ICLR 2025

Proxy Denoising for Source-Free Domain Adaptation

Song Tang[1,2], Wenxin Su[1], Yan Gan[1], Mao Ye[*4], Jianwei Zhang[*5] and Xiatian Zhu[*5]

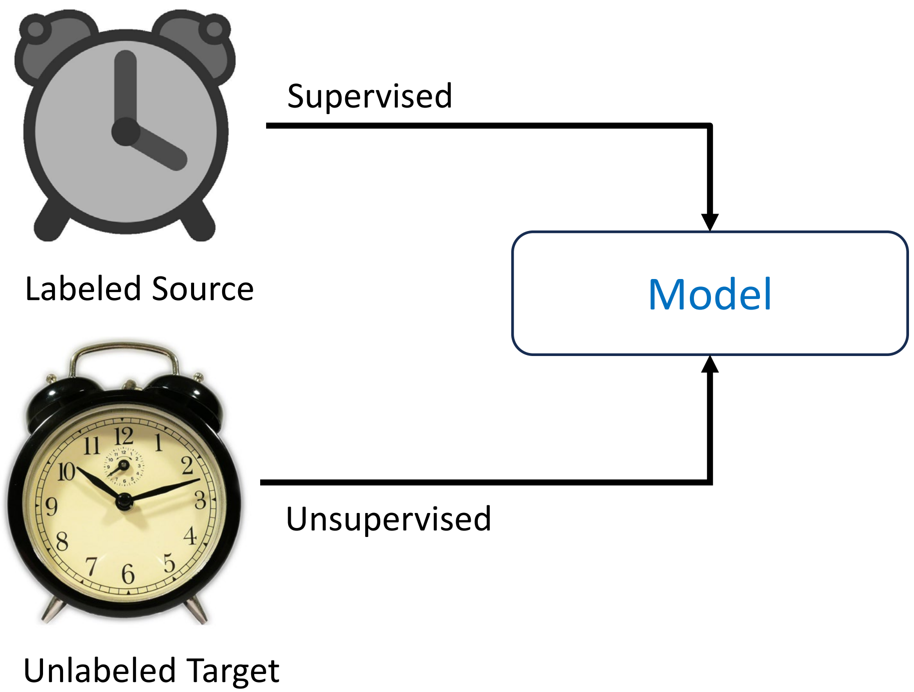
[1]University of Shanghai for Science and Technology [2]Universität Hamburg [3]Chongqing University
[4]University of Electronic Science and Technology of China [5]University of Surrey

* Indicates corresponding author

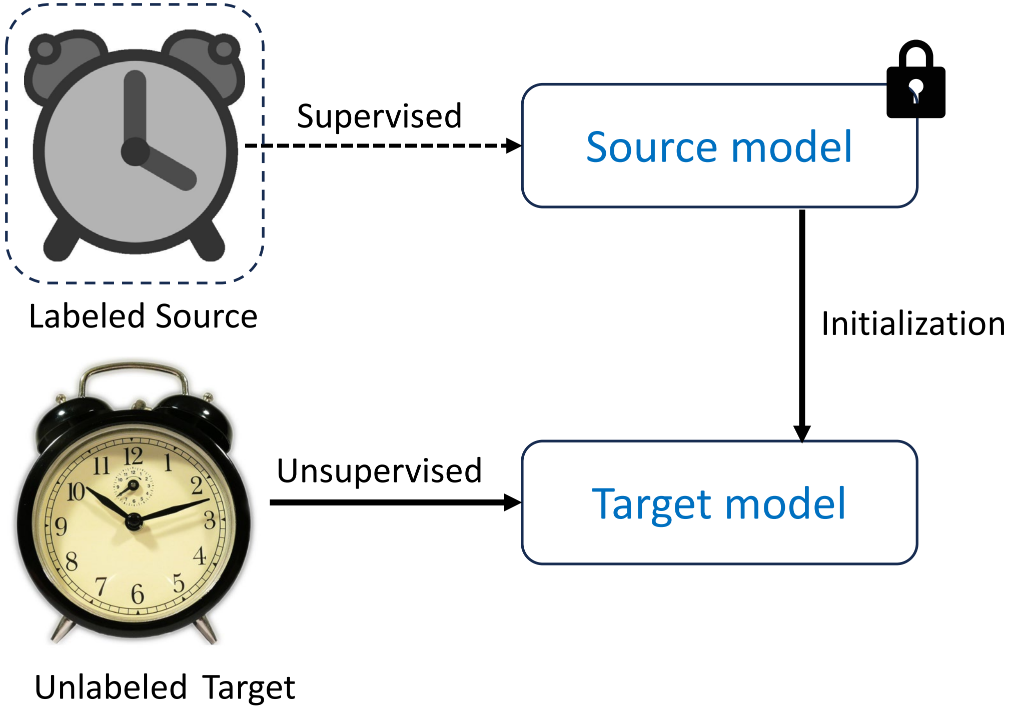
Outlines

Source-Free Domain Adaptation

Unsupervised Domain Adaptation (UDA)



