南京航空航天大學
Nanjing University of Aeronautics and Astronautics

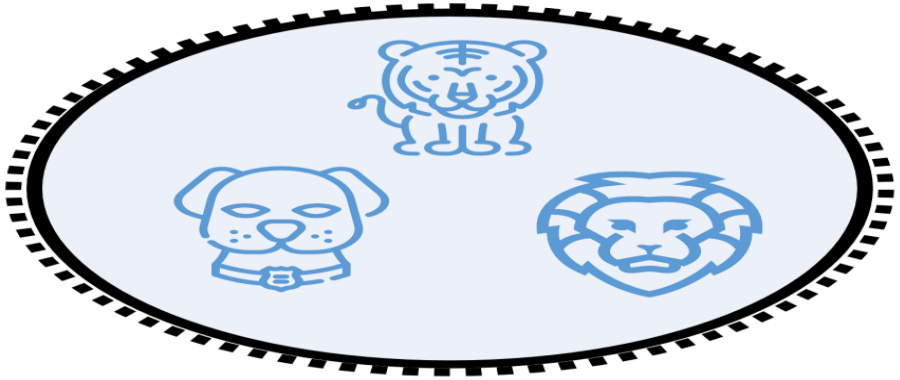
Universal Domain Adaptation

Kaichao You[1], Mingsheng Long[1](✉), Zhangjie Cao[1], Jianmin Wang[1], and Michael I. Jordan[2]
[1]KLiss, MOE; BNRist; School of Software, Tsinghua University, China
[1]Research Center for Big Data, Tsinghua University, China
[1]Beijing Key Laboratory for Industrial Big Data System and Application
[2]University of California, Berkeley, USA
youkaichao@gmail.com, {mingsheng,jimwang}@tsinghua.edu.cn, jordan@cs.berkeley.edu

