Upcycling Models under Domain and Category Shift

Sanqing Qu[1], Tianpei Zou[1], Florian Röhrbein[2], Cewu Lu[3], Guang Chen[1(✉)], Dacheng Tao[4], Changjun Jiang[1]

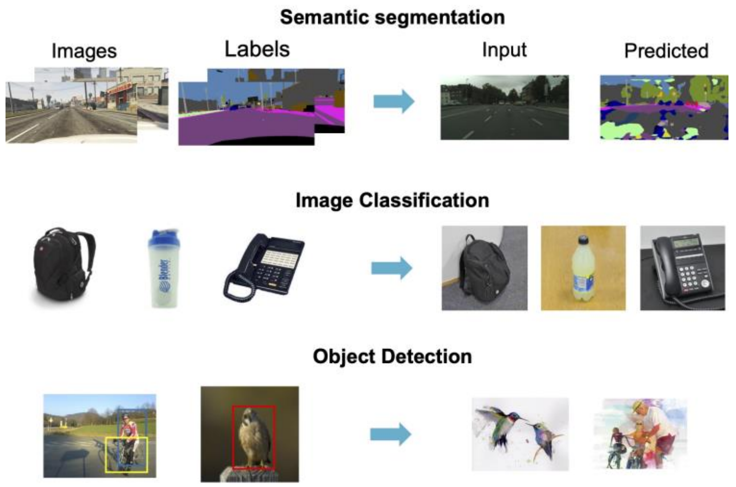
1. Tongji University
2. Chemnitz University of Technology
3. Shanghai Jiao Tong University
4. The University of Sydney

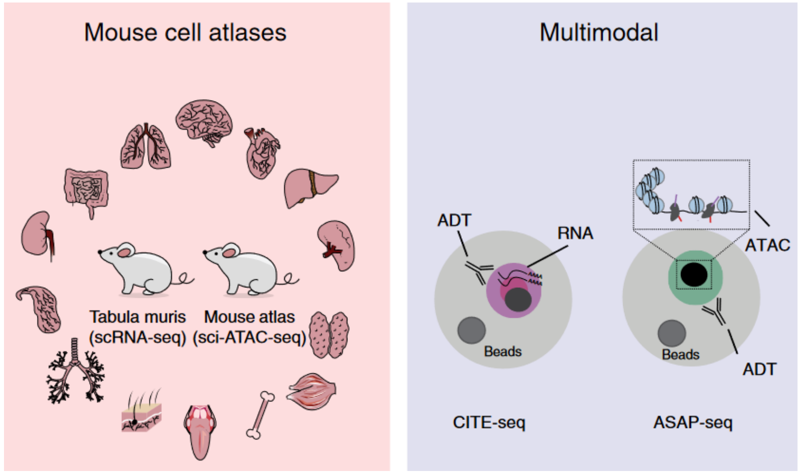
JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Introduction

❑ Quick Preview

- ◆ Unsupervised Domain Adaptation (UDA), especially Source-free Domain Adaptation (SFDA) has become a promising technique to address the Out-of-distribution (OOD) issue of Deep Neural Networks (DNNs).
- ◆ Most of existing methods require that source and target domain share the same label space.
- ◆ To the best of our knowledge, we are the first to exploit and achieve the Source-free Universal Domain Adaptation (SF-UniDA) with only a standard pre-trained closed-set model.

*


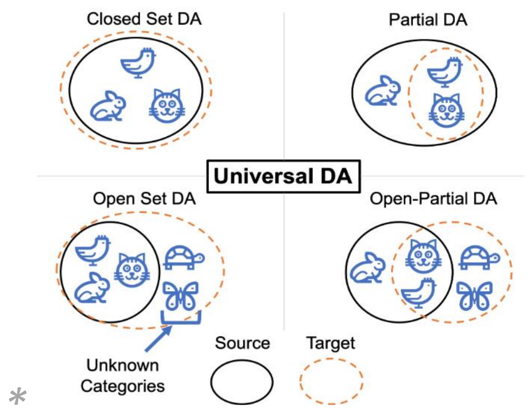
+


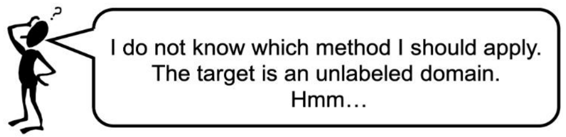
Credit to DANCE [NeurIPS-20]; + scJoint [Nature Biotechnology-22]

