Segment Any Cancer in CT scans through equipping SAM with cross-slice interaction and indicator prompt

Xian Lin[1[0000−0001−8291−4823]*], Zhehao Wang[1[0009−0006−5249−4171]*], Caozhi Shang[1[0009−0001−9491−895X]], Junjie Shi[1[0009−0007−1519−0811]], Zengqiang Yan[1[0000−0002−2039−3863]], and Li Yu[1[0000−0002−5060−2558]]

School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
z_yan@hust.edu.cn

Abstract. CT scanning has become a commonly used imaging method in cancer diagnosis, displaying the anatomical structure of the human body in detail. However, different types of cancer have varying manifestations in CT imaging, posing great challenges for pan-cancer segmentation in CT scans. Recently, SAM has gradually become a landmark in medical image segmentation due to its powerful generality and generalization, providing a new paradigm for universal segmentation. However, current SAM-based segmentation approaches have weak detail perception ability, heavy dependence on manual prompts, and lack of 3D feature interaction. To address these, we have developed a universal cancer segmentation model for CT scans based on the extraordinary segmentation paradigm of SAM. Specifically, a 3D CNN-based U-shaped image encoder and a cross-branch interaction module are developed to increase the detail feature capture and spatial feature interaction of SAM. Besides, a cancer indicator prompt encoder is introduced to remove the dependence of SAM-based approaches on manual prompts. To fully utilize the advantages of SAM-based universal segmentation models and UNet-based specific segmentation models, we have comprehensively considered the prediction results of both, further reducing false positives and omissions in pan-cancer segmentation. In addition, to fully utilize partially annotated data for specific cancers, we use a combination of pseudo labels and partial labels to generate fully annotated data, effectively avoiding data conflict issues. Our method achieved an average score of 30% and 22% for the lesion DSC and NSD on the validation set and the average running time and area under GPU memory-time curve are 18s and 38960MB, respectively.

Keywords: Pan-cancer segmentation · SAM · Auto prompt · Foundation model · Cross-slice interaction.

* Xian Lin and Zhehao Wang should be considered joint first authors.

1 Introduction

Cancer, a disease caused by the loss of normal regulation and excessive proliferation of human cells, poses a serious threat to human health and is one of the leading causes of death worldwide [8,23]. Common cancers include breast cancer, lung cancer, colon cancer, rectal cancer, prostate cancer, skin cancer, and stomach cancer, covering the whole body [34]. Early diagnosis and screening, as well as timely treatment, can reduce the mortality rate of cancer [33]. Computer tomography (CT) can provide important information on human tissue structure and is widely used in cancer diagnosis and treatment. In clinical practice, radiologists and clinical doctors manually identify and measure abnormal areas based on CT images [3]. However, this manual detection method is time-consuming, labor-intensive, subjective, and highly dependent on experts. Consequently, developing an end-to-end cancer automatic detection algorithm using deep learning has extremely high clinical value [4].

Compared with conventional organ segmentation tasks, cancer segmentation mainly faces the following challenges [22]: (1) complex and variable shapes with weak regularity; (2) Some cancers have low contrast and blurred boundaries; (3) The location has diversity and may exist in multiple locations simultaneously. Compared with specialized segmentation tasks, the general segmentation task of whole-body pan-cancer segmentation has the following challenges [21]: (1) difficulty in collecting high-quality datasets. The training data for pan-cancer segmentation usually comes from different sources, with different purposes and annotated cancers. Most data sources have only partially annotated pan-cancer; (2) The differences in pan-cancer characteristics are prominent. The visual appearance of different types of cancer varies greatly, and the same cancer also has significant differences among individuals.

To fully utilize various data from different sources, some efforts have been made to develop model training approaches under partial labels. Chen et al. [37] co-train a het-erogeneous 3D network on multiple partially labeled datasets with a task-shared encoder. Huang et al. [36] introduce weight-averaged models for unified multi-organ segmentation on few-organ datasets. Xie et al. [45] propose TransDoDNet, which introduces a dynamic head to enable the network to accomplish multiple segmentation tasks flexibly and can be trained under partially labeled training data. In addition, to identify cancers of various shapes and appearances in complex backgrounds, plenty of deep learning-based approaches have been proposed for cancer segmentation, demonstrating enormous potential [9,24]. However, these models are tailored for specific cancers, and when applied to other types of cancers, new model parameters need to be trained, which brings great inconvenience to the task of whole-body pan-cancer segmentation [7]. The segment any model (SAM), a foundation model for universal segmentation, has received considerable praise for its excellent segmentation ability across different objects and powerful zero-shot generalization ability [1]. Based on user manual prompts, including points, bounding boxes, and coarse masks, SAM can segment corresponding objects. Therefore, with simple prompts, SAM can effortlessly adapt to various segmentation tasks [43]. This mode can integrate

multiple individual medical image segmentation tasks into a unified framework, greatly facilitating clinical deployment and providing a new perspective for developing the pan-cancer segmentation model.

