AutoPET III Challenge: ResUnet-based nnUnet for Whole body Pet and CT tumor segmentation

Yuxin Gong

Abstract

Automatic segmentation of lesions in FDG and PSMA Whole Body (WB) PET/CT scans are essential for accurate treatment response evaluation, dosimetry optimization, and advancement in theranostic applications. However, PET usually show the unclear Tumor structure comprising lots of noise. Thus, how to improve the PET segmentation remains a challenge. We use nnUnetV2 to get Multi-modal, Pet and CT, tumor segmentation. In the nnUnetV2, we use ResUnet as base Network, and the dataset is AutoPET III challenge dataset, which includes 1200 subjects, our method with a ResUNet-based nnUNetV2 model trained for 250 epochs. Our approach achieved 0.73 Dice in the AutoPET III challenge, demonstrating its effectiveness in Whole body Pet-CT lesion segmentation.

Introduction

PET/CT scans are particularly invaluable due to their ability to provide comprehensive insights into both metabolic activity and anatomical structure. PET scans reveal areas of high radiotracer uptake, which are indicative of increased metabolic activity commonly associated with tumors. Meanwhile, CT scans offer detailed anatomical information, aiding in precise localization of lesions. Despite significant advancements in deep learning algorithms for medical imaging, accurate segmentation of lesions from surrounding normal organs remains a formidable challenge. This difficulty is exacerbated by the high radiotracer uptake in certain normal organs, which can obscure the differentiation between tumor tissues and their surrounding structures. Using CT to provide the context information while improving Pet segmentiaton performance make sense. Our work verify ResUNet-based nnUNetV2 model can get good performance AutoPET III dataset. This approach promises to improve the accuracy of tumor segmentation, thus providing more precise data for treatment planning and monitoring.

Methods:

Data and Preprocessing

The dataset consisted of FDG-PET/CT scans from 1200 patients, provided by the AutoPET III challenge. A held-out test dataset of 58 studies was used to assess model robustness and generalizability. CT and PET scans were resampled to the same spacing and normalized. Annotations for lesions and high-uptake organs were provided by experienced radiologists. The nnUnetV2 split the dataset into training and validation sets automatically

Model Training Methodology

We utilized a ResUNet-based nnUNetV2 model to train the multi-modal segmentation mode. The input images split patchs and the dimensions of them are 128x128x128. We use each patch as input to train the model, and stitch the predicted patches together to obtain the complete segmentation result of the full image, Here, the Base network is ResUnet, which is a refined version of the traditional UNet architecture that integrates residual learning to enhance segmentation performance.

It preserves the UNet's core design, which includes an encoder that progressively extracts features through a series of convolutional layers and downsampling operations, and a decoder that reconstructs the image's spatial resolution using upsampling and convolutional layers. The standout feature of ResUNet is its use of residual blocks, which introduce shortcut connections that bypass one or more layers. This addition helps address the vanishing gradient problem, allowing the network to be deeper and train more effectively. These residual connections enable the network to learn more complex features and improve overall accuracy. Besides, we trained the model using a combination of Dice Similarity Coefficient (DSC) and weighted cross-entropy loss, with data augmentation techniques using Monai. The Data augmentation methods can be seen as follow

1. RandGaussianNoise(prob=0.1),
2. RandGaussianSmooth(sigma_x=(0.5, 1.0), prob=0.2),
3. RandScaleIntensity(factors=0.25, prob=0.15),
4. RandShiftIntensity(offsets=0.1, prob=0.15),
5. RandAdjustContrast(prob=0.15),
6. RandZoom(min_zoom=0.5, max_zoom=1.0, prob=0.25),
7. RandRotate(prob=0.2, range_x=(-30, 30), range_y=(-30, 30), range_z=(-30, 30))

The network is trained for 250 epochs with a batch size of 9 and an SGD optimizer with a learning rate of 0.01. Performance was evaluated using DSC .

Results

ResUnet-based nnUnetV2 performance

Training on a subset of 100 subjects with single-label segmentation achieved DSC of 0.74 for 68 independent test dataset. The post-processing are TTA of nnUnet and Crop.

Discussion and Conclusion

Our submission demonstrates the ResUnet-based nnUnet is effective in tumor segmentation of PET-CT. The result reinforces its potential clinical utility. and contributes valuable insights to the field of medical image analysis.