Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An Efficient Deep Learning Approach for Liver Segmentation in Medical Imaging

Muskan Srivastava[1*], Gaurav Pandey[2]
1M.Tech. Scholar, Department of Computer Science & Information Technology, SHUATS, Uttar Pradesh, India
2 Assistant Professor, Department of Computer Science & Information Technology, SHUATS, Uttar Pradesh, India
Emails: 23mtcst003@shuats.edu.in1, gaurav.pandey@shuats.edu.in2

ARTICLE INFO

Received: 26 Dec 2024

Revised: 14 Feb 2025

Accepted: 22 Feb 2025

ABSTRACT

Introduction: Precise segmentation of the liver from computed tomography (CT) scans is an essential component in the diagnosis and treatment of liver diseases such as hepatocellular carcinoma. Manual segmentation can be time-consuming and subject to variability, spurring the application of automated systems based on deep learning. We introduce a 3D U-Net-based convolutional neural network for end-to-end segmentation of the liver from volumetric CT images. The model is trained and tested on the LiTS dataset using a standard preprocessing pipeline and an in-house composite loss function—AdvancedComboLoss—combining Dice Loss, Binary Cross-Entropy, and Focal Tversky Loss to balance class imbalance and boundary accuracy. The network handles entire 3D volumes with spatial context retained between anatomical planes and is trained on a GPU-enabled setup with the PyTorch environment.

Objectives: To develop a highly accurate and computationally efficient **3D U-Net-based deep learning model** for automated liver segmentation from volumetric CT scans, addressing challenges like class imbalance, anatomical variability, and low tissue contrast, using a custom loss function (AdvancedComboLoss) and a standardized preprocessing pipeline.

Methods: The methodology adopted is volumetric liver segmentation from a specially designed 3D U-Net architecture specific to medical CT scans. The process entails preprocessing of raw CT scans into a common shape and intensity range and training a fully convolutional 3D neural network that learns the spatial relationships among all three anatomical planes. The model uses an encoder-decoder architecture with skip connections to maintain fine-grained structural information, allowing for accurate delineation of liver regions. A hybrid loss function—integrating Dice, Binary Cross-Entropy, and Focal Tversky losses—is used to direct the learning process to optimize overlap of regions, accuracy of classification, and class imbalance.

Results: Experimental outcomes demonstrate that the model has a **Dice score of 0.9604 and voxel-wise accuracy of 0.9979**, indicating excellent performance in segmentation as well as strong agreement with expert annotations. The model presented in this paper presents a solid, scalable, and clinically valuable approach to automated liver segmentation with future potential to integrate into multi-stage liver and tumor analysis workflows.

Conclusions: Herein, we built and tested a 3D U-Net deep learning architecture for unsupervised liver segmentation from abdominal CT volumes. Through utilization of volumetric convolutions as well as through a well-engineered composite loss function, the model was successful in capturing complicated liver anatomy as well as being able to hold high performance even on varied CT volumes. The power of the method is not just in its quantitative performance but also in its capability to generate anatomically consistent and visually plausible segmentations, making it extremely relevant to actual clinical workflows. Although there are some limitations—constrained batch size and absence of data augmentation—there is room for improvement, the findings affirm that deep learning can be an effective tool for automating key tasks in medical image analysis. As a whole, this paper adds a robust and effective segmentation model that would be potentially embedded in wider diagnostic workflows, ending up improving the accuracy, speed, and scalability of liver evaluation in clinical practice. In brief, this paper adds a high-fidelity, scalable, and computationally effective solution for computer-aided liver

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

segmentation, pushing the application of artificial intelligence in clinical decision-making and medical image analysis.

Keywords: 3D U-Net, Liver Segmentation, Medical Image Analysis, Computed Tomography (CT), Deep Learning, Volumetric Segmentation, LiTS Dataset, AdvancedComboLoss, PyTorch

INTRODUCTION

Liver illnesses, and specifically hepatocellular carcinoma (HCC), are among the highest cancer mortality causes globally. Proper and timely segmentation of liver tumors from computed tomography (CT) scans is essential in efficient diagnosis, therapy planning, and monitoring after therapy. Manual delineation of liver tumors by radiologists tends to be lengthy, subjective, and prone to error caused by heterogeneity of tumors, poor contrast between tumor and liver tissues, and anatomical variation. Consequently, deep learning-based automated liver and tumor segmentation has emerged as a key area of study.

Deep learning, particularly convolutional neural networks (CNNs), has shown remarkable ability in medical image analysis by extracting features in hierarchical features directly from raw image data. In the context of liver imaging, CNN-based architectures like U-Net and its variations have proved particularly promising in segmentation. This paper studies the development and optimization of an effective deep learning model for tumor and liver segmentation on the LiTS (Liver Tumor Segmentation) dataset, with respect to enhancing accuracy, robustness, and efficiency in computation. The progress of artificial intelligence (AI) and deep learning approaches has significantly transformed medical imaging, facilitating the automated detection and segmentation of liver tumors at high accuracy. The conventional diagnostic approaches, which are based on subjective interpretation of medical images, are subject to inter-observer variability and labor-intensive manual procedures. Therefore, computational approaches, especially machine learning and deep learning, have become a promising option for improving diagnostic efficacy and efficiency.

A number of studies have examined the use of computational methods to resolve the issues of NAFLD diagnosis. Birjandi et al. [1] considered the application of classification tree models to predict and diagnose NAFLD and determined key associated factors that play a role in disease development and progression. Following up this line of work, Balaguer-Montero et al. [3] used a combinational deep learning method for NAFLD classification on ultrasound images to prove the utility of deep learning to enhance diagnostic accuracy. The creation of large, annotated ultrasound datasets, e.g., those developed by Alshagathrh et al. [4], has played a key role in enabling training and validation of these sophisticated computational models, hence speeding up the translation of research results into the clinic.

Parallel to this is the tremendous improvement in automatic segmentation of liver tumors, an indispensable process in surgery planning and follow-up monitoring. Akash H et al. [2] presented a comparison between decoders for liver and tumor segmentation utilizing the Unet framework, as well as illustrating the effect of architectural design on the performance of segmentation. Song et al. [5] introduced new deep learning models which further improve the accuracy of segmentation. Further, Hettihewa et al. [7] demonstrated the use of U-Net architecture and multi-attention network, respectively, in obtaining accurate liver segmentation from CT scans and proved the potency of these deep neural networks to perceive intricate anatomy. Khoshkhabar et al. [8] used graph convolutional networks in automatically segmenting liver tumors, emphasizing the range of these algorithms for processing challenging medical image information. Based on these developments, the future of single-stage Self-ONN U-Net architectures with Swin Transformer encoders is a hopeful avenue for increasing the efficiency of simultaneous liver and tumor segmentation.

This review seeks to give an overview of the state-of-the-art in computational techniques for liver disease segmentation and diagnosis. Through integration of the results of these varied studies, we hope to bring to the fore the revolutionary potential of these methods to enhance clinical performance and drive medical imaging analysis research. This review will discuss the strengths and limitations of each approach, with implications for future research directions and integration of these technologies into clinical practice.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

RELATED WORK

Emergent developments in deep learning have precipitated the introduction of numerous liver disease diagnostic models and tumor segmentation models. Earlier approaches were based on standard machine learning techniques. For instance, Birjandi et al. [1] used a Classification Tree (CT) model for NAFLD prediction using regular clinical parameters and achieved a 75% success rate in testing data. It gave understandable rules but not image-based diagnostics.

Following the emergence of deep learning, most studies turned their attention to automatic image analysis. Balaguer-Montero et al. [3] introduced a model utilizes a 3D U-Net cascade from the nnU-Net architecture, exhibiting better performance than state-of-the-art transformer models and even radiologist inter-reader agreement, with a Dice Similarity Coefficient (DSC) of 0.760 on tumor-wise external validation and patient-wise detection accuracy of 99.65%. Augmenting this, Alshagathrh et al. [4] also created a vast annotated ultrasound database for NAFLD, which has proven crucial for training and testing AI models.

