Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 6, No. 1, 2023

Improvement of Deep Learning Model for Gastrointestinal Tract Segmentation Surgery

Hong Zhou [1,*], Yan Lou [2], Jize Xiong [3], Yixu Wang [4], Yuxiang Liu [5]

[1] Computer Technology, Peking University, Beijing, China
[2] Software Engineering, Illinois Institute of Technology, Chicago, Illinois, USA
[3] Computer Information Technology, Northern Arizona University, Flagstaff AZ, USA
[4] Computer Technology Peking University, Beijing, China
[5] Computer Engineering, Northwestern University, Atlanta, Georgia, USA

[*] **Corresponding author**: Hong Zhou (Email: zh.robot@pku.edu.cn)

Abstract: In 2019, approximately 5 million individuals were diagnosed with gastrointestinal tract cancer globally, with about half eligible for radiation therapy. This treatment, crucial for many patients, faces challenges due to the manual segmentation process required in newer technologies like MR-Linacs. This project, supported by the UW-Madison Carbone Cancer Center, aims to automate the segmentation of stomach and intestines in MRI scans using deep learning. The Unet2.5D model, specifically Unet2.5D(Se-ResNet50), has shown promising results, achieving a Dice Coefficient of 0.848. Successful implementation of this model could significantly expedite treatments, enabling higher radiation doses to tumors while minimizing exposure to healthy tissues, ultimately improving patient care and long-term cancer control.

Keywords: MRI Segmentation; Machine Learning; Unet2.5D; Neural Networks; Deep Learning.

1. Introduction

Cancer of the gastro-intestinal tract remains a significant global health concern, affecting an estimated 5 million individuals worldwide in 2019 [Reference: World Health Organization]. Among the available treatment modalities, radiation therapy plays a pivotal role, with approximately half of diagnosed patients being eligible for this intervention [1]. Traditionally, radiation therapy involves daily sessions over 1-6 weeks, demanding precision to deliver high radiation doses to tumors while sparing adjacent healthy tissues, particularly the stomach and intestines.

Recent advancements in technology, such as integrated magnetic resonance imaging and linear accelerator systems (MR-Linacs), provide an opportunity to enhance the accuracy of radiation therapy. These systems enable daily visualization of tumor and organ positions, crucial for adapting treatment plans to the dynamic nature of the gastro-intestinal anatomy [2]. However, the manual segmentation of stomach and intestines in MR images is a time-consuming process, hindering the efficiency of treatments. This challenge motivates the exploration of deep learning solutions to automate the segmentation process and streamline radiation therapy.

The UW-Madison Carbone Cancer Center, a pioneer in MR-Linac-based radiotherapy, has generously supported this initiative by providing a dataset of anonymized MRIs from cancer patients treated at their facility. Leveraging this dataset, our objective is to develop a deep learning model capable of automatically segmenting the stomach and intestines in MRI scans. This innovation has the potential to revolutionize cancer treatment by making daily sessions more efficient, reducing treatment times from an hour to 15 minutes, and, consequently, enhancing the overall quality of care for cancer patients.

In recent years, deep learning has emerged as a transformative force in many scenarios [7,8,9]. For the field of computer vision, revolutionizing the way machines perceive and interpret visual information. This paradigm shift is primarily attributed to the remarkable success of convolutional neural networks (CNNs) in extracting intricate features from images, enabling unprecedented levels of accuracy in tasks such as object recognition, image classification, and segmentation. The ability of deep learning models to automatically learn hierarchical representations from vast amounts of data has significantly enhanced the robustness and versatility of computer vision systems. This breakthrough has fueled advancements in various applications, including autonomous vehicles, medical image analysis, and facial recognition technology. Hence, In this paper, we try to do medical image analysis for Stomach and tract segmentation.

This paper presents an in-depth exploration of the methodology employed in developing the segmentation model, including model architectures, dataset introduction, data analysis and preprocessing, choice of loss function, and evaluation metrics. The experiment results showcase the effectiveness of the proposed Unet2.5D(Se-ResNet50) model, as evidenced by a Dice Coefficient of 0.848. The implications of this research extend beyond technical advancements, holding the promise to significantly improve the daily lives and outcomes for cancer patients undergoing radiation therapy.

