Reliable Source Approximation: Source-Free Unsupervised Domain Adaptation for Vestibular Schwannoma MRI Segmentation

Hongye Zeng[1], Ke Zou[2], Zhihao Chen[3], Rui Zheng[1,4], and Huazhu Fu[5]

[1] School of Information Science and Technology, ShanghaiTech University, Shanghai, China
[2] National Key Laboratory of Fundamental Science on Synthetic Vision, College of Computer Science, Sichuan University, Chengdu, China
[3] College of Intelligence and Computing from Tianjin University, China
[4] Shanghai Engineering Research Center of Intelligent Vision and Imaging, Shanghai, China
[5] Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore 138632, Republic of Singapore
{zenghy,zhengrui}@shanghaitech.edu.cn, hzfu@ieee.org

Abstract. Source-Free Unsupervised Domain Adaptation (SFUDA) has recently become a focus in the medical image domain adaptation, as it only utilizes the source model and does not require annotated target data. However, current SFUDA approaches cannot tackle the complex segmentation task across different MRI sequences, such as the vestibular schwannoma segmentation. To address this problem, we proposed Reliable Source Approximation (RSA), which can generate source-like and structure-preserved images from the target domain for updating model parameters and adapting domain shifts. Specifically, RSA deploys a conditional diffusion model to generate multiple source-like images under the guidance of varying edges of one target image. An uncertainty estimation module is then introduced to predict and refine reliable pseudo labels of generated images, and the prediction consistency is developed to select the most reliable generations. Subsequently, all reliable generated images and their pseudo labels are utilized to update the model. Our RSA is validated on vestibular schwannoma segmentation across multi-modality MRI. The experimental results demonstrate that RSA consistently improves domain adaptation performance over other state-of-the-art SFUDA methods. Code is available at https://github.com/zenghy96/Reliable-Source-Approximation.

Keywords: Source-Free Unsupervised Domain Adaptation · Uncertainty estimation · Prediction consistency · MRI segmentation.

---

1   Introduction

Supervised learning models have recently made substantial advancements in medical image analysis and typically assume the same data distribution between train and test data [1]. This assumption challenges them in real-world practice when confronted with data collected from different imaging systems [2]. To address this domain-shift problem, many adaptation approaches have been proposed and their application scenarios are becoming increasingly practical. Unsupervised domain adaptation (UDA) [3,4] is gradually being replaced by source-free unsupervised domain adaptation (SFUDA) [5] since source domain data is often invisible due to privacy concerns.

Entropy minimization is a popular technique in existing SFUDA approaches and aims to produce more confident model predictions [6,7]. TENT [8] is the pioneer in introducing this concept, which minimizes the entropy of model predictions to reduce generalization error. TENT necessitates continuous updates to the model, which potentially incurs significant computational costs and suffers from unstable performance due to small batch size and imbalanced data. ETTA [9] needs lower computation costs by identifying reliable and non-redundant samples. SAR [10] explores obstacles that harm the model performance. Pseudo-labeling is another mainstream SFUDA technique and aims to discard or correct erroneous pseudo-labels, which inevitably appear under the influence of domain shifts. Currently, there exist three distinct solutions: enhancing the quality of pseudo-labels via denoising, screening out inaccurate pseudo-labels, and devising a robust disparity metric for pseudo-labels. Many SFUDA methods directly use the class with the highest prediction probability as the pseudo-label [11,12,13,14,15,16]. To reduce the noise pseudo-labels generated by the argmax operation, several approaches primarily focus on developing diverse filtering mechanisms to utilize dependable pseudo-labels selectively, such as maximum prediction probability [11,12,13,14], self-entropy [15], consistency score [12,16], and weight-averaged predictions [17]. However, most SFUDA methods are validated on domain-shift data caused by different centers, protocols, or devices, and are unable to effectively handle the segmentation task between different MRI sequences, which have relatively large domain shifts.