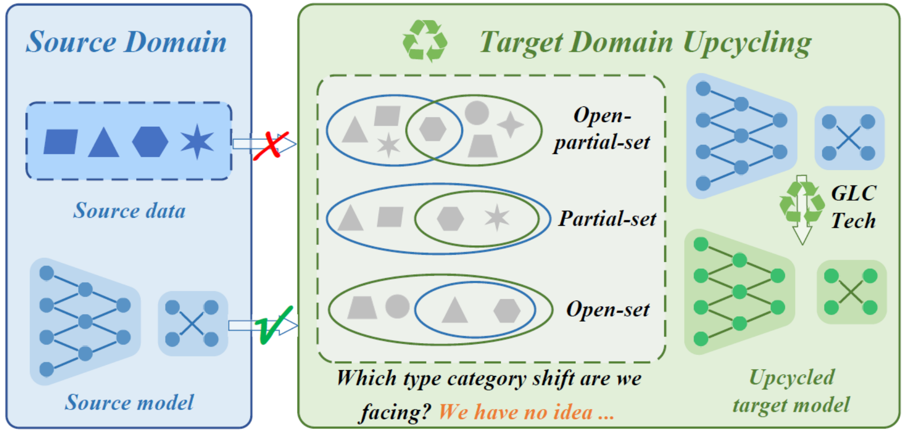
Introduction

❑ Quick Preview

- Unsupervised Domain Adaptation (UDA), especially Source-free Domain Adaptation (SFDA) has become a promising technique to address the Out-of-distribution (OOD) issue of Deep Neural Networks (DNNs).
- Most of existing methods require that source and target domain share the same label space.
- To the best of our knowledge, we are the first to exploit and achieve the Source-free Universal Domain Adaptation (SF-UniDA) with only a standard pre-trained closed-set model.





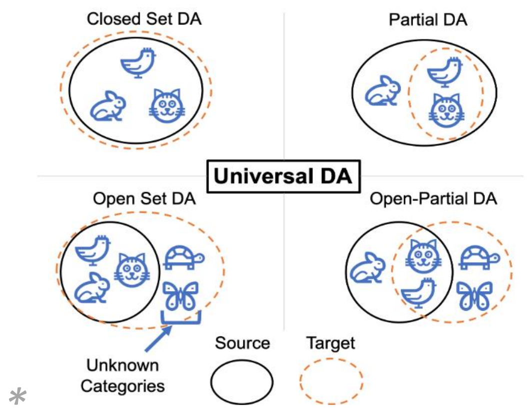


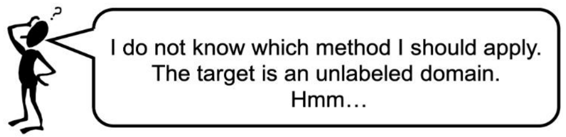
*Credit to DANCE [NeurIPS-20]

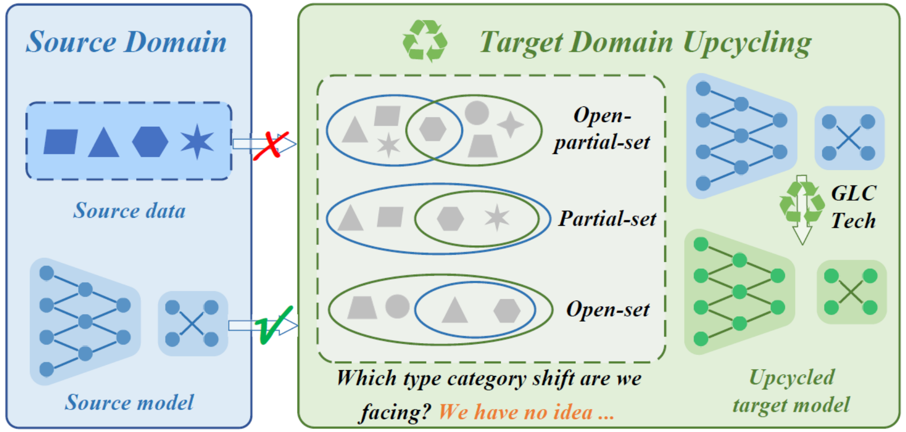
Methodology

❑ Source-free Universal Domain Adaptation (SF-UniDA)

- **Source Model** f_s: $\mathcal{X}_s \rightarrow \mathcal{Y}_s$ pre-trained on Source Domain $\mathcal{D}_s$.
- **Target Domain** $\mathcal{D}_t$: n_t unlabeled data samples $\{x_t^i,\ ?\}_{i=1}^{n_t}$
- **Goal**: Under both domain shift and category shift, i.e., $\mathcal{X}_s \neq \mathcal{X}_t, \mathcal{Y}_s \neq \mathcal{Y}_t$, using $\{x_t^i,\}_{i=1}^{n_t}$ with f_s to learn a target model f_t: $\mathcal{X}_t \rightarrow \mathcal{Y}_t$. f_t is able to identify "known" categories specified in f_s, and reject "unknown" categories not involved in f_s.





