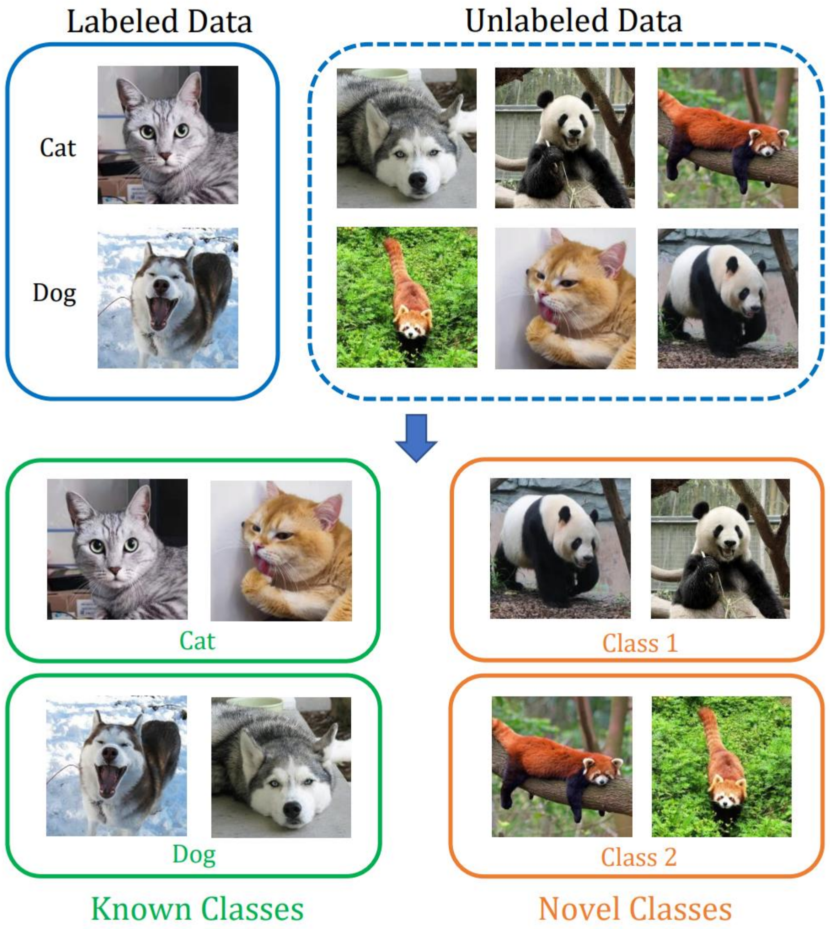
Discover and Align Taxonomic Context Priors for Open-world Semi-Supervised Learning

Yu Wang[1] , Zhun Zhong[2] , Pengchong Qiao[1] , Xuxin Cheng[1] , Xiawu Zheng[3,4] , Chang Liu[5] , Nicu Sebe[6] , Rongrong Ji[3,4] , Jie Chen[1,3*]

[1] Peking University, [2] University of Nottingham, [3]Peng Cheng Laboratory
[4] Xiamen University, [5]Tsinghua University [6] University of Trento

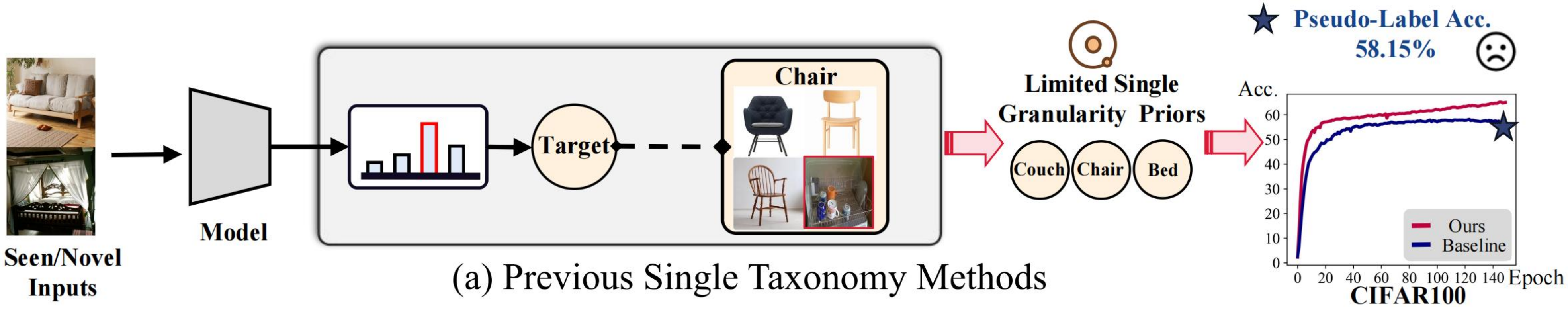
Problem: Open-world Semi-Supervised Learning