Complex medical image segmentation tasks in domain adaptation commonly use source approximation, which aims to generate source-like images and is generally implemented in UDA scenarios through CycleGAN [18]. Cai et al. [19] introduce a cross-modality MR/CT segmentation framework utilizing Cycle-GAN, which integrates a shape consistency loss to maintain anatomical coherence across source and target images. Jiang et al. [20,21] employ CycleGAN to produce MRI images from CT scans for lung tumor segmentation, facilitating precise delineation of tumors adjacent to soft tissues. For SFUDA scenarios, source approximation is implemented by Fourier transform. For example, FSM [22] acquires a preliminary source image by immobilizing the source model and refining a trainable image, which is enhanced through mutual Fourier Transform. The resultant refined source-like image depicts the source data distribution, thereby aiding in domain alignment throughout the adaptation procedure.
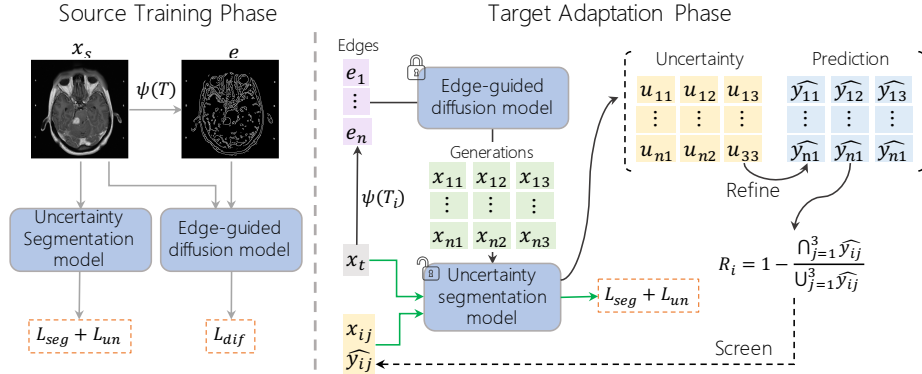
Fig. 1. Overview of the proposed SFUDA method. Two models are pre-trained using source data, and the target adaptation phase consists of Reliable Source Approximation (**RSA**, black arrows) and model fine-tuning (green arrows).

first attains a coarse source image by freezing the source model and training a learnable image, then refines the image via mutual Fourier Transform. The refined source-like image provides a representation of the source data distribution and facilitates domain alignment during the adaptation process.

In this paper, we proposed the Reliable Source Approximation (**RSA**) for vestibular schwannoma segmentation across different MRI sequences. In contrast to previous SFUDA source approximations, our method innovatively uses edge-guided image translation to generate source-like and structure-preserved images from the target domain. To obtain the most reliable generated images and their pseudo-labels, an uncertainty segmentation model is introduced to produce reliable pseudo-labels through uncertainty refinement, and the prediction consistency is proposed to identify the optimal approximation result. Our contributions can be summarized as follows: (1) We propose a novel source approximation method through edge-guided image translation. (2) We introduce the uncertainty refinement and prediction consistency to obtain the most reliable approximation result. (3) Our proposed RSA outperforms other SFUDA methods on adaptive vestibular schwannoma segmentation.

2   Methodology

Figure 1 illustrates our SFUDA framework via reliable source approximation. Given the source data with annotation $D_s = \{x_s^i, y_s^i\}_{i=1}^N$, we first train the edge-guided diffusion model g_θ and the uncertainty segmentation model f_θ. Then, the source-like and structure-preserved images are generated under spatial guidance from edges. The prediction consistency and uncertainty are used to find the optimal generated image x_{ij} and its reliable pseudo label $\hat{y}_{ij}$. Finally, all x_t, y_{ij}, and $\hat{y}_{ij}$ are used to finetune the uncertainty segmentation model for adaptation.