Source-Free Domain Adaptation (SFDA)



- **Condition**: Source model & Unlabeled target data
- **Goal**: Transfer a pretrained source model to the target domain using only unlabeled target data

Challenges

- ☐ **Conventional methods rely on error-prone pseudo-labeling or auxiliary supervision**
  - ➢ Constructing a pseudo source domain to leverage established UDA methods
  - ➢ Mining extra supervision from the source model

The success of Visual-language (ViL) pre-trained model led light on SFDA

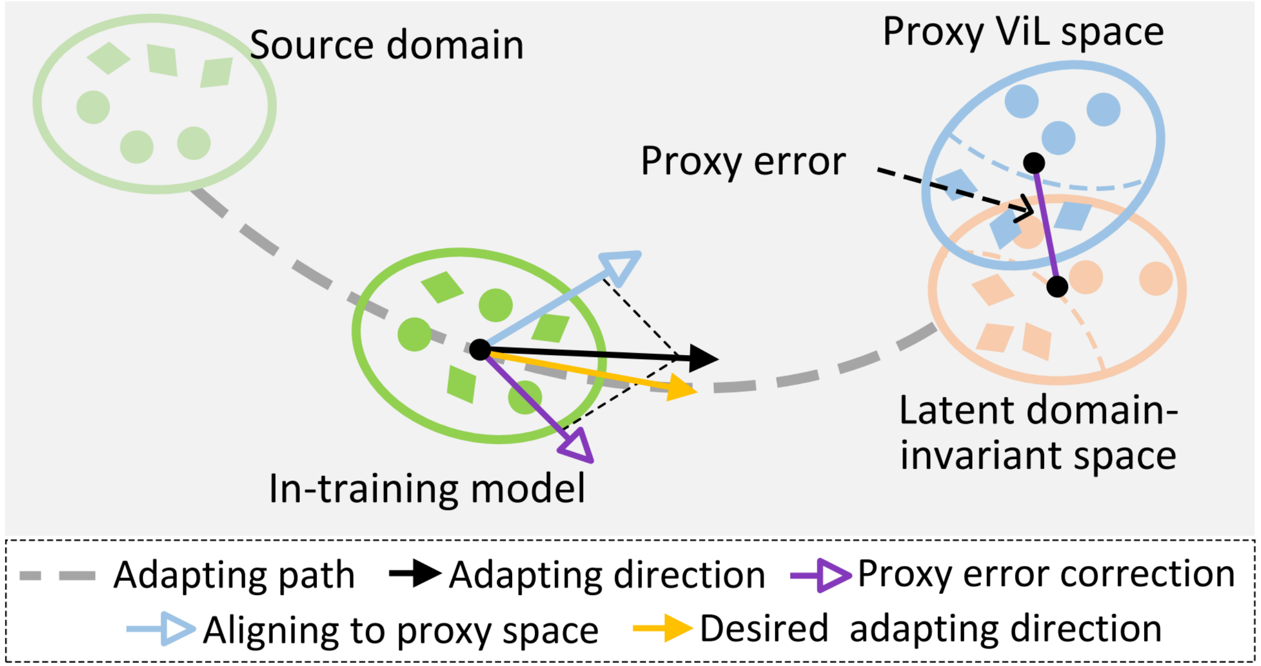
- ☐ **ViL's supervision could be noisy and inaccurate at an unknown rate, introducing additional negative effects during adaption.**

Motivation

Considering the ViL model/space as **a noisy proxy of the latent domain-invariant space**, with a need to be denoised.

☐ Key point

- Exploit the **dynamics of domain adaptation process**, starting at the source model space and terminating presumably in the latent domain invariant space.



ICLR 2025

Proxy Denoising Theory

Modeling **the effect dynamics of proxy error**, i.e., proxy's divergence against the domain-invariant space, in the adaptation process.

