Learn Complementary Pseudo-label for Source-free Domain Adaptive Medical Segmentation

Wanqing Xie
Anhui Medical University&Fanhan Tech&BIDMC/HMS

Mingzhen Li
Washington University in St. Louis

Jinzhou Wu
Harvard Medical School

Yongsong Huang
Tohoku University

Yuanyuan Bu
Shandong University&HMS

Jane You
HK PolyU

Xiaofeng Liu*
MGH/Harvard Medical School

Abstract

Source-free unsupervised domain adaptations (SFUDA) have been a predominant solution for transferring knowledge inherent in the model parameters trained with a privately labeled source domain to apply to an unlabeled target domain. In the case of missing source domain labeled data training, unfortunately, the conventional SFUDA approaches can be easily caught in the pitfall of "winner takes all", i.e., the majority class dominates the predictions of the deep segmentation model in a class-imbalanced task while the minority classes are overlooked. In this work, we provide a complementary self-training (CST) approach for SFUDA segmentation to get over these challenges, since it can be much easier to exclude certain classes with low probabilities than to predict the correct one. Specifically, we resort to the complementary pseudo-label, which can be easier to learn and able to keep low noise level. Its superior performance has been evidenced in a CT-to-MR cardiac anatomical segmentation task with throughout quantitative evaluation.

1 Introduction

Unsupervised domain adaptation (UDA) has been actively developed to transfer knowledge from a labeled source domain to an unlabeled target domain with joint adaptation training [Liu et al., 2022]. Well-labeled source domain data, however, are often inaccessible due to concerns over patients' data privacy or intellectual property [Bateson et al., 2020, Liu et al., 2021b]. As such, it is of great interest and needs to develop an adaptation strategy using a source-free UDA (SFUDA) model, without access to the source data.

In [Chen et al., 2021], the Monte Carlo (MC) dropout is used for pseudo label noise reduction for self-training [Lee, 2013]. However, MC dropout only takes the epistemic uncertainty [Kendall and Gal, 2017] into consideration, while the ignored aleatoric uncertainty can also be common in UDA [Liu et al., 2021a]. In addition, in segmentation (i.e., pixel-wise dense classification) without source domain supervision, the pseudo-label-based models tend to suffer from the state of "winner takes all," which indicates that the model over-fits the majority pixel class but ignores the minority pixel classes, typically in tasks that have an extremely high-proportioned background class.

In this work, we based on a conceptually simple intuition, that is, in a multi-class segmentation task, it can be much easier to exclude certain classes with low probabilities than to predict the correct one.

*Corresponding author: liuxiaofengcmu@gmail.com.

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

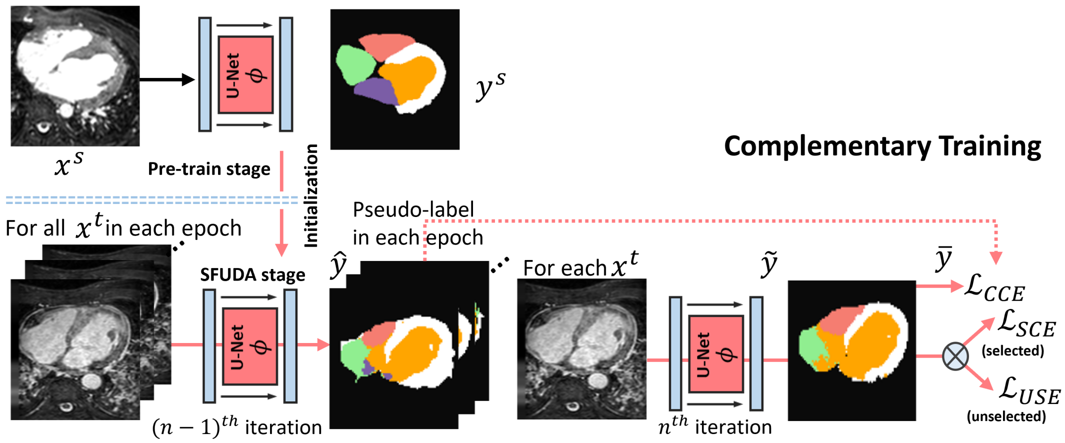

Figure 1: Our proposed CST framework for SFUDA segmentation.

In the registered medical image data, many tissues have their general locations, e.g., the left atrium blood cavity is unlikely to appear on the right.

Therefore, we propose a novel and practicable complementary self-training (CST) framework, in which, to depress the noise in the pseudo label, we propose a novel perspective of CST to generate the complementary labels for each pixel. The complementary label indicates the classes that a pixel is not in. By doing this, we are able to factorize the positive pseudo-label learning in ST to complementary label learning, which can a relatively easier task with lower noise level pseudo-label. The clean and reliable complementary pseudo-label can be efficiently utilized in a complementary learning framework to derive a robust deep model. In addition, the heuristic complementary label selection (HCLS) is specifically proposed to select the relatively reliable and nontrivial complementary label classes for our complementary ST.

