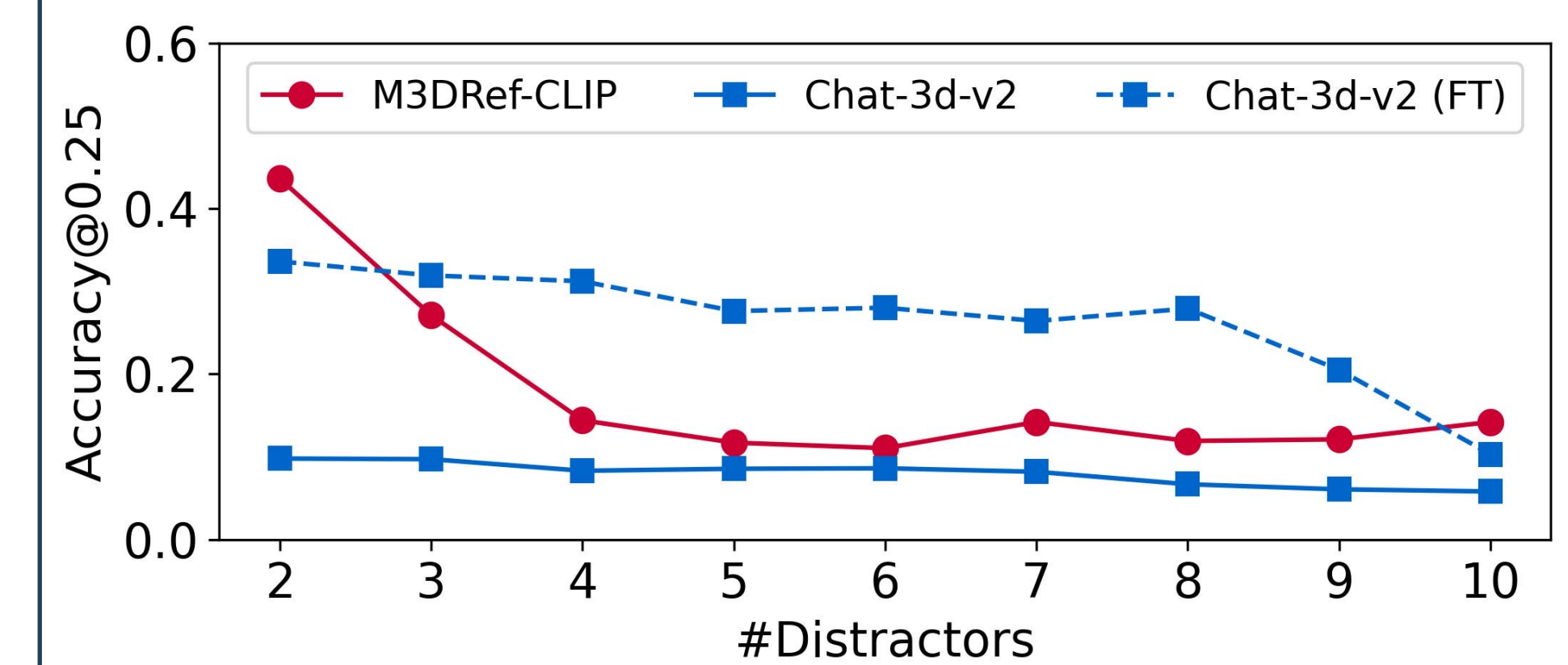
- Leverages LLMs + VLMs to extract the key characteristics that distinguish the target object from others.
- Summarize these as annotations.

Benchmark Summary

3D Benchmark	#Obj	s-t pairs (wd/wod)	Anno (#Obj)
ModelNet	12k	-	Cl
ShapeNet	51k	-	
ScanObjectNN	15k	-	
OmniObject	6k	-	
ScanRefer	-	39k/12k	Cl+s/co/l (1)
Multi3DRef	-	33k/29k	Cl+s/co/l (0/1/2..)
OVE(Ours)	-	75k/12k	Cl+s/co/l (1 wd)

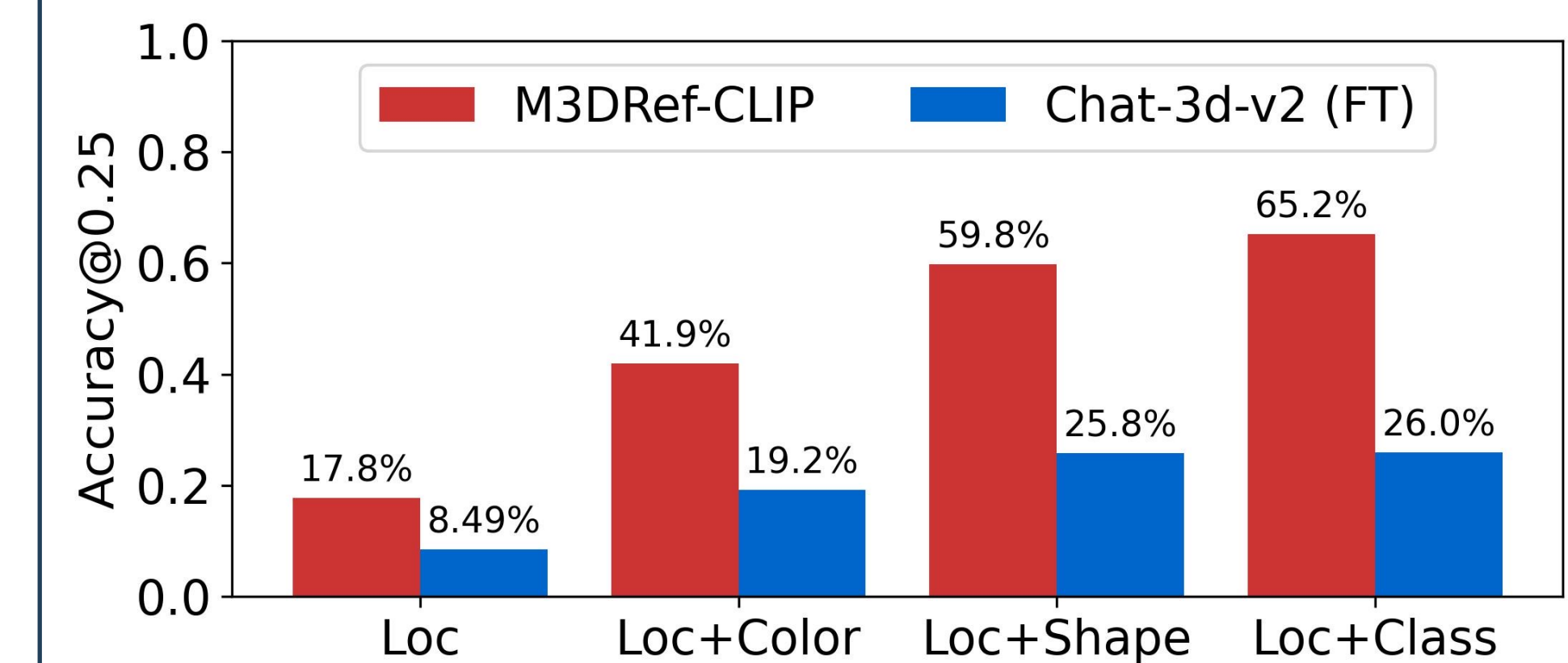
Evaluation on SOTA 3D-LLM

Impact of Number of Distractors



#Distractors ↑ → Model performance ↓

Impact of Number of Distinction Type



For 3D grounding, contributions of Class > Shape > Color > Location

Welcome to scan this code to access our project page !