Open-world Semi-Supervised Learning (OSSL) aims to classify unlabeled samples from both seen and novel classes using partially labeled samples from the seen classes.

Due to the lack of labeled samples for novel classes, **it is vital to exploit some priors as auxiliary supervision** to classifier for all classes.

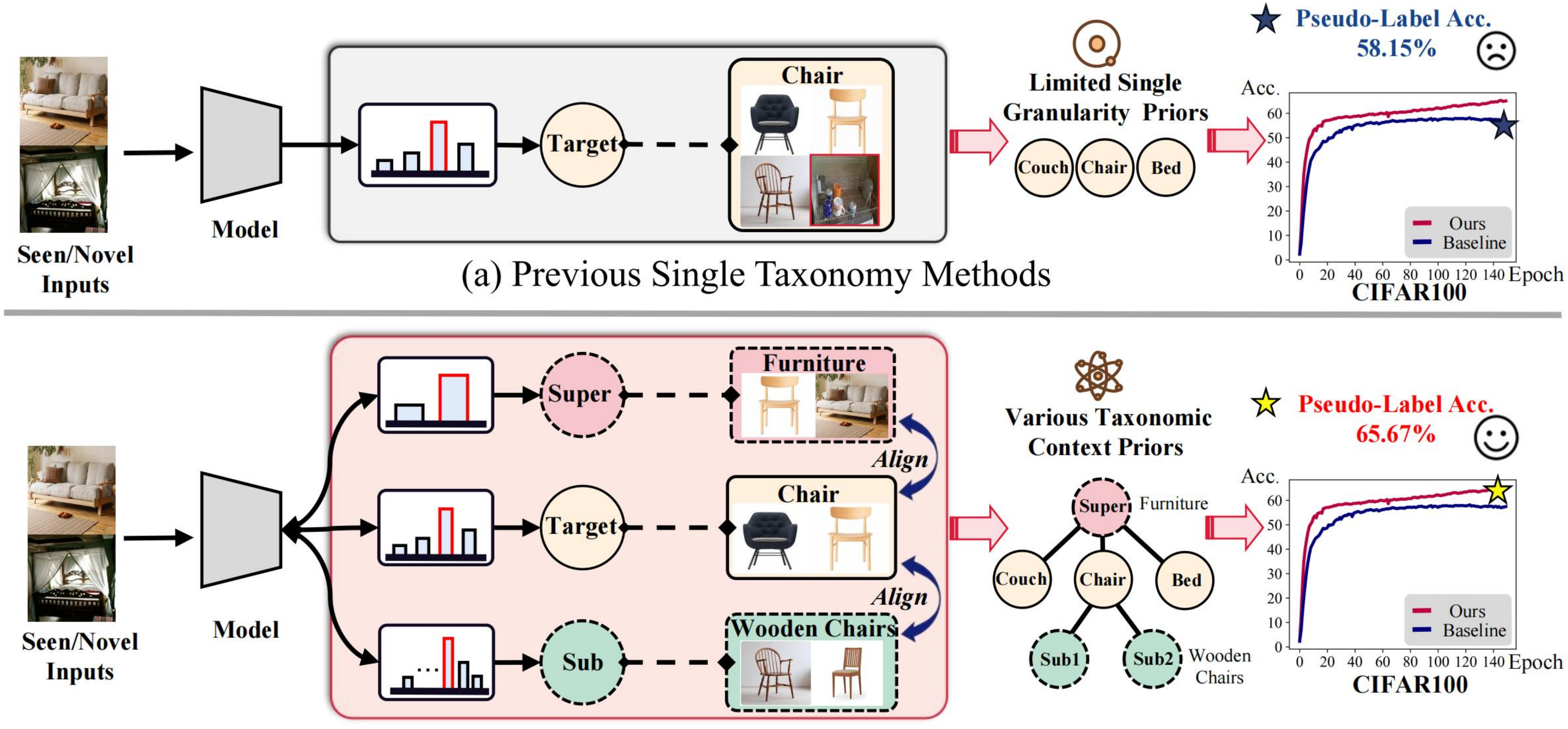
Previous works: only explore priors at a single granularity