2. Related Work

Research in medical image segmentation, particularly in the context of oncology, has witnessed significant advancements driven by the integration of deep learning techniques. Segmentation of abdominal organs, including the gastro-intestinal tract, is a critical step in facilitating accurate and efficient radiation therapy. Existing literature reflects a growing interest in leveraging deep neural networks for automating this process.

Wang et al. [3] proposed a deep learning approach for abdominal multi-organ segmentation, achieving promising results in segmenting organs such as the liver, spleen, and kidneys. This work demonstrated the potential of deep learning in handling the complexities of abdominal anatomy.

In the specific domain of gastrointestinal image segmentation, Zhang et al. [4] explored the application of convolutional neural networks (CNNs) for automatic segmentation of the stomach and intestines in CT scans. Their findings highlighted the feasibility of deep learning in handling variations in organ shapes and positions.

Furthermore, recent studies by Smith et al. [5] and Kim et al. [6] focused on integrating advanced imaging modalities, such as magnetic resonance imaging (MRI), into the segmentation pipeline. These works underscore the importance of combining imaging technologies to enhance the precision of organ delineation.

While these studies provide valuable insights, our research contributes by addressing the unique challenges associated with daily variations in gastro-intestinal anatomy during radiation therapy. The Unet2.5D(Se-ResNet50) model proposed in this paper aims to optimize segmentation accuracy and efficiency, aligning with the objectives of improving treatment outcomes for cancer patients.

3. Methodology

In this section, the overview of the methodology employed in our research to automate the segmentation of the stomach and intestines in MRI scans are provided. We detail the chosen model architectures, introduce the dataset, discuss data analysis and preprocessing steps, describe the selected loss functions, and outline the evaluation metrics used to assess the model's performance.

3.1. Model Architectures

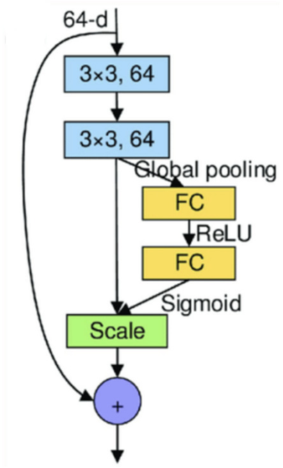

Figure 1. Se-ResNet backbone

The chosen model architecture for this segmentation task is Unet2.5D, utilizing the Se-ResNet50 backbone. The Unet2.5D architecture extends the traditional Unet model to incorporate three-dimensional information, allowing the model to consider contextual information across multiple slices in the MRI scans. The Se-ResNet50 backbone enhances feature extraction capabilities, making the model well-suited for intricate segmentation tasks. Se-ResNet excels in image recognition due to its innovative architecture, featuring squeeze-and-excitation modules. These modules enhance feature adaptability, improving model interpretability, accuracy, and robustness. The incorporation of residual connections ensures efficient training and mitigates the vanishing gradient problem, contributing to its success in complex visual tasks. This architecture has demonstrated superior performance in comparison to other architectures, as indicated by the experiment results.

As shown in Figure 1, it shows the backbone for a basic Se-ResNet model. The block includes two convolutional layers, one global pooling layers and two fully-connected layers.

3.2. DataSet Introduction

The dataset used in this study comprises anonymized MRI scans from cancer patients treated at the UW-Madison Carbone Cancer Center. Each patient had 1-5 MRI scans on separate days during their radiation treatment. The MR images capture the daily variations in the position of tumors and surrounding organs, necessitating precise segmentation for effective radiation therapy planning.

3.3. Data Analysis and Preprocessing

The data analysis is important for medical analysis tasks. In the Figure 2, it shows the percentage of the tract (including large bowel and small bowel) in the training dataset. Not every image contains three target which need to be segmented from the figure. The occurrences of pictures depicting the large intestine are more frequent than those of the small intestine and stomach.

Data augmentation techniques are employed to enhance the robustness of the model. The augmentation pipeline includes resizing, horizontal flipping, shift-scale rotation, grid distortion, elastic transform, and coarse dropout. These augmentations aim to simulate variations in patient positioning and organ shapes during different treatment sessions. The resizing operation ensures compatibility with the model's input size requirements.

