Liver Tumor Segmentation with Deep Learning Architectures

Rahul N. Pandya
rpandya2@depaul.edu

November 20, 2024

Abstract

Liver tumor segmentation is a critical task in medical imaging, enabling precise diagnosis and treatment planning. This project explores the application of U-Net and ResNet50 architectures to automatically segment liver tumors from contrast-enhanced CT scans. U-Net's encoder-decoder structure and skip connections are leveraged for precise segmentation, while ResNet50, as a pre-trained backbone, improves feature extraction. Both architectures are evaluated using Dice coefficient, Jaccard index, and accuracy, achieving strong results on a curated dataset. This work highlights the potential of advanced neural networks to improve segmentation accuracy and robustness in medical imaging.

1 Introduction

Accurate segmentation of liver tumors from CT scans is crucial for effective treatment planning and monitoring in liver cancer management. This task is challenging due to the irregular shapes and diffuse edges of tumors. Manual segmentation, although widely used, is time-intensive and prone to variability.

This study employs deep learning to automate segmentation, focusing on U-Net and ResNet50 architectures. U-Net is known for its success in biomedical segmentation tasks, while ResNet50 serves as a feature extractor with strong generalization capabilities. This project compares their performance on a dataset of liver CT scans, aiming to determine their effectiveness and limitations for medical imaging.

2 Related Work

Deep learning has significantly advanced medical image segmentation, particularly in liver tumor segmentation tasks. Two prominent architectures, U-Net and ResNet, form the foundation of many segmentation models.

U-Net, introduced by **(author?)** [4], has become a benchmark in biomedical image segmentation. Its encoder-decoder structure with skip connections effectively preserves spatial context while capturing hierarchical features. Variants like 3D U-Net [2] extend its capabilities to volumetric data, demonstrating enhanced performance for 3D medical imaging tasks. The adaptability of U-Net has made it a versatile choice for tasks requiring high precision in delineating boundaries.

ResNet, proposed by **(author?)** [3], addresses the vanishing gradient problem in deep neural networks through residual connections. While originally designed for image classification, ResNet has been adapted for segmentation tasks by integrating it with decoder architectures. These adaptations leverage its robust feature extraction capabilities, making it an effective backbone for complex segmentation problems.

The Liver Tumor Segmentation (LiTS) challenge [1] provided a significant benchmark dataset and spurred research in automated liver and tumor segmentation. Methods combining U-Net and ResNet with innovative loss functions and data augmentation techniques have emerged as state-of-the-art. For instance, Tversky loss, proposed by **(author?)** [5], has been employed to address class imbalance in medical image segmentation, achieving improved performance in segmenting small or irregular tumor regions.

This project builds on these foundations by leveraging U-Net and ResNet architectures, enhanced with Tversky loss and robust augmentation techniques, to improve segmentation accuracy and robustness in liver tumor segmentation.

3 Methodology

This project implements two deep learning architectures, U-Net and ResNet50, to segment liver tumors from CT scans. Each architecture was chosen for its unique strengths in image segmentation and feature extraction, allowing us to compare their performance on this critical medical imaging task.

3.1 Architectures

3.1.1 U-Net

The U-Net architecture is specifically designed for biomedical image segmentation . It adopts a symmetric encoder-decoder structure, ensuring that both high-level semantic features and fine-grained spatial details are captured. The encoder path consists of convolutional layers and max-pooling operations to reduce spatial dimensions while increasing feature abstraction. The decoder path employs transposed convolutions (upsampling layers) to restore the original spatial resolution.

A distinctive feature of U-Net is its use of skip connections, which directly transfer feature maps from the encoder to the decoder. This preserves spatial information and mitigates the loss of fine details during downsampling. The final output layer uses a sigmoid activation function to generate a binary segmentation mask.

In our experiments, U-Net exhibited strong performance for liver tumor segmentation, achieving a Dice coefficient of 0.93. The skip connections proved especially effective for delineating tumor boundaries, although the model occasionally struggled with detecting very small or diffuse tumors.

3.1.2 ResNet50

ResNet50 is a deep residual network initially developed for image classification . Its success lies in its residual blocks, which enable deeper architectures by mitigating the vanishing gradient problem through shortcut connections. For segmentation tasks, ResNet50 serves as a feature extractor in an encoder-decoder framework. The encoder (ResNet50 backbone) extracts hierarchical features from input images, while the decoder reconstructs the segmentation mask using upsampling layers.

In this project, we fine-tuned a ResNet50 pre-trained on ImageNet and customized the decoder path for segmentation. The residual connections in ResNet50 facilitated learning complex features, particularly for small and irregular tumor shapes. Compared to U-Net, ResNet50 demonstrated slightly better precision for challenging tumor regions, achieving a Dice coefficient of 0.94.

