UNSW SYDNEY

# Source-free Few-shot Segmentation for Rarer Brain Tumors

**Author:**
Yan, S; Liu, S; Di Ieva, A; Pagnucco, M; Song, Y

**Publication details:**
Proceedings of the International Joint Conference on Neural Networks
v. 00
pp. 1 - 8
2161-4393 (ISSN); 2161-4407 (ISSN)

**Event details:**
2025 International Joint Conference on Neural Networks (IJCNN)
2025-06-30 - 2025-07-05

**Publication Date:**
2025-01-01

**Publisher DOI:**
https://doi.org/10.1109/IJCNN64981.2025.11228625

**License:**
https://creativecommons.org/licenses/by/4.0/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/106746 in https://unsworks.unsw.edu.au on 2026-02-27

# Source-free Few-shot Segmentation for Rarer Brain Tumors

Anonymous

***Abstract*—Since the inception of BraTS challenge, a series of methods has been developed for brain tumor segmentation over the past years. Although these methods achieved promising results, they mostly focus on glioma segmentation, largely due to their relatively high incidence. These fully-supervised methods may not be applicable as they rely on abundant labeled data, which is intrinsically inaccessible for rarer types of brain tumors. Data-efficient transfer learning approaches like few-shot learning and domain adaptation assume full access to source data, which may not be feasible in real-life scenarios due to privacy and confidentiality concerns. In this work, we propose a new *source-free few-shot learning* framework for rarer brain tumor segmentation that adapts source model trained on gliomas to other less common brain tumors such as meningioma, metastasis and pediatric tumors with only a few labeled target data. The proposed framework follows a dual-branch prototypes learning structure that harmonize preservation of common knowledge from source class and learning new features from target. We show that our method gains a 6% increase in Dice score over representative source-free domain adaptation methods, and achieves comparable performance against its fully-supervised counterpart.**

***Index Terms*—3D UNet, brain tumor segmentation, few-shot learning, source-free**

## I. Introduction

Brain tumors can lead to high rates of mortality and severe disabling effects on patients. Glioblastomas, the most common primary brain cancer of glial origin, are almost universally fatal within 2 years of diagnosis despite medical and surgical therapy [1]. Effective treatment is multimodal and relies on accurate diagnosis, and diagnosis requires accurate interpretation of radiological data, among them most commonly magnetic resonance imaging (MRI). Recently, new techniques represented by deep learning have been used to ease the burden of diagnosis. As the focus of this study, brain tumor MRI segmentation allows precise tumor boundary identification and benefits many clinical applications, such as surgical treatment planning and tumor growth monitoring.

Despite its promising benefits, accurate tumor segmentation is a complex task impeded by tumors' varying size, shape, location and complex substructures. Aside from tumors' intrinsic characteristics, data used for developing models is typically acquired across different centers and under varying acquisition protocols, which further hinders a universally effective solution. Methods in this field predominantly adopt UNet-based architectures [2] that combine encoder-decoder design with skip connections to effectively integrate multi-level semantic features while preserving fine-grained details. Many improvements have been made for brain tumor segmentation tasks, such as attention mechanisms [3], and region-based training approaches [4]. Parallel to these developments, other methods [5], [6] have begun leveraging Transformers [7] for their superior capability in modeling long-range dependencies.

Most studies for brain tumor segmentation focus on gliomas, but it is not trivial to directly translate them to other rarer types of tumors, such as meningioma, metastasis and pediatric tumors, as they exhibit different characteristics from gliomas that may lead to a significant performance drop. As shown in Fig.1, gliomas generally arise within the brain's internal structures, whereas meningiomas typically grow on the boundary of brain tissue as they come from external membranous covering of the brain. Metastasis may occur in multiple instances with smaller lesions less than 5mm. Pediatric brain tumors are located in the pons and their necrotic tissue regions are less common or clear. These distinctions pose challenges when transferring a model from glioma cases to segment other types of tumors, as models developed for gliomas tend to look for a single large tumor without particular location priors.

A common approach could be fine-tuning a model trained on a source task with sufficient labeled target data. The shared features between source class and target class help the training process become smoother compared with learning from scratch. However, it assumes there will be enough labeled target data at hand, an assumption hard to realize both because of the data scarcity of less common brain tumors and the costly nature of labeling distinctive features among tumor types. With a few training data available, possible solutions represented by few-shot learning and domain adaptation typically require training with source data. However, this is impractical in real-life scenarios as it is neither easy to access source data due to concerns of personal privacy and confidentiality, nor plausible to store the entire source dataset on devices with limited resources for training purposes.

