

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation

Xinglong Liang^{1,2}, Jiaju Huang³, Luyi Han^{1,2}, Tianyu Zhang^{1,2}, Xin Wang^{1,4}, Yuan Gao^{1,4}, Chunyao Lu^{1,2}, Lishan Cai^{1,4}, Tao Tan^{3,1*}, and Ritse Mann^{1,2}

¹ Department of Radiology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands

² Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands

³ Faculty of Applied Sciences, Macao Polytechnic University, 999078, Macao Special Administrative Region of China

⁴ GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, P. Debyelaan 25, 6202 AZ, Maastricht, The Netherlands
taotan@mpu.edu.mo

Abstract. PET-CT lesion segmentation is challenging due to noise sensitivity, small and variable lesion morphology, and interference from physiological high-metabolic signals. Current mainstream approaches follow the practice of one network solving the segmentation of multiple cancer lesions by treating all cancers as a single task. However, this overlooks the unique characteristics of different cancer types. Considering the specificity and similarity of different cancers in terms of metastatic patterns, organ preferences, and FDG uptake intensity, we propose DpDNet, a Dual-Prompt-Driven network that incorporates specific prompts to capture cancer-specific features and common prompts to retain shared knowledge. Additionally, to mitigate information forgetting caused by the early introduction of prompts, prompt-aware heads are employed after the decoder to adaptively handle multiple segmentation tasks. Experiments on a PET-CT dataset with four cancer types show that DpDNet outperforms state-of-the-art models. Finally, based on the segmentation results, we calculated MTV, TLG, and SUVmax for breast cancer survival analysis. The results suggest that DpDNet has the potential to serve as a valuable tool for personalized risk stratification, supporting clinicians in optimizing treatment strategies and improving outcomes. Code is available at <https://github.com/XinglongLiang08/DpDNet>.

Keywords: Universal PET-CT Segmentation · Dual-Prompt · Survival analysis.

1 Introduction

PET-CT (Positron emission tomography and computed tomography) is widely used for cancer diagnosis and prognosis, offering both anatomical and metabolic

* Corresponding author

insights [7, 4]. PET-CT typically covers whole-body imaging while the lesions of interest are usually small, and fluorodeoxyglucose (FDG) uptake is not confined to tumor tissues but also appears physiologically in areas like the urinary tract and myocardium [8], making PET-CT interpretation a time-consuming task for radiologists. Therefore, developing an accurate PET-CT-based automated lesion segmentation model to assist radiologists in diagnosis, evaluating patient treatment response, and assessing prognosis is valuable.

In recent years, deep learning-based models have significantly advanced medical image segmentation [12, 11, 23, 9, 10, 3]. However, whole-body PET-CT segmentation still faces three major challenges. (1) Lesions in PET-CT range from small metastatic foci to large tumors, exhibiting diverse shapes, sizes, and locations, which makes automatic segmentation challenging. (2) The scarcity of PET-CT data leads to training on mixed PET-CT images without differentiating between cancer types, overlooking the specific characteristics of different cancers. (3) The lack of clinical evaluation of segmentation results, such as their impact on diagnostic accuracy and prognosis prediction, may prevent these models from gaining the trust of healthcare professionals and being adopted in practice.

Prompt engineering has the potential to be a promising tool for addressing these challenges in PET-CT segmentation. On one hand, it helps address the challenges posed by the small size and partially labeled nature of datasets [16]. On the other hand, prompt engineering integrates prior knowledge into models, enabling them to more effectively identify and segment regions of interest. A typical example is CLIP-driven universal models [16, 27, 28], which incorporate the text embeddings of all labels as external knowledge, obtained by feeding medical prompts into CLIP. However, CLIP, trained on natural language text, has limited capacity for understanding medical terminology, and even a slight change in wording can significantly impact its performance [29]. Another groundbreaking work, DoDNet [25, 21] can perform various segmentation tasks by utilizing task encoding and a controller to generate dynamic convolutions, but it suffers from delayed task awareness and the inability to leverage correlations among tasks. UniSeg [23, 22] introduces a learnable universal prompt to capture correlations among tasks. This prompt is converted into a task-specific prompt and fed into the decoder, making the model task-aware early and improving task-specific performance. However, prematurely introducing prompts may lead to information forgetting, and using a shared segmentation head limits task-specific adaptability.

