scientific reports

Check for updates

OPEN

Liver margin segmentation in abdominal CT images using U-Net and Detectron2: annotated dataset for deep learning models

Mohammad Amir Sattari[1], Seyed Abed Zonouri[1], Ali Salimi[2], Saadat Izadi[2], Ali Reza Rezaei[2], Zahra Ghezelbash[3], Mohsen Hayati[1]✉, Mehrdad Seifi[4] & Milad Ekhteraei[5]

The segmentation of liver margins in computed tomography (CT) images presents significant challenges due to the complex anatomical variability of the liver, with critical implications for medical diagnostics and treatment planning. In this study, we leverage a substantial dataset of over 4,200 abdominal CT images, meticulously annotated by expert radiologists from Taleghani Hospital in Kermanshah, Iran. Now made available to the research community, this dataset serves as a rich resource for enhancing and validating various neural network models. We employed two advanced deep neural network models, U-Net and Detectron2, for liver segmentation tasks. In terms of the Mask Intersection over Union (Mask IoU) metric, U-Net achieved an Mask IoU of 0.903, demonstrating high efficacy in simpler cases. In contrast, Detectron2 outperformed U-Net with an Mask IoU of 0.974, particularly excelling in accurately delineating liver boundaries in complex cases where the liver appears segmented into two distinct regions within the images. This highlights Detectron2's advanced potential in handling anatomical variations that pose challenges for other models. Our findings not only provide a robust comparative analysis of these models but also establish a framework for further enhancements in medical imaging segmentation tasks. The initiative aims not just to refine liver margin detection but also to facilitate the development of automated systems for diagnosing liver diseases, with potential future applications extending these methodologies to other abdominal organs, potentially transforming the landscape of computational diagnostics in healthcare.

Keywords Liver segmentation, CT imaging, Deep learning, U-Net, Detectron2, Medical diagnostics, Annotated dataset

In recent decades, advancements in medical imaging have revolutionized the diagnostic and treatment processes for various health conditions. Among the various imaging techniques, CT has become indispensable for providing detailed and accurate anatomical insights. The importance of accurately detecting specific regions, such as the liver, in these images cannot be overstated, given its critical role in both diagnostic and therapeutic settings. However, the variability in appearance and the complex surrounding anatomy present significant challenges in precise segmentation. In this context, deep neural networks have emerged as transformative tools, leveraging their powerful pattern recognition capabilities to overcome the limitations of traditional image processing methods. Equally important is the role of a reliable dataset, as the accuracy of neural networks heavily depends on the quality and the extent of the data used in training. A well-curated and annotated dataset not only enhances the performance of these models but also sets a benchmark for research, facilitating ongoing improvements and innovations in medical imaging technology[1–4].

In the study presented in[1], an efficient method for segmenting liver and tumors from CT image volumes is introduced using a hybrid ResUNet model, which combines elements of ResNet and U-Net architectures. This study employs two overlapping models to achieve liver segmentation and to evaluate the region of interest within

[1]Electrical Engineering Department, Faculty of Engineering, Razi University, Kermanshah, Iran. [2]Department of Computer Engineering, Faculty of Engineering, Razi University, Kermanshah, Iran. [3]Radiology Department, Clinical Research Development Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran. [4]Clinical research development centre, taleghani and imam Ali hospital, Kermanshah university of medical science, Kermanshah, Iran. [5]Medical Biology Research Center, Health Technology Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran. ✉email: mohsen_hayati@yahoo.com

the images. The segmentation process focuses on examining the liver using the volume of abdominal CT images. However, the proposed model is quite complex and has limited applicability due to its intricate architecture. Ref[5]. describes an approach using a convolutional neural network (CNN) that leverages image patches for liver segmentation. In this method, each patch is centered around a target pixel, which is used for classification. The patches of liver tissue are then classified as either normal or malignant based on their content. A patch is labeled as positive if tumor tissue constitutes half or more of its area. This approach achieves an accuracy of approximately 80.6%. In[6], a deep learning technique originally designed for semantic pixel-wise classification in road scene images is adapted for liver CT segmentation and classification. This technique, known as SegNet, utilizes a deep convolutional encoder-decoder architecture that features a hierarchical arrangement of encoder-decoder layers. Although this architecture demonstrated high accuracy on a standard dataset of liver CT scans, its practicality is limited due to the small dataset size and the specific use case of the model.

Another major issue is the low contrast and noise commonly found in CT images. The liver often lacks distinct boundaries from adjacent organs like the stomach or intestines, especially when contrast is low. Additionally, the noise inherent in CT imaging can obscure important details, further complicating the segmentation process[7–9]. These factors make it challenging for models like U-Net[10] and Detectron2[11], which depend on clear feature distinctions, to accurately delineate the liver. Various preprocessing techniques, such as histogram equalization, contrast-limited adaptive histogram equalization, and denoising algorithms, have been explored to enhance image quality before segmentation[12–14]. Multi-phase CT scans, which provide images at different phases of contrast enhancement, also offer valuable additional information that can improve segmentation accuracy. Despite these methods, achieving precise segmentation in low-contrast and noisy environments is still difficult. It should be note that the liver is an asymmetrical, wedge-shaped organ located beneath the diaphragm in the right upper quadrant of the abdominal cavity. It sits in close proximity to the diaphragm, stomach, and gallbladder, and is predominantly shielded by the costal cartilages[15–18].

These challenges highlight the multifaceted nature of developing robust liver segmentation models using deep learning. Overcoming these obstacles requires a combination of innovative algorithmic strategies, enhanced data annotation practices, and methods to improve model generalization, ensuring reliable performance across diverse clinical settings. To better understand the current landscape and efforts to address these challenges, Table 1 provides an overview of recent studies on segmentation in CT and magnetic resonance imaging (MRI), highlighting the applications, methods, accuracy, and complexity of each approach.

U-Net is a type of CNN specifically designed for the task of biomedical image segmentation. Developed initially for segmenting neural structures in microscopy images, U-Net has since been widely adopted for a variety of medical imaging tasks, including CT and MRI image analysis. The architecture of U-Net is particularly noted for its ability to work effectively with a limited amount of training data, a common challenge in the medical field due to privacy concerns and the high cost of data annotation. The fundamental concept behind U-Net is its symmetric structure, which consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. This design allows U-Net to capture both high-level features and fine-grained details in images, making it highly effective for segmenting complex anatomical structures[19,20].

Detectron2 is an advanced framework designed for object detection and segmentation, which extends the capabilities of its predecessor, Detectron, and incorporates features from other influential models like Mask R-CNN and Faster R-CNN. Developed by Facebook AI Research (FAIR), Detectron2 is particularly useful in applications requiring the detection and precise delineation of objects within complex images, such as medical imaging tasks where accurate localization and segmentation of anatomical features are crucial. Detectron2 utilizes a modular design that allows for easy customization and adaptation to a wide range of object detection and segmentation tasks. It supports various backbones for feature extraction, such as ResNet and ResNeXt, and utilizes a Region Proposal Network (RPN) to generate high-quality object proposals[21–25]. The mathematics underlying Detectron2 involves several key components, primarily revolving around the concepts of RPN and the frameworks for segmentation and detection.

This study seeks to advance the field of medical imaging by rigorously comparing the effectiveness of two prominent deep learning architectures, U-Net and Detectron2, in the accurate segmentation of liver margins from CT scans. The decision to compare U-Net and Detectron2 in this study stems from their widespread use and distinct advantages in medical image segmentation. U-Net has long been recognized for its efficacy in

Ref.	Year	Application	Method	Accuracy	Complexity
[1]	2022	Liver tumor in CT scans segmentation	ResNet and UNet	high	high
[5]	2015	Liver tumor in CT scans segmentation	convolutional neural networks	low	medium
[6]	2020	Liver tumor in CT scans segmentation	modified SegNet	medium	medium
[28]	2023	Lung segmentation in CT scans	convolutional neural networks	low	medium
[29]	2021	Noise reduction in CT scans	Generative Adversarial Network (GAN)	high	medium
[30]	2024	Reconstruction of CT scan images	classical filtered back projection reconstruction with deep learning	medium	high
[35]	2021	Breast tumor in CT scans segmentation	annotation-efficient deep learning	medium	medium
[36]	2023	Liver tumor in CT scans segmentation	Coot Extreme Learning Model	high	high
[39]	2022	Liver tumor in MRI scans segmentation	joint adversarial learning and self-learning	medium	high
[40]	2024	Prostate MRI scans segmentation	Manual segmentation	very low	low

Table 1. Summary of recent methods for image segmentation in medical imaging.

biomedical segmentation tasks, particularly when training data is limited, due to its symmetric architecture with skip connections[26]. Conversely, Detectron2 represents a state-of-the-art framework for object detection and instance segmentation, with advanced capabilities such as Region Proposal Networks (RPN) and feature pyramid networks (FPN), making it adept at handling complex and fragmented anatomical structures[27]. This study addresses this gap by evaluating these models' strengths and limitations using a robust, annotated dataset, aiming to provide insights into their relative effectiveness for liver segmentation in CT images.

