Exploiting Latent Classes for Medical Image Segmentation from Partially Labeled Datasets

Xiangyu Zhao[1,2⋆], Xi Ouyang[1⋆], Lichi Zhang[2,3], Zhong Xue[1], and Dinggang Shen[1,4,5(✉)]

[1] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China
[2] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[3] National Engineering Research Center of Advanced Magnetic Resonance Technologies for Diagnosis and Therapy (NERC-AMRT), Shanghai Jiao Tong University, Shanghai, China
[4] School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
[5] Shanghai Clinical Research and Trial Center, Shanghai, China
dgshen@shanghaitech.edu.cn

Abstract. Notable progress has been made in medical image segmentation models due to the availability of massive training data. Nevertheless, a majority of open-source datasets are only partially labeled, and not all expected organs or tumors are annotated in these images. While previous attempts have been made to only learn segmentation from labeled regions of interest (ROIs), they do not consider the *latent classes*, *i.e.*, existing but unlabeled ROIs, in the images during the training stage. Moreover, since these methods rely exclusively on labeled ROIs and those unlabeled regions are viewed as background, they need large-scale and diverse datasets to achieve a variety of ROI segmentation. In this paper, we propose a framework that utilizes latent classes for segmentation from partially labeled datasets, aiming to improve segmentation performance, especially for ROIs with only a small number of annotations. Specifically, we first introduce an ROI-aware network to detect the presence of unlabeled ROIs in images and form the latent classes, which are utilized to guide the segmentation learning. Additionally, ROIs with ambiguous existence are constrained by the consistency loss between the predictions of the student and the teacher networks. By regularizing ROIs with different certainty levels under different scenarios, our method can significantly improve the robustness and reliance of segmentation on large-scale datasets. Experimental results on a public benchmark for partially labeled segmentation demonstrate that our proposed method surpasses previous attempts and has great potential to form a large-scale foundation segmentation model.

Keywords: Medical image segmentation · partially labeled data · latent classes · awareness guidance

⋆ X. Zhao and X. Ouyang contribute equally to this work.

1 Introduction

Accurate segmentation of regions of interest (ROIs) in medical images is an important task for computer-aided diagnosis, which could accelerate clinical workflow and reduce workload for image reading [7]. The recent development of foundation models [2, 12] has witnessed a giant leap in the performance of medical image segmentation on variant organs and lesions attributed to large-scale training data. For example, a number of segmentation models [3, 9, 11, 13] have utilized multiple open-source datasets to form a training set that is large enough to guarantee segmentation performance. However, most of the datasets are partially labeled, *i.e.* only a few ROIs in the images are labeled while the rest organs or lesions are not. Such a partial labeling issue would result in unlabeled organs being treated as the background, and segmentation training could require multiple datasets with a variety of ROIs to overcome such shortcomings. Therefore, partially labeled data pose a challenge in the development of a universal segmentation model.

Previous methods typically utilize multi-head design [3,13] or dynamic paths [9, 21] to surpass the partial labeling issue in the open-source medical image segmentation datasets. Under such circumstances, only the labeled data contributes to the back-propagation, and the unlabeled ROIs in the images are simply ignored. Although these methods could handle the partial labeling issue, they require sufficient data for each ROI to boost performance, as they overlook the unlabeled ROIs that may contribute to the training process. Another commonly applied solution is the combination of dynamic paths and semi-supervised learning [20], which leverages the unlabeled ROIs in the images. However, semi-supervised learning approaches are unaware of the existence of the ROIs in the image, which could lead to less optimal segmentation performance.

