Class Conditional Alignment for Partial Domain Adaptation

Mohsen Kheirandishfard   Fariba Zohrizadeh   Farhad Kamangar

Department of Computer Science and Engineering
The University of Texas at Arlington

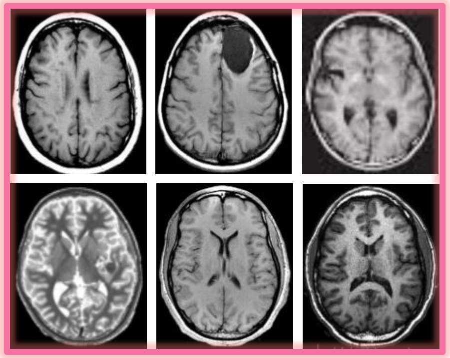
January 2021

Introduction

- The performance of deep models heavily depends on the availability of a large amount of labeled training data.

How to handle unlabeled datasets? 

Traffic Sign Classification

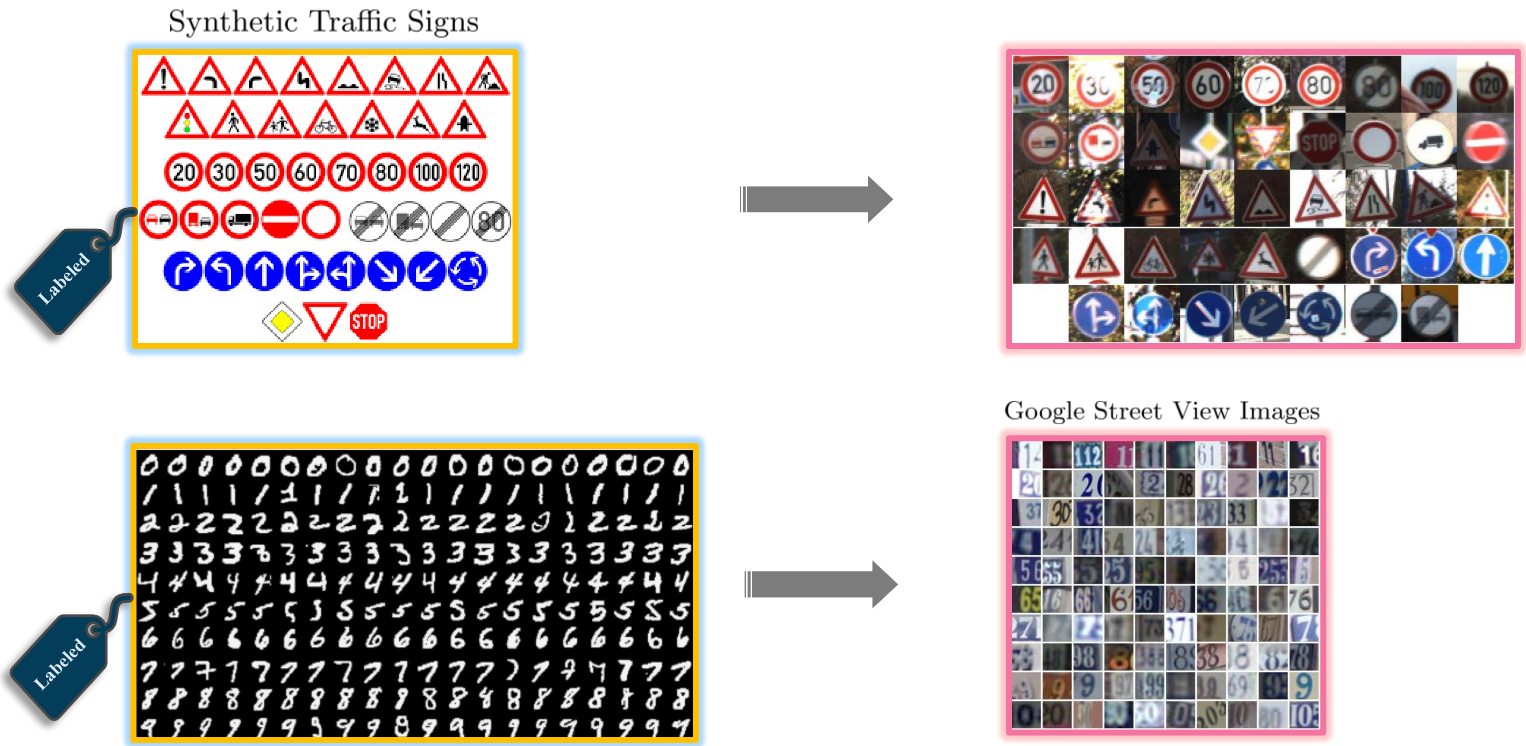


Brain Segmentation

Digit Classification

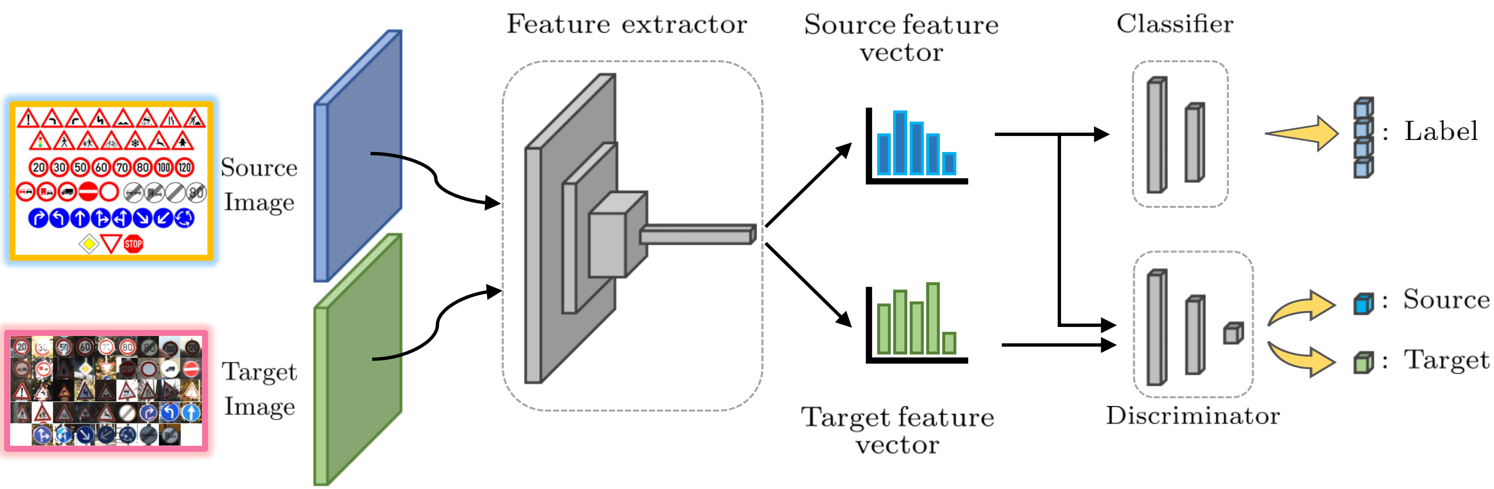
Introduction

- The performance of deep models heavily depends on the availability of a large amount of labeled training data.
- A naive strategy is to transfer knowledge from available labeled source datasets of related but different domains. This strategy may exhibit poor performance due to the **shift** between the distributions of the source and target domains.