Due to the lack of reliable clinical annotations, the performance of SAM in the medical field will rapidly decline [17]. Some foundation models adapt SAM to the field of medical image segmentation by tuning SAM on medical datasets [19,18]. However, these approaches, are prone to disrupting the crucial detail features for identifying small objects and boundaries, making it difficult to segment various cancers with complex shapes, weak boundaries, small sizes, or low contrast . Besides, these SAM-based models require the manual provision of task-related prompts to generate target masks, resulting in semi-automatic pipeline segmentation, which is inconvenient when dealing with pan-cancer segmentation tasks.

In this paper, we propose SAMCancer, a pan-cancer segmentation foundation model that supplements local features to SAM to segment any cancer and introduces a task-indicator prompt encoder for realizing end-to-end automatic segmentation. Specifically, SAMCancer consists of the original SAM, a 3D U-shaped CNN module, a cross-branch interaction module, and a task-indicator prompt encoder. To inherit the powerful feature representation capability of SAM, the structure of SAM has been preserved. To better identify cancers with complex shapes, low contrast, and varying sizes in CT images, we introduced a 3D U-shaped CNN sub-network to capture local features and placed it in parallel with the ViT image encoder of SAM. Then, the cross-branch interaction module is strategically positioned between the ViT-branch and the U-shaped CNN-branch to promote their feature representation ability by exchanging their global semantics and local information with each other. In addition, we extend the SAM-based model to an automatic segmentation model by introducing the task-indicator prompt encoder. Experimental results demonstrate the effectiveness of the proposed SAMCancer.

2 Method

2.1 Preprocessing

Following SAM-based approaches, the 3D CT scan is converted to 2D slices across the coronal plane to match the inputs of SAM-based models. No resampling method is used in our data preprocessing. As the intensity range of CT is usually (-1024, 2048), directly compressing such a large range to (0, 255) may lose valuable information. Therefore, under the guidance of experts, we choose different density windows for objects during training. Specifically, the intensity ranges of tissue and lung are set as (-200, 300) and (-1300, 300). The intensity range of (-200, 300) is used for inference.

2.2 Proposed Method

Network architecture: As depicted in Fig. 1, the proposed SAMCancer consists of the original SAM, a 3D U-shaped CNN image encoder, a cross-branch interaction

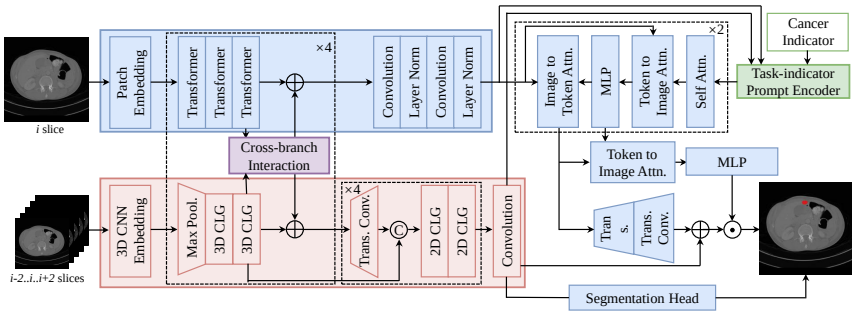

Fig. 1. Network architecture. Based on SAM, a 3D U-shaped CNN image encoder and a cross-branch interaction module are proposed to enhance the feature representation ability of the model, and a task-indicator prompt encoder is introduced to realize end-to-end automatic segmentation.