For the liver and tumor segmentation, Akash et al. [2] experimented on LiTS, 3DIRCADb, and CHAOS dataset and conducted a comparison of performance using different UNet models. The highest DSC Score of 95.45% was achieved by the vanilla U-net model. Likewise, Song et al. [5] designed a better 3D V-Net with pyramidal convolution blocks and Tversky loss, which provided better performance in tumor segmentation on the LiTS dataset.

Others aimed at network architecture improvements. Wesdorp et al. [6] used deep supervision and atrous inception modules in a U-Net structure and further improved with Conditional Random Fields (CRF) and reached a DSC of 97.62% on LiTS. Hettihewa et al. [7] presented MANet, a multi-attention network based on channel and spatial attention, which worked reasonably well for small or irregular tumours.

Graph-based models were also investigated. Khoshkhabar et al. [8] utilized Graph Convolutional Networks (GCNs) with Chebyshev filters and SLIC clustering for solid tumor segmentation under noisy conditions and attained 90% DSC. Hybrid setups like that by Dey and Hong [9], which mixed 2D and 3D CNNs with active contour refinement, yielded Dice scores ranging from 95.6% for liver to 76.1% for tumors on LiTS and IRCAD datasets.

These works demonstrate the continued efforts towards enhancing segmentation accuracy, lowering computational complexity, and improving model generalization across imaging conditions.

Motivation and Novelty

Liver cancer is the fourth most frequent type of cancer in the world. It is the most frequent type of cancer in a number of nations in these regions. Every year, an astonishing figure of more than 800,000 people worldwide are diagnosed with this type of cancer, which eventually claims the lives of more than 700,000 people annually. Segmentation of liver and tumor areas from computed tomography (CT) images is a long-standing difficult task in medical imaging because tumors are small in size, irregular in shape, and have high inter-patient variability, with low tissue contrast. Although deep learning has considerably pushed the state-of-the-art, most approaches still find it difficult to strike a balance between computational efficiency, spatial accuracy, and robustness—particularly when generalizing over volumetric CT scans. Traditional 2D slice-based models like U-Net provide strong baseline performance but fail to utilize the full 3D spatial context across slices, potentially limiting segmentation consistency and accuracy in a clinical setting.

Liver tumor segmentation is a challenging but essential problem in medical image processing because of the intricate nature of hepatic tissue, the low contrast between the liver and neighboring organs, and the wide variability in the shape, size, and intensity of the tumor. Classical image processing or machine learning-based segmentation approaches tend to fail to generalize across a wide range of clinical cases. In addition, most deep learning methods are highly dependent on 2D slice-based models, which break spatial continuity and context between neighboring slices — a critical feature in volumetric medical images such as CT scans.

Against the difficulties of correct liver segmentation of volumetric CT scans, this study formulates a dedicated 3D U-Net architecture for end-to-end liver segmentation. The model architecture follows a symmetric encoder-decoder

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

setup using 3D convolutional blocks coupled with batch normalization, LeakyReLU activation, and dropout regularization to effectively utilize spatial information without overfitting.

The network supports four levels of downsampling and upsampling pathways to facilitate multi-scale feature reconstruction and extraction. Each encoder block down-scales the spatial resolution but up-grades the depth of features, and each decoder block recovers spatial resolution in the form of transposed convolutions with skip connections from similar encoder levels. The skip connections assist in the preservation of subtle anatomical information important for accurate liver boundary contouring. For the purpose of improving learning stability and segmentation precision, a compound loss function—referred to as AdvancedComboLoss—is proposed. It incorporates Dice Loss, Binary Cross-Entropy Loss, and Focal Tversky Loss, encouraging both region overlap and class imbalance penalty. Supervised training of the model is conducted on 3D CT volumes and respective binary liver masks, with preprocessing steps for data maintaining uniform target volume shapes and intensity normalization.

Though this model is currently an independent liver segmentation network and not yet part of a cascaded framework, its design is modular and particularly well-suited to function as stage one in a cascaded liver–tumor segmentation pipeline. Future extensions of this model can use the estimated liver masks as ROI inputs for a follow-up tumor segmentation model to localize the tumor more accurately and avoid false positives by limiting the search space to hepatic tissue only.

Key Novel Contributions of This Work:

- **Volumetric Segmentation Using a 3D Architecture:**

  This work adopts a fully 3D U-Net architecture which is intended to process whole volumetric CT data instead of 2D slices. This ensures that the model preserves spatial continuity across axial, coronal, and sagittal planes, producing more anatomically consistent segmentations.

- **Custom 3D Encoder-Decoder for Liver Segmentation:**

  A well-designed encoder-decoder network using 3D convolutional blocks is constructed from scratch to learn volumetrically rich features. The encoder learns hierarchical spatial features, and the decoder synthesizes fine-grained liver boundaries via upsampling and skip connections.

- **AdvancedComboLoss for Stable and Balanced Learning:**

  A tailored composite loss function that incorporates Dice Loss, Binary Cross-Entropy, and Focal Tversky Loss is used. This hybrid loss architecture not only addresses class imbalance but also maintains precise overlap and penalizes false positives/negatives in boundary predictions.

- **Lightweight Yet Effective Data Augmentation:**

  Though being based on a minimal augmentation policy—primarily resampling and intensity normalization—the model exhibits high generalization in learning the inherent anatomical features from volumetric CT data efficiently.

- **Modular and Extensible Framework:**

  The whole model, ranging from network blocks to loss functions and data processing routines, is modularized in PyTorch to make integration into a future cascaded pipeline smooth. This architecture lends itself to scaling to dual-stage architectures (e.g., liver and then tumor segmentation) and general medical segmentation tasks.

The paper is organized as follows: Section 2 elaborates on the objective of the entire study. Section 3 outlines the dataset, preprocessing, model structure, and training pipeline. Section 4 outlines the experimental results, performance measurement, and visual outputs. Section 5 concludes the research with key findings and future work directions.

OBJECTIVES

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The primary goal of this research is to contribute to the area of medical image segmentation by proposing a highly accurate, robust, and scalable deep learning approach specifically designed for automated liver segmentation from volumetric CT scans. Acknowledging the indispensable role of liver segmentation in clinical applications like diagnosis, treatment planning, and surgical guidance—especially for diseases such as hepatocellular carcinoma—this research aims at addressing ongoing hurdles such as low contrast between liver and surrounding tissues, non-uniform anatomical boundaries, and large class imbalance in volumetric data.

For these purposes, the paper presents a specialized 3D U-Net model which can work directly with volumetric inputs to retain complete spatial context throughout all anatomical planes. In contrast to typical 2D models processing slices independently, this volumetric design allows the model to learn more nuanced and anatomically consistent representations which are paramount in order to obtain clinically sound segmentation.

In addition, the goal is supported by the incorporation of a new, hand-tailored loss function—AdvancedComboLoss—combining Dice Loss, Binary Cross-Entropy, and Focal Tversky Loss. This combined formulation is specifically designed to balance recall and precision, punish misclassified voxels, and maximize segmentation of complex and uncertain boundaries.

METHODS

This part explains the systematic approach taken for building a deep learning-based framework for liver segmentation in volumetric CT scans. The pipeline has four core components: large-scale data preprocessing, a thoughtfully designed 3D U-Net-based model architecture, an advanced composite loss formulation, and a systematic training strategy.