(a) Previous Single Taxonomy Methods

Existing works either exploit pairwise similarity prior or class distribution prior to achieve this. However, these methods only explore priors at a single granularity, which suffer from sub-optimization and inaccurate pseudo labels, due to the limited supervision.

Motivation: Explore various taxonomic context as priors

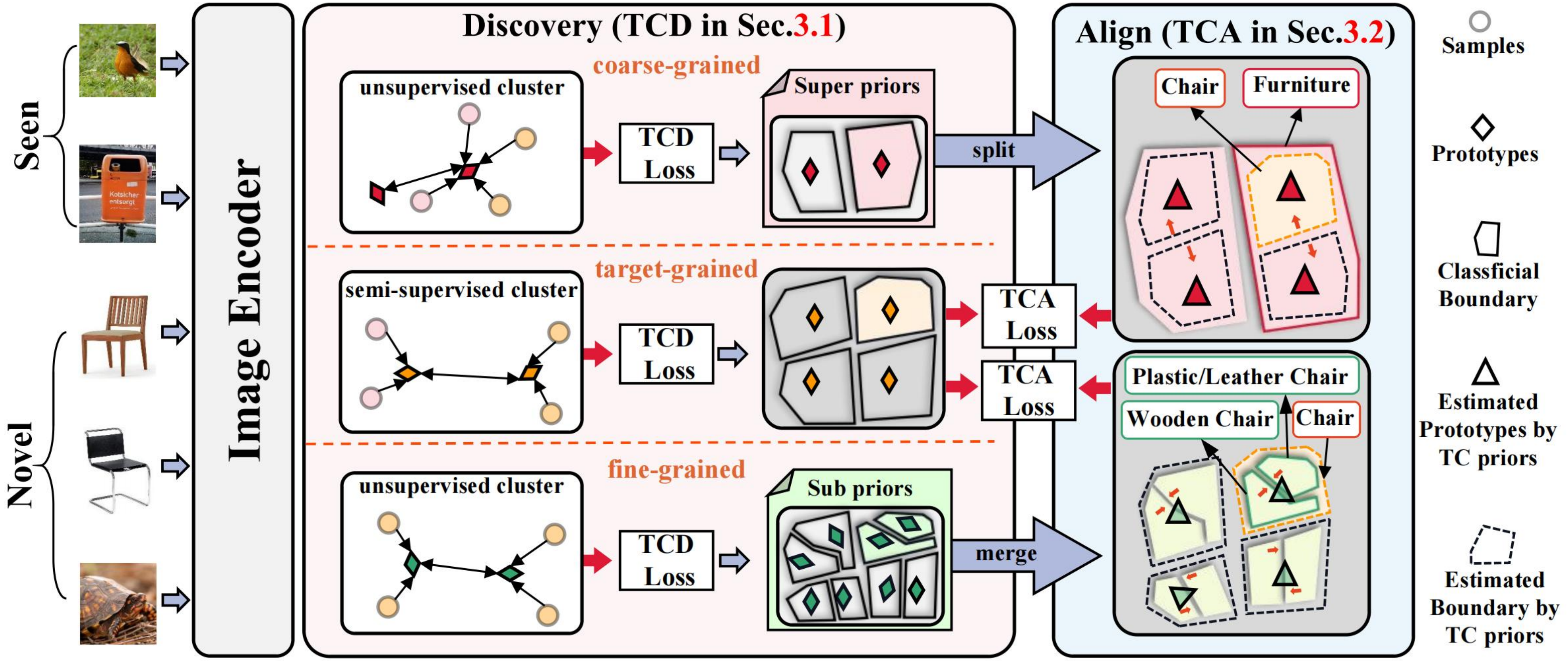


(a) Previous Single Taxonomy Methods

(b) Contextual Taxonomy Method (Ours)

In this paper, we argue that leveraging multiple levels of granularity as semantic priors (e.g., sub-classes, classes, and super-classes, etc.) is a more preferable solution for OSSL.

