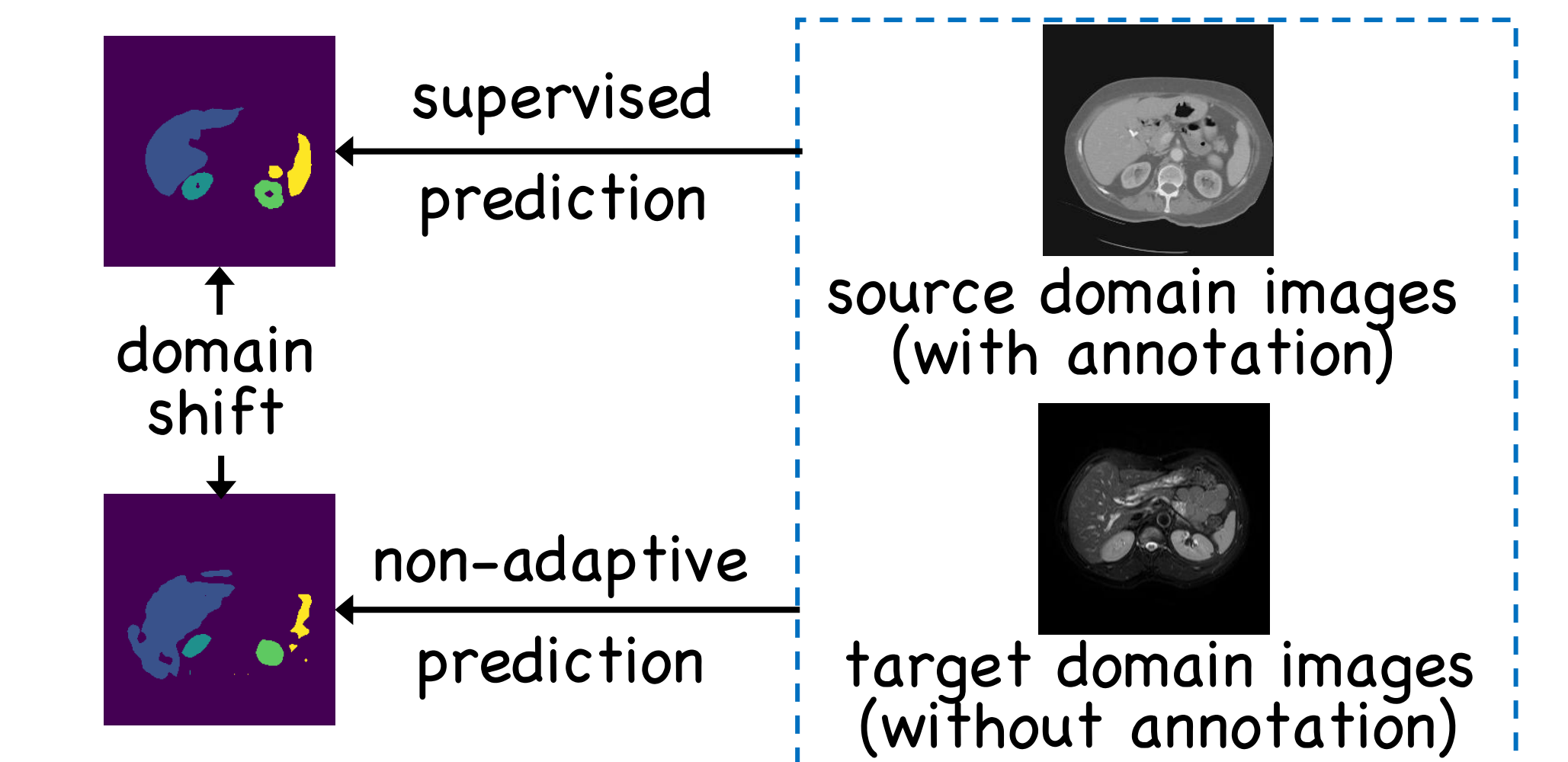


Background and Objectives

Domain shift exists in cross modality image analysis tasks, such as learning the domain shared structural features from MRI and then segmenting multi-organs from CT. When applying the supervised training method to domain adaptation, it is challenging to collect sufficient pixel-wise annotations due to privacy protection or unaffordable expense. Meanwhile, it is hard to utilize source domain's model to get promising performance without any target domain labels. In this case, unsupervised domain adaptation (UDA) becomes a preferring choice.