The methodology adopted is volumetric liver segmentation from a specially designed 3D U-Net architecture specific to medical CT scans. The process entails preprocessing of raw CT scans into a common shape and intensity range and training a fully convolutional 3D neural network that learns the spatial relationships among all three anatomical planes. The model uses an encoder-decoder architecture with skip connections to maintain fine-grained structural information, allowing for accurate delineation of liver regions. A hybrid loss function—integrating Dice, Binary Cross-Entropy, and Focal Tversky losses—is used to direct the learning process to optimize overlap of regions, accuracy of classification, and class imbalance. The model is end-to-end trained with GPU acceleration, and performance is assessed on validation data after each epoch to choose the best checkpoint for deployment. This approach provides the groundwork for future incorporation into a cascaded segmentation model for liver and segmentation. Each step has been carefully crafted to maximize segmentation performance while being computationally efficient.

Dataset Description

This research employs the Liver Tumor Segmentation (LiTS) dataset [11], a publicly available benchmark for assessing liver and tumor segmentation algorithms derived from CT scans. It is composed of 3D volumetric CT images acquired from various clinical sites, together with expert-annotated segmentation masks. The dataset includes 3D CT scans and corresponding segmentation masks in NIfTI (.nii) format. Each mask has: 0 = background, 1 = liver, 2 = tumor

- **Training Set:** 111 CT volumes with corresponding liver and tumor masks.
- **Testing Set:** 20 CT volumes with associated masks (ground truth hidden for challenge evaluation).
- **Data Format:** NIfTI (.nii) format, enabling 3D spatial analysis.
- **Annotations:** Expert-validated masks marking both liver and tumor regions.

The dataset has been widely used across many studies including Unet [2], H-DenseUNet [14], and MANet [7], making it ideal for benchmarking segmentation models in a reproducible way.

Data Preprocessing

Effective preprocessing is an essential requirement in medical image segmentation, especially when using volumetric CT data. CT scans in their raw form tend to display variability in terms of dimensions, intensity distributions, and

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

spatial resolutions, which can all adversely affect model performance if not handled. To normalize the dataset and facilitate model learning, various preprocessing methods were utilized in this research. CT volumes and corresponding liver masks are preloaded in NIfTI format from the nibabel library. For input size standardization, the volumes are resampled to a target size of 160×160×80 by trilinear interpolation. Intensity values are normalized, and a binary mask for liver voxels is created. Data is then divided into training and validation sets, and preloaded into PyTorch DataLoader containers for batch-wise training. Fig. 1 illustrates all the pre-processing activities done prior to training the model. Fig. 2 illustrates a comparison of before and after preprocessing the data.

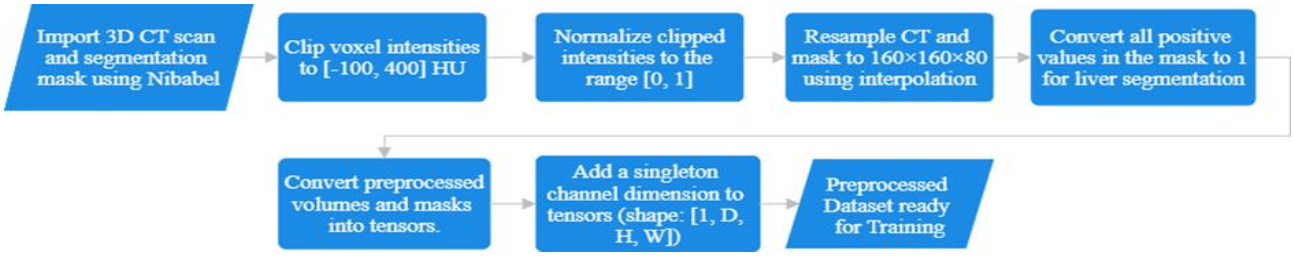

Fig. 1: Shows the steps involved in preprocessing the data to convert it from raw data to processed data.

Volume and Mask Loading

All segmentation masks and volumes of CT scans were saved in the NIfTI format (.nii or.nii.gz). The nibabel library was used to load these with efficiency in processing multidimensional medical image data. The data that were loaded were transformed into NumPy arrays to be processed. This ensured effective extraction of volumetric data and maintaining spatial as well as voxel-wise integrity.

Intensity Clipping and Normalization

CT scan intensity measurements in terms of Hounsfield Units (HU) may extend over a very broad and mostly non-systematic range over patients and acquisitions. To adjust this variability to make it a fixed range applicable clinically, the entire voxel range was clipped from -100 HU to 400 HU. This covers adequately the contrast within soft tissue without accentuating the noise arising in air or bone areas.

After clipping, the min-max normalization was used to normalize the voxel intensities between [0, 1] range. This transformation is specified as:

$$Normalized\ Intensity = \frac{(HU + 100)}{500} \quad (1)$$

This operation ensures that all the CT volumes will have a homogeneously distributed intensity, greatly enhancing convergence when training neural networks.

Spatial Resampling

CT volumes from the dataset had different spatial sizes because of unequal acquisition parameters. For input size standardization and batch-wise training, volumes and masks were all resampled to a common shape of 160 × 160 × 80 voxels. Resampling was done by utilizing the scipy.ndimage.zoom function with varying interpolation strategies:

- **Trilinear interpolation (order = 1)** for continuous CT volumes to maintain intensity gradients.
- **Nearest-neighbor interpolation (order = 0)** for binary segmentation masks to preserve class boundaries without the introduction of interpolation artifacts.

Zoom factors were computed dynamically from the ratio of the target shape to the original volume shape to maintain proportional scaling in all three spatial axes.

Binary Mask Extraction

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The initial mask volumes typically had multiple label values. To facilitate liver segmentation, they were thresholded to binary masks as follows:

$$Liver\ Mask = \begin{cases} 1, & if\ label > 0 \\ 0, & otherwise \end{cases} \quad (2)$$

This provided a uniform target representation for training where the liver region was represented by 1 and the background was represented by 0.

Tensor Conversion and Channel Formatting

Lastly, both preprocessed CT volumes and binary masks were also transferred to PyTorch tensors to be used for model input. A singleton channel dimension was added to meet the input needs of 3D convolutional neural networks. Tensors of the shape **[1, 160, 160, 80]** were then generated for each input-output pair, which can be batched during training and validation.

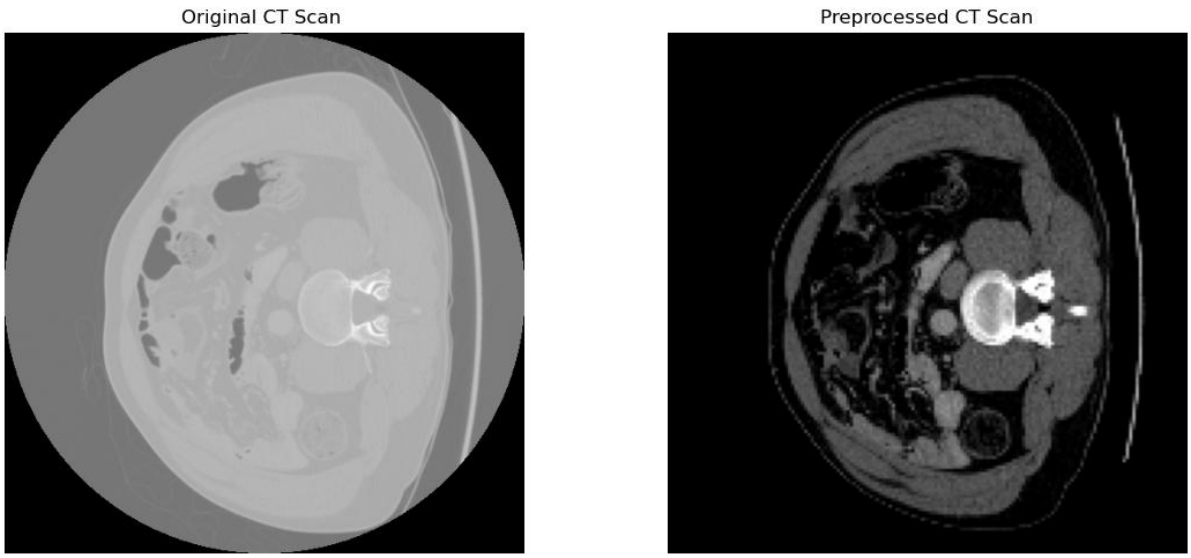

Fig. 2: Visualization of a CT scan before and after preprocessing. The original scan (left) includes wide-range intensity values and scanning artifacts, while the preprocessed scan (right) shows clipped, normalized, and resampled data suitable for training a 3D deep learning model.

