2024 IEEE International Conference on Multimedia and Expo (ICME) | 979-8-3503-9015-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICME57554.2024.10687489

Source-Free Domain Adaptation for Point Cloud Semantic Segmentation

Jianshe Duan[1], Yachao Zhang[2], Yanyun Qu[1†]
[1]Department of Computer Science and Technology, School of Informatics, Xiamen University, Fujian, China
[2]Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, Guangdong, China
jsduancs@foxmail.com, yachaozhang@stu.xmu.edu.cn, yyqu@xmu.edu.cn[†]

***Abstract*—Point cloud semantic segmentation (PCSS) is fundamental in 3D scene understanding. Domain adaptation methods for PCSS enable transferring the knowledge learned from a labeled source domain to an unlabeled target domain with different data distributions. However, they become inapplicable in privacy-preserving scenarios, where we only can leverage a given source model instead of the source data. In this paper, we propose the first source-free domain adaptation PCSS framework to make a given source model generalize to the target domain well. We found existing self-training based source-free domain adaptation methods inevitably lead to confirmation bias and suffer from serious model degradation on PCSS. Therefore, we devise a novel Teacher-Guide Source-Free (TGSF) framework to conquer the above two challenges, including the bidirectional pseudo-label selection and multi-level target consistency learning. Extensive experiments on three benchmarks verify that TGSF achieves state-of-the-art performance, and even outperforms those methods that access the source data.**

***Index Terms*—source-free domain adaptation, point cloud semantic segmentation, domain adaptation**

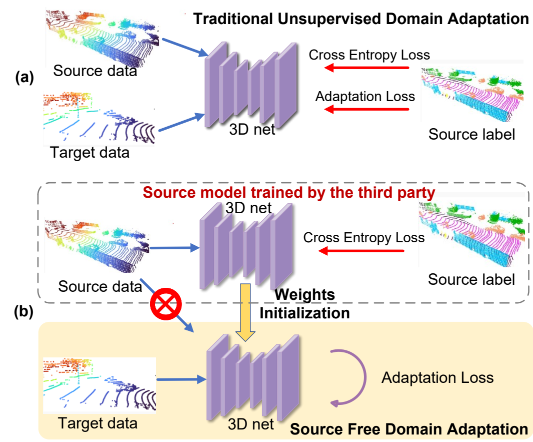

Fig. 1. UDA-PCSS requires source data and target data simultaneously (a). SF-PCSS aims to transfer the knowledge stored in the source model without the help of source data to the target domain (b).

I. Introduction

Point Cloud Semantic Segmentation (PCSS) plays an essential role in 3D computer vision, underpinning critical advancements in fields such as autonomous driving, virtual reality, and robotics. Existing methods [1], [2] have demonstrated excellent performance when the training and test datasets share the same distribution. However, [3], [4] their effectiveness significantly decreases when this assumption does not hold. To address this problem, Unsupervised Domain Adaptation (UDA) methods [3], [4] in PCSS aim to reduce the discrepancy between data distributions through feature alignment. Despite notable improvements, a primary limitation of these methods is their dependence on the accessibility of source data for training, which may be impractical in data privacy-preserving scenarios.

In this paper, we introduce a novel and more practical task setting: Source-Free domain adaptation for Point Cloud Semantic Segmentation (SF-PCSS). Unlike traditional UDA-PCSS methods that depend on source data, SF-PCSS aims to transfer the knowledge from the source model directly to the target model, as illustrated in Fig. 1. This setting is especially crucial in privacy-preserving scenarios, where source data no longer be available. To the best of our knowledge, this is the first exploration of SF-PCSS.

† Corresponding author.