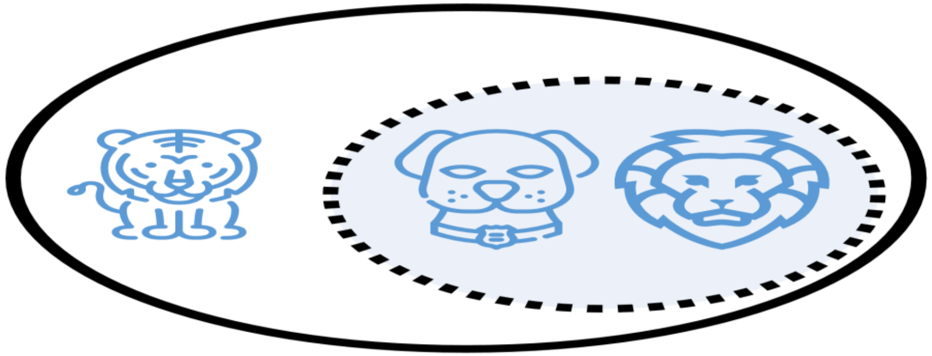
CVPR2019

Domain Adaptation

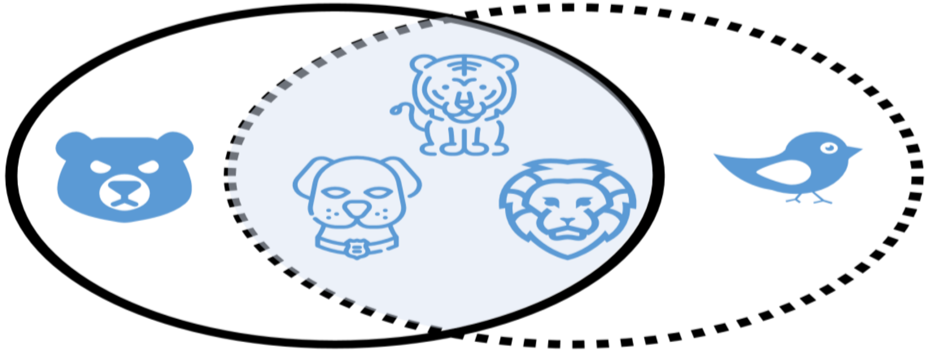
Closed Set DA



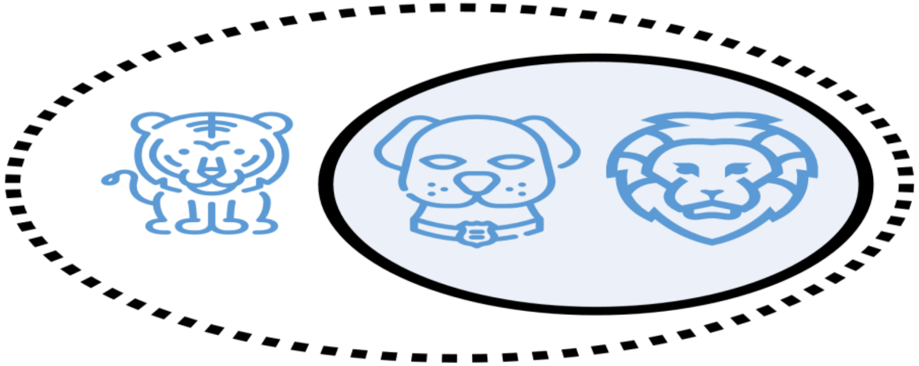
Partial DA



Open Set DA *(Busto et al. 2017)*



Open Set DA *(Saito et al. 2018)*

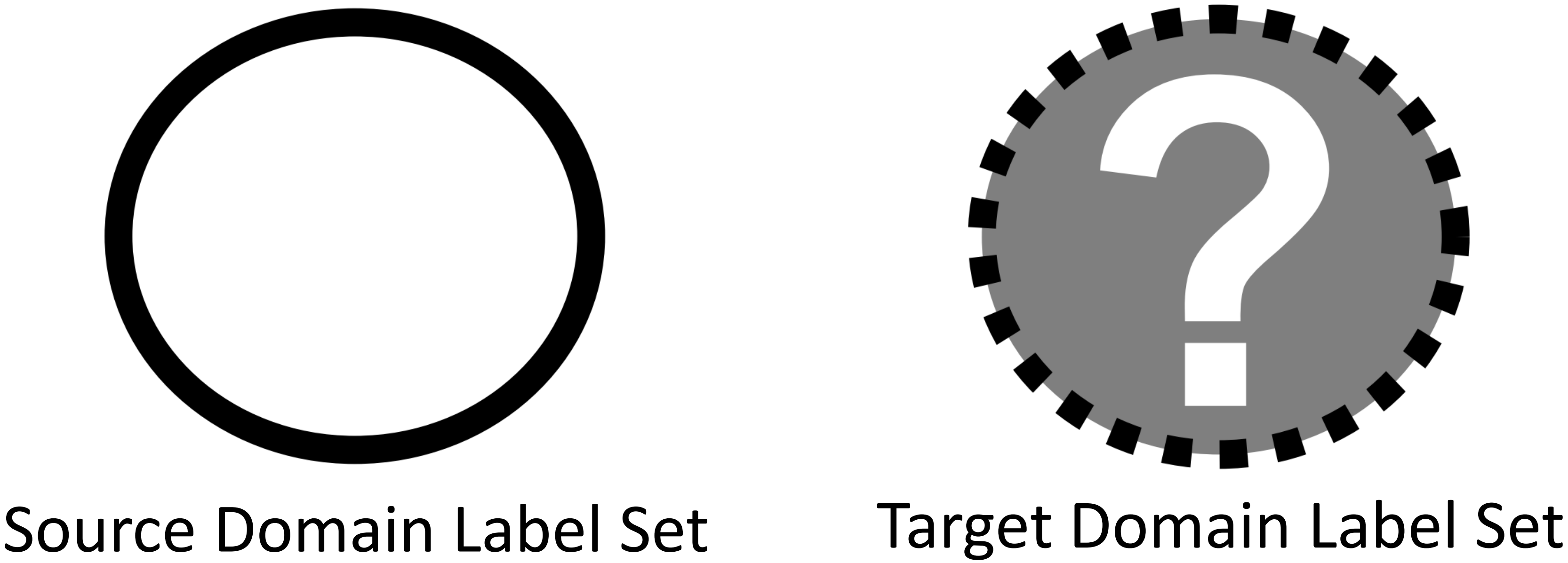


These settings rely on the rich prior knowledge
about the relationship between the label sets of source and target domains

Universal Domain Adaptation (UDA)

- Motivation： the relationship of label sets between the source and target domains is unknown in the presence of a large domain gap.

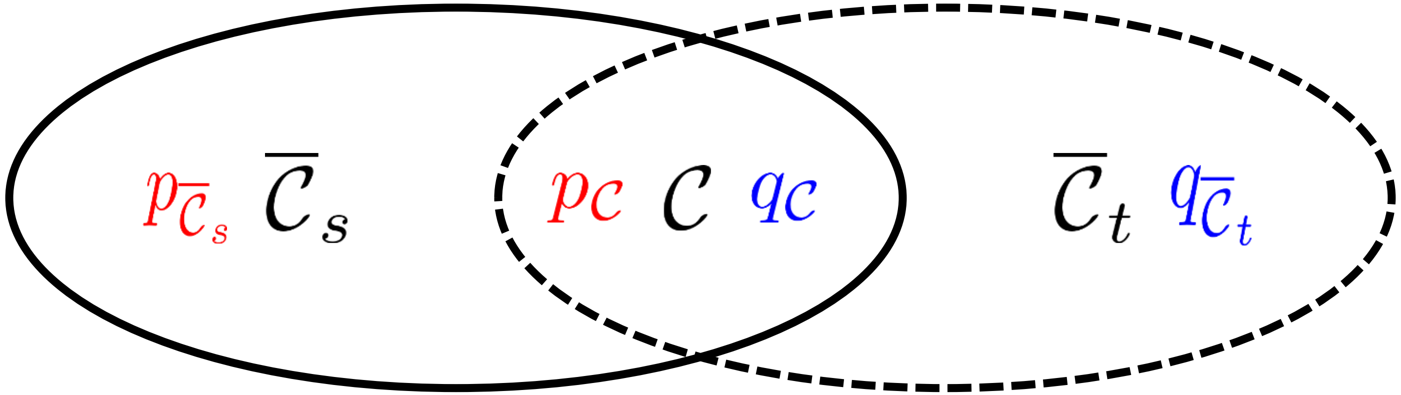
- Learning Goal： classify target samples correctly if it belongs any class in source label set, or mark it as “known” otherwise.