3 D U-Net Segmentation Strategy

The central segmentation model is a 3D U-Net, which applies 3D convolutions to the volumetric input and to learn spatial dependencies in all three anatomical planes. The model is an encoder–decoder structure: the encoder learns multi-scale contextual features via consecutive convolutional layers and downsampling, whereas the decoder carries out transposed convolution-based upsampling and fuses encoder features through skip connections. This architecture facilitates the recovery of fine-grained spatial information required for accurate liver boundary definition.

Proposed Network Architecture

The suggested liver segmentation model is a **3D U-Net**, a fully convolutional neural network tailored to volumetric medical imaging tasks. Unlike conventional 2D networks, which work on two-dimensional slices of a 3D volume, the 3D U-Net processes directly the three-dimensional volumes of CT, allowing it to receive extensive spatial context from the axial, coronal, and sagittal planes. The architecture has a symmetric encoder–decoder pattern with skip connections that connect corresponding levels in the network. The encoder path produces hierarchical features via

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

stacked 3D convolutions, batch normalization, LeakyReLU activation, and max-pooling, while the decoder path reconstructs the segmentation map incrementally via transposed convolutions and feature map concatenation from the encoder. The output is a voxel-wise probability map obtained through a 1×1×1 convolution followed by sigmoid activation. This architecture effectively bridges global contextual perception with high-resolution boundary localization, which makes it highly suitable for accurate organ segmentation of high-variation medical volumes.

Workflow diagram of the proposed network is composed of various custom modules.

(a) 3D Convolutional Layers

The heart of the model proposed is **3D convolutional layers**, which generalize conventional 2D convolutions by acting on three spatial dimensions—height, width, and depth. This allows the network to learn volumetric features, taking contextual information from neighboring slices of the CT volume. In clinical images, particularly in abdominal CT scans, organs like the liver, kidneys, and tumors are not well defined in separate 2D slices. But their space presence is evident when observed in the context of three dimensions. The application of 3D convolutions enables the network to take advantage of these inter-slice relationships, thereby generating more anatomically homogeneous and spatially coherent segmentations.

(b) Encoder (Downsampling Path)

The encoder, or the contracting path, is tasked with extracting hierarchical feature representations from the input CT volume. It is a series of convolutional blocks that successively downsample the spatial resolution of the feature maps but enhance their semantic depth. Each block contains two successive 3D convolutional layers followed by **Batch Normalization** and **LeakyReLU** activation. Batch Normalization normalizes the activation output of each layer, speeding up training and encouraging more stable backpropagation of gradients. LeakyReLU is used as the activation to solve the dying neuron problem by allowing a small, non-zero gradient in the inactive unit. Moreover, 3D dropout layers are incorporated following convolutional blocks to randomly disable a percentage of neurons during training, preventing overfitting and pushing the network to learn generalizable, strong features. Every block ends with a 3D **max-pooling operation**, downsampling the feature maps by half in every dimension while decreasing computational costs and enabling the network to concentrate on abstract representations progressively.

(c) Decoder (Upsampling Path)

The decoder or the expanding path is used to recover the spatial features of the feature maps and produce the end segmentation output. Structurally symmetric to the encoder, it strives to recover lost fine spatial detail that occurs as a result of downsampling. Upsampling is done with **3D transposed convolutions** (deconvolutions), which learn to recover the spatial resolution of the feature maps in a data-driven process. At every decoder level, the upsampled feature maps are concatenated with the respective feature maps of the encoder with skip connections. Skip connections are important for recovering the spatial context and maintaining high-resolution information from previous layers. The **two combined features are forwarded through a subsequent set of 3D convolutional layers using batch normalization and LeakyReLU activations**. With this double-path architecture, the decoder gains both global context and fine spatial detail, which improves the output segmentation's boundary definitions to be more accurate.

(d) Skip Connections

One of the distinctive characteristics of the U-Net architecture, and therefore this suggested model, is the implementation of **skip connections** that connect the encoder and decoder branches at corresponding resolution levels. Such connections enable the network to bypass the bottleneck layer and directly carry high-resolution spatial features from the encoder to the decoder. Skip connections are omitted, and skip connections are avoided, which results in the loss of critical localization information during the downsampling phase. This could cause blurred or erroneous segmentation edges. In the context of medical image segmentation, organ boundaries and pathologies might be subtle with strong localization properties, and having skip connections is the guarantee of preserving fidelity of original anatomical details. They also allow the smoother flow of gradients throughout the network, thus aiding convergence in training.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

(e) Output Layer and Activation

The final layer of the network consists of a **1×1×1 3D convolution**, which projects the multi-channel feature map down to a single output channel representing the probability of each voxel belonging to the liver class. This is followed by a **sigmoid activation function**, which converts the raw logits into probabilities in the range [0, 1]. These probabilities can then be thresholded (e.g., at 0.5) to generate a binary segmentation mask. This architecture allows the model to generate a dense, voxel-wise prediction over the whole input volume, which is critical for medical diagnosis and surgical planning tasks.

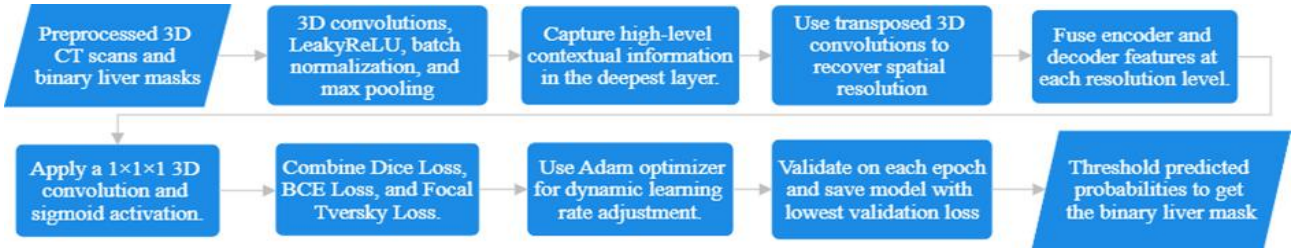

Fig. 3: The entire workflow of the proposed methodology

ENVIRONMENTAL SETUP

The segmentation model is constructed on a system with 11th Gen Intel(R) Core (TM) i5-11320H processor, 16 GB Ram and 64-bit operating system. All segmentation models were implemented using the PyTorch deep learning framework and trained within Kaggle Jupyter Notebook environments.

All experiments in this study were conducted using an interactive Kaggle Notebook environment with GPU acceleration enabled. The segmentation models were developed and executed using Python 3 and implemented with the PyTorch deep learning framework, which offers efficient support for volumetric data processing and GPU-based training. The environment was driven by a Tesla T4 GPU (or similar), enabling quicker training and real-time validation on 3D CT volumes. The hardware setup consisted of 16 GB of RAM and a 64-bit Ubuntu-based operating system delivered via Kaggle's managed environment. GPU acceleration drastically cut down the training time of the 3D U-Net model, which otherwise is computationally expensive because of its high memory requirements from volumetric data and 3D convolution operations.