In this paper, we propose a dual-prompt-driven model for universal whole-body PET-CT segmentation. Previous methods overlook the heterogeneity among different cancer types, such as variations in metastatic patterns, metabolic activity as reflected by FDG uptake, and anatomical preferences for metastases. For instance, certain cancers like breast cancer often exhibit a distinct pattern of bone metastases [13], while others, such as lung cancer, may preferentially spread to the brain or adrenal glands [20]. These differences in biological behavior and imaging characteristics could impact the performance and generalizability of lesion segmentation models. On the other hand, metastases from different cancers

are also interrelated, as they may share similar biological mechanisms of dissemination and exhibit overlapping imaging characteristics [2]. Based on these two considerations, we designed a dual-prompt strategy, comprising task-specific prompts and a common prompt. The specific prompts are tailored to capture the unique characteristics of metastases associated with each cancer type, while the common prompt leverages shared features and interrelations across different cancers to enhance the generalization of the model. These two prompts are fed into the decoder to introduce prompt information to the model early in the process. Additionally, we further designed prompt-aware heads to mitigate information forgetting and adaptively handle multiple segmentation tasks while avoiding a significant increase in model parameters. The head incorporates channel attention and multi-scale branches to enhance small-target feature extraction. Finally, we applied the segmentation results to breast cancer survival analysis to explore its potential clinical applications.

Our contributions are three-fold: (1) We design a dual-prompt strategy with specific and common prompts to capture cancer-specific metastatic patterns while leveraging shared features across different cancer types. (2) Prompt-aware heads are employed to replace the shared segmentation head, enabling adaptive task handling and enhancing small-target feature extraction via channel attention and multi-scale branches. (3) Predefined parameters are automatically calculated from segmentation results on a large dataset for breast cancer survival analysis, demonstrating our model’s generalizability and its potential for risk stratification and survival prediction.

2 Methodology

2.1 Problem Definition

Let $\{C_1, C_2, \dots, C_N\}$ be N datasets of different cancer types. Here, $C_i = \{X_{ij}, Y_{ij}\}_{j=1}^{n_i}$ represents that the i -th dataset has a total of n_i image-label pairs, and X_{ij} and Y_{ij} are the image and the corresponding ground truth, respectively. Typically, PET-CT-based cancer segmentation models are trained either by treating all cancer datasets as a single task or by focusing on a specific cancer type. The former approach fails to capture the distinct characteristics of different cancers, while the latter is constrained by limited data availability. To address these challenges, we propose a universal PET-CT segmentation framework, termed DpDNet. The overall architecture of DpDNet is shown in Fig. 1, with details introduced in the following subsections.

2.2 Encoder-decoder backbone

The backbone of DpDNet is based on STU-Net [11], a scalable and transferable large-scale medical segmentation model. STU-Net is based on the nnU-Net architecture but incorporates key modifications to the encoder, decoder, and upsampling processes. Residual connections are added to the basic blocks to address gradient diffusion issues, while downsampling is integrated into the first

