Implicit Semantic Response Alignment for Partial Domain Adaptation

Wenxiao Xiao1* Zhengming Ding[2] Hongfu Liu[1]

[1]Brandeis University [2]Tulane University

*wenxiaoxiao@brandeis.edu

2021/12/07

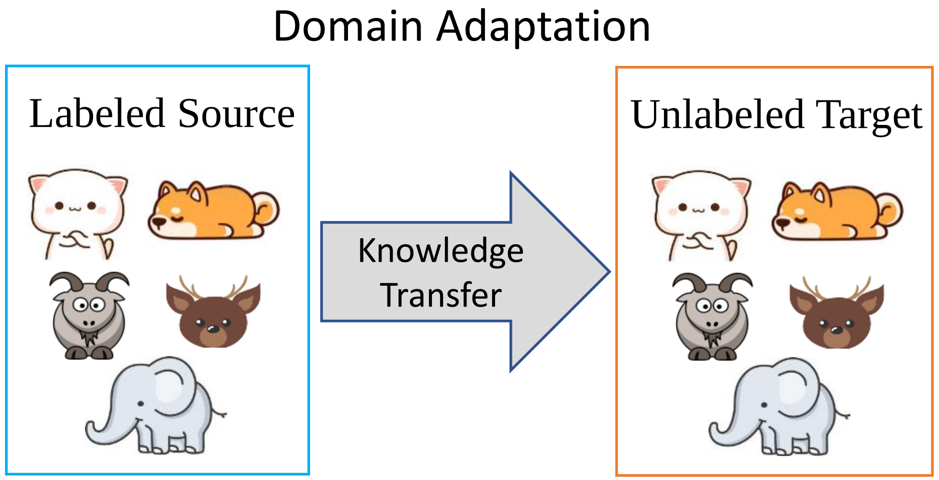
Partial Domain Adaptation

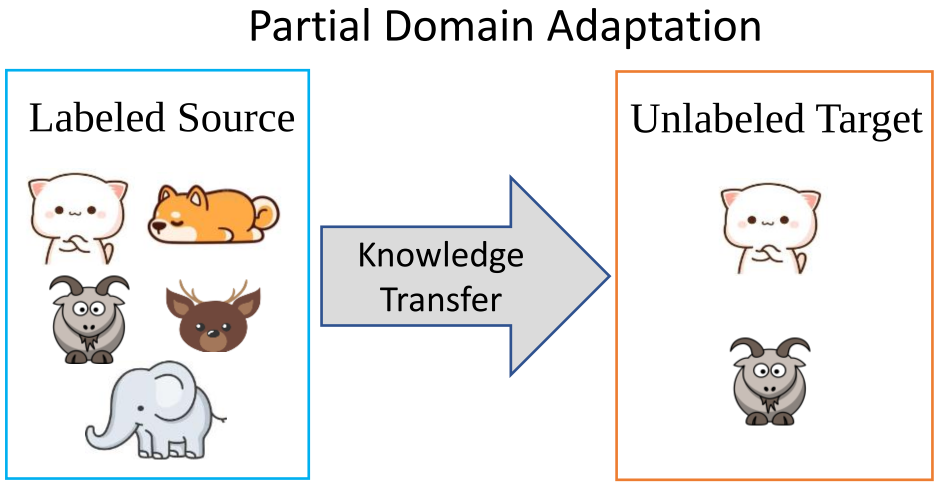
- **Domain Adaptation** (DA) aims to learn transferable representations from a well-labeled source domain to a different but related unlabeled target domain

- A Domain Adaptation model trained on the source and target data needs to learn how to classify the target samples without accessing the target labels