Although prior Source Free Domain Adaptation (SFDA) methods [5], [6] have shown success in classification and 2D image segmentation, their application to 3D semantic segmentation faces two significant challenges. Firstly, 3D point cloud data, being irregular, orderless, and continuous, is fundamentally different from 2D images. While images contain rich semantic information, point clouds mainly convey 3D geometric details. These disparities hinder the effectiveness of existing SFDA methods in generating accurate pseudo-labels in PCSS task, leading to the confirmation bias [5] which refers to overfitting to incorrect pseudo-labels. Secondly, we observe that existing single-structure SFDA methods are particularly vulnerable to parameter changes, resulting in severe model degradation over iterations, which is the phenomenon known as catastrophic forgetting [7] in transfer learning.

To address these challenges, we introduce a novel method, the Teacher-Guide Source-Free framework, which includes two key components: Bidirectional Pseudo-Label Selection (BPLS) and Multi-level Target Consistency Learning (MTCL). Motivated by the observation that: source model tends to annotate accurate pseudo-label for those target data that is similar to the source data, BPLS is designed to mitigate confirmation bias by dividing target data into source-similar and source-dissimilar parts for selecting more accurate pseudo-labels.

979-8-3503-9015-5/24/$31.00 ©2024 IEEE
Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.

Meanwhile, MTCL aims to prevent catastrophic forgetting by leveraging a dual network structure, exploring complementary supervision between teacher and student models in the teacher-guide framework. We summarize our main contributions are three-fold:

1. To the best of our knowledge, we are the first to explore source-free domain adaptation for point cloud semantic segmentation, where existing SFDA methods on 2D image are inapplicable due to confirmation bias and catastrophic forgetting.
2. We propose a novel and effective TGSF framework which contains the BPLS strategy and MTCL module. BPLS is designed to mitigate confirmation bias and MTCL can effectively avoid catastrophic forgetting.
3. Extensive experimental results demonstrate that TGSF achieves state-of-the-art performance without using source data, and even outperforms UDA methods with source data assistance.

II. Related Work

Source free domain adaptation [8], [9] aims to adapt unlabeled target domain by using a pre-trained source model without accessing any source data. Existing SFDA works can be divided into two categories: data generation methods and model fine-tuning methods. Data generation methods attempt to directly generate source-like images by batch normalization statistics [10] or generate networks [11], [12] and apply traditional UDA methods to achieve source knowledge transfer. There also some works [13], [14] propose to generate source distribution and then align it with target distribution. Different with data generation methods, many studies utilize target data to fine-tuning source model with self-supervised ways. For instance, source hypothesis transfer (SHOT) [8] attempt to use information enrtopy for mine the feature structure of the target domain. [6], [15] use cluster-aware pseudo labeling for target data by exploiting intrinsic feature structures of target data. [16] use contrastive learning to capture discriminative representations of target data.

Although these methods have achieved good performance on image classification and image semantic segmentation tasks, they cannot directly deal with 3D vision tasks due to the differences between image and point cloud.

III. Method

A. Problem Definition

Due to privacy-preserving and storage limitation, SF-PCSS cannot utilize the n_s labeled data $\left\{x_s^i, y_s^i\right\}_{i=1}^{n_s}$ from source domain D_s during the adaptation stage, which is the main difference between traditional UDA-PCSS and SF-PCSS setting. Referring to SHOT [5], we are given a source model $M_s : \mathcal{X}_s \rightarrow \mathcal{Y}_s$ on D_s with the standard cross-entropy loss:

$$\ell_s\left(M_s; \mathcal{X}_s, \mathcal{Y}_s\right) = -\mathbb{E}_{(x_s, y_s) \in \mathcal{X}_s, \mathcal{Y}_s} \sum_{k=1}^{K} q_k log\delta\left(M_s\left(x_s\right)\right) \tag{1}$$

where $\delta_k(a) = \frac{exp(a_k)}{\sum_i exp(a_i)}$ denotes the k-th element in the softmax output of a K-th dimensional vector a, and q_k is the one-of-K encoding of y_s.

The goal of SF-PCSS is to learn a target model $M_t : \mathcal{X}_t \rightarrow \mathcal{Y}_t$, only with the source model M_s and n_t unlabeled data $\left\{x_t^i\right\}_{i=1}^{n_t}$ from target domain D_t, then predict the labels $\left\{y_t^i\right\}_{i=1}^{n_t}$ of target data. The label space of source domain y_s and target domain y_t are the same, where $y \in Y \subseteq \mathbb{R}^K$ is the one hot ground-truth and K represents the number of the classes.

