Teacher-Student Competition for Unsupervised Domain Adaption

Ruixin Xiao[1], Zhilei Liu[1], and Baoyuan Wu[2]

Email: zhileiliu@tju.edu.cn

[1] College of Intelligence and Computing
Tianjin University, China
[2] School of Data Science, Chinese University of Hong Kong, Shenzhen, China

January 15, 2021

Outline

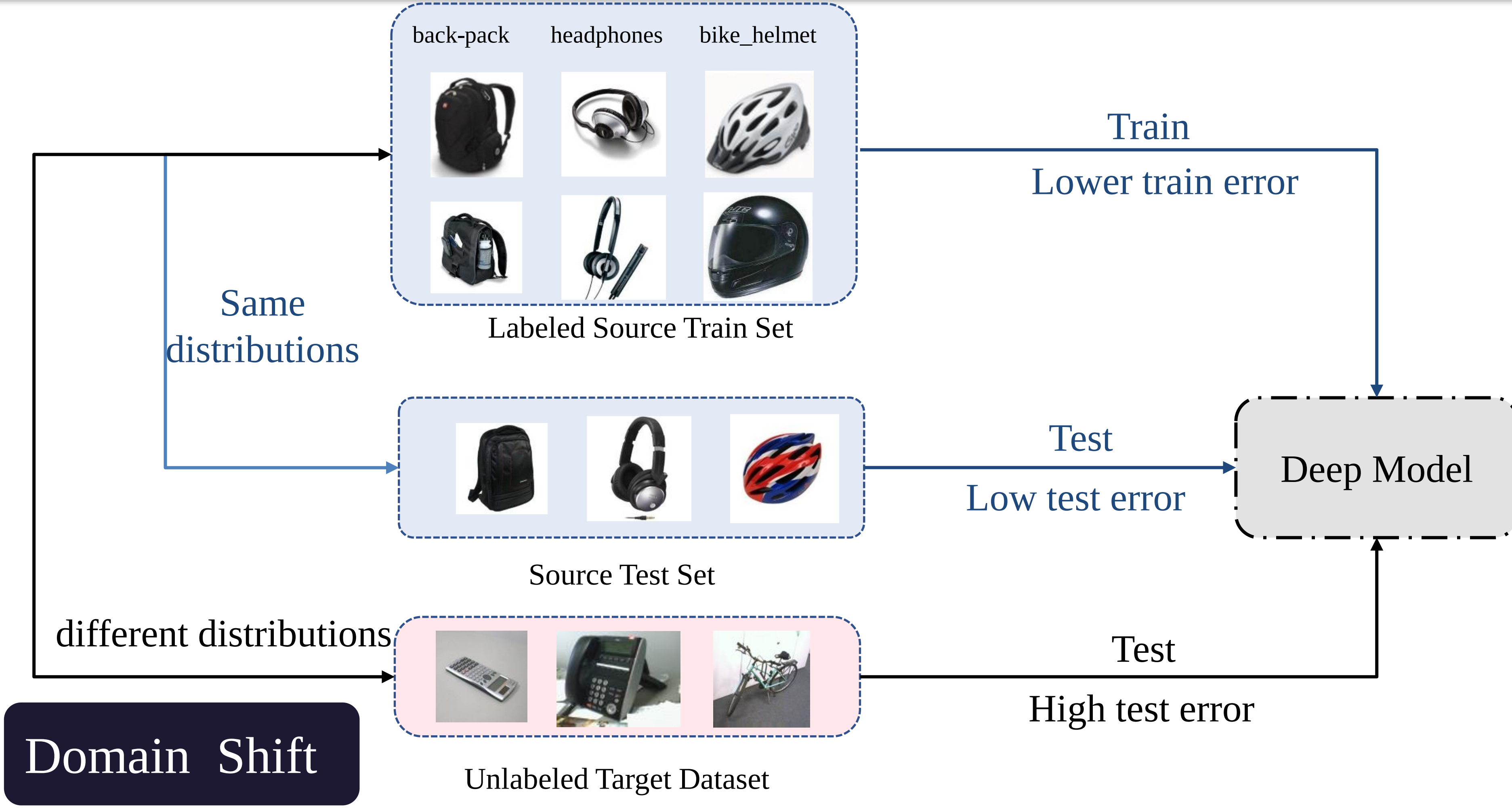
- Introduction
- Proposed Method
- Experiment and Discussion
- Conclusion