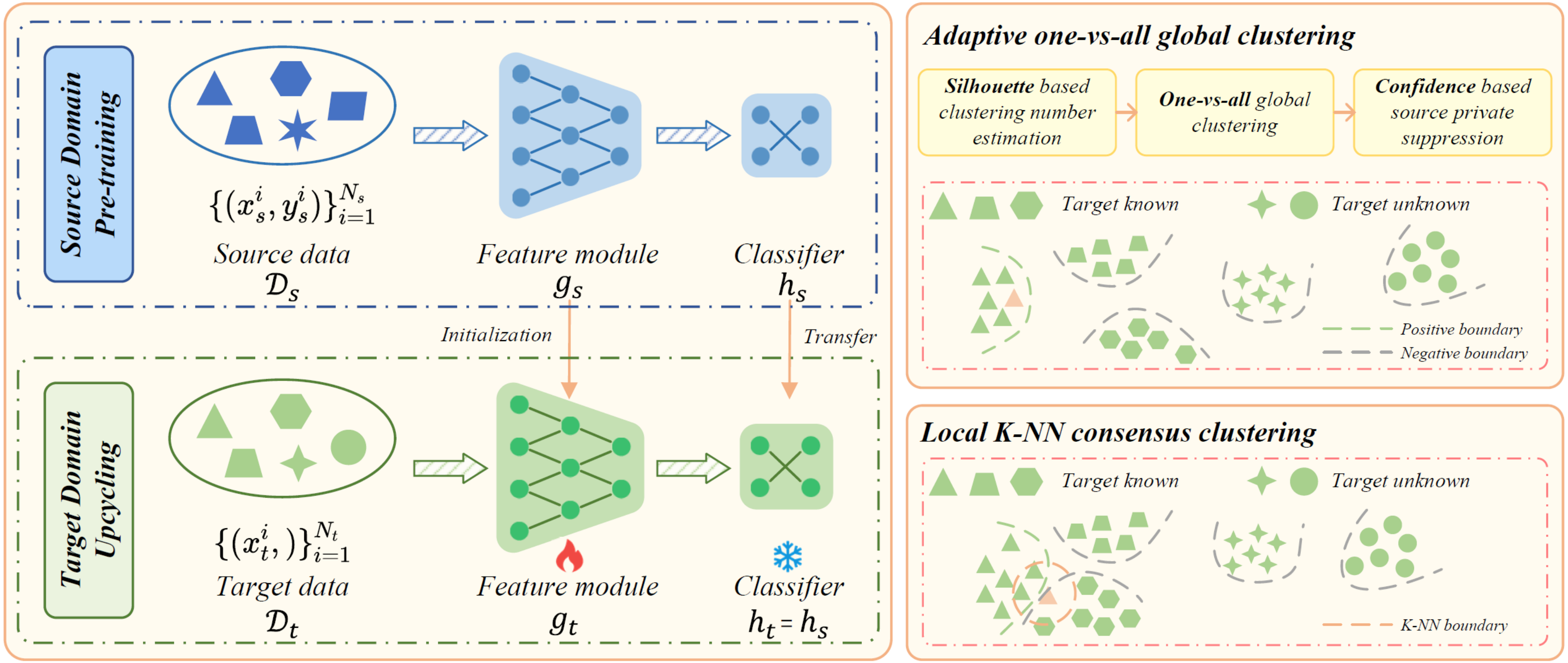


*Credit to DANCE [NeurIPS-20]

Methodology

❑ Global and Local Clustering (GLC)

- **Global Clustering:** Global adaptive one-vs-all clustering for pseudo labeling.
- **Local Clustering:** Local KNN consensus clustering to alleviate negative transfer.



Methodology

❑ Global one-vs-all clustering

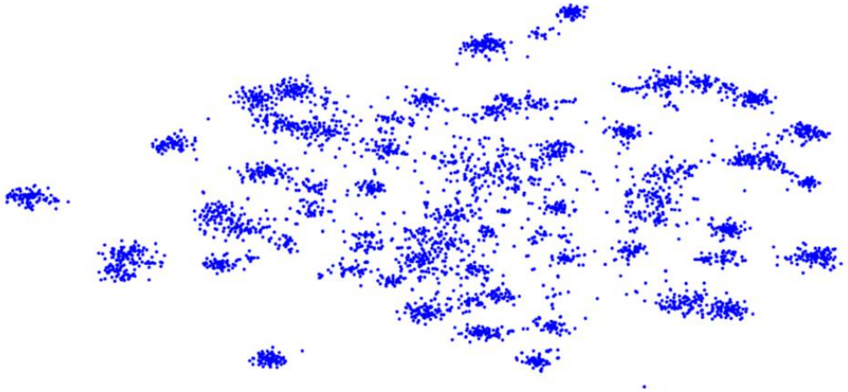
- **Existing solutions:** Assigning pseudo labels based on sample-level prediction with argmax operation, or introducing feature prototype with weighted k-means clustering. **Only applicable to closed-set setting**.
- ***How to achieve pseudo labeling with inconsistent label space?***
  - To accommodate the unknown categories, given a specified “known” category, we need to figure out what is and what is not the category.
  - Introducing **one-vs-all** clustering, we convert the K-way classification task to K times binary classification. For a particular category, ***we aggregate the Top-K sampled instances as the positive prototype for this category, and apply K-means for the rest instances to construct negative prototypes.***
  - To achieve adaptive clustering, we first estimate the number of categories involved in target domain with the assistance of the Silhouette metric. Besides, we design a strategy to filter out source-private categories.