Method



To achieve it, we develop two modules: i) Taxonomic Context Discovery (TCD) module to discovers the underlying taxonomic context priors; ii) Taxonomic Context-based prediction Alignment (TCA) module to enforces consistency across hierarchical predictions.

Taxonomic Context Discovery (TCD)

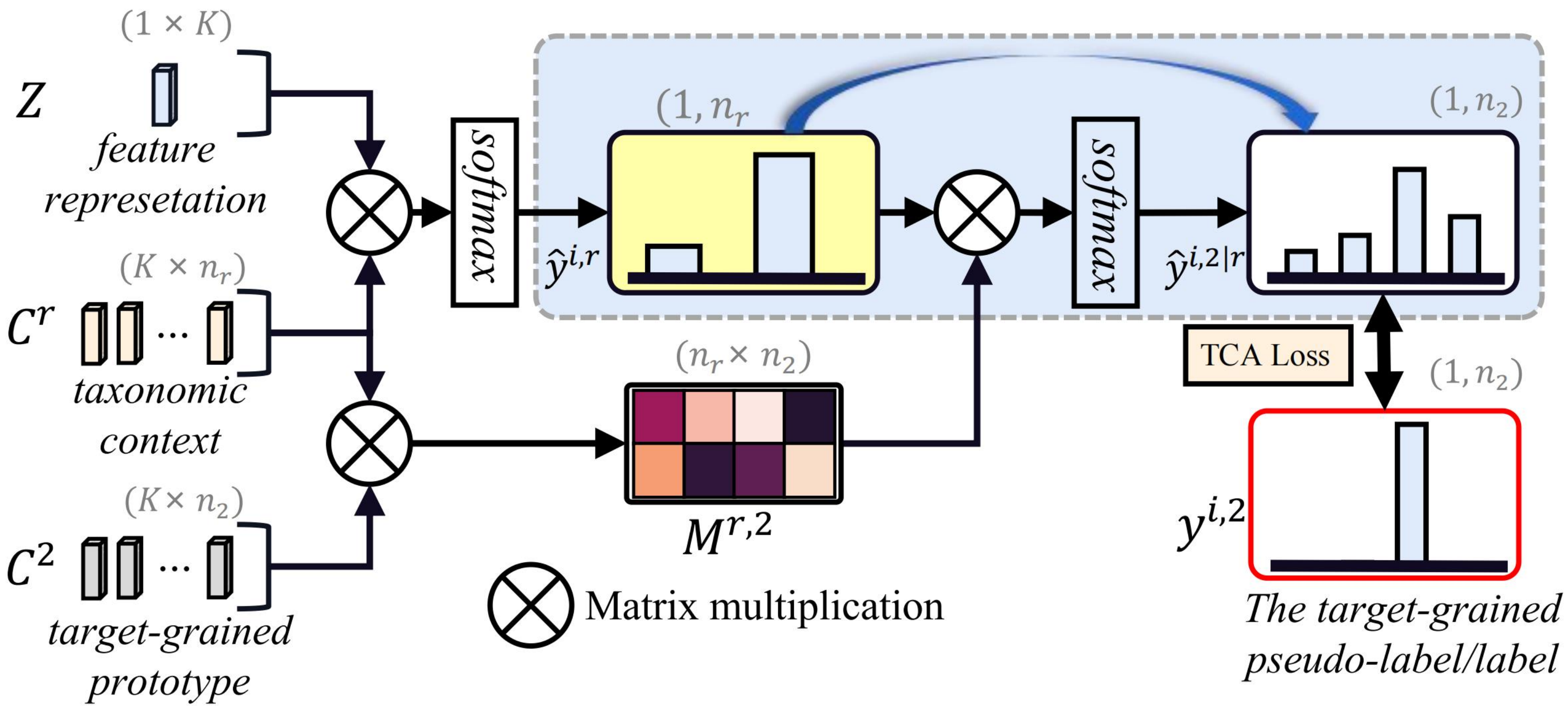
TIDA first obtains normalized feature representation $z_i \in \mathbb{R}^K = f_\theta(x_i)$ for the i-th sample x_i by image encoder f_θ. Then, we use z_i and prototypes on each hierarchy to perform clustering by optimizating $\mathcal{L}_{tcd}$.