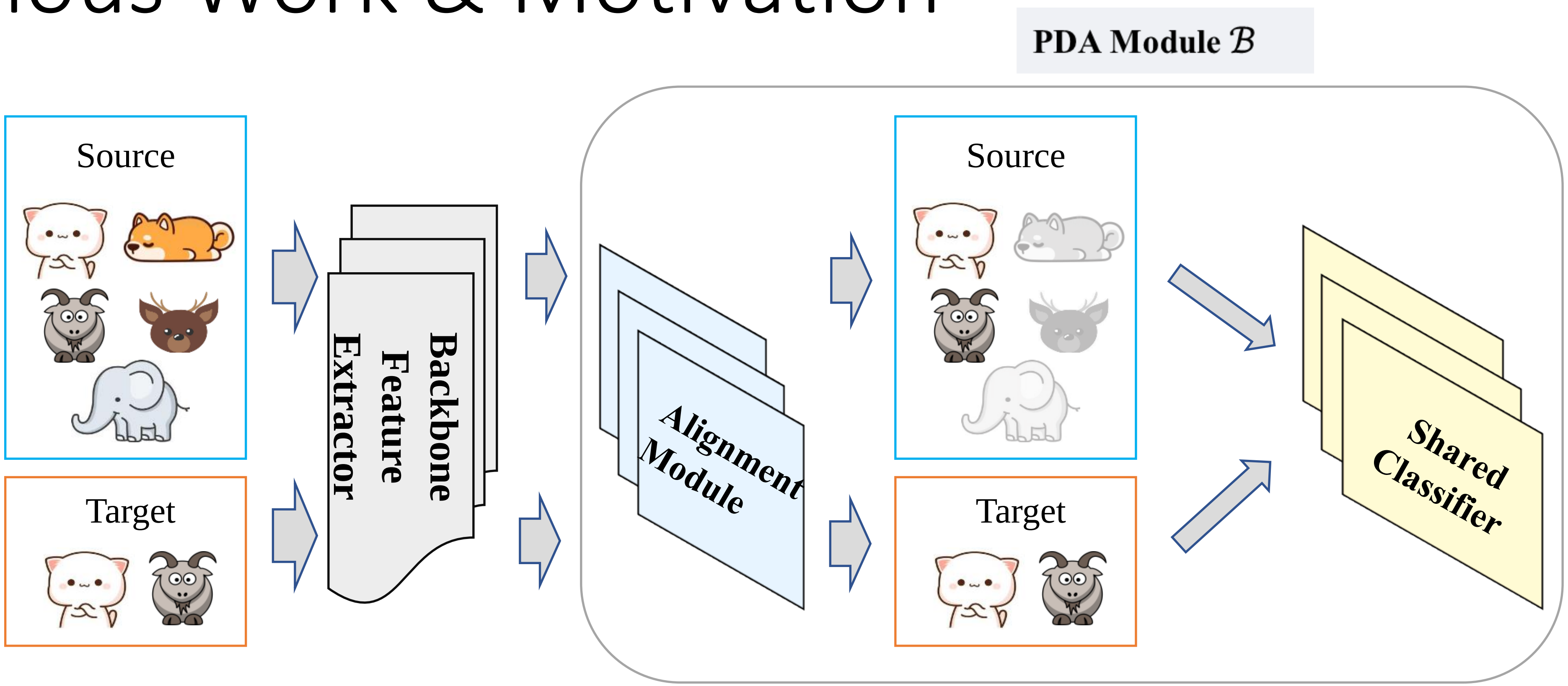
- Tradition DA methods require the source and target domain share the exact same set of object categories

- While Partial Domain Adaptation focuses on a more realistic situation where target label space is only a subset of source labels space