Seeking to reduce the reliance on source data, we investigate a new *source-free few-shot learning* setting for rarer brain tumor segmentation that only requires access to a pre-trained model without the source data. We simulate the source-free few-shot segmentation of less common brain tumors by taking gliomas as the source class and meningioma, metastasis and pediatric brain tumors as target classes, and aim to achieve good generalization ability with a few annotated target cases.

In this work, we propose a source-free few-shot segmentation network for rarer brain tumors. Without access to sufficient source data but only a pre-trained source model, we designed a confidence-based dual-branch learning method. As brain tumors share some common traits and the brain structure

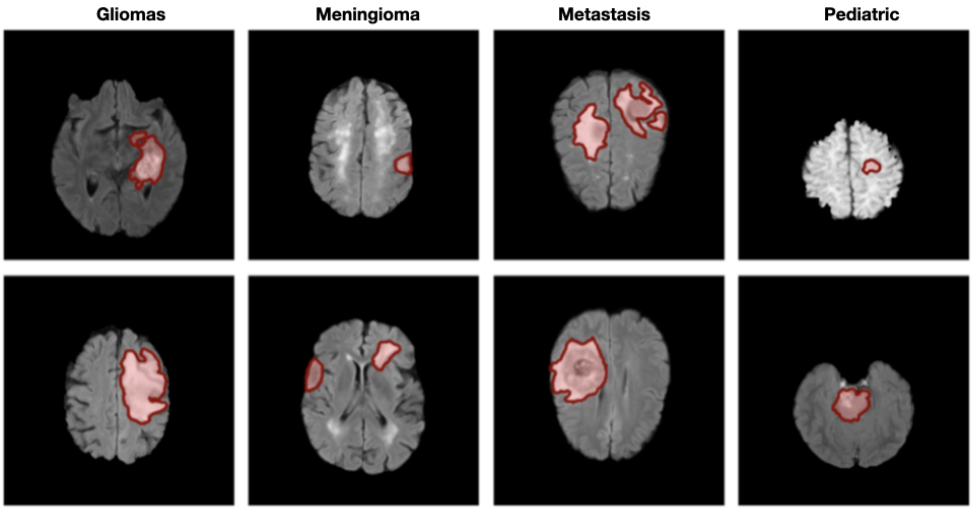



Fig. 1. Comparison of gliomas, meningiomas, metastasis and pediatric brain tumors. Whole tumors are highlighted. Different type of brain tumors vary in size, location and distributions.

remains the same, the model trained on source tumors can also offer robust and reliable initial predictions for target tumors. We retain the shared knowledge learned from source domain by updating the source adaptation branch using the highly confident predictions of target tumors. However, given that tumors of different type exhibit distinctive features, a target learning branch is dedicated to specifically learn from less confident target samples and capture the unique target features. As training exclusively on the few-shot labeled target data can lead to overfit easily, we incorporate unlabeled target data into the training by a prototype-based self-training framework. The network extracts representative prototypes for foreground and background, and generates the pseudo labels for unlabeled data by calculating the cosine similarity between features and prototypes. The model maintains a robust set of prototypes by updating them at every iteration with the new prototypes. In the end, we unify the results of two branches at the final stage of the training by minimizing the discrepancy between the two outputs.

Our contributions are summarized as follows:

1) We propose our Confidence-based Dual-branch Learning (CDL) method that first divide the target domain data according to the output confidence of the source model, and then conquer them through either the source adaptation branch or target learning branch. The dual-branch design retains the knowledge from source domain while learns the new features from target domain.
2) We present a Prototype-based Self-training Framework (PSF) that dynamically updates the prototype memory bank and generates pseudo labels for unlabeled target data. With only a few examples, the class prototypes can effectively represent different classes and remain consistent and robust through updates.
3) The efficacy of our method is demonstrated through extensive experiments on three different types of rare brain tumors. Our experiment results showcase the superior performance of our methods against other state-of-the-arts.

## II. Related Works

### A. Brain Tumor Segmentation

Early attempts [8], [9] to tackle the problem of brain tumor segmentation process volumes are based on single slices, due to the computational constraints. These attempts are natural extensions from effective model structures in general images, such as FCN [10]. If computational resources allow, most 2D models can be extended to 3D models by replacing the 2D convolutional kernels with 3D ones, which strengthens the inter-slice consistency. As a remedy to exorbitant memory increase, model takes sampled 3D cubes as input instead of the whole brain volumes.