Hence, our motivation for this research is driven by the requirement for neural networks to first accurately delineate the liver region as a fundamental step towards future applications and disease diagnosis. To this end, we utilized a comprehensive dataset of over 4,200 abdominal CT images, meticulously annotated by expert radiologists from Taleghani Hospital in Kermanshah, Iran. This dataset, now accessible to the research community, serves as a rich resource for enhancing and validating various neural network models. We employed U-Net and Detectron2 for liver segmentation tasks, revealing their respective strengths in different scenarios. U-Net demonstrated strong efficacy in simpler cases, achieving an Mask Intersection over Union (Mask IoU) score of 0.903. Meanwhile, Detectron2 showcased superior performance with an Mask IoU of 0.974, particularly excelling in accurately delineating liver boundaries in complex cases where the liver appears segmented into two distinct regions within the images. By providing a detailed analysis of these models against this robustly annotated dataset, we aim to highlight the significant potential of deep learning in improving diagnostic processes. The findings from this work not only contribute to the ongoing enhancement of automated medical imaging but also pave the way for the development of more sophisticated diagnostic tools that could drastically improve patient care and treatment outcomes. Through this paper, we demonstrate the critical role of high-quality data and advanced computational techniques in the broader context of healthcare innovation.

The remainder of this paper is organized as follows: "Related work" Section reviews related works, providing an overview of previous studies in the field of liver segmentation and deep learning. "Dataset information" Section presents detailed information about the dataset used in this study, including its characteristics and preparation process. In "Proposed method" Section, we describe the proposed methodology, outlining the implementation of the U-Net and Detectron2 architectures and their application to the segmentation task. "Evaluation" Section discusses the results, highlighting the performance metrics and comparing the efficacy of the two models. "Conclusion" Section concludes the study. Finally, "Supplementary information" Section provides the Supplementary Information, including the mathematical formulations and additional architectural details of the models.

Realated work

In the section, studies related to liver segmentation methods are reviewed and discussed. In[28], various image preprocessing techniques, including histogram equalization, contrast-limited adaptive histogram equalization, and hybrid methods, are developed to enhance segmentation performance. The study utilizes a dataset of 3,616 images and demonstrates that incorporating a hybrid approach with a CNN yields the best segmentation results. This technique improves segmentation accuracy by 1.23% during training and 3.22% during testing, specifically for lung CT scan images. Ref[29]. introduces a generative adversarial network designed to optimize noise reduction in CT scans. The authors evaluated their method by simulating low-dose CT images using a Poisson noise distribution and also employed a human chest phantom at varying exposure levels. However, this study primarily focuses on noise reduction and does not address image segmentation. In contrast, Ref[30]. utilizes a label denoising network, while Ref[31]. implements a neural network trained on a different pre-training task—Gaussian noise removal in natural grayscale images (such as photographs)—rather than directly on CT data. The process of data annotation and ground truth generation is another critical challenge. Training deep learning models requires accurately annotated datasets, where the liver boundary in each CT slice must be carefully delineated. This process is labor-intensive and time-consuming, often requiring manual input from experienced radiologists. The annotation process is not only tedious but also prone to variability due to different radiologists' subjective interpretations, leading to inconsistencies in ground truth data. Efforts to address this challenge include developing semi-automatic or fully automatic annotation tools that provide initial segmentation proposals, reducing the manual effort needed from radiologists. Additionally, consensus-based annotation methods, where multiple experts annotate the same images to produce a unified ground truth, can reduce variability. However, constructing large-scale, consistently annotated datasets remains a bottleneck in the development of effective liver segmentation models[32–34].

In[35], an annotation deep learning framework is introduced as an open-source tool designed to handle incomplete training datasets. While this model has improved segmentation accuracy, its performance remains below the desired ideal levels. In[36], a novel deep learning-based segmentation technique using the Coot extreme learning model is proposed, leveraging publicly available liver scan data. This method not only enhances tumor diagnosis but also achieves high efficiency in segmentation outcomes. The study introduces a geodesic distance encoding technique, which is applied to the initial segmentation and refined using a minimal number of additional users clicks to improve segmentation quality by incorporating inner border points. Furthermore, the classifier's parameters are optimally selected through the Coot optimization method, followed by the application of an extreme learning model for classification. Lastly, the generalization of models across different scanners and imaging protocols poses a significant challenge. CT images can vary widely depending on the scanner type, imaging protocols (such as slice thickness and contrast agent use), and the specific settings applied during scans. This variability affects the appearance of the liver and surrounding tissues, making it challenging for models trained on one set of conditions to perform well on another. This lack of generalization is particularly problematic in clinical environments where multiple scanners and protocols are in use. To overcome this, researchers are increasingly using domain adaptation techniques to enable models to adapt to new data distributions without extensive retraining[37,38]. In[39], the authors propose a method that combines semantic-aware adversarial learning

and joint-aware entropy with a post-situ identification technique. This approach is designed to implicitly align the distribution of task-specific features extracted from the target domain with those from the source domain. The study proposed in[40] investigates the learning progress in manual prostate MRI segmentation by readers with varying levels of experience. It highlights how experience and specialized training significantly impact the accuracy and quality of prostate MRI segmentation. The authors analyze the improvement in performance over time and the differences between readers with different experience levels.

Building upon the limitations identified in prior works, this study compares the performances of U-Net and Detectron2 for liver segmentation in CT images. The proposed evaluation demonstrates the distinct capabilities of each method in handling complex anatomical variations and achieves high Mask IoU scores, showcasing their potential for advancing automated liver segmentation and enhancing diagnostic accuracy.

Dataset information

Data specifications for each case are detailed in Table 2. The dataset utilized in this study comprises a comprehensive collection of CT scan images focusing on the abdominal regions. This dataset is specifically designed to facilitate the investigation and diagnosis of liver diseases, serving as a valuable resource for developing and evaluating deep learning algorithms for liver detection and segmentation tasks. The dataset includes 25 distinct cases, each identified by a unique case number rather than personal information. It encompasses a total of 4,211 CT scan images. Each case is structured into two primary components: one folder containing the CT scan images and another folder housing the corresponding segmentation masks. This organization allows researchers to access both the raw imaging data and the pre-processed segmentation information, which is crucial for training and validating machine learning models while maintaining patient anonymity. One notable feature of this dataset is the variation in the number of images and masks across cases. The number of CT images per case ranges from 66 (Case 20) to 707 (Case 10), reflecting the diversity of the dataset and potentially capturing a wide spectrum of anatomical variations and pathological conditions. Similarly, the number of masks varies, with some cases having more masks than images, indicating potential overlapping or multi-layer segmentations. The segmentation masks are particularly valuable as they provide pixel-level annotations of regions of interest within the liver. These masks typically cover a subset of the total images in each case, focusing on the most relevant slices for liver analysis. The mask ranges vary across cases, with some starting as early as image 1 (Case 9) and others extending to image 292 (Case 10).

To further enhance the dataset's utility and facilitate more comprehensive analyses, additional demographic and clinical information has been incorporated into the dataset description. This includes columns for age, gender, health status (healthy/unhealthy), and type of disease (Table 3). Once populated, this additional data will enable researchers to conduct more nuanced analyses considering factors such as age-related changes in liver structure, gender-specific variations, and the impact of different liver diseases on imaging characteristics. The inclusion of the "Type of Disease" column is particularly significant. It allows for the categorization of cases based on specific liver pathologies. This information, when available, will be invaluable for training disease-specific models or for studying the imaging characteristics associated with different liver disorders. The composition of this dataset makes it exceptionally suitable for a wide range of medical imaging research applications. It can be used to train deep learning models for automated liver segmentation, to develop computer-aided diagnosis systems for liver diseases, or to investigate the efficacy of various image processing techniques in enhancing CT scan interpretations. The potential addition of demographic and clinical data will further expand its utility, allowing for more sophisticated analyses that consider patient-specific factors in liver disease diagnosis and progression. Crucially, all images, labels, and segmentations underwent thorough review and approval by a radiologist with specialized expertise in liver imaging. This ensures the accuracy and reliability of the annotations, making the dataset a robust resource for medical imaging research.