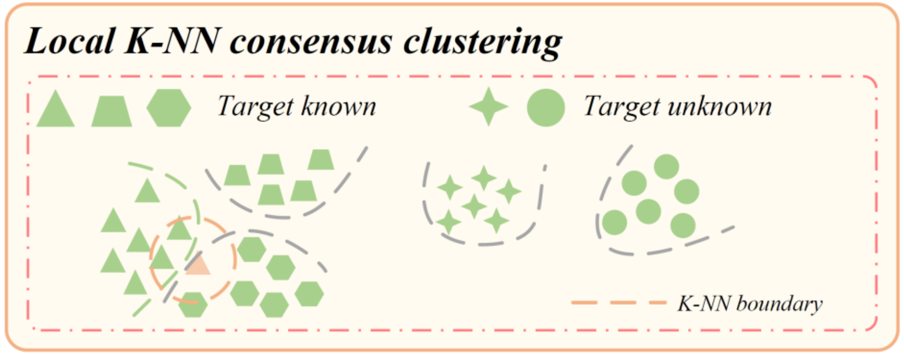
Methodology

❑ Local KNN consensus clustering

- **Limitations of pseudo labeling:** Due to the presence of both domain and category shift, there are still a lot of mis-assigned pseudo labels.
- ***How to alleviate the negative transfer caused by those incorrectly assigned pseudo labels?***
  - ➢ Data samples that are close to each other in the embedding manifold space should contain similar semantics.
  - ❑ Existing works have applied this mechanism to the closed-set setting, e.g., SFDA. We find that it is also beneficial for SF-UniDA in cases involving “unknown” categories.

*





* Credit to NRC [NeurIPS-19]

Methodology

❑ Optimization and Inference

◆ Optimization objective:

$$l_c^i = \frac{1}{|L^i|} \sum_{x_t \in L^i} \delta_c(f_t(x_t)),$$

$$\mathcal{L}_{tar}^{glb} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C_s} \hat{q}_c^i \log \delta_c(f_t(x_t^i)),$$

$$\mathcal{L}_{tar}^{loc} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C_s} l_c^i \log \delta_c(f_t(x_t^i)).$$

$$\mathcal{L}_{tar} = \eta \mathcal{L}_{tar}^{glb} + \mathcal{L}_{tar}^{loc}.$$

◆ Inference:

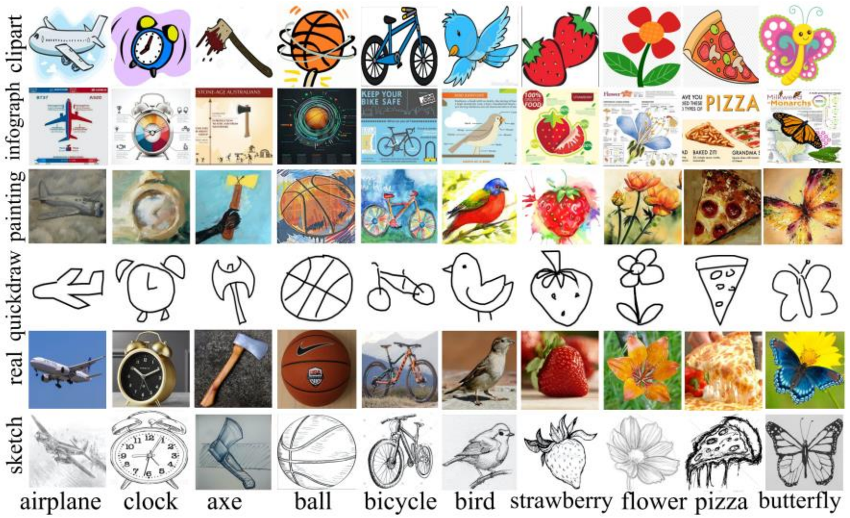
$$I(x_t) = -\frac{1}{\log C_s} \sum_{c=1}^{C_s} \delta_c(f_t(x_t)) \log \delta_c(f_t(x_t))$$

$$y(x_t) = \begin{cases} \text{unknown}, & \text{if } I(x_t) \geq \omega \\ \text{argmax}(f_t(x_t)), & \text{if } I(x_t) < \omega \end{cases}$$

Experiments

❑ Setup

- **Datasets:** Office-31 , Office-Home, VisDA, and DomainNet.
- **Setting:** Partial-set Domain Adaptation (PDA), Open-set Domain Adaptation (OSDA), and Open-partial Domain Adaptation (OPDA).
- **Metrics:** Hscore $= \frac{2\times ACC_{ukn}\times ACC_{knw}}{ACC_{ukn}+Acc_{knw}}$



*

Table 3. Details of class split. Here, $\mathcal{Y}$, $\bar{\mathcal{Y}}_s$, and $\bar{\mathcal{Y}}_t$ denotes the source-target-shared class, the source-private class, and the target-private class, respectively.

