

# Liver Tumor Segmentation using Image Segmentation Techniques

Shreyas Kulkarni(ssk16),Eranki Vasistha(veranki2),Shantanu Roy(roy33)

## Abstract

This project aims to implement and evaluate advanced image segmentation algorithms for accurately segmenting liver tumors in medical imaging data. The study involves preprocessing and cleaning the dataset to ensure optimal input quality, followed by the implementation of state-of-the-art segmentation models, such as U-Net. The performance of these models is assessed using metrics like Dice coefficient and Intersection over Union (IoU) to quantify segmentation accuracy. A detailed analysis of the results is provided, highlighting the strengths and weaknesses of each model. Additionally, intuitive visualizations of the segmentation outputs are developed to facilitate interpretation by non-technical stakeholders, contributing to better diagnostic and treatment planning support in clinical settings.

## 1 Introduction

Medical imaging has become an indispensable tool in the diagnosis, monitoring, and treatment of various diseases, including cancer. Among the many diagnostic techniques available, CT scans, MRI, and ultrasound are commonly used to detect and analyze liver tumors. A critical step in utilizing these imaging techniques is image segmentation, a process that partitions images into meaningful regions, allowing for the precise identification and localization of tumors. Accurate segmentation of liver tumors plays a vital role in assessing tumor characteristics such as size, shape, and location, which are essential for planning effective treatment strategies.

This project focuses on leveraging advanced image segmentation techniques to analyze and segment liver tumors from medical imaging data. Using the "Liver Tumor Segmentation" dataset from Kaggle, which includes both raw CT images and annotated tumor masks, we aim to develop and evaluate models that assist healthcare professionals

in identifying tumor boundaries with high accuracy. The dataset provides a solid foundation for implementing state-of-the-art segmentation methods and ensuring robust model performance.

Liver tumor segmentation is particularly challenging due to the complex anatomy of the liver and the similarities in intensity and texture with neighboring organs such as the stomach and kidneys. Nonetheless, advancements in deep learning and image processing techniques have opened new avenues for tackling these challenges. By developing and evaluating multiple segmentation models, this project seeks to improve diagnostic accuracy, enhance treatment planning, and contribute valuable insights to the field of medical imaging.

This report details the methodology, experiments, and results of our efforts to achieve accurate liver tumor segmentation. Through rigorous evaluation using metrics such as the Dice coefficient and Intersection over Union (IoU), we aim to identify the most effective approach and provide an in-depth analysis of its strengths and limitations. Additionally, intuitive visualizations are presented to make the results accessible to technical and non-technical audiences alike.

## 2 Background and related work

### 2.1 Deep learning techniques for liver and liver tumor segmentation: A review

This study explores advancements in liver and tumor segmentation from 3D volumetric medical images, a challenging task due to similarities in shape, texture, and intensity between the liver and neighboring organs. It highlights the growing role of deep learning techniques in medical image segmentation, emphasizing their effectiveness in handling complex volumetric data. The research provides an overview of various deep learning methods, their evaluation metrics (e.g., accuracy, Dice coefficient, volume overlap error), and specialized architec-

tures for semantic segmentation. Additionally, it compares existing approaches from previous challenges in liver and tumor segmentation, aiming to identify trends and best practices in the field.

## 2.2 Liver Tumour Segmentation based on ResNet Technique

The referenced paper focuses on using deep learning techniques, specifically ResNet models, for liver and tumor segmentation from CT scans. It evaluates four ResNet-based models—ResNet-18, ResNet-34, ResNet-50, and ResNet-101—using a dataset of 130 CT images. The models were assessed based on performance metrics like training and testing accuracy, epochs, validation loss, and training loss. Among the models, ResNet-34 achieved the highest accuracy (99.2%), followed by ResNet-50. ResNet-101 demonstrated the highest efficiency, while ResNet-18 was the fastest.

This work is related to the project, which also aims to implement image segmentation techniques for liver tumor segmentation in CT images. Like the paper's approach, this project focuses on segmenting liver tumors using advanced methods, and it similarly aims to assess model performance based on key metrics such as accuracy. While the paper uses ResNet models, the project can incorporate similar architectures (e.g., U-Net or ResNet-based models) for comparison and evaluation.

## 3 Data Preprocessing

The preprocessing steps are crucial for preparing CT scan images and their segmentation masks for machine learning. They ensure that the data is clean, structured, and enhanced to highlight the most relevant features for segmentation tasks. The combination of windowing, histogram scaling, and multi-channel conversion ensures that the model has a rich and consistent input, enabling it to focus on tumor boundaries and other critical features. These steps lay the foundation for accurate and reliable liver tumor segmentation.

