Half-UNet: A Simplified U-Net Architecture for Medical Image Segmentation

Jai Raj Choudhary
jchoudhary@hawk.iit.edu

Laylo Karimova
lkarimova@hawk.iit.edu

Chao Yang
cyang72@hawk.iit.edu

Introduction

Medical image analysis, particularly medical image segmentation, plays a vital role in modern diagnostic and treatment methodologies. This process is crucial for accurate diagnosis, treatment planning, and monitoring disease progression. The integration of deep learning methods, especially Convolutional Neural Networks (CNNs), has significantly enhanced the accuracy and efficiency of these analyses. 0.0

Current State-of-the-art

Current state-of-the-art image segmentation models include fully convolutional networks (FCN) (Long et al., 2015), U-Net (Ronneberger et al., 2015), SegNet (Badrinarayanan et al., 2017), PSPNet (Zhao et al., 2017), and several DeepLab versions (Chen et al., 2017a,b, 2018). These models make use of unique information from various scales. U-Net is among the most commonly employed methods in the medical image segmentation and it is believed that the success of this method depends on its U-shaped structure. This architecture integrates high-level semantic information from the decoder with detailed features from the encoder via skip connections. Numerous models based on the U-Net structure have been proposed, each offering distinct enhancements and optimizations.

Residual Units U-net

Introduced by Kerfoot et al. (2018), the architecture is built from residual units and aims to enhance the segmentation of the left ventricle.

H-DenseUNet

Li et al. (2018) adapted the U-Net model by using dense convolutions for more effective liver and tumor segmentation tasks.

UNet++

Proposed by Zhou et al. (2018), this model introduces nested and densely connected skip pathways, reducing the semantic gap between encoder and decoder components. However, its nested network structure can become overly complex.

NAS-UNet

Weng et al. (2019) presented a novel approach, employing neural architecture search to automatically discover optimal architectures for both downsampling (DownSC) and upsampling (UpSC) paths. This resulted in a significant reduction of parameters while improving performance.

UNet3+

With an aim to leverage all-scale feature maps, Huang et al. (2020) developed UNet3+, which combines features at every level of fusion, enhancing the usage of full-scale feature information. This approach offers competitive results with fewer parameters than the traditional U-Net.

DC-UNet

Proposed by Lou et al. (2021), this version analyzed the classic U-Net and MultiResUNet architectures. It introduces a Dual-Channel CNN block, aiming to provide more effective feature extraction with a reduced parameter count.

Despite the variety and innovation of these U-Net variants, they all rely on to the fundamental U-shaped structure of the original U-Net. However, each model attempts to balance complexity, efficiency, and performance, often focusing on reducing the number of parameters and floating-point operations (FLOPs). Considering, the limitations Half-UNet have been introduced. In our paper we used three medical image segmentation datasets to compare U-Net, Half-UNet and other variants of U-Net architectures: 1) Mammography dataset, 2) CT images in Lung nodule dataset, 3) Left ventricular MRI image dataset.

Problem Description

Traditional U-Net, while effective, often requires a significant number of parameters and computational resources. This complexity becomes a challenge, especially in resource-limited settings or where fast processing is crucial. Variants of U-Net, like Residual U-Net, H-DenseUNet, UNet++, NAS-UNet, UNet3+, and

DC-UNet, have made improvements by modifying the architecture to enhance performance and efficiency. However, these models still predominantly follow the original U-shaped design and continue to grapple with issues of complexity and high computational demand.

Half-UNet proposes a solution by challenging a common belief: that U-Net's success relies heavily on its symmetric, U-shaped structure, particularly the feature fusion in the decoder. Instead, Half-UNet suggests that the performance mainly comes from the divide-and-conquer approach used in the encoder. Thus, it introduces an asymmetric design, focusing on simplifying the architecture and reducing complexity. This approach aims to maintain, or even enhance, the segmentation performance while significantly cutting down on the number of parameters and the computational load. By doing so, Half-UNet offers a more efficient and practical model for medical image segmentation, addressing the critical need for balance between performance, speed, and resource usage in various clinical and research applications.

Another aspect the Half-UNet addresses is the adaptability and flexibility of the segmentation model. Medical images come in various forms and resolutions, ranging from high-resolution MRI scans to varying qualities of X-rays and CT scans. A model proficient in segmenting one type of image may not perform equally well on another due to differences in image characteristics and the nature of the tissues or structures to be segmented. Hence, a versatile model capable of handling different types of medical images with minimal adjustments would be highly beneficial.

The Half-UNet model's introduction, focusing on an asymmetric architecture with a streamlined design, aims to mitigate these issues. It suggests a new direction away from the established symmetric structures, promoting a potentially more versatile and broadly applicable approach in the medical imaging field. This innovative perspective hopes to balance the critical aspects of model performance, adaptability to various medical imaging types, and the ever-present need for efficiency and practicality in clinical and research environments.

In conclusion, the "Half-UNet: A Simplified U-Net Architecture for Medical Image Segmentation" addresses critical challenges in the field of medical image segmentation. By innovatively balancing efficiency with accuracy, it presents a compelling solution that could reshape the landscape of medical imaging analysis, paving the way for broader and more efficient healthcare service delivery.

Description of Data

In the evaluation of the Half-UNet model for medical image segmentation, three distinct datasets, each from a different domain of medical imaging, will be employed. These datasets will be pivotal in establishing the model's accuracy, efficiency, and adaptability across various medical imaging scenarios.

Mammography Dataset from DDSM

Source: The Digital Database for Screening Mammography (DDSM) from the University of South Florida, USA. Content: This dataset included 483 regions of interest (ROIs) containing mammographic masses. Usage: Of these, 400 images will be allocated for the training set and 83 for the testing set. This dataset will primarily assist in gauging the model's performance in identifying and segmenting masses in mammographic images, a key task in breast cancer screening and diagnosis.

