# Supplementary materials

## 1 Image segmentation

All the CT images were exported in Joint Photographic Experts Group (JPEG) format. Two experienced radiologists in abdominal imaging diagnosis (3 years and 13 years, respectively) participated in the segmentation of the entire tumor. The first radiologist manually delineated the ROIs of the entire tumor layer by layer on venous phase CT images, whilst the other senior radiologist confirmed that the ROIs met the requirements for tumor segmentation, ensuring that the entire tumor was included while avoiding blood vessels, calcification, and hemorrhagic areas. Then, the final image segmentations were completed by the senior radiologist (Figure 1). Next, images were converted to NII format, and the spline interpolation scaling method was employed to standardize the 3D data input, setting the target depth to 64. Following this, a rectangular box that could contain the largest tumor was uniformly cropped in the depth direction and resized to a fixed image size of 224 × 224 × 64. Finally, voxel normalization was performed on CT images to limit pixel values to a specific range, such as [0,1].

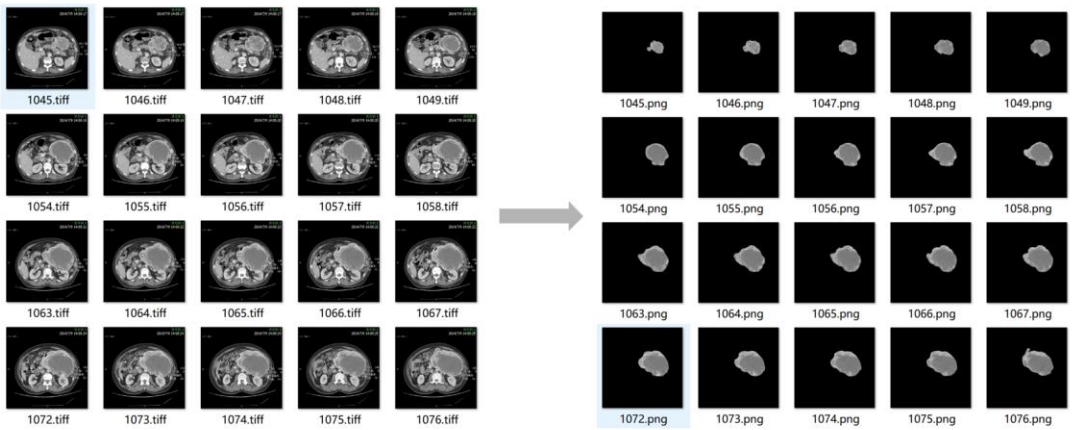


Figure 1 Representation of image segmentation. The first radiologist manually delineated the ROIs of the entire tumor layer by layer on venous phase CT images, while the senior radiologist confirmed that the ROIs met the requirements for tumor segmentation and completed the final image segmentation.

## 2 Training details

In our study, the network architecture in the Pytorch framework was implemented and trained on NVIDIA GPU to accelerate training speed. The model was trained using the Adam optimizer with an initial learning rate of 0.01 to minimize cross-entropy loss. The batch size for this experiment was set to 8, with a maximum number of epochs of 120, and an average pooling factor of 2. Additionally, a multi-layer perceptron (MLP) was employed to classify and predict feature vectors generated from the integration of imaging

data and clinical information. The MLP architecture consisted of three fully connected layers, with a batch normalization layer added after each linear layer to enhance the convergence speed of the neural network.