### 3.1 Data Loading and Organization

For each CT scan file (volume-i.nii), the code identifies its associated mask file (segmentation-i.nii). These pairs are recorded in separate columns, mask\_dirname and mask\_filename. Rows without a corresponding mask file are removed, leaving only valid image-mask pairs. This step ensures clean and reliable input data for the segmen-

tation pipeline. By organizing the data into a structured format, the preprocessing pipeline avoids mismatches and missing data, which could adversely affect model training and evaluation.

### 3.2 Reading .nii Files

The function `read_nii(filepath)` loads .nii files, which are a standard format for medical images, using the nibabel library. The `get_fdata()` method extracts the pixel array from the file, representing the 3D volumetric image. The array is then rotated (`np.rot90`) to ensure proper orientation, aligning with the expected anatomical view. A sample CT scan and its corresponding mask are loaded from the organized DataFrame to test the pipeline. Their shapes are printed to confirm compatibility and integrity. This step converts medical images into pixel arrays that can be processed by machine learning models. Proper orientation ensures consistency in model input.

### 3.3 Image Enhancement and Windowing

The `windowed()` function adjusts image intensity values based on a specified "window width" ( $w$ ) and "window level" ( $l$ ). Intensity values outside the range  $[l-w/2, l+w/2]$  are clipped to improve focus on specific anatomical structures, such as the liver and tumor. The remaining values are normalized to the range  $[0, 1]$ . Windowing enhances contrast for specific tissues, allowing models to better distinguish between organs and abnormalities like tumors. It simulates how radiologists analyze CT images by focusing on relevant intensity ranges.

### 3.4 Frequency Histogram-Based Scaling

**Frequency Histogram Binning** The `freqhist_bins()` function divides pixel intensity values into bins, ensuring each bin contains approximately the same number of pixels. This is done by sorting the intensity values and selecting boundaries that equalize the distribution.

**Histogram Scaling** The `hist_scaled()` function normalizes pixel intensities to a  $[0, 1]$  range using the bins generated earlier. This ensures that intensity variations are uniformly represented across the image.

These steps enhance image contrast and normalize intensity values, making them robust to variations in scanning equipment or settings. This improves model performance by standardizing input data.

### 3.5 Multi-Channel Conversion

**Creating Multi-Channel Images** The `to_nchan()` function generates multi-channel images by combining different windowed versions of the image with a histogram-scaled version. These channels are stacked to form a richer representation of the input image. Multi-channel images provide the model with diverse views of the same data, capturing different levels of detail and enhancing feature extraction for segmentation.

### 3.6 Saving Preprocessed Images

The `save_jpg()` function converts the multi-channel tensor into a format suitable for saving as a JPG image. The image data is scaled to an 8-bit range (0-255), and the channels are arranged for compatibility with standard image formats like RGB or CMYK. The file is saved with adjustable quality. Saving preprocessed images as JPG files makes them accessible to models or systems that require standardized formats. It also enables easy visualization and manual inspection.

### 3.7 Data Preparation for Image Segmentation

This preprocessing step prepares the raw medical imaging data for deep learning model training and evaluation. The dataset, comprising 3D volumetric CT scans and their corresponding segmentation masks in .nii format, is split into training and testing subsets to ensure robust model evaluation. The 3D scans are sliced along the z-axis to generate 2D images, reducing data redundancy by saving every alternate slice. These slices are then saved as .jpg files for CT scans and .png files for masks, ensuring compatibility with deep learning frameworks. By standardizing and organizing the data into structured directories, this step optimizes it for model training, testing, and evaluation, supporting the development of an accurate and efficient liver tumor segmentation pipeline.

### 3.8 Resizing Images

This step prepares the test dataset for evaluating the segmentation model's performance. By resizing, normalizing, and structuring the data appropriately, the code ensures that the model receives consistent and correctly formatted inputs, enabling accurate testing and performance metrics computation.

## 4 Model Training

### 4.1 ResNet-50 Model

#### What is the ResNet-50 Model?

ResNet-50 is a 50-layer deep convolutional neural network (CNN) architecture introduced by Microsoft Research in the paper "Deep Residual Learning for Image Recognition". It uses a concept called "residual learning", which introduces shortcut connections between layers, allowing the model to effectively train very deep networks without the problem of vanishing or exploding gradients.