$$\mathcal{L}_{tcd} = -\sum_{l=1}^{L}\sum_{j=1}^{n_l} y_i^{j,l} \log \hat{y}_i^{j,l} = -\sum_{l=1}^{L}\sum_{j=1}^{n_l} y_i^{j,l} \log \frac{\exp(\mathcal{S}(z_i, c_j^l)/\tau)}{\sum_{m=1}^{n_l} \exp(\mathcal{S}(z_i, c_m^l)/\tau)},$$

where $y_i^l \in \{0,1\}^{|\mathcal{Y}^l + \mathcal{Y}^u|}$ is the pseudo-label of x_i on l-th hierarchy generated by Sinkhorn-Knopp algorithm, $\hat{y}_i^l$ denotes the prediction of model for x_i and $\tau = 0.1$ is the temperature.

Taxonomic Context-based prediction Alignment (TCA)

To achieve taxonomic context consistency, we propose the Taxonomic Context-based prediction Alignment (TCA), which aims to establish reliable affinity relationships across hierarchies.



Experiment

Performance on four generic datasets.

Methods	CIFAR10			CIFAR100			ImageNet-100			Tiny ImageNet		
	Seen	Novel	All	Seen	Novel	All	Seen	Novel	All	Seen	Novel	All
DTC [29]	42.7	31.8	32.4	22.1	10.5	13.7	24.5	17.8	19.3	13.5	12.7	11.5
RankStats [28]	71.4	63.9	66.7	20.4	16.7	17.8	41.2	26.8	37.4	9.6	8.9	6.4
UNO [23]	86.5	71.2	78.9	53.7	33.6	42.7	66.0	42.2	53.3	28.4	14.4	20.4
ORCA [5]	82.8	85.5	84.1	52.5	31.8	38.6	83.9	60.5	69.7	–	–	–
OpenNCD [50]	83.5	86.7	85.3	53.6	33.0	41.2	84.0	65.8	73.2	–	–	–
TRSSL [65]	94.9	89.6	92.2	68.5	52.1	60.3	82.6	67.8	75.4	39.5	20.5	30.3
OpenLDN [64]	92.4	93.2	92.8	55.0	40.0	47.7	–	–	–	–	–	–
TIDA (Ours)	**94.2**	**93.4**	**93.8**	**73.3**	**56.6**	**65.3**	**83.4**	**71.2**	**77.6**	**45.7**	**28.4**	**37.2**

Performance on three fine-grained datasets.

Methods	Oxford-IIIT Pet			FGVC Aircraft			Stanford-Cars		
	Seen	Novel	All	Seen	Novel	All	Seen	Novel	All
DTC [29]	20.7	16.0	13.5	16.3	16.5	11.8	12.3	10.0	7.7
RankStats [28]	12.6	11.9	11.1	13.4	13.6	11.1	10.4	9.1	6.6
UNO [23]	49.8	22.7	34.9	44.4	24.7	31.8	49.0	15.7	30.7
TRSSL [65]	70.9	36.1	53.9	69.5	41.2	55.4	83.5	37.1	60.4
OpenLDN [64]	66.8	33.1	50.4	–	–	45.7	–	–	38.7
TIDA (Ours)	**75.7**	**39.2**	**59.9**	**71.1**	**43.7**	**57.4**	**85.9**	**43.5**	**66.0**