In this paper, we are aware of the latent classes in partially labeled segmentation and develop a novel approach that leverages the latent classes in the medical images to improve the segmentation performance and reduce the demand for large-scale training data. Specifically, we design a teacher-student framework with an extra ROI-aware classifier that identifies the existence of unlabeled ROIs in the image, while the supervision of the student network is determined by the prediction certainty from the classifier. Given an unlabeled ROI, it is regarded as a *latent class* when the certainty level of its existence is high enough, otherwise it is regarded as an *ambiguous class*. During training, the student network is supervised by the segmentation losses from both labeled ROIs and *latent classes*. Additionally, the student network is further regularized by the consistency loss with the teacher network for *ambiguous classes*. In this way, the segmentation fully benefits from all of the available training data regardless of the partial labeling issue. This is more advantageous for categories with only a small number of annotations to benefit from potential unlabeled information. The proposed method has been extensively evaluated on a partially labeled multi-organ and tumor segmentation dataset. Experimental results demonstrate the effectiveness of the proposed method by mining latent classes in the images, which outperforms previous partially labeled segmentation methods.

2 Method

2.1 Preliminaries: Partially Labeled Segmentation

Considering a universal segmentation of an ROI set $\mathbf{C}$ with K ROIs (including organs and tumors), the ideal implementation is a fully-supervised multi-class segmentation on all the ROIs. However, it is not practical in the real scenario to collect large-scale datasets with fully-labeled ROIs. Therefore, it is a common practice to assemble m partially labeled datasets, where each dataset contains annotations for K_m ROIs out of total K ROIs ($K_m < K$). Actually, this is consistent with the annotation strategy of most open-source datasets which only annotate the targets of interest.

Dynamic heads [21] have proved to be successful in dealing with partially labeled segmentation. Inspired by [21] and [9], we build a segmentation network $\mathcal{F}$ with a dynamic segmentation head parameterized by the ROI text embeddings generated from BiomedCLIP [22]. Compared with the one-hot task encoding proposed in [21], CLIP text embeddings have more flexibility and could handle the cases where different ROIs are partly related [9]. During training, only the labeled ROIs are considered for back-propagation, while the ROIs that are not labeled do not contribute to the loss calculation.

Although the network with dynamic heads can learn segmentation from multiple partially labeled datasets, the current learning objective [9,21] does not take into consideration the latent classes that are not labeled in the images. Thus, the training data is not fully investigated during the training process, leading to less optimal segmentation performance especially when the available training data are not large enough for some target ROI. However, many images contain these ROI regions but without gold standard.

2.2 Latent Class Identification

The ROIs that exist in the image but remain unlabeled, which are termed *latent classes*, could be leveraged to improve the segmentation performance, especially for the ROIs with limited annotations. In order to utilize the latent classes, it is necessary to first identify the ROIs existing in the image. Liu *et al.* [9] claimed that CLIP embeddings as task encoding provide information on the relationships of different ROIs. However, the related ROIs hinted by CLIP embeddings do not imply spatial location adjacency in a medical scan. For instance, "lower limbs artery" and "carotid artery" may have a very high similarity in CLIP embeddings as they are both vital arteries in the human body. However, they are located in completely different parts of the human body, which means it is almost impossible for them to appear in the same patch of CT or MRI images.

Here we adopt a simple yet effective approach to surpass this problem. Specifically, we propose to train a classification network that identifies whether a certain ROI is present in the input image (or image patch), called ROI-aware network. We adopt a 3D ResNet-18 [5] for the identification of the latent ROIs when given an input image, which is a lightweight and effective classifier on medical image

