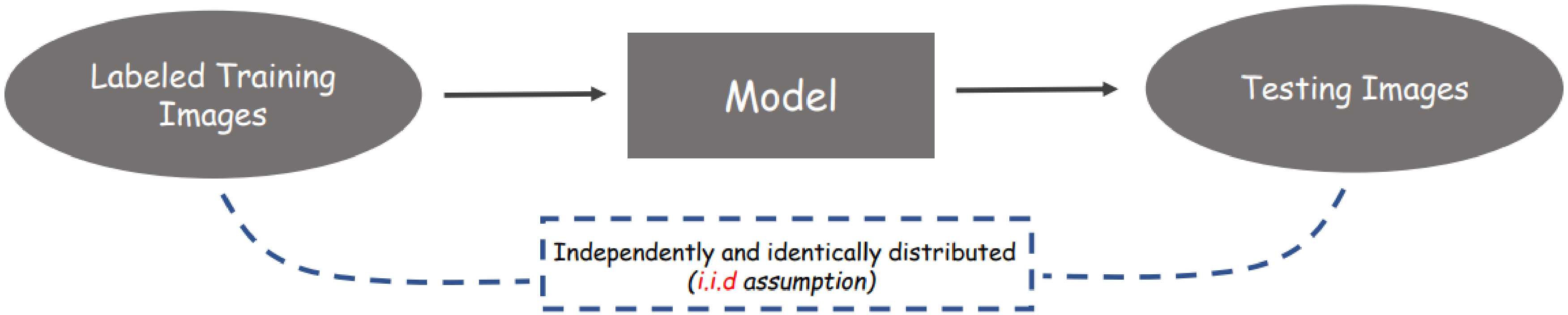
南京航空航天大學
Nanjing University of Aeronautics and Astronautics

# Dirichlet-based Uncertainty Calibration for Active Domain Adaptation

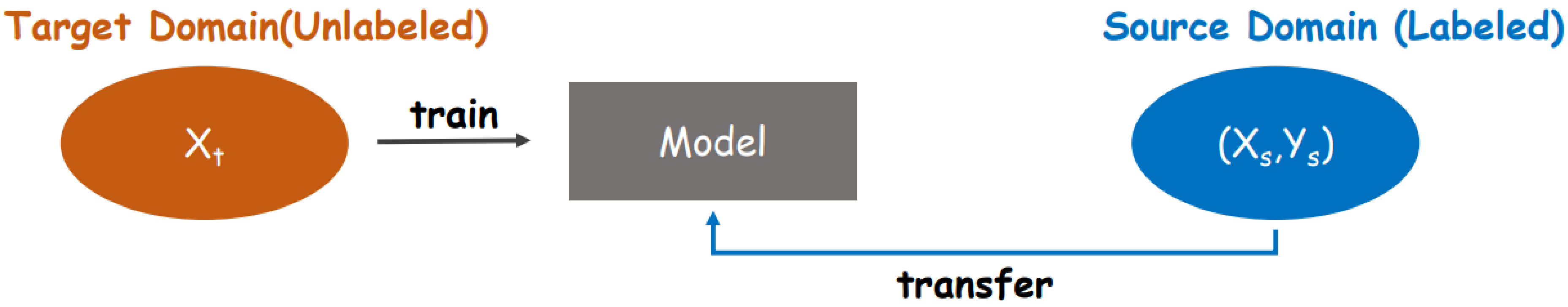
ICLR 2023

# Unsupervised Domain Adaptation

- Supervised Learning




- Unsupervised Domain Adaptation(UDA)




- ◻ Domain shift: $P_s(X,Y) \neq P_t(X,Y)$ or $P_s(x_s) \neq P_t(x_t)$
- ◻ Feature space and label space are consistent: $x_s, x_t \in \mathbb{R}^d \quad y_s, y_t \in L$

# Unsupervised Domain Adaptation

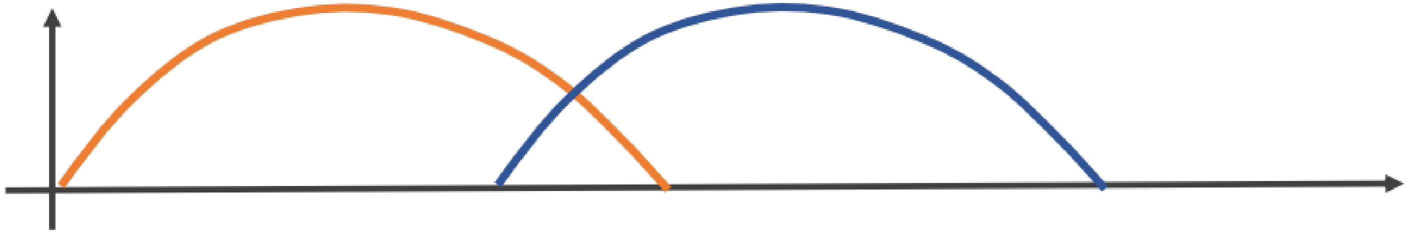
- Unsupervised Domain Adaptation(UDA)

Bike:

Keyboard:

Source domain
with annotations

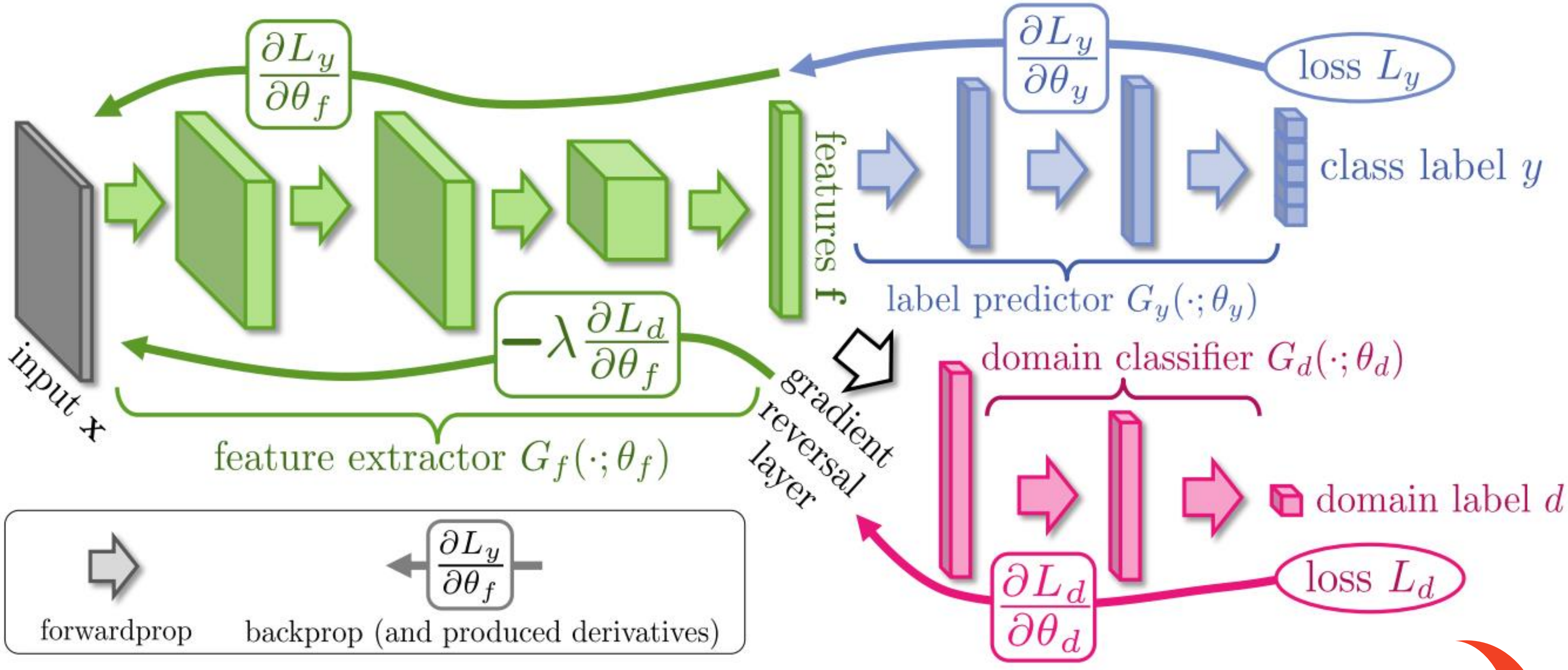
Target domain
without annotations



**Solution: Domain Alignment**

✓ **Goal**: Achieving good performance on the target domain.

# Unsupervised Domain Adaptation

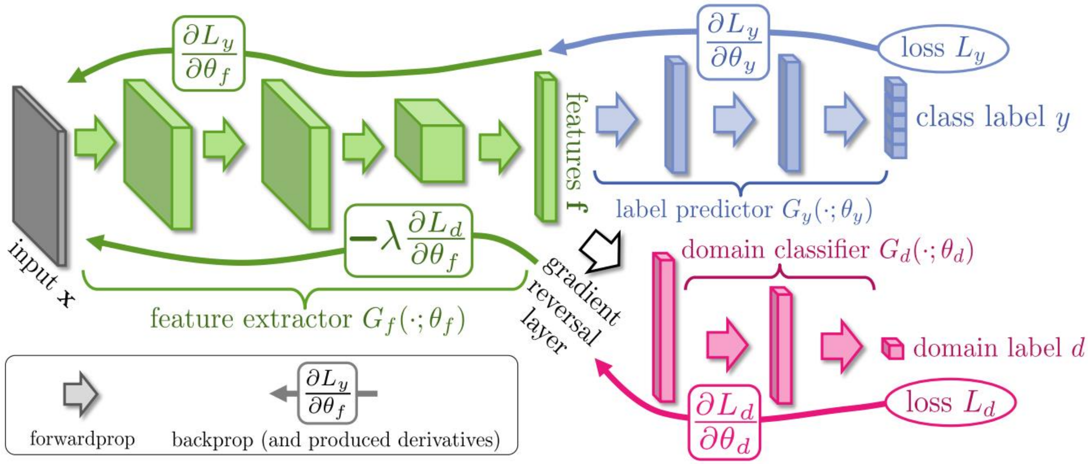



$$\min_{\theta_f,\theta_y} \max_{\theta_d} \mathcal{L}_c(G_y(G_f(x)), y) + \lambda \mathcal{L}_d$$

$$\mathcal{L}_d = \mathbb{E}_{x \sim p_S(x)} \left[ \log G_d(G_f(x)) \right] + \mathbb{E}_{x \sim p_T(x)} \left[ \log(1 - G_d(G_f(x))) \right]$$

# Active Domain Adaptation (ADA)


Goal: Identify more valuable target samples that, once labeled and used for training, improve the model's accuracy and generalization performance significantly.