#### Key Features:

**1. Residual Blocks:** ResNet-50 uses residual blocks to learn the identity mapping  $H(x) = F(x) + x$ , where  $F(x)$  is the output of convolutional layers and  $x$  is the input.

**2. Pre-trained Weights:** It can be pre-trained on large datasets like ImageNet, allowing it to capture general features such as edges, textures, and shapes, which are reusable for many tasks.

**3. Transfer Learning:** ResNet-50 is widely used for transfer learning, enabling fine-tuning for specific applications (e.g., liver tumor segmentation) with smaller datasets.

ResNet-50 significantly enhances liver tumor segmentation by serving as a robust backbone for feature extraction, leveraging its ability to learn hierarchical features from input images. Pre-trained on ImageNet, ResNet-50 provides a strong initialization, enabling efficient training and generalization even with limited medical imaging data. In this task, the fully connected classification layers are excluded (`'include_top=False'`), focusing on the convolutional layers to extract spatial and structural patterns. These features are then adapted for liver tumor segmentation through fine-tuning, making ResNet-50 applicable to medical imaging despite being originally trained on natural images. To generate segmentation masks, the architecture adds upsampling and convolutional layers that progressively restore the spatial resolution of the feature maps.

This allows precise pixel-level predictions, where each pixel is assigned a probability of belonging to the tumor region. By capturing complex tumor shapes and textures, ResNet-50 improves the model's ability to localize and delineate tumors accurately. This combination of transfer learning, hierarchical feature extraction, and segmentation-specific layers makes ResNet-50 a powerful tool

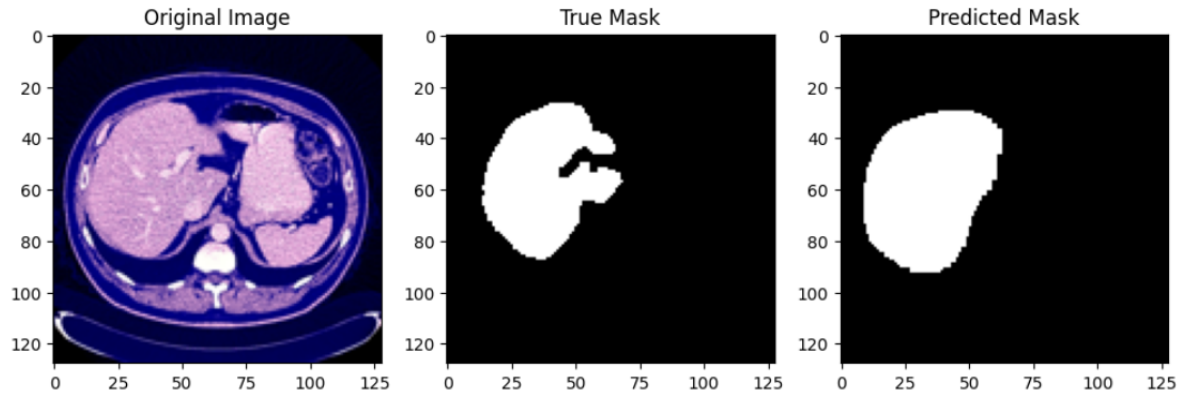


Figure 1: Comparison of original image, true mask, and predicted mask after training a ResNet model for 10 epochs.

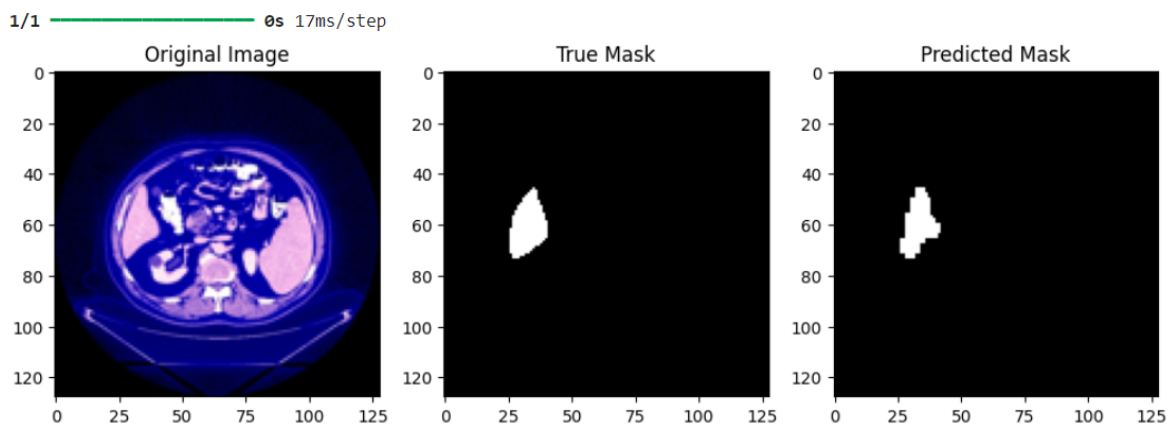


Figure 2: Comparison of original image, true mask, and predicted mask after training a U-Net model for 50 epochs.