2.1   Edge-guided Image Translation

The edge-guided image translation aims to translate image style and simultaneously preserve the original spatial structure. To achieve this, we first train a basic denoising diffusion probabilistic model (DDPM) [23] for source image synthesis. It can generate source-like images by progressive denoising an input Gaussian noise, but generated images are completely random without any spatial guidance. Here, we consider the edge e as spatial guidance, which is derived from Canny algorithm ψ:

$$e = \psi(x_s, T) \tag{1}$$

where T is the Canny threshold and produced based on annotation. Following the ControlNet [24] strategy, we connect the locked DDPM and its trainable copy by zero convolution layers, and decode the edge into the model to learn spatial guidance. The learning objective of the entire edge-guided diffusion model f_θ is defined as:

$$L_{dif} = \mathbb{E}_{x_s, t, e, \epsilon \sim \mathcal{N}(0,1)} \left[\left\| \epsilon - \epsilon_f \left(x_s^t, t, e \right) \right\| \right] \tag{2}$$

where ϵ is real noise and ϵ_f is estimated noise. x_s^t is the noisy image, obtained by progressively adding the random noise t times to a source image x_s. The optimization of the edge-guided diffusion model involves minimizing the discrepancy between real and estimated noises.

2.2   Uncertainty segmentation model

The uncertainty segmentation model is employed to produce segmentation masks, with the uncertainty estimation being utilized for pseudo-label refinement. We use the basic U-Net [25] for the segmentation task and the learning objective consist of the commonly used cross-entropy loss and Dice loss:

$$L_{seg} = \sum_{n=1}^{C} -y_s^n \log(\hat{y}) + 1 - \frac{2 * y_s * \hat{y} + 1}{y_s + \hat{y} + 1} \tag{3}$$

where $\hat{y} = f_\theta(x_s)$ is the predicted mask.

In our work, we expanded the U-Net architecture by introducing an additional branch aimed at estimating prediction uncertainty. Importantly, this extension seamlessly integrates into the existing U-Net structure without necessitating any fundamental alterations. This new branch is dedicated to directly assessing uncertainty by predicting the parameters of the Normal-Inverse-Gamma (NIG) distribution [26]. The NIG distribution serves as a conjugate prior distribution for the source domain distribution. We consider the relationship as follows: y_s follows a normal distribution $\mathcal{N}(\mu, \sigma^2)$, where μ follows $\mathcal{N}(\gamma, \sigma^2 \omega^{-1})$ and σ^2 follows $\Gamma^{-1}(\alpha, \beta)$, where Γ^{-1} represents the gamma function. This formulation characterizes the distribution of y_s as an NIG distribution with parameters γ, ω, α, and β.

$$p(\mu, \sigma^2 | \gamma, \omega, \alpha, \beta) = \frac{\beta^\alpha \sqrt{\omega}}{\Gamma(\alpha)\sqrt{2\pi\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left\{-\frac{2\beta + \omega(\gamma - \mu)^2}{2\sigma^2}\right\} \quad (4)$$

where $\gamma \in \mathbb{R}$, $\omega > 1$ and $\beta > 0$. In the training phase, we apply the following negative log-likelihood loss to account for the NIG distribution

$$L_{NIG} = \frac{1}{2}\log\left(\frac{\pi}{\omega}\right) - \alpha\log(\Omega) + (\alpha + \frac{1}{2})\log((y - \gamma)^2\omega + \Omega) + \log(\Theta) \quad (5)$$

where $\Omega = 2\beta(1 + \omega)$ and $\Theta = \left(\frac{\Gamma(\alpha)}{\Gamma(\alpha + \frac{1}{2})}\right)$. Furthermore, we introduce a regularization term $L_R = |y_i - \gamma| \cdot (2\omega + \alpha)$ into the total loss to penalize incorrect evidence

$$L_{un} = L_{NIG} + \lambda L_R \quad (6)$$

We introduce a coefficient λ to balance the contributions of the two loss terms. Ultimately, we replace the deterministic output of the model with a NIG distribution $f_\theta(x_s) = NIG(\gamma, \omega, \alpha, \beta)$. The prediction $\hat{y}$ and the model uncertainty map u are defined as:

$$\hat{y} = \mathbb{E}(u) = \gamma \quad (7)$$

$$u = \frac{\beta}{\omega(\alpha - 1)} \quad (8)$$

2.3 Reliable Source Approximation

Due to domain shift, it is impossible to directly find an edge of the target image and obtain a high-quality source approximation through the edge-guided diffusion model learned from the source domain. Therefore, we generate multiple edges with n thresholds $\{e_i | e_i = r(x_t, T_i), i = 1, 2, \ldots, n\}$ for a target image x_t, and $n \times 3$ images are generated under the guidance of edges. The key challenge is to find the optimal source approximation: the source-like and structure-preserved generated image with its reliable pseudo-label.