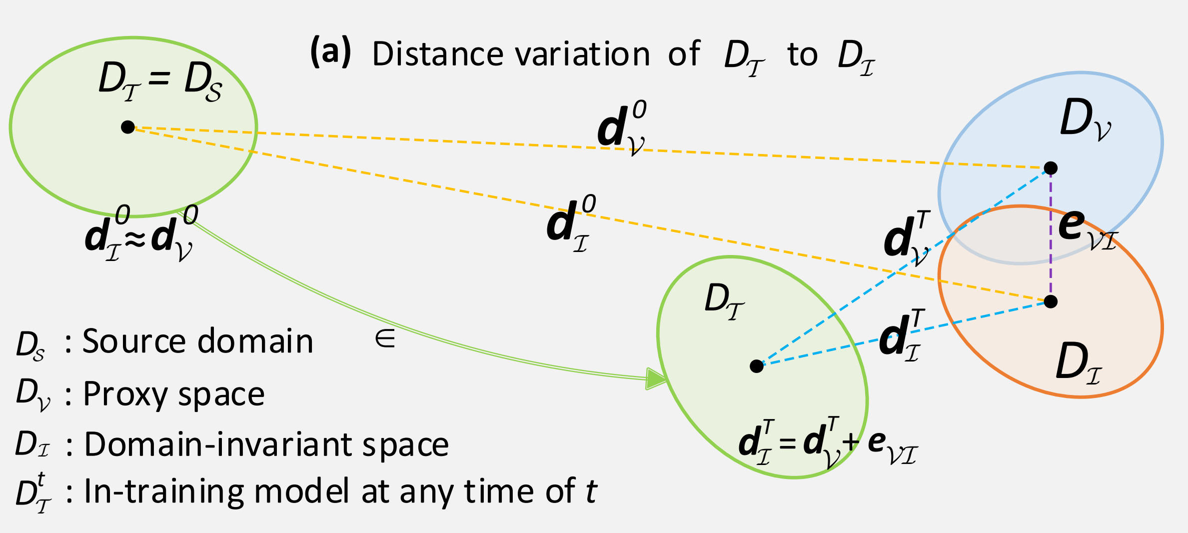
- **Distance variation of intermediate space $D_{\mathcal{T}}^{t}$ (presented by in-training model) to domain-invariant space $D_{\mathcal{I}}$**

Case-1: When $D_{\mathcal{T}}^{t}$ is way far from $D_{\mathcal{V}}$, the beginning of adaptation (t=0)

$$\boldsymbol{d}_{\mathcal{I}}^{0} \approx \boldsymbol{d}_{\mathcal{V}}^{0} \gg \boldsymbol{e}_{\mathcal{V}\mathcal{I}}$$

Case-2: When $D_{\mathcal{T}}^{t}$ approaches $D_{\mathcal{V}}$, the later phase in the adaptation (t=U >> 0)

$$\boldsymbol{d}_{\mathcal{I}}^{U} = \boldsymbol{d}_{\mathcal{V}}^{U} + \boldsymbol{e}_{\mathcal{V}\mathcal{I}}$$



$$\eta_t = \frac{|\boldsymbol{d}_{\mathcal{I}}^{t}|}{|\boldsymbol{d}_{\mathcal{V}}^{t}|} = \frac{|\boldsymbol{d}_{\mathcal{V}}^{t} + \boldsymbol{e}_{\mathcal{V}\mathcal{I}}|}{|\boldsymbol{d}_{\mathcal{V}}^{t}|} \leq \frac{|\boldsymbol{d}_{\mathcal{V}}^{t}| + |\boldsymbol{e}_{\mathcal{V}\mathcal{I}}|}{|\boldsymbol{d}_{\mathcal{V}}^{t}|} = 1 + \frac{|\boldsymbol{e}_{\mathcal{V}\mathcal{I}}|}{|\boldsymbol{d}_{\mathcal{V}}^{t}|}$$

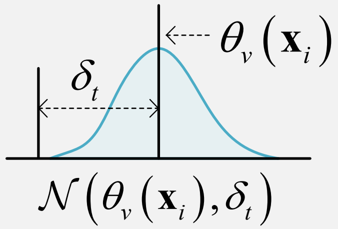
The effect of proxy error is gradually increasing

Proxy Denoising Theory

The effect of proxy error is gradually increasing

In probabilistic view

(b) Confidence variation of ViL prediction



Prediction with confidence at *t*

Treat the ViL predictions that approximate a Gaussian distribution $\mathcal{N}\left(\theta_v\left(x_i\right), \delta_t\right)$:

- Mean $\theta_v\left(x_i\right)$ is the ViL prediction
- Variance $\delta_t \propto \eta_t$