The model was trained in a Jupyter-based notebook environment hosted on Kaggle, which allowed the use of pre-installed libraries and easy data import through Kaggle Datasets. The pipeline for training was optimized for low-resource usage while ensuring high fidelity of the model through methods like mixed precision training (when possible), optimized data loading through PyTorch DataLoader, and dynamic memory management.

Parameter	**Liver Segmentation**
Learning Rate	1e-4
Optimizer	Adam
Loss Function	AdvancedComboLoss (Dice + BCE + Focal Tversky)
Batch Size	1
Input Dimensions	1×160×160×80 (CT Volume Only)
Input Channels	1
Epochs	120
Normalization	Min–Max Scaling to [0, 1] after HU Clipping [-100, 400]

Table 1. Model parameters for Liver Segmentation

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Methodology and Functionality

3D U-Net Encoder Architecture

The 3D U-Net encoder is crucial to hierarchically learning semantic and spatial information from volumetric medical imaging data like CT or MRI. It is a series of convolution blocks and downsampling processes that successively reduce the spatial size of the input and augment the representational capacity of the feature maps. This path of contraction allows the network to obtain contextual information with bigger receptive fields, which is important for precise segmentation in three-dimensional space. Each stage of encoding consists of a repeated unit called a convolutional block, which is formally defined as:

$$\mathit{Block(Cin,Cout)=Conv3D\rightarrow BatchNorm3D\rightarrow LeakyReLU\rightarrow Dropout3D\rightarrow Conv3D\rightarrow BatchNorm3D\rightarrow LeakyReLU}$$

Here:

- Conv3D layers employ a kernel size of 3×3×3 with padding of 1 voxel in each dimension to maintain spatial resolution.
- BatchNorm3D normalizes feature maps to speed up convergence and offer regularization.
- LeakyReLU is employed as the activation function to prevent dying neurons, especially useful in medical datasets where data imbalance is prevalent.
- Dropout 3D is used between convolutions with dropout probability 0.2 to prevent overfitting while maintaining spatial coherence in the 3D structure.

Every convolutional block is equivalent to a hierarchical feature extractor, where earlier blocks extract low-level edges and textures and deeper blocks extract higher-level features meaningful for organ or lesion localization. The encoder contains four hierarchical levels, where each level consists of a convolutional block followed by a downsampling operation. A summary of what each level and the involved components do is outlined in Table 2.

Encoding Level	Input Channels	Output Channels	Spatial Resolution (relative)	Operation Sequence
Level 1	1	32	1× (full)	Block
Level 2	32	64	1/2×	MaxPool → Block
Level 3	64	128	1/4×	MaxPool → Block
Level 4	128	256	1/8×	MaxPool → Block

Table 2. Detailed functionality of each level of Encoder Blocks.

This progressive and sequential feature extraction during the encoder maintains fine-grained as well as high-level semantic details for transmission to the decoder path, making use of skip connections to fuse encoder outputs and upsampled corresponding decoder features. Skip connections are instrumental to exact localization and are especially critical in applications like tumor boundary separation or organ delineation in healthcare imaging. Fig. 4, illustrates 3d U-Net encoder workflow.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

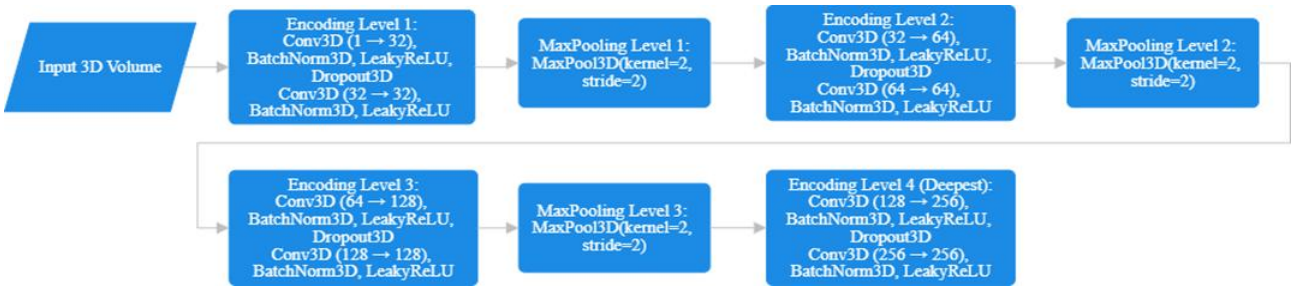

Fig. 4: Workflow of the Encoder Architecture of 3d U-Net Model

3D U-Net Decoder Architecture

The decoder branch of the 3D U-Net aims to reconstruct the segmentation map from the high-level feature representations of the encoder as it progressively retrieves the original spatial resolution. Such an expansive path is done using upsampling and followed by convolution operations, and it plays a major role in reinstating fine-grained anatomical details required in precise medical segmentation.

To restore and reintegrate lost spatial context during downsampling, the decoder uses skip connections from the matching encoder levels. These skip connections concatenate the encoder feature maps with upsampled features of the decoder so that the network can combine semantic information from deeper layers with localization cues from shallow layers. Each level of the decoder involves two main operations:

- Upsampling (Transposed Convolution): ConvTranspose3d increases the spatial size of the feature map by a factor of 2 along each axis (depth, height, and width), essentially undoing the downsampling done in the encoder.
- Decoder Block (Post-Concatenation Convolutions): A similar structure convolutional block to the encoder blocks but one that has been modified to accommodate decoding. Decoder blocks accept concatenated inputs—the upsampled decoder features and their corresponding encoder output.

Decoding Level	Input Channels	Output Channels	Operation	Output Spatial Resolution
Level 3	256 (128+128)	128	Transposed Conv → Concat → Block	1/4×
Level 2	128 (64+64)	64	Transposed Conv → Concat → Block	1/2×
Level 1	64 (32+32)	32	Transposed Conv → Concat → Block	1× (original size)
Final Output	32	1 (or C_out)	1×1×1 Conv3D	1×

Table 3. Detailed functionality of each level of Decoder Blocks.

The decoder in the 3D U-Net model successfully reconstructs the original volume resolution while synthesizing semantic features with spatial context through skip connections. This localization-abstraction balance is crucial to the success of the model in volumetric medical image segmentation tasks.

Role of Skip Connection –

One of the most important features of the U-Net architecture is the use of skip connections that bypass the pass encoder feature maps to decoder layers. This operation offers a remedy to the common issue of spatial loss of information due to down sampling since it maintains edge and boundary information critical for correct segmentation.

Formally, e_i denotes the encoder output at level , i_i and u_i is the upsampled decoder feature at the same level, the skip connection is defined as:

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$c_i = Concat\ (u_i, e_i)$$

Such connections enable easier flow of gradients during backpropagation, enabling faster convergence and better accuracy, especially in segmenting small or complicated structures such as tumors, vessels, or lesions in 3D medical images.