for liver tumor segmentation.

ResNet-50 serves as a robust backbone for liver tumor segmentation by leveraging its feature extraction capabilities. By combining pre-trained weights with upsampling and task-specific layers, this method provides an effective pipeline for segmenting tumors in medical images.

## 4.2 U-Net Model

The U-Net model is a convolutional neural network (CNN) architecture designed for biomedical image segmentation. It is widely used in medical imaging tasks due to its ability to precisely localize features and delineate structures, such as tumors, from surrounding tissues. U-Net's strength lies in its **encoder-decoder structure with skip connections**, which combines semantic and spatial information for accurate pixel-level segmentation.

### How the U-Net Model Work

We implemented a tumor segmentation model using an encoder-decoder architecture with skip connections. The encoder extracts features like

edges and textures through convolutional layers and reduces spatial dimensions with max-pooling, allowing the model to focus on high-level features efficiently. The decoder reconstructs the image by upsampling these features back to the original resolution, refining them with additional convolutional layers. Skip connections link encoder and decoder layers, combining detailed spatial information from the encoder with semantic features in the decoder, preserving fine-grained details necessary for accurate segmentation. The output layer generates a binary segmentation mask, assigning each pixel a probability of belonging to a tumor. This approach effectively balances global context and local detail for precise medical image segmentation.

### How U-Net Helps in Liver Tumor Segmentation

U-Net is highly effective for liver tumor segmentation due to its pixel-wise segmentation capability, assigning a class label to each pixel for accurate delineation of irregular tumor shapes. Its skip connections integrate high-level context, such as tumor

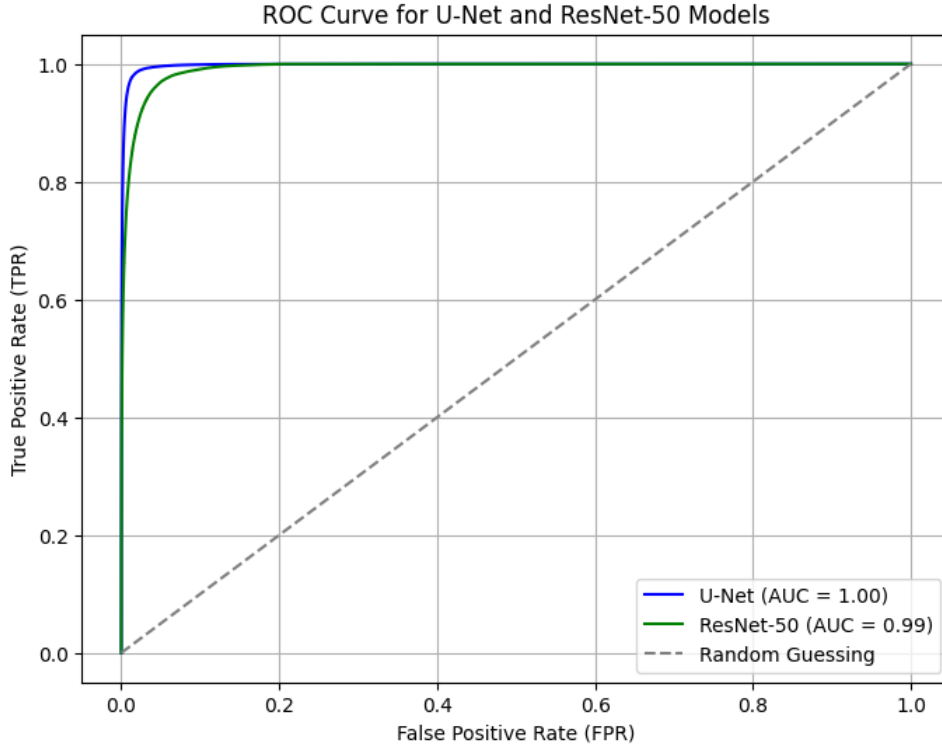


Figure 3: ROC Curve for ResNet-50 Model, demonstrating an AUC of 0.99. The curve shows the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) for the model’s performance.

location, with fine-grained details, like boundaries, to improve accuracy. The encoder extracts complex features, while the decoder ensures precise localization, enabling it to handle small or indistinct tumors. U-Net performs well even with limited datasets, a common scenario in medical imaging, and is adaptable to various imaging modalities like CT and MRI, making it a versatile tool for tumor segmentation tasks.