Domain Shift

Labeled Source Dataset

Unlabeled Target Dataset

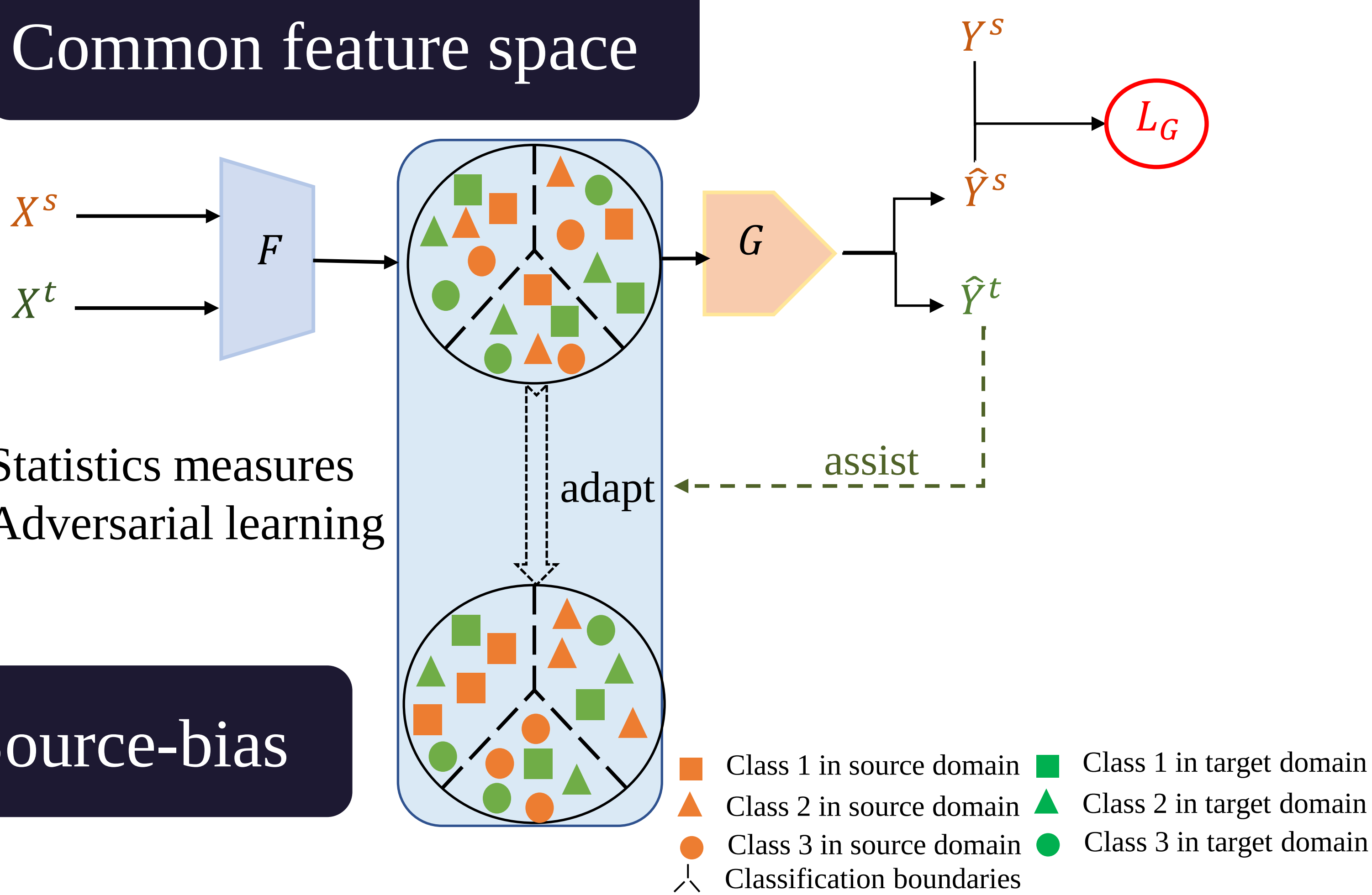
Domain Shift



Unsupervised Domain Adaptation(UDA)