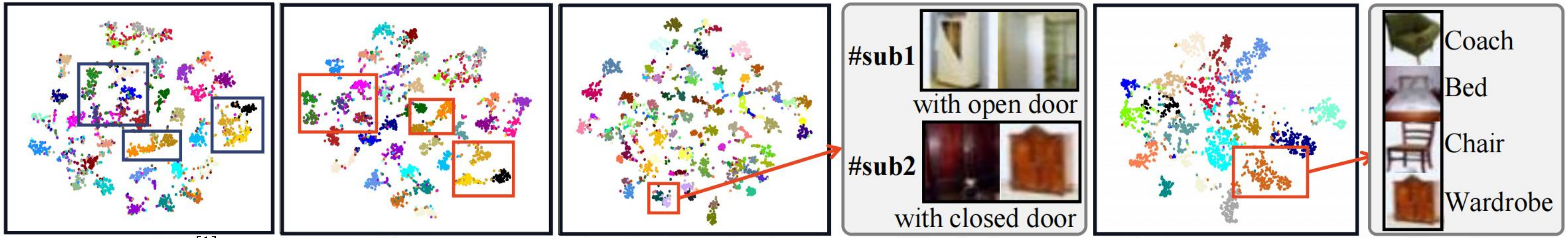
Ablation study

C-TCP: Coarse-grained Taxonomic Context Priors; **F-TCP**: Fine-grained Taxonomic Context Priors; **TCA**: Taxonomic Context-based prediction Alignment (When using **TCA** only, the model is equipped with three target-grained classifiers that are aligned by **TCA**).

#	C-TCP	F-TCP	TCA	CIFAR100			Tiny ImageNet			Stanford-Cars		
				Seen	Novel	All	Seen	Novel	All	Seen	Novel	All
a)				67.0	48.9	57.9	39.7	21.8	31.1	84.8	38.3	61.3
b)	✔			63.7	47.0	55.5	36.4	19.2	28.0	81.2	41.2	61.1
c)	✔		✔	71.3	51.6	61.6	43.7	27.1	35.8	84.9	40.8	62.9
d)		✔		65.3	45.6	55.3	36.3	20.4	28.8	78.1	31.2	54.8
e)		✔	✔	71.5	54.6	63.1	44.6	27.3	36.8	85.3	42.3	64.1
f)			✔	71.8	49.1	60.5	43.1	21.1	32.5	82.8	39.5	61.9
g)	✔	✔		69.9	45.4	57.7	36.3	19.8	29.0	76.4	30.2	53.8
h)	✔	✔	✔	**73.3**	**56.6**	**65.3**	**45.7**	**28.4**	**37.2**	**85.9**	**43.5**	**66.0**

As we can see, the proposed TCA is indispensable to our TIDA. This indicates that the key to performance improvements is having **consistent taxonomic context** for OSSL problems.