## 5 Results and Error analysis

The U-Net model demonstrated superior performance compared to ResNet-50 in liver tumor segmentation, as evidenced by key metrics such as the Dice Coefficient and IoU. U-Net achieved a mean Dice Coefficient of 0.821 and a mean IoU of 0.7918, surpassing ResNet-50, which recorded values of 0.7711 and 0.7318, respectively. This highlights U-Net’s advantage in pixel-wise segmentation tasks, likely due to its encoder-decoder architecture and skip connections that effectively preserve spatial information. Furthermore, the ROC curve analysis revealed that both models achieved high AUC scores, with U-Net scoring 1.00 and ResNet-50 closely following at 0.99. These results affirm the models’ strong classification capabilities,

although U-Net’s slightly higher AUC reinforces its superior segmentation accuracy.

Loss analysis over the training epochs further underscores U-Net’s stability and convergence. U-Net exhibited a consistent reduction in both training and validation loss, with only minor fluctuations in validation loss, suggesting efficient learning with minimal overfitting. In contrast, ResNet-50 showed more pronounced fluctuations in validation loss, reflecting potential instability and challenges in generalization. These observations align with ResNet-50’s architectural limitations for pixel-level tasks, originally designed for classification tasks rather than segmentation. Overall, U-Net’s tailored design and robust performance metrics make it a more effective model for liver tumor segmentation.

## 6 Conclusion

In this project, U-Net and ResNet-50 architectures were evaluated for liver tumor segmentation using metrics such as the Dice Coefficient, IoU, and AUC of ROC curves. The U-Net model demonstrated superior performance, achieving a mean Dice Coefficient of 0.821 and a mean IoU of 0.7918, compared to ResNet-50’s 0.7711 and 0.7318, respectively. The ROC curve analysis showed that both mod-

Model	Mean Dice Coefficient	Mean IoU
U-Net	0.8215	0.7976
ResNet-50	0.7711	0.7318

Table 1: Comparison of U-Net and ResNet-50 Models Based on Mean Dice Coefficient and Mean IoU.

els performed exceptionally well in distinguishing tumor regions, with U-Net achieving an AUC of 1.00 and ResNet-50 scoring 0.99. These results underscore U-Net's strength in pixel-wise segmentation, attributed to its encoder-decoder structure with skip connections that effectively capture and preserve spatial features. Loss analysis revealed a smoother convergence for U-Net, with minimal overfitting, while ResNet-50 exhibited greater fluctuations in validation loss, reflecting challenges in stability and generalization. The findings indicate that while both models are capable of accurate tumor segmentation, U-Net's architecture, designed explicitly for segmentation tasks, makes it more robust and reliable for clinical applications. This project highlights the importance of model selection tailored to task-specific requirements and underscores the potential of deep learning models like U-Net in advancing medical image analysis, ultimately improving diagnostic accuracy and aiding in early detection and treatment planning.

## References

- [1] H. Watanabe, M. Bouazizi and T. Ohtsuki, "Hate Speech on Twitter: A Pragmatic Approach to Collect Hateful and Offensive Expressions and Perform Hate Speech Detection," in *IEEE Access*, vol. 6, pp. 13825-13835, 2018, doi: 10.1109/ACCESS.2018.2806394. keywords: Speech;Feature extraction;Twitter;Voice activity detection;Task analysis;Sentiment analysis;Twitter;hate speech;machine learning;sentiment analysis,Galen Andrew and Jianfeng Gao. 2007. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.
- [2] A. Sirco, A. Almisreb, N. M. Tahir and J. Bakri, "Liver Tumour Segmentation based on ResNet Technique," 2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICC-SCE), Penang, Malaysia, 2022, pp. 203-208, doi: 10.1109/ICCSCCE54767.2022.9935636. keywords: Deep learning;Training;Image segmentation;Liver cancer;Computed tomography;Computational modeling;Liver;ResNet;Liver cancer;deep learning;tumour,