様式 8

論 文 内 容 要 旨

報告番号	甲 先 第 469 号	氏 名	周 宇翔
学位論文題目	Research on Medical Image Segmentation Based on Deep Learning Methods (深層学習に基づく医用画像セグメンテーションに関する研究)		

内容要旨

Medical imaging involves the technique and process of generating visual representations of a patient's body for clinical analysis and medical intervention. Healthcare professionals heavily depend on medical images for accurate diagnosis and treatment. In clinical practice, segmentation is typically performed manually. However, when processing a vast number of medical images, the quality of segmentation can vary based on the expertise of the medical professional. This variability underscores the need for a more consistent and efficient method to enhance the performance of segmentation tasks. Amassing such datasets is a complex task and is often beyond the capacity of a single institution within a limited timeframe. As a result, medical datasets tend to be smaller in size and can exhibit inconsistencies. Another notable characteristic of medical images is the imbalance between the foreground and background areas. Unlike natural scene images, where the foreground and background might be more balanced, medical images often have a much larger background compared to the foreground.

Extensive research has been conducted over the years to achieve fully automatic segmentation of the region of interest in medical images, aiming to improve efficiency and accuracy in comprehending such images. Based on deep learning and the server's powerful data processing capabilities, pixel-level processing and segmentation methods are usually used to process medical images, especially for brain tumor segmentation in 3D MRI scans, and for organ and lesion segmentation in 2D images. With the continuous development of deep learning, various neural network models have made remarkable achievements in semantic segmentation, stimulating research interest in medical image segmentation using deep learning.

The currently popular medical image segmentation method is still a network with an encoder-decoder structure. However, current method implementations do not fully explore feature fusion between different features. Furthermore, a notable challenge with Transformer-based models is their inherent complexity. They often necessitate large datasets for training to achieve optimal performance. Consequently, Transformers may not be the ideal choice for every medical segmentation task.

In this work, we propose an MDSU-Net, a variation of the U-Net, for 2D medical image segmentation. MDSU-Net incorporates both multi-attention and DSC layer for improved performance. The multi-attention module within our framework utilizes dual attention and

attention gates to capture rich contextual information and fuse features of different convolutional layers. MDSU-Net uses DSC layer to reduce model complexity without degrading model performance, which is suitable for different segmentation tasks. MDSU-Net registers a Dice score of 0.7055 on ATLAS, 0.9760 on CHAOS, and 0.8883 on NERVE. Notably, these scores surpass the current state-of-the-art (SoTA) benchmarks.

Additionally, in terms of 3D medical image segmentation, we propose Dual Encoder U-Net (DEU-Net), which uses Transformer and CNN respectively to extract 3D medical image features in the encoder. Transformer is a pre-trained model in BTCV, which improves its ability to capture contextual features of medical images and increases the learning speed. Besides, we introduce CBAM with each convolutional layer in the encoder part to enhance CNN's feature extraction capabilities for 3D medical images. To fuse the two kinds of features, we proposed a Dual Feature Fusion Module (DFFM) to fuse the features extracted from the Transformer and CNN in the encoder, making full use of the feature extraction capabilities of the two extractors for datasets of different sizes. We compared the proposed method with other advanced Methods on the BraTS 2020 and BraTS 2021 datasets, and the results show that the performance of DEU-Net has improved in 3D medical image segmentation task.

In the final section, we comprehensively conclude the advantage and weakness of this thesis based on the existing research. In the future, we will further explore a universal model that combines tasks from both dimensions. Additionally, regardless of whether it is 2D or 3D medical images, the parameter magnitudes of the two models we proposed are relatively large. For datasets with smaller volumes, we will further reduce model complexity to alleviate computational pressure and enhance the practicality of the model in real clinical environments.