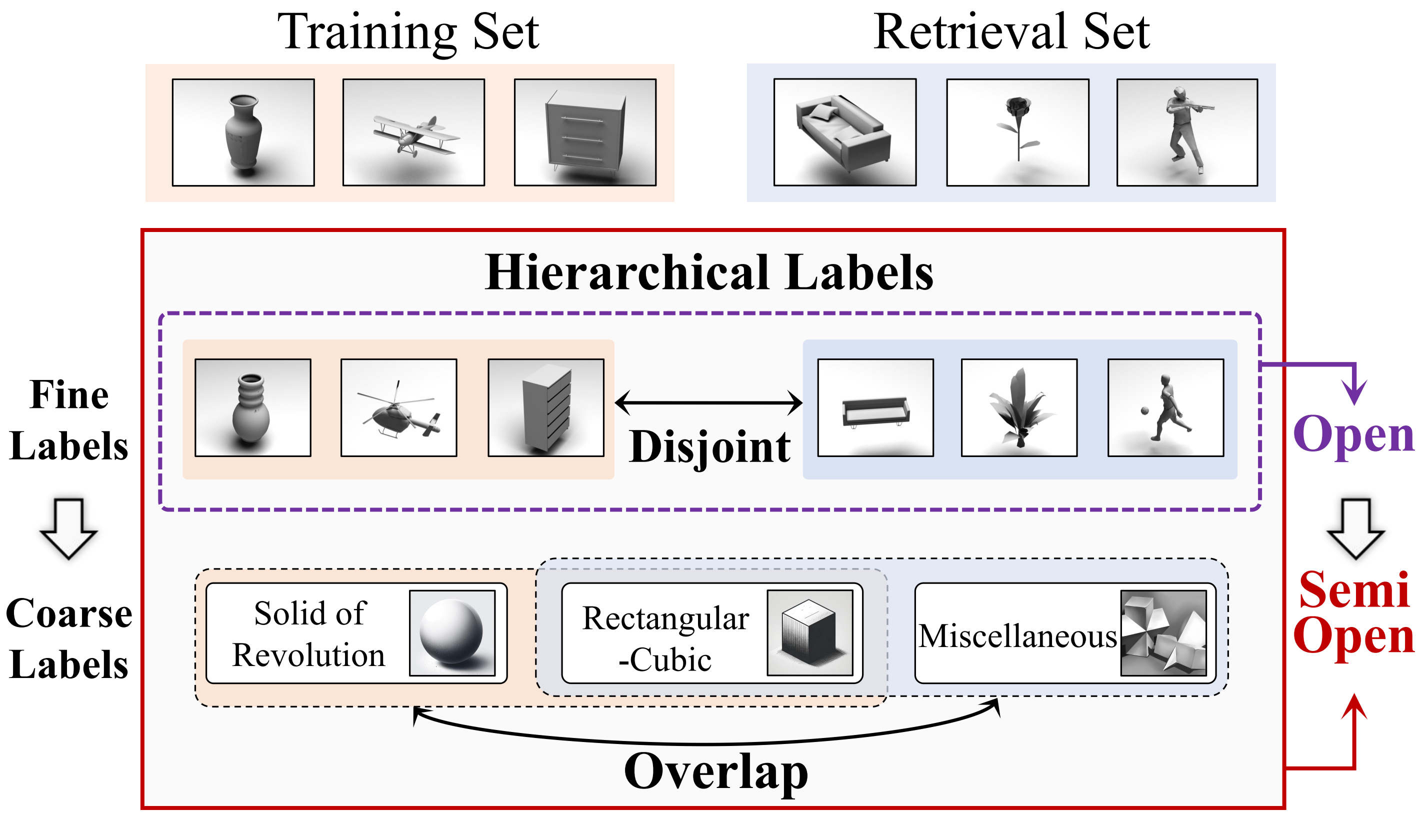
NEURAL INFORMATION PROCESSING SYSTEMS