A common yet effective model design for brain tumor segmentation is inspired by UNet [2], an encoder-decoder structure that extracts both local and semantic features in a hierarchial manner. The model perform segmentation by exploiting the high-level semantic information while taking into account the local properties in low-levels. U-Net's success inspired numerous architectural improvements. Dense U-Net [11] enhanced the network's capability by implementing densely connected convolutions to better propagate and reuse features. U-Net++ [12] bridged the gap between encoder and decoder features, enabling simultaneous processing of both simple surface-level features and complex abstract ones. The field further evolved with innovations like atrous convolution [13] and pyramid pooling [14], which expanded the networks' ability to understand more context and represent features at multiple scales. However, despite these advances, these models struggle when faced with situations where labeled training data is scarce.

### B. Few-shot Medical Image Segmentation

Few-shot learning investigates model's ability to understand new classes with a number of examples, in contrast with the sufficient amount of training examples commonly seen in the traditional supervised learning paradigm. While it is still infeasible to train a well-performing models within limited examples, some studies [15] suggest it is possible to reduce the training samples by training in a principled manner, episodic training. Episodic training rearranges the way data is fed into the model, where each episode focuses on learning a different class. In this way, the goal of training shifts from capturing characteristic class features distribution to utilising new examples effectively. With limited training examples for each class, one way to infer test examples is to learning representative features collectively from training examples, known as prototypes. Prototype learning [16] first gathers certain amount of prototypes from training examples, then matches prototypes with features from test examples. Some works [17]–[19] improve model's strategy to learn prototypes. Other works [20] propose practical matching mechanisms between query and prototypes.

### C. Domain Adaptation

Many researches [21], [22] have explored using deep learning to transfer knowledge from source to target domains, using both supervised and unsupervised approaches. While supervised methods can be effective, they require extensive labeled data that is time-consuming and labor-intensive to collect. This has led to increased interest in unsupervised domain adaptation, which can be categorized into three main approaches: feature alignment based on difference measurement, domain feature disentanglement-based methods, and approaches that combine pixel-level and feature alignment using Generative Adversarial Networks (GANs) [23].

Adversarial learning has emerged as a powerful tool for addressing domain differences at multiple levels - from raw images to abstract features. GANs introduced the ability to transform images between source and target domains. This capability was enhanced by CycleGAN [24], which enabled unpaired image transformation while preserving structural integrity through cycle consistency loss. Building on this, researchers demonstrated practical applications by using CycleGAN to convert CT images to MR images, combining the synthetic data with a small set of real MR images for semi-supervised tumor segmentation. However, while these methods successfully generate high-quality transformed images, they lack explicit semantic constraints to keep structural elements consistent during the transformation process.

## III. Methodology

### A. Problem Definition

Given a source model $\mathcal{M}_s$ fully trained on abundant labeled support images of glioma data $\mathcal{D}_s^l = \left\{(x_s^l, y_s^l)_i\right\}_{i=1}^{N_s}$, and a set of few-shot target labeled data $\mathcal{D}_t^l = \left\{(x_t^l, y_t^l)_i\right\}_{i=1}^{N_l}$ and unlabeled target images $\mathcal{D}_t^u = \left\{(x_t^u)_i\right\}_{i=1}^{N_u}$ of less common brain tumors ($N_l << N_s$) , we aim to find a mapping $\mathcal{F} = f(X_m; \mathcal{M}_s, \mathcal{D}_t^l, \mathcal{D}_t^u)$ that maps input $X_t$ to pixel-wise segmentation masks $Y_t$. Note this source-free few-shot learning setup is different from conventional few-shot learning, which focus on model's learn-to-learn ability for unseen target classes in training. This setup also differs from domain adaptation where the classes remain the same between source and target domains, whereas in our setting, source images and target images are of different class of tumors.

### B. Confidence-based dual-branch learning

Our proposed framework is shown in Fig.2, which follows a dual-branch structure that focuses on retaining source knowledge and learning target features, respectively. The two branches are both in the form of 3D ResUNet and share a feature encoder, followed by different decoders that output the final segmentation results. Both the encoder and decoders are first initialized with the weight of pre-trained source model $\mathcal{M}_s$. We train the model first with labeled target data $\mathcal{D}_t^l$ and then with unlabeled target data $\mathcal{D}_t^u$.

The two-branch structure divides target data into two groups and explicitly learns the intra-class discrepancy from source-alike samples and target-alike samples. Through a first pass on the source model, input data is partitioned based on the confidence ranks of the outputs. The half samples with higher confidence that are similar to source class are fed into the source adaptation branch. In contrast, the lower half samples are passed through target learning branch, which focuses on learning target class-specific characteristics.