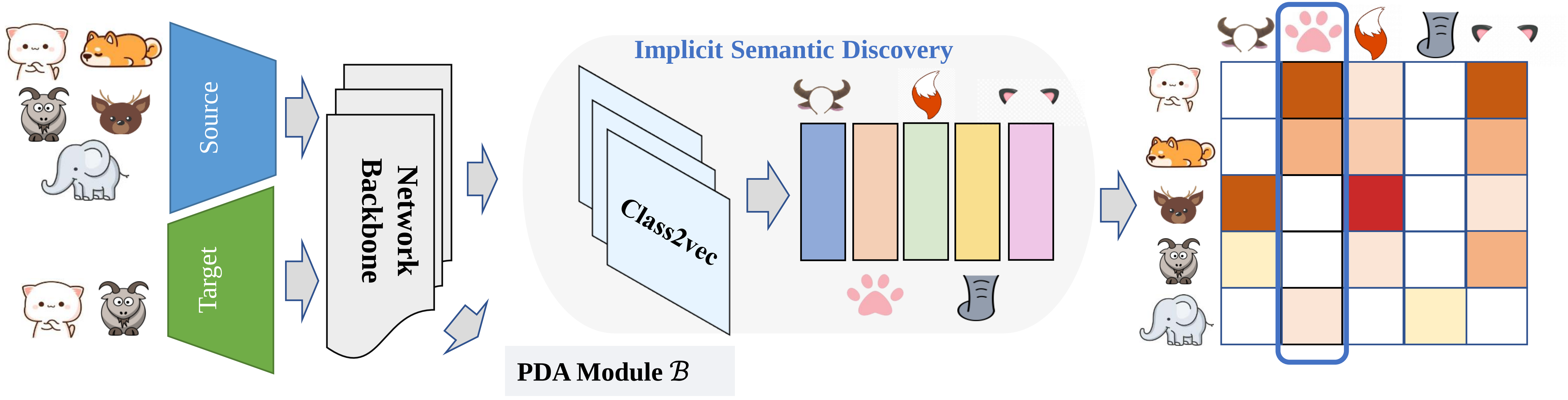


Previous Work & Motivation



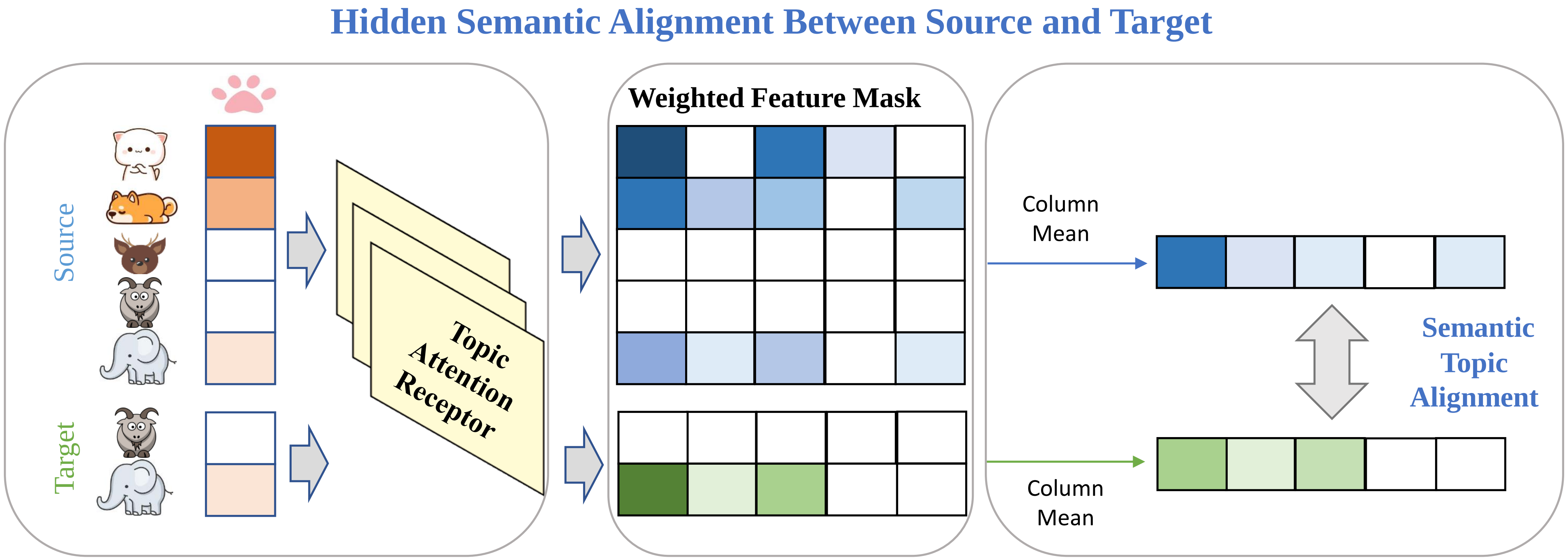
- Previous PDA methods aims to align source and target domains by down-weighting irrelevant categories
- However, we believe the irrelevant categories still contain important information for positive transfer
- For example, cats and dogs have clear distinguished features for class separation
- On the other hand, they also share many common semantic topics including fur and four legs
- We want to extract these semantics and align them between two domains by weighting on the feature level