This setting imposes no prior knowledge on the label sets of source and target domains

Problem Setting

- Source data: $\mathcal{D}_s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s}$
- Source label set: $\mathcal{C}_s$
- Source distribution: p
- Source private label set: $\overline{\mathcal{C}}_s$
- Target data: $\mathcal{D}_t = \{\mathbf{x}_i^t\}_{i=1}^{n_t}$
- Target label set: $\mathcal{C}_t$
- Target distribution: q
- Target private label set: $\overline{\mathcal{C}}_t$

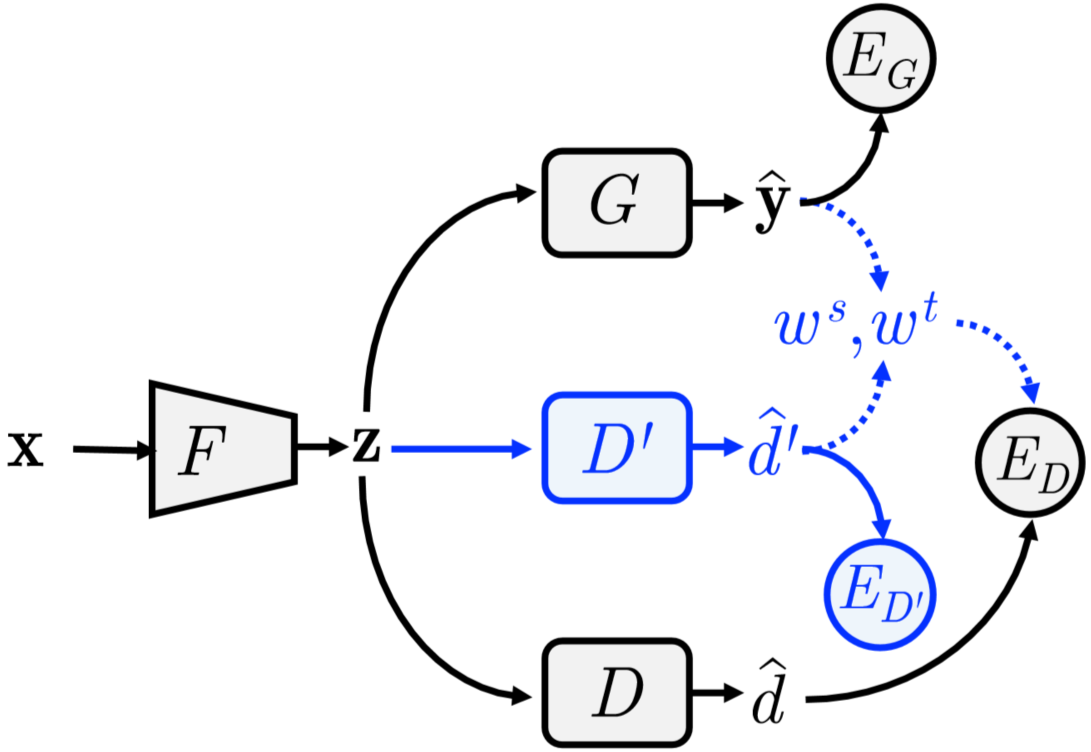


Align distributions of source and target data in the common label set

Challenges of UDA

- **Category gap:** $\mathcal{C}$, $\overline{\mathcal{C}}_s$ and $\overline{\mathcal{C}}_t$ are unknown
  - Automatically identify the source and target data from $\mathcal{C}$, such that feature alignment can be done in the auto-discovered common label set.
- **Domain gap:** $p \neq q$ and $p_{\mathcal{C}} \neq q_{\mathcal{C}}$
  - Domain adaptation should be applied to align distributions of source and target data in common label set $\mathcal{C}$.
- **Detect unknown class:** confidence thresholding
  - The predictions by neural networks are usually overconfident but less discriminative due to the underlying domain gap.

Universal Adaptation Network (UAN)



G : Label classifier

F : Feature extractor

D : Adversarial domain discriminator

D': Non-adversarial domain discriminator

$$E_G = \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim p} L\left(\mathbf{y}, G(F(\mathbf{x}))\right)$$

$$\begin{aligned} E_{D'} = &-\mathbb{E}_{\mathbf{x}\sim p} \log D'(F(\mathbf{x})) \\ &- \mathbb{E}_{\mathbf{x}\sim q} \log\left(1 - D'(F(\mathbf{x}))\right) \end{aligned}$$

