Cross-Modal Interactive Perception Network with Mamba for Lung Tumor Segmentation in PET-CT Images

Supplementary Material

8. Experiment Appendix

8.1. Qualitative Comparison on PCLT20K

As illustrated in Fig. 8, we provide additional qualitative comparison examples between our proposed CIPA and other segmentation methods. These examples include lung tumors of varying sizes, locations, shapes, and structures, highlighting the challenges of lesion segmentation based on PET-CT images. It can be observed that our method accurately locates the lesions and segments complete tumors with well-preserved boundary details. For example, in the 9th row of Fig. 8, the tumor is located in the left lung and is relatively large. CIPA successfully segments the tumor completely, whereas other methods show missing regions in the segmented lesion and produce some false-positive pixels. These cases validate the effectiveness of our CIPA in lung tumor segmentation using PET-CT images.

8.2. Robustness for Tumors of Varying Sizes

As described in Sec. 3.3 of the paper, the size distribution of tumors in the PCLT20K dataset is divided into four levels: [0,100], [101,500], [501,1000], and [1001,6000]. We evaluate the robustness of the models for tumors of different sizes, as shown in Table 7. CIPA achieves accuracies of 76.59%, 84.20%, 90.33%, and 89.86% on test sets with tumor sizes of [0,100], [101,500], [501,1000], and [1001,6000], respectively, which outperforms other methods. This demonstrates the robustness of CIPA in handling tumors of varying sizes.

9. Code Appendix

The complete benchmark, including the training and testing code for CIPA, training weights, and prediction results, will be released at https://github.com/mj129/CIPA. We hope this will help advance research in lung tumor segmentation based on PET-CT images.

10. Data Appendix

The complete PCLT20K dataset will be released at https://github.com/mj129/CIPA.

10.1. More Examples of PCLT20K Dataset

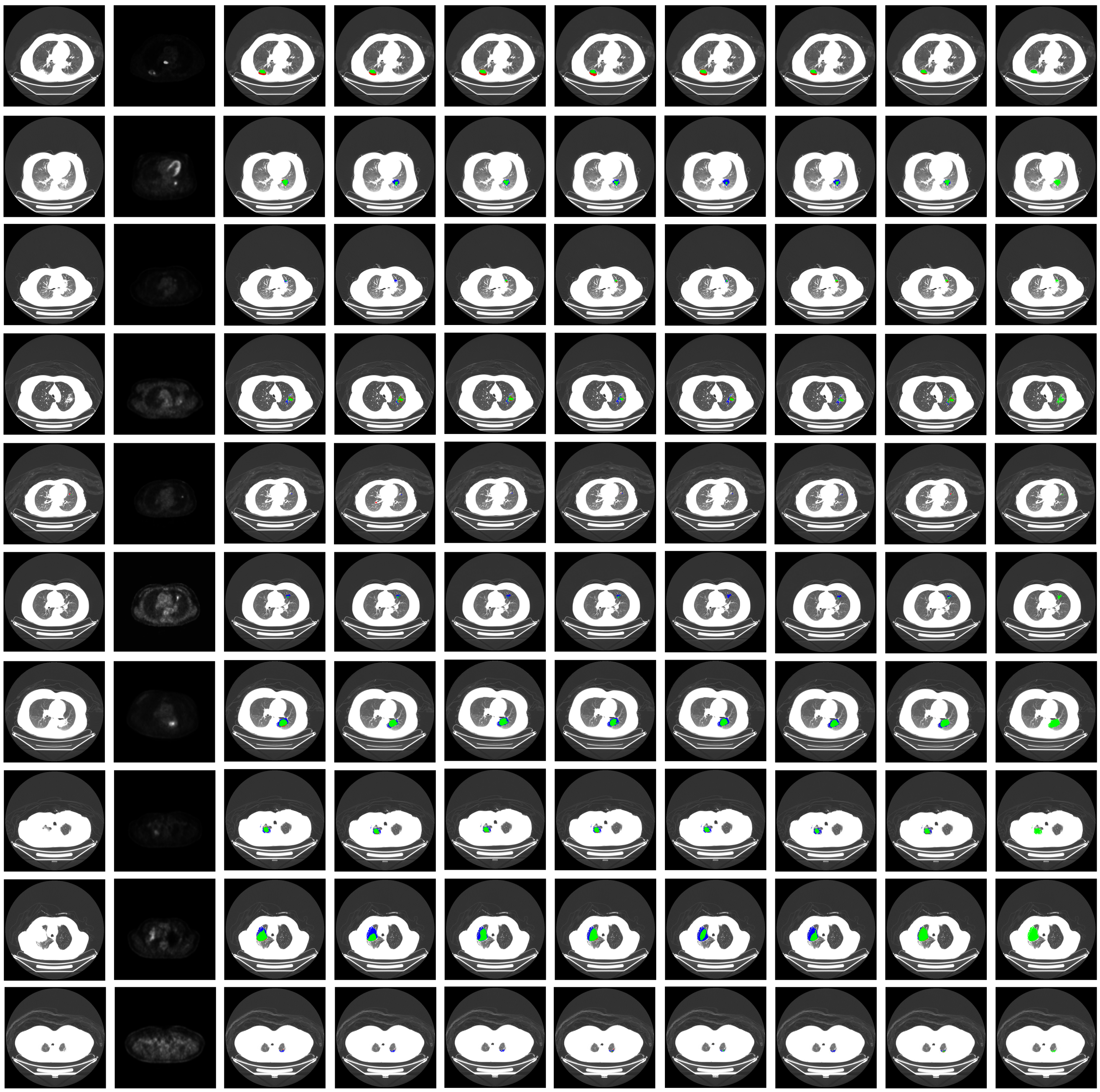
More examples of the PET-CT images and the corresponding tumor picked from the PCLT20K dataset are presented in Fig. 9. For four different tumor sizes, *i.e.,* [0, 100], [101, 500], [501, 1000], and [1001, 6000], we present six examples for each. Each example includes a CT image and a PET image, along with corresponding zoomed-in views of the tumor region, both with and without annotations.