Related Works

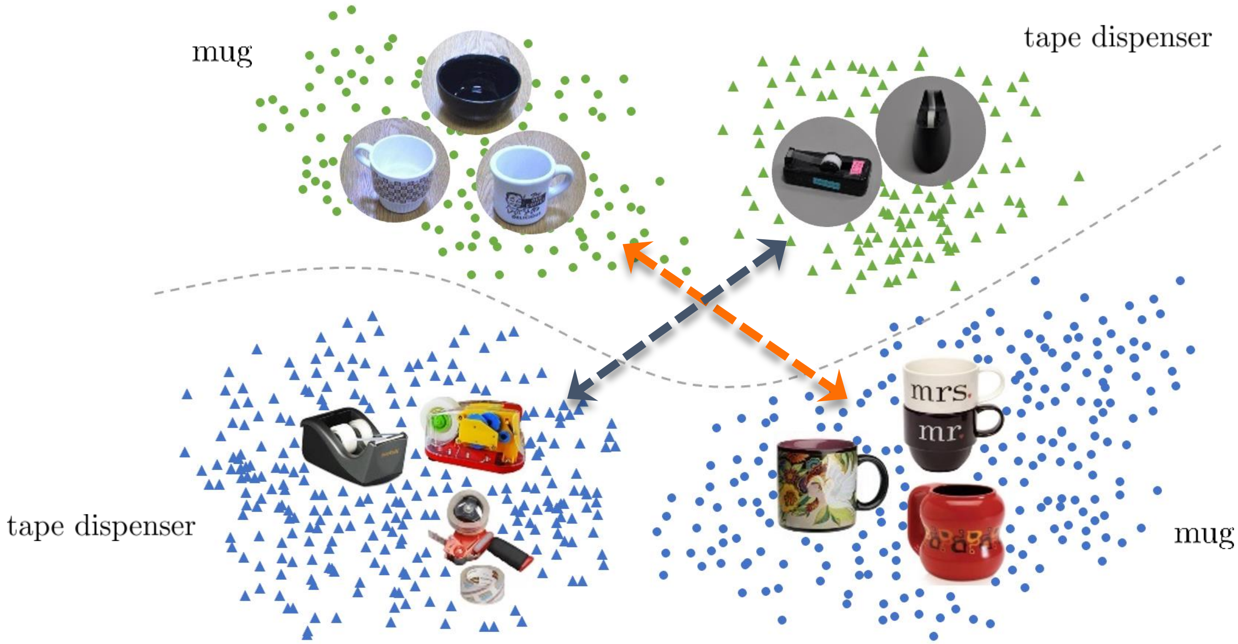
- **Unsupervised Domain Adaptation** (UDA) aims to eliminate the shift between the domains and train a classifier on source labeled samples that generalizes well on the target domain.



Related Works

- **Unsupervised Domain Adaptation** (UDA) aims to eliminate the shift between the domains and train a classifier on source labeled samples that generalizes well on the target domain.

- UDA relies on the assumption that the domains share the same label space.

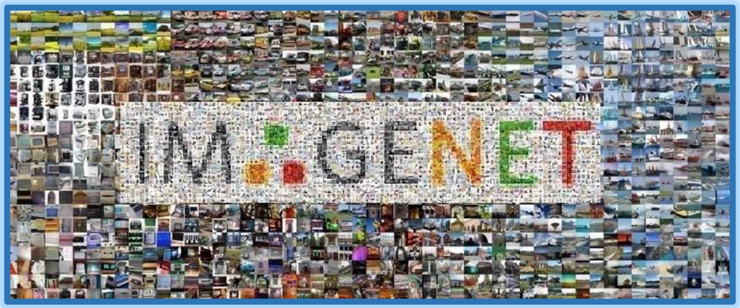


Related Works

- **Unsupervised Domain Adaptation** (UDA) aims to eliminate the shift between the domains and train a classifier on source labeled samples that generalizes well on the target domain.

- UDA relies on the assumption that the domains share the same label space.