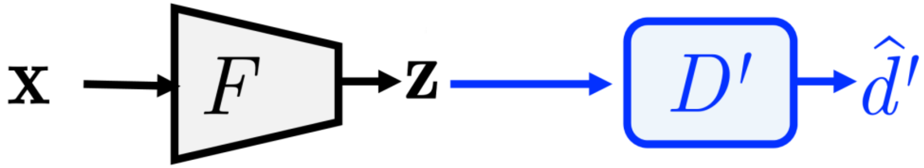
$$\begin{aligned} E_D = &-\mathbb{E}_{\mathbf{x}\sim p} w^s(\mathbf{x}) \log D(F(\mathbf{x})) \\ &- \mathbb{E}_{\mathbf{x}\sim q} w^t(\mathbf{x}) \log\left(1 - D(F(\mathbf{x}))\right) \end{aligned}$$

- The training of UAN can be written as a minmax game:

$$\max_{D} \min_{F,G} E_G - \lambda E_D$$

$$\min_{D'} E_{D'}$$

Domain Similarity



- Hypothesize:

$$\mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} \hat{d}'$$

- Reasonability:
  - For source data $d' = 1$, and for target data $d' = 0$, so:

$$\mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} \hat{d}', \mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} \hat{d}', \mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} \hat{d}'$$

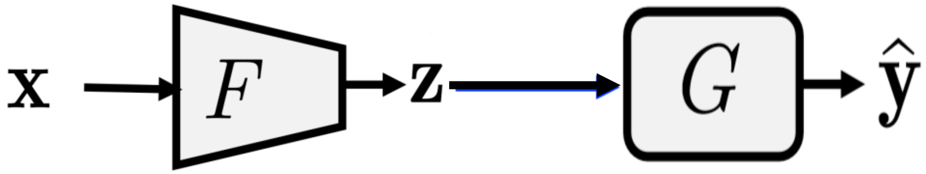
  - $p_{\mathcal{C}}$ is closer to $q_{\mathcal{C}}$ compared with $p_{\overline{\mathcal{C}}_s}$, so:

$$\mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} \hat{d}'$$

  - Similarly:

$$\mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} \hat{d}' > \mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} \hat{d}'$$

Prediction Uncertainty



- Hypothesize:

$$\mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} H(\hat{\mathbf{y}})$$

- Reasonability:
  - Predictions are certain for source data and uncertain for target data.

$$\mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} H(\hat{\mathbf{y}}), \mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} H(\hat{\mathbf{y}}), \mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} H(\hat{\mathbf{y}})$$

  - Similar samples from $p_{\mathcal{C}}$ and $q_{\mathcal{C}}$ can attract each other.

$$\mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} H(\hat{\mathbf{y}})$$

$$\mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} H(\hat{\mathbf{y}}) > \mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} H(\hat{\mathbf{y}})$$

Transferability Criterion

$$E_D = -\mathbb{E}_{\mathbf{x}\sim p} w^s(\mathbf{x}) \log D(F(\mathbf{x})) - \mathbb{E}_{\mathbf{x}\sim q} w^t(\mathbf{x}) \log(1 - D(F(\mathbf{x})))$$

- A well-established sample-level transferability criterion should satisfy:

$$\mathbb{E}_{\mathbf{x}\sim p_{\mathcal{C}}} w^s(\mathbf{x}) > \mathbb{E}_{\mathbf{x}\sim p_{\overline{\mathcal{C}}_s}} w^s(\mathbf{x})$$

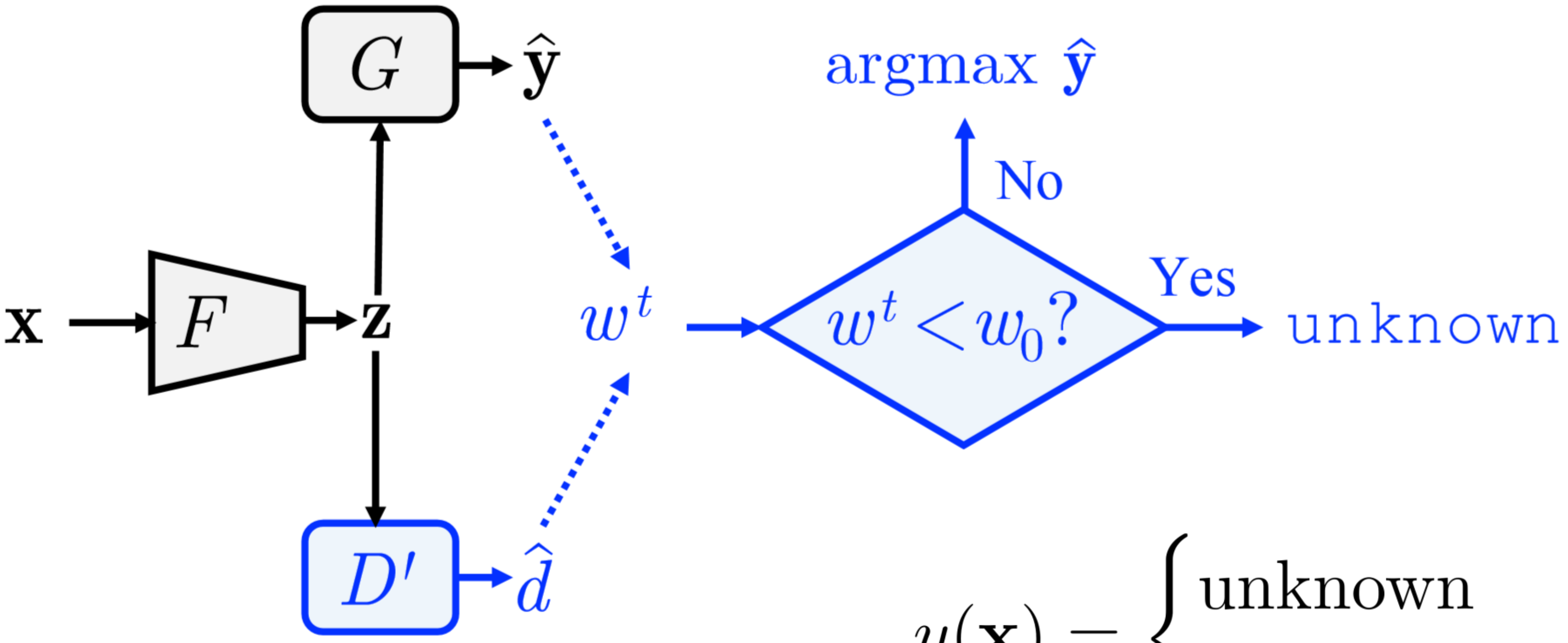
$$\mathbb{E}_{\mathbf{x}\sim q_{\mathcal{C}}} w^t(\mathbf{x}) > \mathbb{E}_{\mathbf{x}\sim q_{\overline{\mathcal{C}}_t}} w^t(\mathbf{x})$$