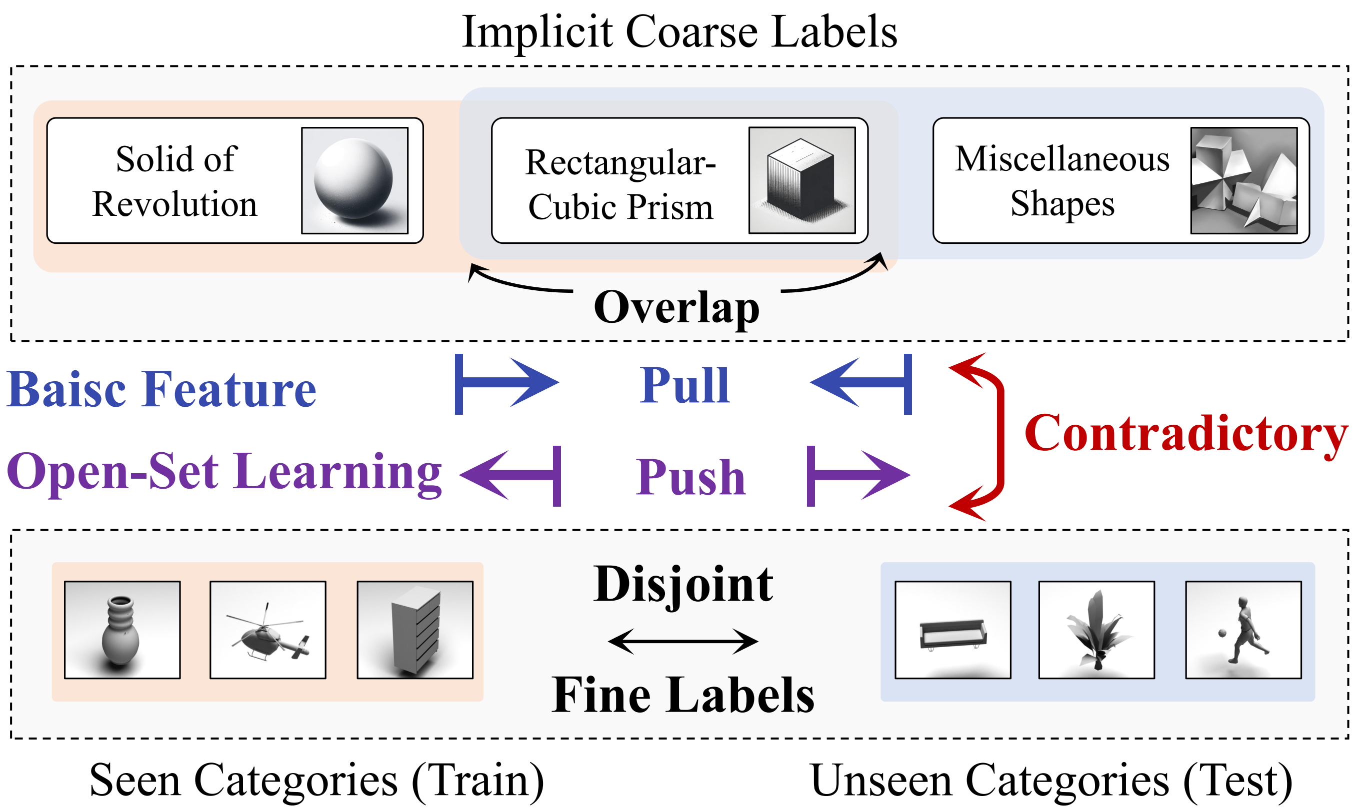
Semi-Open 3D Object Retrieval via Hierarchical Equilibrium on Hypergraph