Where F_{cani} denotes the prompt features belonging to the i -th cancer and F denotes the sample-specific features from the bottleneck. $\text{cat}(\cdot, \cdot)$ represents a concatenation operation, $f(\cdot)$ denotes the feed-forward processes, and $\text{Split}(\cdot)^N$ indicates splitting features along the channel to obtain N cancer-specific prompts. Then, we select the target prompt according to the ongoing task from $\{F_{can1}, \dots, F_{canN}\}$. Subsequently, we perform gated fusion on the remaining prompts, excluding the current task, to obtain F_{comi} , defined as:

$$F_{ini} = f_{2i}(\text{cat}(F_{can1}, \dots, F_{cani-1}, F_{cani+1}, \dots, F_{canN})) \quad (2)$$

$$F_{comi} = \text{Softmax}(F_{\text{cat}(F_{can1}, \dots, F_{cani-1}, F_{cani+1}, \dots, F_{canN})}) \odot F_{ini} \quad (3)$$

Finally, we concatenate F , the selected F_{cani} and common prompt F_{comi} as the decoder input. In this way, cancer-related prior information is introduced into the model to enhance the training of the entire decoder.

2.4 Prompt-aware Heads

Unlike the dynamic convolution heads in DoDNet and CLIP-Driven models, we designed prompt-aware heads, which replace the final stage of the decoder by selecting segmentation heads based on cancer-type prompts. Prompt-aware heads incorporate channel attention to capture inter-channel relationships, multi-scale branches improve feature extraction: the $1 \times 1 \times 1$ branch preserves original information to prevent loss, the $3 \times 3 \times 3$ branch captures local features for small-target edges and textures, and the $5 \times 5 \times 5$ convolution branch expands the receptive field. Compared to dynamic convolution heads (which apply only a single convolution after the decoder and may not fully capture diverse cancer characteristics), our design adapts earlier to different cancer types, enabling better feature learning while also simplifying training, especially for limited datasets. In contrast to methods that assign a separate decoder for each task, our approach balances task adaptability and efficiency without significantly increasing model parameters.

3 Experiments and Results

3.1 Datasets and Evaluation Metric

Datasets For this study, we collected whole-body PET-CT data from four types of cancer: lung cancer, lymphoma, breast cancer, and melanoma to train and test DpDNet. The lung cancer, lymphoma, and melanoma datasets were sourced from AutoPET [6], while the breast cancer dataset came from our private collection. The combined dataset includes a total of 697 cases, with 557 cases used for training and 140 cases for validation. For the survival prediction, we collected 1,210 nonmetastatic breast cancer patients, without ground truth. Median values of SUVmax, MTV, and TLG were computed from 480 patients and used to stratify the remaining 780 patients into high- and low-risk groups.

Evaluation Metric We evaluated the segmentation performance using four metrics: Dice similarity coefficient (DSC) and Intersection over Union (IoU). For the survival prediction, we used C-index and Hazard Ratio (HR) as the evaluation metric.

3.2 Implementation Details

During training, we adopted the SGD optimizer and set the batch size to 2, the initial learning rate to 0.0001, the default patch size to $112 \times 160 \times 128$ (while SwinUNet used $128 \times 128 \times 128$ due to network limitations), and the maximum training epoch to 1000 with a total of 250,000 iterations. All experiments were conducted on the NVIDIA Quadro RTX A6000 GPU.

3.3 Results

Comparing to Non-Prompt-Based and Prompt-Based Models Our DpD-Net was compared with the state-of-the-art Non-Prompt-Based and Prompt-Based models. Additionally, we trained nnU-Net separately on each cancer type for comparison. The Non-Prompt-Based models includes nnU-Net [12], STU-Net [11], UNETR [19], 3DUX-Net [14], SwinUNETRV2 [19], U-Mamba [17]. The Prompt-Based models includes CLIP-driven Universal Model [16], DoD-Net[25], UniSeg [23]. To balance network complexity and performance, we adopted STU-Net-S as the backbone for all prompt-based models. As shown in Table 1, our method achieves the highest average segmentation performance across multiple cancer types. Specifically, the average DSC surpasses the second-best method by 1.32% and the average IoU improves by 1.56%, while our approach achieves the best segmentation results in lung cancer, lymphoma, and breast cancer. Additionally, our model maintains a parameter size of approximately 15M and FLOPs around 140G. These results highlight the advantages of our method in both segmentation accuracy and efficiency. Moreover, prompt-based models generally outperform non-prompt-based models, demonstrating the effectiveness of the cancer-level prompts we introduced for PET-CT segmentation tasks across whole-body lesions.

