

# Supplementary for Mask-Enhanced Segment Anything Model for Tumor Lesion Semantic Segmentation

Hairong Shi<sup>1</sup> Songhao Han<sup>1</sup> Shaofei Huang<sup>2(✉)</sup> Yue Liao<sup>3</sup> Guanbin Li<sup>4(✉)</sup>  
Xiangxing Kong<sup>5</sup> Hua Zhu<sup>5</sup> Xiaomu Wang<sup>6</sup> Si Liu<sup>1</sup>

<sup>1</sup> Beihang University, Beijing, China

<sup>2</sup> Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China  
nowherespfly@gmail.com

<sup>3</sup> The Chinese University of Hong Kong, Hong Kong, China

<sup>4</sup> Sun Yat-sen University, Guangzhou, China  
liguanbin@mail.sysu.edu.cn

<sup>5</sup> Department of Nuclear Medicine, Peking University Cancer Hospital & Institute,  
Peking University, Beijing, China

<sup>6</sup> Nanjing University, Nanjing, China

## 1 Dataset Description

KiTS19, MSD Lung, MSD Pancreas and LiTS datasets contain 300, 96, 420, 201, 443 and 100 CT volumes with 3, 1, 2, 2, 2 and 1 classes respectively, while the BraTS21 dataset contains 1251 MRI volumes with 3 classes. Lung100\* is a private dataset we constructed, consisting of 100 lung CT images along with their corresponding tumor segmentation masks. All other datasets used in our study are publicly available.

## 2 Evaluation Metrics

The Dice score and Intersection over Union (IoU) are pivotal metrics for assessing the accuracy of segmentation models by quantifying the overlap between the predicted segmentation ( $X$ ) and the actual ground truth ( $Y$ ). Both metrics use  $|\cdot|$  to denote the operation for calculating the size or count of a set. The Dice score is defined as

$$Dice(X, Y) = \frac{2 \times |X \cap Y|}{|X| + |Y|}, \quad (1)$$

highlighting the relative overlap by considering the size of the intersection twice over the sum of the sizes of both sets. In contrast, the IoU measures the ratio of the intersection to the union of the predicted and ground truth segmentations, computed as

$$IoU(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}, \quad (2)$$

offering a direct comparison of segmentation accuracy.