B. Bidirectional Pseudo-Label Selection

In this paper, we propose the Bidirectional Pseudo-Label Selection strategy to select amounts of reliable data with the correct pseudo-label for target model training, as illustrated in Fig.2. Based on the observation that the source model tends to predict accurate pseudo-label for the part of target data whose distribution is similar to the source domain, we use entropy to measure the uncertainty of model output for target data and divide it into source-similar data and source-dissimilar data. Since the distribution of source-similar data is more biased towards source data distribution, the uncertainty of model output for source-similar data is less, the pseudo-labels are more accurate. However, source-similar data is limited and doesn't fully represent the target domain's characteristics. Inspired by [17] that dynamically measures the difficulty of data and provides it for model training, we can select a part of reliable data with correct pseudo-label from source-dissimilar data to train the network. Next, we will introduce the details of dividing target data and selecting reliable data.

Firstly, based on entropy, we measure target data uncertainty and select source-similar data. The outputs of the source model for source-similar data are usually with lower entropy and more likely correct. On the contrary, the output of source-dissimilar data has high entropy and contains noise. The entropy value of the teacher model's output for the target data is calculated by:

$$E\left(x_t^i\right) = -\sum_{k=1}^{K} \delta\left(M^T\left(x_t^i\right)\right) log\delta\left(M^T\left(x_t^i\right)\right) , \tag{2}$$

where E is the Entropy function. Then, we set an entropy threshold λ to divide target data and select the part whose entropy value is less than threshold as the source-similar data:

$$\hat{y}_{x_t^i}^{s} = \begin{cases} 1, argmax\ P_{x_t^i} = k\ and\ E(x_t^i) < \lambda\ , \\ 0, otherwise\ , \end{cases} \tag{3}$$

where $P_{x_t^i} = \delta(M^T(x_t^i))$ is the prediction of M^T for target data x_t^i. The remainder with higher entropy than λ is divided into source-dissimilar data:

$$\hat{y}_{x_t^i}^{ds} = \begin{cases} 1, argmax\ P_{x_t^i} = k\ and\ E(x_t^i) >= \lambda\ , \\ 0, otherwise\ , \end{cases} \tag{4}$$

Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.