3.2 Loss Functions

Tversky Loss: To handle the inherent class imbalance in medical imaging datasets (where tumor regions occupy a smaller fraction of the image), we employed the Tversky loss function:

$$L_{Tversky} = 1 - \frac{TP}{TP + \alpha \cdot FP + \beta \cdot FN}$$

where TP, FP, and FN represent true positives, false positives, and false negatives, respectively. Hyperparameters α and β control the penalization of false positives and false negatives, and were set to 0.7 and 0.3, respectively. This loss function improved segmentation accuracy, especially for smaller tumor regions.

3.3 Dataset and Preprocessing

The dataset comprises 130 CT scans with annotated liver and tumor regions from the Liver Tumor Segmentation (LiTS) challenge [1]. Each CT scan was preprocessed to standardize intensity values and resize images to 128×128 resolution. Data augmentation, including random rotations, flips, and intensity scaling, was applied to improve model generalization and mitigate overfitting.

3.4 Training and Hyperparameter Tuning

Both models were trained using the Adam optimizer with a learning rate of 10^{-4}. A ReduceLROnPlateau scheduler dynamically reduced the learning rate when the validation loss plateaued. Training was conducted over 20 epochs with a batch size of 16. The validation set consisted of 20% of the dataset, ensuring robust evaluation during training.

3.5 Performance Analysis

The models were evaluated using three key metrics: Dice coefficient, Jaccard index, and accuracy. U-Net excelled in computational efficiency and provided rapid convergence during training. However, it occasionally missed fine details in small tumor regions. ResNet50, with its powerful feature extraction capabilities, demonstrated superior precision in segmenting irregular and small tumors, albeit at the cost of higher computational requirements.

Metric	**U-Net**	**ResNet50**
Dice Coefficient	0.93	0.94
Jaccard Index	0.92	0.94
Validation Accuracy	97.18%	97.15%
Training Time (per epoch)	88 seconds	120 seconds

Table 1: Comparison of U-Net and ResNet50 performance.

Visualizations of the segmentation masks revealed that ResNet50's ability to retain high-level features improved its performance in capturing small tumor regions. U-Net, while efficient, required additional post-processing to address noise and improve segmentation fidelity.

3.6 Post-Processing

To enhance segmentation outputs, morphological operations such as closing were applied to the binary masks. This step reduced noise and improved boundary smoothness, particularly for U-Net's predictions. ResNet50 required minimal post-processing, further highlighting its robustness.

4 Experimental Evaluation

4.1 Performance Metrics

- **Dice Coefficient:** Measures the overlap between predicted and true masks:

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|}$$

- **Jaccard Index:** Represents intersection-over-union (IoU).
- **Accuracy:** Evaluates the proportion of correctly segmented pixels.

4.2 Results

Model	**Dice Coefficient**	**Jaccard Index**	**Validation Accuracy**
U-Net	0.93	0.92	97.18%
ResNet50	0.94	0.94	97.15%

Table 2: Performance metrics for U-Net and ResNet50.

ResNet50 exhibited marginally better segmentation accuracy, particularly for smaller tumor regions. Visualizations revealed that U-Net occasionally missed fine details, while ResNet50 captured them more effectively.

4.3 Qualitative Analysis

Overlay plots of predicted and ground truth masks demonstrated close alignment. Post-processing with morphological operations reduced noise, further improving segmentation accuracy.

5 Conclusion

5.1 Findings

Both U-Net and ResNet50 achieved high accuracy in liver tumor segmentation. U-Net is computationally efficient and suitable for rapid deployment, while ResNet50 excels in precision for complex datasets.

5.2 Limitations

Dataset size and variability limited generalizability. ResNet50's computational demands may hinder real-time applications.

5.3 Future Directions

- Expanding the dataset with cross-institutional scans to improve robustness.
- Integrating attention mechanisms, such as Transformers [**?**], to enhance feature learning.
- Exploring hybrid architectures combining U-Net's decoder with ResNet50's encoder.

References

[1] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, Grzegorz Chlebus, Hao Chen, Qi Dou, Chi-Wing Fu, Xiao Han, Pheng-Ann Heng, et al. The liver tumor segmentation benchmark (lits), 2019. Available at `https://arxiv.org/abs/1901.04056`.

[2] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016*, 9901:424–432, 2016.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, 9351:234–241, 2015.

[5] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour. Tversky loss function for image segmentation using 3d fully convolutional deep networks. In *International Workshop on Machine Learning in Medical Imaging (MLMI)*, volume 10541, pages 379–387. Springer, 2017.