

Low-shot Object Learning with Mutual Exclusivity Bias

Anh Thai, Ahmad Humayun*, Stefan Stojanov*, Zixuan Huang, Bikram Boote, James M. Rehg



Motivation from Developmental Psychology

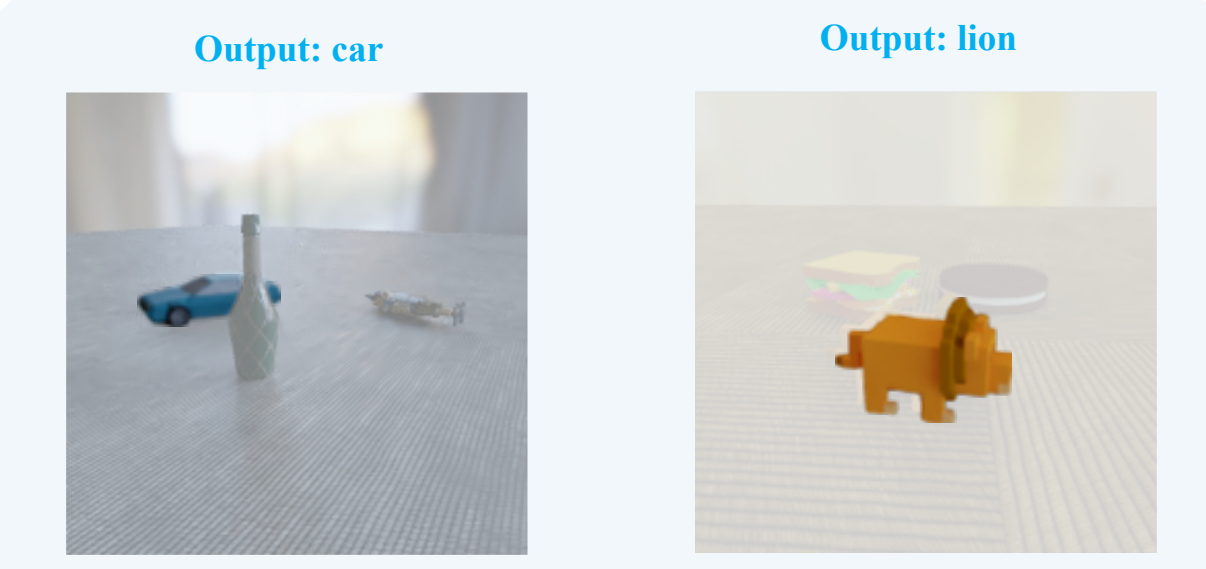
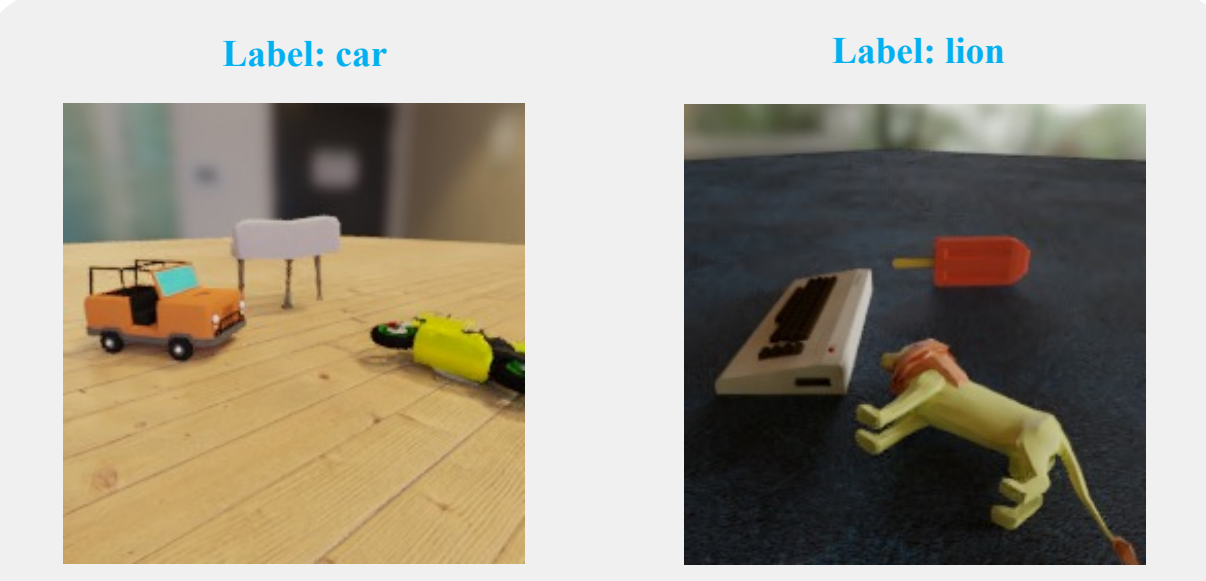