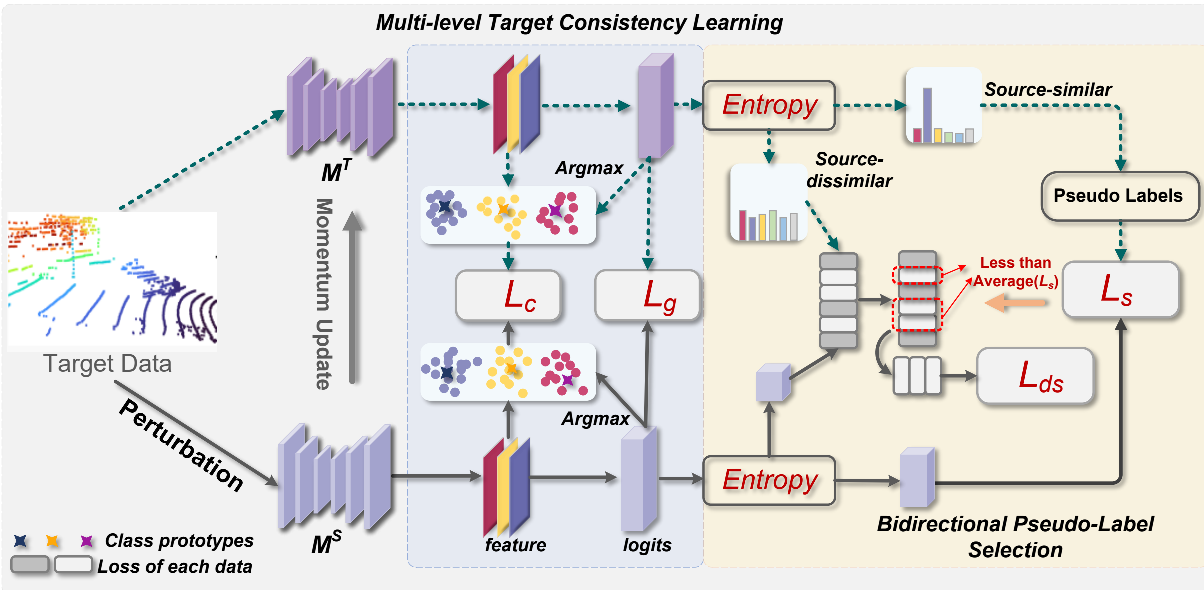

Fig. 2. Overview of the proposed Teacher-Guide Source Free (TGSF) framework for SF-PCSS.

where k is the k-th class of K. For source-similar data with reliable pseudo-labels, we utilize cross entropy loss for student model M^S training:

$$\mathcal{L}_s = -\sum_{i=1}^{n_t^s} \hat{y}_{x_t^i}^s log\delta\left(M^S\left(x_t^i\right)\right) \ . \tag{5}$$

To ensure the model comprehensive learning from the target data, we design a training scheduler η. It means with the increase of training iteration, M^S adapts to the target domain well, the model's prediction accuracy for target data is improving and we can choose more reliable data from source-dissimilar data set for M^S training by dynamically increasing threshold ρ_t:

$$\rho_t = \gamma^{-(t-1)}\rho \ , \tag{6}$$

where $\gamma > 1$ is a constant to control the rate of ρ_t changes, t is a function of training iterations, $t = \frac{\text{all iterations}}{2} -$ current iteration. But ρ_t cannot be increased all the time, its maximum value is ρ. Then, training scheduler is composed of ρ_t and ρ :

$$\eta = min(\rho_t, \rho) \ . \tag{7}$$

Now, it is important to estimate the value of ρ. Refer to the theoretical analysis [18], ρ can be estimated as averaged loss from the source-similar data in our task:

$$\rho = \frac{1}{n_t^s}\sum_{i=1}^{n_t^s} \hat{y}_{x_t^i}^s log\delta\left(M^S\left(x_t^i\right)\right) \ . \tag{8}$$

To select reliable data with smaller loss from source-dissimilar data, we firstly calculate cross entropy value for them:

$$\ell_{ce}^i = \hat{y}_{x_t^i}^{ds} log\delta\left(M^S\left(x_t^i\right)\right) \ . \tag{9}$$

Then, we choose reliable data from source-dissimilar data set which cross entropy value less than η combined to $\mathcal{L}_{ds}$:

$$\mathcal{L}_{ds} = -\sum \ell_{ce}^i \mathbb{1}(\ell_{ce}^i < \eta) \ . \tag{10}$$

where $\mathbb{1}$ is the indicator function.

C. Multi-level Target Consistency Learning

After applying the Bidirectional Pseudo-Label Selection strategy, we effectively obtain a portion of the target data with reliable pseudo-labels, which helps in reducing confirmation bias. However, further analysis reveals a significant challenge that existing SFDA methods based on the single model structure inevitably caused model degradation in PCSS task (Fig. 4(b)). During the adaptation stage, continual adjustments to the model parameters can lead to instability and even forgetting the learned source domain knowledge. This is not conducive to the model learning the target data with the help of source domain knowledge. To alleviate catastrophic forgetting, we propose teacher-guide networks with Multi-level Target Consistency Learning. In the teacher-guide framework, teacher model M^T not only provides reliable pseudo-labels for M^S, but also prevents the performance degradation of M^S by exploring target consistency from global and class prototype views.