As target brain tumors share some common characteristics with source brain tumors, target data that are similar to source tumors can benefit directly from the knowledge retained in the source model. As for the target data that exhibit different characteristics from source tumors, the target learning branch will focus on learning the target-specific features. This division follows a divide-and-conquer strategy and allows efficient adaptation. Compared to a single branch design which strives to balance the knowledge from source and new features from target, the dual branch model retains the knowledge learned from source to a large extent, and learns the target-specific characteristics independently.

On this stage, we freeze the shared encoder and only update the decoder part of each branch with Dice loss $L_{Dice}$.

### C. Prototype-based self-training framework

The utilization of unlabeled data is crucial in the source-free few-shot learning setting, as without access to source data, learning on the few shot target data alone easily leads to overfit. For the unlabeled target data $x_t^u$ without ground truth labels, we adopt self-training and obtain their dice loss by generating pseudo labels as ground truths through a progressive prototype learning process. This process is inspired by prototype-based few-shot learning [25], where prototypes can efficiently represent different classes with access to only a few samples. We take the source adaptation branch as an example to illustrate the prototype learning process as the

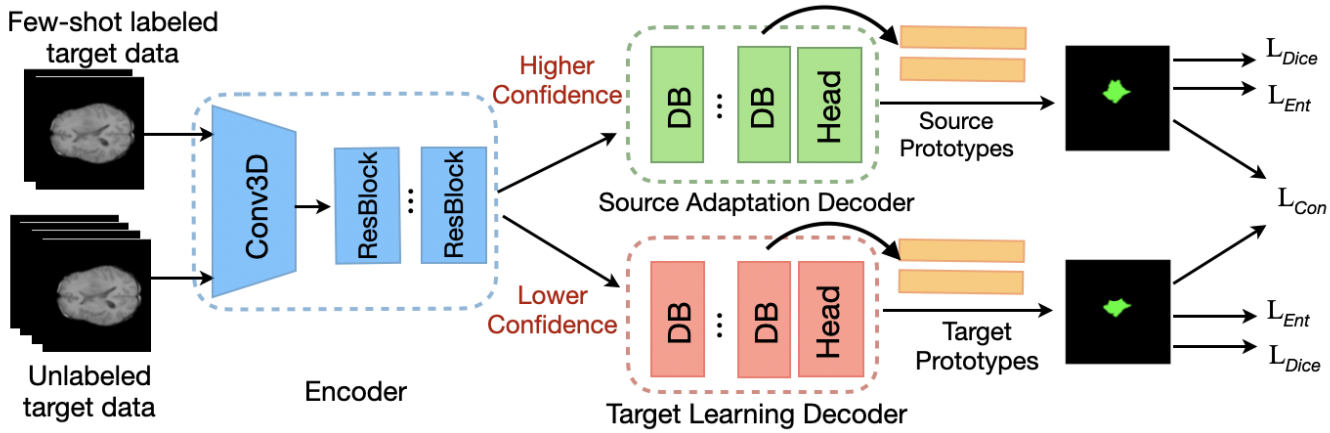



Fig. 2. The proposed *source-free few-shot learning* strategy. DB stands for decoder block. ResBlock stands for residual block. The model takes few-shot labele target data and unlabeled target data as input. Based on the initial prediction confidence, samples with higher confidence are fed into the source adaptation decoder, while the lower confidence ones are processed through target learning decoder. For unlabeled target data, the model also generate pseudo labels for Dice loss with a set of dynamically updated prototypes.

other branch follows the same procedure. During training, each branch maintains a memory bank of two representative prototypes, namely the foreground prototype and the background prototype. These prototypes are first initialized by few-shot labeled target data $x_t^l$ with the following procedure. We obtain the background prototype $\mathcal{P}_0$ and foreground prototype $\mathcal{P}_1$ by calculating the masked weighted average features derived from the binary ground truth mask $\mathcal{T}$ and feature map $\mathcal{F}$ from the output of second-to-last layer of decoder before the prediction head, i.e:

$$\mathcal{P}_1 = \frac{\sum_x \mathcal{F}_x \mathcal{T}_x}{\sum_x \mathcal{T}_x}, \mathcal{P}_0 = \frac{\sum_x \mathcal{F}(1-\mathcal{T}_x)}{\sum_x (1-\mathcal{T}_x)}, \quad x \in \{x^l\}.$$