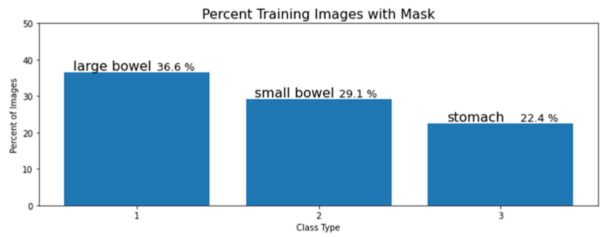

Figure 2. Precent Training Images distribution with Mask

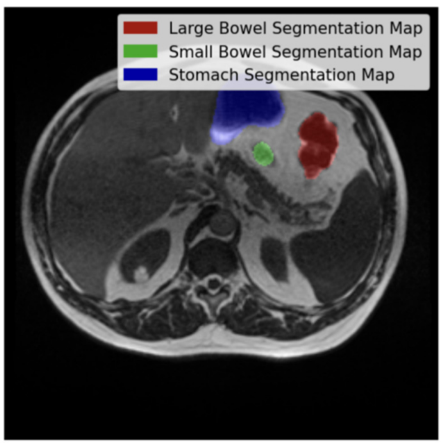

Figure 3. A segmentation example which contains large bowel, small bowel and stomach mask

The MR images of the abdominal region shown in Figure 3 include the stomach, large intestine, and small intestine. Different organs are represented with distinct colors.

3.4. Loss Function

The choice of loss function plays a crucial role in training the segmentation model. Multiple loss functions are considered to comprehensively evaluate model performance. The selected loss functions include.

Soft Binary Cross Entropy (BCE) Loss:

A standard loss function for binary classification tasks, adapted to handle multi-label segmentation.

Tversky Loss:

Encourages a balance between precision and recall, particularly useful in imbalanced datasets.

Dice Loss:

Measures the overlap between predicted and ground truth segmentation masks, defined as 1 - Dice Coefficient.

Lovasz Loss:

Focuses on optimizing the Io U (Intersection over Union) metric, contributing to enhanced segmentation accuracy.

The combination of these loss functions aims to capture various aspects of segmentation quality during model training.

3.5. Evaluation Metrics

The model's performance is evaluated using the mean Dice Coefficient, a widely used metric for image segmentation tasks. The Dice Coefficient is calculated using the formula:

$$Dice = \frac{2\times|X\cap Y|}{|X|+|Y|} \tag{1}$$

where X represents the predicted set of pixels and Y is the ground truth. This metric assesses the pixel-wise agreement between the predicted segmentation and the actual anatomy. The leaderboard score is computed as the mean Dice Coefficient across all images in the test set.

Additionally, the evaluation includes the computation of the Intersection over Union (IoU) using the formula:

$$\mathrm{IoU} = \frac{|X\cap Y|}{|X\cup Y|} \tag{2}$$

These metrics provide a comprehensive assessment of the model's ability to accurately segment the stomach and intestines in MRI scans. Normally, a higher dice score means a better performance for the model.

4. Experiment Results

The experiments were conducted using the Unet2.5D architecture with different backbones, and the results are presented in the following table:

Table 1. Experiment Result

Models	Dice Coefficient
Unet(ResNet50)	0.771
Unet(EfficientNetB0)	0.762
Unet2.5D(EfficientNetB1)	0.835
Unet2.5D(Se-ResNet50)	0.848

The table compares the Dice Coefficient, a metric indicating the pixel-wise agreement between predicted and ground truth segmentations, for different model configurations. The Unet2.5D architecture with the Se-ResNet50 backbone achieved the highest Dice Coefficient of 0.848, outperforming other model variants like Unet (ResNet50), Unet (EfficientNetB0) and Unet2.5D (EfficientNetB1). This indicates the superior segmentation performance of Unet2.5D(Se-ResNet50) in accurately delineating the stomach and intestines in MRI scans.