Background



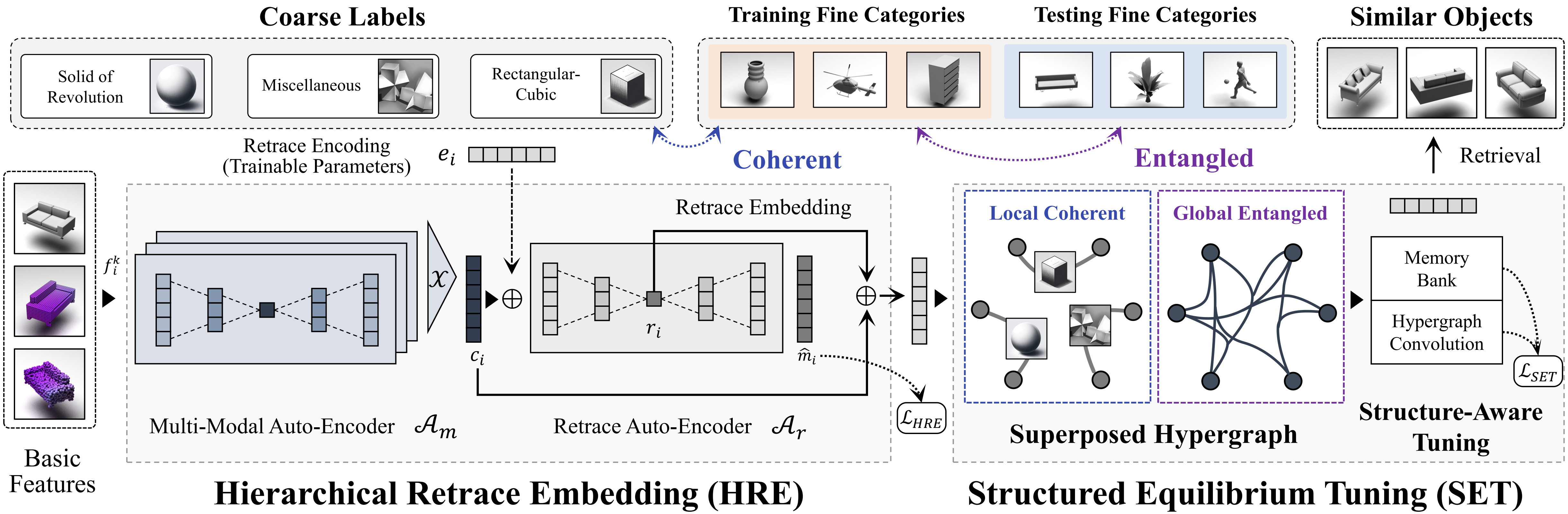
Training Set
Retrieval Set
Hierarchical Labels
Fine Labels
Disjoint
Open
Coarse Labels
Solid of Revolution
Rectangular -Cubic
Miscellaneous
Semi Open
Overlap

Motivation



Implicit Coarse Labels
Solid of Revolution
Rectangular-Cubic Prism
Miscellaneous Shapes
Overlap
Baisc Feature
Pull
Open-Set Learning
Push
Contradictory
Disjoint
Fine Labels
Seen Categories (Train)
Unseen Categories (Test)

Framework



Coarse Labels
Solid of Revolution
Miscellaneous
Rectangular-Cubic
Training Fine Categories
Testing Fine Categories
Similar Objects
Retrace Encoding
(Trainable Parameters)
e_i
Coherent
Entangled
Retrieval
f_i^k
Retrace Embedding
$\mathcal{X}$
c_i
r_i
$\hat{m}_i$
$\mathcal{L}_{HRE}$
Multi-Modal Auto-Encoder $\mathcal{A}_m$
Retrace Auto-Encoder $\mathcal{A}_r$
Basic Features
Local Coherent
Global Entangled
Superposed Hypergraph
Memory Bank
Hypergraph Convolution
$\mathcal{L}_{SET}$
Structure-Aware Tuning
Hierarchical Retrace Embedding (HRE)
Structured Equilibrium Tuning (SET)

Datasets

		SO-ESB	SO-NTU	SO-MN40	SO-ABO
Categories	Coarse	3	3	3	3
	Fine	41	67	40	21
	Seen	17	13	8	4
	Unseen	24	54	32	17
Number	Training	98	378	2821	1082
	Retrieval	457	1232	7591	4432
	Query	96	216	128	68
	Target	361	1016	7463	4364