Difficulty

Given that the proxy error is unknown, we cannot formulate these dynamics explicitly

Our solution

$$\mathcal{N}\left(\theta_v\left(x_i\right), \delta_t\right) \Longrightarrow P\left(G_{P(\mathcal{V})} = True, t\right) P\left(\mathcal{V}\right)$$

- $P(\mathcal{V})$ is the probability distribution of the proxy space $D_{\mathcal{V}}$
- $G_{P(\mathcal{V})}$ is a random event that the sampling result (i.e., a ViL prediction) from $P(\mathcal{V})$ is confident.
- $P\left(G_{P(\mathcal{V})} = True, t\right)$ is proxy confidence, indicating the probability of the event $G_{P(\mathcal{V})}$ being true at a time t.

Proxy Denoising Theory

Theorem 1 *We note that the source domain ($D_{\mathcal{S}}$), the domain-invariant space ($D_{\mathcal{I}}$), the proxy space ($D_{\mathcal{V}}$) and the in-training model ($D_{\mathcal{T}}^t$) follow the probability distributions $P(\mathcal{S})$, $P(\mathcal{I})$, $P(\mathcal{V})$ and $P(\mathcal{T}^t)$, respectively, where $\mathcal{S}$, $\mathcal{I}$, $\mathcal{V}$ and $\mathcal{T}^t$ are corresponding random variables. With our proxy alignment idea (see* `Sec. 3.1`*), the proxy confidence can be expressed as:*

$$P\left(G_{P(\mathcal{V})} = True, t\right) \propto \frac{P(\mathcal{T}^t)}{P(\mathcal{S})}$$

The effect of ViL prediction errors on domain adaptation can be approximately estimated by contrasting the distributions of the source model and the current in-training model

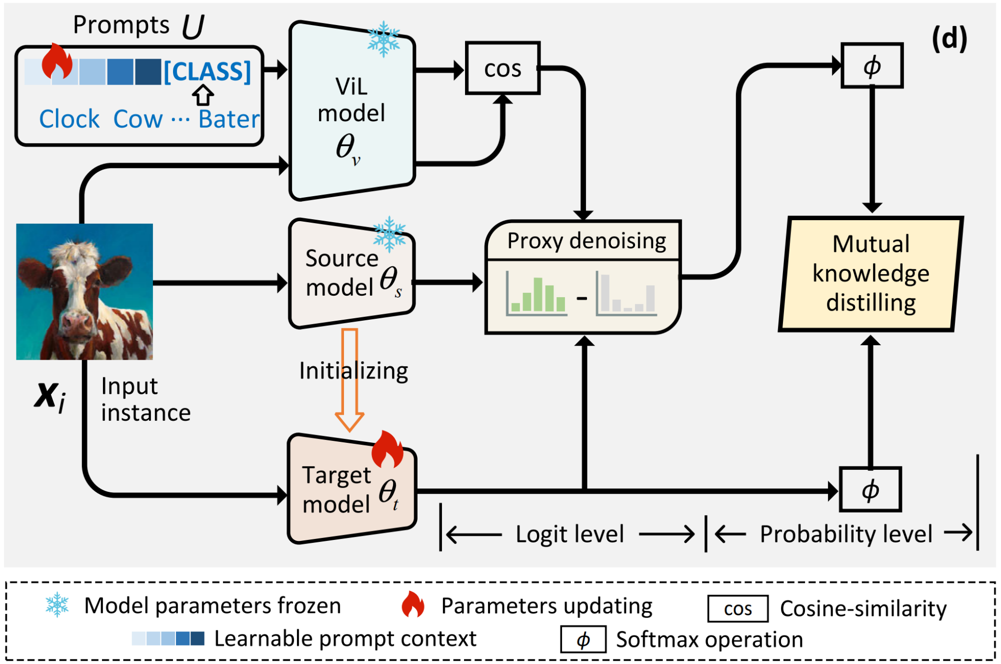
Capitalizing on Corrected Proxy

☐ Design of proxy denoising mechanism

$$\mathcal{N}\left(\theta_v\left(x_i\right), \delta_t\right) \Longrightarrow P\left(G_{P(\mathcal{V})} = True, t\right) P\left(\mathcal{V}\right) \quad + \text{ Theorem 1}$$

$$\log\left(\frac{P(\mathcal{T}^t)}{P(\mathcal{S})} P(\mathcal{V})\right) = \log P(\mathcal{V}) - \left[\log P(\mathcal{S}) - \log P(\mathcal{T}^t)\right]$$

$$\boldsymbol{p}_i' = \mathrm{softmax}\left(\theta_v\left(\boldsymbol{x}_i, \boldsymbol{v}\right) - \omega\left[\theta_s\left(\boldsymbol{x}_i\right) - \theta_t\left(\boldsymbol{x}_i\right)\right]\right)$$