The pseudo label $\mathcal{Y}_p$ can be obtained by classifying each voxel $i$ into the more similar prototype $i$ through cosine distance, i.e:

$$\mathcal{Y}_p = \underset{k}{\operatorname{argmin}} \cos(\mathcal{F}_i, \mathcal{P}_k), \quad k \in \{0, 1\}.$$

We further utilize unlabeled data for robust and consistent prototypes generation. These prototypes will be updated per batch by calculating the exponential moving average (EMA) using the current batch prototypes $\mathcal{P}_c$ with a coefficient $\theta$, i.e:

$$\mathcal{P}_k^t = \theta \cdot \mathcal{P}_k^{t-1} + (1-\theta) \cdot \mathcal{P}_c,$$

where $\mathcal{P}_c$ can be derived for unlabeled data by simply replacing the ground truth mask to pseudo label.

### D. Loss Function

We train the model using both labeled data with ground truth and unlabeled data with pseudo labels. For labeled data, we adopt dice loss $L_{Dice}$ as supervised loss. For unlabeled data, we treat pseudo label generated by prototype matching as ground truth for dice loss and introduce an additional loss term $L_{Ent}$ to minimize the entropy of the network's raw prediction $\hat{\mathcal{Y}}$,

$$\mathcal{L}_{Ent} = -\sum_x \hat{\mathcal{Y}} \log \hat{\mathcal{Y}}.$$

At the final stage of training, we integrate the complementary information learned through both branches to make comprehensive predictions. We achieve this goal by further incorporating a loss term $L_{Con}$to minimize the discrepancy between output $\hat{\mathcal{Y}}_s$ from support adaptation branch and output $\hat{\mathcal{Y}}_t$ from target learning branch, while freeze decoders and only update the shared encoder, i.e:

$$\mathcal{L}_{\text{Con}} = \sum_x \|\hat{\mathcal{Y}}_s - \hat{\mathcal{Y}}_t\|_2.$$

The overall loss for training our model is

$$\mathcal{L} = \mathcal{L}_{\text{Dice}} + \mathcal{L}_{\text{Ent}} + \mathcal{L}_{\text{Con}}.$$

## IV. Experiment

### A. Data and Evaluation Metrics

We evaluate our proposed method on the public available dataset provided by the Brain Tumor Segmentation (BraTS) 2023 challenge [26], which contains 1251 glioma cases, 1000 meningioma cases, 238 metastasis cases, and 100 pediatric brain tumor cases. Every case is composed of four brain MRI sequences, namely pre-contrast T1-weighted, post-contrast T1-weighted, T2-weighted, and T2-weighted Fluid Attenuated Inversion Recovery (FLAIR) series. All four modalities are used in the proposed method, being co-registered to the same anatomical template, interpolated to the same resolution of $240 \times 240 \times 155$ voxels and skull-stripped. The segmentation accuracy is measured by the Dice score, and the Hausdorff distance (HD, 95%) metrics for the region of whole tumor.

### B. Implementation Details

We use ResUNet-50 as our network backbone. The source model is fully trained on glioma data. In the adaptation stage, for each tumor type, we keep 70% of cases for training, 10% for validation and 20% for testing. From training set, we randomly choose 10 cases as labeled training data, and the remaining as unlabeled data. We repeat this process for three times for all the experiments. All the MRI data was cropped into $128 \times 128 \times 128$ patches. Common data augmentation techniques such as rotation, scaling, brightness, Gamma augmentation were applied. Adam optimizer and cosine annealing scheduler were deployed. The initial learning rate was set as 1e-4. $\theta$ is set as $0.9$ for labeled cases and $0.99$ for unlabeled ones. Our method was implemented on the PyTorch framework with a single GPU (NVIDIA RTX 4090). The maximum epoch is set to 500. During testing, the branch with higher confidence is selected as the final prediction.

### C. Evaluation of Different Training Strategies

In this section, we compare our method with other training methods visually and quantitatively. Two oracle models are constructed to assess the upper bound performance of our method. The Oracle transfer learning (TL) model is acquired by training firstly on gliomas and then fine-tuned on abundant target tumors. The Oracle Scratch model is directly trained on abundant target tumors from scratch. For comparison with source-free adaptation methods, we re-implemented two representative methods, including intra-domain adversarial alignment ($A^2$Net) [27] and self-training with pseudo labeling (NRC) [28] method. For comparison with similar few-shot learning methods, we implemented few-shot TL by first training the model on gliomas and then fine-tuning it on few-shot target tumors. We also include comparisons with SIGN [29], a state-of-the-art method for few-shot transfer learning.