In this study, we use a novel Deep Learning-based approach for preprocessing medical images, particularly CT scans, that simultaneously reduces noise and enhances resolution. Our method employs a CNN architecture inspired by U-Net and ResNet, which we term DnCNN-SR (Denoising Convolutional Neural Network with Super-Resolution). The preprocessing pipeline begins with normalizing input images to a [0,1] range. The DnCNN-SR network consists of three main components: a contracting path for feature extraction, a bottleneck for noise mapping, and an expanding path for super-resolution. The contracting path uses stridden convolutions to downsample the image while extracting hierarchical features. The bottleneck employs dilated convolutions to effectively map and remove noise patterns. The expanding path utilizes transposed convolutions for upsampling, guided by skip connections from the contracting path to preserve fine details. This end-to-end approach allows

Characteristic	Description
Data Type	CT scan images of the abdomen
Number of Cases	25
Total CT Images	4211
Image Content	Abdominal regions
Accompanying Data	Segmentation masks for each case
Purpose	Segmenting the border of the liver in CT images
Imaging Location	Ayatollah Taleghani Hospital, Kermanshah, Iran

Table 2. Data specifications.

Case Number	Number of CT Images	Number of Masks	Mask Range	Age	Gender	Health Status	Type of Disease
1	108	120	image 8 to 47	50	male	unhealthy	liver lesion
2	305	72	image 10 to 33	34	male	healthy	-
3	100	102	image 8 to 41	38	female	healthy	
4	95	102	image 7 to 40	52	female	unhealthy	fatty liver
5	90	129	image 3 to 45	61	female	healthy	
6	96	105	image 6 to 40	70	female	unhealthy	liver lesion
7	118	114	image 7 to 44	14	male	healthy	
8	94	96	image 8 to 39	42	female	healthy	
9	110	114	image 1 to 38	29	male	unhealthy	fatty liver
10	707	717	image 54 to 292	70	female	healthy	
11	115	93	image 30 to 60	18	male	healthy	
12	99	132	image 4 to 47	35	male	healthy	
13	591	534	image 45 to 222	39	male	healthy	
14	104	102	image 9 to 42	22	male	healthy	
15	109	96	image 14 to 45	22	male	healthy	
16	89	81	image 6 to 32	20	female	healthy	
17	85	84	image 6 to 33	76	male	unhealthy	liver lesion
18	521	495	image 36 to 200	41	female	unhealthy	liver lesion
19	99	99	image 9 to 41	17	male	healthy	
20	66	78	image 4 to 29	4	male	healthy	
21	87	99	image 4 to 36	75	female	healthy	
22	109	111	image 13 to 49	16	male	healthy	
23	101	111	image 4 to 40	15	female	healthy	
24	103	120	image 8 to 47	33	female	healthy	
25	110	90	image 12 to 41	34	male	healthy	

Table 3. Demographic and clinical characteristics of cases in the dataset.

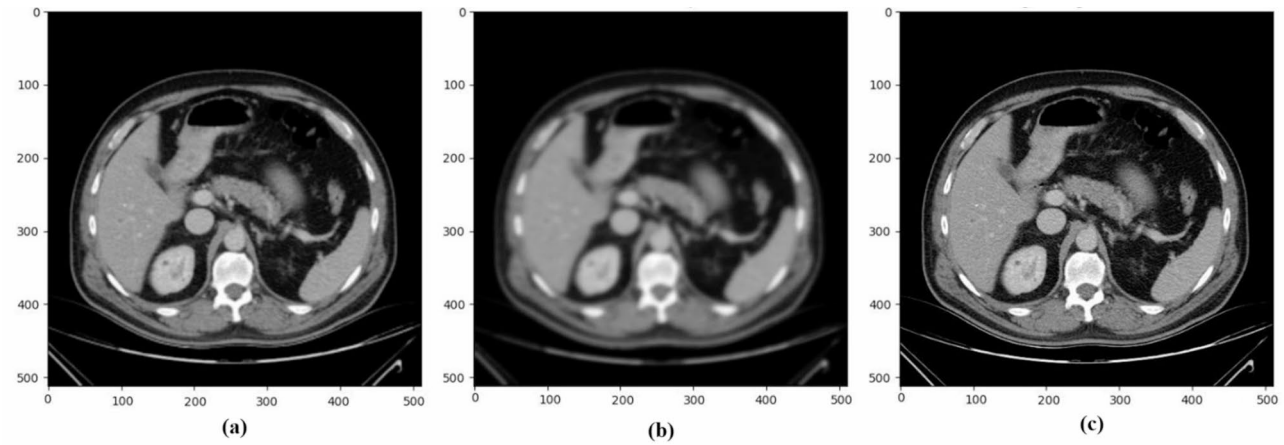


Fig. 1. Demonstration of noise reduction using DnCNN-SR: (**a**) Low-resolution CT image, (**b**) CT image processed using DnCNN-SR for noise reduction, and (**c**) Original high-resolution CT image for comparison.

the network to learn the complex relationship between noisy, low-resolution inputs and clean, high-resolution outputs. We train the network using a combination of Mean Squared Error (MSE) and Structural Similarity Index (SSIM) as the loss function, balancing pixel-wise accuracy with perceptual quality. Quantitative evaluation using Peak Signal-to-Noise Ratio (PSNR) and SSIM, along with qualitative assessment by radiologists, demonstrates significant improvements in both noise reduction and image clarity. This preprocessing technique shows promise in enhancing the diagnostic value of medical images, potentially improving the accuracy and efficiency of clinical assessments.

The effectiveness of the noise reduction method is demonstrated in Fig. 1, where (a) shows a low-resolution CT image, (b) illustrates the noise-reduced output processed using DnCNN-SR, and (c) presents the original high-resolution CT image for comparison. This comparison highlights the substantial enhancement in image quality achieved through the DnCNN-SR method. Also, Fig. 2 presents a typical CT scan image from the dataset,

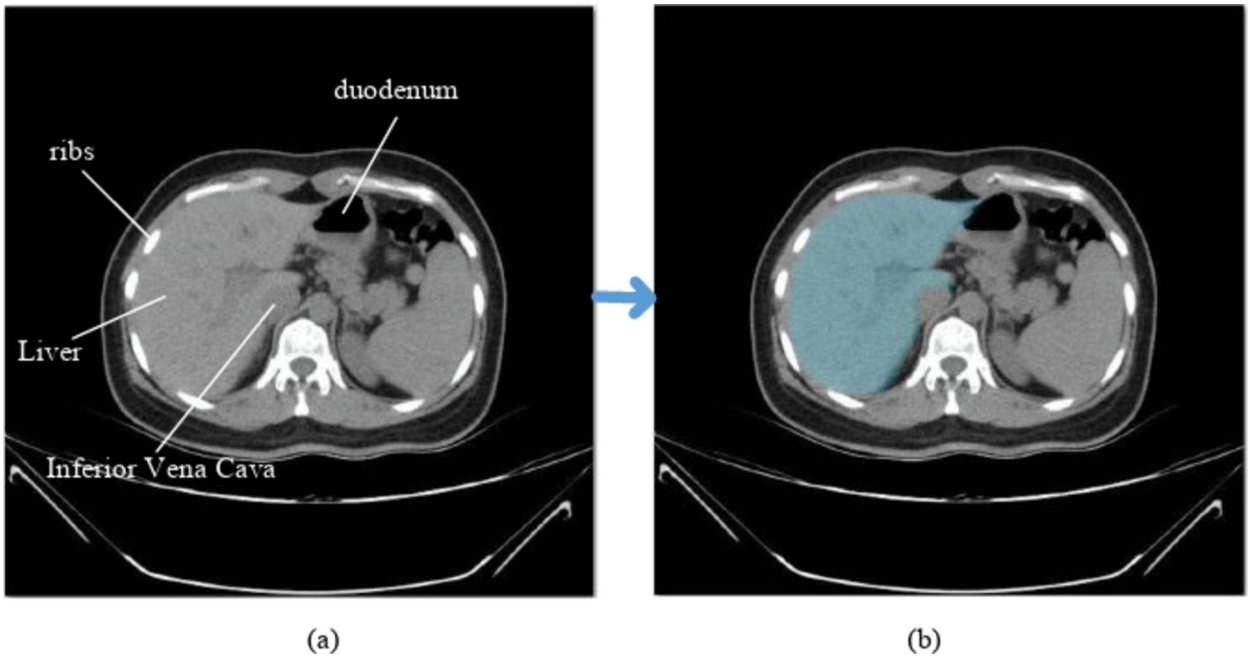


Fig. 2. (**a**) Original CT image with annotated anatomical structures (liver, ribs, duodenum, and inferior vena cava). (**b**) Segmented liver region (blue) demonstrating accurate boundary delineation by the proposed model.

highlighting the liver margin annotated by a medical expert. This image exemplifies the detailed annotation process, where the liver, along with adjacent anatomical structures such as the gallbladder, inferior vena cava, ribs, stomach, and, duodenum are clearly marked. The highlighted liver area in blue demonstrates the precise segmentation achieved, which is critical for training and validating the performance of deep learning models used in the study. This level of detail ensures that the models learn to recognize and delineate the liver accurately from surrounding tissues, which is essential for effective medical diagnostics. The detailed dataset preparation process, including annotation protocols and preprocessing techniques, is described in Supplementary Information Section.