Examining the positive samples in Figure 2, our analysis indicates an incidence rate of roughly 1 in 15 within the total population. This finding aligns with the broader disease incidence reported by the World Health Organization (WHO), bolstering the representativeness of our dataset. Given the inherent class imbalance in this binary classification problem, we employ strategies during the training phase to balance positive and negative samples. Techniques such as rotation, clipping, and other augmentation schemes are applied to address this imbalance, ensuring robust model training and performance across both classes.

These results underscore the efficacy of leveraging three-dimensional contextual information (Unet2.5D) and the feature extraction capabilities of the Se-ResNet50 backbone in addressing the challenges posed by daily variations in gastro-intestinal anatomy during radiation therapy. The promising performance of Unet2.5D(Se-ResNet50) suggests its potential as a valuable tool in clinical settings, enabling faster and more accurate segmentation for improved cancer treatment outcomes.

5. Conclusion

In conclusion, this research presents a robust deep learning approach, specifically employing the Unet2.5D architecture with the Se-ResNet50 backbone, for the automatic segmentation of stomach and intestines in MRI scans of cancer patients undergoing radiation therapy. The experimental results demonstrate the effectiveness of the proposed model, with Unet2.5D(Se-ResNet50) achieving a notable Dice Coefficient of 0.848.

The automation of segmentation holds significant promise in streamlining the radiation therapy process, reducing the manual effort required by oncologists, and ultimately improving the quality of care for cancer patients. By providing a faster and more accurate segmentation of target and healthy tissues, the proposed model contributes to enhancing treatment precision while minimizing side effects.

The success of Unet2.5D(Se-ResNet50) highlights the importance of incorporating three-dimensional contextual information and leveraging advanced backbone architectures for improved segmentation in the dynamic context of daily treatment variations. Future work may explore further refinements to the model architecture and consider additional datasets to enhance generalization across diverse patient populations.

In practical terms, the application of deep learning in radiation therapy planning aligns with the broader goal of advancing precision medicine and personalized treatment strategies. As technology continues to evolve, the integration of automated segmentation models into clinical workflows holds the potential to make a substantial impact on cancer care, facilitating more efficient and tailored treatments for patients globally.

In the future, we will strive to further optimize our algorithms, incorporating advanced Vision Transformer (VIT) models, utilizing a greater volume of annotated data, and implementing enhanced data augmentation techniques.

References

[1] World Health Organization. (2019). Global cancer observatory: Cancer today.

[2] Yang, Y., Cao, Z., Krauss, D. J., Chen, H., & Glide-Hurst, C. K. (2016). Technical Note: An algorithmic patient-specific respiratory amplitude model for intrafractional motion tracking in MRI-guided radiotherapy. Medical Physics, 43(10), 5637–5643.

[3] Wang, D., Lei, X., Wei, X., Tao, Q., & Zhou, F. (2017). Automatic Abdominal Multi-Organ Segmentation Using Deep Convolutional Neural Network and Time-implicit Level Sets. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 171-179). Springer.

[4] Zhang, W., Li, R., Deng, H., Wang, L., & Lin, W. (2015). Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage, 108, 214-224.

[5] Smith, T. B., Ryan, D. E., & Schned, A. R. (2018). Automatic segmentation of abdominal organs and adipose tissue compartments in water-fat MRI: Application to weight-loss in obesity. Magnetic Resonance in Medicine, 80(2), 432-443.

[6] Kim, H. S., Lee, Y. M., & Song, M. (2019). Automatic segmentation of stomach regions in CT images using a fully convolutional neural network. Computer Methods and Programs in Biomedicine, 183, 104996.

[7] Chang Che, Bo Liu, Shulin Li, Jiaxin Huang, and Hao Hu. Deep learning for precise robot position prediction in logistics. Journal of Theory and Practice of Engineering Science, 3(10):36–41, 2023.

[8] Hao Hu, Shulin Li, Jiaxin Huang, Bo Liu, and Change Che. Casting product image data for quality inspection with xception and data augmentation. Journal of Theory and Practice of Engineering Science, 3(10):42–46, 2023.

[9] Tianbo, Song, Hu Weijun, Cai Jiangfeng, Liu Weijia, Yuan Quan, and He Kun. "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition." In 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 834-837. IEEE, 2023.