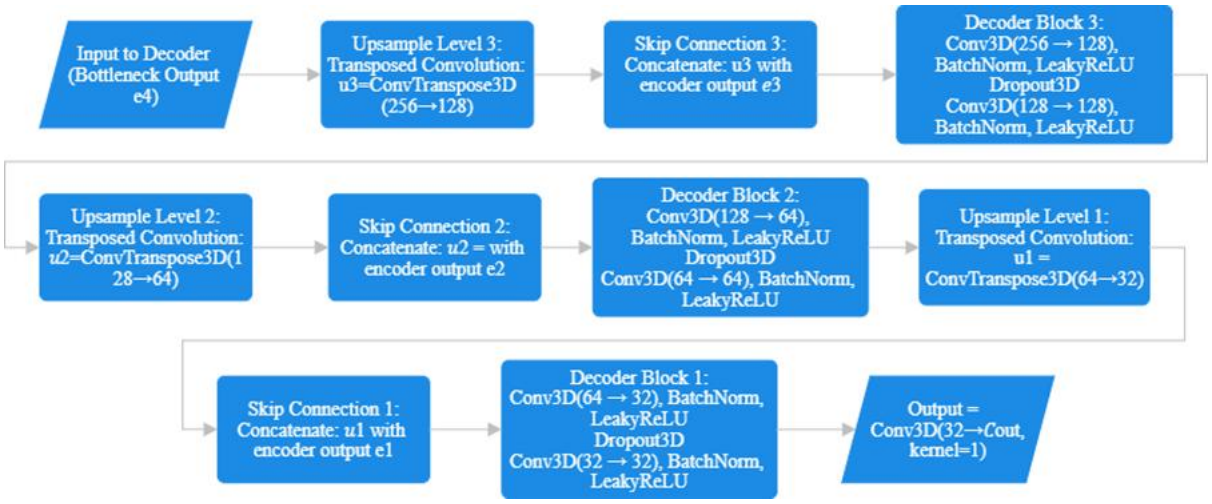

Fig. 5: Workflow of the Decoder Architecture of 3d U-Net Model

Pseudo Code

function train_model(train_loader, val_loader, model, optimizer, loss_fn, device);

Input:

train_loader – PyTorch DataLoader for training samples

val_loader – PyTorch DataLoader for validation samples

model – 3D U-Net initialized with in_channels=1, out_channels=1

optimizer – Adam optimizer (learning rate = 1e-4)

loss_fn – AdvancedComboLoss (Dice + BCE + Focal Tversky)

device – 'cuda' or 'cpu'

Output: Best-performing trained model based on validation Dice score

1. best_dice = 0.0
2. num_epochs = 120
3. for epoch in range(num_epochs) do
4. model.train()
5. train_loss = 0.0
6. for batch in train_loader do
7. volume, mask = batch['image'].to(device), batch['mask'].to(device)
8. optimizer.zero_grad()
9. output = model(volume)
10. loss = loss_fn(output, mask)

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

```
    loss.backward()
    optimizer.step()
    train_loss += loss.item()
  end for
  model.eval()
  val_dice = 0.0
  val_acc = 0.0
  num_val_samples = 0
  with torch.no_grad() do
    for batch in val_loader do
      volume, mask = batch['image'].to(device), batch['mask'].to(device)
      output = model(volume)
      pred = (output > 0.5).float()
      dice = dice_score(pred, mask)
      acc = voxel_accuracy(pred, mask)
      val_dice += dice
      val_acc += acc
      num_val_samples += 1
    end for
  end with
  avg_dice = val_dice / num_val_samples
  avg_acc = val_acc / num_val_samples
  if avg_dice > best_dice then
    best_dice = avg_dice
    Save model as 'best_model.pth'
  end if
  Print epoch, train_loss, avg_dice, avg_acc
end for
return model
end
```

Loss Function

For a precise and solid segmentation of liver areas from 3D volumes of CT data, this paper utilizes a bespoke composite loss function called **AdvancedComboLoss**. Segmenting the liver voxels comes with a great class imbalance scenario—where the liver voxels form a limited portion of the overall volume compared to the background. Moreover, anatomical contours can be blurred or poorly defined. Thus, one may find that a single loss function is not able to capture well the varied needs of overlap accuracy, class balance, and boundary sensitivity. To combat such

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

difficulties, AdvancedComboLoss incorporates three complementary loss components: **Dice Loss, Binary Cross-Entropy (BCE) Loss, and Focal Tversky Loss,** each having its own unique contribution to the learning process.

Dice Loss

Dice Loss is formulated to directly optimize the spatial overlap between the ground truth and the predicted segmentation mask. Dice Loss is especially suitable for medical image segmentation with class imbalance. In mathematics, Dice Loss is derived from the Dice Similarity Coefficient (DSC), and it focuses on areas where the prediction and ground truth are in agreement. By reducing Dice Loss, the model is incentivized to increase the overall overlap, which is of key importance when it comes to delineating organ boundaries in voluminous 3D spaces.

Binary Cross-Entropy (BCE) Loss

BCE Loss treats segmentation as a voxel-wise binary classification task. For each voxel, it computes the cross-entropy between the predicted probability and the ground truth label. BCE provides stable gradient feedback during training, especially in early epochs when predictions are noisy. Even though it does not handle class imbalance inherently, it works well when paired with other spatial losses and provides a probabilistic learning framework.

Focal Tversky Loss

The Focal Tversky Loss generalizes the Tversky Index, a Dice score generalization that adds tunable parameters (α and β) to differently weight false positives and false negatives. This makes it especially useful for cases with severe class imbalance. The "focal" component further emphasizes hard-to-segment voxels by raising the Tversky term to a power (γ), helping the model focus on difficult examples—such as fuzzy liver boundaries or small missed regions.

Formulation of AdvancedComboLoss

The final loss used during training is a weighted combination of the three components:

$$L_{AdvancedCombo} = 0.3 \times L_{Dice} + 0.3 \times L_{BCE} + 0.4 \times L_{Focal\ Tversky} \quad (3)$$

This formulation ensures a balance between:

- Global spatial overlap (Dice Loss),
- Local voxel-wise accuracy (BCE),
- And sensitivity to class imbalance and hard examples (Focal Tversky).

Evaluation Criteria

In order to thoroughly evaluate the performance of the suggested 3D U-Net-based liver segmentation model, two metrics were used: the **Dice Similarity Coefficient** (DSC) and **voxel-wise accuracy**. The metrics were used because they are applicable in medical image segmentation, especially in cases of class imbalance and anatomical boundary accuracy.

Dice Similarity Coefficient (DSC)

The Dice Similarity Coefficient, or the F1 score for segmentation problems, is used as the main evaluation metric. Dice measures the spatial overlap between the forecasted segmentation mask and ground truth annotation. Dice works especially well in medical image segmentation problems, where the foreground class (e.g., liver tissue) normally covers a small percentage of the entire image volume.

Mathematically, the Dice score can be expressed as:

$$Dice = \frac{2 \times |P \cap G|}{|P| + |G|} \quad (4)$$

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

where P is the estimated mask and G is the ground truth mask. A Dice value of 1 represents perfect overlap, whereas a value of 0 means zero overlap. Here, in the present research work, the Dice value is calculated for every sample within the validation set and taken average across the dataset to provide an epoch level performance measure. The performance of the model is tracked and checkpointed according to the best validation Dice score so that the best-performing model is preserved.

Voxel-wise Accuracy

As a secondary measure, voxel-wise accuracy offers a broad indication of the number of correctly classified voxels within the CT volume. Voxel-wise accuracy incorporates both true negatives and true positives and serves as a helpful indicator of general model correctness in preliminary training periods.

While less sensitive than Dice score to boundary accuracy and class imbalance, accuracy serves to complement Dice in providing a view of model action over the volume as a whole. Accuracy is calculated as:

$$Accuracy = \frac{Number\ of\ Correct\ Voxels}{Total\ Number\ of\ Voxels} \quad (5)$$

In this implementation, both Dice and accuracy scores are calculated at the end of each training epoch using a dedicated evaluate () function. These metrics are printed per epoch and used to track training progress, validation performance, and model convergence.

RESULTS

Overview of Experimental Setup

All experiments within this research were performed in a **cloud-based Kaggle Notebook** environment with GPU acceleration turned on. The hardware setup was an **NVIDIA Tesla T4 GPU** with 16 GB of GDDR6 memory, augmented by 16 GB of system RAM. This setup provided a scalable and stable platform to train computationally intensive 3D deep learning models on volumetric medical imaging data.