☐ Mutual knowledge distilling

$$L_{\text{ProDe}} = \min_{\theta_t, \boldsymbol{v}} \alpha \overbrace{\left(-\mathbb{E}_{\boldsymbol{x}_i \in \mathcal{X}_t} \mathbf{MI}\left(\boldsymbol{p}_i', \boldsymbol{p}_i\right) + \gamma \sum_{c=1}^{C} \bar{q}_c \log \bar{q}_c\right)}^{L_{\text{Apt}}} \overbrace{-\beta\, \mathbb{E}_{\boldsymbol{x}_i \in \mathcal{X}_t} \sum_{c=1}^{C} \mathbb{1}\left[c = y_i'\right] \log p_{i,c}}^{L_{\text{Ref}}},$$

Experiment

☐ Vanilla Closed-set SFDA

Table 1: Closed-set SFDA results (%) on **Office-31**. **SF** means source-free.

Method	Venue	SF	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Source	–	–	79.1	76.6	59.9	95.5	61.4	98.8	78.6
SHOT	ICML20	✓	93.7	91.1	74.2	98.2	74.6	**100.**	88.6
NRC	NIPS21	✓	96.0	90.8	75.3	99.0	75.0	**100.**	89.4
GKD	IROS21	✓	94.6	91.6	75.1	98.7	75.1	**100.**	89.2
HCL	NIPS21	✓	94.7	92.5	75.9	98.2	77.7	**100.**	89.8
AaD	NIPS22	✓	96.4	92.1	75.0	**99.1**	76.5	**100.**	89.9
AdaCon	CVPR22	✓	87.7	83.1	73.7	91.3	77.6	72.8	81.0
CoWA	ICML22	✓	94.4	95.2	76.2	98.5	77.6	99.8	90.3
ELR	ICLR23	✓	93.8	93.3	76.2	98.0	76.9	**100.**	89.6
PLUE	CVPR23	✓	89.2	88.4	72.8	97.1	69.6	97.9	85.8
CPD	PR24	✓	96.6	94.2	77.3	98.2	78.3	**100.**	90.8
TPDS	IJCV24	✓	97.1	94.5	75.7	98.7	75.5	99.8	90.2
DIFO-R	CVPR24	✓	93.6	92.1	78.5	95.7	78.8	97.0	89.3
DIFO-V	CVPR24	✓	**97.2**	95.5	83.0	**97.2**	**83.2**	98.8	92.5
ProDe-R	–	✓	94.4	92.1	79.8	95.6	79.0	98.6	89.9
ProDe-V	–	✓	96.8	**96.4**	**83.1**	97.0	82.5	99.8	**92.6**

Table 2: Closed-set SFDA results (%) on **Office-Home** and **VisDA**. **SF** means source-free. The full results on **VisDA** are provided in Appendix E.1.