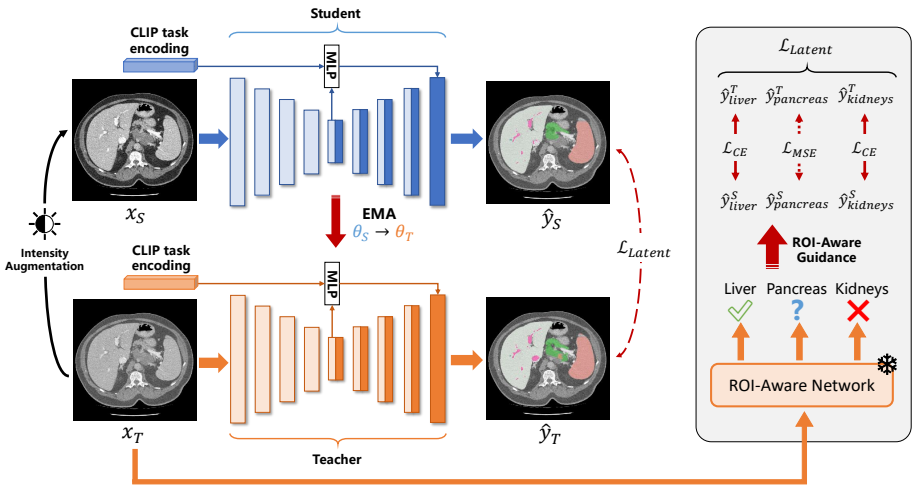

Fig. 1. The overall framework for partially labeled segmentation that leverages latent classes, where the unlabeled ROIs are identified by an ROI-aware network. Existing and non-existing ROIs are classified as *latent classes* and regularized with CE loss, while ambiguous classes are regularized with MSE loss.

classification without adding noticeable extra computational cost. To enhance the network's ability to recognize ROI, we use the TotalSegmentator dataset [17] to train the classifier, which contains 104 anatomical structures (27 organs, 59 bones, 10 muscles, 8 vessels). The learning objective for the ROI-aware network is to minimize the multi-class binary classification loss given the existence label of ROIs from TotalSegmentator dataset. Based on that, our ROI-aware network can identify almost all the organs of the human body. It should be noted that in this dataset, tumors are not presented in these images. Thus, when the ROI-aware network is identifying whether a certain kind of tumor is a latent class, the decision is based on the existence of the related organ. For example, when "liver" is identified as a latent class, "liver tumor" is viewed as a latent class as well, even if we are not aware of the existence of liver tumors.

2.3 Latent Class Mining for Segmentation

In order to leverage the latent classes during segmentation learning, we propose to allow the network to learn the segmentation of latent classes from the pseudo labels. Thus, we adopt a teacher-student framework to generate the psuedo labels for unlabeled ROIs. The main difference between the proposed latent class mining framework and vanilla mean-teacher framework [15] is that our method leverages the ROI-awareness information, *i.e.*, only the identified latent classes contribute to the segmentation learning. Specifically, the proposed framework consists of a student network $\mathcal{F}_S$ and a teacher network $\mathcal{F}_T$. During training, the parameters θ_T of teacher network $\mathcal{F}_T$ is updated using the expo-

nential moving average (EMA) of the parameters θ_S of the student model $\mathcal{F}_S$: $\tilde{\theta}_T = \alpha\theta_T + (1-\alpha)\theta_S$, where $\alpha = 0.999$ is an empirical decay rate for EMA update.

Considering the segmentation of the ROI set $\mathbf{C}$ with K ROIs, let $x \in \mathbb{R}^{H\times W\times D}$ denote a partially labeled input image, where the labeled ROI set $\mathbf{C}_l$ contains K_m ROIs ($K_m < K$). The unlabeled ROIs are denoted as $\mathbf{C}_u$ with $K - K_m$ kinds of ROIs, where $\mathbf{C}_l \cup \mathbf{C}_u = \mathbf{C}$. During training, the input image x is augmented with spatial and intensity transformations to reduce overfitting. The transformed version of x is denoted as x_T, which is then fed to the teacher network $\mathcal{F}(\cdot;\theta_T)$ for generating the corresponding prediction $\hat{y}_T = \mathcal{F}(x;\theta_T) \in \mathbb{R}^{K\times H\times W\times D}$. Also, x_T is further augmented by intensity transformations to generate x_S, which is then fed to the student network $\mathcal{F}(\cdot;\theta_S)$ for generating the corresponding prediction $\hat{y}_S = \mathcal{F}(x;\theta_S) \in \mathbb{R}^{K\times H\times W\times D}$. During the segmentation training, the student network is optimized by the supervised segmentation loss of the labeled ROIs $\mathbf{C}_l$:

$$\min \sum_{c_i\in\mathbf{C}_l} \mathcal{L}_{seg}(\hat{y}_S^{c_i}, y^{c_i}), \text{ where } \hat{y}_S = \mathcal{F}(x_S;\theta_S), \tag{1}$$

where c_i denotes a specific category out of labeled ROI, and y^{c_i} denotes the corresponding segmentation label.