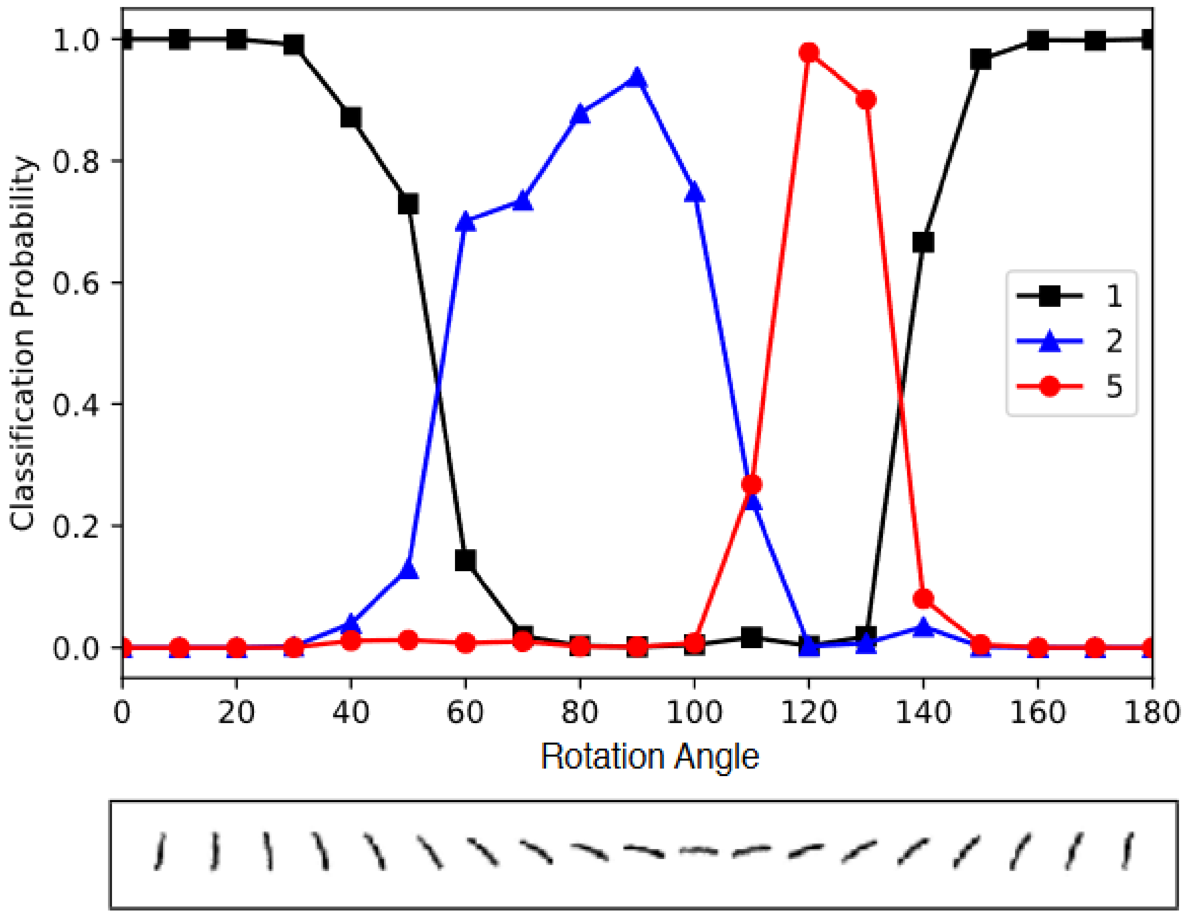



**Algorithm 1** AADA

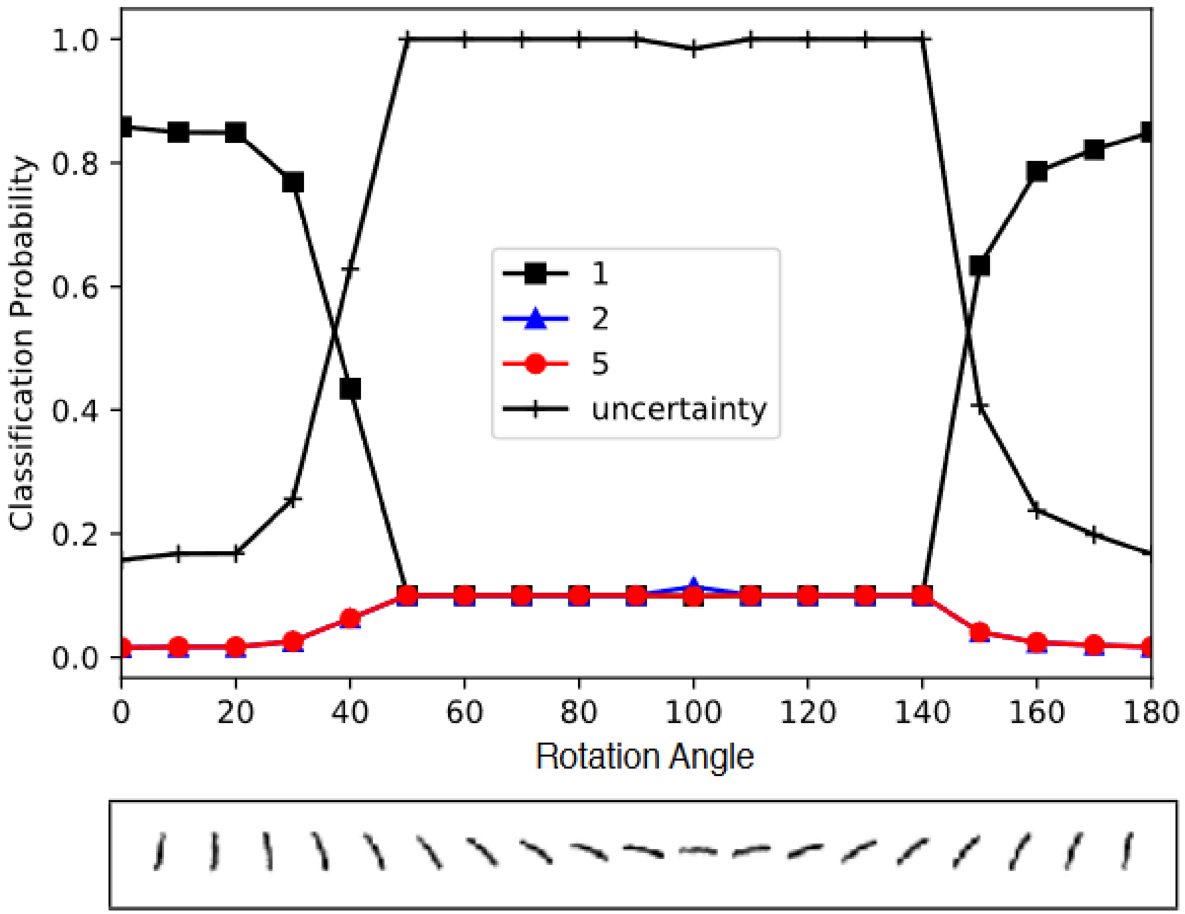
**Input:** labeled source $L_s$; unlabeled target $U_t$;
labeled target $L_t = \emptyset$; budget per round $b$
**Model:** $\mathcal{M}$={$G_f$, $G_y$, $G_d$}; feature extractor $G_f$;
class predictor $G_y$; discriminator $G_d$
Train $\mathcal{M}$ with $(L_s, U_t)$
**for** round $\leftarrow$ 1 to MaxRound **do**
Compute $s(x)$ $\forall x \in U_t$ via (5)
Select a set of $b$ images $z$ from $U_t$ according to $s(z)$
Get labels $y_z$ from oracle
$L_t \leftarrow L_t \cup (z, y_z)$
$U_t \leftarrow U_t \setminus (z, y_z)$
Train $\mathcal{M}$ with $(L_s \cup L_t, U_t)$

# Evidential Model






Softmax-based Model



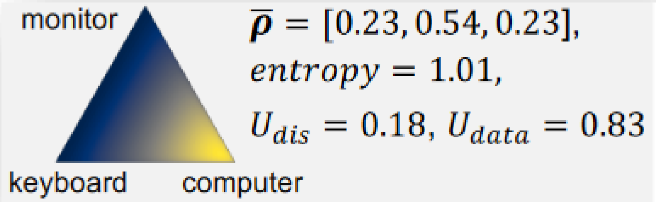

Evidential Model

# Evidential Model

$$p(\boldsymbol{\rho}|\boldsymbol{x_i},\boldsymbol{\theta}) = Dir(\boldsymbol{\rho}|\boldsymbol{\alpha}_i) = \begin{cases} \frac{\Gamma(\sum_{c=1}^{C}\alpha_{ic})}{\prod_{c=1}^{C}\Gamma(\alpha_{ic})}\prod_{c=1}^{C}\rho_c^{\alpha_{ic}-1}, & if\ \boldsymbol{\rho}\in\triangle^C \\ 0, & otherwise \end{cases}, \quad \alpha_{ic} > 0.$$

[1] $\boldsymbol{\rho} = [\rho_1, \rho_2, \cdots, \rho_C]^\top = [P(y=1), P(y=2), \cdots, P(y=C)]^\top$ is a vector of class probabilities.

$$\boldsymbol{\alpha}_i = g(f(\boldsymbol{x}_i,\boldsymbol{\theta})) \qquad \boldsymbol{e}_i = \boldsymbol{\alpha}_i - \mathbf{1} \qquad \triangle^C = \{\boldsymbol{\rho}|\sum_{c=1}^{C}\rho_c = 1 \text{ and } \forall\rho_c, 0\le\rho_c\le 1\}$$