Proposed method

This paper aims to rigorously compare the effectiveness of two prominent deep learning architectures, U-Net and Detectron2, in accurately segmenting liver from CT scans. U-Net and Detectron2 were chosen for this study due to their complementary strengths in medical image segmentation tasks. U-Net is widely recognized for its efficiency in biomedical image segmentation, particularly in handling pixel-wise predictions with limited data. Detectron2, on the other hand, excels in instance segmentation and object detection, making it suitable for complex cases where anatomical structures exhibit significant variability. While previous studies have evaluated U-Net or Detectron2 individually, a direct comparison of these two models for liver segmentation in CT images has not been thoroughly explored[41–46]. This study aims to fill that gap by providing a robust comparative analysis of their performance in both simple and complex liver segmentation tasks.

U-Net, developed for biomedical image segmentation, is widely used due to its ability to capture both high-level and fine-grained features through its contracting and expanding path architecture[19,20]. Detectron2, on the other hand, is a modular framework for object detection and instance segmentation, leveraging state-of-the-art methods like Mask R-CNN and Feature Pyramid Networks (FPN) for robust performance in complex segmentation tasks. Detailed mathematical formulations for both models can be found in Supplementary Information[21–25]. The mathematical formulations lay the groundwork for U-Net's ability to perform precise and efficient image segmentation, crucial for medical imaging applications. Implementing U-Net involves setting up the network architecture with alternating convolutional and max-pooling layers in the contracting path, followed by up-convolutional layers and concatenations with corresponding cropped feature maps from the contracting path in the expanding path. Training a U-Net typically requires preparing a dataset of images and corresponding segmentation masks, which are used as ground truth. The network is trained using a suitable loss function, such as the cross-entropy to measure the discrepancy between the predicted and actual segmentations. During training, the parameters of the network are adjusted t;,o minimize the loss function using optimization algorithms like stochastic gradient descent. Data augmentation techniques such as rotation, scaling, and elastic deformations are often employed to increase the diversity of training samples and prevent overfitting, enhancing the network's generalization capabilities.

Detectron2 and U-Net are designed for different tasks within the field of image analysis. U-Net is specifically tailored for image segmentation, which involves classifying each pixel of an image as belonging to a particular class. It is particularly suited for medical image segmentation due to its efficient use of data through skip

connections and its ability to capture fine details through upscaling paths. In contrast, Detectron2 is more focused on object detection and instance segmentation, which requires identifying each distinct object in an image and delineating its boundaries. Detectron2 excels in scenarios where multiple object classes are present and where precise localization of each object is critical. To implement and train a Detectron2 model, one would typically follow these steps:

- **Preparation**: Prepare a dataset annotated with bounding boxes and, for segmentation tasks, pixel-wise masks.
- **Configuration**: Set up the configuration settings, which include selecting the backbone architecture, RPN settings, and other hyperparameters.
- **Model Training**: Train the model using the prepared dataset. Detectron2 utilizes a variety of loss functions to optimize the detection and segmentation tasks, such as cross-entropy loss for classification and smooth L1 loss for bounding box regression.
- **Evaluation and Fine-tuning**: After training, evaluate the model's performance on a validation set, and adjust hyperparameters or model settings based on performance metrics.

Detectron2 provides a flexible and powerful platform for developing state-of-the-art object detection and segmentation models, which can be particularly advantageous in medical imaging scenarios where high precision and reliability are required.

Figure 3 illustrates the comprehensive workflow utilized in our study for the segmentation of liver margins in CT images. The process begins with the collection and labeling of CT images, which involves meticulous annotation by expert radiologists to ensure the accuracy of training data. These images then undergo a crucial noise reduction step to enhance clarity and reduce potential artifacts, preparing them for effective segmentation analysis. Following noise reduction, the images are segmented manually using the KITTI 1.0 format, which provides a standardized framework for annotating medical images. The core of our study involves the implementation of two advanced deep learning models, U-Net and Detectron2, which are applied to the prepared images to perform automated segmentation. The final step in the workflow involves comparing the performance of these models to evaluate their effectiveness in accurately delineating liver boundaries. The segmented outputs from U-Net and Detectron2 are shown, highlighting the precision of each model in handling complex liver structures within the CT images.

A detailed overview of the architecture and mathematical framework of both U-Net and Detectron2 is provided in Supplementary Information.

Evaluation

In the evaluation section, the results of the two networks are presented, including performance metrics, implementation details, and an in-depth discussion of their effectiveness. This analysis highlights the strengths and limitations of each model in the liver segmentation task.

Results

Upon entering the U-Net network, the CT images are processed, and at the conclusion of the network, they are compared against their corresponding segmentation masks. This comparison is facilitated by a predefined loss function that quantifies the discrepancies between the predicted segmentation and the ground truth masks. Subsequent to this, the backpropagation algorithm updates the network's weights and biases to minimize the loss, an iterative process that continues until the network parameters converge to their optimal values, thus maximizing the model's performance. It is important to note that all input images are standardized to a resolution of 512×512 pixels. This standardization ensures consistency during training and enables the network to focus more effectively on structural details critical for accurate liver segmentation, while reducing computational demands. The U-Net architecture is designed to enhance the capability of CNNs by utilizing multiple layers. Each layer in the network captures increasingly complex features, enhancing the model's ability to perform detailed segmentation tasks. Specifically, the network architecture employs a 3D CNN approach, where the input feature map consists of a CT image block of size $512 \times 512 \times 1$. This is followed by nineteen convolutional layers,

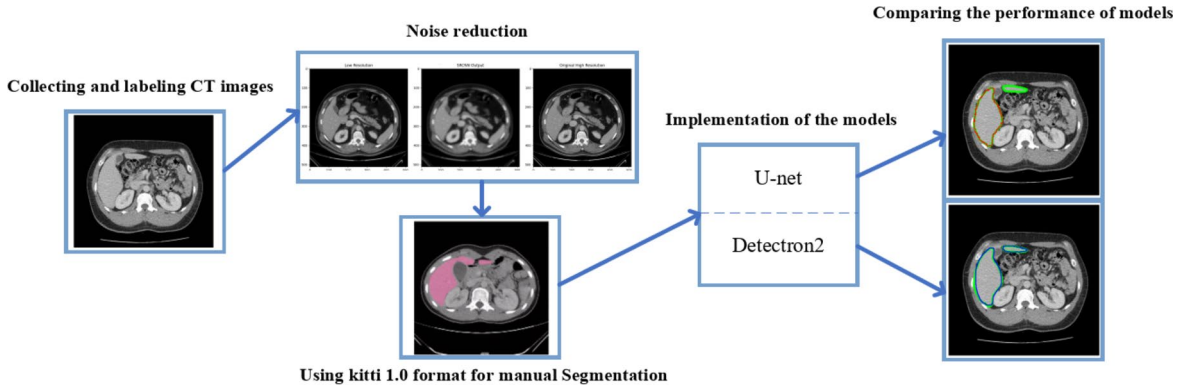


Fig. 3. Research workflow diagram.

each equipped with various filters and succeeded by a ReLU activation function to facilitate faster training. As detailed in Table 4, the U-Net architecture also includes a concatenation step where the cropped feature map from the contraction path is combined with the upsampled output. At each upsampling step, the number of feature channels is doubled, thereby enhancing the network's capacity to learn and retain detailed features from the high-resolution contracting path. Despite the advanced capabilities of the U-Net architecture, it exhibits some limitations, particularly evident in scenarios where the liver is segmented into two distinct regions within the images. In these cases, the U-Net model sometimes fails to accurately delineate and segment the liver areas, struggling with the precise identification of boundaries. This limitation highlights the challenges posed by complex anatomical variations that are not uniformly addressed by standard segmentation models. The difficulty in adapting to the unique contours and boundaries of divided liver images suggests a need for model refinement to improve segmentation accuracy in complex imaging conditions.