Refine pseudo-label using uncertainty Domain shift can compromise the accuracy of pseudo-labels, potentially hindering overall model performance if the model is directly fine-tuned using them. To mitigate this issue, we employ uncertainty estimation to discern pixels with dependable pseudo-labels, subsequently utilizing only refined pseudo-labels for model supervision. Since the uncertainty segmentation model always exhibits high uncertainty at the edges of tumors when predicting generated images, we include the non-tumor pixels with high uncertainty at the edges in our analysis:

$$\hat{y_{ij}} = \mathbb{1}\{u_{ij} > T_{un}\} \oplus \hat{y_{ij}} \cup \hat{y_{ij}} \quad (9)$$

where T_{un} is the threshold of uncertainty, and $\mathbb{1}\{u_{ij} > T_{un}\}$ find the high uncertainty prediction. The operation $\oplus$ finds non-tumor pixels with high uncertainty, and $\cup$ incorporates them into the predicted mask.

Prediction consistency Under the guidance of same edge e_i, the spatial structures of all generated images $\{x_{i1}, x_{i2}, x_{i3}\}$ should be consistent if e_i provides a precise spatial context. Further, their predictions $\{\hat{y_{i1}}, \hat{y_{i2}}, \hat{y_{i3}}\}$ should be similar and the prediction consistency R_i is defined as:

$$R_i = 1 - \frac{\bigcap_{j=1}^{3} \hat{y_{ij}}}{\bigcup_{j=3}^{3} \hat{y_{ij}}} \tag{10}$$

the edges with prediction consistency above T_R are filtered out and best edge index is found by $argmin\{R[i] \mid R[i] \leq T_R, \ i \in [0, n]\}$. The final approximation $\{x_{ij}, \hat{y_{ij}}\}$ is found based the smallest variance prediction in $\{\hat{y_{i1}}, \hat{y_{i2}}, \hat{y_{i3}}\}$.

Batch-based and centralized fine-tuning We offer two fine-tuning methods here. Batch-based fine-tuning solely utilizes the current batch to train the model and directly predicts the results for the current batch. Batch-based fine-tuning only trains one epoch, and all batches are processed once. Centralized fine-tuning requires aggregating all batches and conducts multiple epochs of training.
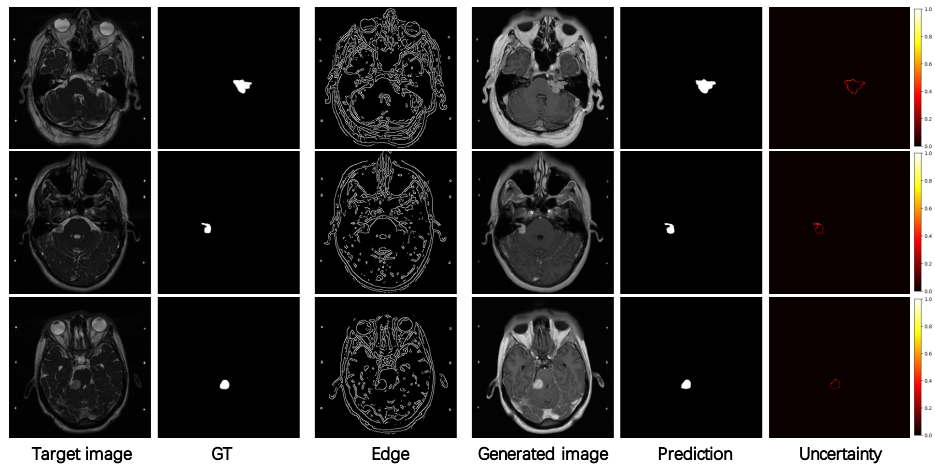
3   Experiment