Dataset	Class Split($\mathcal{Y}/\bar{\mathcal{Y}}_s/\bar{\mathcal{Y}}_t$)		
	OPDA	OSDA	PDA
Office-31 [42]	10/10/11	10/0/11	10/21/0
Office-Home [53]	10/5/50	25/0/40	25/40/0
VisDA [39]	6/3/3	6/0/6	6/6/0
DomainNet [38]	150/50/145	-	-

* Credit to DomainNet [ICCV-19]

Experiments

❑ Results

◆ OPDA Setting

- ➢ On the VisDA benchmark, our GLC outperforms the SOTA by 14.8% H-score.
- ➢ On the DomainNet benchmark, our GLC outperforms the SOTA by 3.0% H-score.

Table 2. H-score (%) comparison in OPDA scenario on Office-31, VisDA, and DomainNet. Some results are cited from UMAD [23].

Methods	SF	OPDA	OSDA	PDA	Office-31							VisDA	DomainNet						
					A2D	A2W	D2A	D2W	W2A	W2D	Avg	S2R	P2R	P2S	R2P	R2S	S2P	S2R	Avg
UAN [47]	✗	✓	✗	✗	59.7	58.6	60.1	70.6	60.3	71.4	63.5	34.8	41.9	39.1	43.6	38.7	38.9	43.7	41.0
CMU [11]	✗	✓	✗	✗	68.1	67.3	71.4	79.3	72.2	80.4	73.1	32.9	50.8	45.1	52.2	45.6	44.8	51.0	48.3
DCC [20]	✗	✓	✓	✓	**88.5**	78.5	70.2	79.3	75.9	88.6	80.2	43.0	56.9	43.7	50.3	43.3	44.9	56.2	49.2
OVANet [39]	✗	✓	✓	✗	85.8	79.4	80.1	**95.4**	84.0	**94.3**	86.5	53.1	56.0	47.1	51.7	44.9	47.4	**57.2**	50.7
GATE [7]	✗	✓	✓	✓	**87.7**	**81.6**	84.2	**94.8**	83.4	94.1	**87.6**	56.4	57.4	**48.7**	**52.8**	**47.6**	**49.5**	56.3	**52.1**
Source-only	✓	-	-	-	70.9	63.2	39.6	77.3	52.2	86.4	64.9	25.7	57.3	38.2	47.8	38.4	32.2	48.2	43.7
SHOT-O [22]	✓	✗	✓	✗	73.5	67.2	59.3	88.3	77.1	84.4	75.0	44.0	35.0	30.8	37.2	28.3	31.9	32.2	32.6
UMAD [23]	✓	✓	✓	✗	79.1	77.4	**87.4**	90.7	**90.4**	**97.2**	87.0	**58.3**	**59.0**	44.3	50.1	42.1	32.0	55.3	47.1
GLC	✓	✓	✓	✓	81.5	**84.5**	**89.8**	90.4	**88.4**	92.3	**87.8**	**73.1**	**63.3**	**50.5**	**54.9**	**50.9**	**49.6**	**61.3**	**55.1**

Experiments

❑ Results

◆ OSDA Setting

- On the VisDA benchmark, GLC outperforms UMAD by 5.3% H-score.
- On the Office-Home benchmark, GLC outperforms UMAD by 3.4% H-score.

Table 4. H-score (%) comparison in OSDA scenario on Office-Home, Office-31, and VisDA. (Best in red and second best in blue)