In order to identify the latent classes in input image x, the input image patch x_T is fed to the ROI-aware network for outputting the scores p_{c_i} of each ROI c_i in $\mathbf{C}$. The latent classes are classified by a set of thresholds $\{\phi_l, \phi_h\}$. When the predicted score p_{c_i} of ROI c_i is greater than the upper bound ϕ_h, c_i is regarded as a positive latent class in the image. It is worth noticing that when the predicted score p_{c_i} of ROI c_i is smaller than the lower bound ϕ_l, c_i is viewed as a negative latent class, since the ROI-aware network is very confident about its existence in the image patch. In this case, the student network is optimized by the cross-entropy loss between student prediction and the pseudo labels $\bar{y}_{c_i}$, where the pseudo labels are generated from teacher's output and "sharpened" [18] to reduce the entropy in student's prediction.

$$\min \sum_{c_i\in\mathbf{C_{latent}}} \mathcal{L}_{CE}(\hat{y}_S^{c_i}, \bar{y}^{c_i}), \text{ where } \bar{y}^{c_i} = \frac{(\hat{y}_T^{c_i})^{1/\tau}}{(\hat{y}_T^{c_i})^{1/\tau} + (1-\hat{y}_T^{c_i})^{1/\tau}}, \tag{2}$$

where $\mathbf{C_{latent}}$ denotes the set of latent classes, and $\tau = 0.1$ is the temperature parameter for "sharpening" operation. Note that the labeled ROIs are excluded from $\mathbf{C_{latent}}$ and do not involve the teacher-student loss calculation.

However, the ROI-aware network will not necessarily output a score that meets the requirement of thresholds $\{\phi_l, \phi_h\}$. When the output score for ROI c_j is greater than ϕ_l but smaller than ϕ_h, the ROI c_j is regarded as an ambiguous class. This phenomenon could occur when the ROI size is too small to be identified by the ROI-aware network. In such cases, we do not use the teacher's output as pseudo labels as the teacher's predictions could be inaccurate. Inspired by [19], we require the predictions from both the student network and teacher

network to be consistent on the ambiguous class c_j:

$$\min \sum_{c_j \in \mathbf{C_{ambiguity}}} \mathcal{L}_{MSE}(\hat{y}_S^{c_j}, \hat{y}_T^{c_j}), \tag{3}$$

where $\mathbf{C_{ambiguity}}$ denotes the set of ambiguous classes. It should be noted that $\mathbf{C_{ambiguity}} \cup \mathbf{C_{latent}} = \mathbf{C}_u$.

Thus, both latent classes and ambiguous classes are considered during segmentation learning, and different strategies are employed for best segmentation learning. Overall, the segmentation loss for latent classes $\mathcal{L}_{Latent}$ could be written as follows:

$$\mathcal{L}_{Latent} = w_{CE}\mathcal{L}_{CE} + w_{MSE}\mathcal{L}_{MSE}. \tag{4}$$

where w_{CE} and w_{MSE} are weighting factors of the losses.

3 Experiments

3.1 Data

Following [21], we build the MOTS benchmark from the LiTS [1], KiTS [6], and MSD [14] challenges. The MOTS benchmark consists of 920 training images and 235 testing images for the segmentation of the following 11 organs and tumors: liver, liver tumor, kidneys, kidney tumor, hepatic vessels, hepatic vessel tumor, pancreas, pancreatic tumor, colon cancer, lung tumor, and spleen. We follow the same data split setting described in [21] for fair comparison. In addition, we utilize TotalSegmentator [17] dataset for the training of ROI-aware network, as all the above organs are labeled in this fully-labeled dataset. We use 90% of the data in TotalSegmentator to train the ROI-aware network, while the remaining 10% are used for the performance evaluation. It should be noted that only the ROI-classifier is trained on the TotaoSegmentator dataset, and the segmentation network is not trained on the TotalSegmentator to maintain a fair comparison.