Method	Venue	SF	**Office-Home**													**VisDA**
			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.	Sy→Re
Source	–	–	43.7	67.0	73.9	49.9	60.1	62.5	51.7	40.9	72.6	64.2	46.3	78.1	59.2	49.2
SHOT	ICML20	✓	56.7	77.9	80.6	68.0	78.0	79.4	67.9	54.5	82.3	74.2	58.6	84.5	71.9	82.7
NRC	NIPS21	✓	57.7	80.3	82.0	68.1	79.8	78.6	65.3	56.4	83.0	71.0	58.6	85.6	72.2	85.9
GKD	IROS21	✓	56.5	78.2	81.8	68.7	78.9	79.1	67.6	54.8	82.6	74.4	58.5	84.8	72.2	83.0
AaD	NIPS22	✓	59.3	79.3	82.1	68.9	79.8	79.5	67.2	57.4	83.1	72.1	58.5	85.4	72.7	88.0
AdaCon	CVPR22	✓	47.2	75.1	75.5	60.7	73.3	73.2	60.2	45.2	76.6	65.6	48.3	79.1	65.0	86.8
CoWA	ICML22	✓	56.9	78.4	81.0	69.1	80.0	79.9	67.7	57.2	82.4	72.8	60.5	84.5	72.5	86.9
ELR	ICLR23	✓	58.4	78.7	81.5	69.2	79.5	79.3	66.3	58.0	82.6	73.4	59.8	85.1	72.6	85.8
PLUE	CVPR23	✓	49.1	73.5	78.2	62.9	73.5	74.5	62.2	48.3	78.6	68.6	51.8	81.5	66.9	88.3
CPD	PR24	✓	59.1	79.0	82.4	68.5	79.7	79.5	67.9	57.9	82.8	73.8	61.2	84.6	73.0	85.8
TPDS	IJCV24	✓	59.3	80.3	82.1	70.6	79.4	80.9	69.8	56.8	82.1	74.5	61.2	85.3	73.5	87.6
DAPL-R	TNNLS23	✗	54.1	84.3	84.8	74.4	83.7	85.0	74.5	54.6	84.8	75.2	54.7	83.8	74.5	86.9
PADCLIP-R	ICCV23	✗	57.5	84.0	83.8	77.8	85.5	84.7	76.3	59.2	85.4	78.1	60.2	86.7	76.6	88.5
ADCLIP-R	ICCVW23	✗	55.4	85.2	85.6	76.1	85.8	86.2	76.7	56.1	85.4	76.8	56.1	85.5	75.9	87.7
PDA-R	AAAI24	✗	55.4	85.1	85.8	75.2	85.2	85.2	74.2	55.2	85.8	74.7	55.8	86.3	75.3	86.4
DAMP-R	CVPR24	✗	59.7	88.5	86.8	76.6	88.9	87.0	76.3	59.6	87.1	77.0	61.0	89.9	78.2	88.4
DIFO-R	CVPR24	✓	62.6	87.5	87.1	79.5	87.9	87.4	78.3	63.4	88.1	80.0	63.3	87.7	79.4	88.6
DIFO-V	CVPR24	✓	70.6	90.6	88.8	82.5	90.6	88.8	80.9	70.1	88.9	**83.4**	70.5	91.2	83.1	90.3
ProDe-R	–	✓	64.0	90.0	88.3	81.1	90.1	88.6	79.8	65.4	89.0	80.9	65.5	90.2	81.1	88.7
ProDe-V	–	✓	**72.7**	**92.3**	**90.5**	**82.5**	**91.5**	**90.7**	**82.5**	**72.5**	**90.8**	83.0	**72.6**	**92.2**	**84.5**	**91.0**

Table 3: Closed-set SFDA results (%) on **DomainNet-126**. **SF** means source-free.

Method	Venue	SF	C→P	C→R	C→S	P→C	P→R	P→S	R→C	R→P	R→S	S→C	S→P	S→R	Avg.
Source	–	–	44.6	59.8	47.5	53.3	75.3	46.2	55.3	62.7	46.4	55.1	50.7	59.5	54.7
SHOT	ICML20	✓	63.5	78.2	59.5	67.9	81.3	61.7	67.7	67.6	57.8	70.2	64.0	78.0	68.1
GKD	IROS21	✓	61.4	77.4	60.3	69.6	81.4	63.2	68.3	68.4	59.5	71.5	65.2	77.6	68.7
NRC	NIPS21	✓	62.6	77.1	58.3	62.9	81.3	60.7	64.7	69.4	58.7	69.4	65.8	78.7	67.5
AdaCon	CVPR22	✓	60.8	74.8	55.9	62.2	78.3	58.2	63.1	68.1	55.6	67.1	66.0	75.4	65.4
CoWA	ICML22	✓	64.6	80.6	60.6	66.2	79.8	60.8	69.0	67.2	60.0	69.0	65.8	79.9	68.6
PLUE	CVPR23	✓	59.8	74.0	56.0	61.6	78.5	57.9	61.6	65.9	53.8	67.5	64.3	76.0	64.7
TPDS	IJCV24	✓	62.9	77.1	59.8	65.6	79.0	61.5	66.4	67.0	58.2	68.6	64.3	75.3	67.1
DAPL-R	TNNLS23	✗	72.4	87.6	65.9	72.7	87.6	65.6	73.2	72.4	66.2	73.8	72.9	87.8	74.8
ADCLIP-R	ICCVW23	✗	71.7	88.1	66.0	73.2	86.9	65.2	73.6	73.0	68.4	72.3	74.2	89.3	75.2
DAMP-R	CVPR24	✗	76.7	88.5	71.7	74.2	88.7	70.8	74.4	75.7	70.5	74.9	76.1	88.2	77.5
DIFO-R	CVPR24	✓	73.8	89.0	69.4	74.0	88.7	70.1	74.8	74.6	69.6	74.7	74.3	88.0	76.7
DIFO-V	CVPR24	✓	76.6	87.2	74.9	80.0	87.4	75.6	80.8	77.3	75.5	80.5	76.7	87.3	80.0
ProDe-R	–	✓	79.3	91.0	75.3	80.0	90.9	75.6	80.4	78.9	75.4	80.4	79.2	91.0	81.5
ProDe-V	–	✓	**83.2**	**92.4**	**79.0**	**85.0**	**92.3**	**79.3**	**85.5**	**83.1**	**79.1**	**85.5**	**83.4**	**92.4**	**85.0**