Lung Nodule Dataset from LIDC-IDRI

Source: The LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative) public database. Content: The dataset encompasses 1,018 cases, amounting to 4,104 images of CT scans with lung nodules. Selection Criteria: Only nodules with a diameter ¿= 3 mm will be chosen, based on the detailed contour coordinates available. Standardization: The ground truth for lung nodule segmentation will be established according to the 50 percent agreement principle, where a lung nodule is confirmed if at least two out of four radiologists agree on its presence. Split: The dataset will be divided into training and test sets in a 7:3 ratio. This dataset is crucial in evaluating the model's capability in detecting and segmenting lung nodules from CT images, a challenging task due to the variability in nodule size, shape, and appearance.

Left Ventricular MRI Dataset from MICCAI 2009

Source: Provided by the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) 2009 challenge. Content: This dataset contains short-axis cardiac MRI scans from 45 cases. Categorization: The cases will be divided into three groups, each containing 15 cases, including those with ischemic heart failure, non-ischemic heart failure, myocardial hypertrophy, and normal cases. Split: 30 cases (542 images) will be used as the training set, and the remaining 15 cases (265 images) will constitute the test set. Focus: The segmentation focus will be on the endocardium, with some cases also involving the epicardium. This dataset will provide a comprehensive platform to assess the model's performance in segmenting complex cardiac structures, crucial for diagnoses involving heart diseases.

There are relatively few images available in mammography and left ventricular MRI data sets. Recognizing the limited number of images in these datasets, we will implement a comprehensive data augmentation strategy to enhance our training data's volume and diversity. This approach is critical in improving the model's learning efficiency and its ability to generalize from limited data. The augmentation techniques to be used are as follows: Rotation: Each image in the mammography and left ventricular MRI training sets will be rotated clockwise in increments of 45 degrees, resulting in seven additional images per original image. This method aims to make the model robust against orientation variations in medical images. Flipping: Beyond rotation, each image will be also subjected to two flipping operations: Horizontal Flip: An image will be flipped along its vertical axis, creating a mirror image. This step helps the model learn to recognize structures irrespective of their horizontal orientation. Vertical Flip: Similarly, flipping the image along its horizontal axis ensures that the model does not become biased towards the vertical orientation of anatomical structures. By augmenting the images in these manners, we effectively increased the number of images in the training sets by tenfold. This significant enhancement in the dataset size aims to offset the challenges posed by the original limited number of images, providing a more robust and varied set of training examples.

Each of these datasets presents its unique challenges and complexities, making them ideal for testing the Half-UNet model's versatility in handling diverse medical image segmentation tasks. The variety of images, from mammograms and CT scans of lung nodules to cardiac MRIs, enabled a thorough examination of the model's performance across different imaging modalities and medical conditions.

Finished Work

In the initial phase of our project, we encountered several challenges that shaped our dataset selection process. While we had multiple sources for potential datasets, the performance limitations of individual team members' workstations led us to independently choose project datasets. One significant challenge we faced was the unexpected dataset size. The LIDC-IDRI dataset, in particular, exceeded a substantial 128GB. Even after splitting it into training and testing sets, the combined storage requirements approached the storage capacity limits of our team members' hard drives. As a result, we had to make the difficult decision to exclude certain datasets. This decision, while necessary, could potentially pose challenges during model fitting and future testing phases.

In the data preparation phase, our primary objectives were to create, augment, and organize our datasets. We implemented a versatile function to generate both training and testing data. Striving for a balanced dataset, we allocated a 50:50 size ratio between the training and testing sets. Given that our medical data was exclusively in Digital Imaging and Communications in Medicine (DICOM) format, we undertook the additional task of converting it into PNG format to facilitate future utilization. Furthermore, we developed a data augmentation function to enhance our training dataset. Data augmentation serves several essential roles in deep learning:

Increased Dataset Size: Data augmentation enables us to expand our training dataset by introducing various transformations, effectively amplifying its size without the need for additional data collection.

Regularization: It acts as a regularization technique, mitigating the risk of overfitting by introducing diversity and complexity into the training data.

Enhanced Invariance: By applying transformations such as rotations, translations, and flips, data augmentation helps the model become more invariant to these variations, fostering robustness.

In our augmentation process, we specifically employed image rotation and flipping to diversify our training dataset. Additionally, we created a function to generate validation data, extracting 500 samples from our training dataset.

Remained Work

Our project's next phase involves the implementation of our deep learning model. As we move forward, we anticipate challenges related to device performance and potentially lengthy training times. Given the vast size of our dataset, our desktop workstations are expected to require substantial time for both model training and testing during each run.

In terms of coding and logic, our model implementation will encompass the following key functions:

Data Generator: We will create a data generator responsible for efficiently loading our dataset.

Image Quality Checks: To ensure the integrity and suitability of our data, we will implement image quality checks, verifying pixel values and dimensions.

Half-UNet Architecture with Batch Normalization: Our model will incorporate a half-UNet architecture with the addition of batch normalization layers to enhance training stability and convergence.

UNet Architecture: We will implement the UNet architecture, a well-established model for image segmentation tasks.
Nested Half-UNet Model: Additionally, we will explore a nested half-UNet model to assess its potential for improved performance.

Model Visualization: Finally, we will develop techniques for visualizing the results produced by our model.

Undoubtedly, these tasks will be demanding, but we remain confident that our dedication and commitment to this project will lead to successful outcomes. Our efforts are directed toward addressing critical challenges in medical image segmentation, and we are steadfast in our pursuit of achieving our research objectives.