The Detectron2 network utilized in this study represents a robust architecture optimized for object detection and segmentation tasks. As outlined in Table 5, the network structure begins with an input layer that processes images of size 512 × 512 pixels. The backbone of the network is a ResNet-101 model enhanced with a FPN,

Layer Type	Filters	Kernel Size	Stride	Activation	Feature Size	Output Shape
Input	-	-	-	-	512 × 512 × 1	512 × 512 × 1
Convolution	16	3 × 3	1	ReLU	512 × 512 × 1	512 × 512 × 16
Dropout	-	-	-	-	512 × 512 × 16	512 × 512 × 16
Convolution	16	3 × 3	1	ReLU	512 × 512 × 16	512 × 512 × 16
MaxPooling	-	2 × 2	2	-	512 × 512 × 16	256 × 256 × 16
Convolution	32	3 × 3	1	ReLU	256 × 256 × 16	256 × 256 × 32
Dropout	-	-	-	-	256 × 256 × 32	256 × 256 × 32
Convolution	32	3 × 3	1	ReLU	256 × 256 × 32	256 × 256 × 32
MaxPooling	-	2 × 2	2	-	256 × 256 × 32	128 × 128 × 32
Convolution	64	3 × 3	1	ReLU	128 × 128 × 32	128 × 128 × 64
Dropout	-	-	-	-	128 × 128 × 64	128 × 128 × 64
Convolution	64	3 × 3	1	ReLU	128 × 128 × 64	128 × 128 × 64
MaxPooling	-	2 × 2	2	-	128 × 128 × 64	64 × 64 × 64
Convolution	128	3 × 3	1	ReLU	64 × 64 × 64	64 × 64 × 128
Dropout	-	-	-	-	64 × 64 × 128	64 × 64 × 128
Convolution	128	3 × 3	1	ReLU	64 × 64 × 128	64 × 64 × 128
MaxPooling	-	2 × 2	2	-	64 × 64 × 128	32 × 32 × 128
Convolution	256	3 × 3	1	ReLU	32 × 32 × 128	32 × 32 × 256
Dropout	-	-	-	-	32 × 32 × 256	32 × 32 × 256
Convolution	256	3 × 3	1	ReLU	32 × 32 × 256	32 × 32 × 256
Conv2DTranspose	128	2 × 2	2	-	32 × 32 × 256	64 × 64 × 128
Concatenate	-	-	-	-	64 × 64 × 128	64 × 64 × 256
Convolution	128	3 × 3	1	ReLU	64 × 64 × 128	64 × 64 × 256
Dropout	-	-	-	-	64 × 64 × 256	64 × 64 × 128
Convolution	128	3 × 3	1	ReLU	64 × 64 × 128	64 × 64 × 128
Conv2DTranspose	64	2 × 2	2	-	64 × 64 × 128	128 × 128 × 64
Concatenate	-	-	-	-	128 × 128 × 64	128 × 128 × 128
Convolution	64	3 × 3	1	ReLU	128 × 128 × 64	128 × 128 × 128
Dropout	-	-	-	-	128 × 128 × 128	128 × 128 × 64
Convolution	64	3 × 3	1	ReLU	128 × 128 × 64	128 × 128 × 64
Conv2DTranspose	32	2 × 2	2	-	128 × 128 × 64	256 × 256 × 32
Concatenate	-	-	-	-	256 × 256 × 32	256 × 256 × 64
Convolution	32	3 × 3	1	ReLU	256 × 256 × 32	256 × 256 × 64
Dropout	-	-	-	-	256 × 256 × 64	256 × 256 × 32
Convolution	32	3 × 3	1	ReLU	256 × 256 × 32	256 × 256 × 32
Conv2DTranspose	16	2 × 2	2	-	256 × 256 × 32	512 × 512 × 16
Concatenate	-	-	-	-	512 × 512 × 16	512 × 512 × 32
Convolution	16	3 × 3	1	ReLU	512 × 512 × 16	512 × 512 × 32
Dropout	-	-	-	-	512 × 512 × 32	512 × 512 × 16
Convolution	16	3 × 3	1	ReLU	512 × 512 × 16	512 × 512 × 16
Convolution	1	1 × 1	1	Sigmoid	512 × 512 × 16	512 × 512 × 1

Table 4. Detailed architecture of CNNs used in U-Net.

Layer Type	Details
Input	Image of size 512 × 512 × 1
Backbone	ResNet-101 with Feature Pyramid Network (FPN)
Conv1	7 × 7, 64 filters, stride 2
MaxPool	3 × 3, stride 2
Res2 Block	[1 × 1, 64], [3 × 3, 64], [1 × 1, 256] × 3
Res3 Block	[1 × 1, 128], [3 × 3, 128], [1 × 1, 512] × 4
Res4 Block	[1 × 1, 256], [3 × 3, 256], [1 × 1, 1024] × 23
Res5 Block	[1 × 1, 512], [3 × 3, 512], [1 × 1, 2048] × 3
FPN	Build a top-down feature pyramid from the ResNet feature maps
RPN	3 × 3 convolution, 256 channels, followed by two sibling 1 × 1 convolutions for classification and regression
ROI Align	Align the proposed regions to fixed size (e.g., 7 × 7)
Bounding Box Head	2 fully connected layers, 1024 units each, followed by class and box regression layers
Mask Head	4 convolutional layers, 256 filters each, followed by a deconvolution layer for mask prediction

Table 5. Detailed architecture of CNNs used in Detectron2.

which is crucial for detecting objects at different scales. The network comprises multiple convolutional layers arranged into blocks (Res2 to Res5), each with an increasing number of filters and decreasing spatial dimensions, allowing for deep feature extraction at various levels of granularity. The FPN within Detectron2 constructs a top-down architecture where higher-level semantic information is merged with lower-level features, enabling precise localization and segmentation capabilities. The Region Proposal Network (RPN) then generates region proposals, which are defined by the ROI Align layer that aligns the proposed regions to a fixed size, ensuring that the features extracted are accurately scaled to the network's subsequent layers. The final stages of Detectron2 include the Bounding Box Head and the Mask Head. The Bounding Box Head consists of two fully connected layers followed by class and box regression layers, which determine the object classes and their bounding boxes. Meanwhile, the Mask Head comprises four convolutional layers followed by a deconvolution layer, specifically designed for mask prediction. This allows Detectron2 to not only localize objects within an image but also generate precise pixel-wise segmentations.

Contrary to the U-Net model, Detectron2 has demonstrated remarkable capability in handling images where the liver is segmented into two distinct regions. The advanced architecture and sophisticated feature integration in Detectron2 enable it to effectively delineate liver boundaries, even in complex cases where the liver appears in disjointed sections. This ability marks a significant improvement over the U-Net model, particularly in scenarios that challenge traditional segmentation models.

As depicted in Fig. 4, the performance of the Detectron2 and U-Net models in segmenting liver margins from CT images is quantitatively assessed using Mask IoU and loss function metrics, with results that are highly ideal and demonstrate the significant value of this study. Detectron2 achieves a peak Mask IoU of 0.974 during validation (Fig. 4(a)), nearly perfectly aligning with the ground truth in most cases. This outstanding performance underscores Detectron2's exceptional capability in handling complex liver geometries, which can be attributed to its sophisticated architecture that combines deep feature extraction with precise spatial localization. Additionally, its training and validation loss curves (Fig. 4(c)) converge rapidly to minimal values, highlighting its efficient optimization and strong generalization ability to unseen data.

Similarly, U-Net achieves a peak Mask IoU of 0.903 during validation (Fig. 4(b)), which, while slightly lower than Detectron2, still reflects a high level of accuracy in liver segmentation. This performance showcases the robustness of U-Net, despite its simpler architecture. However, the slightly slower convergence and marginally higher validation loss compared to training loss (Fig. 4(d)) suggest limitations in capturing intricate boundaries and handling ambiguous regions as effectively as Detectron2.

These results are ideal and validate the effectiveness of the segmentation methods evaluated in this work, with both models achieving high accuracy and robust optimization. The clear superiority of Detectron2, particularly in boundary delineation and overall precision, demonstrates the value of this study and its contributions to advancing segmentation methodologies for challenging medical imaging tasks.

Mask Intersection over Union (Mask IoU) is a common metric used to evaluate the accuracy of object detection models on certain datasets, especially in tasks where precise localization is critical, such as in medical image segmentation. Mask IoU is calculated by dividing the area of overlap between the predicted segmentation and the ground truth by the area of the union of the predicted and actual labels. Mathematically, it is expressed as:

$$Mask\ IoU = \frac{\text{Area of Overlap between the Prediction and Ground Truth}}{\text{Area of Union of the Prediction and Ground Truth}} \quad (1)$$

This metric provides a quantitative measure of how well the predicted boundaries align with the true boundaries of objects, with a higher Mask IoU indicating better model performance.