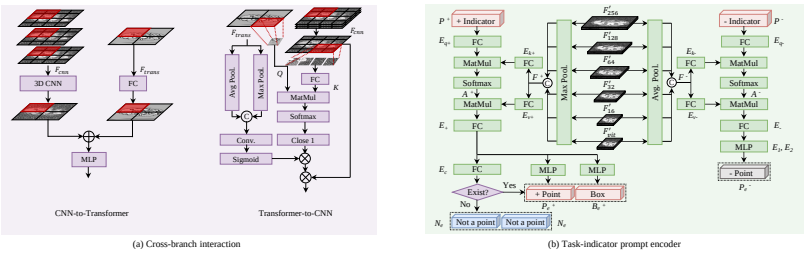

Fig. 2. Details of the cross-branch interaction and task-indicator prompt encoder.

module, and a task-indicator prompt encoder. The input of SAMCancer consists of the target slice and its four adjacent slices above and below it. The ViT-branch image encoder is responsible for extracting the feature of the target slice, while the 3D U-shaped CNN image encoder is used to extract abundant detail features of all input slices. By introducing local features of adjacent slices, the model can better cope with cancers with difficult appearances. In addition, we introduce a cross-branch interaction module as a bridge for information exchange between ViT-branch and CNN-branch, as depicted in Fig. 2, to further enhance the feature extraction capability of the model. After encoding image features, a task-indicator prompt encoder (details are presented in Fig. 2) is applied to prompt the mask decoder by the learnable cancer indicators. In addition to the mask decoder, we also adopt a segmentation head to predict the segmentation result based on the output features of the CNN-branch. Finally, the segmentation results of cancer are determined by the combined prediction results of the mask decoder and segmentation head.

Loss function: we use the summation between Dice loss and cross-entropy loss because compound loss functions have proven robust in various medical image segmentation tasks [25].

Table 1. Development environments and requirements.

System	Ubuntu 20.04.6 LTS
CPU	Intel(R) Xeon(R) Gold 6143 CPU @ 2.80GHz
RAM	32×8GB; 2666MT/s
GPU (number and type)	Two NVIDIA RTX3090 24G
CUDA version	11.1
Programming language	Python 3.8
Deep learning framework	torch 1.8.0, torchvision 0.9.0
Specific dependencies	N/A
Code	https://github.com/xianlin7/SAMCT

Other strategies: We reduce false positives on CT scans from healthy patients by introducing a classifier in the task-indicator prompt encoder. For the inputs of healthy patients, we do not provide positive prompt embeddings for the mask decoder. We use the partial labels to train the model first and obtain the pseudo labels of each scan. Then, we combine the combination of pseudo labels and partial labels to generate the fully annotated data. The whole process is simple, and our focus is more on the model structure. Unlabeled images were not used in our solution. We did not use the pseudo labels generated by the FLARE23 winning algorithm. We did not adopt any additional acceleration strategies for the inference process. However, by using only a small number of adjacent slices, our method can achieve acceptable inference speed and resource consumption.

2.3 Post-processing

We directly used the output of the model as the result without any post-processing.

3 Experiments

3.1 Dataset and evaluation measures

The segmentation targets cover various lesions. The training dataset is curated from more than 50 medical centers under the license permission, including TCIA [6], LiTS [5], MSD [38], KiTS [13,15,14], autoPET [12,11], TotalSegmentator [39], and AbdomenCT-1K [31], FLARE 2023 [30], DeepLesion [42], COVID-19-CT-Seg-Benchmark [28], COVID-19-20 [35], CHOS [20], LNDB [32], and LIDC [2]. The training set includes 4000 abdomen CT scans where 2200 CT scans with partial labels and 1800 CT scans without labels. The validation and testing sets include 100 and 400 CT scans, respectively, which cover various abdominal cancer types, such as liver cancer, kidney cancer, pancreas cancer, colon cancer, gastric cancer, and so on. The lesion annotation process used ITK-SNAP [44], nnU-Net [16], MedSAM [26], and Slicer Plugins [10,27].

The evaluation metrics encompass two accuracy measures—Dice Similarity Coefficient (DSC) and Normalized Surface Dice (NSD)—alongside two efficiency

measures—running time and area under the GPU memory-time curve. These metrics collectively contribute to the ranking computation. Furthermore, the running time and GPU memory consumption are considered within tolerances of 45 seconds and 4 GB, respectively.

Table 2. Training protocols.

Network initialization	He
Batch size	16
Patch size	5×256×256
Total epochs	80
Optimizer	Adam
Initial learning rate (lr)	0.0005
Lr decay schedule	Periodic decay
Training time	120 hours
Loss function	Dice loss and cross-entropy loss
Number of model parameters	153.53M*
Number of flops	106.83G*
CO_2eq	26.54 Kg*

Table 3. Quantitative evaluation results.

Methods	Public Validation		Online Validation		Testing	
	DSC(%)	NSD(%)	DSC(%)	NSD(%)	DSC(%)	NSD (%)
SAMCancer	27.19 ± 21.20	15.07 ± 15.79	29.92	21.92	25.21	15.72

Table 4. Ablation study.