Connection with softmax-based DNNs:

$$P(y=c|\boldsymbol{x}_i,\boldsymbol{\theta}) = \int p(y=c|\boldsymbol{\rho})p(\boldsymbol{\rho}|\boldsymbol{x}_i,\boldsymbol{\theta})d\boldsymbol{\rho} = \frac{\alpha_{ic}}{\sum_{k=1}^{C}\alpha_{ik}} = \frac{g(f_c(\boldsymbol{x}_i,\boldsymbol{\theta}))}{\sum_{k=1}^{C}g(f_k(\boldsymbol{x}_i,\boldsymbol{\theta}))} = \mathbb{E}[Dir(\rho_c|\boldsymbol{\alpha}_i)]$$

$$S = \sum_{i=1}^{K}(e_i+1) \qquad u = \frac{K}{S}$$

Evidential Deep Learning to Quantify Classification Uncertainty

# Evidential Model



$$\underbrace{\mathcal{I}[y,\boldsymbol{\theta}|\boldsymbol{x}^*,\mathcal{D}]}_{Model\ Uncertainty} = \underbrace{\mathcal{H}[\mathbb{E}_{\mathrm{p}(\boldsymbol{\theta}|\mathcal{D})}[\mathrm{P}(y|\boldsymbol{x}^*,\boldsymbol{\theta})]]}_{Total\ Uncertainty} - \underbrace{\mathbb{E}_{\mathrm{p}(\boldsymbol{\theta}|\mathcal{D})}[\mathcal{H}[\mathrm{P}(y|\boldsymbol{x}^*,\boldsymbol{\theta})]]}_{Expected\ Data\ Uncertainty}$$

Predictive Uncertainty Estimation via Prior Networks

Data uncertainty:

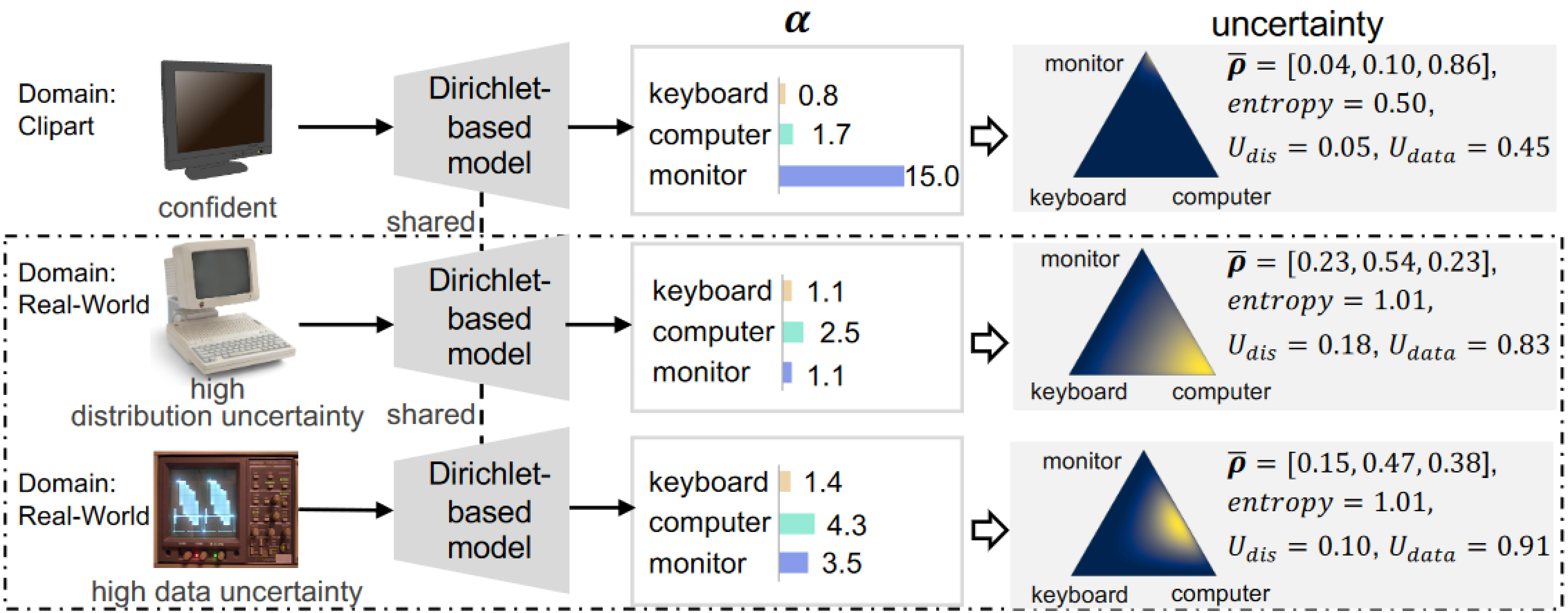
$$U_{data}(\boldsymbol{x}_j,\boldsymbol{\theta}) \triangleq \mathbb{E}_{p(\boldsymbol{\rho}|\boldsymbol{x}_j,\boldsymbol{\theta})}\left[H[P(y|\boldsymbol{\rho})]\right] = \sum_{c=1}^{C}\bar{\rho}_{jc}\left(\psi(\sum_{k=1}^{C}\alpha_{jk}+1)-\psi(\alpha_{jc}+1)\right)$$

Evidential Deep Learning to Quantify Classification Uncertainty

Distribution uncertainty:

$$U_{dis}(\boldsymbol{x}_j,\boldsymbol{\theta}) \triangleq I[y,\boldsymbol{\rho}|\boldsymbol{x}_j,\boldsymbol{\theta}] = \sum_{c=1}^{C}\bar{\rho}_{jc}\left(\psi(\alpha_{jc}+1)-\psi(\sum_{k=1}^{C}\alpha_{jk}+1)\right) - \sum_{c=1}^{C}\bar{\rho}_{jc}\log\bar{\rho}_{jc}$$

# Evidential Model




$\boldsymbol{\alpha}$: parameters of Dirichlet distribution $\bar{\boldsymbol{\rho}}$: expected prediction $U_{dis}$ / $U_{data}$ : distribution / data uncertainty