In the teacher-guide networks, the input of the teacher branch is original target data x_t and the input $\tilde{x}_t$ of the student model is perturbed by scene wise transformation and point wise displacement. To aggregate global level semantic information, we use Kullback-Leibler divergence to enforce the constraint of target data consistency:

$$\mathcal{L}_g = \frac{1}{n_t}\sum_{i=1}^{n_t} KL\left(M^S(\tilde{x}_t^i) \,||\, M^T(x_t^i)\right) \ , \tag{11}$$

Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.

TABLE I

QUANTITATIVE COMPARISONS (MEAN INTERSECTION OVER UNION, MIOU) WITH UDA AND SFDA METHODS FOR POINT CLOUD SEMANTIC SEGMENTATION. THE BEST PERFORMANCE OF METHODS IS IN BOLD FONT.

methods	source data	usa→singapore	day→night	A2D2→SemanticKITTI
source only	×	46.5	41.2	35.9
MinEnt [19]	✓	47.0	43.5	38.0
FCNs [20]	✓	46.8	42.3	43.5
AdaptSegNet [21]	✓	47.7	44.6	44.3
CLAN [22]	✓	51.2	43.7	44.7
TPPDA [23]	×	46.4	41.7	41.2
SHOT [5]	×	46.7	45.2	41.1
GSSF [24]	×	45.9	41.3	38.2
GPUE [25]	×	47.6	44.2	42.7
TGSF	×	**51.7**	**47.9**	**44.4**

where $M^S(\tilde{x}_t^i)$ is the predicted probability of M^S to the perturbed target data $\tilde{x}_t$, $M^T(x_t^i)$ is the predicted probability of M^T to the original target data x_t.

Since the class prototype is insensitive to noisy labels and treats different classes equally when data is class imbalance (in PCSS task, background points are much more than other classes), we hope to use class prototypes to explore semantic information of different categories between original and perturbed data. Firstly, we compute the feature class prototypes for the two branches:

$$\tilde{c}_{(k)} = \frac{\sum_i f^S\left(\tilde{x}_t^i\right) * \mathbb{1}\left(\tilde{y}_t^{(i,k)} == 1\right)}{\sum_i \mathbb{1}\left(\tilde{y}_t^{(i,k)} == 1\right)}\ , \tag{12}$$

$$c_{(k)} = \frac{\sum_i f^T\left(x_t^i\right) * \mathbb{1}\left(y_t^{(i,k)} == 1\right)}{\sum_i \mathbb{1}\left(y_t^{(i,k)} == 1\right)}\ , \tag{13}$$

where $\mathbb{1}$ is the indicator function, y_t and $\tilde{y}_t$ are the labels predicted by M^T and M^S, f^t and f^s are feature extractor respectively.

Then, we introduce $\mathcal{L}_c$ to reduce the distance between prototypes of the same category and expand the distance between prototypes of different categories:

$$\mathcal{L}_c = \frac{1}{K}\sum_{\mathrm{i=j}} KL\left(\tilde{c}_{\mathrm{i}} \,||\, c_{\mathrm{j}}\right) - \frac{1}{K*(K-1)}\sum_{\mathrm{i\neq j}} KL\left(\tilde{c}_{\mathrm{i}} \,||\, c_{\mathrm{j}}\right)\ , \tag{14}$$

where K and $K*(K-1)$ denote to balance the two part that calculate K times KL divergence for same category and $K*(K-1)$ times for different categories.

In summary the overall optimization goal for SF-PCSS is stated as:

$$\mathcal{L} = \mathcal{L}_s + \omega\mathcal{L}_{ds} + \mathcal{L}_g + \mu\mathcal{L}_c\ , \tag{15}$$

where ω and μ are balancing hyper-parameters.

IV. EXPERIMENTS

A. Implementation Details

In our framework, we use SparseConvNet with U-Net architecture that downsamples six times to process point cloud. For fair comparisons, all the methods are trained and evaluated by using the same backbone on one NVIDIA RTX 3090 with 24GB RAM and we use PyTorch to implement our method. We train the network with a batch size of 8 and the Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$. We set $\omega = 0.5$,$\mu = 0.1$,$\gamma = 1.01$,$\lambda = 0.01$ for all experiments except $\lambda = 0.5$ for Day→Night benchmark. We use the exponential moving average scheme to gradually update the parameters of M^T. For each experiment, the tables display the average results obtained under three random seeds.