Methods	Public Validation		Online Validation	
	DSC(%)	NSD(%)	DSC(%)	NSD(%)
one slice	26.72 ± 19.82	14.25 ± 13.51	27.96	18.91
five slices	27.19 ± 21.20	15.07 ± 15.79	29.92	21.92

3.2 Implementation details

Environment settings The development environments and requirements are presented in Table 1.

Training protocols Due to the amount of partial labeled data is large, we did not use unlabeled data. For the partial labels, We divide them into 5 folds for training and obtain pseudo labels for the validation data of each fold, separately. By combining the pseudo labels with partial labels, we obtain fully annotated data. Then, we train the SAMCancer with the fully annotated data. For data augmentation, methods including contrast adjustment, gamma augmentation, random rotation, and scaling are adopted. We take turns using each slice as the target slice and then combine it with its four adjacent slices as a patch after resizing their sizes into 256×256. No special patch sampling strategy was adopted. The best-performing model on the local validation set is selected as the optimal model. More details of the training protocol are presented in Table 2.

4 Results and discussion

4.1 Quantitative results on validation set

Quantitative results are summarized in Table 3. Our method achieves a mean DSC of 27.19% and a NSD of 15.07% on the FLARE 2024 public validation dataset. On the FLARE 2024 online validation dataset, the proposed approach achieves a mean DSC of 29.92% and a NSD of 21.92%.

Table 5. Quantitative evaluation of segmentation efficiency in terms of the running them and GPU memory consumption. Total GPU denotes the area under GPU Memory-Time curve. Evaluation GPU platform: NVIDIA QUADRO RTX5000 (16G).

Case ID	Image Size	Running Time (s)	Max GPU (MB)	Total GPU (MB)
0001	(512, 512, 55)	30.45	3563	69021
0051	(512, 512, 100)	12.99	3563	22074
0017	(512, 512, 150)	13.44	3563	22249
0019	(512, 512, 215)	13.71	3563	23359
0099	(512, 512, 334)	15.87	3563	31983
0063	(512, 512, 448)	13.18	3563	22280
0048	(512, 512, 499)	29.67	3563	79407
0029	(512, 512, 554)	25.26	3563	64418

Ablation study results are summarized in Table 4. Compared to the single slice input of SAM-based methods, the structure that support multiple slice inputs achieve better performance, indicating the effectiveness of the proposed SAMCancer.

4.2 Qualitative results on validation set

Qualitative results are depicted in Fig. 3. By injecting more planar and spatial details into SAM, the proposed method can effectively reduce false positives on the background and improve the accuracy of cancer recognition under low

contrast. However, due to the limitations of the dataset, the proposed method performs poorly on discrete small cancers and rare regional cancers.