Methods	SF	OPDA	OSDA	PDA	Office-Home													Office31	VisDA
					Ar2Cl	Ar2Pr	Ar2Re	Cl2Ar	Cl2Pr	Cl2Re	Pr2Ar	Pr2Cl	Pr2Re	Re2Ar	Re2Cl	Re2Pr	Avg	Avg	Avg
OSBP [41]	✗	✗	✓	✗	55.1	65.2	72.9	64.3	64.7	70.6	63.2	53.2	73.9	66.7	54.5	72.3	64.7	83.7	52.3
ROS [2]	✗	✗	✓	✗	60.1	69.3	76.5	58.9	65.2	68.6	60.6	56.3	74.4	68.8	60.4	75.7	66.2	85.9	66.5
CMU [11]	✗	✓	✗	✗	55.0	57.0	59.0	59.3	58.2	60.6	59.2	51.3	61.2	61.9	53.5	55.3	57.6	65.2	54.2
DANCE [38]	✗	✓	✓	✓	6.5	9.0	9.9	20.4	10.4	9.2	28.4	12.8	12.6	14.2	7.9	13.2	12.9	79.8	67.5
DCC [20]	✗	✓	✓	✓	56.1	67.5	66.7	49.6	66.5	64.0	55.8	53.0	70.5	61.6	57.2	71.9	61.7	72.7	59.6
OVANet [39]	✗	✓	✓	✗	58.6	66.3	69.9	62.0	65.2	68.6	59.8	53.4	69.3	68.7	59.6	66.7	64.0	91.7	66.1
GATE [7]	✗	✓	✓	✓	63.8	70.5	75.8	66.4	67.9	71.7	67.3	61.5	76.0	70.4	61.8	75.1	69.0	89.5	70.8
Source-only	✓	-	-	-	46.1	63.3	72.9	42.8	54.0	58.7	47.8	36.1	66.2	60.8	45.3	68.2	55.2	69.6	29.1
SHOT-O [22]	✓	✗	✓	✗	37.7	41.8	48.4	56.4	39.8	40.9	60.0	41.5	49.7	61.8	41.4	43.6	46.9	77.5	28.1
UMAD [23]	✓	✓	✓	✗	59.2	71.8	76.6	63.5	69.0	71.9	62.5	54.6	72.8	66.5	57.9	70.7	66.4	89.8	66.8
GLC	✓	✓	✓	✓	65.3	74.2	79.0	60.4	71.6	74.7	63.7	63.2	75.8	67.1	64.3	77.8	69.8	89.0	72.5

Results

PDA Setting

- GLC achieves the competitive results compared to methods specially designed for PDA.
- GLC achieves the SOTA performance compared to existing universal methods.

Table 5. Accuracy (%) comparison in PDA scenario on Office-Home, Office-31, and VisDA. (Best in red and second best in blue)

Methods	SF	OPDA	OSDA	PDA	OfficeHome Ar2Cl	Ar2Pr	Ar2Re	Cl2Ar	Cl2Pr	Cl2Re	Pr2Ar	Pr2Cl	Pr2Re	Re2Ar	Re2Cl	Re2Pr	Avg	Office31 Avg	VisDA Avg
ETN [6]	✗	✗	✗	✓	59.2	77.0	79.5	62.9	65.7	75.0	68.3	55.4	84.4	75.7	57.7	84.5	70.4	96.7	59.8
BA3US [25]	✗	✗	✗	✓	60.6	83.2	88.4	71.8	72.8	83.4	75.5	61.6	86.5	79.3	62.8	86.1	76.0	97.8	54.9
DANCE [38]	✗	✓	✓	✓	53.6	73.2	84.9	70.8	67.3	82.6	70.0	50.9	84.8	77.0	55.9	81.8	71.1	86.0	73.7
DCC [20]	✗	✓	✓	✓	54.2	47.5	57.5	83.8	71.6	86.2	63.7	65.0	75.2	85.5	78.2	82.6	70.9	93.3	72.4
OVANet [39]	✗	✓	✓	✗	34.1	54.6	72.1	42.4	47.3	55.9	38.2	26.2	61.7	56.7	35.8	68.9	49.5	74.6	34.3
GATE [7]	✗	✓	✓	✓	55.8	75.9	85.3	73.6	70.2	83.0	72.1	59.5	84.7	79.6	63.9	83.8	74.0	93.7	75.6
Source-only	✓	-	-	-	45.9	69.2	81.1	55.7	61.2	64.8	60.7	41.1	75.8	70.5	49.9	78.4	62.9	87.8	42.8
SHOT-P [22]	✓	✗	✗	✓	64.7	85.1	90.1	75.1	73.9	84.2	76.4	64.1	90.3	80.7	63.3	85.5	77.8	92.2	74.2
UMAD [23]	✓	✓	✓	✗	51.2	66.5	79.2	63.1	62.9	68.2	63.3	56.4	75.9	74.5	55.9	78.3	66.3	89.5	68.5
GLC	✓	✓	✓	✓	55.9	79.0	87.5	72.5	71.8	82.7	74.9	41.7	82.4	77.3	60.4	84.3	72.5	94.1	76.2

Experiments

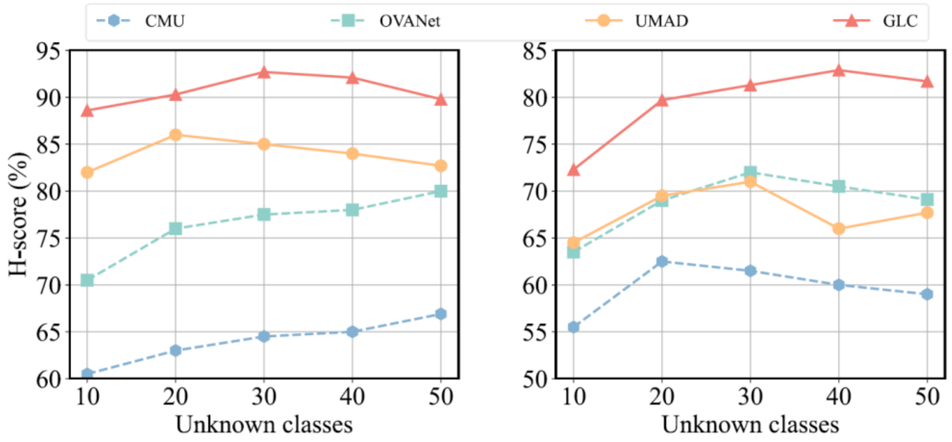
Performance Analysis

Ablation and Robustness Analysis

- Ablation study demonstrates the effectiveness of our global and local clustering.
- Additional experiments varying the number of unknown categories in target domain support the robustness of our algorithm.