3.2 Implementation Details

The framework is implemented with PyTorch 1.12.1 on a Debian Linux server with 4 NVIDIA A40 40G GPUs. The segmentation network is built upon SwinUNETR [4], where the segmentation head is replaced by a dynamic head proposed in [9, 21]. During training, the input images are cropped into patches with a fixed spatial size of $160 \times 160 \times 96$ and fed to the segmentation network as well as the ROI-aware network. The batch size is set to 12 and the model is trained for 2000 epochs. The segmentation network is optimized by an AdamW [10] optimizer with an initial learning rate 1×10^{-4} decayed by cosine annealing and a weight decay of 1×10^{-5}. For hyperparameter choosing, we set $\phi_h = 0.8$, $\phi_l = 0.1$, $w_{CE} = 1.0$, and $w_{MSE} = 1.0$ in the experiments. Due to the GPU memory restrictions, the ROI-aware network is trained in advance and frozen during segmentation training. We optimize the ROI-aware network with the same configurations and train for 2000 epochs.

Table 1. Performance (Dice/%) of different segmentation methods on MOTS benchmark. The best two results are highlighted in bold. We also show the number of training and testing images ("train"/"test") in the bracket after each task name.

Method	Liver (104/27)		Kidneys (168/42)		Hepatic Vessels (242/61)	
	Organ	Tumor	Organ	Tumor	Organ	Tumor
1) Baseline	93.26	60.04	90.32	69.16	58.79	72.63
2) MT	94.81	64.32	94.06	76.44	**64.34**	74.05
3) w/o $\mathcal{L}_{MSE}$	94.77	62.05	93.89	**77.68**	63.12	**76.77**
4) DoDNet [21]	**96.87**	**65.47**	**96.52**	77.59	62.42	73.39
5) Ours	**95.78**	**68.07**	**94.74**	**78.46**	**64.30**	**76.48**

Method	Pancreas (224/57)		Colon (100/26)	Lungs (50/13)	Spleen (32/9)	Average
	Organ	Tumor	Tumor	Tumor	Organ	
1) Baseline	73.16	54.88	42.66	63.62	92.01	70.05
2) MT	78.15	56.01	41.58	66.15	**93.94**	73.08
3) w/o $\mathcal{L}_{MSE}$	77.44	56.98	46.59	69.88	92.76	73.81
4) DoDNet [21]	**82.64**	**60.45**	**51.55**	**71.25**	93.91	**75.64**
5) Ours	**81.51**	**62.15**	**49.66**	**75.52**	**94.59**	**76.48**

3.3 Evaluation

We have evaluated the proposed method on MOTS benchmark [21] and provided a thorough comparison with several partially-labeled segmentation methods. The comparison includes: 1) Baseline: a SwinUNETR segmentation network with dynamic segmentation heads; 2) MT: a mean teacher framework, which does not involve the identification of latent classes compared with our method; 3) w/o $\mathcal{L}_{MSE}$: our framework without the MSE loss on ambiguous classes; 4) DoDNet [21]: current state-of-the-art method on MOTS benchmark, and 5) Ours: our proposed framework with latent class mining. It should be noted that DoDNet employs different training and post-processing strategies from ours, but the quantitative comparison could prove the capacity and potential of our method on the MOTS benchmark.