Method	[0,100]	[101,500]	[501,1000]	[1001,6000]
TokenFusion [34]	67.03	79.99	88.80	84.53
CEN [35]	57.89	77.42	86.72	86.84
CMX [40]	66.19	79.23	87.02	84.43
CMNeXt [41]	63.78	80.64	89.27	86.23
AsymFormer [5]	58.90	72.01	78.25	77.30
DFormer [39]	61.72	76.73	84.86	85.43
Sigma [31]	69.68	81.02	87.08	86.07
GeminiFusion [13]	68.32	80.62	87.54	85.62
CIPA (ours)	**76.59**	**84.20**	**90.33**	**89.86**

Table 7. Comparison of our proposed CIPA with other methods for tumor segmentation of varying sizes based on the evaluation metric "accuracy".

10.2. Morphological Complexity of Tumors

For lung tumor segmentation in PET-CT images, it faces substantial challenges due to the complex nature of tumors, such as complex morphology, irregular shape, and unclear tumor boundaries. In the PCLT20K dataset, we classify tumor morphological complexity into three levels based on Perimeter-to-Area Ratio and Fractal Dimension: Low (5,390 pairs), Medium (10,534 pairs), and High (6,006 pairs).



(a) CT (b) PET (c) TokenFusion (d) CMX (e) CMNeXt (f) GeminiFusion (g) DFormer (h) Sigma (i) Ours (j) G.T.

Figure 8. Additional qualitative comparison of our CIPA with other segmentation methods on the PCLT20K dataset. Green: true positive, red: false positive, blue: false negative.