# Evidential Model Learning

## Negative logarithm of the marginal likelihood

$$\mathcal{L}_{nll} = \frac{1}{n_s}\sum_{\boldsymbol{x}_i \in \mathcal{S}} -\log\left(\int p(y=y_i|\boldsymbol{\rho})p(\boldsymbol{\rho}|\boldsymbol{x}_i,\boldsymbol{\theta})d\boldsymbol{\rho}\right) + \frac{1}{|\mathcal{T}^l|}\sum_{\boldsymbol{x}_j \in \mathcal{T}^l} -\log\left(\int p(y=y_j|\boldsymbol{\rho})p(\boldsymbol{\rho}|\boldsymbol{x}_j,\boldsymbol{\theta})d\boldsymbol{\rho}\right)$$

$$= \frac{1}{n_s}\sum_{\boldsymbol{x}_i \in \mathcal{S}}\sum_{c=1}^{C}\Upsilon_{ic}\left(\log\Big(\sum_{c=1}^{C}\alpha_{ic}\Big) - \log\alpha_{ic}\right) + \frac{1}{|\mathcal{T}^l|}\sum_{\boldsymbol{x}_j \in \mathcal{T}^l}\sum_{c=1}^{C}\Upsilon_{jc}\left(\log\Big(\sum_{c=1}^{C}\alpha_{jc}\Big) - \log\alpha_{jc}\right)$$

where $\Upsilon_{ic}/\Upsilon_{jc}$ is the $c$-th element of the one-hot label vector $\boldsymbol{\Upsilon}_i/\boldsymbol{\Upsilon}_j$ of sample $\boldsymbol{x}_i/\boldsymbol{x}_j$.

## KL divergence

$$\mathcal{L}_{kl} = \frac{1}{C \cdot n_s}\sum_{\boldsymbol{x}_i \in \mathcal{S}} KL\left[Dir(\boldsymbol{\rho}|\tilde{\boldsymbol{\alpha}}_i)\|Dir(\boldsymbol{\rho}|\mathbf{1})\right] + \frac{1}{C \cdot |\mathcal{T}^l|}\sum_{\boldsymbol{x}_j \in \mathcal{T}^l} KL\left[Dir(\boldsymbol{\rho}|\tilde{\boldsymbol{\alpha}}_j)\|Dir(\boldsymbol{\rho}|\mathbf{1})\right]$$

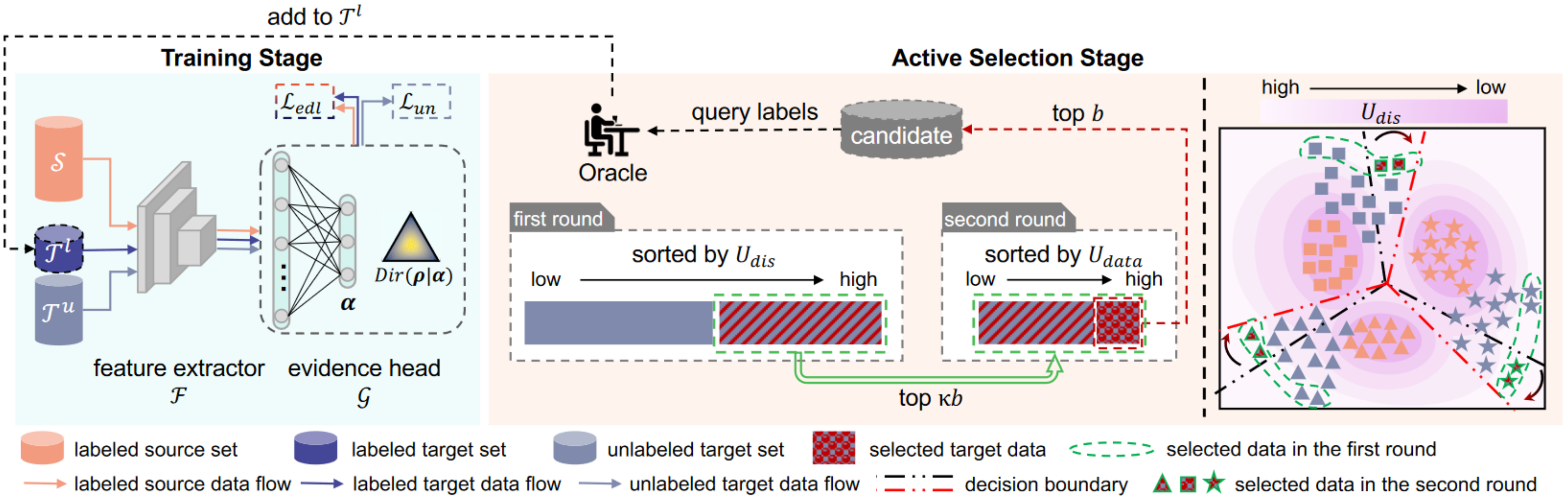
where $\tilde{\boldsymbol{\alpha}}_{i/j} = \boldsymbol{\Upsilon}_{i/j} + (\mathbf{1} - \boldsymbol{\Upsilon}_{i/j}) \odot \boldsymbol{\alpha}_{i/j}$ and $\odot$ is the element-wise multiplication.

# Evidential Model Learning

## Uncertainty loss

$$\mathcal{L}_{un} = \beta\mathcal{L}_{U_{dis}} + \lambda\mathcal{L}_{U_{data}} = \frac{\beta}{|\mathcal{T}^u|}\sum_{\boldsymbol{x}_k \in \mathcal{T}^u} U_{dis}(\boldsymbol{x}_k, \boldsymbol{\theta}) + \frac{\lambda}{|\mathcal{T}^u|}\sum_{\boldsymbol{x}_k \in \mathcal{T}^u} U_{data}(\boldsymbol{x}_k, \boldsymbol{\theta})$$

$$\mathcal{L}_{total} = \mathcal{L}_{edl} + \mathcal{L}_{un} = (\mathcal{L}_{nll} + \mathcal{L}_{kl}) + (\beta\mathcal{L}_{U_{dis}} + \lambda\mathcal{L}_{U_{data}})$$



# Experiment

Table 1: Accuracy (%) on miniDomainNet with 5% target samples as the labeling budget (ResNet-50).