Experiment

- **Other SFDA settings**
  - Comparison with CLIP
  - Partial-set SFDA
  - Open-set SFDA
  - Generalized SFDA
  - Source-Free Multi-Target DA
  - Source-Free Multi-Source DA
  - Test-Time Adaptation

Table 4: Comparison results with CLIP (%). Appendix E.1 presents the full results.

Method	Office-31	Office-Home	VisDA	DomainNet-126
CLIP-R	71.4	72.1	83.7	72.7
ProDe-R	**89.9**	**81.1**	**88.7**	**81.5**
CLIP-V	79.8	76.1	82.9	76.3
ProDe-V	**92.6**	**84.5**	**91.0**	**85.0**

Table 5: Partial-set and open-set results (%) on **Office-Home**. Appendix E.1 presents the full results.

Partial-set	Venue	Avg.	Open-set	Venue	Avg.
Source	–	62.8	Source	–	46.6
SHOT	ICML20	79.3	SHOT	ICML20	72.8
HCL	NIPS21	79.6	HCL	NIPS21	72.6
CoWA	ICML22	83.2	CoWA	ICML22	73.2
AaD	NIPS22	79.7	AaD	NIPS22	71.8
CRS	CVPR23	80.6	CRS	CVPR23	73.2
DIFO-V	CVPR24	84.1	DIFO-V	CVPR24	75.9
ProDe-V	–	**84.2**	**ProDe**-V	–	**82.6**

Table 6: Generalized SFDA results (%) on **Office-Home**. S, T are the results of the adapted target model on the source and target domains, i.e., Acc_s, Acc_t, respectively; **WAD** means With Anti-forgetting Design. Appendix E.1 presents the full results.

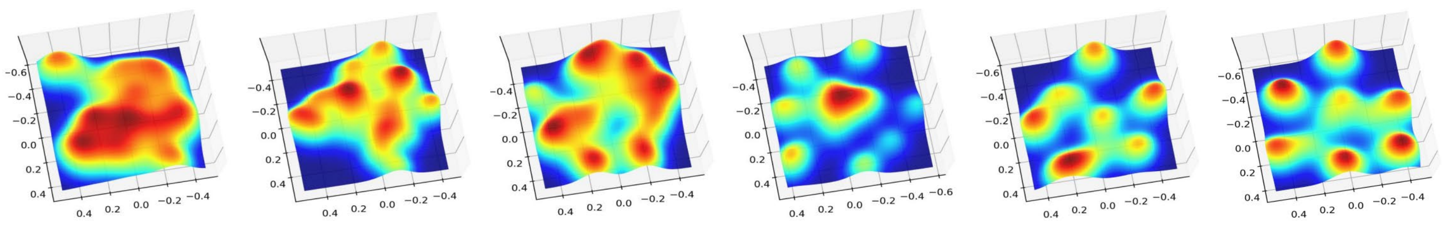
Method	Venue	WAD	Avg. S (98.1-S)	Avg. T	Avg. H
Source	–	✗	98.1	59.2	73.1
SHOT	ICML20	✗	84.2 (13.9)	71.8	77.5
GKD	IROS21	✗	86.8 (11.3)	72.5	79.0
NRC	NIPS21	✗	91.3 (6.8)	72.3	80.7
AdaCon	CVPR22	✗	88.2 (9.9)	65.0	74.8
CoWA	ICML22	✗	91.8 (6.3)	72.4	81.0
PLUE	CVPR23	✗	**96.3** (1.8)	66.9	79.0
TPDS	IJCV24	✗	83.8 (14.3)	73.5	78.3
GDA	ICCV21	✓	80.0 (18.1)	70.2	74.4
PSAT-ViT	TMM24	✓	86.4 (11.7)	83.6	**85.0**
DIFO-R	CVPR24	✗	78.3 (19.8)	79.4	78.8
DIFO-V	CVPR24	✗	78.0 (20.1)	83.1	80.5
ProDe-R	–	✗	84.9 (13.2)	81.1	82.9
ProDe-V	–	✗	85.1 (13.0)	**84.5**	84.8

Table 7: SF-MTDA, SF-MSDA and TTA results (%) on **Office-Home**. The full results of TTA are provided in Appendix E.1.