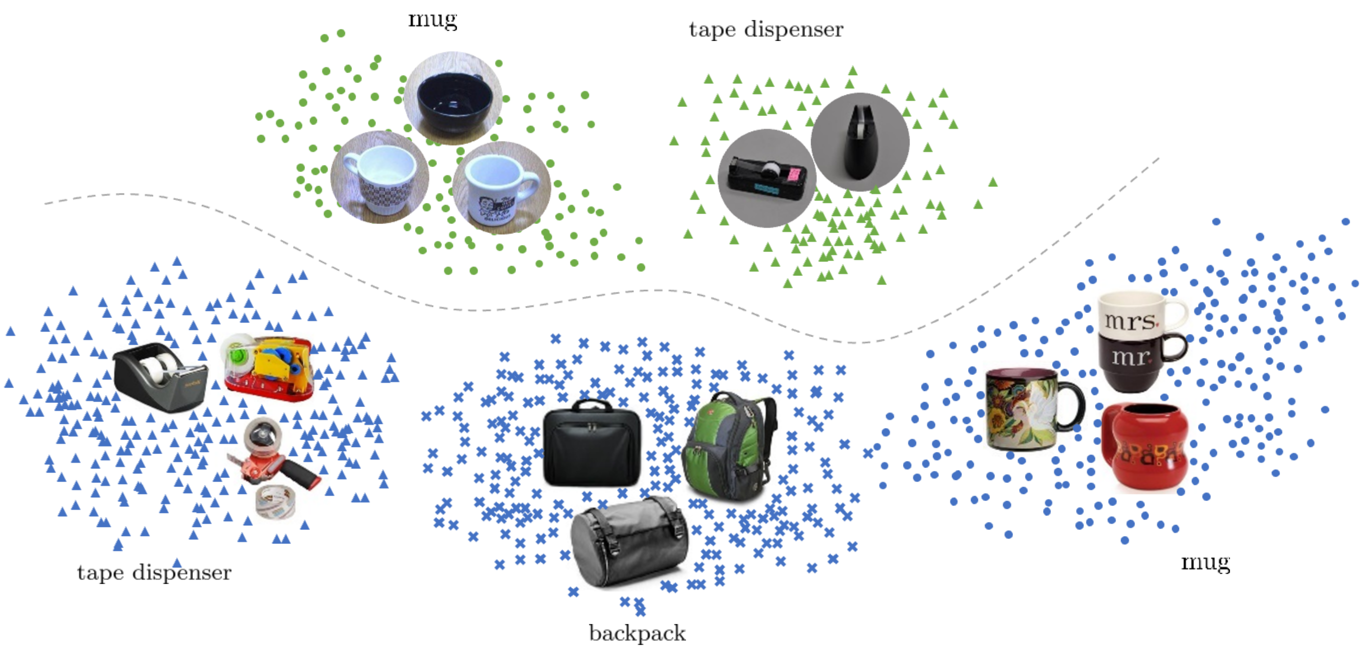
What if the domains do not share the same label space?



Problem Definition

- **Partial Domain Adaptation** (PDA) is an adaptation scenario in which the target label space is a subset of the source label space.