Most UDA methods try to align domains by minimizing the discrepancy between source domain and target domain in the feature space, where annotations of target samples are absent for training. In collaboration with the above process, pseudo labeling is the self-training method to generate pseudo labels based on the learned domain-invariant representation. Then the pseudo labels are utilized as annotations to fine-tune the adapted model in return.

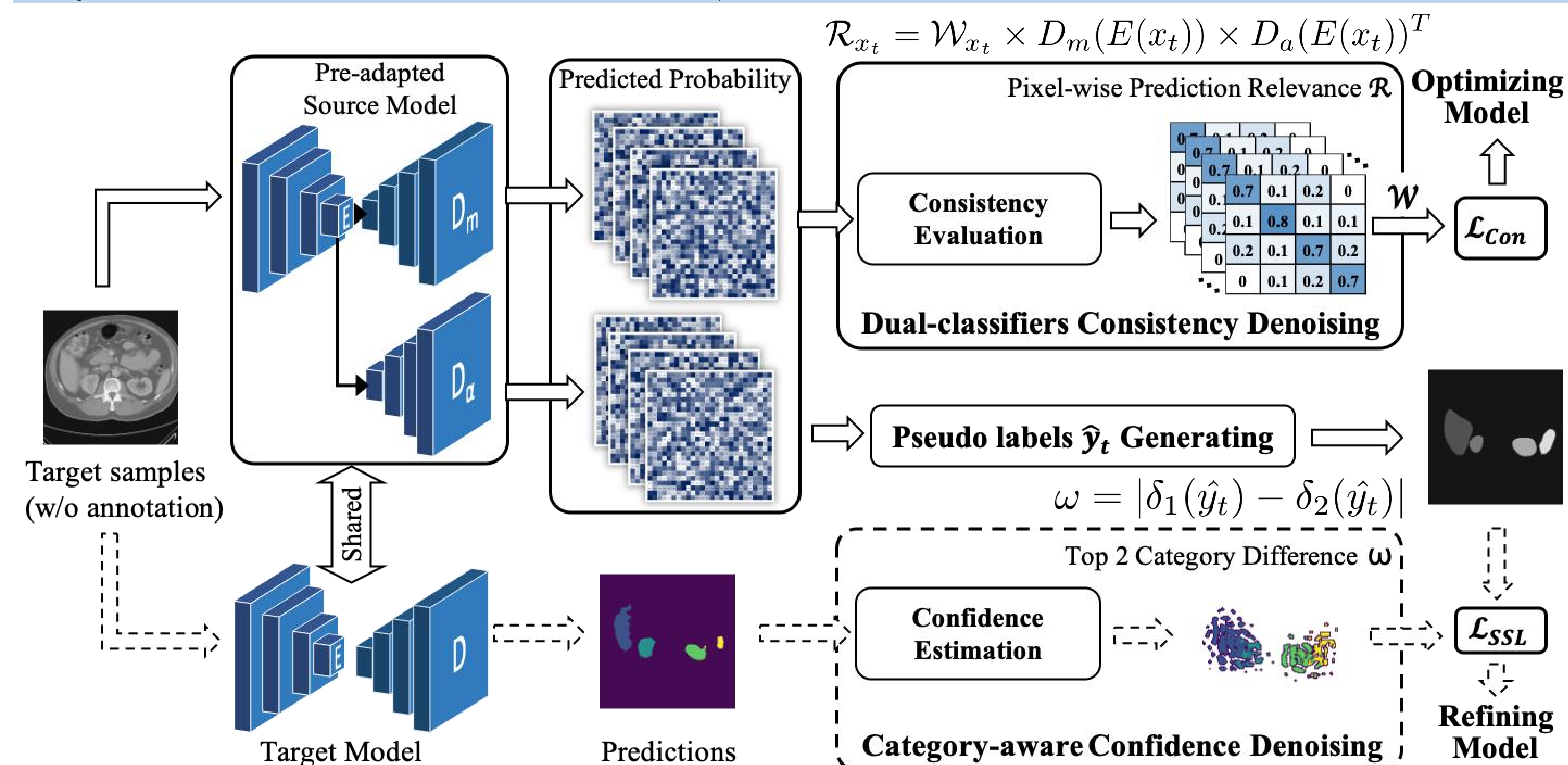
Motivation

Most threshold-based methods generated confident pseudo labels through handcrafted thresholds. However, they ignored that there was still noise inside the prediction. Both noise and semantic information in the prediction were treated equally, which compromised the subsequent training. Meanwhile, the optimal handcrafted thresholds are different for different tasks. Other methods still need to utilize source domain samples for regularization. So it was prone to overfit the source domain-specific features during pseudo labeling. Besides, previous works failed to take advantage of the complementary relationship between confidence and consistency.

Method and Graph Abstract

We investigate two critical properties for the generated pseudo labels, i.e., consistency and confidence.

- **Consistency** is measured by calculating the variance of predictions from different branches and evaluated as the internal reliability of the pseudo label. The dual-classifiers consistency helps promote the robustness of pseudo labels, which is calculated through the weighted relevance matrix.