B. Comparison with State-of-the-art Methods

In Table I, our method achieves significant improvements of 5.2%, 6.7%, 8.5% mIoU over source only on three benchmarks, respectively. Source only is the commonly used baseline in the SFDA work, it means directly using source model to test target data. Compared with state-of-the-art SFDA methods, our method can perform better and improved by 4.1%, 2.7%, 1.7% over the second best SFDA methods on three different benchmarks. It is worth noting that our method even performs better than the UDA methods that with the help of source data. Our method outperforms CLAN by 0.5% on the usa→singapore and improved by 3.3% than AdaptSegNet on day→night.

TABLE II

ABLATION STUDY (MIOU, %) ON EFFECT OF TGSF IN THE DAY→NIGHT AND A2D2→SEMANTICKITTI BENCHMARKS.

Loss/Dataset	Day/Night	A2D2/SemKITTI
source only	41.2	36.6
$\mathcal{L}_s$	43.2	37.5
$\mathcal{L}_s + \mathcal{L}_{ds}$	44.5	39.8
$\mathcal{L}_s + \mathcal{L}_{ds} + \mathcal{L}_g$	46.8	43.7
$\mathcal{L}_s + \mathcal{L}_{ds} + \mathcal{L}_g + \mathcal{L}_c$	47.9	44.4

C. Ablation study on loss function

As shown in Table II, to demonstrate the effect of each component, we perform ablation studies in loss function $\mathcal{L}$ on Day→Night and A2D2→SemanticKITTI. The results demonstrate that the BPLS strategy and MTCL module are effective for model performance improvement.

D. Visualization results

As shown in Fig.3, on A2D2→SemanctiKITTI, the baseline and SHOT incorrectly classify the big area of nature as a

Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.

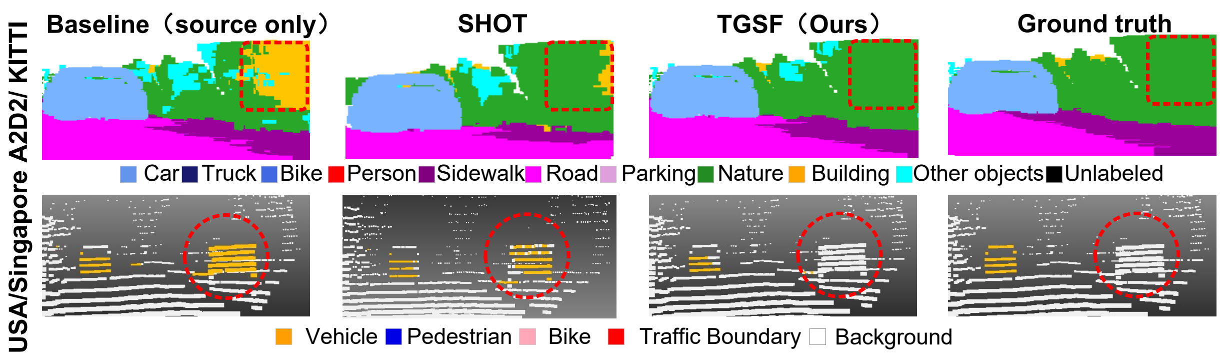

Fig. 3. Qualitative results of baseline, SHOT and TGSF on the three adaptation benchmarks: A2D2→SemanticKITTI, USA→Singapore. We utilize the red box to distinguish the differences between the visualization results.