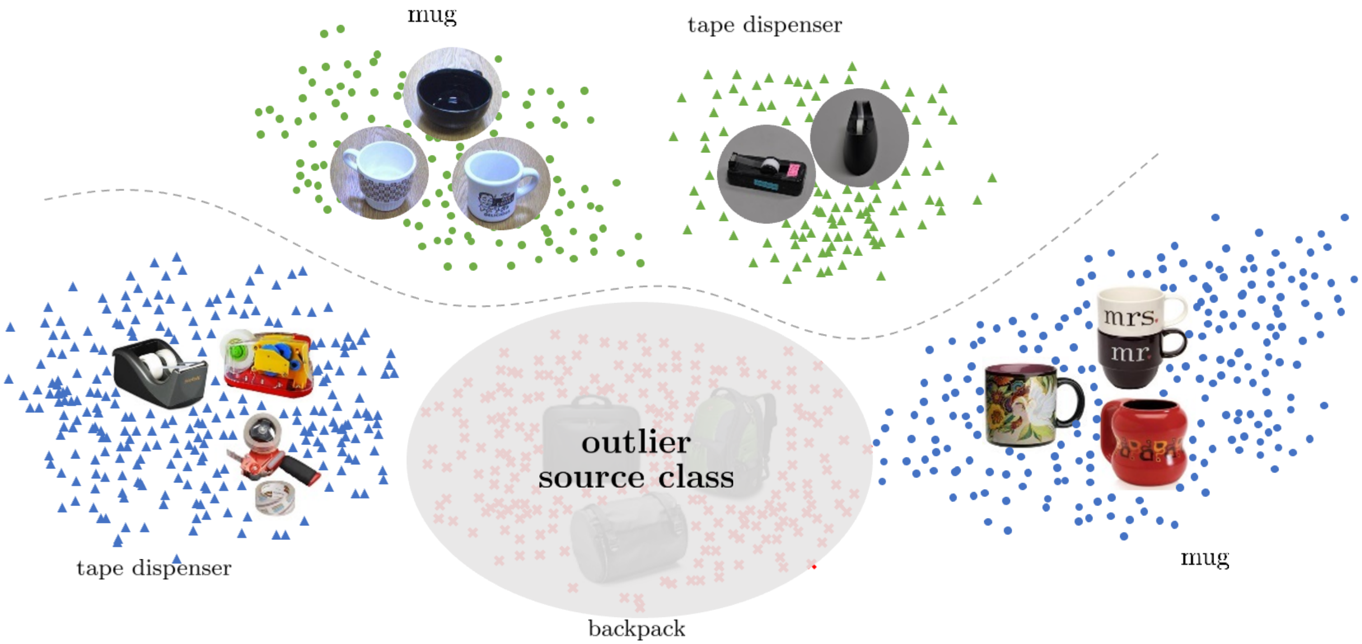
- The main purpose of PDA methods is to identify and reject the outlier source classes and align the domain distributions across the shared label space.



Problem Definition

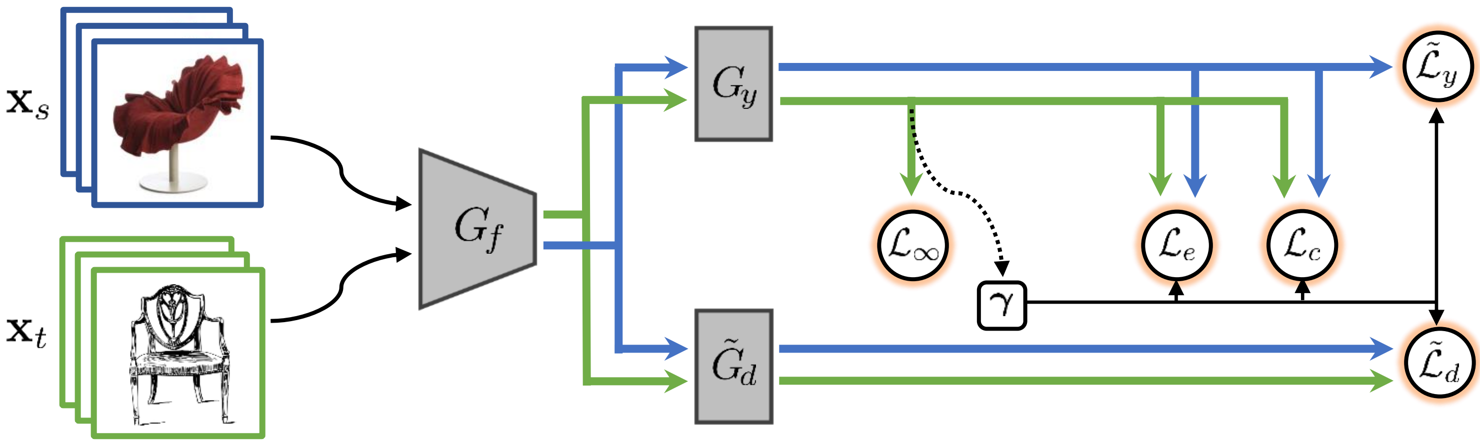
- **Partial Domain Adaptation** (PDA) is an adaptation scenario in which the target label space is a subset of the source label space.

- The main purpose of PDA methods is to identify and reject the outlier source classes and align the domain distributions across the shared label space.



Proposed Method

- Propose PDA method can
  - identify and reject the outlier source classes
  - increase the confidence level of the classifier
  - align the marginal and class-conditional distributions of both domains.