2 Methodology

In SFUDA segmentation, we are given a labeled source domain $\{(x_i^s, y_i^s)\}_{i=1}^{N^s} \sim \mathcal{D}^s$, and a unlabeled target domain $\{(x_i^t)\}_{i=1}^{N^t} \sim \mathcal{D}^t$. In source domain pre-training stage, only the source domain data $\{(x_i^s, y_i^s)\}_{i=1}^{N^s}$ are utilized. For simplicity, we adopt the conventional cross-entropy (CE) loss for training, following the previous works [Liu et al., 2021b]. Next, the pre-trained model will be adapted to the target domain, and only the target domain data $\{(x_i^t)\}_{i=1}^{N^t}$ are present.

Self-training [Lee, 2013] assigns unlabeled data with a pseudo label $\hat{y}$ with the previous prediction. Following [Chen et al., 2021], we can use the pre-trained source domain segmentor for x_i^t to predict $\hat{y}_i^t$, and use $\{x_i^t, \hat{y}_i^t\}_{i=1}^{N^t}$ to train the target domain segmentor. However, applying the pre-trained source domain model directly to the target domain can leads to undesirable noisy-labeled pixels due to the domain gap [Zou et al., 2019]. Moreover, without the source domain data supervision in SFUDA adaptation stage, we found that the majority classes, which dominate the deep segmentor, prevail over the minority classes. As a result, the model only predict the background class, with the most important features ignored.

2.1 Self-training with Pseudo Label Selection

In conventional self-training, a straightforward solution to filter out the unreliable and noisy pseudo-labeled pixels $\hat{y}_{i:p}^t$ is to select the pixels with their softmax prediction probabilities higher than a threshold, i.e., to winnow out the low-confidence classes. Specifically, we index all the pixel in $\{\hat{y}_i^t\}_{i=1}^{N^t}$ with p, and sort the probability with $\arg\max p$ in increasing order for each pixel in all training samples independently. Then, we set the confidence threshold τ as the minimum p of the top $\alpha\%$ rank, i.e.,

$$\tau = \min\{\text{top } \alpha\% \text{ sorted } \arg\max p\}. \quad (1)$$

Table 1: Numerical evaluations w.r.t. DSC of the CT-to-MR SFUDA segmentation in MM-WHS.

Method	Source data	Dice Score [%] ↑				
		MYO	LA	LV	RA	RV
Source only[Wu and Zhuang, 2021]	no UDA	0.0811±0.272	3.08±11.63	0.00±0.00	0.742±2.44	23.9±29.2
CRUDA [Bateson et al., 2020]	Partial[2]	28.6±12.5	52.3±14.0	60.1±13.7	54.2±10.8	42.7±8.9
DPL[Chen et al., 2021]	**no**	28.9±13.1	41.5±12.8	60.2±12.4	53.6±9.7	30.7±8.6
OSUDA[Liu et al., 2021b]	**no**	30.7±12.9	57.8±13.7	59.9±10.6	55.3±10.2	51.6±9.5
CST	**no**	**30.9**±13.5	**57.9**±14.0	**60.4**±13.7	**55.8**±9.9	**53.2**±8.4
SIFA[Chen et al., 2020]	yes	37.1±16.0	65.7±18.9	61.2±27.5	51.9±23.3	18.5±19.5

Then, the selected pseudo-label will be applied to the supervised training with cross-entropy loss $\mathcal{L}_{SCE}$ [Liu et al., 2021b], while the un-selected pixels are will be applied to the unsupervised training with self-entropy minimization [Liu et al., 2021b] with $\mathcal{L}_{USE}$, which leads to the prediction approximating to a one-hot distribution, i.e., confident prediction [Grandvalet and Bengio, 2005, Vu et al., 2019].

Though the wrong $\hat{y}^t_{i:p}$ can be removed, the predicament of "winner takes all" still exists, as the model would be biased towards the majority classes while the minority classes are largely overlooked.

2.2 Complementary Self-Training with Heuristic Label Selection

We propose to tackle the noisy pseudo label in SFUDA from the perspective of complementary learning. Compared to determining the correct class in a multi-class segmentation task, it is more feasible to find the classes that are certainly not the correct one and winnow out the classes that have a low probability value. This philosophy can be more promising when applied to SFUDA, which lacks both source and target domain labels supervision. In the previous negative classification method [Kim et al., 2019], the negative label $\overline{y}$ is randomly assigned by choosing any class other than $\hat{y}$. However, in SFUDA segmentation, this would result in significantly noisy pseudo labels due to the domain gap. Intuitively, the "head" classes (high probability other than the pseudo label) are still likely to be the correct one, and assigning a negative label can cause an accumulation of errors. Moreover, The trivial or too easy samples might lead to slow convergence [Liu et al., 2019]. In contrast, the "tail" classes that have low probabilities indicate the network has learned to be quite confident that the pixel does not correspond to these classes.