The quantitative results on the test set of the MOTS benchmark have been shown in Table 1. It can be seen that our proposed method outperforms other alternatives in the multi-organ and tumor segmentation tasks, which demonstrates the effectiveness of the proposed training diagram. Overall, the introduction of latent class mining and the utilization of ambiguous classes could provide moderate performance gain in all kinds of organs or tumors. It should be noted that the performance gain in spleen segmentation is more significant, as the training data is insufficient (only 32 cases). Another interesting observation is that the segmentation performance in tumors, such as lung tumors, has also been further improved. Since we utilize BiomedCLIP text embeddings as task encoding for segmentation, the relations between different kinds of tumors may benefit the overall performance in tumor segmentation. Another plausible explanation

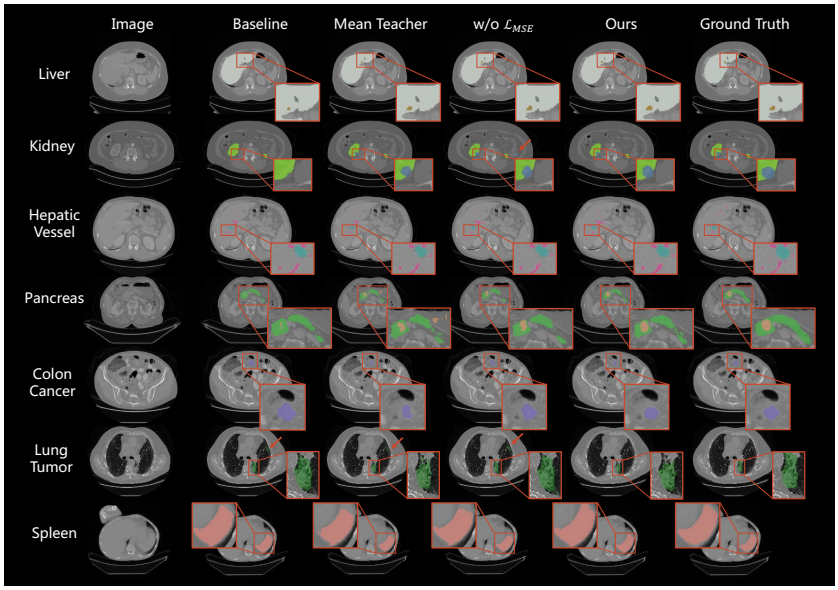

Fig. 2. Visualization of segmentation results of different algorithms on the MOTS dataset.

is that the segmentation model benefits unlabeled lung tumors in liver cancer scans, as lung cancer is a common metastasis of liver tumors [8, 16].

In addition, we build up two additional experimental settings to investigate the effectiveness of latent classes and ambiguous classes in the images, which are shown as No.2 and No.3 in the table. The mean teacher framework does not distinguish latent classes or ambiguous classes and regularizes the predictions of all classes with consistency loss. This naive implementation could further mislead the model when the student model has not been well supervised. By utilizing the ROI-aware network, our method could identify latent classes and apply CE loss on student predictions and "sharpened" pseudo labels, which could reduce the entropy in student's prediction and promote faster convergence. Thus, when comparing experiments No.2 and No.4 experiments, our method could achieve a better performance compared to the vanilla mean teacher. On the other hand, model performance could still benefit ambiguous classes as well. Although the existence of ambiguous classes is not clear, the consistency regularization of students and teachers could still force the model to learn a robust segmentation, which improves the overall quality of segmentation as well. Thus, in No.3 experiment, the segmentation performance deteriorates when the consistency loss $\mathcal{L}_{MSE}$ in ambiguous classes is removed.

In Fig. 2, we have visualized the segmentation comparison of our proposed segmentation method and other alternatives. The qualitative results in Fig. 2 indicate that our method could achieve more accurate segmentation with better quality.

4 Conclusion

In this paper, we have presented a novel framework that leverages latent classes in the partially labeled medical images for training the segmentation model. Our method utilizes existing but unlabeled ROIs in the input images to further supervise the segmentation training in an ROI-aware manner. The ROI-aware network makes possible the accurate and robust identification of latent classes in the input images, while the teacher-student framework is optimized by applying different constraints of the predictions based on ROI existence. Experimental results on the MOTS benchmark demonstrate the effectiveness of our proposed method. Since the development of future foundation models in medical image segmentation could be demanding in annotated data, our approach has great potential to accelerate the development of medical image segmentation and promote the success of the universal foundation model in medical image analysis.