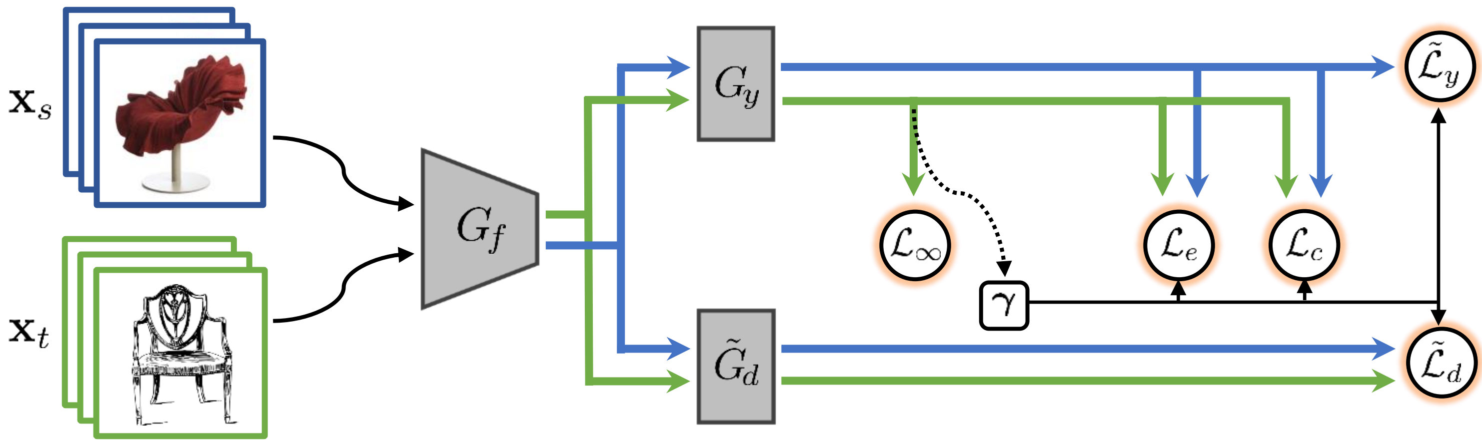


Multiclass Discriminator generates class domain vectors of size $2K$ where $K = |\mathcal{C}_s|$.

$$(\mathbf{x}^i, \tilde{\mathbf{d}}^i),\ \tilde{\mathbf{d}}^i = \begin{cases} [\mathbf{y}_s^i, \mathbf{0}] & \mathbf{x}^i \in \mathcal{X}_s \\ [\mathbf{0}, \tilde{\mathbf{y}}_t^i] & \mathbf{x}^i \in \mathcal{X}_t \end{cases}$$

Proposed Method

- Propose PDA method can
  - identify and reject the outlier source classes
  - increase the confidence level of the classifier
  - align the marginal and class-conditional distributions of both domains.



$$\boxed{\max_{\boldsymbol{\theta}_d} \min_{\boldsymbol{\theta}_y,\boldsymbol{\theta}_f} \; \mathcal{L}_y(\boldsymbol{\theta}_f,\boldsymbol{\theta}_y) + \lambda\,\tilde{\mathcal{L}}_d(\boldsymbol{\theta}_f,\tilde{\boldsymbol{\theta}}_d,\boldsymbol{\theta}_y) + \mathcal{L}_c(\boldsymbol{\theta}_f,\boldsymbol{\theta}_y) + \mu\,\mathcal{L}_\infty(\boldsymbol{\theta}_f,\boldsymbol{\theta}_y) + \zeta\,\mathcal{L}_e(\boldsymbol{\theta}_f,\boldsymbol{\theta}_y)}$$

Numerical Results

- Experiments on two commonly used datasets: Office-31 and Office-Home.