- **Confidence** is measured by calculating the discrepancy of pseudo label and evaluated by the external reliability of the pseudo label. The category-aware confidence is utilized as adaptive pixel-wise weight to avoid handcrafted thresholds, which is obtained by calculating the probability difference of the top two classes.
- * Higher consistency and confidence denote better prediction.



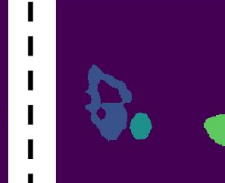
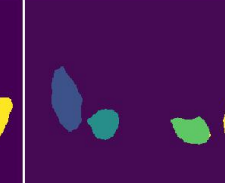



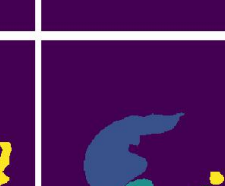
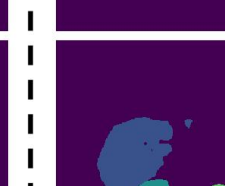
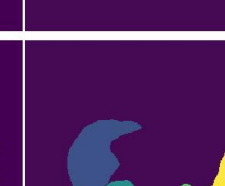




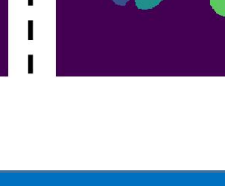
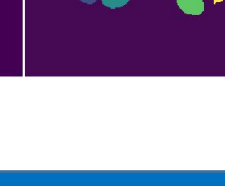










$$\min_{\theta_{D_m, D_a}} \mathcal{L}_{Con} = \sum_{x_t}^{X_T} [\sum_{i,j=1}^C \mathcal{R}_{x_t}^{i,j} - \sum_{i=1}^C \mathcal{R}_{x_t}^{i,i}]$$

$$\min_{\theta_E} \mathcal{L}_{SSL} = \sum_{x_t \in X_T, \hat{y}_t \in \hat{Y}_T} \omega \cdot [-\hat{y}_t \log D(E(x_t))]$$

Results

UDA Methods																
Dice (%)	W/o Adapt	Supervised Train	SynSegNet ^[1]	SynSegNet ^[1] +APL	AdaOutput ^[2]	AdaOutput ^[2] +APL	CycleGAN ^[3]	CycleGAN ^[3] +APL	CycADA ^[4]	CycADA ^[4] +APL	ADVENT ^[5]	ADVENT ^[5] +APL	SIFA-v1 ^[6]	SIFA-v1 ^[6] +APL	SIFA-v2 ^[7]	SIFA-v2 ^[7] +APL
liver	73.1	92.8	85	85.19	85.4	87.65	83.4	86.79	84.5	89.06	89.28	90.37	87.9	89.56	88	89.56
right kidney	47.3	86.4	82.1	81.26	79.7	83.85	79.3	78.54	78.6	81.43	77.05	85.1	83.7	89.74	83.3	87.88
left kidney	57.3	87.4	72.7	79.24	79.7	83.88	79.4	79.24	80.3	83.3	81.37	81.48	80.1	83.66	80.9	89.19
spleen	55.1	88.2	81	84.41	81.7	80.77	77.3	84.62	76.9	81.33	83.45	84.82	80.5	82.31	82.6	82.8
Avg	58.2	88.7	80.2	82.53	81.6	84.04	79.9	82.3	80.1	83.78	82.79	85.44	83.1	86.32	83.7	87.11