- Integrates both the **domain similarity** and the **prediction uncertainty** of each sample.

$$w^s(\mathbf{x}) = \frac{H(\hat{\mathbf{y}})}{\log|\mathcal{C}_s|} - \hat{d}'(\mathbf{x})$$

$$w^t(\mathbf{x}) = \hat{d}'(\mathbf{x}) - \frac{H(\hat{\mathbf{y}})}{\log|\mathcal{C}_s|}$$

Test Phase



$$y(\mathbf{x}) = \begin{cases} \text{unknown} & w^t < w_0 \\ \text{argmax}(\hat{\mathbf{y}}) & w^t \geq w_0 \end{cases}$$

- Either rejects the target sample $\mathbf{x}$ as "unknown" class or classifiers it to one of the source classes.

Experiment Setup

- Compared Methods:
  - Convolution neural networks: $\mathbf{ResNet}$
  - Close-set domain adaptation methods: $\mathbf{DaNN}, \mathbf{RTN}$
  - Partial domain adaptation methods: $\mathbf{IWAN}, \mathbf{PADA}$
  - Open-set domain adaptation methods: $\mathbf{ATI}, \mathbf{OSBP}$
- Two variants of UAN
  - $\mathbf{UAN\ w/o\ d}$: Only use domain similarity as weight
  - $\mathbf{UAN\ w/o\ y}$: Only use prediction uncertainty as weight
- Evaluation Protocols: the average of per-class accuracy for all the $|\mathcal{C}|_s + 1$ classes.

Experiment Result1

- $\mathcal{C}$: the first 10 classes
- $\overline{\mathcal{C}}_s$: the next 5 classes
- $\overline{\mathcal{C}}_t$: the rest 50 classes

Commonness $\xi = \frac{|\mathcal{C}_s \cap \mathcal{C}_t|}{|\mathcal{C}_s \cup \mathcal{C}_t|}$

Table 1. Average class accuracy (%) of universal domain adaptation tasks on **Office-Home** ($\xi = 0.15$) dataset (ResNet)

Method	Office-Home												
	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 [13]	59.37	76.58	87.48	69.86	71.11	81.66	73.72	56.30	86.07	78.68	59.22	78.59	73.22
OSBP [35]	47.75	60.90	76.78	59.23	61.58	74.33	61.67	44.50	79.31	70.59	54.95	75.18	63.90
UAN w/o d	61.60	81.86	87.67	74.52	73.59	84.88	73.65	57.37	86.61	81.58	62.15	79.14	75.39
UAN w/o y	56.63	77.51	87.61	71.96	69.08	83.18	71.40	56.10	84.24	79.27	60.59	78.35	72.91
UAN	**63.00**	**82.83**	**87.85**	**76.88**	**78.70**	**85.36**	**78.22**	**58.59**	**86.80**	**83.37**	**63.17**	**79.43**	**77.02**