Learnable Prompt Visualization We conducted a visualization analysis of the learnable cancer-specific and common prompts using T-SNE. As shown in Fig. 2, red, blue, purple, and green points represent different cancer-specific prompts, with the common prompts positioned centrally among them, indicating that they encode shared features across cancer types. These findings demonstrate that our method effectively separates cancer-specific features, capturing their correlations while preserving their distinctiveness.

Ablation Study Table 2 shows the contribution of each component to the model performance. Compared to the baseline (70.24%), adding the cancer-specific prompt increases DSC to 72.93%, while further introducing the common prompt boosts it to 73.79%. Prompt-aware heads (PA-Heads) alone achieve

Table 1. Comparison of different methods for cancer lesion segmentation with Weighted Averages. Bold values indicate the highest score, and underlined values indicate the second highest. All comparison results yielded $p < 0.05$. ("Single" denotes single-task training, "+" indicates nnU-Net with residual connections, and "*" represents STU-Net-B).

Category	Metrics	Non-Prompt-Based Methods								Prompt-Based Methods			
		Single	nnU-Net	nnU-Net ⁺	STU-Net	STU-Net*	UNETR	SwinUNET	U-Mamba	3DUX-Net	Dod-Net	CLIP-driven	UniSeg Proposed
Lung cancer	DSC	80.45	78.57	78.83	79.55	<u>80.68</u>	72.92	75.94	79.76	80.05	79.26	80.39	80.81
	IoU	68.31	65.86	66.07	67.20	<u>68.51</u>	59.37	62.99	67.67	68.05	66.91	68.49	68.98
Lymphoma	DSC	71.30	73.66	71.93	70.15	74.07	68.85	70.65	72.49	73.01	<u>76.80</u>	74.38	75.87
	IoU	59.57	61.51	59.75	58.15	61.90	56.92	58.83	60.18	61.08	<u>64.02</u>	62.95	66.38
Melanoma	DSC	72.19	72.44	75.46	70.53	75.03	61.23	68.71	70.34	76.96	75.57	75.82	72.70
	IoU	60.43	60.17	63.70	58.32	62.52	48.27	55.87	58.38	<u>64.03</u>	64.10	64.36	65.15
Breast cancer	DSC	<u>65.89</u>	60.23	61.51	62.65	63.26	57.02	61.79	61.60	64.21	64.41	64.45	66.27
	IoU	<u>51.76</u>	45.45	47.14	48.17	48.91	42.78	47.68	47.36	49.61	50.09	50.10	51.85
AVG	DSC	72.07	70.66	71.54	70.25	72.79	64.33	68.81	70.46	73.22	<u>73.55</u>	73.35	72.93
	IoU	59.58	57.63	58.74	57.42	59.92	51.11	55.82	57.76	60.50	<u>61.00</u>	60.99	62.56
Parameters	M	30.79	30.79	46.71	14.55	58.16	96.19	46.49	42.12	83.81	15.14	15.34	15.12
FLOPs	GFLOPs	526.29	526.29	833.94	138.70	548.61	196.92	202.58	1067.69	1509.02	139.26	139.26	138.79

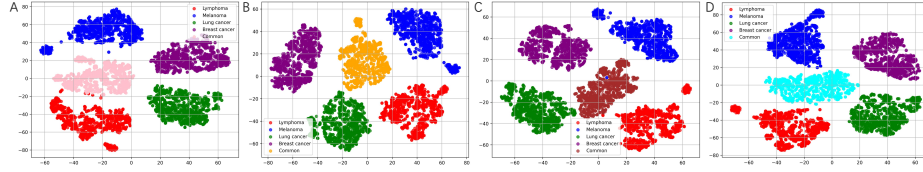


Fig. 2. T-SNE visualization of learnable cancer-specific and common prompts for (A) Lung Cancer, (B) Lymphoma, (C) Melanoma, and (D) Breast Cancer. The common prompt for each cancer type is derived from the fusion of cancer-specific prompts from other cancers, leading to distinct representations. To illustrate these differences, we present four separate visualizations.