To provide a more comprehensive evaluation of the segmentation models, additional performance metrics beyond Mask IoU were included. While Mask IoU is commonly used to assess the overall overlap between

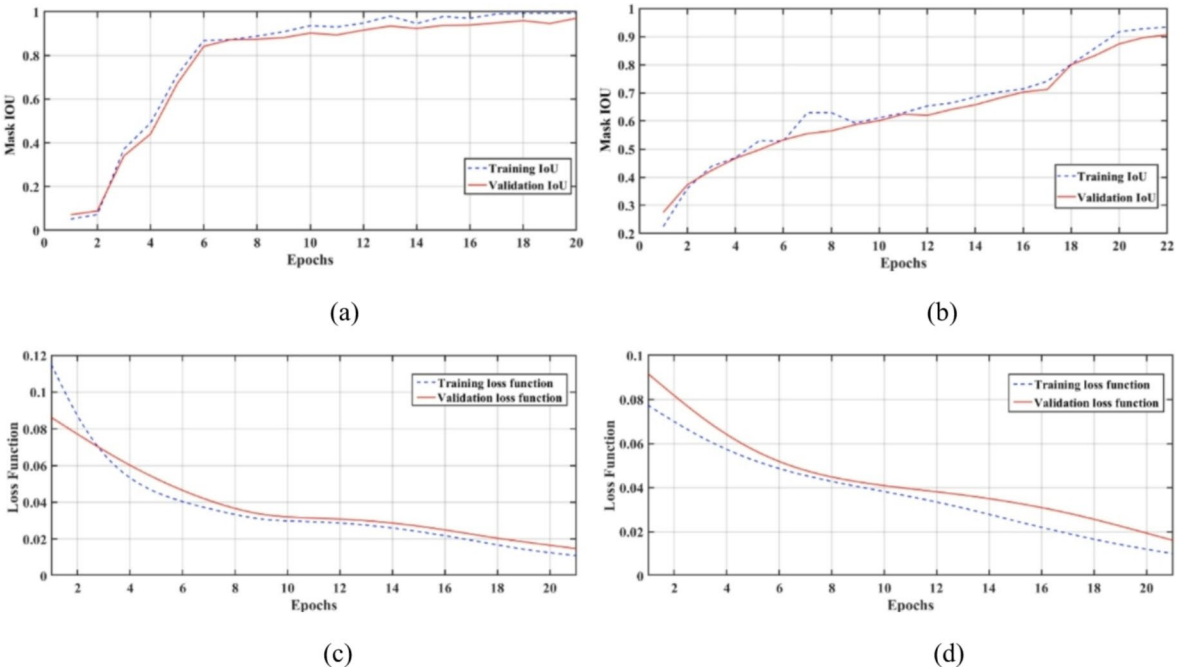


Fig. 4. Training and validation Mask IoU and loss curves for (**a**), (**c**) Detectron2, (**b**), and (**d**) U-Net.

Metric	U-Net	Detectron2
Mask IoU	0.903	0.974
Trimap IoU	0.8891	0.92.51
F-Measure	0.9517	0.9902
Boundary IoU	0.8469	0.9122

Table 6. Comparison of U-Net and Detectron2 performance on key segmentation metrics.

predicted segmentation masks and the ground truth, it does not fully capture boundary accuracy, which is critical in medical imaging applications such as liver segmentation[47,48].

Trimap IoU was introduced to specifically evaluate segmentation performance near object boundaries. This metric considers a narrow region around the ground truth boundary to measure how well the predicted segmentation aligns with it, providing insights into the model's ability to capture fine structural details.

F-Measure, a harmonic mean of precision and recall, was also incorporated to assess the alignment accuracy between the predicted and actual segmentation boundaries. This metric ensures that both the completeness and correctness of boundary predictions are considered.

Additionally, Boundary IoU was included to focus on the segmentation quality at the object edges by only considering pixels within a predefined distance from the actual boundary. This metric helps in evaluating how effectively the model delineates precise anatomical structures, which is particularly useful in complex cases where the liver margins may be irregular or fragmented.

The results of these metrics, which provide a comparative analysis of the two models, are summarized in Table 6. The findings highlight the superior performance of Detectron2 in boundary-sensitive evaluations compared to U-Net, further emphasizing its effectiveness in medical image segmentation tasks. These metrics were calculated using an independent dataset of CT images, which was not used during training or validation. This approach ensured that the reported results accurately reflect the real-world generalizability of the models. By evaluating the most well-trained models on unseen data, we confirmed that Detectron2 consistently outperforms U-Net across all metrics, particularly in metrics focused on boundary quality.

Discussion

The significance of this research lies in its substantial contribution to the field of medical imaging and diagnostics. By leveraging advanced deep learning models like U-Net and Detectron2 for liver segmentation, this study not only enhances the precision of medical imaging analyses but also paves the way for more accurate and early diagnosis of liver diseases. The high performance of the Detectron2 model, in particular, demonstrates its potential in handling complex segmentation tasks which are often encountered in medical scans involving pathological conditions that alter normal anatomy. The research utilized a robust computational setup to ensure the efficient processing of complex neural network models. The proposed method was trained on a machine

equipped with two GPUs: an NVIDIA RTX 3080Ti and an NVIDIA RTX 1080Ti, each with 8GB of RAM. This setup was powered by an Intel Core i9-7800X processor with a base frequency of 3.50 GHz and a maximum turbo frequency of 4.00 GHz, supported by 128 GB of DDR4 RAM. The operating system used was Ubuntu 24.04, and the development environment included JupyterLab 3.6.7 IDE alongside Python, which offers extensive libraries for neural network training and image processing tasks.

Furthermore, the open access availability of the meticulously annotated dataset used in this study represents a valuable resource for the research community. Making these labeled images publicly accessible underpins a collaborative framework that encourages the development of improved and more diverse analytical models. Researchers can utilize this dataset to benchmark their algorithms, fostering innovation and potentially leading to breakthroughs in automated disease diagnosis. Looking ahead, the implications of this research are manifold. The dataset's utility extends beyond liver diseases to potentially include other abdominal pathologies. As machine learning models evolve, the scope for using this dataset to train algorithms capable of detecting and diagnosing a wide range of diseases will expand, thereby enhancing the efficacy of medical imaging in clinical settings. The open-access nature of the dataset ensures that it can serve as a foundation for future research, aiding in the development of algorithms that are more adept at interpreting complex medical images.

To evaluate the robustness of the models in challenging scenarios, five complex liver cases were selected where the liver appeared fragmented, irregularly shaped, or poorly defined. Figure 5 presents these cases in three states: (1) the ground truth mask annotated by the radiologist (left), (2) the segmentation predicted by Detectron2 (middle), and (3) the segmentation predicted by U-Net (right). These cases were chosen to emphasize scenarios where precise boundary delineation and segmentation accuracy are critical. The results show that both Detectron2 and U-Net deliver exceptional segmentation performance, showcasing their mastery in handling complex liver cases. Detectron2 demonstrates a slight advantage in challenging scenarios, consistently achieving accurate boundary delineation and maintaining segmentation continuity, even when the liver appears fragmented or irregular. This can be attributed to its advanced feature extraction and region-based prediction capabilities.

U-Net, while slightly less effective in these specific cases, remains a robust and reliable model, achieving highly accurate results in most scenarios. Its performance in delineating liver structures with ambiguous or poorly defined edges highlights the strength of its encoder-decoder architecture and its suitability for medical imaging tasks. Both models represent significant advancements in automated liver segmentation, with Detectron2 excelling in particularly complex cases and U-Net demonstrating strong overall performance. These findings underscore the value of both approaches for clinical applications, where precise and reliable segmentation is essential.

The segmentation models presented in this study, particularly Detectron2, demonstrate robust performance in handling complex anatomical structures, making them promising candidates for broader applications in medical imaging. Beyond liver segmentation, these models could be adapted to segment other abdominal organs such as the pancreas, spleen, or kidneys, which also present significant challenges due to their anatomical variability and close proximity to surrounding tissues. Additionally, our approach could be extended to address liver-specific pathologies, including segmentation of hepatic lesions, cysts, or tumors. Accurate segmentation of these pathological structures is crucial for disease diagnosis, treatment planning, and surgical navigation. Detectron2's capability to handle fragmented and low-contrast regions could make it particularly valuable in such applications. Future work could involve integrating domain-specific information or training on datasets representing a broader range of pathologies to further enhance the generalization and effectiveness of these models. By broadening the scope of applications, this research not only underscores the utility of deep learning models in liver segmentation but also sets the stage for developing advanced tools to improve diagnostics and treatment across various clinical contexts.