3.1   Experimental Setup

Dataset and evaluation metrics We validate our SFUDA approach on the public vestibular schwannoma dataset including ceT1 and hrT2 MRI sequences from 242 patients [27]. We consider ceT1 MRI as the source domain and hrT2 as the target domain. Following the data split in [27] and slicing the MRI volume, we finally obtain 1599 source training images, 1493 target training images, and 423 target testing images. In the evaluation process, the Dice similarity coefficient (Dice) and the average symmetric surface distance (ASSD) are employed to quantitatively assess the segmentation outcomes. [27].

Implementation details The experiments are conducted on the PyTorch platform with an NVIDIA A100 GPU. We first train DDPM for 400 epochs using source images, and then update the entire edge-guided diffusion model for 100 epochs simultaneously using source images and their edges. The diffusion model training uses the AdamW optimizer with a learning rate of 1e-4 and the batch size is 16. The sampler is Denoising Diffusion Implicit Models (DDIM) [28] with 50 steps. We train the uncertainty segmentation model for 100 epochs and use the Adam optimizer with a learning rate of 1e-4 and a batch size of 32. Both the generated images and inputs of the uncertainty segmentation model are 320×320 resolutions. In the adaptation phase, we equally divide the range from 30 to 80 to obtain n Canny thresholds. The hyper-parameters are set as $n = 2$, $T_{un} = 0.2$, $T_R = 0.3$ through ablation study.

Table 1. Quantitative results of comparison with different methods. The best scores of batch-based and centralized fine-tuning are highlighted.

Strategy	Method	Fine-tuning	Dice (%) ↑	ASSD (mm)↓
No adaptation	Source-only	-	6.63 ± 1.85	2.10 ± 0.29
Supervised	Target-only	-	84.49 ± 0.96	0.87 ± 0.06
Entropy minimization	TENT[8]	Batch-based	26.53 ± 4.73	9.56 ± 0.70
		Centralized	26.55 ± 5.69	9.54 ± 0.80
	EATA[9]	Batch-based	26.53 ± 4.73	9.55 ± 0.71
		Centralized	26.54 ± 5.70	9.53 ± 0.78
	SAR[10]	Batch-based	26.31 ± 5.82	9.52 ± 0.80
		Centralized	26.30 ± 5.80	9.51 ± 0.81
Pseudo -labeling	COTTA[17]	Batch-based	26.52 ± 4.83	9.50 ± 0.60
		Centralized	26.54 ± 5.80	9.42 ± 0.77
Source approximation	FSM[22]	Centralized	56.12 ± 1.32	4.56 ± 0.54
	Ours	Batch-based	**45.55 ± 3.01**	**8.23 ± 1.11**
		Centralized	**77.83 ± 0.98**	**1.24 ± 0.11**

Target image   GT   Edge   Generated image   Prediction   Uncertainty

Fig. 2. Visulazition of reliable source approximation and segmentation results.

3.2 Experimental Results

Quantitative comparison In our experiments, the "source-only" method refers to training a model solely on the source domain and evaluating it directly on the target domain without any adaptation. Conversely, the "target-only" method involves training and testing exclusively on the target domain. We conducted comparisons between our approaches and recent SFUDA methods employing various strategies, with many of these methods employing dual fine-tuning techniques.

Table 1 presents quantitative comparison results. There is a huge performance gap between source-only and target-only methods due to the domain shift between ceT1 and hrT2 MRI. The Dice score and ASSD of our method significantly outperform the entropy minimization and pseudo-labeling methods in two fine-tuning methods (Dice improvements $> 45\%$ and $> 15\%$, ASSD improvements $> 8mm$ and $> 1mm$). These improvements can be attributed to the source approximation ability in tackling large domain shifts. Our method still outperforms the other source approximation FSM by a margin of 21.71% in Dice and 3.32 mm in ASSD. All entropy minimization methods achieve similar results, which means that the latest improvements over TENT have not been very effective for VS segmentation tasks. Moreover, our method demonstrates greater stability compared to other approaches, since it achieves the least performance variance. Figure 2 shows the best source approximations and their segmentation results, the edge-guided approximations are source-like and structure-preserved.