| Method | clp→pnt | clp→rel | clp→skt | pnt→clp | pnt→rel | pnt→skt | rel→clp | rel→pnt | rel→skt | skt→clp | skt→pnt | skt→rel | Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Source-only | 52.1 | 63.0 | 49.4 | 55.9 | 73.0 | 51.1 | 56.8 | 61.0 | 50.0 | 54.0 | 48.9 | 60.3 | 56.3 |
| Random | 61.6 | 78.7 | 61.6 | 64.0 | 78.7 | 63.7 | 60.5 | 64.3 | 61.1 | 64.8 | 58.7 | 75.2 | 66.1 |
| BvSB (Joshi et al., 2009) | 63.2 | 77.9 | 62.7 | 66.7 | 80.5 | 64.9 | 64.3 | 67.0 | 62.2 | 67.6 | 62.5 | 77.8 | 68.1 |
| Entropy (Wang & Shang, 2014) | 63.3 | 78.3 | 61.0 | 65.7 | 81.4 | 63.2 | 63.3 | 66.2 | 63.0 | 67.9 | 60.5 | 78.3 | 67.7 |
| CoreSet (Sener & Savarese, 2018) | 62.6 | 78.3 | 60.2 | 62.1 | 79.9 | 63.6 | 63.6 | 65.2 | 59.1 | 63.1 | 62.3 | 78.1 | 66.5 |
| WAAL (Shui et al., 2020) | 63.2 | 80.2 | 62.1 | 60.6 | 80.3 | 64.6 | 62.9 | 64.1 | 59.5 | 65.4 | 61.8 | 78.6 | 66.9 |
| BADGE (Ash et al., 2020) | 64.3 | 80.8 | 63.5 | 65.2 | 80.2 | 63.8 | 65.9 | 65.4 | 63.4 | 66.7 | 63.3 | 79.2 | 68.5 |
| AADA (Su et al., 2019) | 62.4 | 77.5 | 61.7 | 61.9 | 79.7 | 61.1 | 65.6 | 66.0 | 60.8 | 65.1 | 62.1 | 80.0 | 67.0 |
| DBAL (Deheeger et al., 2021) | 62.9 | 79.2 | 60.8 | 64.6 | 78.1 | 62.5 | 65.6 | 65.2 | 59.2 | 66.3 | 61.3 | 80.3 | 67.2 |
| TQS (Fu et al., 2021) | **67.8** | **82.0** | 65.4 | 67.5 | **84.8** | 66.1 | 63.8 | 67.2 | 62.5 | 71.1 | 64.4 | **81.6** | 70.4 |
| CLUE (Prabhu et al., 2021) | 57.6 | 77.5 | 58.6 | 58.9 | 76.8 | 65.9 | 66.3 | 60.2 | 60.5 | 66.2 | 58.7 | 76.0 | 65.3 |
| EADA (Xie et al., 2021) | 66.0 | 80.8 | 63.5 | 69.4 | 83.0 | 65.1 | 71.1 | 68.6 | 65.7 | 71.0 | 64.3 | 81.0 | 70.8 |
| DUC | 67.1±0.4 | 81.1±0.5 | **67.1**±0.5 | **74.0**±0.6 | 83.5±0.3 | **67.6**±0.3 | **72.4**±0.7 | **70.3**±0.4 | **66.5**±0.4 | **73.5**±0.3 | **70.0**±0.5 | 81.1±0.3 | **72.9**±0.4 |
| Fully-supervised | 74.8 | 89.2 | 73.8 | 82.9 | 89.2 | 75.1 | 82.4 | 75.6 | 74.9 | 82.7 | 73.8 | 88.7 | 80.3 |

For miniDomainNet, since these compared baselines do not report the results on this dataset, we report our own runs based on their open source code.

# Experiment



Table 2: Accuracy (%) on Office-Home and VisDA-2017 with 5% target samples as the labeling budget (ResNet-50).

| Method | VisDA-2017 | Office-Home | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Synthetic→Real | 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 |
| Source-only | 44.7 ± 0.1 | 42.1 | 66.3 | 73.3 | 50.7 | 59.0 | 62.6 | 51.9 | 37.9 | 71.2 | 65.2 | 42.6 | 76.6 | 58.3 |
| Random | 78.1 ± 0.6 | 52.5 | 74.3 | 77.4 | 56.3 | 69.7 | 68.9 | 57.7 | 50.9 | 75.8 | 70.0 | 54.6 | 81.3 | 65.8 |
| BvSB (Joshi et al., 2009) | 81.3 ± 0.4 | 56.3 | 78.6 | 79.3 | 58.1 | 74.0 | 70.9 | 59.5 | 52.6 | 77.2 | 71.2 | 56.4 | 84.5 | 68.2 |
| Entropy (Wang & Shang, 2014) | 82.7 ± 0.3 | 58.0 | 78.4 | 79.1 | 60.5 | 73.0 | 72.6 | 60.4 | 54.2 | 77.9 | 71.3 | 58.0 | 83.6 | 68.9 |
| CoreSet (Sener & Savarese, 2018) | 81.9 ± 0.3 | 51.8 | 72.6 | 75.9 | 58.3 | 68.5 | 70.1 | 58.8 | 48.8 | 75.2 | 69.0 | 52.7 | 80.0 | 65.1 |
| WAAL (Shui et al., 2020) | 83.9 ± 0.4 | 55.7 | 77.1 | 79.3 | 61.1 | 74.7 | 72.6 | 60.1 | 52.1 | 78.1 | 70.1 | 56.6 | 82.5 | 68.3 |
| BADGE (Ash et al., 2020) | 84.3 ± 0.3 | 58.2 | 79.7 | 79.9 | 61.5 | 74.6 | 72.9 | 61.5 | 56.0 | 78.3 | 71.4 | 60.9 | 84.2 | 69.9 |
| AADA (Su et al., 2019) | 80.8 ± 0.4 | 56.6 | 78.1 | 79.0 | 58.5 | 73.7 | 71.0 | 60.1 | 53.1 | 77.0 | 70.6 | 57.0 | 84.5 | 68.3 |
| DBAL (Deheeger et al., 2021) | 82.6 ± 0.3 | 58.7 | 77.3 | 79.2 | 61.7 | 73.8 | 73.3 | 62.6 | 54.5 | 78.1 | 72.4 | 59.9 | 84.3 | 69.6 |
| TQS (Fu et al., 2021) | 83.1 ± 0.4 | 58.6 | 81.1 | 81.5 | 61.1 | 76.1 | 73.3 | 61.2 | 54.7 | 79.7 | 73.4 | 58.9 | 86.1 | 70.5 |
| CLUE (Prabhu et al., 2021) | 85.2 ± 0.4 | 58.0 | 79.3 | 80.9 | 68.8 | 77.5 | 76.7 | 66.3 | 57.9 | 81.4 | 75.6 | 60.8 | 86.3 | 72.5 |
| EADA (Xie et al., 2021) | 88.3 ± 0.1 | 63.6 | 84.4 | 83.5 | 70.7 | **83.7** | 80.5 | 73.0 | 63.5 | 85.2 | 78.4 | 65.4 | 88.6 | 76.7 |
| **DUC** | 88.9 ± 0.2 | **65.5**±0.3 | **84.9**±0.2 | **84.3**±0.4 | **73.0**±0.4 | 83.4±0.2 | **81.1**±0.3 | **73.9**±0.3 | **66.6**±0.5 | **85.4**±0.2 | **80.1**±0.2 | **69.2**±0.3 | **88.8**±0.1 | **78.0**±0.3 |
| Fully-supervised | 93.3 | 95.6 | 99.5 | 99.5 | 99.3 | 99.6 | 99.5 | 99.3 | 95.8 | 99.5 | 99.5 | 95.6 | 99.5 | 98.5 |

# Experiment



Table 3: mIoU (%) comparisons on the task GTAV → Cityscapes.