Conclusion

In this study, we have demonstrated the application of advanced deep learning models for the segmentation of liver margins in CT images, a task that presents significant challenges due to the complexity of liver morphology and its implications for medical diagnostics and treatment planning. Employing two neural network models, U-Net and Detectron2, we conducted a robust comparative analysis based on a dataset meticulously annotated by expert radiologists from Taleghani Hospital in Kermanshah, Iran. Our results show that the U-Net model achieved an Mask IoU of 0.903, while Detectron2 demonstrated superior performance with an Mask IoU of 0.974. This notable improvement is particularly evident in cases where the liver appears segmented into two distinct regions within the CT images. Detectron2's architecture, adept at handling complex and fragmented anatomical structures, proved more effective than U-Net in these scenarios. This capability significantly enhances the accuracy of liver boundary detection, which is crucial for precise diagnostic assessments and effective treatment planning. By making our dataset and findings available to the broader research community, we aim to encourage the utilization of this valuable resource to further explore and refine neural network models for medical imaging. This initiative is designed not only to improve liver margin detection but also to facilitate the development of automated diagnostic systems that can be extended to other abdominal organs. Future research should focus on leveraging the strengths of Detectron2 to enhance segmentation techniques for other complex organ structures in medical imaging. This could potentially transform the landscape of computational diagnostics in healthcare, leading to more accurate, timely, and personalized medical treatments.

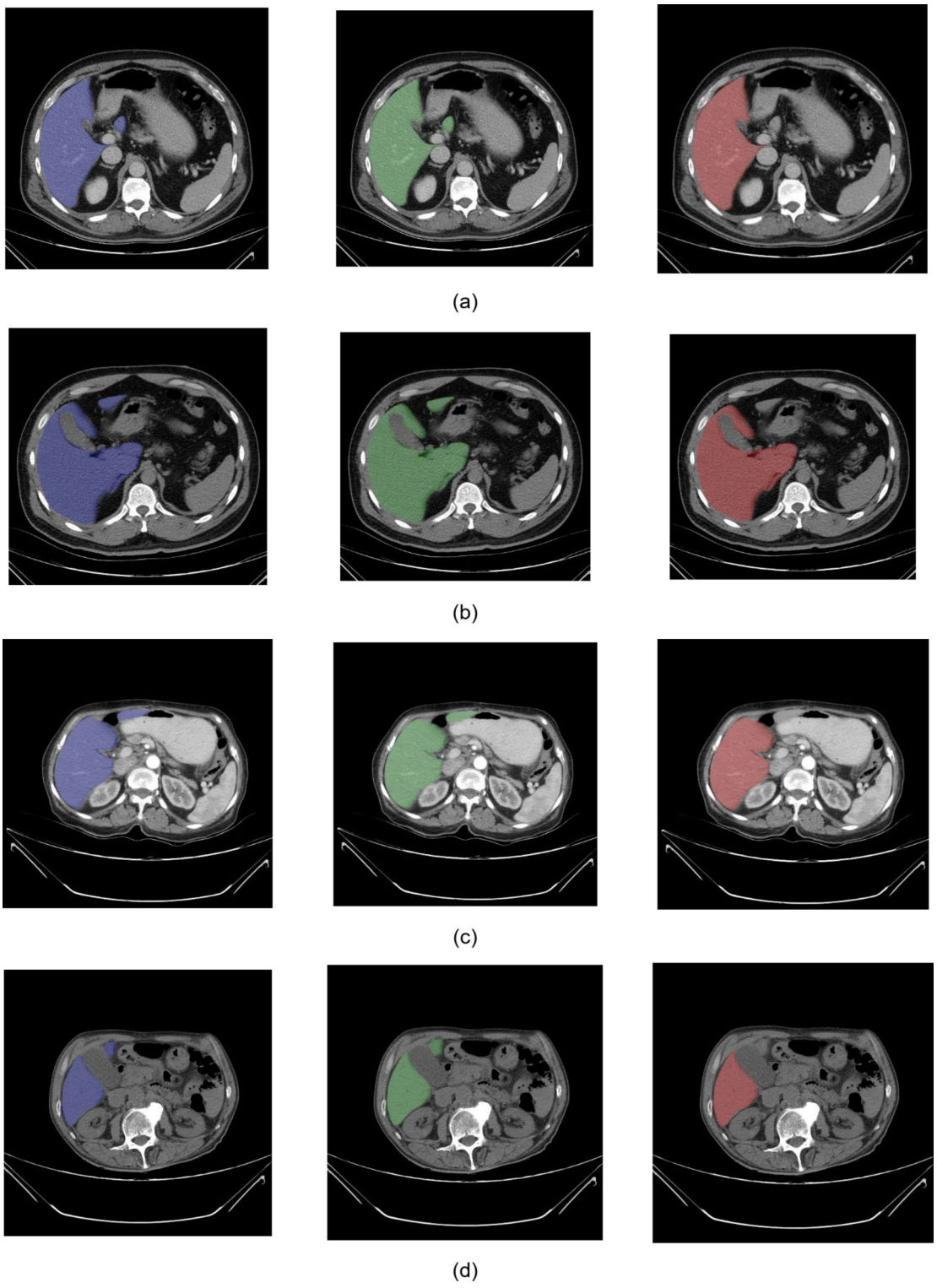
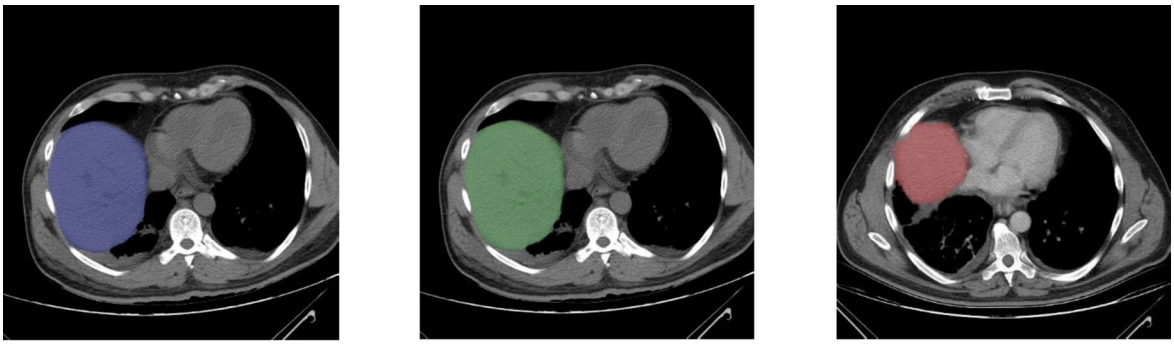


Fig. 5. Segmentation results for five challenging liver cases: (**a**) to (**e**) represent cases 1 to 5, respectively, with radiologist's mask (blue), Detectron2 prediction (green), and U-Net prediction (red), demonstrating both models' strong performance in complex scenarios.



(e)

Figure 5. (continued)

Data availability

The datasets generated during the study are available from the corresponding author on reasonable request.

Received: 2 November 2024; Accepted: 27 February 2025
Published online: 13 March 2025