Pseudo Methods	liver	right kidney	left kidney	spleen	Avg	Target Image	W/o Adaptation	Supervised Train	CycleGAN ⇒ +APL	SIFA ⇒ +APL	Ground Truth
Threshold-based ^[8]	81.44	75.18	73.18	79.98	77.59						
CBST ^[9]	89.49	80.6	82.14	84.19	84.11						
MRNet ^[10]	87.77	86.67	81.97	83.1	84.88						
APL (L_Con)	88.12	89.7	81.87	81.77	85.37						
APL (+L_SSL)	89.56	89.74	83.66	82.31	86.32						

Summary

The proposed adaptive pseudo labeling method APL focuses on the unsupervised source-free domain adaptation for cross-modality medical images segmentation. Specifically, the dual-classifiers consistency and predictive category-aware confidence are combined to form a novel regularization for pseudo-label denoising.

Algorithm

Algorithm 1 Optimizing process with the proposed APL.

Input: The pre-adapted feature extractor E and classifiers D_m, D_a ; The un-annotated target domain samples, X_T .

Output: Adaptive annotated pseudo labels of target domain \hat{Y}_T and refined models;

- 1: **for** $i = 1$ to N **do** (do Semi-supervised Learning)
- 2: Predict the target domain samples ($x_t \in X_T$);
- 3: Obtain the predictions $D_m(E(x_t)), D_a(E(x_t))$ from the dual-classifiers;
- 4: Evaluate the the weighted Dual-Classifiers Prediction Relevance \mathcal{R} of predictions with Eq.(1);
- 5: Merge the predictions and obtain the pseudo label \hat{y}_t ;
- 6: Evaluate the Category-aware Confidence with Eq.(3);
- 7: Optimize the models θ_{D_m, D_a} by maximizing the Dual-classifiers Consistency with Eq.(2);
- 8: Update the ensemble of target pseudo label $\hat{y}_t \in \hat{Y}_T$;
- 9: Refine the models θ_E by maximizing the the Category-aware Confidence with Eq.(4);
- 10: **end for**
- 11: **return** \hat{Y}_T, E, D_m, D_a ;

Acknowledgment

This research was funded by the National Key Research and Development Program of China (No. 2018YFB0204301).

References

- [1] Yuankai Huo, et.al, "Synseg-net: Synthetic segmentation without target modality ground truth," IEEE TMI 2019.
- [2] Qi Dou, et.al, "Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss," IJCAI 2018.
- [3] Jun-Yan Zhu, et.al, "Unpaired image-to-image translation using cycle-consistent adversarial networks," ICCV 2017.
- [4] Judy Hoffman, et.al, "CycADA: Cycle-consistent adversarial do main adaptation," ICML 2018.
- [5] Tuan-Hung Vu, et.al, "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation," CVPR 2019.
- [6] Cheng Chen, et.al, "Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation," AAAI 2019.
- [7] Cheng Chen, et.al, "Unsupervised bidirectional cross modality adaptation via deeply synergistic image and feature alignment for medical image segmentation," IEEE TMI 2020.
- [8] Dong-Hyun Lee et al, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," WREPL 2013.
- [9] Yang Zou, et al, "Unsupervised domain adaptation for se mantic segmentation via class-balanced self-training," ECCV 2018.
- [10] Zhedong Zheng, et al, "Unsupervised scene adaptation with memory regularization in vivo," IJCAI 2020.