Visualization

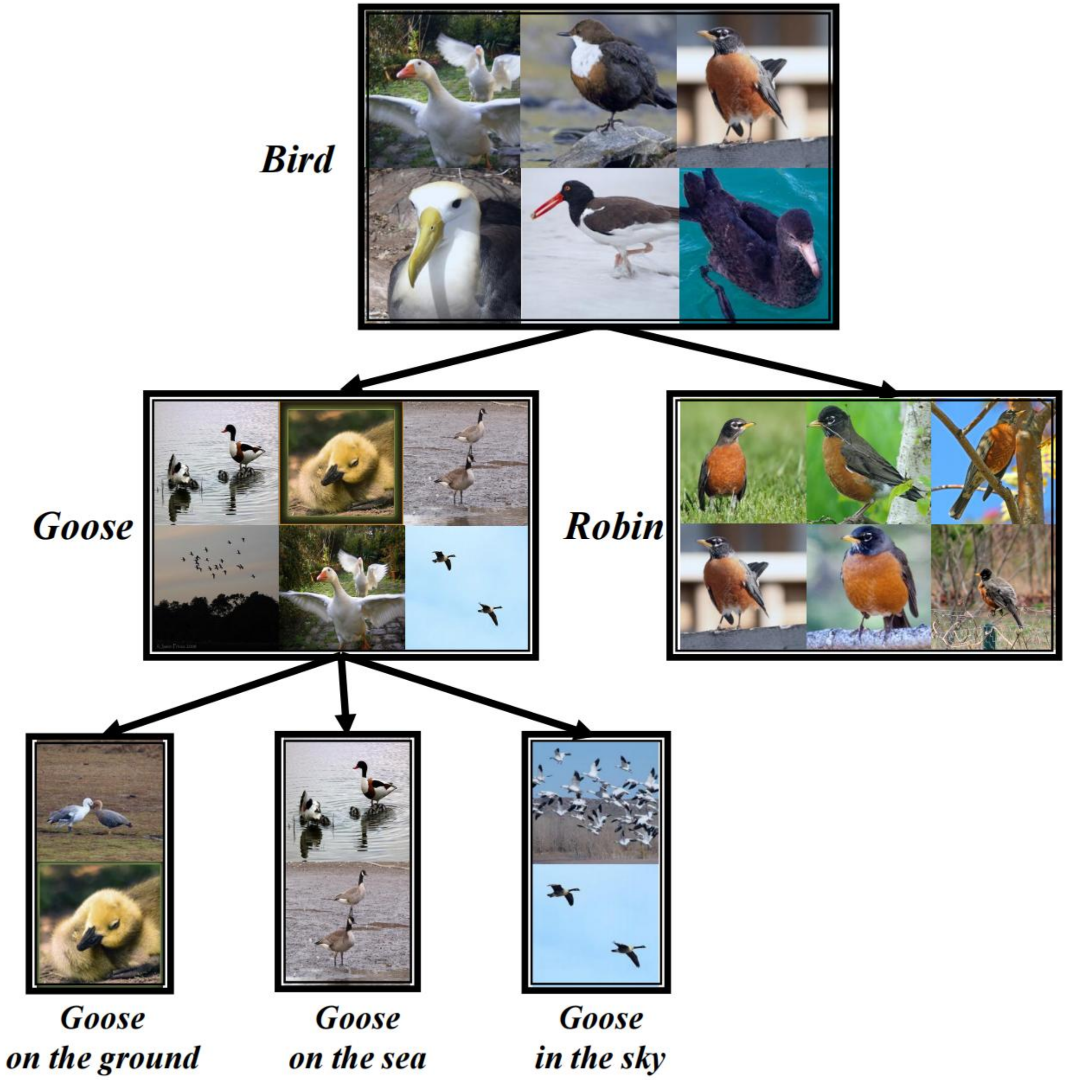


(a) TRSSL[1] (b) Ours (c) Fined-Grained TC Wardrobe (d) Coarse-Grained TC Furniture

We use t-SNE to visualize the features learned by TIDA and the baseline[1]. As shown in above, TIDA produces more discriminative features than TRSSL[1], where the samples are generally better clustered.

[1] *Mamshad Nayeem Rizve, Navid Kardan, and Mubarak Shah. Towards realistic semi-supervised learning. In ECCV, 2022.*

Visualization



The visualization of hierarchical semantic structure learned by TIDA on ImageNet100.

Conclusion

- In this paper, we identify the importance of multi-granularity priors for Open-world Semi-Supervised Learning (OSSL) and introduce a new type of prior knowledge, i.e., taxonomic context priors.

- Moreover, we introduce a uniformed OSSL framework, named by TIDA, which can discover taxonomic context priors without any extra supervision.

Conclusion

- Our study uncovers a significant observation that incorporating taxonomic context as priors can enhance the performance of our model in challenging real-world scenarios with limited supervision and unknown semantic concepts.

- This provides a new idea for utilizing unlabeled data, not only limited to open-set semi-supervised learning.

Thank you!