- **Traditional methods**

Common feature space

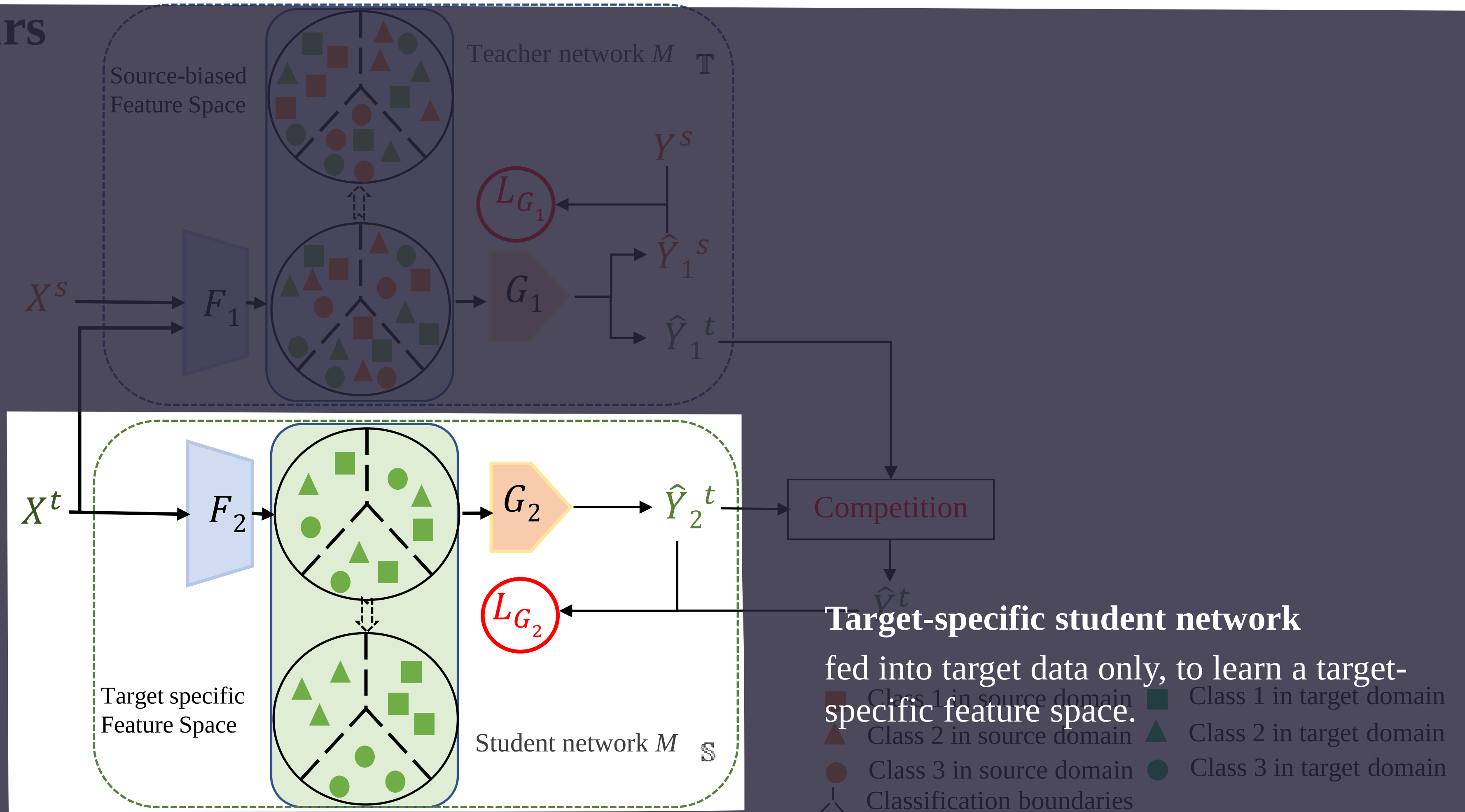


- Statistics measures
- Adversarial learning

Source-bias

Motivation

❑ Ours

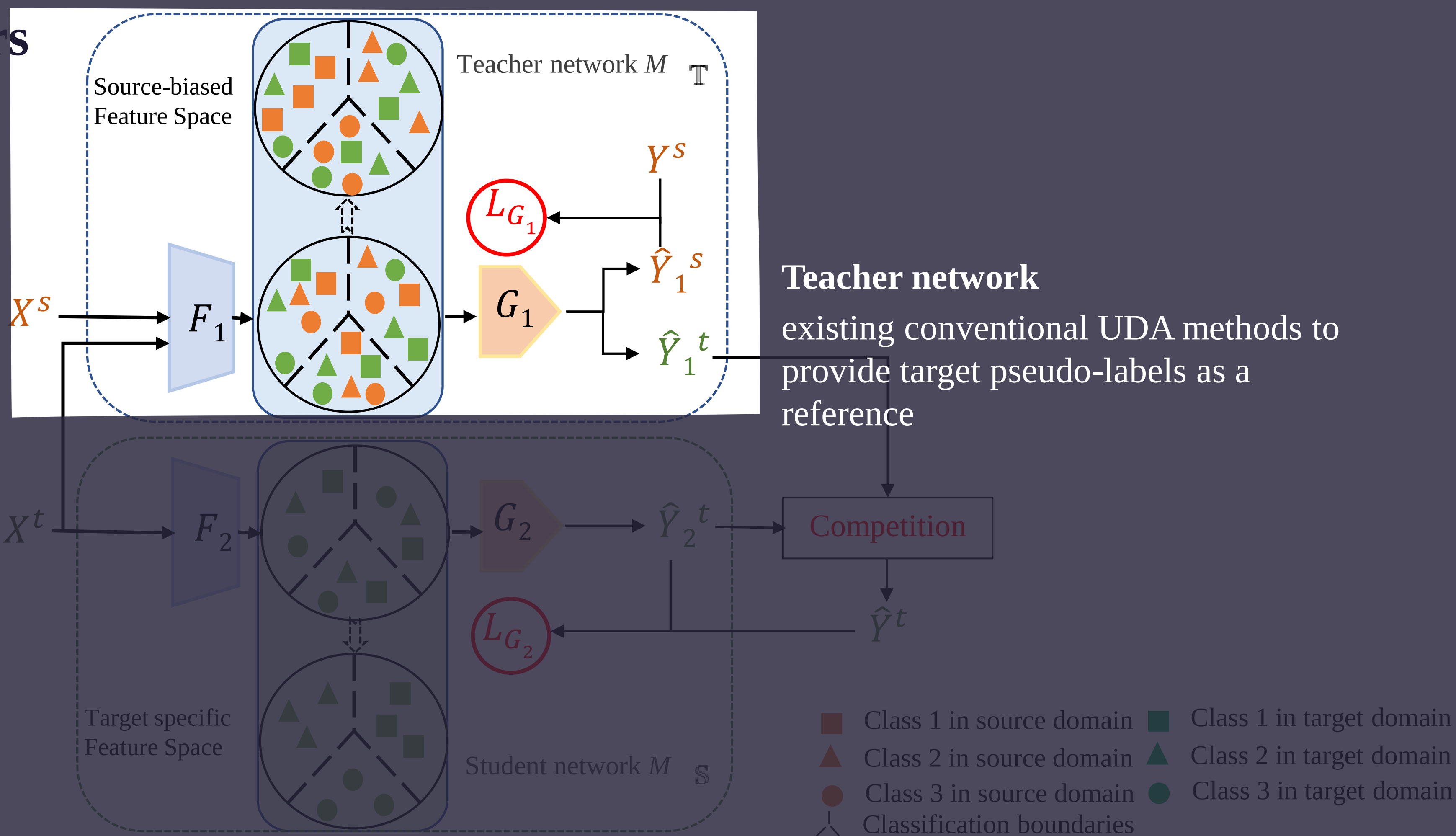


Target-specific student network

fed into target data only, to learn a target-specific feature space.

Motivation

❑Ours

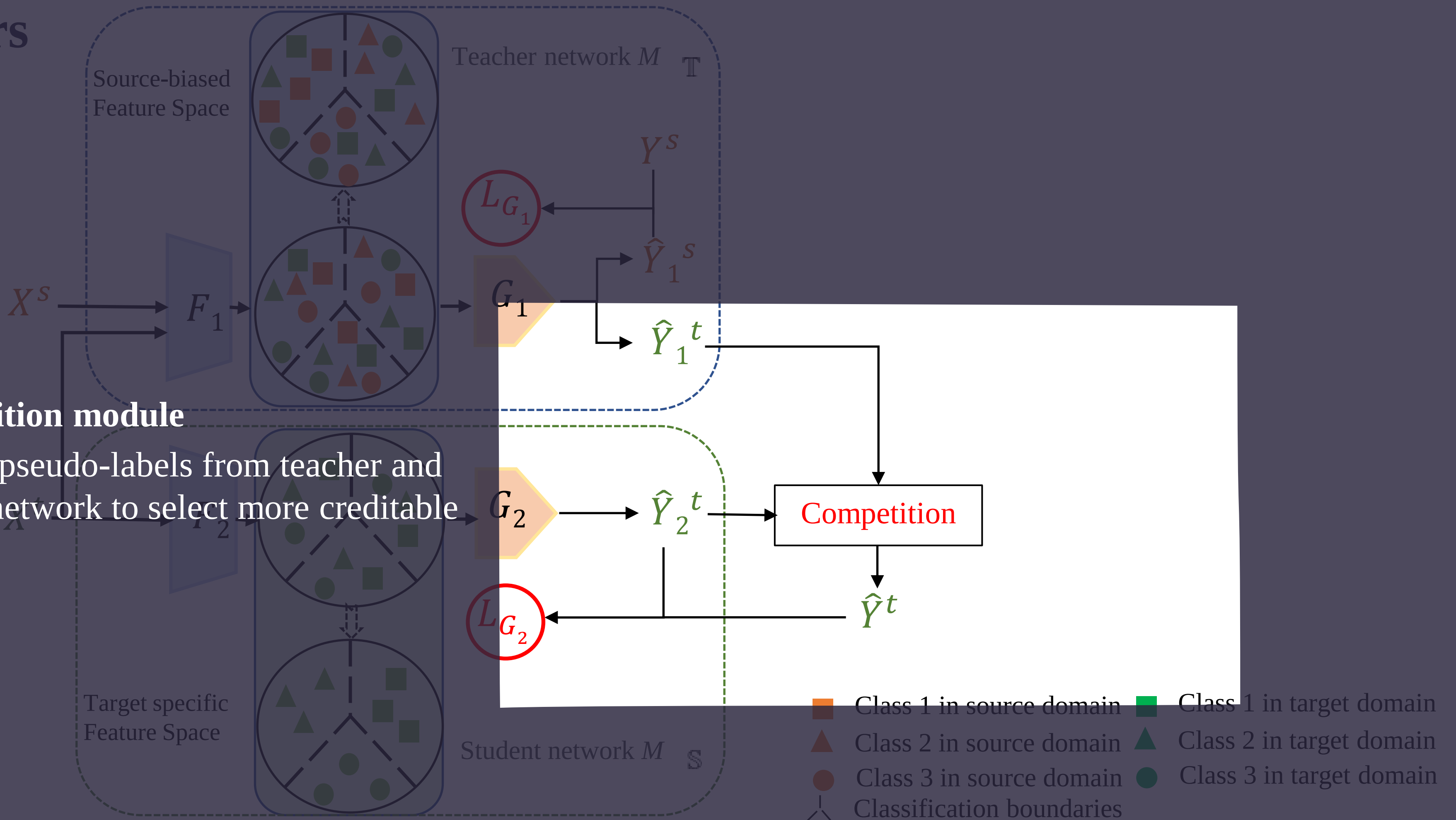


Teacher network

existing conventional UDA methods to provide target pseudo-labels as a reference

Motivation

❑ Ours



Competition module

Receive pseudo-labels from teacher and student network to select more creditable ones