The execution was done in **Python 3** and the **PyTorch** deep learning platform, both of which were pre-installed and GPU-optimized for use within the Kaggle environment. The Jupyter interface provided an interactive development environment through which code, output, and visualizations were integrated seamlessly. The model was learnt on the **Liver Tumor Segmentation (LiTS)** dataset consisting of high-resolution 3D abdominal CT scans along with manually segmented liver segmentation masks in NIfTI format. All scans underwent a standard preprocessing pipeline that included:

- **Clipping voxel intensities** to the Hounsfield Unit range [−100,400]
- **Normalizing intensities** to the range [0,1] using min–max scaling,
- **Resampling the volumes and masks** to a standardized shape of **160 × 160 × 80**,
- **Converting the data into PyTorch tensors** with an added channel dimension to match model input requirements.

For the task of segmentation, a **3D U-Net architecture** was employed. The model is composed of **3D convolutional blocks** in encoder–decoder format, with skip connections to maintain spatial information. Supervised training was performed with a composite loss function (AdvancedComboLoss), which is a combination of **Dice Loss, Binary Cross-Entropy Loss,** and **Focal Tversky Loss** to offer well-balanced learning signals, especially in the context of class imbalance and intricate organ boundaries.

The network was trained for **120 epochs** with the **Adam optimizer** and a learning rate of 1e-4 and batch size of 1. Evaluation was performed at every epoch end using **Dice Similarity Coefficient** and voxel-wise accuracy as the main performance measures. The model checkpoints were saved according to the best validation Dice score to achieve the best segmentation quality.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Qualitative Analysis of Segmentation Results

For visually evaluating the performance of the segmentation of the suggested model, comparative qualitative analysis was performed between input CT scan, expert-annotated ground truth, and model-predicted liver mask. A sample illustration is presented in Fig. 4.

The native CT slice depicts intricate abdominal anatomy; wherein liver boundary differentiation may be difficult due to similarity in intensity with the adjacent organs. Notwithstanding, the model's prediction is outstandingly anatomically accurate. The predicted liver area accurately reflects the global shape, location, and organization of the liver with great fidelity. Specifically, the model maintains internal variability, e.g., vessel voids or non-hepatic areas, not over-segmenting in those regions where liver intensity overlaps with adjacent tissues. The segmentation result is visually highly aligned with ground truth, which indicates that the model has learned to generalize between patients and anatomy variations well. The contours of the liver are regular, continuous, and closely track the boundaries marked by the experts, which is important for the clinical reliability required in downstream operations like surgical planning or volumetric analysis.

This output indicates that the model has reached clinically reliable segmentation performance. The predicted mask not only closely resembles the ground truth both in size and shape, but also shows excellent anatomical realism — a fundamental characteristic in the application of deep learning models in medical diagnosis. The visual contrast in the figure above illustrates that the proposed **3D U-Net model** yields very precise liver boundaries. In short, qualitative assessment supports quantitative measures by displaying the model to generate precise, smooth, and context-sensitive segmentation results that reflect expert annotations and are in correspondence with different presentations of anatomy.

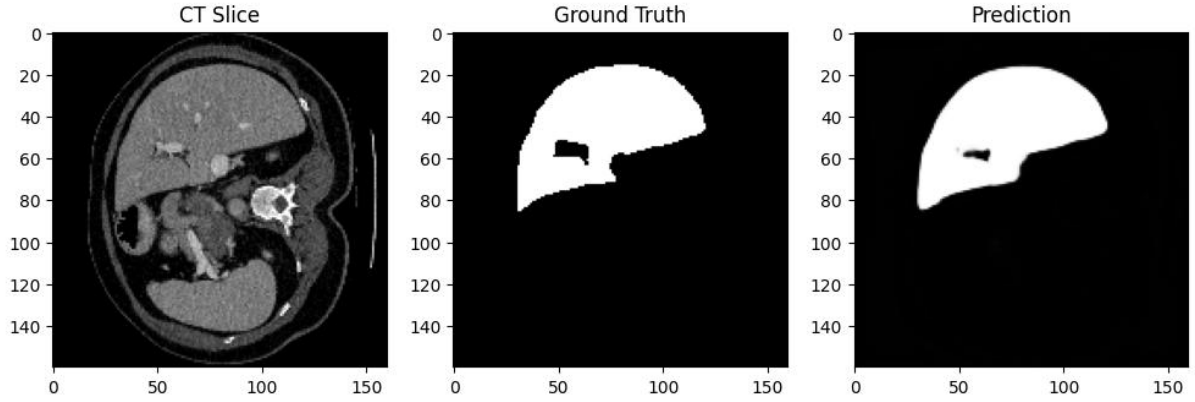

Fig. 4: Qualitative comparison of segmentation results on a representative axial CT slice. The first panel shows the original CT image with liver anatomy visible. The middle panel displays the ground truth liver mask, manually annotated by experts. The final panel shows the predicted segmentation generated by the proposed 3D U-Net model. The predicted mask closely aligns with the annotated region, demonstrating accurate liver boundary detection and strong spatial consistency.

Quantitative Results and Performance Evaluation

The suggested 3D U-Net model was quantitatively assessed with respect to two of the most significant metrics: Dice Similarity Coefficient (DSC) and voxel-wise accuracy, both being commonly applied metrics in medical image segmentation. With a number of training epochs, the model registered a **Dice value of 0.9604**, which meant a spatial overlap of 96.04% between the ground truth liver mask and the predicted liver mask. This top score indicates the model's good capability to define accurate liver boundaries over different anatomical structures. The model also achieved a **voxel-wise accuracy of 0.9979**, which verifies that more than **99.7% of the voxels in the CT volumes were accurately classified**.

These findings illustrate the model's superior segmentation performance, balancing fine boundary detection with overall classification accuracy. The agreement between Dice and accuracy also confirms that the model generalizes

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

well, does not overfit, and is robust across various samples. The union of a high Dice score and close to perfect accuracy implies that the model not only reproduces the overall structure of the liver but is also highly adept at separating subtle anatomical borders in volumetric CT data. The model generalizes robustly throughout the dataset, demonstrating stable predictions throughout heterogeneous patient scans with minimal artifacts or segmentation drift.

Epoch-wise evaluation logs reflected a uniform rise in Dice and accuracy, reflecting uniform model convergence and absence of overfitting. Application of composite loss function and standard preprocessing pipeline also added to the stability and robustness of the model.

Metric	Value
Dice Coefficient	0.9604
Accuracy	0.9979

Table 2. Summary of Quantitative Results of Liver Segmentation

DISCUSSION

This paper proposes a deep learning-based method for volumetric liver segmentation based on a 3D U-Net model, trained and tested on the publicly released LiTS dataset. The model performed very well, achieving a **Dice Similarity Coefficient of 0.9604 and voxel-wise accuracy of 0.9979 following 120 epochs of training**. These outcomes indicate the model's capability to accurately duplicate expert-annotated liver masks and suggest its potential as a reliable tool for automated medical image analysis.

The model's success lies in several fundamental methodological design decisions. To begin with, the implementation of 3D convolutions enables the network to pick up spatial relationships between axial, coronal, and sagittal planes, providing an integrated liver anatomical perception within the volumetric CT data. This integrated spatial cognition facilitates the model to preserve anatomical coherence and reject the discontinuities typically present in 2D slice-by-slice segmentation. In addition, the standardized preprocessing pipeline, such as HU clipping, normalization, and resampling, made all input volumes uniform and model-ready, stabilizing training further.

The loss function specifically designed, AdvancedComboLoss, was of key importance in the segmentation accuracy. With a combination of Dice Loss, Binary Cross-Entropy, and Focal Tversky Loss, the model got well-balanced gradient signals to support overlap accuracy, voxel-wise classification, and class imbalance robustness at the same time. This multi-objective loss formulation perhaps helped the model generate smooth, accurate contours without over-segmentation or missed areas.