Mutual Exclusivity Bias in children: tendency to associate *novel word* with *novel object*

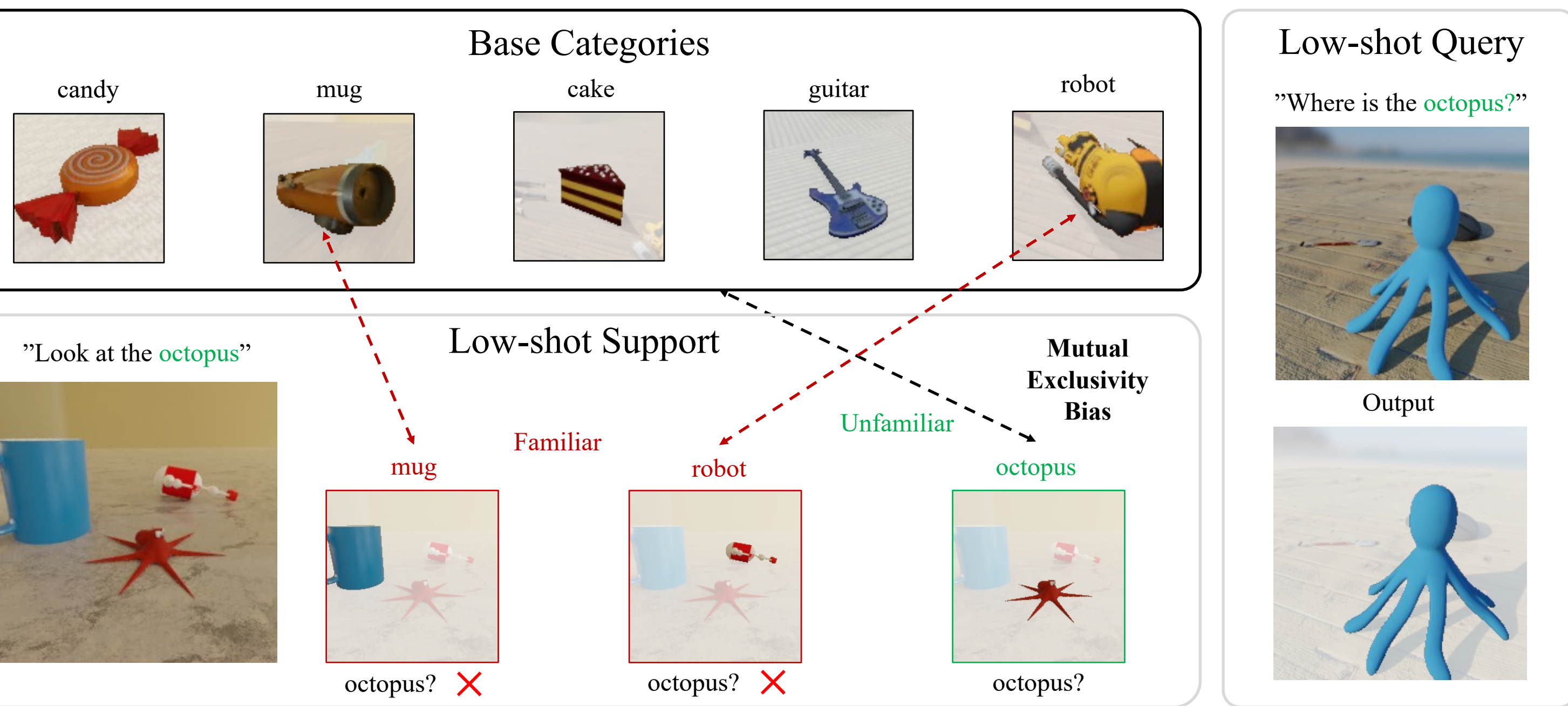
Standard Low-shot Learning