Methodology



- To this end, we propose Implicit Semantic Response Alignment for Partial Domain Adaptation as an add-on module
- The implicit semantic discovery module extracts semantics from the backbone features with a class2vec machine
- Each data points will be represented by an embedding vector corresponding to extracted semantics
- Each semantic topic guides the following source and target feature space alignment as an intermediate signal
- Next, we will demonstrate the semantic alignment for one semantic topic (pawls for example)

Methodology



- For each semantic, a topic attention receptor retrieves the attention corresponding to the backbone features
- The attention map has the same dimension as features and can be used as feature-level weights
- Weighted feature masks is calculated by taking the dot product between features and attention weights
- The column-wise mean vectors for the source and target feature masks are then aligned together with l_2 loss

Experiments & Results

- We add our module to the state-of-art partial domain adaptation model BA^{3}US and conduct comprehensive experiments on three PDA benchmarks: *Office-Home*, *ImageNet-Caltech* and *Office31*
- As shown in Table 1&2, our method achieves best prediction accuracy in 8 out of 12 task on the challenging *Office-Home* and improves the BA^{3}US by 2.22%
- For the large-scale *ImageNet-Caltech* we also get the state-of-art results in both tasks and improve task I->C by 1.28%, where the source domain contains a large number of irrelevant categories
- For *Office31*, our method achieves best or second best for all tasks and improves BA^{3}US in 5 out of 6 tasks