Experiment Result2

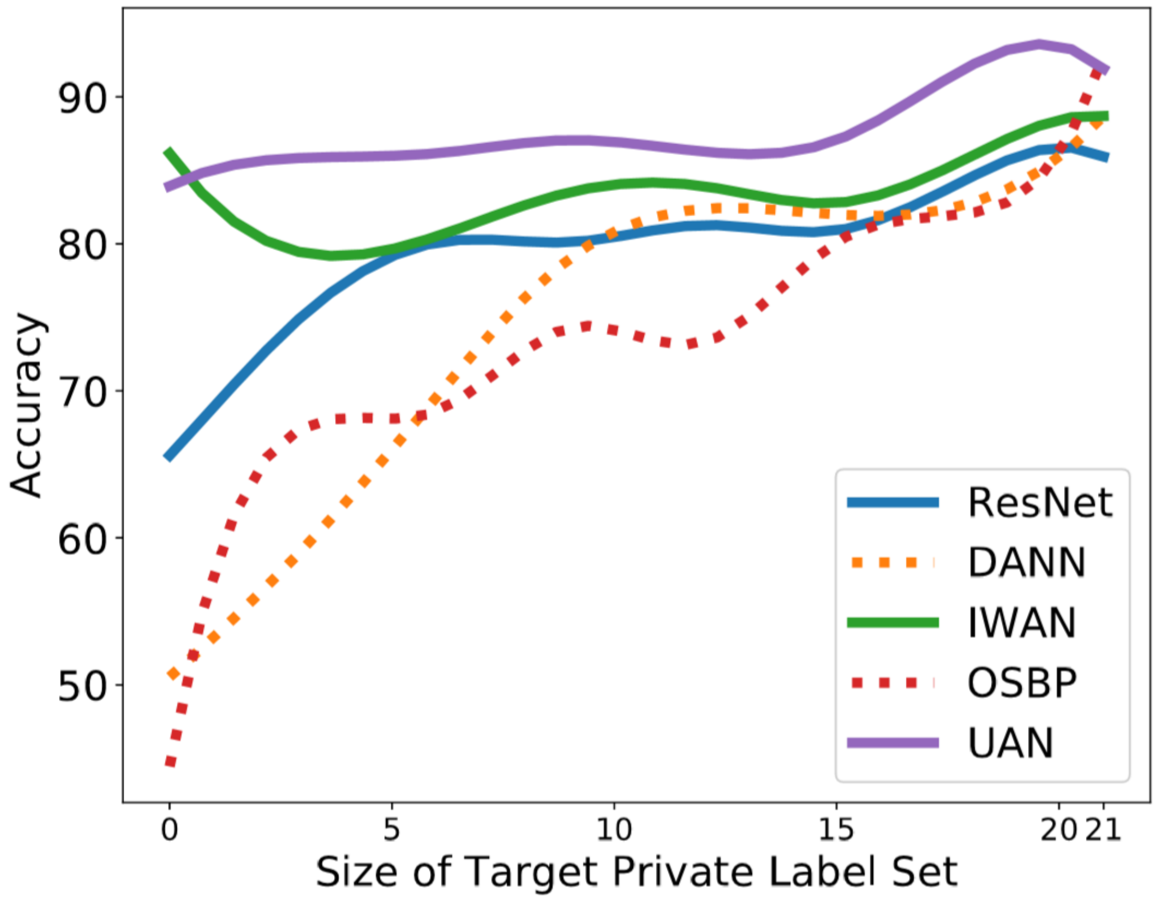
南京航空航天大学
Nanjing University of Aeronautics and Astronautics

Table 2. Average class accuracy (%) on **Office-31** ($\xi = 0.32$), **ImageNet-Caltech** ($\xi = 0.07$) and **VisDA2017** ($\xi = 0.50$) (ResNet)

Method	Office-31							ImageNet-Caltech		VisDA
	A → W	D → W	W → D	A → D	D → A	W → A	Avg	I → C	C → I	
ResNet [13]	75.94	89.60	90.91	80.45	78.83	81.42	82.86	70.28	65.14	52.80
DANN [6]	80.65	80.94	88.07	82.67	74.82	83.54	81.78	71.37	66.54	52.94
RTN [23]	85.70	87.80	88.91	82.69	74.64	83.26	84.18	71.94	66.15	53.92
IWAN [45]	85.25	90.09	90.00	84.27	84.22	86.25	86.68	72.19	66.48	58.72
PADA [45]	85.37	79.26	90.91	81.68	55.32	82.61	79.19	65.47	58.73	44.98
ATI [28]	79.38	92.60	90.08	84.40	78.85	81.57	84.48	71.59	67.36	54.81
OSBP [35]	66.13	73.57	85.62	72.92	47.35	60.48	67.68	62.08	55.48	30.26
UAN	**85.62**	**94.77**	**97.99**	**86.50**	**85.45**	**85.12**	**89.24**	**75.28**	**70.17**	**60.83**