Acknowledgments. This research is supported by the National Key R&D Program of China (Grant No. 2023YFB4706300) and the Fundamental Research Funds for the Central Universities (YG2023LC07).

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

References

1. Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G., et al.: The liver tumor segmentation benchmark (lits). Medical Image Analysis **84**, 102680 (2023)
2. Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., et al.: On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258 (2021)
3. Fang, X., Yan, P.: Multi-organ segmentation over partially labeled datasets with multi-scale feature abstraction. IEEE Transactions on Medical Imaging **39**(11), 3619–3629 (2020)
4. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI Brainlesion Workshop. pp. 272–284. Springer (2021)
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
6. Heller, N., Sathianathen, N., Kalapara, A., Walczak, E., Moore, K., Kaluzniak, H., Rosenberg, J., Blake, P., Rengel, Z., Oestreich, M., et al.: The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint arXiv:1904.00445 (2019)
7. Kaur, G., Chhaterji, J.: A survey on medical image segmentation. International Journal of Science and Research **6**(4), 1305–1311 (2017)
8. Lee, Y.T.M., Geer, D.A.: Primary liver cancer: pattern of metastasis. Journal of surgical oncology **36**(1), 26–31 (1987)

9. Liu, J., Zhang, Y., Chen, J.N., Xiao, J., Lu, Y., A Landman, B., Yuan, Y., Yuille, A., Tang, Y., Zhou, Z.: Clip-driven universal model for organ segmentation and tumor detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 21152–21164 (2023)
10. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101 (2017)
11. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024)
12. Moor, M., Banerjee, O., Abad, Z.S.H., Krumholz, H.M., Leskovec, J., Topol, E.J., Rajpurkar, P.: Foundation models for generalist medical artificial intelligence. Nature **616**(7956), 259–265 (2023)
13. Shi, G., Xiao, L., Chen, Y., Zhou, S.K.: Marginal loss and exclusion loss for partially supervised multi-organ segmentation. Medical Image Analysis **70**, 101979 (2021)
14. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., Van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., et al.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063 (2019)
15. Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in neural information processing systems **30** (2017)
16. Tsilimigras, D.I., Brodt, P., Clavien, P.A., Muschel, R.J., D'Angelica, M.I., Endo, I., Parks, R.W., Doyle, M., de Santibanes, E., Pawlik, T.M.: Liver metastases. Nature reviews Disease primers **7**(1), 27 (2021)
17. Wasserthal, J., Breit, H.C., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., et al.: Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images. Radiology: Artificial Intelligence **5**(5) (2023)
18. Wu, Y., Ge, Z., Zhang, D., Xu, M., Zhang, L., Xia, Y., Cai, J.: Mutual consistency learning for semi-supervised medical image segmentation. Medical Image Analysis **81**, 102530 (2022)
19. Xu, Z., Wang, Y., Lu, D., Luo, X., Yan, J., Zheng, Y., Tong, R.K.y.: Ambiguity-selective consistency regularization for mean-teacher semi-supervised medical image segmentation. Medical Image Analysis **88**, 102880 (2023)
20. Zhang, G., Yang, Z., Huo, B., Chai, S., Jiang, S.: Automatic segmentation of organs at risk and tumors in ct images of lung cancer from partially labelled datasets with a semi-supervised conditional nnu-net. Computer methods and programs in biomedicine **211**, 106419 (2021)
21. Zhang, J., Xie, Y., Xia, Y., Shen, C.: Dodnet: Learning to segment multi-organ and tumors from multiple partially labeled datasets. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1195–1204 (2021)
22. Zhang, S., Xu, Y., Usuyama, N., Bagga, J., Tinn, R., Preston, S., Rao, R., Wei, M., Valluri, N., Wong, C., et al.: Large-scale domain-specific pretraining for biomedical vision-language processing. arXiv preprint arXiv:2303.00915 (2023)