Table 6. **Ablation Study**. Results for OPDA on Office-31, Office-Home, and VisDA with different variants of GLC.

Method	Office-31	Office-Home	VisDA
Source model	64.9	60.9	25.7
GLC (w/o $\mathcal{L}_{tar}^{loc}$)	86.1	74.8	66.0
GLC (w/o $\mathcal{L}_{tar}^{glb}$)	87.4	67.2	57.3
GLC (full)	**87.8**	**75.6**	**73.1**



(a) Ar →Re in OPDA (b) Cl → Pr in OPDA

Figure 4. **H-score (%) of OPDA** when varying the number of unknown classes in Office-Home. GLC shows stable and much superior performance against existing methods.

❑ Performance Analysis

◆ Hyper-parameter Analysis

➢ All kinds of hyper-parameters analysis results show that GLC is not sensitive to hyper-parameter selection.

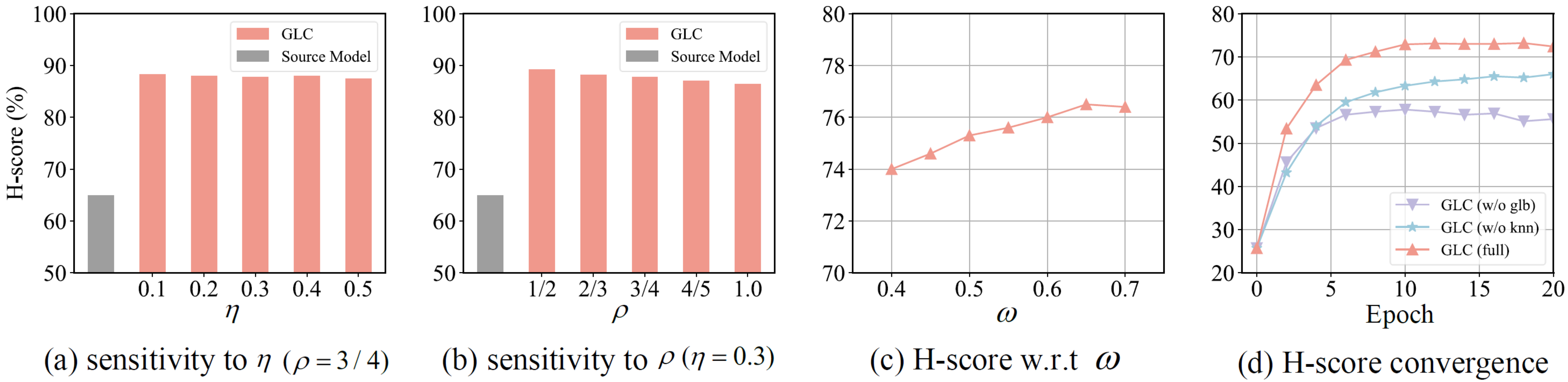


Figure 3. **Analysis of GLC.** (a-b) present the hyper-parameter sensitivity of η and ρ on Office-31 in OPDA. (c) plots the H-score with respect to ω on Office-Home in OPDA. (d) shows the H-score curves on VisDA in OPDA during the training process. Here GLC (w/o glb) refers to *GLC* w/o $\mathcal{L}_{tar}^{glb}$ and GLC w/o knn denotes *GLC w/o* $\mathcal{L}_{tar}^{loc}$.

❑ More results

◆ CLDA setting

- ➢ Most existing methods designed for category shift are not applicable to the vanilla closed-set DA (CLDA) setting.
- ➢ GLC is also applicable to the CLDA, and even outperforms methods specifically designed for CLDA.

Table 1. Accuracy (%) comparison in CLDA scenario on Office-Home and Office-31. (Best in **bold**)