Quantitative results are shown in Table 1. Oracle methods indicates the upper bound of all the methods, where transfer learning shows a superior performance than training from scratch. Out of all the other methods, Few-shot TL scores the lowest Dice score of 40.38%. This is possibly because of the few labeled target data is too few to shift the learned feature distribution from a biased source class. Both $A^2$Net and NRC showed decent performances with an average Dice scores of 66.62% and 66.82% respectively. This shows that that adversarial alignment can adapt well to the target class by first explicitly learning the differences and then minimizing their discrepancies. It should be noted that our proposed method outperforms all the other methods in terms of Dice score and come very close to SIGN for the HD score. Our method also achieved a comparable performance against the oracle methods, while only require a few labeled supervision.

One of the challenging aspect of brain tumor segmentation is the small tumors. Smaller tumors are hard to locate, esay to miss and can reflect the overall robustness of methods. We select several cases with small tumors in Figure 3. It demonstrates intuitively how different methods perform when segmenting small rarer brain tumors. For this challenging task, all methods can roughly locate the tumors but struggle to precisely delineate tumor boundaries. This can also be reflected in their high HD scores compared to oracle methods. Notably, A²Net and NRC made false positives predictions outside the ground truth area in the case 1, whereas in the case 3, A²Net and NRC were overly conservative and made false negative predictions. In case 4, A²Net and NRC falsely incorporate part of the background area as their prediction and fail to follow the edge of the brain tissue. SIGN missed the tumor completely in case 1, but aligned with the boundary of the ground truths relatively better. Our approach achieved consistently superior predictions with accurate location and shape, as it combines task-specific features from target class with common features from source class.

### D. Ablation Study

In this section, we conducted ablation study to assess the effectiveness of our proposed Confidence-based Dual-branch Learning (CDL) and Prototype-based Self-training Framework (PSF) method. The results of the ablation study were summarized in Table II.

Without inclusion of CDL and PSF, the network revert to a simple single-branch 3D-UNet. The network takes every target sample and updates regradless of the initial prediction result from the source model. As for the unlabeled samples, the model only benefits from the entropy loss. The segmentation performance is the lowest among the four different variants, with an averge Dice score of 47.8% and HD score of 1.602. With the addition of CDL, the segmentation performance improved significantly with a Dice score of 64.15% and HD score of 1.218. This showcased that the model benefited from the dual-branch learning method. The counterpart single-branch model may not be able to retain the robust knowledge from source because it prioritize the new updates from target domain, which could contain greater noise and be less robust as the supervision from target domain is very limited. On the other hand, the inclusion of PSF also leads to a decent improvement of the segmentation performance, with a Dice score of 60.41% and HD score of 1.38. As a data-efficient approach, the dynamically updated prototypes track the feature distribution of foreground and background classes, and reflect robust representations of target tumors and normal tissues. The pseudo labels generated based on these prototypes enables additional supervision from unlabeled data. The most significant improvement was observed when both CDL and PSF were included. The model achieved an average Dice score of 72.87% and HD score of 0.893. This result also demonstrate the two components work well together and have a synergistic impact.

## V. Conclusion

We proposed a source-free few-shot learning framework for rarer brain tumor segmentation with gliomas as source class. With only access to a pre-trained source model and no source data, we gradually adapt the model from source to target by retaining the knowledge from source model in one

TABLE I
SEGMENTATION PERFORMANCE OF TARGET TUMORS

| Methods | Metastasis | | Pediatric Tumors | | Meningiomas | | Average | |
|---|---|---|---|---|---|---|---|---|
| | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] |
| Oracle TL | 72.56 | 0.943 | 88.35 | 0.812 | 82.32 | 0.594 | 81.08 | 0.783 |
| Oracle Scratch | 68.41 | 1.108 | 80.10 | 0.935 | 81.65 | 0.603 | 76.72 | 0.882 |
| $A^2$Net | 61.33 | 1.141 | 78.74 | 1.374 | 59.79 | 0.686 | 66.62 | 1.064 |
| NRC | 59.61 | 1.203 | 72.12 | 1.758 | 68.72 | **0.619** | 66.82 | 1.194 |
| Few-shot TL | 40.38 | 1.63 | 58.16 | 1.571 | 42.73 | 1.120 | 47.09 | 1.440 |
| SIGN | 61.08 | 1.179 | 80.81 | 1.056 | 70.67 | 0.684 | 70.85 | 0.907 |
| Ours | **63.02** | **0.979** | **82.79** | **1.03** | **72.8** | 0.678 | **72.87** | **0.893** |