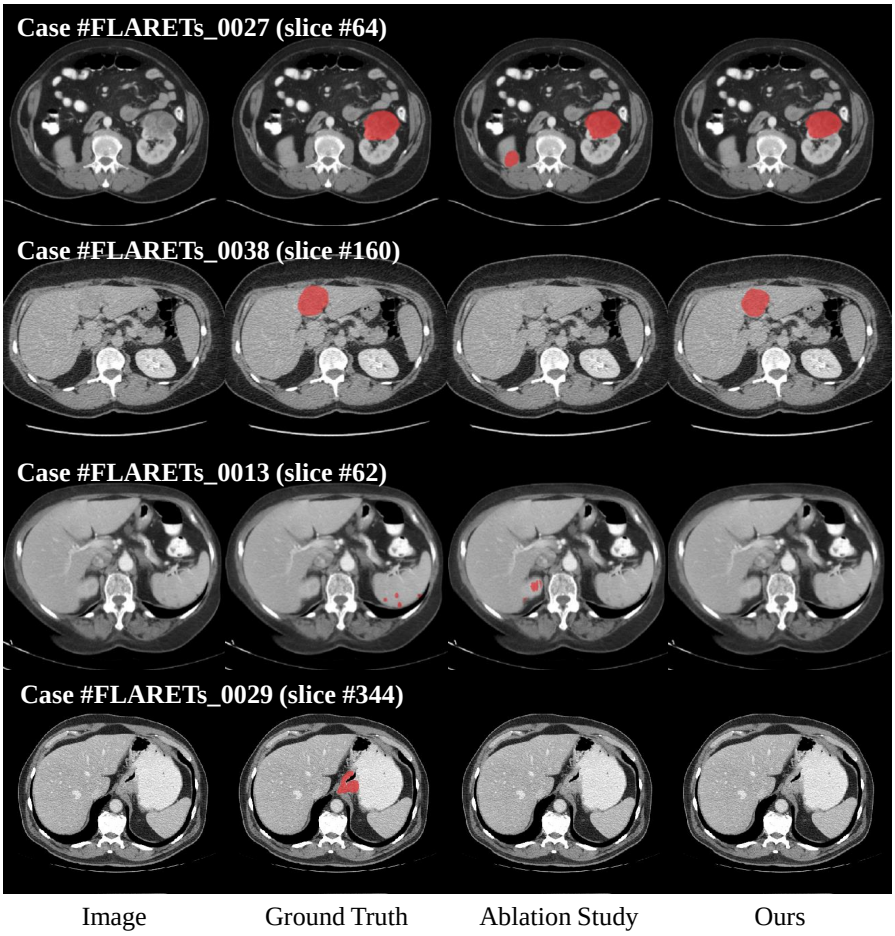

Fig. 3. Qualitative results.

4.3 Segmentation efficiency results on validation set

The average running time in online validation dataset is 18.05s per case in inference phase, and average used GPU memory is 3563 MB. The area under GPU memory-time curve is 38960. Table 5 lists segmentation efficiency of some typical cases. In addition, the false positive rate of the proposed approach on the healthy CT scans is 0.04 ± 0.05.

4.4 Results on final testing set

This is a placeholder. We will send you the testing results during MICCAI.

4.5 Limitation and future work

Our method did not utilize the unlabeled data and did not fully utilize the partially labeled data. In addition, our data preprocessing at the 3D level is not rich. In the future, we will explore how to fully utilize available data for model training under any annotation type. In addition, it is also important for universal medical image segmentation to fully preprocess various 3D data within a unified framework and with the automatic hyperparameter setting.

5 Conclusion

In this work, we design a pan-cancer segmentation foundation model based on SAM to segment various cancers. By introducing the 3D U-shaped CNN encoder and the cross-branch interaction module, we can promote the model to recognize various cancers with complex appearances. Besides, introducing the task-indicator prompt encoder makes the SAM-based model an end-to-end automatic pipeline. These designs may be helpful for other universal medical image segmentation tasks.

Acknowledgements The authors of this paper declare that the segmentation method they implemented for participation in the FLARE 2024 challenge has not used any pre-trained models nor additional datasets other than those provided by the organizers. The proposed solution is fully automatic without any manual intervention. We thank all data owners for making the CT scans publicly available and CodaLab [41] for hosting the challenge platform.

Disclosure of Interests

The authors declare no competing interests.

References

1. Alexander, K., Eric, M., Nikhila, R., Hanzi, M., Chloe, R., Laura, G., Tete, X., Spencer, W., Alexander C, B., Wan-Yen, L., et al.: Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4015–4026 (2023) 2
2. Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al.: The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical Physics **38**(2), 915–931 (2011) 5