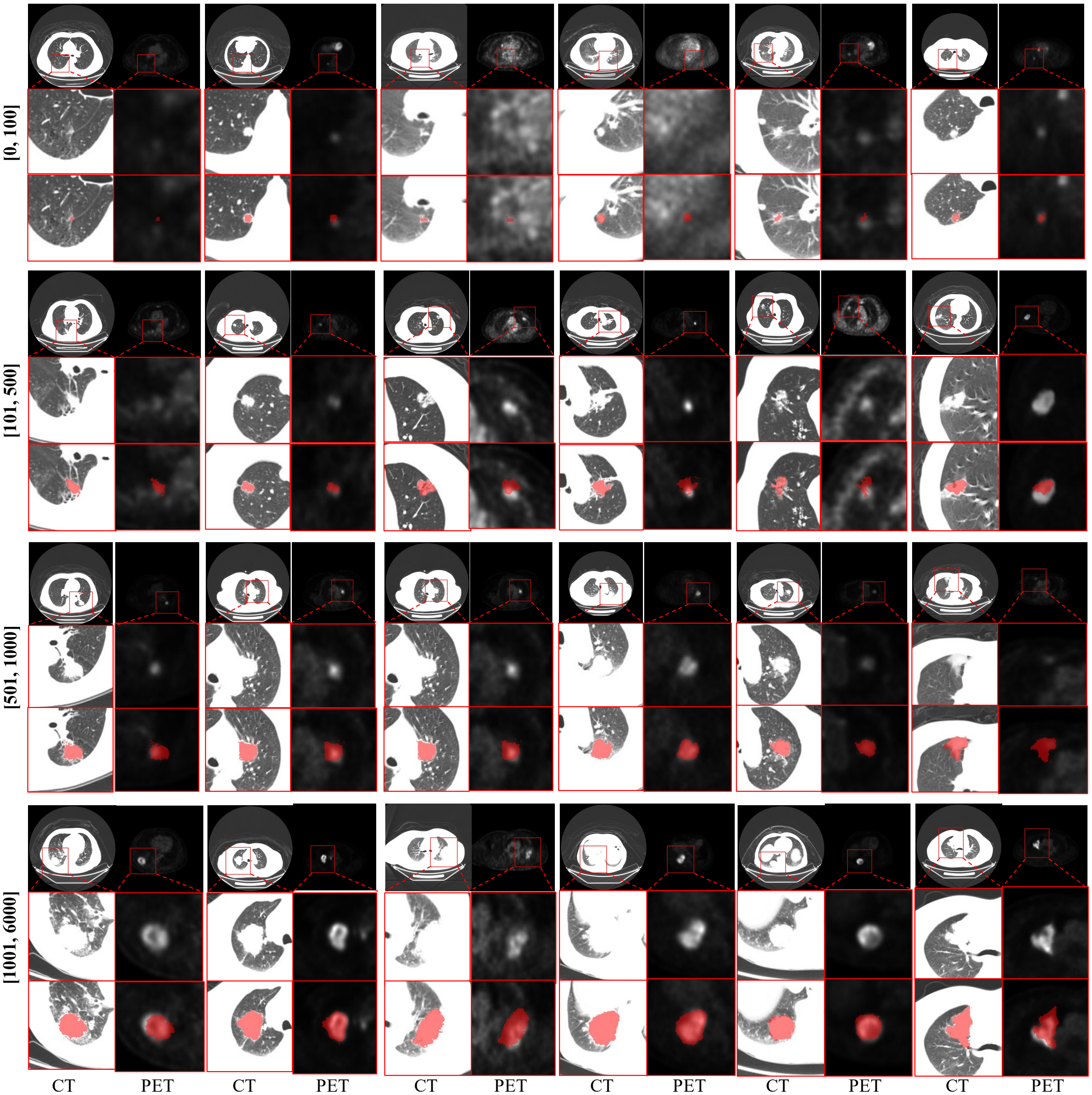


Figure 9. More examples of the PET-CT images and the corresponding tumor mask picked from the PCLT20K dataset. PET primarily highlights areas of increased metabolic activity, and CT provides detailed anatomical images, showing the precise location, size, and shape of tumors. The range of numbers on the left side represents the pixel count of the tumor region.