Low-shot Learning with Mutual Exclusivity Bias

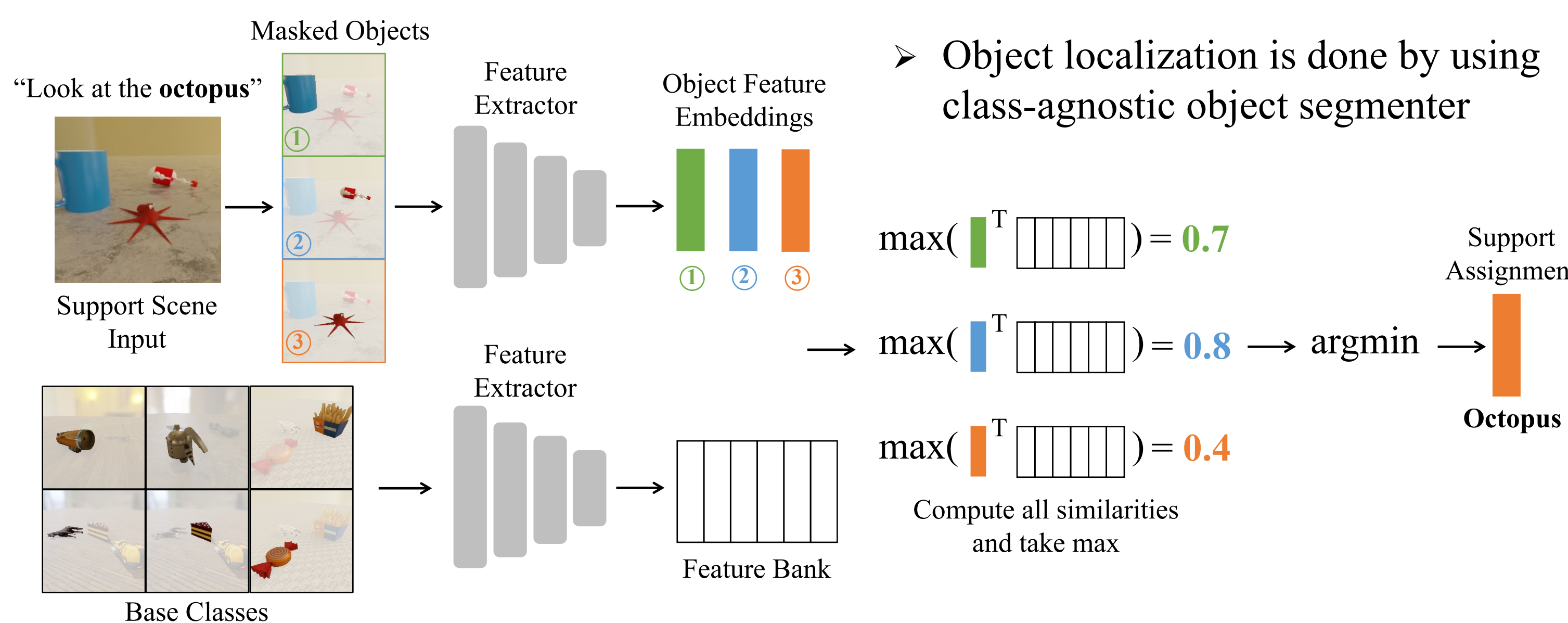


Low-shot Object Learning with Mutual Exclusivity Bias (LSME)