Motivation

❑ Target-specific student network

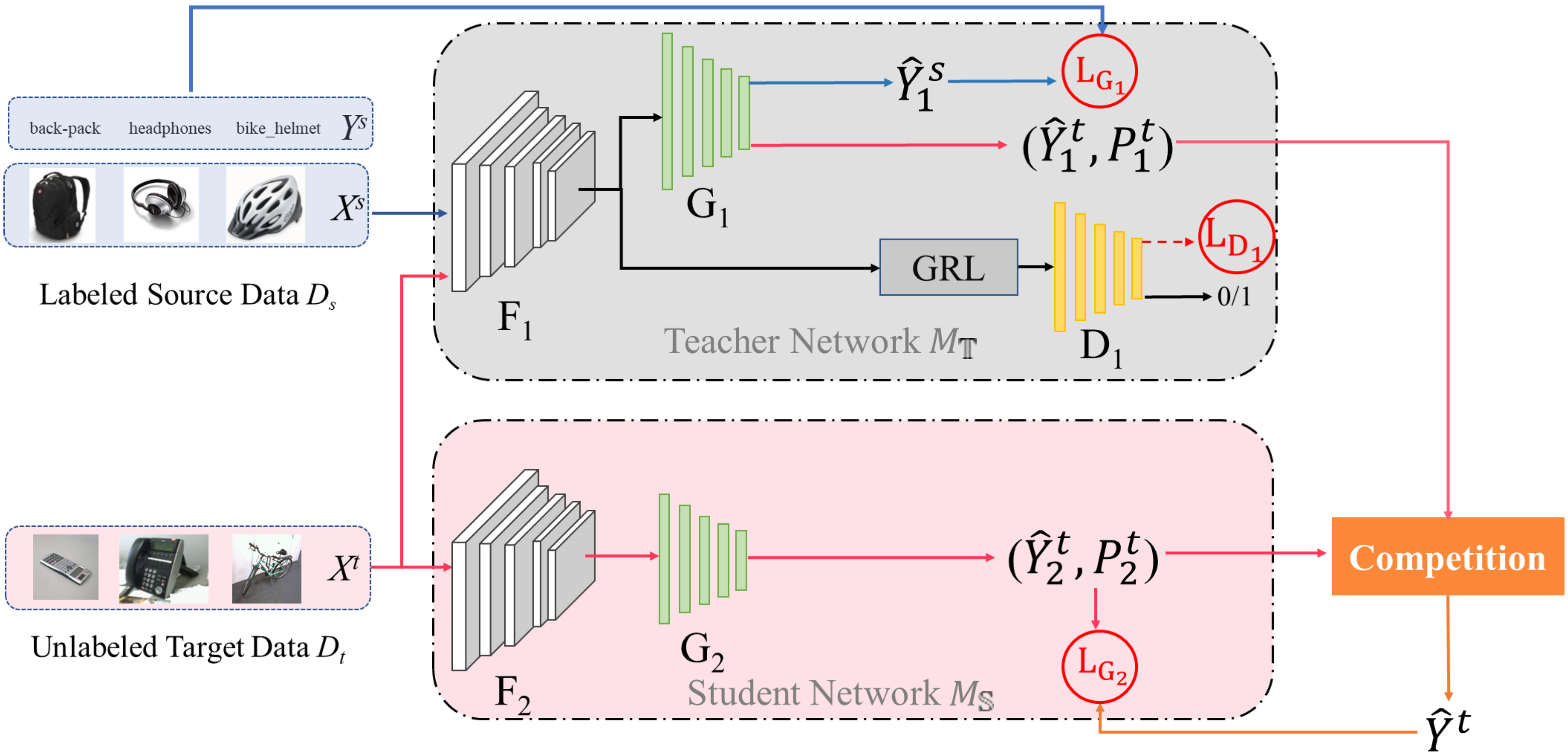
- ❑ To alleviate the source-biased problem, a target-specific student network is built to learn a target-specific feature space.

❑ Competition mechanism

- ❑ A novel pseudo-label selection strategy, in which pseudo-labels from teacher network and student network compete to be the final pseudo-label training for student network.

Proposed Method

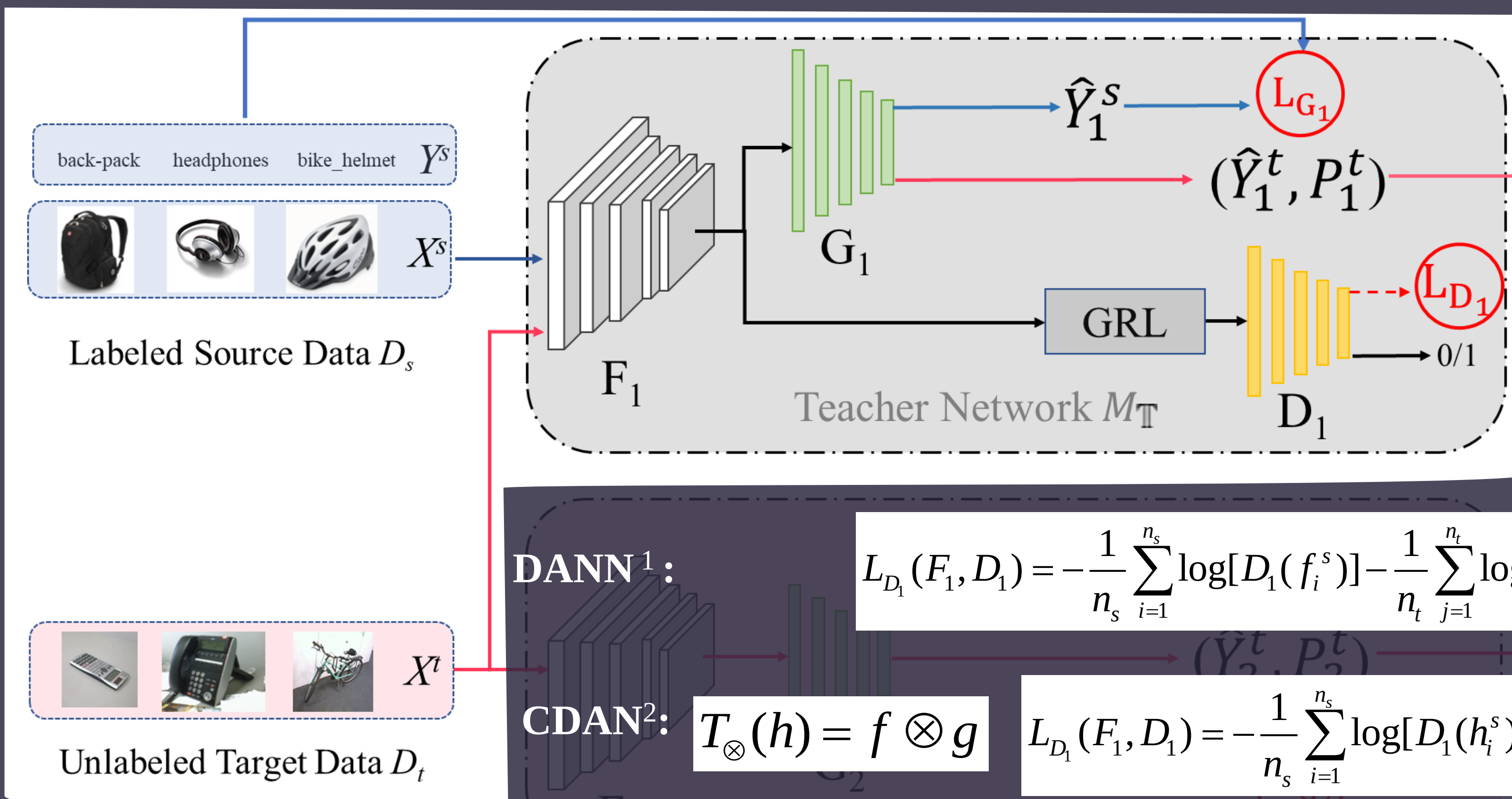
❑TSC



Proposed Method

Teacher network

$$L_{G_1}(F_1,G_1)=\frac{1}{n_s}\sum_{i=1}^{n_s}\ell(G_1(F_1(x_i^s)),y_i^s)$$



DANN[1]:

$$L_{D_1}(F_1,D_1)=-\frac{1}{n_s}\sum_{i=1}^{n_s}\log[D_1(f_i^s)]-\frac{1}{n_t}\sum_{j=1}^{n_t}\log[1-D_1(f_j^t)]$$