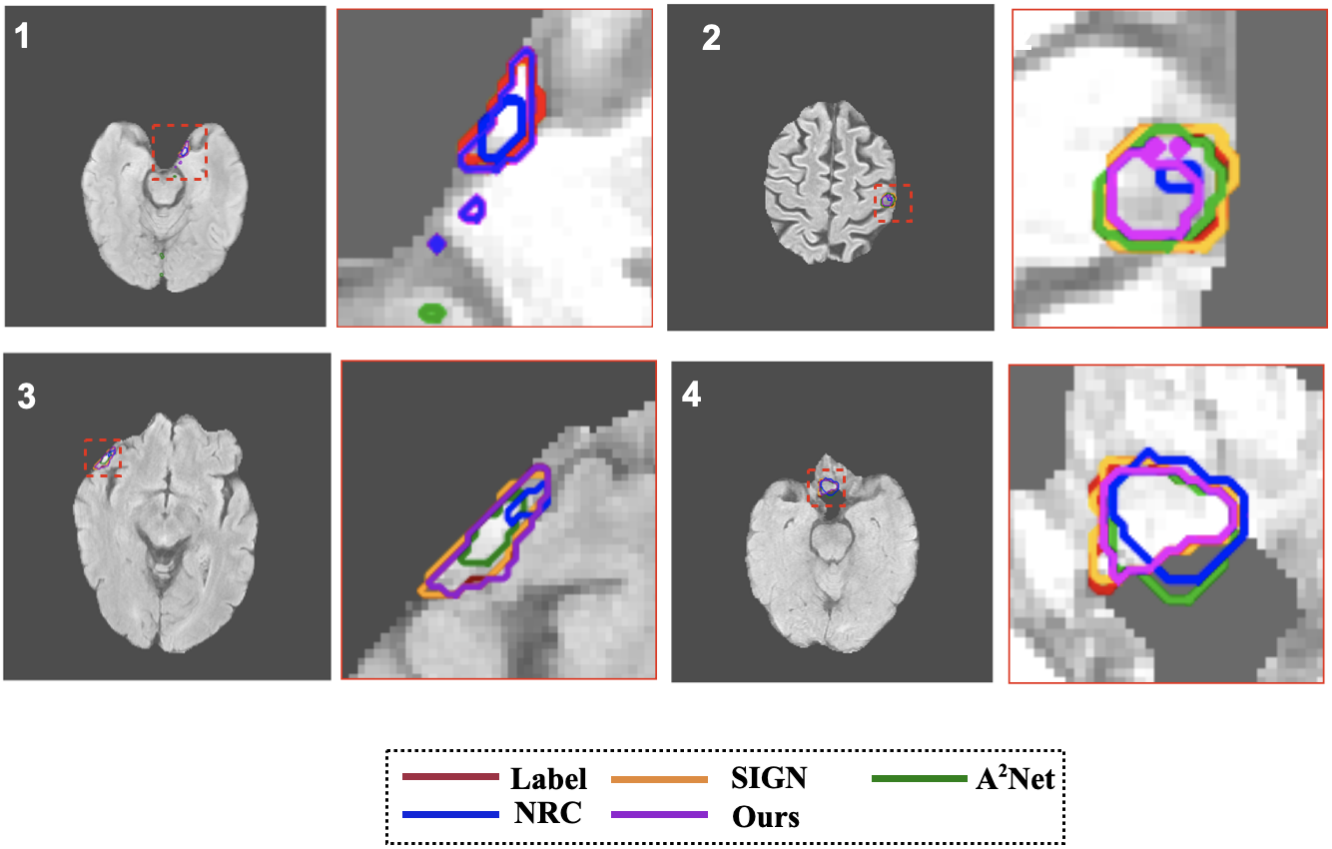



Fig. 3. Visual comparison for small brain tumors

branch for source-similar target cases and learning the unique target features in another branch for source-dissimilar target cases. We also impose additional supervision from unlabeled target data by maintaining a set of representative prototypes for pseudo label generation and self-training. We train our model by first explicitly learning the difference between source and target data, then minimizing their discrepancies by unifying the outputs from two branches.

Without any access to the source class data but only source model parameters, our method effectively ultilizes source model knowledge and few-shot target class data. Our experiments suggest the proposed method achieved better results compared to other representative source-free adaptation and few-shot learning methods, which demonstrates its effectiveness for rarer brain tumor segmentation.

