Table A: Navigation results on PInNED on the environments of HM3D dataset, comparing categorical and fine-grained textual modalities.

			Navigation Metrics					**Detection Metrics**			
	Backbone	Modality	SR↑	SPL↑	CE↓	D2G↓	Steps	%Match↑	TM↑	CM↓	NM↓
Modular Agents											
CLIP	ViT-B/16	Categorical	3.52	2.75	10.23	7.98	148.1	95.47	5.12	15.73	79.15
CLIP	ViT-B/16	Fine-Grained	3.10	1.82	9.31	7.94	503.1	62.95	20.07	22.07	57.86
OWL	ViT-B/32	Categorical	7.79	3.73	19.96	7.96	780.6	38.81	26.50	58.49	15.01
OWL	ViT-B/32	Fine-Grained	7.29	3.36	12.66	7.90	871.7	22.97	62.60	32.88	4.52
End-to-end Agents											
RIM	ResNet-50	Categorical	4.61	3.78	14.25	9.23	336.0	-	-	-	-
RIM	ResNet-50	Fine-Grained	7.12	6.67	10.44	8.43	409.3	-	-	-	-

Table B: Navigation results on PInNED on the environments of HM3D dataset, considering the presence of distractors from the same category. **Bold** text denotes the best performance among each category of approaches.

			Navigation Metrics					**Detection Metrics**			
	Backbone	Modality	SR↑	SPL↑	CE↓	D2G↓	Steps	%Match↑	TM↑	CM↓	NM↓
Modular Agents											
CLIP	ViT-B/16	Textual	3.10	1.82	9.31	7.94	503.1	62.95	20.07	22.07	57.86
CLIP-Grad	ViT-B/32	Textual	4.53	2.42	6.95	7.94	465.8	77.95	4.65	7.21	84.14
OWL	ViT-B/32	Textual	7.29	3.36	12.66	7.90	871.7	22.97	**62.60**	32.88	4.52
SuperGlue	-	Visual	3.27	1.28	7.38	8.36	804.0	29.42	16.96	3.44	79.60
IEVE	-	Visual	3.52	3.07	12.25	7.73	712.1	30.03	32.39	16.01	51.60
PerSAM	ViT-B/16	Visual	2.77	1.81	6.53	8.20	362.5	**81.98**	1.15	10.43	88.42
PerSAM-F	ViT-B/16	Visual	1.93	1.28	**5.63**	8.12	321.3	36.13	0.60	13.48	85.92
DINO	ViT-B/16	Visual	4.02	1.71	6.88	8.28	826.0	23.89	62.73	**1.36**	35.91
CLIP	ViT-B/16	Visual	9.64	5.39	13.33	7.79	623.5	58.51	32.53	16.35	51.12
DINOv2	ViT-B/14	Visual	**14.84**	**7.94**	26.15	**7.28**	658.7	55.74	55.33	42.00	**2.67**
End-to-end Agents											
RIM	ResNet-50	Textual	7.12	6.67	**10.44**	8.43	409.3	-	-	-	-
RIM	ResNet-50	Visual	8.80	6.80	13.41	8.48	402.1	-	-	-	-
ZSON	ResNet-50	Visual	**9.14**	**7.18**	21.12	**7.00**	389.9	-	-	-	-