Table 1: Accuracy for Partial Domain Adaptation on *Office-Home*

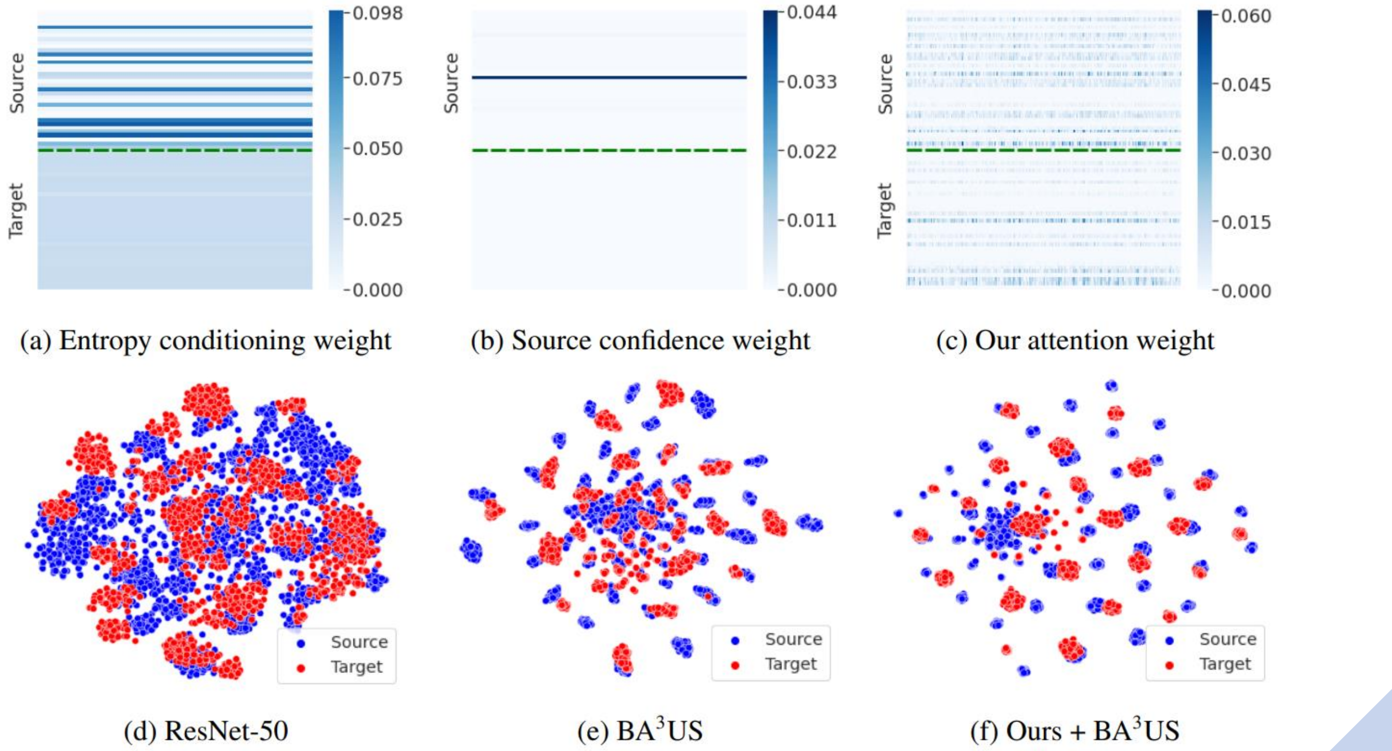
Method	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Pr	Avg.
ResNet-50 [11]	46.33	67.51	75.87	59.14	59.94	62.73	58.22	41.79	74.88	67.40	48.18	74.17	61.35
CDAN+E [26]	47.52	65.91	75.65	57.07	54.12	63.42	59.60	44.30	72.39	66.02	49.91	72.80	60.73
IWAN [50]	53.94	54.45	78.12	61.31	47.95	63.32	54.17	52.02	81.28	76.46	56.75	82.90	63.56
SAN [2]	44.42	68.68	74.60	67.49	64.99	77.80	59.78	44.72	80.07	72.18	50.21	78.66	65.30
PADA [3]	51.95	67.00	78.74	52.16	53.78	59.03	52.61	43.22	78.79	73.73	56.60	77.09	62.06
MWPDA [14]	55.39	77.53	81.27	57.08	61.03	62.33	68.74	56.42	86.67	76.70	57.67	80.06	68.41
ETN [4]	59.20	77.03	79.54	62.92	65.73	75.01	68.29	55.37	84.37	75.72	57.66	84.50	70.45
DRCN [19]	54.00	76.40	83.00	62.10	64.50	71.00	70.80	49.80	80.50	77.50	59.10	79.90	69.00
AFN [47]	58.93	76.25	81.42	70.43	72.97	77.78	72.36	55.34	80.40	75.81	60.42	79.90	71.83
SLM [37]	56.54	**83.75**	**90.40**	**76.03**	73.99	80.95	72.97	56.60	87.32	**82.55**	59.76	82.52	75.29
BA^{3}US [23]	60.62	83.16	88.39	71.75	72.79	83.40	75.45	61.59	86.53	79.25	62.80	86.05	75.98
Ours + BA^{3}US	**64.66**	82.97	89.12	75.67	**75.52**	**85.36**	**78.51**	**64.24**	**88.07**	81.27	**65.31**	**86.67**	**78.20**

Table 2: Accuracy for Partial Domain Adaptation on *Office31* and *ImageNet-Caltech*

Method	*Office31*							*ImageNet-Caltech*		
	A→D	A→W	D→A	D→W	W→A	W→D	Avg.	I→C	C→I	Avg.
ResNet-50 [11]	83.44	75.59	83.92	96.27	84.97	98.09	87.05	69.69	71.29	70.49
CDAN+E [26]	77.07	80.51	93.58	98.98	91.65	98.09	89.98	72.45	72.02	72.24
IWAN [50]	90.45	89.15	95.62	99.32	94.26	99.36	94.69	78.06	73.33	75.70
SAN [2]	94.27	93.90	94.15	99.32	88.73	99.36	94.96	77.75	75.26	76.51
PADA [3]	82.17	86.54	92.69	99.32	95.41	**100.00**	92.69	77.03	70.48	73.76
MWPDA [14]	95.12	96.61	95.02	**100.00**	95.51	**100.00**	97.04	-	-	-
ETN [4]	95.03	94.52	96.21	**100.00**	94.64	**100.00**	96.73	83.23	74.93	79.08
DRCN [19]	86.00	88.05	95.60	**100.00**	95.80	**100.00**	94.24	75.30	78.90	77.10
SLM [37]	98.73	**99.77**	**96.1**	**100.00**	**95.89**	99.79	**98.38**	82.31	81.41	81.86
BA^{3}US [23]	**99.36**	98.98	94.82	**100.00**	94.99	98.73	97.80	84.00	83.35	83.68
Ours + BA^{3}US	98.73	99.32	95.41	**100.00**	95.41	**100.00**	98.15	**85.28**	**83.73**	**84.50**