CDAN[2]:

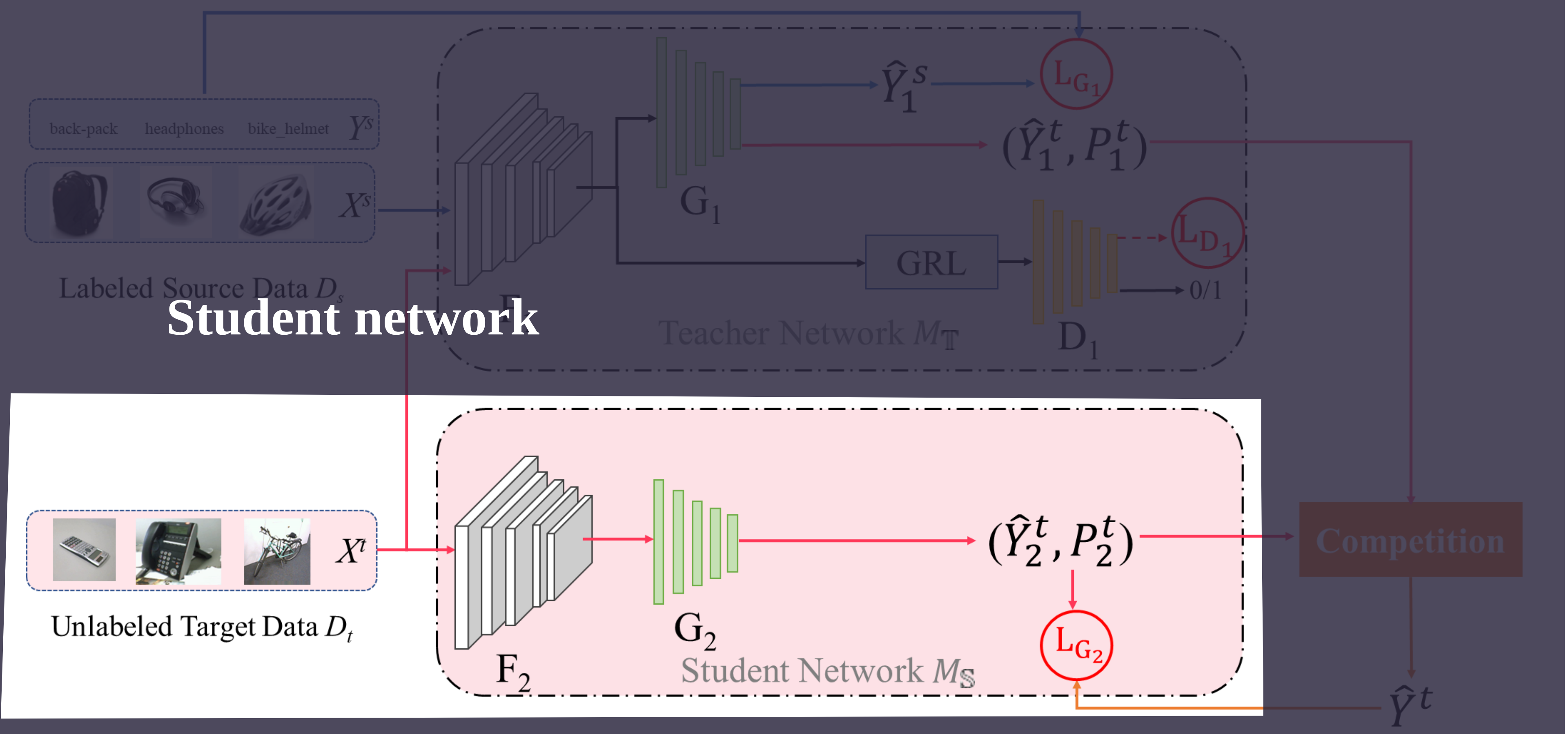
$$T_{\otimes}(h)=f\otimes g$$

$$L_{D_1}(F_1,D_1)=-\frac{1}{n_s}\sum_{i=1}^{n_s}\log[D_1(h_i^s)]-\frac{1}{n_t}\sum_{j=1}^{n_t}\log[1-D_1(h_j^t)]$$

[1] Y. Ganin and V. Lempitsky.Unsupervised domain adaptation by backpropagation[C].ICML2015:1180–1189.
[2] M. Long, Z. Cao, J. Wang, and M. I. Jordan. Conditional adversarial domain adaptation[C]. NIPS 2018:1640–1650..