3. Benjamin, H., Sumeet, H., Richard W, L.: The role of artificial intelligence in early cancer diagnosis. Cancers **14**(6), 1524 (2022) 2
4. Bhavneet, B., Coryandar, G., Neel S, M., Olivier, E.: Artificial intelligence in cancer research and precision medicine. Cancer Discovery **11**(4), 900–915 (2021) 2
5. Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G., Lohöfer, F., Holch, J.W., Sommer, W., Hofmann, F., Hostettler, A., Lev-Cohain, N., Drozdzal, M., Amitai, M.M., Vivanti, R., Sosna, J., Ezhov, I., Sekuboyina, A., Navarro, F., Kofler, F., Paetzold, J.C., Shit, S., Hu, X., Lipková, J., Rempfler, M., Piraud, M., Kirschke, J., Wiestler, B., Zhang, Z., Hülsemeyer, C., Beetz, M., Ettlinger, F., Antonelli, M., Bae, W., Bellver, M., Bi, L., Chen, H., Chlebus, G., Dam, E.B., Dou, Q., Fu, C.W., Georgescu, B., i Nieto, X.G., Gruen, F., Han, X., Heng, P.A., Hesser, J., Moltz, J.H., Igel, C., Isensee, F., Jäger, P., Jia, F., Kaluva, K.C., Khened, M., Kim, I., Kim, J.H., Kim, S., Kohl, S., Konopczynski, T., Kori, A., Krishnamurthi, G., Li, F., Li, H., Li, J., Li, X., Lowengrub, J., Ma, J., Maier-Hein, K., Maninis, K.K., Meine, H., Merhof, D., Pai, A., Perslev, M., Petersen, J., Pont-Tuset, J., Qi, J., Qi, X., Rippel, O., Roth, K., Sarasua, I., Schenk, A., Shen, Z., Torres, J., Wachinger, C., Wang, C., Weninger, L., Wu, J., Xu, D., Yang, X., Yu, S.C.H., Yuan, Y., Yue, M., Zhang, L., Cardoso, J., Bakas, S., Braren, R., Heinemann, V., Pal, C., Tang, A., Kadoury, S., Soler, L., van Ginneken, B., Greenspan, H., Joskowicz, L., Menze, B.: The liver tumor segmentation benchmark (lits). Medical Image Analysis **84**, 102680 (2023) 5
6. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F.: The cancer imaging archive (tcia): maintaining and operating a public information repository. Journal of Digital Imaging **26**(6), 1045–1057 (2013) 5
7. Dengao, L., Xiaohui, C., Yanfen, C., Jumin, Z., Kenan, Z., Xiaotang, Y.: Improved u-net based on contour prediction for efficient segmentation of rectal cancer. Computer Methods and Programs in Biomedicine **213**, 106493 (2022) 2
8. Dingwen, Z., Jiajia, Z., Qiang, Z., Jungong, H., Shu, Z., Junwei, H.: Automatic pancreas segmentation based on lightweight dcnn modules and spatial prior propagation. Pattern Recognition **114**, 107762 (2021) 2
9. Essam H, H., Doaa A, A., Marwa M, E., Mohamed Abdel, H., Mina, Y.: An efficient image segmentation method for skin cancer imaging using improved golden jackal optimization algorithm. Computers in Biology and Medicine **149**, 106075 (2022) 2
10. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al.: 3d slicer as an image computing platform for the quantitative imaging network. Magnetic Resonance Imaging **30**(9), 1323–1341 (2012) 5
11. Gatidis, S., Früh, M., Fabritius, M., Gu, S., Nikolaou, K., Fougère, C.L., Ye, J., He, J., Peng, Y., Bi, L., Ma, J., Wang, B., Zhang, J., Huang, Y., Heiliger, L., Marinov, Z., Stiefelhagen, R., Egger, J., Kleesiek, J., Sibille, L., Xiang, L., Bendazolli, S., Astaraki, M., Schölkopf, B., Ingrisch, M., Cyran, C., Küstner, T.: The autopet challenge: towards fully automated lesion segmentation in oncologic pet/ct imaging. Nature Machine Intelligence (2023) 5
12. Gatidis, S., Hepp, T., Früh, M., La Fougère, C., Nikolaou, K., Pfannenberg, C., Schölkopf, B., Küstner, T., Cyran, C., Rubin, D.: A whole-body fdg-pet/ct dataset with manually annotated tumor lesions. Scientific Data **9**(1), 601 (2022) 5
13. Heller, N., Isensee, F., Maier-Hein, K.H., Hou, X., Xie, C., Li, F., Nan, Y., Mu, G., Lin, Z., Han, M., Yao, G., Gao, Y., Zhang, Y., Wang, Y., Hou, F., Yang, J.,

Xiong, G., Tian, J., Zhong, C., Ma, J., Rickman, J., Dean, J., Stai, B., Tejpaul, R., Oestreich, M., Blake, P., Kaluzniak, H., Raza, S., Rosenberg, J., Moore, K., Walczak, E., Rengel, Z., Edgerton, Z., Vasdev, R., Peterson, M., McSweeney, S., Peterson, S., Kalapara, A., Sathianathen, N., Papanikolopoulos, N., Weight, C.: The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. Medical Image Analysis **67**, 101821 (2021) 5