3 sub-tasks: ➤ Object Localization ➤ Open-world Recognition ➤ Low-shot Learning

Computational Modeling of Mutual Exclusivity Bias



➤ Assigning the **least familiar** object with the **novel word**

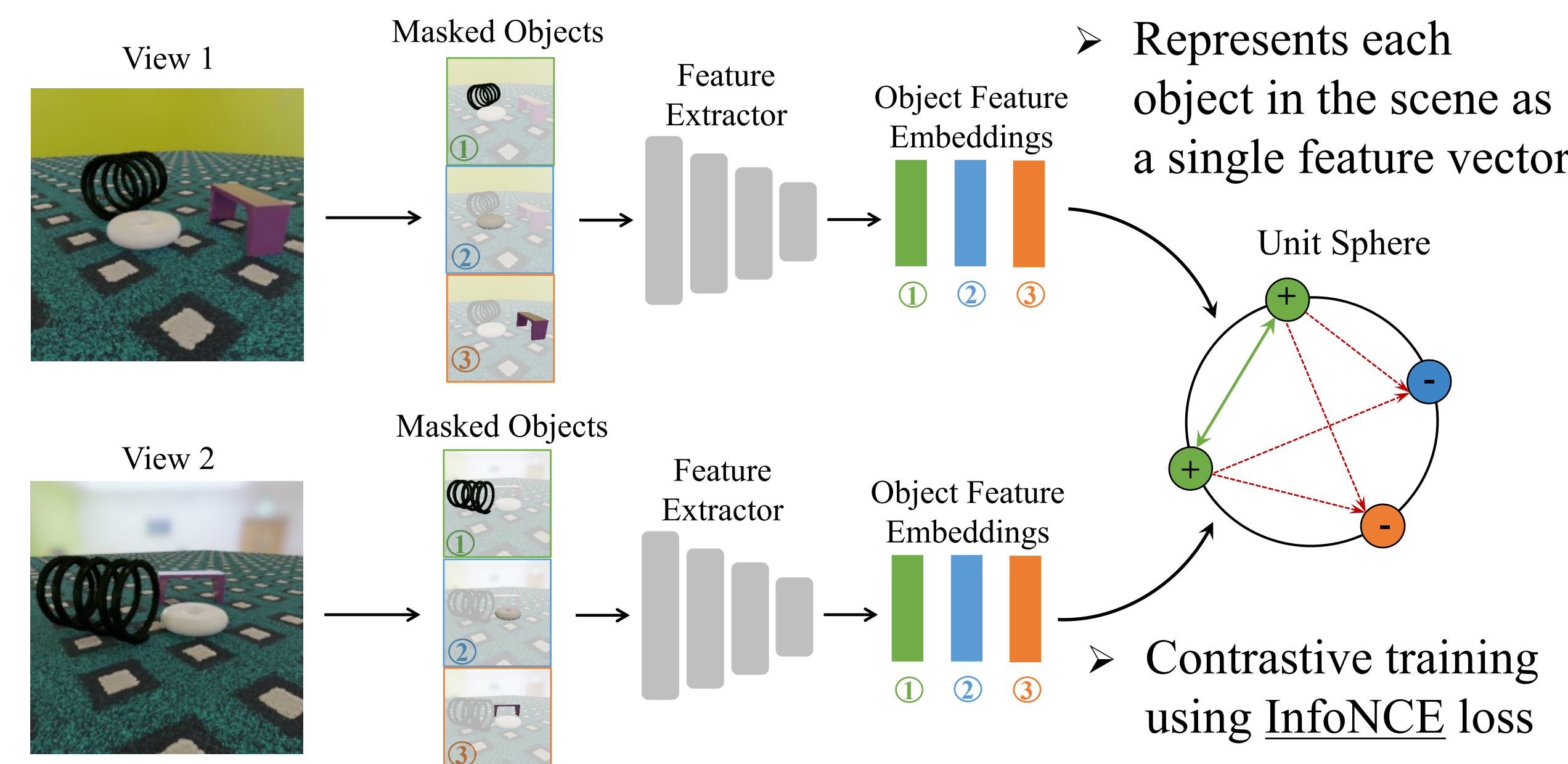
Related Tasks

	Open Vocabulary Detection	Object Discovery	Open Set Detection	LSME
Mutual Exclusivity Bias	✗	✗	✗	✓
Discover Novel Classes	✓	✗	✓	✓
Label Novel Classes	✓	✗	✗	✓
No Pretrained LLM	✗	✓	✓	✓