73.92%, highlighting their individual effectiveness. Combining all components yields the highest DSC of 74.87%, emphasizing their complementary roles.

Survival analysis Biomarkers derived from radiological imaging have been effectively utilized to predict breast cancer prognosis [5, 26]. Several studies have investigated the prognostic value of metabolic activity (SUVmax), tumor volume (MTV), and total lesion glycolysis (TLG) [24, 15, 1, 18]. In this study, we leverage our proposed method to perform inference on a large dataset without ground truth and automatically extract predefined parameters for survival analysis.

The results are presented in Table 3. We selected the three best-performing segmentation models to calculate MTV, TLG, and SUVmax and evaluated their prognostic performance using the C-index. The findings indicate that better segmentation accuracy enhances survival prediction precision, demonstrating the generalizability of our model. Patients were then stratified into high-risk and low-risk groups based on the median values of SUVmax, MTV, and TLG. As shown in Fig. 3, patients in the high-risk group had worse outcomes than those in the low-risk group. A Cox proportional hazards model was applied based on this stratification, revealing that MTV provided the strongest prognostic

Table 2. Ablation study results on the average DSC (%).

Experimental Setting				Category				
Baseline	S-Prompt	C-Prompt	PA-Heads	Lung	Lymphoma	Melanoma	Breast	AVG
✓				79.55	70.15	70.53	62.65	70.24
	✓			80.45	75.87	72.70	64.85	72.93
✓		✓		81.05	76.63	74.59	64.98	73.79
✓			✓	79.66	76.48	75.54	65.71	73.92
✓	✓	✓	✓	80.81	77.76	76.44	66.27	74.87

Table 3. Comparison of three best-performing segmentation models for Survival Prediction (C-index %)

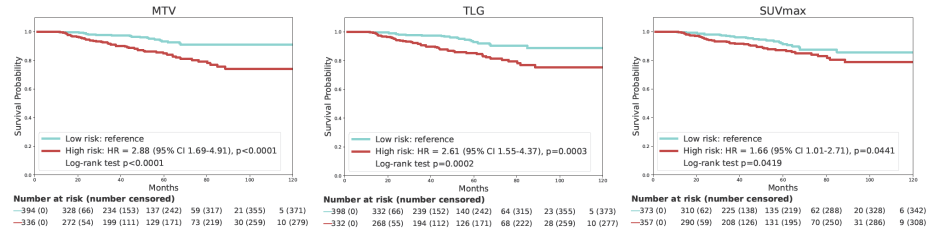
Method	MTV		TLG		SUVmax	
Dod-Net	68.51	[61.79, 75.30]	66.12	[59.12, 72.35]	62.58	[56.09, 68.44]
CLIP-driven	68.83	[61.62, 75.69]	66.07	[59.23, 72.43]	62.44	[56.10, 68.97]
Proposed	69.33	[62.05, 75.83]	66.63	[59.52, 72.84]	62.74	[55.65, 68.89]

stratification. Specifically, for overall survival, MTV had the highest hazard ratio (HR = 2.88), compared to TLG (HR = 2.61) and SUVmax (HR = 1.66).

Our study demonstrates the potential of the proposed universal segmentation model as a reliable and efficient clinical tool for breast cancer prognosis. By automatically extracting key parameters from PET-CT, it eliminates the need for time-consuming manual annotations. The model stratifies patients into high- and low-risk groups, enabling targeted monitoring and early interventions while minimizing unnecessary treatments. These findings underscore its value in enhancing cancer diagnosis, treatment planning, and patient management.