Table 2. Ablation study for hyperparameter T_R.

T_R	Approximation		Performance	
	Quality (%)	Quantity (n)	Dice (%)	ASSD (mm)
0.1	**86.86**	278	77.47	1.49
0.3	78.58	706	**77.63**	1.53
0.5	73.85	910	76.55	**1.46**
0.7	69.88	1064	73.11	1.62
0.9	65.33	**1226**	72.06	1.61

Table 3. Ablation study for hyperparameter T_{un}.

T_{un}	Approximation		Performance	
	Quality (%)	Quantity (n)	Dice (%)	ASSD (mm)
0.002	72.66	910	74.95	1.47
0.02	73.85	910	75.59	**1.33**
0.2	**73.92**	910	**76.32**	1.41

Hyper-parameter Sensitivity Analysis We further investigate the impact of different hyper-parameters through source approximation and model performance. The quality of source approximation could be measured by the Dice score between pseudo-labels and ground truth.

Table 2 presents the results with different prediction consistency thresholds T_R. It presents that both too-loose and too-tight thresholds would cause a drop in performance: a loose threshold of 0.9 generates a large amount of low-quality

Table 4. Ablation study for hyperparameter n.

n	Approximation		Performance	
	Quality (%)	Quantity (n)	Dice (%)	ASSD (mm)
2	**76.77**	786	**78.53**	**1.17**
4	73.85	910	75.59	1.33
8	68.72	**1107**	73.21	1.35

data, which is detrimental to model fine-tuning; a tight threshold of 0.1 yields better-quality data but in smaller quantities, potentially preventing the model from acquiring new knowledge. Table 3 presents the results with different uncertainty threshold T_{un}. We find that a more relaxed T_{un} is more beneficial to results, as too stringent T_{un} may lead to an excessive number of pixels being added to the predicted mask and damages the pseudo-label quality. We also find that T_{un} only affects the approximation quality but does not impact the quantity. Table 4 presents the results with different numbers of edges n. Excessive edges could lead to a surge in low-quality approximations and noisy pseudo-labels, resulting in decreased performance.

4 Conclusion

This study introduces a novel SFUDA method tailored for the vestibular schwannoma MRI segmentation. We implement a novel source-free source approximation method via edge-guided image translation. The uncertainty segmentation model and prediction consistency are introduced to obtain Reliable Source Approximation results. Experimental results on cross-sequence MRI image segmentation demonstrate that our method outperforms state-of-the-art SFUDA approaches.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher, "Deep learning-enabled medical computer vision," *npj Digit. Med.*, vol. 4, no. 1, pp. 1–9, Jan. 2021.
2. I. Redko, E. Morvant, A. Habrard, M. Sebban, and Y. Bennani, *Advances in Domain Adaptation Theory.* Elsevier, Aug. 2019.
3. Y. Ganin *et al.*, "Domain-adversarial training of neural networks," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, 2016.