## REFERENCES

[1] A P Patel, J L Fisher, E Nichols, F Abd-Allah, J Abdela, A Abdelalim, H N Abraha, D Agius, F Alahdab, T Alam, et al., "Global, regional, and national burden of brain and other cns cancer, 1990–2016: a systematic

TABLE II
ABLATION STUDY TO ASSESS THE EFFECTIVENESS OF OUR PROPOSED METHOD. THE BEST RESULTS WERE HIGHLIGHTED IN BLACK BOLD.

| CDL | PSF | Metastasis | | Pediatric Tumors | | Meningiomas | | Average | |
|---|---|---|---|---|---|---|---|---|---|
| | | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] | Dice[%] | HD[voxels] |
| ✗ | ✗ | 41.39 | 1.730 | 60.17 | 1.858 | 41.83 | 1.220 | 47.80 | 1.602 |
| ✓ | ✗ | 55.61 | 1.303 | 71.12 | 1.580 | 65.71 | 0.771 | 64.15 | 1.218 |
| ✗ | ✓ | 53.39 | 1.427 | 66.49 | 1.927 | 60.41 | 0.784 | 60.1 | 1.38 |
| ✓ | ✓ | **63.02** | **0.979** | **82.79** | **1.03** | **72.8** | **0.678** | **72.87** | **0.893** |


analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 4, pp. 376–393, 2019.

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[3] Q. Jia and H. Shu, "BiTr-Unet: A CNN-Transformer Combined Network for MRI Brain Tumor Segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12963, pp. 3–14. Springer, 2022.

[4] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, pp. 118–132. Springer, 2021.

[5] W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, and J. Li, "TransBTS: Multimodal Brain Tumor Segmentation Using Transformer," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, vol. 12901, pp. 109–119. Springer, 2021.

[6] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12962, pp. 272–284. Springer, 2022.

[7] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[8] Pedro Pinheiro and Ronan Collobert, "Recurrent convolutional neural networks for scene labeling," in *International conference on machine learning*. PMLR, 2014, pp. 82–90.

[9] Clement Farabet, Camille Couprie, Laurent Najman, and Yann LeCun, "Learning hierarchical features for scene labeling," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915–1929, 2012.

[10] Jifeng Dai, Yi Li, Kaiming He, and Jian Sun, "R-fcn: Object detection via region-based fully convolutional networks," *Advances in neural information processing systems*, vol. 29, 2016.

[11] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*. Springer, 2016, pp. 424–432.

[12] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, 2018, pp. 3–11.

[13] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[15] Jake Snell, Kevin Swersky, and Richard Zemel, "Prototypical networks for few-shot learning," *Advances in neural information processing systems*, vol. 30, 2017.

[16] Kaixin Wang, Jun Hao Liew, Yingtian Zou, Daquan Zhou, and Jiashi Feng, "Panet: Few-shot image semantic segmentation with prototype alignment," in *proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 9197–9206.

[17] Xiang Li, Tianhan Wei, Yau Pun Chen, Yu-Wing Tai, and Chi-Keung Tang, "Fss-1000: A 1000-class dataset for few-shot segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2869–2878.

[18] Shuo Lei, Xuchao Zhang, Jianfeng He, Fanglan Chen, Bowen Du, and Chang-Tien Lu, "Cross-domain few-shot semantic segmentation," in *European Conference on Computer Vision*. Springer, 2022, pp. 73–90.

[19] Xiaolin Zhang, Yunchao Wei, Yi Yang, and Thomas S Huang, "Sg-one: Similarity guidance network for one-shot semantic segmentation," *IEEE transactions on cybernetics*, vol. 50, no. 9, pp. 3855–3865, 2020.

[20] Cheng Ouyang, Carlo Biffi, Chen Chen, Turkay Kart, Huaqi Qiu, and Daniel Rueckert, "Self-supervision with superpixels: Training few-shot medical image segmentation without annotation," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIX 16*. Springer, 2020, pp. 762–780.

[21] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International conference on machine learning*. PMLR, 2017, pp. 1126–1135.

[22] Xun Huang and Serge Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1501–1510.

[23] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[24] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[25] Z Cheng, S Wang, T Xin, et al., "Few-shot medical image segmentation via generating multiple representative descriptors," *IEEE Transactions on Medical Imaging*, 2024.

[26] B. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, pp. 1993–2024, 2014.

[27] H Xia, H Zhao, and Z Ding, "Adaptive adversarial network for source-free domain adaptation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 9010–9019.

[28] S Yang, J Van de W, L Herranz, S 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.

[29] D. Zhang, R. Confidence, and U. Anazodo, "Stroke lesion segmentation from low-quality and few-shot mris via similarity-weighted self-ensembling framework," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 87–96.