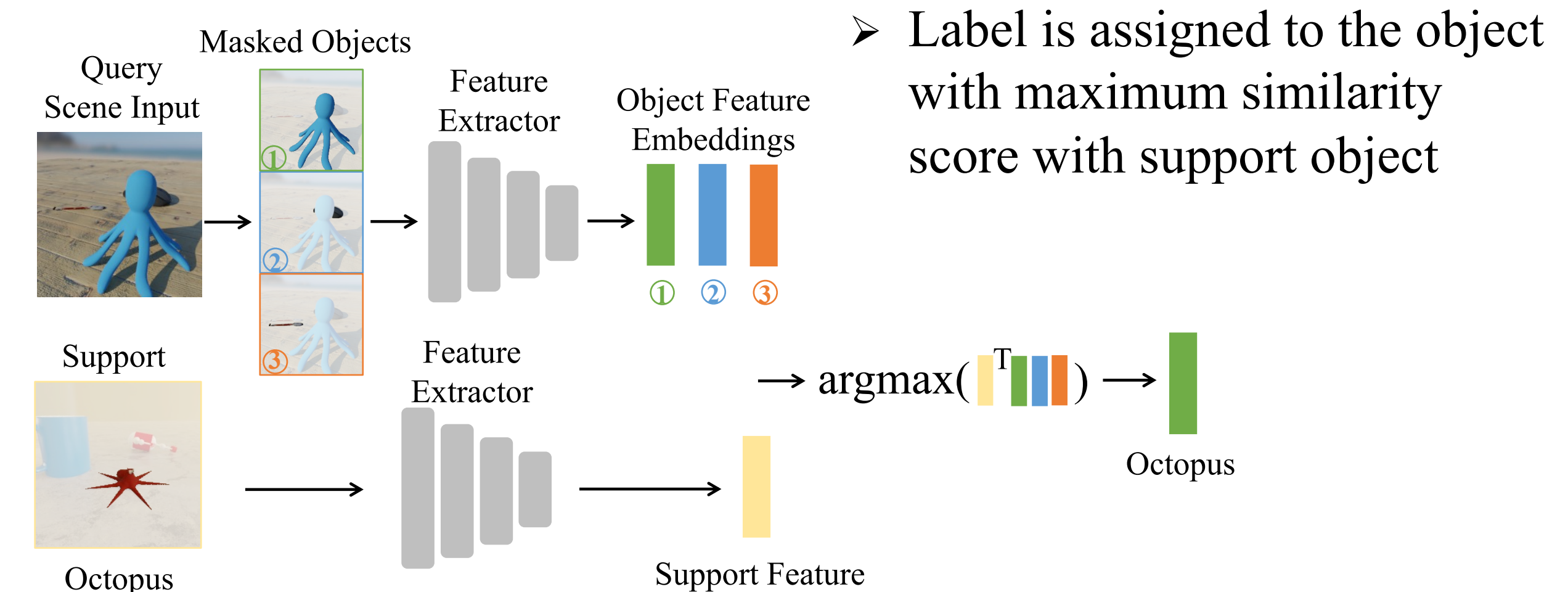
Settings Building Up to LSME

	Multi Object	Support Assignment	LSME
Base	No base classes since ME is not required		
Support			
Query			
Localization	✗	✗	✓
Mutual Exclusivity	✗	✓	✓