4 Conclusion

In this study, we propose a dual-prompt-driven model designed for whole-body PET-CT segmentation across multiple cancer types. To address the heterogene-

**Fig. 3.** Kaplan-Meier survival curves stratified by MTV, TLG, and SUVmax, showing the survival differences between high-risk and low-risk groups.

ity and interrelations among different cancers, we introduce a dual-prompt strategy that integrates cancer-specific and common prompts, enabling the model to distinguish unique metastatic patterns while leveraging shared features. Additionally, we incorporate prompt-aware heads to adaptively handle multiple segmentation tasks while maintaining efficiency. With these designs, DpDNet achieves state-of-the-art performance across multiple PET-CT datasets and outperforms various existing segmentation models. Furthermore, its application to breast cancer survival analysis demonstrates its potential for risk stratification and clinical decision support. These findings indicate that DpDNet could be a scalable and effective tool for advancing personalized cancer management.

Acknowledgements

Xinglong Liang was funded by Chinese Scholarship Council (CSC) scholarship. This work was supported by the Science and Technology Development Fund of Macao (0041/2023/RIB2).

Disclosure of Interests

The authors declare no competing interests.

References

1. Bouron, C., Mathie, C., Seegers, V., Morel, O., Jézéquel, P., Lasla, H., Guillerminet, C., Girault, S., Lacombe, M., Sher, A., et al.: Prognostic value of metabolic, volumetric and textural parameters of baseline [18f] fdg pet/ct in early triple-negative breast cancer. *Cancers* **14**(3), 637 (2022)
2. Budczies, J., von Winterfeld, M., Klauschen, F., Bockmayr, M., Lennerz, J.K., Denkert, C., Wolf, T., Warth, A., Dietel, M., Anagnostopoulos, I., et al.: The landscape of metastatic progression patterns across major human cancers. *Oncotarget* **6**(1), 570 (2014)
3. Chen, J., Chen, C., Huang, W., Zhang, J., Debattista, K., Han, J.: Dynamic contrastive learning guided by class confidence and confusion degree for medical image segmentation. *Pattern Recognition* **145**, 109881 (2024)
4. Endo, K., Oriuchi, N., Higuchi, T., Iida, Y., Hanaoka, H., Miyakubo, M., Ishikita, T., Koyama, K.: Pet and pet/ct using 18 f-fdg in the diagnosis and management of cancer patients. *International journal of clinical oncology* **11**, 286–296 (2006)
5. Gao, Y., Ventura-Diaz, S., Wang, X., He, M., Xu, Z., Weir, A., Zhou, H.Y., Zhang, T., van Duijnhoven, F.H., Han, L., et al.: An explainable longitudinal multi-modal fusion model for predicting neoadjuvant therapy response in women with breast cancer. *Nature Communications* **15**(1), 9613 (2024)
6. Gatidis, S., et al.: A whole-body fdg-pet/ct dataset with manually annotated tumor lesions sci. *Data* **9**(1), 601 (2022)
7. Groheux, D., Espié, M., Giacchetti, S., Hindié, E.: Performance of fdg pet/ct in the clinical management of breast cancer. *Radiology* **266**(2), 388–405 (2013)