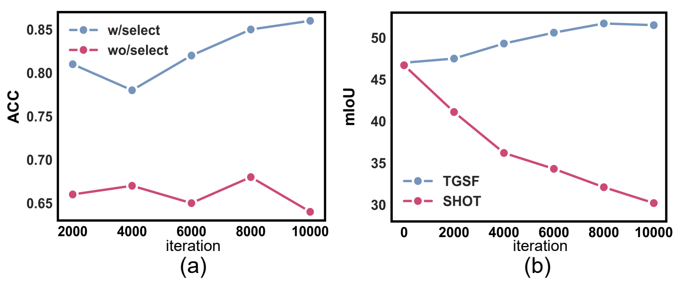

Fig. 4. Analysis on pseudo-label accuracy (a) and model stability (b).

building, but TGSF does not. And we can observe that TGSF is good for classifying the edge part of the two categories. On USA→Singapore, TGSF correctly classifies the vehicle and background, but the baseline does not.

TABLE III
QUANTITATIVE COMPARISONS WITH DIFFERENT PSEUDO LABEL SELECTION STRATEGIES.

Method	USA→Singapore	Day→Night
source only	46.5	41.2
Confidence	45.7	43.2
Entropy	48.6	43.2
TPPDA [23]	46.4	41.7
GSSF [24]	45.9	41.3
BPLS	**50.6**	**44.5**

E. Analysis on Bidirectional Pseudo-Label Selection

In this section, we analyze and verify the effectiveness of bidirectional pseudo-label selection. We calculate the pseudo-label accuracy of source-dissimilar data that with or without training scheduler select. In the adaptation stage, we collect pseudo-labels for each valid point and measure the average accuracy over all categories from 2000 to 10000 iterations. In Fig. 4(a), we observe that with the training scheduler select, student model M^S can obtain more accurate data for training. As shown in Table III, we also compare it with other different pseudo-label strategies: confidence means using the maximum confidence predicted by the model to generate pseudo labels, entropy means means using the entropy to measure the model's predictions and generate pseudo labels. TPPDA [23] and GSSF [24] are existing SFDA methods. It can be seen that BPLS achieves the best performance compared to other pseudo-label strategies.

F. Stability of TGSF

To verify the stability of TGSF, we show the performance of SHOT and TGSF in Fig. 4(b). It is obvious that SHOT occurs severe model degradation during model training stage. As the model parameters are adjusted during the training stage, SFDA methods based on a single model structure become unstable due to catastrophic forgetting. On the contrary, it is obvious that TGSF with MTCL can effectively alleviate model degradation. Because the target model can obtain more target data information and source knowledge from the teacher model by exploring multi-level consistency.

V. Conclusion

We propose a challenge setting: Source-Free domain adaptation for Point Cloud Semantic Segmentation which considers the inaccessibility of source data in privacy protection scenarios. We first identify the limitations of existing source free domain adaptation methods for SF-PCSS which are confirmation bias and catastrophic forgetting. Then, we propose an effective Teacher-Guide Source Free framework that contains BPLS strategy and MTCL module. Extensive experimental results demonstrate that TGSF achieves significant performance and we show the effectiveness of BPLS and MTCL by providing extensive ablation studies and analysis.

Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grant Grant 62176224, Grant 62222602, Grant 62306165; in part by the China Postdoctoral Science Foundation under Grant 2023M731957; and in part by the China Computer Federation (CCF)-Lenovo Blue Ocean Research Fund.

Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.

References

[1] Andres Milioto, Ignacio Vizzo, Jens Behley, and Cyrill Stachniss, "Rangenet++: Fast and accurate lidar semantic segmentation," in *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2019, pp. 4213–4220.

[2] Zijin Du, Hailiang Ye, and Feilong Cao, "A novel local-global graph convolutional method for point cloud semantic segmentation," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[3] Bichen Wu, Xuanyu Zhou, Sicheng Zhao, Xiangyu Yue, and Kurt Keutzer, "Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 4376–4382.

[4] Yikai Bian, Le Hui, Jianjun Qian, and Jin Xie, "Unsupervised domain adaptation for point cloud semantic segmentation via graph matching," *arXiv preprint arXiv:2208.04510*, 2022.

[5] Jian Liang, Dapeng Hu, Yunbo Wang, Ran He, and Jiashi Feng, "Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8602–8617, 2021.