Proposed Method

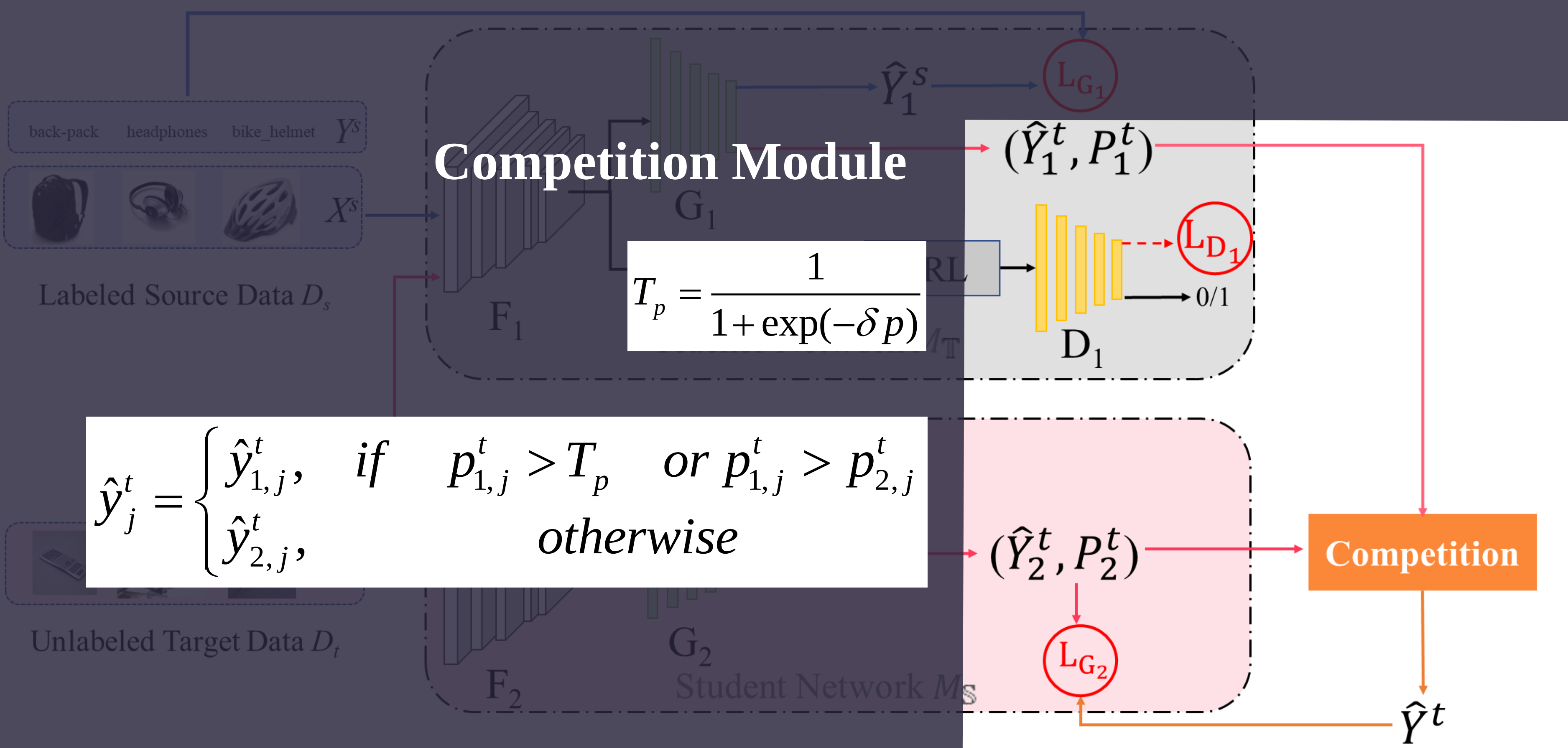
❑TSC



Student network

Proposed Method

❑TSC



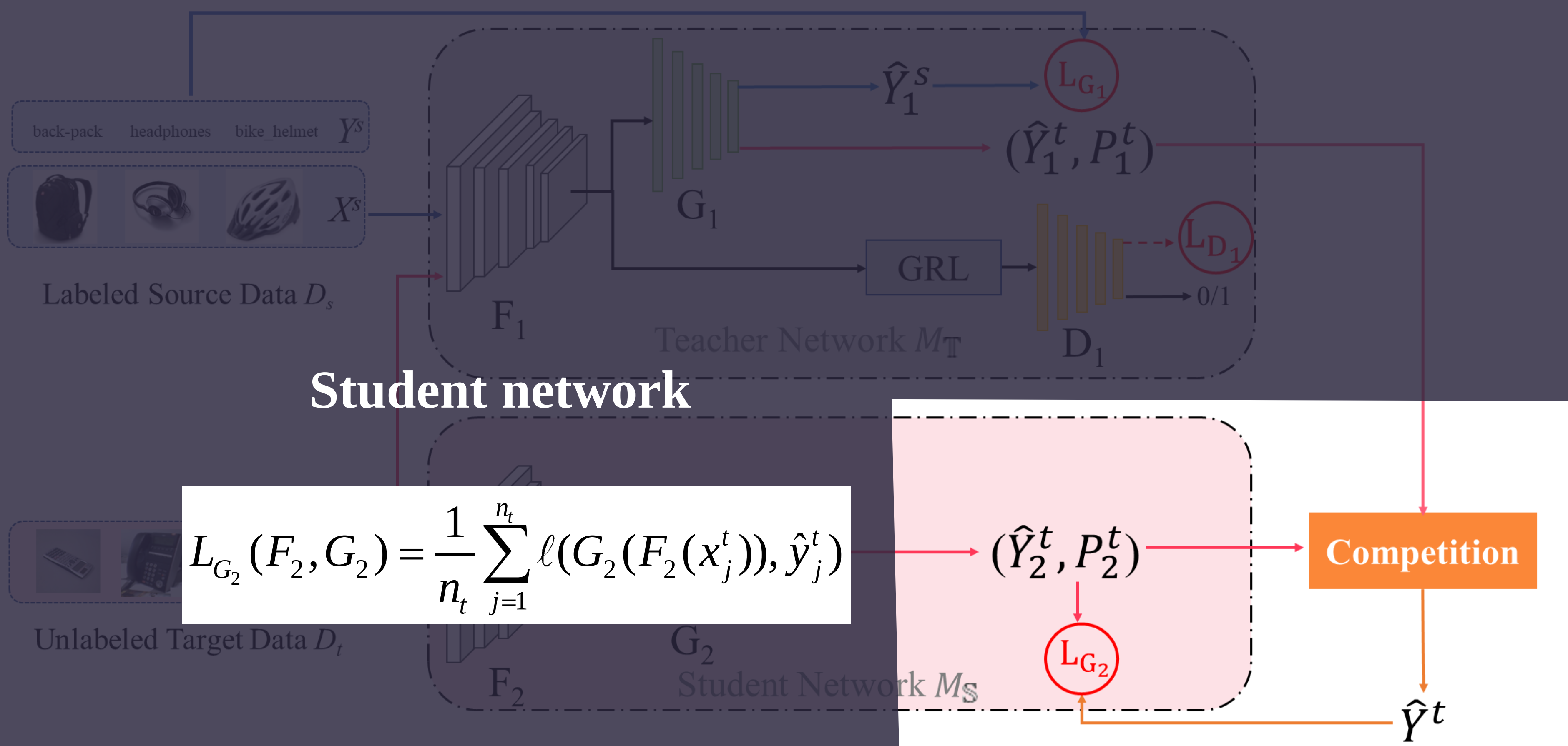
Competition Module

$$T_p = \frac{1}{1+\exp(-\delta p)}$$

$$\hat{y}_j^t = \begin{cases} \hat{y}_{1,j}^t, & if \quad p_{1,j}^t > T_p \quad or \; p_{1,j}^t > p_{2,j}^t \\ \hat{y}_{2,j}^t, & otherwise \end{cases}$$

Proposed Method

❑TSC



$$L_{G_2}(F_2, G_2) = \frac{1}{n_t}\sum_{j=1}^{n_t} \ell(G_2(F_2(x_j^t)), \hat{y}_j^t)$$

Proposed Method

❑ TSC

❑ Overall Objective Function

$$L = L_{G_1}(F_1, G_1) - \lambda L_{D_1}(F_1, D_1) + \beta L_{G_2}(F_2, G_2)$$

Experimental Conditions

❑ Dataset

- ❑ **ImageCLEF-DA: I**, **P** and **C**. 12 categories. 50 images per category.
- ❑ **Office-31: A**, **W**, and **D.** 4,652 images in 31 categories.