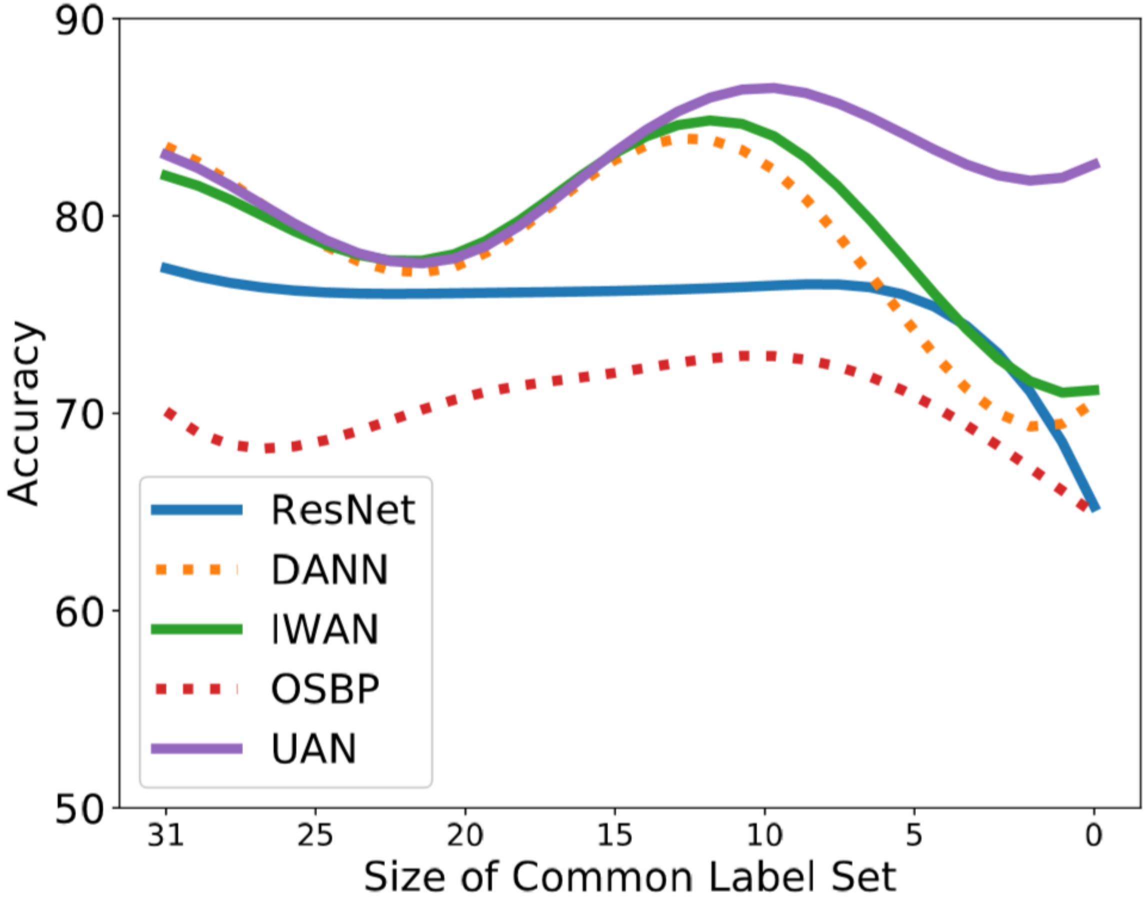
Experiment Result3

Vary size of $\overline{\mathcal{C}}_s$ and $\overline{\mathcal{C}}_t$



(a) Accuracy w.r.t. $|\overline{\mathcal{C}}_t|$

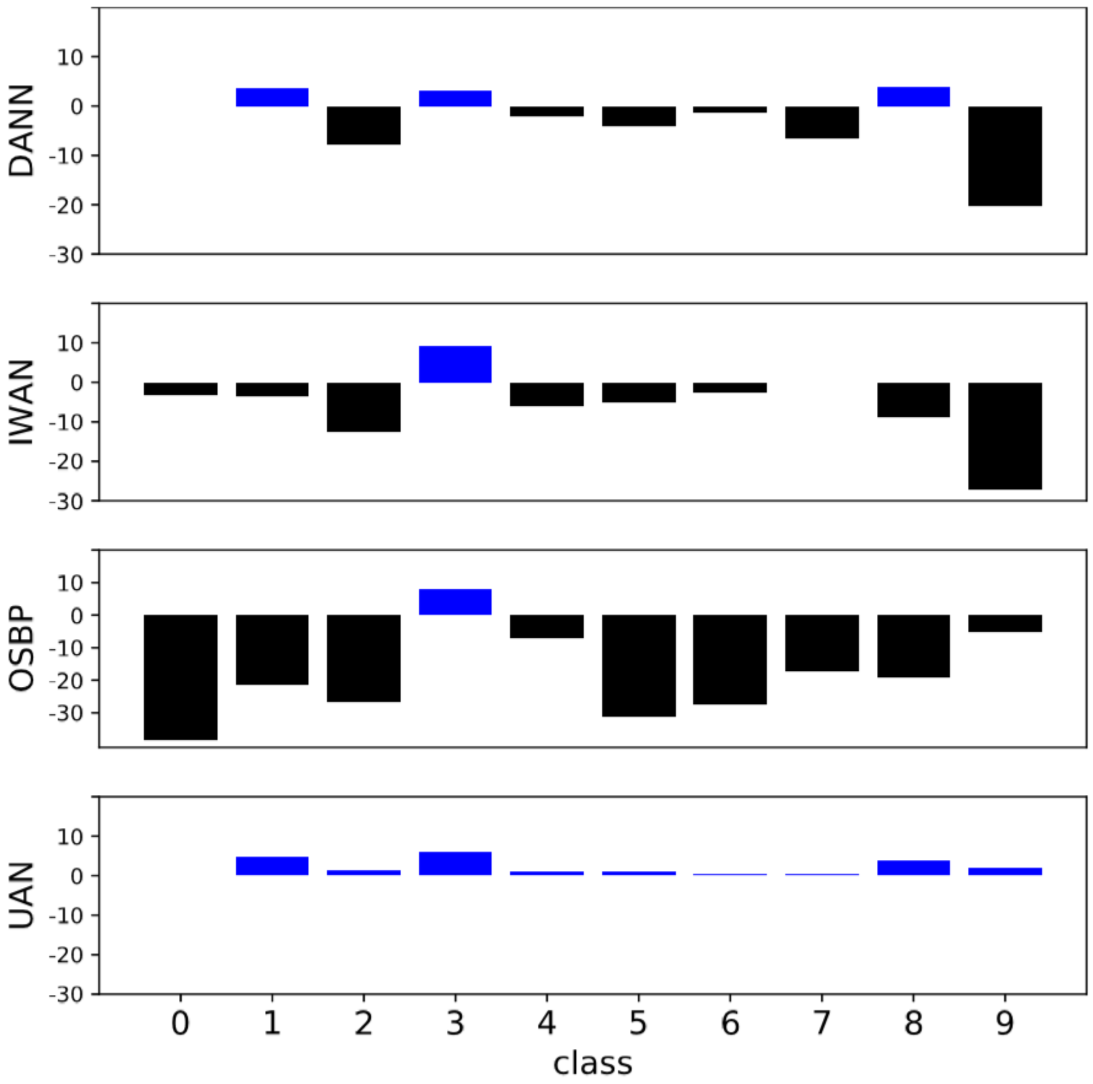
Vary size of common label set $\mathcal{C}$