8. Häggström, I., Leithner, D., Alvé, J., Campanella, G., Abusamra, M., Zhang, H., Chhabra, S., Beer, L., Haug, A., Salles, G., et al.: Deep learning for [18f] fluorodeoxyglucose-pet-ct classification in patients with lymphoma: a dual-centre retrospective analysis. *The Lancet Digital Health* **6**(2), e114–e125 (2024)
9. 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)
10. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: Unetr: Transformers for 3d medical image segmentation. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. pp. 574–584 (2022)
11. Huang, Z., Wang, H., Deng, Z., Ye, J., Su, Y., Sun, H., He, J., Gu, Y., Gu, L., Zhang, S., et al.: Stu-net: Scalable and transferable medical image segmentation models empowered by large-scale supervised pre-training. *arXiv preprint arXiv:2304.06716* (2023)
12. 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)
13. Kennecke, H., Yerushalmi, R., Woods, R., Cheang, M.C.U., Voduc, D., Speers, C.H., Nielsen, T.O., Gelmon, K.: Metastatic behavior of breast cancer subtypes. *Journal of clinical oncology* **28**(20), 3271–3277 (2010)
14. Lee, H.H., Bao, S., Huo, Y., Landman, B.A.: 3d ux-net: A large kernel volumetric convnet modernizing hierarchical transformer for medical image segmentation. *arXiv preprint arXiv:2209.15076* (2022)
15. Lee, M.I., Jung, Y.J., Kim, D.I., Lee, S., Jung, C.S., Kang, S.K., Pak, K., Kim, S.J., Kim, H.Y.: Prognostic value of suvmax in breast cancer and comparative analyses of molecular subtypes: A systematic review and meta-analysis. *Medicine* **100**(31), e26745 (August 2021)
16. 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)
17. Ma, J., Li, F., Wang, B.: U-mamba: Enhancing long-range dependency for biomedical image segmentation. *arXiv preprint arXiv:2401.04722* (2024)
18. Quartuccio, N., Alongi, P., Urso, L., Ortolan, N., Borgia, F., Bartolomei, M., Arnone, G., Evangelista, L.: 18f-fdg pet-derived volume-based parameters to predict disease-free survival in patients with grade iii breast cancer of different molecular subtypes candidates to neoadjuvant chemotherapy. *Cancers* **15**(10), 2715 (2023)
19. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 20730–20740 (2022)
20. Wood, S.L., Pernemalm, M., Crosbie, P.A., Whetton, A.D.: The role of the tumor-microenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets. *Cancer treatment reviews* **40**(4), 558–566 (2014)
21. Xie, Y., Zhang, J., Xia, Y., Shen, C.: Learning from partially labeled data for multi-organ and tumor segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023)

22. Ye, Y., Chen, Z., Zhang, J., Xie, Y., Xia, Y.: Meduniseg: 2d and 3d medical image segmentation via a prompt-driven universal model. arXiv preprint arXiv:2410.05905 (2024)
23. Ye, Y., Xie, Y., Zhang, J., Chen, Z., Xia, Y.: Uniseg: A prompt-driven universal segmentation model as well as a strong representation learner. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 508–518. Springer (2023)
24. Yue, Y., Cui, X., Bose, S., Audeh, W., Zhang, X., Fraass, B.: Stratifying triple-negative breast cancer prognosis using 18f-fdg-pet/ct imaging. *Breast Cancer Research and Treatment* **153**, 607–616 (2015)
25. 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)
26. Zhang, T., Tan, T., Samperna, R., Li, Z., Gao, Y., Wang, X., Han, L., Yu, Q., Beets-Tan, R.G., Mann, R.M.: Radiomics and artificial intelligence in breast imaging: a survey. *Artificial Intelligence Review* **56**(Suppl 1), 857–892 (2023)
27. Zhang, Y., Li, X., Chen, H., Yuille, A.L., Liu, Y., Zhou, Z.: Continual learning for abdominal multi-organ and tumor segmentation. In: International conference on medical image computing and computer-assisted intervention. pp. 35–45. Springer (2023)
28. Zhang, Z., Han, L., Zhang, T., Lin, Z., Gao, Q., Tong, T., Sun, Y., Tan, T.: Unimrisegnet: Universal 3d network for various organs and cancers segmentation on multi-sequence mri. *IEEE Journal of Biomedical and Health Informatics* (2024)
29. Zhou, K., Yang, J., Loy, C.C., Liu, Z.: Learning to prompt for vision-language models. *International Journal of Computer Vision* **130**(9), 2337–2348 (2022)