	Model	Venue	Ar→	Cl→	Pr→	Rw→	Avg.
SF-MTDA	CoNMix	WACV23	75.6	81.4	71.4	73.4	75.4
	ProDe-V	–	**83.3**	**89.2**	**80.9**	**81.2**	**83.6**
	Method	Venue	→Rw	→Pr	→Cl	→Ar	Avg.
SF-MSDA	SHOT-Ens	ICML20	82.9	82.8	59.3	72.2	74.3
	DECISION	CVPR21	83.6	84.4	59.4	74.5	75.5
	ProDe-V-Ens	–	**91.1**	**92.5**	**73.4**	**83.0**	**85.0**

	Method	Venue	Avg.
TTA	Tent	ICLR20	61.7
	T3A	NeurIPS21	63.8
	CoTTA	CVPR22	60.5
	EATA	ICML22	60.7
	SAR	ICLR23	60.3
	ProDe-V	–	**76.5**

Experiment

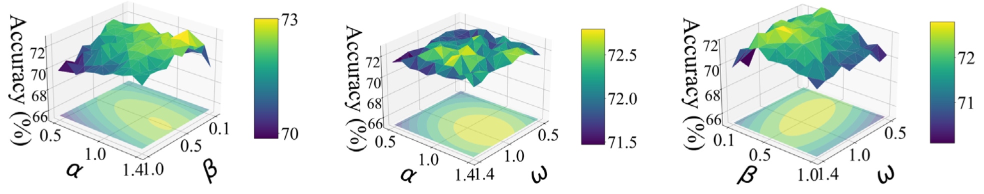


(a) CLIP-V (b) SHOT (c) TPDS (d) DIFO-V (e) ProDe-V (f) Oracle

Figure 3: Feature visualization comparison in 3D density charts.

☐ Model analysis

- Feature distribution
- Training resource demands
- Parameter sensitivity
- Reliance on ViL models



(a) α–β (b) α–ω (c) β–ω

Figure 5: Sensitivity analysis of hyper-parameters α, β and ω.

Table 22: Comparison of training resource demands (per iter.) on Ar→Cl in **Office-Home**.

#	Item / Method	SHOT	AaD	**ProDe**
1	GPU memory consumption↓ (G)	**7.868**	9.622	9.851
2	Training times↓ (s)	**0.407**	0.547	0.491

Table 19: Reliance analysis results (%) on **DomainNet-126** in the Closed-set SFDA setting.

Method	Venue	C→P	C→R	C→S	P→C	P→R	P→S	R→C	R→P	R→S	S→C	S→P	S→R	Avg.
DIFO w/ CLIP	CVPR24	76.6	87.2	74.9	80.0	87.4	75.6	80.8	77.3	75.5	80.5	76.7	87.3	80.0
ProDe w/ CLIP	–	**83.2**	**92.4**	**79.0**	**85.0**	**92.3**	**79.3**	**85.5**	**83.1**	**79.1**	**85.5**	**83.4**	**92.4**	**85.0**
DIFO w/ OpenCLIP	CVPR24	91.2	91.5	79.4	85.2	91.2	79.7	85.7	82.7	80.5	85.9	81.3	91.4	84.6
ProDe w/ OpenCLIP	–	**86.7**	**93.7**	**84.4**	**89.2**	**93.7**	**84.5**	**89.6**	**86.6**	**84.4**	**89.5**	**86.7**	**93.7**	**88.6**

Contribution

- We for the first time investigate the inaccurate predictions of ViL models in the SFDA context.
- We formulate a novel ProDe method that reliably corrects the ViL model's predictions under the guidance of a proxy confidence theory.
- A mutual knowledge distilling regularization is introduced for better capitalizing on refined proxy predictions.

Takeaways

- **New view**: Understanding domain adaptation in a dynamic process is a promising solution；
- **Multi-source challenge**: Efficient strategies that adapts from heterogeneous generic external knowledge；
- **Black box challenge**: Models in the cloud, our proxy denoising may not work well since all details of the model are transparent to us.

Thank you !

Code and data can be accessed at:
https://github.com/tntek/source-free-domain-adaptation