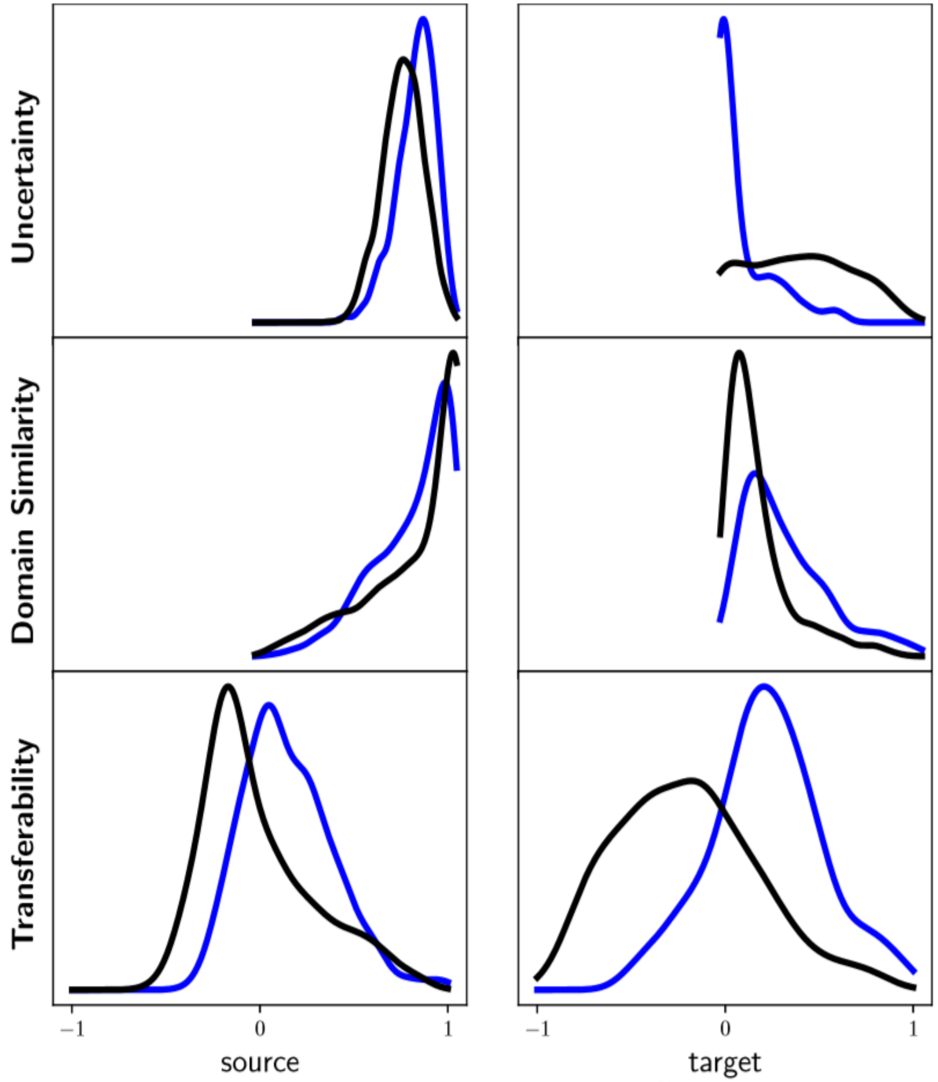
(b) Accuracy w.r.t. $|\mathcal{C}|$

Experiment Result4

Negative Transfer in UAN



Hypotheses Quality
(blue for common and black for private)



THANKS