14. Heller, N., Isensee, F., Trofimova, D., Tejpaul, R., Zhao, Z., Chen, H., Wang, L., Golts, A., Khapun, D., Shats, D., Shoshan, Y., Gilboa-Solomon, F., George, Y., Yang, X., Zhang, J., Zhang, J., Xia, Y., Wu, M., Liu, Z., Walczak, E., McSweeney, S., Vasdev, R., Hornung, C., Solaiman, R., Schoephoerster, J., Abernathy, B., Wu, D., Abdulkadir, S., Byun, B., Spriggs, J., Struyk, G., Austin, A., Simpson, B., Hagstrom, M., Virnig, S., French, J., Venkatesh, N., Chan, S., Moore, K., Jacobsen, A., Austin, S., Austin, M., Regmi, S., Papanikolopoulos, N., Weight, C.: The kits21 challenge: Automatic segmentation of kidneys, renal tumors, and renal cysts in corticomedullary-phase ct. arXiv preprint arXiv:2307.01984 (2023) 5
15. Heller, N., McSweeney, S., Peterson, M.T., Peterson, S., Rickman, J., Stai, B., Tejpaul, R., Oestreich, M., Blake, P., Rosenberg, J., et al.: An international challenge to use artificial intelligence to define the state-of-the-art in kidney and kidney tumor segmentation in ct imaging. American Society of Clinical Oncology **38**(6), 626–626 (2020) 5
16. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (2021) 5
17. Jun, M., Yuting, H., Feifei, L., Lin, H., Chenyu, Y., Bo, W.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024) 3
18. Junde, W., Wei, J., Yuanpei, L., Huazhu, F., Min, X., Yanwu, X., Yueming, J.: Medical sam adapter: Adapting segment anything model for medical image segmentation. arXiv preprint arXiv:2304.12620 (2023) 3
19. Kaidong, Z., Dong, L.: Customized segment anything model for medical image segmentation. arXiv preprint arXiv:2304.13785 (2023) 3
20. Kavur, A.E., Gezer, N.S., Barış, M., Aslan, S., Conze, P.H., Groza, V., Pham, D.D., Chatterjee, S., Ernst, P., Özkan, S., Baydar, B., Lachinov, D., Han, S., Pauli, J., Isensee, F., Perkonigg, M., Sathish, R., Rajan, R., Sheet, D., Dovletov, G., Speck, O., Nürnberger, A., Maier-Hein, K.H., Bozdağı Akar, G., Ünal, G., Dicle, O., Selver, M.A.: CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. Medical Image Analysis **69**, 101950 (2021) 5
21. Ke, Y., Jinzheng, C., Youjing, Z., Adam P, H., Dakai, J., Youbao, T., Yuxing, T., Lingyun, H., Jing, X., Le, L.: Learning from multiple datasets with heterogeneous and partial labels for universal lesion detection in ct. IEEE Transactions on Medical Imaging **40**(10), 2759–2770 (2020) 2
22. Kleppe, A., Ole-Johan, S., Sepp, D.R., Knut, L., David J, K., E, H.: Designing deep learning studies in cancer diagnostics 2
23. Luís A Vale, S., Karl, R.: Pan-cancer prognosis prediction using multimodal deep learning. In: International Symposium on Biomedical Imaging. pp. 568–571 (2020) 2
24. Luqman, A., Muhammad Munwar, I., Hamza, A., Shehzad, K., Yasir, S., Saqib, S.: Images data practices for semantic segmentation of breast cancer using deep neural network. Journal of Ambient Intelligence and Humanized Computing **14**(11), 15227–15243 (2023) 2

25. Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., Martel, A.L.: Loss odyssey in medical image segmentation. Medical Image Analysis **71**, 102035 (2021) 4
26. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**, 654 (2024) 5
27. Ma, J., Kim, S., Li, F., Baharoon, M., Asakereh, R., Lyu, H., Wang, B.: Segment anything in medical images and videos: Benchmark and deployment. arXiv preprint arXiv:2408.03322 (2024) 5
28. Ma, J., Wang, Y., An, X., Ge, C., Yu, Z., Chen, J., Zhu, Q., Dong, G., He, J., He, Z., Cao, T., Zhu, Y., Nie, Z., Yang, X.: Towards data-efficient learning: A benchmark for covid-19 ct lung and infection segmentation. Medical Physics **48**(3), 1197–1210 (2021) 5
29. Ma, J., Zhang, Y., Gu, S., An, X., Wang, Z., Ge, C., Wang, C., Zhang, F., Wang, Y., Xu, Y., Gou, S., Thaler, F., Payer, C., Štern, D., Henderson, E.G., McSweeney, D.M., Green, A., Jackson, P., McIntosh, L., Nguyen, Q.C., Qayyum, A., Conze, P.H., Huang, Z., Zhou, Z., Fan, D.P., Xiong, H., Dong, G., Zhu, Q., He, J., Yang, X.: Fast and low-gpu-memory abdomen ct organ segmentation: The flare challenge. Medical Image Analysis **82**, 102616 (2022)
30. Ma, J., Zhang, Y., Gu, S., Ge, C., Wang, E., Zhou, Q., Huang, Z., Lyu, P., He, J., Wang, B.: Automatic organ and pan-cancer segmentation in abdomen ct: the flare 2023 challenge. arXiv preprint arXiv:2408.12534 (2024) 5
31. Ma, J., Zhang, Y., Gu, S., Zhu, C., Ge, C., Zhang, Y., An, X., Wang, C., Wang, Q., Liu, X., Cao, S., Zhang, Q., Liu, S., Wang, Y., Li, Y., He, J., Yang, X.: Abdomenct-1k: Is abdominal organ segmentation a solved problem? IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(10), 6695–6714 (2022) 5
32. Pedrosa, J., Aresta, G., Ferreira, C., Atwal, G., Phoulady, H.A., Chen, X., Chen, R., Li, J., Wang, L., Galdran, A., et al.: Lndb challenge on automatic lung cancer patient management. Medical Image Analysis **70**, 102027 (2021) 5
33. Rebecca C, F., Antonis C, A., Ljiljana, F., Nitzan, R.: The future of early cancer detection. Nature Medicine **28**(4), 666–677 (2022) 2
34. Reem, N., Horacio, B.: Current and future development in lung cancer diagnosis. International Journal of Molecular Sciences **22**(16), 8661 (2021) 2
35. Roth, H.R., Xu, Z., Tor-Díez, C., Jacob, R.S., Zember, J., Molto, J., Li, W., Xu, S., Turkbey, B., Turkbey, E., et al.: Rapid artificial intelligence solutions in a pandemic—the covid-19-20 lung ct lesion segmentation challenge. Medical Image Analysis **82**, 102605 (2022) 5
36. Rui, H., Yuanjie, Z., Zhiqiang, H., Shaoting, Z., Hongsheng, L.: Multi-organ segmentation via co-training weight-averaged models from few-organ datasets. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 146–155 (2020) 2
37. Sihong, C., Kai, M., Yefeng, Z.: Med3d: Transfer learning for 3d medical image analysis. arXiv preprint arXiv:1904.00625 (2019) 2
38. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., Bilic, P., Christ, P.F., Do, R.K.G., Gollub, M., Golia-Pernicka, J., Heckers, S.H., Jarnagin, W.R., McHugo, M.K., Napel, S., Vorontsov, E., Maier-Hein, L., Cardoso, M.J.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063 (2019) 5

39. Wasserthal, J., Breit, H.C., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., Bach, M., Segeroth, M.: Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images. Radiology: Artificial Intelligence **5**(5), e230024 (2023) 5
40. Wu, Y., Wang, E., Shao, Z.: Fast abdomen organ and tumor segmentation with nn-unet. In: MICCAI Challenge on Fast and Low-Resource Semi-supervised Abdominal Organ Segmentation, pp. 1–14. Springer (2023)
41. Xu, Z., Escalera, S., Pavão, A., Richard, M., Tu, W.W., Yao, Q., Zhao, H., Guyon, I.: Codabench: Flexible, easy-to-use, and reproducible meta-benchmark platform. Patterns **3**(7), 100543 (2022) 9
42. Yan, K., Wang, X., Lu, L., Summers, R.M.: Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. Journal of Medical Imaging **5**(3), 036501–036501 (2018) 5
43. Yuhao, H., Xin, Y., Lian, L., Han, Z., Ao, C., Xinrui, Z., Rusi, C., Junxuan, Y., Jiongquan, C., Chaoyu, C., et al.: Segment anything model for medical images? Medical Image Analysis **92**, 103061 (2024) 2
44. Yushkevich, P.A., Gao, Y., Gerig, G.: Itk-snap: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 3342–3345 (2016) 5
45. Yutong, X., Jianpeng, Z., Yong, X., Chunhua, S.: Learning from partially labeled data for multi-organ and tumor segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(12), 14905–14919 (2023) 2

Table 6. Checklist Table. Please fill out this checklist table in the answer column.

Requirements	Answer
A meaningful title	Yes
The number of authors ($\leq$6)	6
Author affiliations and ORCID	Yes
Corresponding author email is presented	Yes
Validation scores are presented in the abstract	Yes
Introduction includes at least three parts: background, related work, and motivation	Yes
A pipeline/network figure is provided	Figures 1&2
Pre-processing	Page 5
Strategies to use the partial label	Pages 5 & 6
Strategies to use the unlabeled images.	Page 5
Strategies to improve model inference	Page 5
Post-processing	Page 5
The dataset and evaluation metric section are presented	Page 5
Environment setting table is provided	Table 1
Training protocol table is provided	Table 2
Ablation study	Page 7
Efficiency evaluation results are provided	Table 5
Visualized segmentation example is provided	Figure 3
Limitation and future work are presented	Yes
Reference format is consistent.	Yes