❑ Implementation Details

- ❑ **PyTorch using the SGD optimizer**: minibatch size of 36, a weight decay of 0.0005, and a momentum of 0.95.
- ❑ learning rate of the layers trained from scratch are set to be **10 times** those of fine-tuned layers
- ❑ $\lambda = 1.0$ and $\beta = 0.3$

Quantitative Evaluation

Table 4-1 Accuracy (%) on ImageCLEF-DA for unsupervised domain adaptation (ResNet-50).

Methods	I→P	P→I	I→C	C→I	C→P	P→C	**AVG**
RestNet-50	74.8±0.3	91.5±0.3	83.9±0.1	78.0±0.2	65.5±0.3	91.2±0.3	80.7
DAN	74.5±0.4	82.2±0.2	92.8±0.2	86.3±0.4	69.2±0.4	89.8±0.4	82.5
DANN	75.0±0.6	86.0±0.3	96.2±0.4	87.0±0.5	74.3±0.5	91.5±0.6	85.0
JAN	76.8±0.4	88.0±0.2	94.7±0.2	89.5±0.3	74.2±0.3	91.7±0.3	85.8
MADA	75.0±0.3	87.9±0.2	96.0±0.3	88.8±0.3	75.2±0.2	92.2±0.3	85.8
CDAN	76.7±0.3	90.6±0.3	97.0±0.4	90.5±0.4	74.5±0.3	93.5±0.4	87.1
CDAN+E	77.7±0.3	90.7±0.2	**97.7**±0.3	91.3±0.3	74.2±0.2	94.3±0.3	87.7
iCAN	79.5	89.7	94.7	89.9	78.5	92.0	87.4
rDANN+CAT	77.2±0.2	91.0±0.3	95.5±0.3	91.3±0.3	75.3±0.6	93.6±0.5	87.3
TSC+DANN	78.3±0.2	92.8±0.3	96.8±0.2	90.3±0.5	74.5±0.7	96.0±0.2	88.1
TSC+CDAN	**79.0**±0.3	**93.2**±0.5	**97.2**±0.4	**92.7**±0.2	**77.4**±0.4	**96.5**±0.3	**89.3**

Quantitative Evaluation

Table 4-2 Accuracy (%) on Office-31 for unsupervised domain adaptation (ResNet-50).

Methods	A→W	D→W	W→D	A→D	D→A	W→A	AVG
RestNet-50	68.4±0.2	96.7±0.1	99.3±0.1	68.9±0.2	62.5±0.3	60.7±0.3	76.1
DAN	80.5±0.4	97.1±0.2	99.6±0.1	78.6±0.2	63.6±0.3	62.8±0.2	80.4
DANN	82.0±0.4	96.9±0.2	99.1±0.1	79.7±0.4	68.2±0.4	67.4±0.5	82.2
JAN	85.4±0.3	97.4±0.2	99.8±0.2	84.7±0.3	68.6±0.3	70.0±0.4	84.3
MADA	90.0±0.1	97.4±0.1	99.6±0.1	87.8±0.2	70.3±0.3	66.4±0.3	85.2
GTA	89.5±0.5	97.9±0.3	99.8±0.4	87.7±0.5	72.8±0.3	71.4±0.4	86.5
CDAN	93.1±0.2	98.2±0.2	**100.0**±0.0	89.8±0.3	70.1±0.4	68.0±0.4	86.6
CDAN+E	94.1±0.1	98.6±0.1	**100.0**±0.0	92.9±0.2	71.0±0.3	69.3±0.3	87.7
iCAN	92.5	**98.8**	**100.0**	90.1	72.1	68.9	87.2
rDANN+CAT	94.4±0.1	98.0±0.2	**100.0**±0.0	90.8±1.8	72.2±0.6	70.2±0.1	87.6
TSC+DANN	85.0±0.3	98.0±0.1	**100.0**±0.0	80.3±0.4	69.3±0.2	67.7±0.1	83.4
TSC+CDAN	**94.6**±0.3	98.2±0.2	**100.0**±0.0	**94.7**±0.1	**74.0**±0.2	**71.6**±0.7	**88.9**