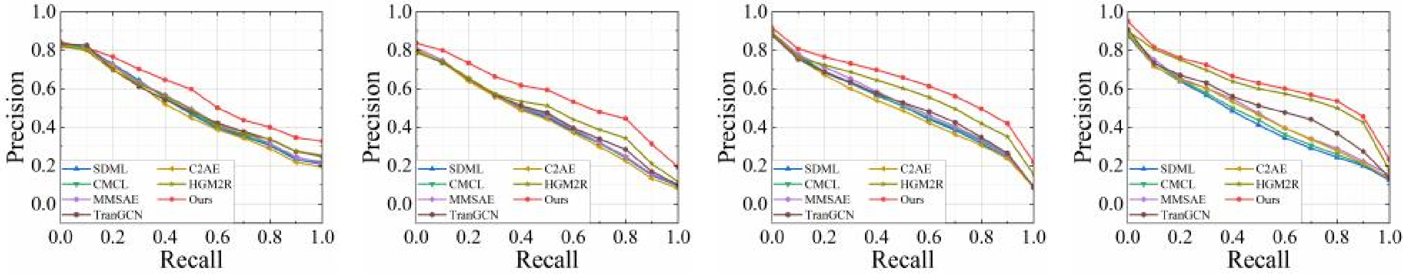
Metrics

- **Average Precision (mAP)**
- **Recall**
- **Normalized Discounted Cumulative Gain (NDCG)**
- **Average Normalized Modified Retrieval Rank (ANMRR)**
- **Precision-Recall Curve (PR-Curve)**

Rertrieval Performance

Method	SO-MN40				SO-ABO			
	mAP↑	Recall↑	NDCG↑	ANMRR↓	mAP↑	Recall↑	NDCG↑	ANMRR↓
SDML	0.5018	0.3241	0.6082	0.5106	0.4380	0.3425	0.4726	0.5564
CMCL	0.5086	0.3281	0.6128	0.5060	0.4520	0.3657	0.4816	0.5458
MMSAE	0.5189	0.3335	0.6226	0.4938	0.4783	0.3863	0.4929	0.5264
TranGCN	0.5188	0.3358	0.6131	0.4957	0.5175	0.3956	0.5127	0.4801
C2AE	0.4865	0.3152	0.5977	0.5231	0.4669	0.3674	0.4794	0.5313
HGM^2R	0.5779	0.3698	0.6482	0.4407	0.6069	0.4675	0.5463	0.4154
Ours	**0.6336**	**0.3993**	**0.6874**	**0.3972**	**0.6339**	**0.4793**	**0.5622**	**0.3836**

Method	SO-ESB				SO-NTU			
	mAP↑	Recall↑	NDCG↑	ANMRR↓	mAP↑	Recall↑	NDCG↑	ANMRR↓
SDML	0.4947	0.8027	0.1858	0.5430	0.4384	0.7009	0.1937	0.5764
CMCL	0.4990	0.8154	0.1880	0.5457	0.4440	0.7053	0.1946	0.5721
MMSAE	0.5036	0.8503	0.1931	0.5523	0.4454	0.7046	0.1935	0.5745
TranGCN	0.5063	0.9011	0.1968	0.5408	0.4548	0.7121	0.1961	0.5624
C2AE	0.4809	0.7863	0.1824	0.5501	0.4303	0.6987	0.1915	0.5828
HGM^2R	0.5049	0.8831	0.1939	0.5551	0.4821	0.7364	0.2026	0.5438
Ours	**0.5756**	**0.9346**	**0.2045**	**0.4874**	**0.5678**	**0.8116**	**0.2251**	**0.4677**

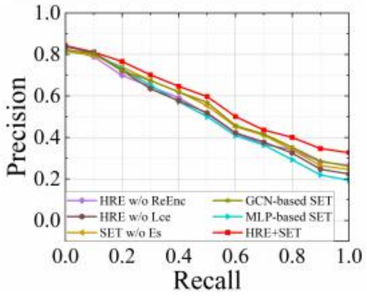


(a) PR-C on SO-ESB. (b) PR-C on SO-NTU. (c) PR-C on SO-MN40. (d) PR-C on SO-ABO.