References

1. Rahman, H. et al. A deep learning approach for liver and tumor segmentation in CT images using resunet. *Bioengineering* **9** (8), 368 (2022).
2. Sedghi, M., Zolfaghari, M., Mohseni, A. & Nosratian-Ahour, J. Real-time transient stability Estimation of power system considering nonlinear limiters of excitation system using deep machine learning: an actual case study in Iran. *Eng. Appl. Artif. Intell.* **127**, 107254 (2024).
3. Thundiyil, S., Shalamzari, S. S., Picone, J. & McKenzie, S. Transformers for Modeling Long-Term Dependencies in Time Series Data: A Review. In 2023 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1–5). IEEE. (2023), December.
4. Mohseni, A., Doroudi, A. & Karrari, M. Data mining based excitation system parameters Estimation considering DCS and PMU measurements using Cubature Kalman filter. *Int. Trans. Electr. Energy Syst.* **30**(11), e12619, 1–15. (2020).
5. Li, W., Jia, F. & Hu, Q. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. *J. Comput. Commun.* **3** (11), 146–151 (2015).
6. Almotairi, S., Kareem, G., Aouf, M., Almutairi, B. & Salem, M. A. M. Liver tumor segmentation in CT scans using modified SegNet. *Sensors* **20** (5), 1516 (2020).
7. Park, S. et al. Image quality in liver CT: low-dose deep learning vs standard-dose model-based iterative reconstructions. Eur. Radiol. **32**, 2865–2874 (2022).
8. Cao, L. et al. A study of using a deep learning image reconstruction to improve the image quality of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions. *The British journal of radiology*, *94*(1118), p.20201086. (2021).
9. Yasaka, K., Akai, H., Abe, O. & Kiryu, S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. *Radiology* **286** (3), 887–896 (2018).
10. Liu, Z. et al. Liver CT sequence segmentation based with improved U-Net and graph cut. *Expert Syst. Appl.* **126**, 54–63 (2019).
11. Kourounis, G., Elmahmudi, A. A., Thomson, B., Nandi, R., Tingle, S. J., Glover, E.K., … Wilson, C. (2024). Deep learning for automated boundary detection and segmentation in organ donation photography. Innovative Surgical Sciences, (0).
12. S. S., K., R. S., V.K. Literature survey on deep learning methods for liver segmentation from CT images: a comprehensive review. *Multimed Tools Appl* **83**, 71833–71862 (2024).
13. Singh, G., & Mittal, A. Various image enhancement techniques-a critical review. International Journal of Innovation and Scientific Research, **10**(2), 267–274 (2014)
14. Gul, S. et al. Deep learning techniques for liver and liver tumor segmentation: A review. *Computers in Biology and Medicine*, **147**, p.105620. (2022).
15. Ananda, S., Jain, R. K., Li, Y., Iwamoto, Y., Han, X. H., Kanasaki, S., … Chen, Y.W. (2023). A Boundary-Enhanced Liver Segmentation Network for Multi-Phase CT Images with Unsupervised Domain Adaptation. Bioengineering, **10**(8), 899.
16. Mahadevan, V. (2020). Anatomy of the liver. Surgery (Oxford), **38**(8), 427–431.
17. Kuang, H., Yang, X., Li, H., Wei, J. & Zhang, L. *Adaptive Multi-Phase Liver Tumor Segmentation With Multi-Scale Supervision* (IEEE Signal Processing Letters, 2024).
18. Chen, L. et al. Adoption value of support vector machine algorithm-based computed tomography imaging in the diagnosis of secondary pulmonary fungal infections in patients with malignant hematological disorders. *Open Life Sciences*, **18**(1), p.20220765. (2023).
19. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18 (pp. 234–241). Springer International Publishing. (2015).
20. Özcan, F., Uçan, O. N., Karaçam, S. & Tunçman, D. Fully automatic liver and tumor segmentation from CT image using an AIM-Unet. *Bioengineering* **10** (2), 215 (2023).
21. Pham, V., Pham, C. & Dang, T. Road damage detection and classification with detectron2 and faster r-cnn. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 5592–5601). IEEE. (2020), December.
22. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961–2969). (2017).
23. Lin, T. Y. et al. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117–2125). (2017).
24. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39** (6), 1137–1149 (2016).

25. Krizhevsky, A., Sutskever, I., & Hinton, G. E. . ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90. (2017)
26. Azad, R. et al. Medical image segmentation review: The success of u-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. (2024).
27. CV, A. Deep Learning-Based Instance Segmentation of Aircraft in Aerial Images using Detectron2. (2023).
28. Saifullah, S. & Dreżewski, R. Modified histogram equalization for improved CNN medical image segmentation. *Procedia Comput. Sci.* **225**, 3021–3030 (2023).
29. Gajera, B. et al. CT-scan denoising using a Charbonnier loss generative adversarial network. *IEEE Access.* **9**, 84093–84109 (2021).
30. Wang, S. et al. Iterative label denoising network: segmenting male pelvic organs in CT from 3D bounding box annotations. *IEEE Trans. Biomed. Eng.* **67** (10), 2710–2720 (2020).
31. Shalamzari, S. S., Bagritsevich, M., Melles, A., Obeid, I., Picone, J., Connolly,D., … Wu, H. (2023, December). Big Data Resources for Digital Pathology. In 2023 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1–25). IEEE.
32. Bilic, P., Christ, P., Li, H. B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., … Menze,B. (2023). The liver tumor segmentation benchmark (lits). Medical Image Analysis, 84, 102680.
33. Thai, B., McNicholas, S., Shalamzari, S. S., Meng, P. & Picone, J. Towards a More Extensible Machine Learning Demonstration Tool. In 2023 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1–5). IEEE. (2023), December.
34. Bai, P., Udupa, J. K., Tong, Y., Xie, S. & Torigian, D. A. Body region localization in whole-body low-dose CT images of PET/CT scans using virtual landmarks. *Med. Phys.* **46** (3), 1286–1299 (2019).
35. Wang, S., Li, C., Wang, R., Liu, Z., Wang, M., Tan, H., … Zheng, H. (2021). Annotation-efficient deep learning for automatic medical image segmentation. Nature communications, 12(1),5915.
36. Sridhar, K., Lai, W. C. & Kavin, B. P. Detection of liver tumour using deep learning based segmentation with Coot extreme learning model. *Biomedicines* **11** (3), 800 (2023).
37. Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D. I., Wang, G., … Vercauteren,T. (2018). NiftyNet: a deep-learning platform for medical imaging. Computer methods and programs in biomedicine, 158, 113–122.
38. Dou, Q., Coelho de Castro, D., Kamnitsas, K. & Glocker, B. Domain generalization via model-agnostic learning of semantic features. Adv. Neural. Inf. Process. Syst. 32, (2019)
39. Hong, J., Yu, S. C. H. & Chen, W. Unsupervised domain adaptation for cross-modality liver segmentation via joint adversarial learning and self-learning. *Appl. Soft Comput.* **121**, 108729 (2022).
40. Langkilde, F., Masaba, P., Edenbrandt, L., Gren, M., Halil, A., Hellström, M., … Jäderling,F. (2024). Manual prostate MRI segmentation by readers with different experience:a study of the learning progress. European Radiology, 1–9.
41. Niño, S. B., Bernardino, J. & Domingues, I. Algorithms for liver segmentation in computed tomography scans: A historical perspective. *Sensors* **24** (6), 1752 (2024).
42. Ansari, M. Y., Abdalla, A., Ansari, M. Y., Ansari, M. I., Malluhi, B., Mohanty, S.,… Dakua, S. P. (2022). Practical utility of liver segmentation methods in clinical surgeries and interventions. BMC medical imaging, 22(1), 97.
43. Ansari, M. Y., Yang, Y., Balakrishnan, S., Abinahed, J., Al-Ansari, A., Warfa, M.,… Dakua, S. P. (2022). A lightweight neural network with multiscale feature enhancement for liver CT segmentation. Scientific reports, 12(1), 14153.
44. Sheela, K. S., Justus, V., Asaad, R. R., & Kumar, R. L. (2025). Enhancing liver tumor segmentation with UNet-ResNet: Leveraging ResNet's power. *Technology and Health Care* **24**(1), 1–15.
45. Mohseni, A., Doroudi, A. & Karrari, M. Excitation system parameters Estimation by different input signals and various Kalman filters. *Electromechanical Energy Convers. Syst.* **2** (2), 40–51 (2022).
46. Brianna, D. F., Kesuma, L. I., Geovani, D., & Sari, P. (2025). Combination of Image Enhancement and Double U-Net Architecture for Liver Segmentation in CT-Scan Images. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 7(1), 208–219.
47. Cheng, B., Girshick, R., Dollár, P., Berg, A. C. & Kirillov, A. Boundary IoU: Improving object-centric image segmentation evaluation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 15334–15342). (2021).
48. Reinke, A., Tizabi, M. D., Baumgartner, M., Eisenmann, M., Heckmann-Nötzel, D., Kavur,A. E., … Maier-Hein, L. (2024). Understanding metric-related pitfalls in image analysis validation. *Nature methods* **21**(2), 182–194.

Author contributions

M.A.S. and S. A. Z.: Software, Validation, Formal analysis, A.S. and S.I. and A.R.R. : Methodology, Conceptualization, M.H. and M.S.: Supervision, and Z.GH.: Validation and M.E.: Project administration.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

This research was conducted in full compliance with the ethical standards of both the institutional and national research committees. Ethics approval was obtained from the Research Ethics Board at Taleghani Hospital, Kermanshah, Iran, prior to the commencement of the study. The approval, identified by the code IR.KUMS.REC.1402.586, encompassed all aspects of data collection, handling, analysis, and publication to ensure adherence to ethical guidelines for medical research. All data were collected with informed consent from patients, ensuring transparency and voluntary participation. To safeguard patient privacy, all personally identifiable information was removed from the dataset prior to analysis. Additionally, procedures involving the use of CT images strictly followed regulations to protect patient confidentiality.

Statement confirming

This study was conducted in full compliance with ethical standards, with approval from the relevant institutional ethics committee. Informed consent was obtained from all participants or their legal guardians, who expressed their complete satisfaction with the anonymous use of their CT images. Each participant generously donated their data to support advancements in medical research and the promotion of scientific knowledge. By contributing their anonymized images, they have helped to further research efforts aimed at improving diagnostics and medical treatments, ensuring both scientific progress and respect for their privacy.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92423-9.

Correspondence and requests for materials should be addressed to M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.