Quantitative Evaluation

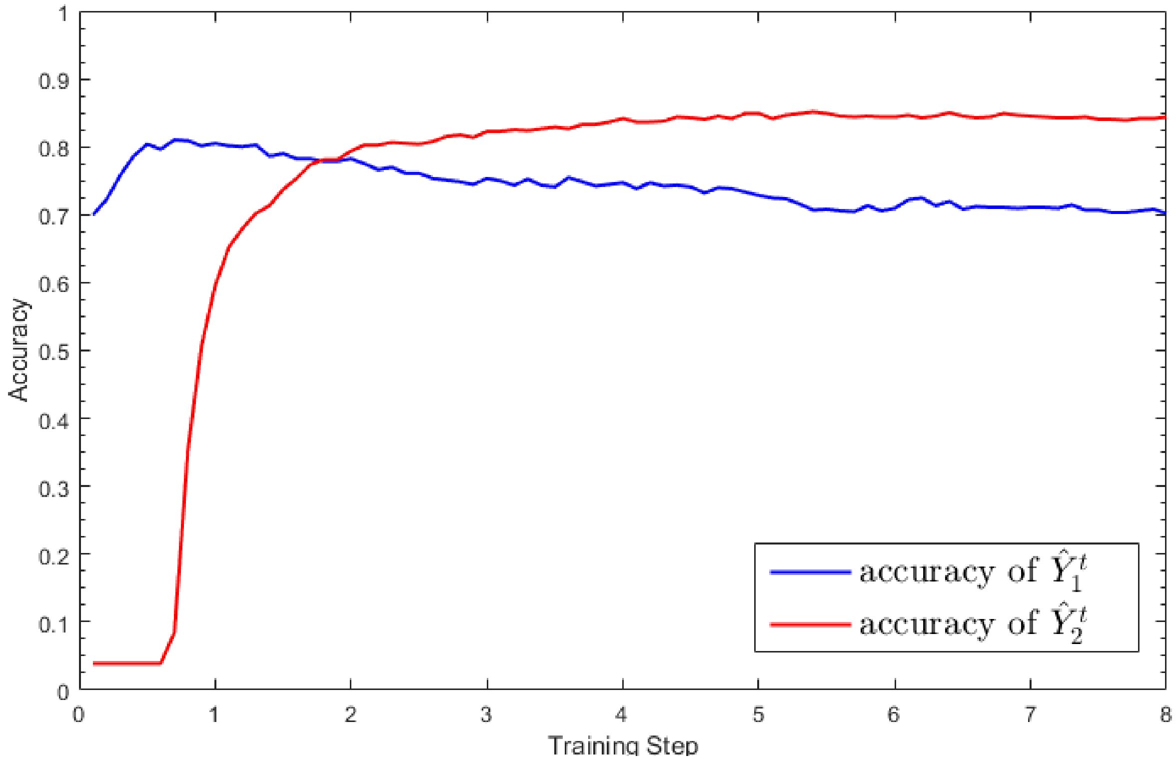
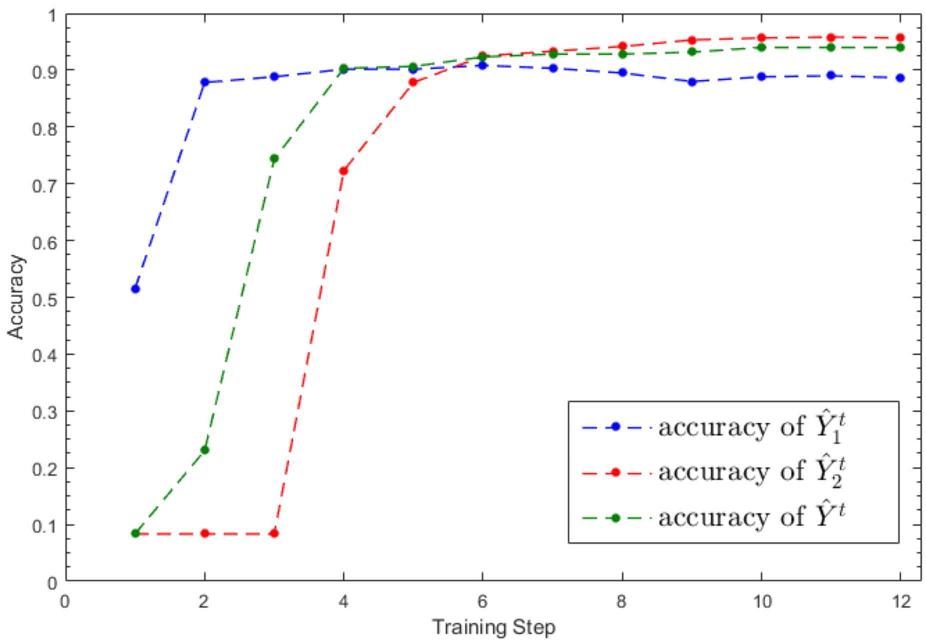
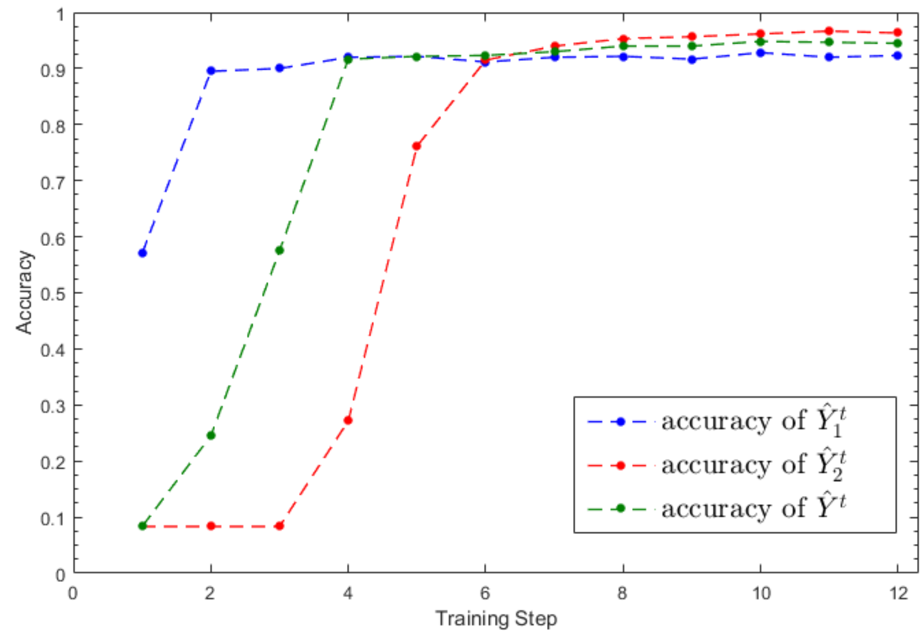


Fig. 4-1:Accuracy of TSC+DANN on A→W in office-31

Quantitative Evaluation



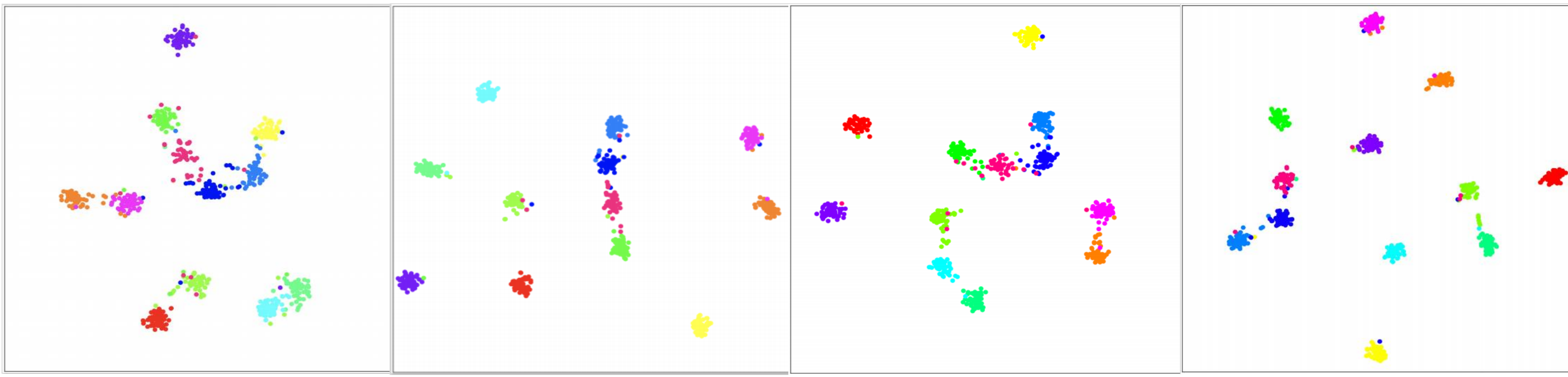
(a) TSC+DANN



(b) TSC+CDAN

Fig. 4-2: Pseudo-labeling accuracy comparision on task P→C in ImageCLEF-DA.

Visualization



(a) M_T in TSC+DANN (b) M_S in TSC+DANN (c) M_T in TSC+CDAN (d) M_S in TSC+CDAN

Fig. 4-3: The t-SNE visualization of target features on task P→C in ImageCLEF-DA.

Conclusion

❑ Conclusion.

- ❑ Our proposed TSC significantly outperforms the state-of-the-art domain adaption methods.
- ❑ More separable target feature space can be achieved by introducing our competition model to tackle the source-bias problem.

❑ Future Work

- ❑ consider to set the hyper parameter T_p dynamically and smoothly to improve the performance further

Teacher-Student Competition for Unsupervised Domain Adaption

Thanks

Q&A

Ruixin Xiao[1], **Zhilei Liu**[1], and Baoyuan Wu[2]

Email: zhileiliu@tju.edu.cn

[1] College of Intelligence and Computing
Tianjin University, China
[2]School of Data Science, Chinese University of Hong Kong, Shenzhen, China

January 15, 2021