Topic Attention Weighting

- Here we use task Ar→Cl on *Office-Home* to visualize the effect of our topic attention weighting
- In figure (a-c), we visualize the weights on the features of one mini-batch. As shown in figure (c), our weighting schema discoveries the information that responds to the same implicit topic on the feature level
- The t-SNE visualizations of features in figure (d-f) demonstrate that our proposed method divides existing clusters in ResNet and BA^{3}US into smaller and well-separated clusters related to implicit semantic topics



(a) Entropy conditioning weight (b) Source confidence weight (c) Our attention weight
(d) ResNet-50 (e) BA^{3}US (f) Ours + BA^{3}US

(a-c) Entropy conditioning weight and source confidence weight of BA^{3}US in one mini-batch of task the Ar→Cl ; (c) Ours attention map for the same mini-batch; (d-f) t-SNE visualization of features from Resnet-50, BA^{3}US and ours of task Ar→Cl

Cross-Class Interaction

- Remember our motivation is to use the shared semantic in the extra classes to promoter the positive transfer for related target classes
- Thus, we draw the similarity matrix among 9 extra source classes and 4 shared classes in the right figure. And check if our method benefits the target classes that are similar to the extra classes
- According to this figure, Computer and Clipboards are highly correlated to the extra source classes. And Table 3 shows they benefit most from our method
- On the other hand, the accuracy of Candles class decreases by 9.09% in our method, which indicates that semantic alignment may introduce noisy for the classes that do not share semantic with extra classes

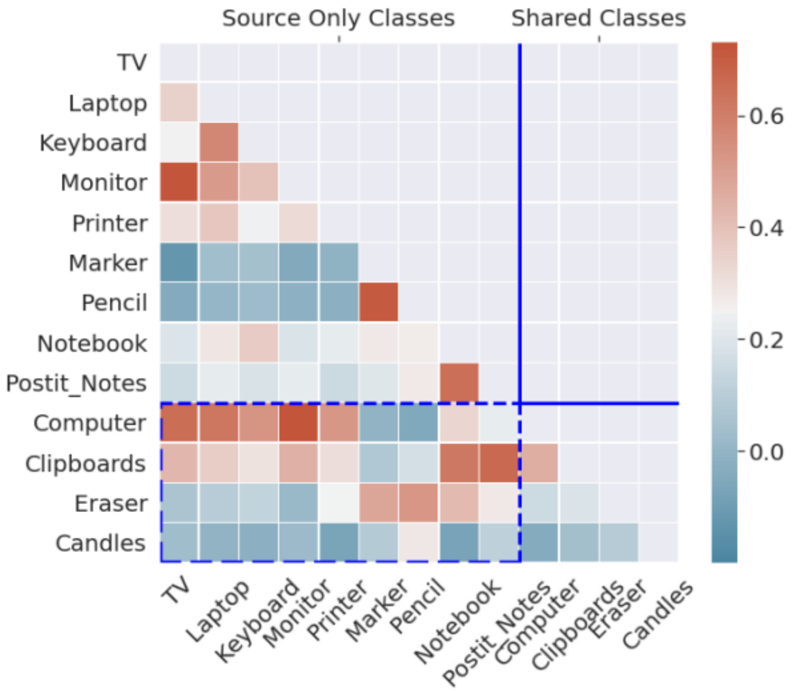


Table 3: Partial domain adaptation on individual class for task Ar→Cl on *Office-Home*

Class	n_s	n_t	BA^3US	Ours	Improv.(%)
Computer	99	44	12.12	59.60	47.48
Clipboards	40	25	67.50	87.50	20.00
Eraser	40	18	0.00	0.00	0.00
Candles	99	76	79.80	70.71	-9.09
All classes	2427	1675	60.62	64.66	4.04

Thanks!