In total, the model presented in this study is shown to have a high ability for precise and trustworthy liver segmentation. Its high quantitative score, coupled with uniform visual quality, promises practical utility in many clinical applications, such as surgery planning, radiological reporting, and longitudinal liver tracking. With further testing and improvement, the model may become a base building block for fully automated liver analysis pipelines.

Impact of the 3D U-Net Model

The proposed 3D U-Net-based segmentation model demonstrated exceptional performance in accurately identifying liver regions from abdominal CT scans. With a **Dice Similarity Coefficient of 0.9604** and **voxel-wise accuracy of 0.9979**, the model effectively mimics expert-level annotations, making it a reliable tool for automated liver segmentation.

The effect of using a 3D convolutional model is especially large. By evaluating whole volumetric data rather than 2D slices, the model is better able to sense richer anatomical context and spatial continuity in every dimension. This is especially important for organs like the liver that have shape variability and structural heterogeneity between neighbor slices. Further, the inclusion of a composite loss function (AdvancedComboLoss) enabled the model to harmonize spatial overlap, class imbalance, and voxel-level accuracy—benefiting its resilience across a diverse range of CT volumes. Experimental results affirm the model's promise to maximize medical image analysis efficiency and

consistency and provide an auspicious move toward clinical decision-making in surgeries such as planning, radiotherapy, and monitoring of diseases.

Limitations

Although the model returns a strong outcome, some of its limitations need to be mentioned. To begin with, training was with a batch size of 1, which was unavoidable due to the immense memory requirements of volumetric data and 3D convolutions with the Tesla T4 GPU available for use. While performance was not adversely affected in this instance, a bigger batch size would even stabilize learning and converge faster.

Second, the model was solely trained on the LiTS dataset, which while well- curated, perhaps may not fully cover the range of variability that might be seen with real-world clinical data within multiple institutions, scanners, or populations. This would potentially limit immediate generalizability to outside datasets or different imaging modalities.

Finally, some minor segmentation artifacts like small false positives or boundary abnormalities were present in some slices. Although they were rare and did not contribute significantly to overall metrics, they indicate potential for improvement in training strategy as well as post processing.

Future Work

Building upon the successful results of this study, many possibilities for follow-on work can be envisioned. Firstly, it may be extended to multi-organ segmentation procedures where liver segmentation acts as an initialization step prior to tumor or vessel detection. This would pave the way towards more holistic clinical support systems. Secondly, utilizing state-of-the-art data augmentation mechanisms—i.e., elastic deformation, affine transform, intensity, etc. -- etc.—may also provide a boosting impact on generalizability ability. Investigating semi-supervised learning would also enable the model to take advantage of unlabelled data for additional performance improvements.

One possible direction is adding attention mechanisms or residual connections in order to emphasize the network over important anatomical structures and increase boundary accuracy. Concurrently, post processing procedures such as conditional random fields (CRFs) or morphological filtering might be used to adjust output masks to remove small-scale artifacts.

REFRENCES

[1] Birjandi, M., Ayatollahi, S. M. T., Pourahmad, S., & Safarpour, A. R. (2016). Prediction and Diagnosis of Non-Alcoholic Fatty Liver Disease (NAFLD) and Identification of Its Associated Factors Using the Classification Tree Method. *Iranian Red Crescent Medical Journal*, *18*(8), e32858. https://doi.org/10.5812/ircmj.32858

[2] Halder, A., Sau, A., Majumder, S., Kaplun, D., & Sarkar, R. (2025). An Experimental Study of U-Net Variants on Liver Segmentation from CT Scans. Journal of Intelligent Systems, 34, 20240185. https://doi.org/10.1515/jisys-2024-0185

[3] Balaguer-Montero, M., Morales, A. M., Ligero, M., Zatse, C., Leiva, D., Atlagich, L. M., Staikoglou, N., Viaplana, C., Monreal, C., Mateo, J., Hernando, J., García-Álvarez, A., Salvà, F., Capdevila, J., Elez, E., Dienstmann, R., Garralda, E., & Perez-Lopez, R. (2025). A CT-based deep learning-driven tool for automatic liver tumor detection and delineation in patients with cancer. Cell Reports Medicine, 6(4), 102032. https://doi.org/10.1016/j.xcrm.2025.102032

[4] Alshagathrh, F., Alzubaidi, M., Alswat, K., Aldhebaib, A., Alahmadi, B., Alkubeyyer, M., ... & Househ, M. (2024). Large annotated ultrasound dataset of non-alcoholic fatty liver from Saudi hospitals for analysis and applications. *Data in Brief*, *53*, 111266. https://doi.org/10.1016/j.dib.2024.111266

[5] Song, Z., Wu, H., Chen, W., & Slowik, A. (2024). Improving automatic segmentation of liver tumor images using a deep learning model. *Heliyon*, *10*(7), e28538. https://doi.org/10.1016/j.heliyon.2024.e28538

[6] Wesdorp, N. J., Zeeuw, J. M., Postma, S. C. J., et al. (2023). Deep learning models for automatic tumor segmentation and total tumor volume assessment in patients with colorectal liver metastases. European Radiology Experimental, 7, 75. https://doi.org/10.1186/s41747-023-00383-4

Research Article

[7] Hettihewa, K., Kobchaisawat, T., Tanpowpong, N., & Chalidabhongse, T. H. (2023). MANet: A Multi-Attention Network for Automatic Liver Tumor Segmentation in CT Imaging. *Scientific Reports*, *13*(1), 20119. https://doi.org/10.1038/s41598-023-46580-4

[8] Khoshkhabar, M., Meshgini, S., Afrouzian, R., & Danishvar, S. (2023). Automatic Liver Tumor Segmentation from CT Images Using Graph Convolutional Network. *Sensors*, *23*(17), 7561. https://doi.org/10.3390/s23177561

[9] Dey, R., & Hong, Y. (2020). Hybrid Cascaded Neural Network for Liver Lesion Segmentation. *IEEE International Symposium on Biomedical Imaging (ISBI)*. https://doi.org/10.1109/ISBI45749.2020.9098501

[10] Heker, M., & Greenspan, H. (2020). Joint Liver Lesion Segmentation and Classification via Transfer Learning. *International Journal of Computer Assisted Radiology and Surgery*, *15*, 1697–1707. https://doi.org/10.1007/s11548-020-02195-2

[11] Bilic, P., Christ, P. F., Li, H. B., Vorontsov, E., Ben-Cohen, A., et al. (2022). The Liver Tumor Segmentation Benchmark (LiTS). *Medical Image Analysis*, *84*, 102680. https://doi.org/10.1016/j.media.2022.102680

[12] Jin, Q., Meng, Z., Sun, C., Cui, H., & Su, R. (2020). RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. Frontiers in Bioengineering and Biotechnology, 8, 605132. https://doi.org/10.3389/fbioe.2020.605132

[13] Ozcan, F., Ucan, O. N., Karacam, S., & Tuncman, D. (2023). Fully Automatic Liver and Tumor Segmentation from CT Image Using an AIM-Unet. *Expert Systems with Applications*, *222*, 119596. https://doi.org/10.1016/j.eswa.2023.119596

[14] Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.-W., & Heng, P.-A. (2018). H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. *IEEE Transactions on Medical Imaging*, *37*(12), 2663–2674. https://doi.org/10.1109/TMI.2018.2845918

[15] Zhang, J., Luo, S., Qiang, Y., Tian, Y., Xiao, X., Li, K., & Li, X. (2022). Edge Constraint and Location Mapping for Liver Tumor Segmentation from Nonenhanced Images. *Computers in Biology and Medicine*, *144*, 105357. https://doi.org/10.1016/j.compbiomed.2022.105357

[16] Bi, L., Kim, J., Kumar, A., & Feng, D. (2017). Automatic Liver Lesion Detection Using Cascaded Deep Residual Networks. *Neurocomputing*, *320*, 176–185. https://doi.org/10.1016/j.neucom.2018.09.046