Thus, we propose a heuristic complementary label selection (HCLS) scheme to choose the relatively reliable and nontrivial pixel-wise $\overline{y}^t_{i:p}$, i.e., the "body" classes, which alleviates the aforementioned risks. Specifically, for each pixel, we set $\overline{y}^t_{i:p}$ as the classes with the top δ softmax value, which can be formulated as:

$$\delta = [C//2] + rand[-\epsilon, +\epsilon] \tag{2}$$

where // represents integer division, and band parameter ϵ is used to define the selection range. Then, the complementary CE loss can be formulated as:

$$\mathcal{L}_{CCE} = -\sum_{\forall p} \overline{y}^t_{i:p} \log(1 - f^t(x^t_{i:p})). \tag{3}$$

The proposed CST is optimized in a unified framework as shown in Fig. 1.

3 Experiments and Results

We used the MM-WHS dataset for the whole heart segmentation [Zhuang and Shen, 2016], and follow the standard evaluation protocol [Wu and Zhuang, 2021, Chen et al., 2020] for CT-to-MRI segmentation in a subject-independent manner. For a fair comparison, we adopted the same segmentation network backbones in [Wu and Zhuang, 2021]. The quantitative evaluation results of 2D MR slice segmentation are provided in Table 1. The proposed CST-based methods largely outperformed CRUDA [Bateson et al., 2020], DPL [Chen et al., 2021] and OSUDA [Liu et al., 2021b]. In addition, our CST-based methods outperformed the source-available UDA methods, e.g., SIFA [Chen et al., 2020], for the segmentation of MYO, RA, and RV w.r.t. DSC. In qualitative checking, we can see that without CST, there can be a relatively large black area in the middle of the prediction. Similarly, there can be a large RV area at the bottom for DPL, and a large LV area between RA and RV for DPL/OSUDA.

4 Potential negative societal impact

Though unsupervised domain adaptation can significantly alleviate the costly labeling in the target domain and has been applied to many medical image analysis tasks, the performance in cases with large domain gap is usually limited. The inaccurate delineation may lead to misinterpretation of the patient's status. Nevertheless, The trained model can be used to provide the draft of a segmentation mask or for computer-aided systems, instead of fully automated diagnosis systems. A careful check by the experienced clinician is necessary in practice.

References

Mathilde Bateson, Hoel Kervadec, Jose Dolz, Hervé Lombaert, and Ismail Ben Ayed. Source-relaxed domain adaptation for image segmentation. In *MICCAI*, pages 490–499, 2020.

Cheng Chen, Qi Dou, Hao Chen, Jing Qin, and Pheng Ann Heng. Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation. *IEEE transactions on medical imaging*, 39(7):2494–2505, 2020.

Cheng Chen, Quande Liu, Yueming Jin, Qi Dou, and Pheng-Ann Heng. Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In *MICCAI*, pages 225–235, 2021.

Yves Grandvalet and Yoshua Bengio. Semi-supervised learning by entropy minimization. In *NIPS*, 2005.

Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30, 2017.

Youngdong Kim, Junho Yim, Juseung Yun, and Junmo Kim. Nlnl: Negative learning for noisy labels. In *ICCV*, pages 101–110, 2019.

Dong-Hyun Lee. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In *ICML Workshop on Challenges in Representation Learning*, 2013.

Xiaofeng Liu, Chaehwa Yoo, Fangxu Xing, Hyejin Oh, Georges El Fakhri, Je-Won Kang, Jonghye Woo, et al. Deep unsupervised domain adaptation: A review of recent advances and perspectives. *APSIPA Transactions on Signal and Information Processing*, 11(1).

Xiaofeng Liu, BVK Vijaya Kumar, Ping Jia, and Jane You. Hard negative generation for identity-disentangled facial expression recognition. *Pattern Recognition*, 88:1–12, 2019.

Xiaofeng Liu, Fangxu Xing, Maureen Stone, Jiachen Zhuo, Timothy Reese, Jerry L Prince, Georges El Fakhri, and Jonghye Woo. Generative self-training for cross-domain unsupervised tagged-to-cine mri synthesis. In *MICCAI*, pages 138–148, 2021a.

Xiaofeng Liu, Fangxu Xing, Chao Yang, Georges El Fakhri, and Jonghye Woo. Adapting off-the-shelf source segmenter for target medical image segmentation. In *MICCAI*, pages 549–559, 2021b.

Xiaofeng Liu, Fangxu Xing, Georges El Fakhri, and Jonghye Woo. Memory consistent unsupervised off-the-shelf model adaptation for source-relaxed medical image segmentation. *Medical Image Analysis*, page 102641, 2022.

Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *CVPR*, 2019.

Fuping Wu and Xiahai Zhuang. Unsupervised domain adaptation with variational approximation for cardiac segmentation. *IEEE Transactions on Medical Imaging*, 40(12):3555–3567, 2021.

Xiahai Zhuang and Juan Shen. Multi-scale patch and multi-modality atlases for whole heart segmentation of mri. *Medical image analysis*, 31:77–87, 2016.

Yang Zou, Zhiding Yu, Xiaofeng Liu, BVK Kumar, and Jinsong Wang. Confidence regularized self-training. In *ICCV*, pages 5982–5991, 2019.