Office-31 dataset

Webcam (**W**)

Digital camera (**D**)

Amazon (**A**)

Accuracy of partial domain adaptation tasks on Office-31

Method	**A → W**	**D → W**	**W → D**	**A → D**	**D → A**	**W → A**	**Avg**
ResNet	75.59	96.27	98.09	83.44	83.92	84.97	87.05
DANN	73.56	96.27	98.73	81.53	82.78	86.12	86.50
ADDA	75.67	95.38	99.85	83.41	83.62	84.25	87.03
RTN	78.98	93.22	85.35	77.07	89.25	89.46	85.56
IWAN	89.15	99.32	99.36	90.45	95.62	94.26	94.69
SAN	93.90	99.32	99.36	94.27	94.15	88.73	94.96
PADA	86.54	99.32	**100.0**	82.17	92.69	95.41	92.69
ETN	94.52	**100.0**	**100.0**	95.03	**96.21**	94.64	96.73
CCPDA	**99.66**	**100.0**	**100.0**	**97.45**	95.72	**95.71**	**98.09**

Numerical Results

- Experiments on two commonly used datasets: Office-31 and Office-Home.

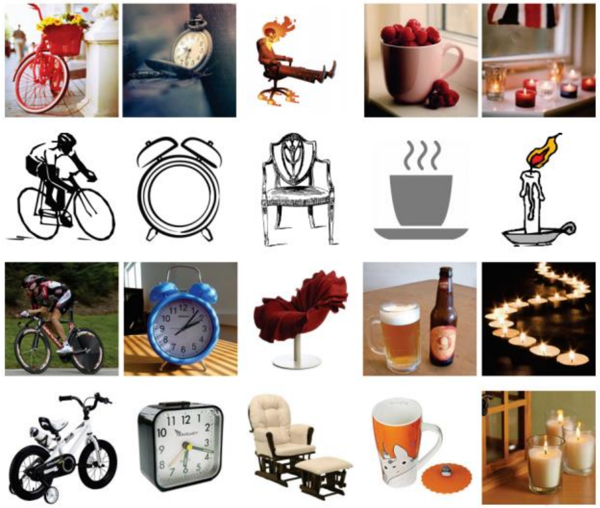
Office-Home dataset

Art (**Ar**)

Clipart (**Cl**)

Real-World (**Rw**)

Product (**Pr**)



Accuracy of partial domain adaptation tasks 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	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
DANN	43.76	67.90	77.47	63.73	58.99	67.59	56.84	37.07	76.37	69.15	44.30	77.48	61.72
ADDA	45.23	68.79	79.21	64.56	60.01	68.29	57.56	38.89	77.45	70.28	45.23	78.32	62.82
RTN	49.31	57.70	80.07	63.54	63.47	73.38	65.11	41.73	75.32	63.18	43.57	80.50	63.07
IWAN	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	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	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
ETN	**59.24**	77.03	79.54	62.92	65.73	75.01	68.29	**55.37**	**84.37**	75.72	**57.66**	84.54	70.45
CCPDA	55.31	**80.11**	**88.07**	**73.28**	**71.21**	77.63	**71.89**	52.97	81.41	**81.81**	56.21	**85.15**	**72.92**

Ablation Study

- **Ablation Study** shows the benefits brought by different components of our method.

Accuracy of CCPDA and its variants for Partial Domain Adaptation on Office-31 dataset

Method	**A→W**	**D→W**	**W→D**	**A→D**	**D→A**	**W→A**	**Avg**
CCPDA_{∞}	95.12	99.32	100.0	93.21	96.03	95.19	96.48
CCPDA_{e}	97.45	96.64	100.0	96.47	94.92	93.86	96.56
$\text{CCPDA}_{d,c}$	93.42	97.62	100.0	90.43	93.45	95.53	95.07
CCPDA	99.66	100.0	100.0	97.45	95.72	95.71	98.09

Conclusion

- The proposed approach adopts a multi-class adversarial loss function to jointly align the marginal and class-conditional distributions across the shared classes between the domains.

- The regularization terms reduce the effects of outlier classes and can be directly incorporated into many adversarial architectures.

- Experiments on a benchmark dataset demonstrate the high potential of our approach for different PDA tasks.

Thank You