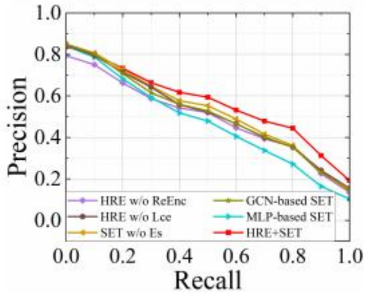
Ablation Studies

Method	SO-ESB				SO-NTU			
	mAP↑	Recall↑	NDCG↑	ANMRR↓	mAP↑	Recall↑	NDCG↑	ANMRR↓
HRE w/o ReEnz	0.5159	0.9086	0.1953	0.5431	0.4913	0.7534	0.2053	0.5355
HRE w/o $\mathcal{L}_{ce}$	0.5133	0.8738	0.1934	0.5365	0.5161	0.7902	0.2162	0.5162
SET w/o $\mathcal{E}_c$	0.5358	0.8957	0.1975	0.5184	0.5285	0.7898	0.2184	0.4986
GCN-based SET	0.5405	0.8999	0.2003	0.5192	0.5144	0.7703	0.2140	0.5138
MLP-based SET	0.5014	0.8483	0.1930	0.5476	0.4689	0.7304	0.2023	0.5561
HRE+SET	**0.5756**	**0.9346**	**0.2045**	**0.4874**	**0.5678**	**0.8116**	**0.2251**	**0.4677**

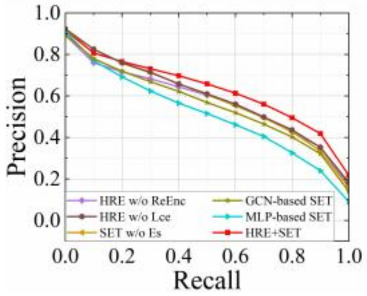
Method	SO-MN40				SO-ABO			
	mAP↑	Recall↑	NDCG↑	ANMRR↓	mAP↑	Recall↑	NDCG↑	ANMRR↓
HRE w/o ReEnz	0.5791	0.3710	0.6479	0.4410	0.6055	0.4523	0.5535	0.4062
HRE w/o $\mathcal{L}_{ce}$	0.5967	0.3783	0.6756	0.4309	0.5885	0.4269	0.5413	0.4230
SET w/o $\mathcal{E}_c$	0.5913	0.3757	0.6669	0.4347	0.6006	0.4263	0.5494	0.4132
GCN-based SET	0.5602	0.3573	0.6410	0.4628	0.5686	0.4253	0.5314	0.4415
MLP-based SET	0.5088	0.3290	0.6149	0.5073	0.4880	0.3787	0.5023	0.5159
HRE+SET	**0.6336**	**0.3993**	**0.6874**	**0.3972**	**0.6339**	**0.4793**	**0.5622**	**0.3836**



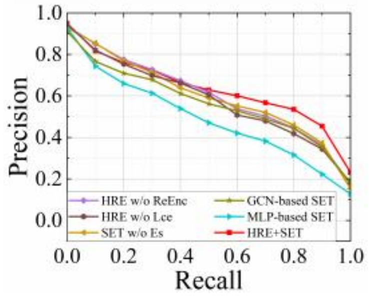
(a) PR-C on SO-ESB.



(b) PR-C on SO-NTU.



(c) PR-C on SO-MN40.



(d) PR-C on SO-ABO.

Conclusion

- In this paper, we introduce a more practical *Semi-Open Environment* setting for open-set 3D object retrieval with hierarchical labels, in which the training and testing set share a partial label space for coarse categories but are completely disjoint from fine categories. We propose the Hypergraph-Based Hierarchical Equilibrium Representation (HERT) framework for semi-open 3D object retrieval. Specifically, to overcome the global disequilibrium of unseen categories, we propose the Hierarchical Retrace Embedding (HRE) module to fully leverage the multi-level category information. Besides, we perform the Structured Equilibrium Tuning (SET) module to tackle the feature overlap and class confusion problem. This module utilizes more equilibrial correlations among objects and generalizes to unseen categories, by constructing a superposed hypergraph based on the local coherent and global entangled correlations. Furthermore, we construct four 3D object datasets with multi-level category labels for semi-open 3DOR tasks, *i.e.*, SO-ESB, SO-NTU, SO-MN40, and SO-ABO. Results demonstrate that the proposed method can effectively generate and generalize the hierarchical embeddings of 3D objects towards the semi-open environment. However, due to dataset limitations, we are currently unable to verify the balanced representation effect on more than three levels of labels, which is one of our future research directions. We believe this paper can provide new insights for future research in more practical scenarios of open-set learning.