4. M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional Adversarial Domain Adaptation," Dec. 2018.
5. J. Liang, D. Hu, and J. Feng, "Do We Really Need to Access the Source Data? Source Hypothesis Transfer for Unsupervised Domain Adaptation," Jun. 2021.
6. M. Bateson, H. Kervadec, J. Dolz, H. Lombaert, and I. Ben Ayed, "Source-relaxed domain adaptation for image segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*.   Cham: Springer International Publishing, 2020, pp. 490–499.
7. M. Bateson, H. Kervadec, J. Dolz, H. Lombaert, and B. AyedIsmail, "Source-free domain adaptation for image segmentation," *Medical Image Analysis*, vol. 82, p. 102617, Nov. 2022.
8. D. Wang, E. Shelhamer, S. Liu, B. Olshausen, and T. Darrell, "Tent: Fully Test-time Adaptation by Entropy Minimization," Mar. 2021.
9. S. Niu, J. Wu, Y. Zhang, Y. Chen, S. Zheng, P. Zhao, and M. Tan, "Efficient Test-Time Model Adaptation without Forgetting," in *Proceedings of the 39th International Conference on Machine Learning*.   PMLR, Jun. 2022, pp. 16 888–16 905.
10. S. Niu, J. Wu, Y. Zhang, Z. Wen, Y. Chen, P. Zhao, and M. Tan, "Towards Stable Test-Time Adaptation in Dynamic Wild World," Feb. 2023.
11. L. Zhou, M. Ye, and S. Xiao, "Domain adaptation based on source category prototypes," *Neural Computing and Applications*, vol. 34, no. 23, pp. 21 191–21 203, 2022.
12. V. Prabhu, S. Khare, D. Kartik, and J. Hoffman, "Augco: Augmentation consistency-guided self-training for sourcefree domain adaptive semantic segmentation," *arXiv preprint arXiv:2107.10140*, 2022.
13. D. Kothandaraman, R. Chandra, and D. Manocha, "Ss-sfda: Selfsupervised source-free domain adaptation for road segmentation in hazardous environments," in *Proc. ICCV Workshops*, 2021.
14. A. Dasgupta, C. Jawahar, and K. Alahari, "Overcoming label noise for source-free unsupervised video domain adaptation," in *Indian Conference on Computer Vision, Graphics, and Image Processing (ICVGIP)*.   ACM, 2022.
15. L. Tian, L. Zhou, H. Zhang, Z. Wang, and M. Ye, "Robust selfsupervised learning for source-free domain adaptation," *Signal, Image and Video Processing*, pp. 1–9, 2023.
16. Y. Hou and L. Zheng, "Visualizing adapted knowledge in domain transfer," in *Proc. CVPR*, 2021, pp. 13 824–13 833.
17. Q. Wang, O. Fink, L. Van Gool, and D. Dai, "Continual Test-Time Domain Adaptation," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.   New Orleans, LA, USA: IEEE, Jun. 2022, pp. 7191–7201.
18. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," Aug. 2020.
19. J. Cai *et al.*, "Towards cross-modal organ translation and segmentation: A cycle- and shape-consistent generative adversarial network," *Med. Image Anal.*, vol. 52, pp. 174–184, 2019.
20. J. Jiang, J. Hu, N. Tyagi, A. Rimner, S. L. Berry, J. O. Deasy, and H. Veeraraghavan, "Integrating cross-modality hallucinated mri with ct to aid mediastinal lung tumor segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Intervention*, 2019, pp. 221–229.
21. J. Jiang *et al.*, "Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Intervention*, 2018, pp. 777–785.

22. C. Yang, X. Guo, Z. Chen, and Y. Yuan, "Source free domain adaptation for medical image segmentation with fourier style mining," *Medical Image Analysis*, vol. 79, p. 102457, Jul. 2022.

23. J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," Dec. 2020.

24. L. Zhang and M. Agrawala, "Adding Conditional Control to Text-to-Image Diffusion Models," Feb. 2023.

25. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds.   Cham: Springer International Publishing, 2015, vol. 9351, pp. 234–241.

26. A. Amini, W. Schwarting, A. Soleimany, and D. Rus, "Deep Evidential Regression," in *Advances in Neural Information Processing Systems*, vol. 33. Curran Associates, Inc., 2020, pp. 14 927–14 937.

27. J. Shapey, G. Wang, R. Dorent, A. Dimitriadis, W. Li, I. Paddick, N. Kitchen, S. Bisdas, S. R. Saeed, S. Ourselin, R. Bradford, and T. Vercauteren, "An artificial intelligence framework for automatic segmentation and volumetry of vestibular schwannomas from contrast-enhanced T1-weighted and high-resolution T2-weighted MRI," *Journal of Neurosurgery*, vol. 134, no. 1, pp. 171–179, Dec. 2019.

28. J. Song, C. Meng, and S. Ermon, "Denoising Diffusion Implicit Models," Oct. 2022.