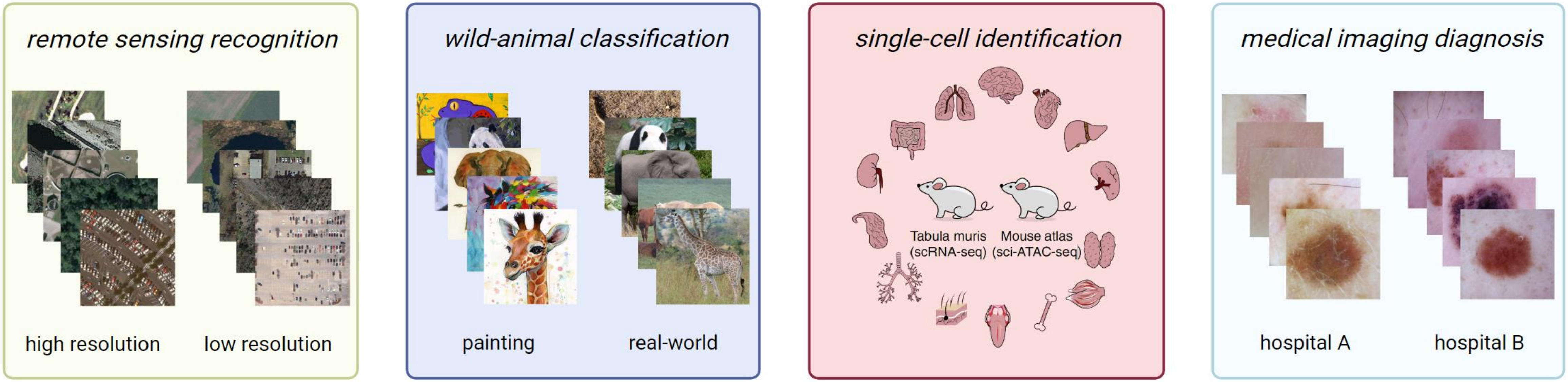
Methods	SF	OPDA	OSDA	PDA	CLDA	Office-Home													Office-31
						Ar2Cl	Ar2Pr	Ar2Re	Cl2Ar	Cl2Pr	Cl2Re	Pr2Ar	Pr2Cl	Pr2Re	Re2Ar	Re2Cl	Re2Pr	Avg	Avg
CDAN [10]	✗	✗	✗	✗	✓	49.0	69.3	74.5	54.4	66.0	68.4	55.6	48.3	75.9	68.4	55.4	80.5	63.8	86.6
MDD [26]	✗	✗	✗	✗	✓	**54.9**	73.7	77.8	60.0	71.4	71.8	61.2	53.6	78.1	72.5	**60.2**	82.3	68.1	**88.9**
UAN [23]	✗	✓	✗	✗	✗	45.0	63.6	71.2	51.4	58.2	63.2	52.6	40.9	71.0	63.3	48.2	75.4	58.7	84.4
CMU [3]	✗	✓	✗	✗	✗	42.8	65.6	74.3	58.1	63.1	67.4	54.2	41.2	73.8	66.9	48.0	78.7	61.2	79.9
DANCE [15]	✗	✓	✓	✗	✗	54.3	75.9	78.4	64.8	72.1	73.4	63.2	53.0	79.4	**73.0**	58.2	82.9	69.1	85.5
DCC [6]	✗	✓	✓	✓	✗	35.4	61.4	75.2	45.7	59.1	62.7	43.9	30.9	70.2	57.8	41.0	77.9	55.1	87.4
OVANet [16]	✗	✓	✓	✗	✗	34.5	55.8	67.1	40.9	52.8	56.9	35.4	26.2	61.8	53.8	35.4	70.8	49.3	70.4
Source-only	✓	-	-	-	-	44.8	67.4	74.2	53.0	63.3	65.1	53.7	40.5	73.5	65.6	46.3	78.3	60.5	78.8
UMAD [8]	✓	✓	✓	✗	✗	48.0	65.1	73.0	58.6	65.3	67.9	58.2	47.3	74.0	69.4	53.0	77.8	63.1	81.7
GLC	✓	✓	✓	✓	✗	51.2	**76.0**	**79.9**	**65.4**	**78.6**	**78.7**	**65.6**	**54.1**	**81.6**	70.9	58.4	**84.2**	**70.4**	88.1

Experiments

❑ More results

◆ Real-world applications

➢ In addition to experiments on standard computer science benchmarks, we have also validated the effectiveness of GLC in realistic applications. Results are included in the Appendix of our paper.



Real-world applications evaluation

Conclusion

❑ Conclusion

- To the best of our knowledge, we are the first to exploit and achieve the Source-free Universal Domain Adaptation (SF-UniDA) with only a standard pre-trained closed-set model.

- We propose a generic global and local clustering technique (GLC) to address the SF-UniDA. The global one-vs-all clustering can achieve pseudo labeling under various category-shift.

- Extensive experiments on four standard benchmarks demonstrate the superiority of GLC. Remarkably, in the open-partial-set DA (OPDA) situation, GLC attains an H-score of 73.1% on the VisDA benchmark, which is 14.8% and 16.7 higher that UMAD and GATE, respectively.

Thanks for you watching!

- If you have any further question or require any further information, please feel free to contact me. Email: 2011444@tongji.edu.cn ; Wechat ID: sanqing2020
- Code is already available at https://github.com/ispc-lab/GLC
- **Ad:** We also published a paper on single domain generalization titled “Modality Agnostic Debiasing for Single Domain Generalization” in CVPR-2023.