| Method | budget | road | side. | buil. | wall | fence | pole | light | sign | veg. | terr. | sky | pers. | rider | car | truck | bus | train | motor | bike | mIoU |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Source-only | - | 75.8 | 16.8 | 77.2 | 12.5 | 21.0 | 25.5 | 30.1 | 20.1 | 81.3 | 24.6 | 70.3 | 53.8 | 26.4 | 49.9 | 17.2 | 25.9 | 6.5 | 25.3 | 36.0 | 36.6 |
| MRKLD (Zou et al., 2019) | - | 91.0 | 55.4 | 80.0 | 33.7 | 21.4 | 37.3 | 32.9 | 24.5 | 85.0 | 34.1 | 80.8 | 57.7 | 24.6 | 84.1 | 27.8 | 30.1 | 26.9 | 26.0 | 42.3 | 47.1 |
| Seg-Uncertainty (Zheng & Yang, 2021) | - | 90.4 | 31.2 | 85.1 | 36.9 | 25.6 | 37.5 | 48.8 | 48.5 | 85.3 | 34.8 | 81.1 | 64.4 | 36.8 | 86.3 | 34.9 | 52.2 | 1.7 | 29.0 | 44.6 | 50.3 |
| TPLD (Shin et al., 2020) | - | 94.2 | 60.5 | 82.8 | 36.6 | 16.6 | 39.3 | 29.0 | 25.5 | 85.6 | 44.9 | 84.4 | 60.6 | 27.4 | 84.1 | 37.0 | 47.0 | 31.2 | 36.1 | 50.3 | 51.2 |
| ProDA (Zhang et al., 2021) | - | 87.8 | 56.0 | 79.7 | 46.3 | 44.8 | 45.6 | 53.5 | 53.5 | 88.6 | 45.2 | 82.1 | 70.7 | 39.2 | 88.8 | 45.5 | 59.4 | 1.0 | 48.9 | 56.4 | 57.5 |
| EADA (Xie et al., 2021) | 5% | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 65.2 |
| EADA* (Xie et al., 2021) | 5% | 96.5 | 73.8 | 88.6 | 51.3 | 44.8 | 40.9 | 47.4 | 56.5 | 89.1 | 55.0 | 91.3 | 69.2 | 47.6 | 90.7 | 66.4 | 64.9 | 53.1 | 52.4 | 66.6 | 65.6 |
| **DUC** | 5% | **96.8** | **76.2** | 89.2 | **53.2** | 46.0 | 42.5 | 48.5 | 57.6 | 89.6 | **58.5** | 92.1 | 72.9 | 51.3 | 92.0 | 62.8 | 72.2 | 48.5 | 52.8 | 70.3 | 67.0 |
| Fully-supervised | 100% | 97.2 | 78.1 | 90.6 | 54.5 | 52.7 | 43.2 | 54.2 | 65.1 | 90.5 | 59.9 | 92.4 | 72.8 | 50.7 | 91.8 | 74.0 | 77.2 | 67.6 | 56.3 | 70.9 | 70.5 |
| AADA# (Su et al., 2019) | 5% | 92.2 | 59.9 | 87.3 | 36.4 | 45.7 | 46.1 | 50.6 | 59.5 | 88.3 | 44.0 | 90.2 | 69.7 | 38.2 | 90.0 | 55.3 | 45.1 | 32.0 | 32.6 | 62.9 | 59.3 |
| MADA# (Ning et al., 2021) | 5% | 95.1 | 69.8 | 88.5 | 43.3 | **48.7** | 45.7 | 53.3 | 59.2 | 89.1 | 46.7 | 91.5 | 73.9 | 50.1 | 91.2 | 60.6 | 56.9 | 48.4 | 51.6 | 68.7 | 64.9 |
| **DUC#** | 5% | 95.9 | 70.6 | **89.8** | 50.7 | 48.3 | **47.8** | **53.7** | **59.7** | **90.3** | 56.8 | **93.1** | **74.7** | **55.1** | **92.8** | **74.8** | **77.9** | **63.4** | **59.5** | **71.6** | **69.8** |
| Fully-supervised# | 100% | 96.8 | 80.4 | 90.2 | 48.6 | 56.8 | 52.3 | 58.6 | 68.3 | 90.2 | 59.4 | 93.3 | 75.8 | 54.2 | 92.5 | 74.9 | 79.1 | 71.6 | 56.8 | 71.8 | 72.2 |

Methods with # are based on DeepLab-v3+ (Chen et al., 2018) and others are based on DeepLab-v2 (Chen et al., 2018). Method with budget "-" are the source-only or UDA methods. EADA* denotes the results are based on our own runs according to the corresponding open source code.

# Experiment

Table 4: mIoU (%) comparisons on the task SYNTHIA → Cityscapes. mIoU* is reported according to the average of 13 classes, excluding the "wall", "fence" and "pole".

| Method | budget | road | side. | buil. | wall* | fence* | pole* | light | sign | veg. | sky | pers. | rider | car | bus | motor | bike | mIoU | mIoU* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Source-only | - | 64.3 | 21.3 | 73.1 | 2.4 | 1.1 | 31.4 | 7.0 | 27.7 | 63.1 | 67.6 | 42.2 | 19.9 | 73.1 | 15.3 | 10.5 | 38.9 | 34.9 | 40.3 |
| MRKLD (Zou et al., 2019) | - | 67.7 | 32.2 | 73.9 | 10.7 | 1.6 | 37.4 | 22.2 | 31.2 | 80.8 | 80.5 | 60.8 | 29.1 | 82.8 | 25.0 | 19.4 | 45.3 | 43.8 | 50.1 |
| TPLD (Shin et al., 2020) | - | 80.9 | 44.3 | 82.2 | 19.9 | 0.3 | 40.6 | 20.5 | 30.1 | 77.2 | 80.9 | 60.6 | 25.5 | 84.8 | 41.1 | 24.7 | 43.7 | 47.3 | 53.5 |
| Seg-Uncertainty (Zheng & Yang, 2021) | - | 87.6 | 41.9 | 83.1 | 14.7 | 1.7 | 36.2 | 31.3 | 19.9 | 81.6 | 80.6 | 63.0 | 21.8 | 86.2 | 40.7 | 23.6 | 53.1 | 47.9 | 54.9 |
| ProDA (Zhang et al., 2021) | - | 87.8 | 45.7 | 84.6 | 37.1 | 0.6 | 44.0 | **54.6** | 37.0 | 88.1 | 84.4 | 74.2 | 24.3 | 88.2 | 51.1 | 40.5 | 45.6 | 55.5 | 62.0 |
| **DUC** | 5% | 96.1 | 73.1 | 88.7 | 43.3 | 39.0 | 42.2 | 49.9 | 55.5 | **90.7** | **92.8** | 73.7 | 49.2 | 91.9 | 67.9 | 45.9 | **71.1** | 66.9 | 72.8 |
| Fully-supervised | 100% | 97.3 | 79.4 | 89.6 | 52.8 | 54.0 | 46.7 | 53.4 | 62.6 | 90.5 | 92.9 | 71.3 | 50.8 | 92.1 | 77.9 | 55.4 | 68.7 | 71.0 | 75.5 |
| AADA# (Su et al., 2019) | 5% | 91.3 | 57.6 | 86.9 | 37.6 | **48.3** | 45.0 | 50.4 | 58.5 | 88.2 | 90.3 | 69.4 | 37.9 | 89.9 | 44.5 | 32.8 | 62.5 | 61.9 | 66.2 |
| MADA# (Ning et al., 2021) | 5% | **96.5** | **74.6** | 88.8 | 45.9 | 43.8 | 46.7 | 52.4 | 60.5 | 89.7 | 92.2 | 74.1 | 51.2 | 90.9 | 60.3 | 52.4 | 69.4 | 68.1 | 73.3 |
| **DUC#** | 5% | 96.3 | **74.6** | **89.4** | **46.8** | 47.6 | **46.8** | 49.7 | **63.1** | 90.3 | 91.3 | **74.7** | **53.8** | **93.1** | **78.9** | **57.0** | 71.0 | **70.3** | **75.6** |
| Fully-supervised# | 100% | 97.0 | 80.4 | 90.9 | 48.6 | 56.2 | 52.1 | 58.5 | 67.4 | 91.3 | 93.4 | 75.5 | 54.2 | 92.3 | 78.5 | 56.1 | 71.3 | 72.7 | 77.4 |