Multi-view Multi-object Representation Learning

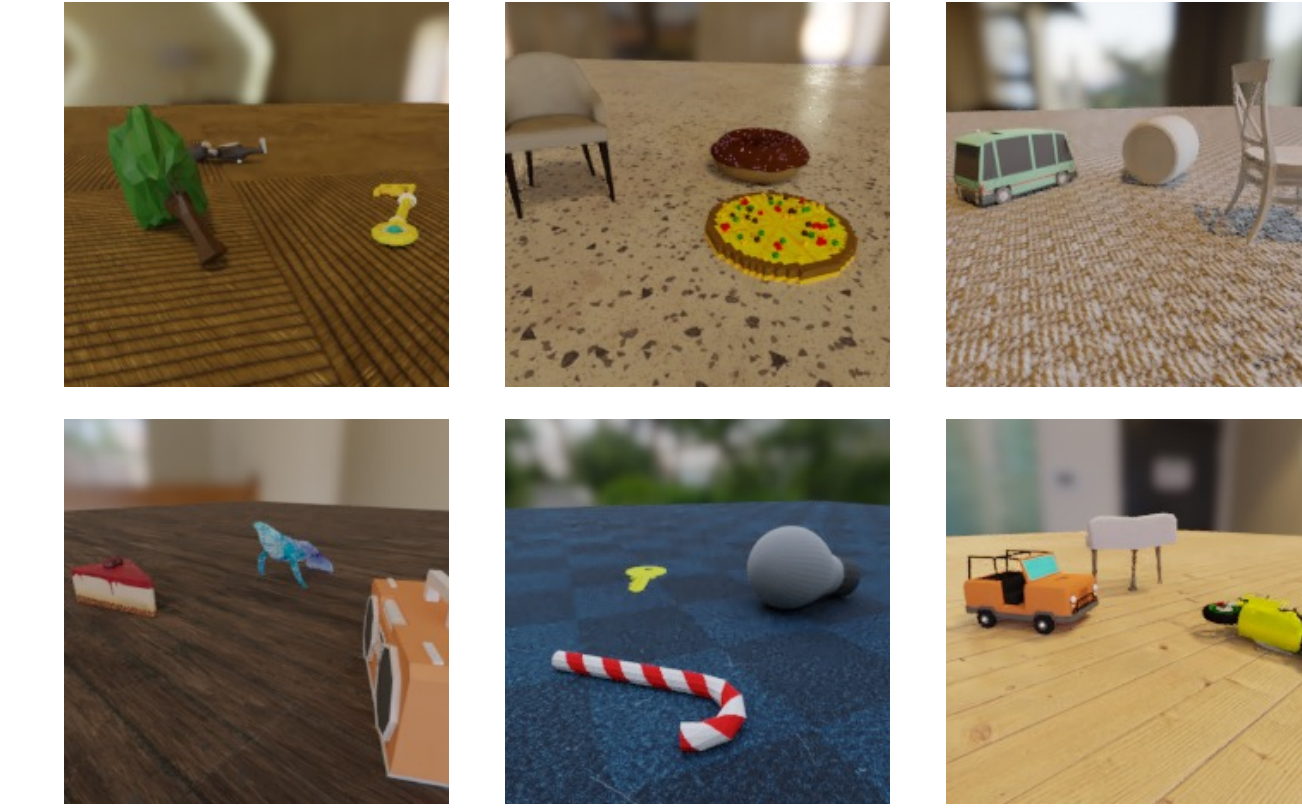


Low-shot Learning



Data for Studying LSME

There is currently no real-world dataset to directly test and study LSME

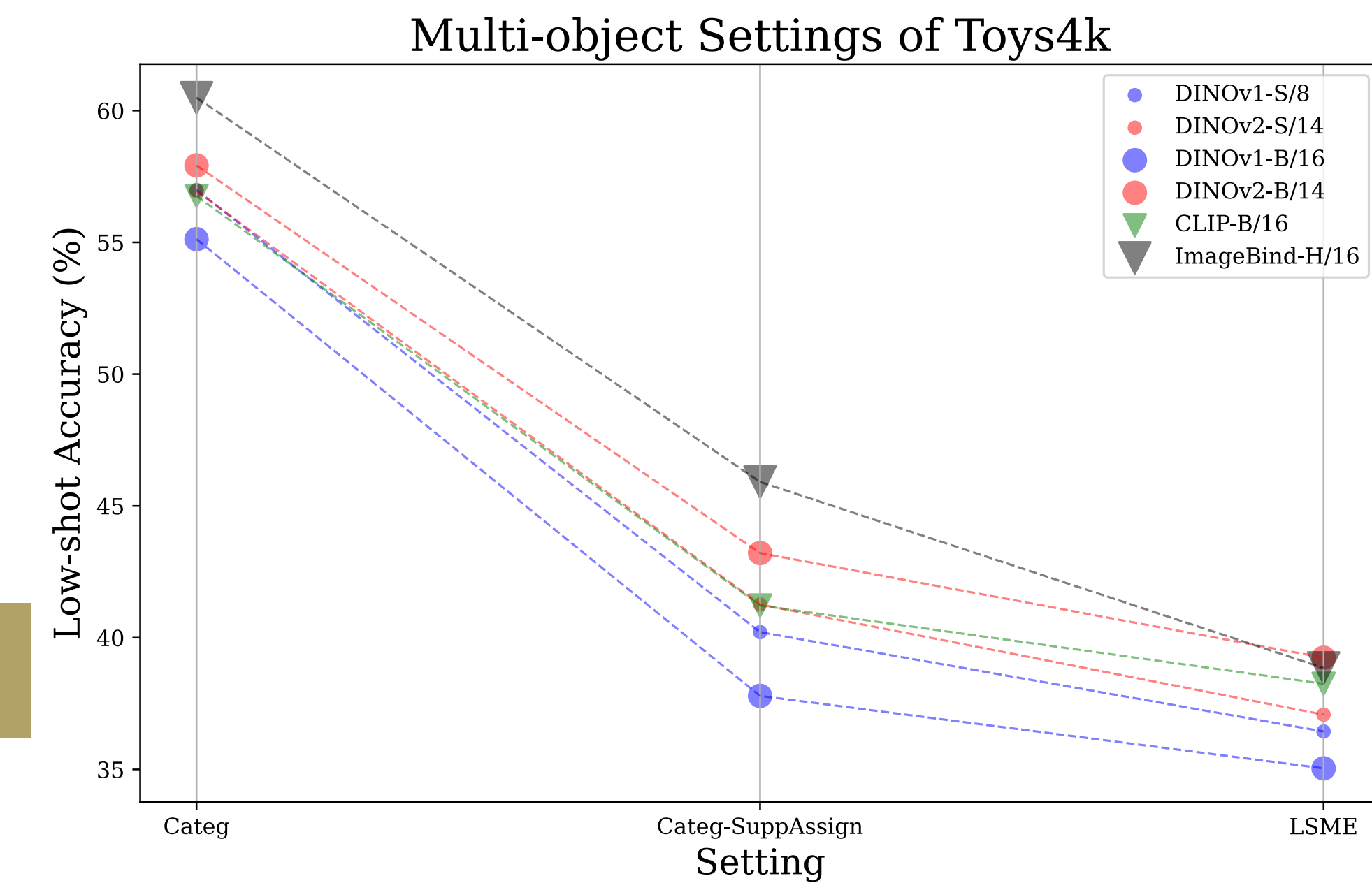


- Multiple objects in a scene
- Photorealistic lighting and environment
- Realistic object poses
- Diverse object/scene/pose combinations
- Works with any 3D asset

- 3 objects/scene
- Low-shot assumption: exactly 1 novel object in the scene



Results



	DINOv1 ViT S/8	DINOv2 ViT S/14	DINOv2 ViT B/14
Single-object Scene (W/o occlusion)	68.84	73.07	75.18
Multi-object Scene (With Occlusion)	56.99	56.95	57.92
Δ	-11.85	-16.12	-17.26

- Performance of all baselines decreases when ME and localization are required
- Occlusion impacts baselines performances

Low-shot Setup	DINOv2 B/14	DINOv2 B/14-ABC
1-shot 5-way	39.24	47.70
5-shot 5-way	55.03	63.52
1-shot 10-way	28.32	35.66
5-shot 10-way	43.26	51.72

- Our approach improves low-shot accuracy on various low-shot setups of Toys4k

- Our approach improves performance on ShapeNetCore.v2

	mIoU		DINOv2 S/14-ABC		DINOv2 B/14-ABC	
	Supp.	Query	LSA	SA	LSA	SA
FreeSOLO	0.52	0.54	33.99	44.84	35.50	48.92
CutLER	0.61	0.63	36.34	46.08	39.42	52.04
SAM	0.72	0.73	38.58	52.04	42.38	56.92

- Low-shot accuracy and support assignment accuracy improves with better object localization performance

Contributions

- **LSME** computational framing for learning with mutual exclusivity bias
- **Data pipeline generation** for studying LSME
- **Performance benchmarking** for multiple foundation model baselines
- **Novel self-supervised learning model** that achieves SOTA