[6] Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, et al., "Attracting and dispersing: A simple approach for source-free domain adaptation," in *Advances in Neural Information Processing Systems*, 2022.

[7] Ian J Goodfellow, Mehdi Mirza, Da Xiao, Aaron Courville, and Yoshua Bengio, "An empirical investigation of catastrophic forgetting in gradient-based neural networks," *arXiv preprint arXiv:1312.6211*, 2013.

[8] Jian Liang, Dapeng Hu, and Jiashi Feng, "Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation," in *International Conference on Machine Learning*. PMLR, 2020, pp. 6028–6039.

[9] Rui Li, Qianfen Jiao, Wenming Cao, Hau-San Wong, and Si Wu, "Model adaptation: Unsupervised domain adaptation without source data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 9641–9650.

[10] Yuang Liu, Wei Zhang, and Jun Wang, "Source-free domain adaptation for semantic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 1215–1224.

[11] Vinod K Kurmi, Venkatesh K Subramanian, and Vinay P Namboodiri, "Domain impression: A source data free domain adaptation method," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2021, pp. 615–625.

[12] Yunzhong Hou and Liang Zheng, "Visualizing adapted knowledge in domain transfer," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 13824–13833.

[13] Ning Ding, Yixing Xu, Yehui Tang, Chao Xu, Yunhe Wang, and Dacheng Tao, "Source-free domain adaptation via distribution estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 7212–7222.

[14] Zhen Qiu, Yifan Zhang, Hongbin Lin, Shuaicheng Niu, Yanxia Liu, Qing Du, and Mingkui Tan, "Source-free domain adaptation via avatar prototype generation and adaptation," *arXiv preprint arXiv:2106.15326*, 2021.

[15] Shiqi Yang, Joost van de Weijer, Luis Herranz, Shangling Jui, et al., "Exploiting the intrinsic neighborhood structure for source-free domain adaptation," *Advances in neural information processing systems*, vol. 34, pp. 29393–29405, 2021.

[16] Jiaxing Huang, Dayan Guan, Aoran Xiao, and Shijian Lu, "Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data," *Advances in Neural Information Processing Systems*, vol. 34, pp. 3635–3649, 2021.

[17] Xin Wang, Yudong Chen, and Wenwu Zhu, "A survey on curriculum learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 4555–4576, 2021.

[18] Yi Xu, Lei Shang, Jinxing Ye, Qi Qian, Yu-Feng Li, Baigui Sun, Hao Li, and Rong Jin, "Dash: Semi-supervised learning with dynamic thresholding," in *International Conference on Machine Learning*. PMLR, 2021, pp. 11525–11536.

[19] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez, "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2517–2526.

[20] Judy Hoffman, Dequan Wang, Fisher Yu, and Trevor Darrell, "Fcns in the wild: Pixel-level adversarial and constraint-based adaptation," *arXiv preprint arXiv:1612.02649*, 2016.

[21] Yi-Hsuan Tsai, Wei-Chih Hung, Samuel Schulter, Kihyuk Sohn, Ming-Hsuan Yang, and Manmohan Chandraker, "Learning to adapt structured output space for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7472–7481.

[22] Yawei Luo, Liang Zheng, Tao Guan, Junqing Yu, and Yi Yang, "Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2507–2516.

[23] Youngeun Kim, Donghyeon Cho, and Sungeun Hong, "Towards privacy-preserving domain adaptation," *IEEE Signal Processing Letters*, vol. 27, pp. 1675–1679, 2020.

[24] Jonghyun Lee, Dahuin Jung, Junho Yim, and Sungroh Yoon, "Confidence score for source-free unsupervised domain adaptation," in *International Conference on Machine Learning*. PMLR, 2022, pp. 12365–12377.

[25] Alessio Litrico and Pietro Morerio, "Guiding pseudo-labels with uncertainty estimation for source-free unsupervised domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 7640–7650.

Authorized licensed use limited to: North China Electric Power University. Downloaded on April 17,2025 at 12:55:04 UTC from IEEE Xplore. Restrictions apply.