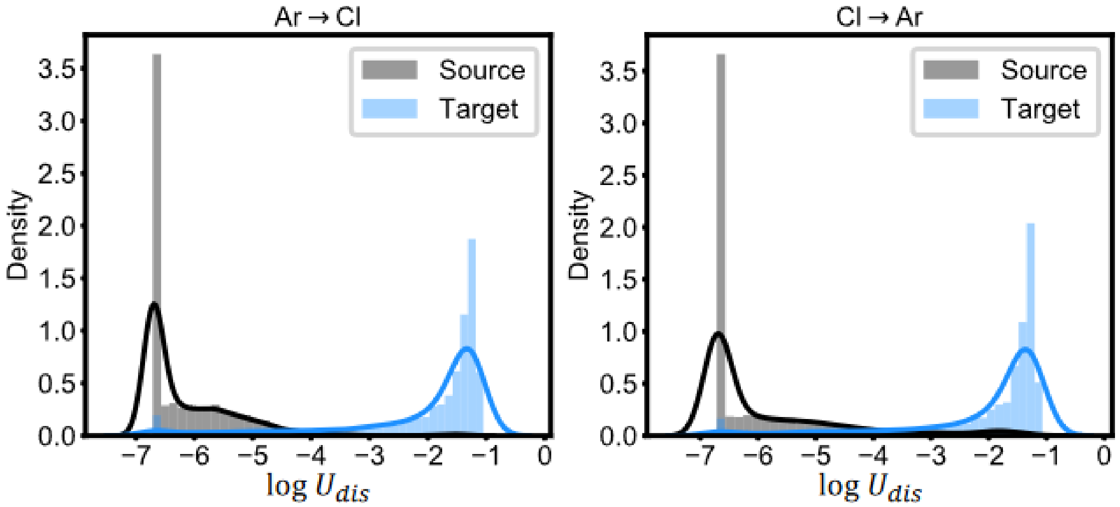
Methods with # are based on DeepLab-v3+ (Chen et al., 2018) and others are based on DeepLab-v2 (Chen et al., 2015). Method with budget "-" are the source-only or UDA methods.

# Experiment

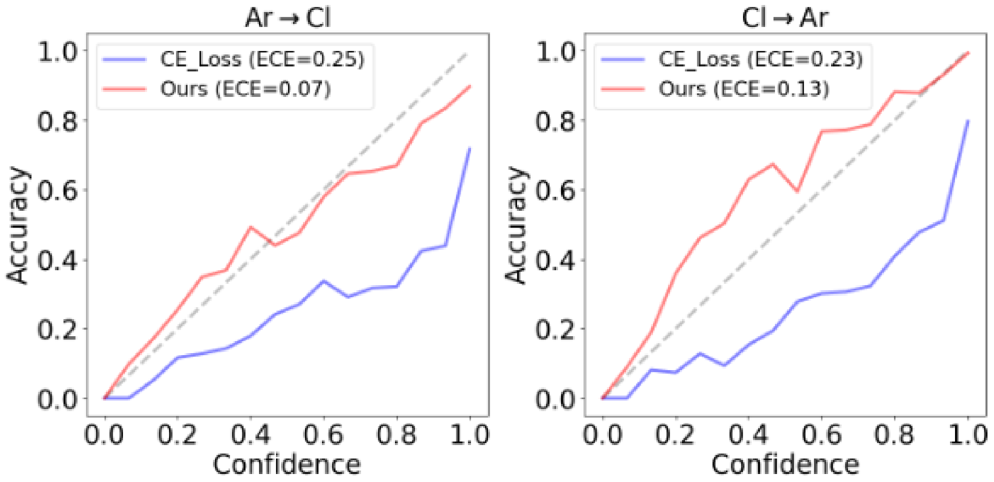


Table 5: Ablation study of DUC on Office-Home.

| Method | Loss | | Active Selection Criterion | | | | Avg |
|---|---|---|---|---|---|---|---|
| | $\mathcal{L}_{U_{dis}}$ | $\mathcal{L}_{U_{data}}$ | random | entropy | $U_{dis}$ | $U_{data}$ | |
| EDL | - | - | - | - | - | - | 61.5 |
| EDL | - | - | ✓ | - | - | - | 71.1 |
| EDL | - | - | - | ✓ | - | - | 73.3 |
| Variant A | - | - | - | - | ✓ | ✓ | 74.1 |
| Variant B | ✓ | - | - | - | ✓ | ✓ | 76.6 |
| Variant C | - | ✓ | - | - | ✓ | ✓ | 74.2 |
| Variant D | ✓ | ✓ | - | - | - | - | 68.6 |
| Variant E | ✓ | ✓ | ✓ | - | - | - | 75.0 |
| Variant F | ✓ | ✓ | - | ✓ | - | - | 76.7 |
| Variant G | ✓ | ✓ | - | - | ✓ | - | 77.1 |
| Variant H | ✓ | ✓ | - | - | - | ✓ | 76.9 |
| DUC | ✓ | ✓ | - | - | ✓ | ✓ | **78.0** |




(a) Distribution of $\log U_{dis}$




(b) Expected calibration error (ECE) of target data

# THANKS