SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I12P105

Volume 12 Issue 12, 44-59, December 2025
© 2025 Seventh Sense Research Group®

Original Article

An Attention-Enhanced U-Net Model for Precise Liver Tumor Segmentation Using Convolutional Block Attention and Efficient Neural Network: Advancing Automated Imaging for Improved Liver Tumor Diagnosis

B. Shashikanth[1], K. Sivani[2]

[1]*Department of Electronics, Communication and Instrumentation Engineering, KITS, Warangal, Kakatiya University, Telangana, India.*
[2]*Department of Electronics, Communication and Instrumentation Engineering, KITS, Warangal, Telangana, India.*

[1]*Corresponding Author : shashiboorla@gmail.com*

Received: 05 October 2025 Revised: 07 November 2025 Accepted: 06 December 2025 Published: 27 December 2025

Abstract *- Liver Tumor Segmentation in CT images is an essential step in the medical imaging process, which is generally difficult due to the variability of tumor shapes, sizes, and ambiguous boundaries with healthy liver tissue. This paper proposes a new model that integrates the U-Net structure and Convolutional Block Attention Module along with EfficientNet-B0 to promote both segmentation performance and computational efficiency. Our proposed model is distinctive because it incorporates CBAM to recalibrate the spatial and channel-wise attention maps adaptively, focusing on the informative tumor regions for improved segmentation performance. EfficientNet-B0, an encoder backbone characterized by the compound scaling method and lightweight structure, is employed to enhance hierarchical feature extraction and computational efficiency. The model achieved a mean Intersection over Union (IoU) of 0.9356 on a public dataset. This result significantly outperforms strong previous methods, such as PSP (0.9113) and MANet (0.6555). Our findings show that the incorporation of attention modules with lightweight encoders is effective for precise liver tumor segmentation, and the proposed method has high potential for clinical applications. This paper paves the way for the potential of innovative and scalable diagnostics in the field of health care.*

Keywords *- Liver Tumor Segmentation, U-Net, CBAM, EfficientNet-B0, Medical Imaging, IoU.*

1. Introduction

The liver is an essential organ that is involved in a variety of biochemical functions, including metabolism, detoxification, and bile formation [1]. Disorders such as liver tumors can seriously affect liver function, which may endanger human health [2]. Hepatocellular Carcinoma (HCC) is one of the most prevalent and fatal liver neoplasms, accounting for a considerable proportion of all cancer-induced deaths globally. Therefore, early diagnosis and accurate localization of liver tumors are critical for the treatment and prognosis of the disease, prompting the use of more advanced detection and localization tools. Imaging techniques, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), are important for diagnosing liver tumors and have been commonly used in daily practice due to their ability to provide complete data regarding the location and spread of the lesion [3]. However, it remains a challenging task to interpret these images and segment the liver tumors [4], and the widely used manual segmentation method is time-consuming, error-prone, and relies on the radiologists' experience. To address this, automated alternatives have been continuously studied.

Despite advances in imaging and radiology, accurate liver tumor segmentation is still a challenging task, as tumors can vary widely in shape, size, position, and contrast across different imaging modalities [5]. In addition, in most images, the tumor margins are not well defined and are hard to contrast with the normal liver [6].

These challenges underscore the importance of reliable and transferable segmentation techniques. Recent developments in the field of medical image analysis, particularly in the provision of automated diagnostic aids, could help reduce the workload of radiologists and improve the accuracy of tumor detection. The traditional image processing methods have been shown to be ineffective in fulfilling the complex procedure of liver tumor segmentation, which provides space for computational-based solutions.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Deep learning, one of the AI's subfields, has become an increasingly powerful method for analysing medical images. As described in the paper, deep learning models, such as Convolutional Neural Networks (CNNs) and their extensions, are able to learn hierarchical features directly from raw data, which compares favourably to traditional methods for segmentation [8]. Deep learning models have demonstrated good potential in learning very accurate representations from large databases, such as in liver tumour segmentation. This methodology is regarded as a comprehensive solution to the problem of liver tumor segmentation, which consequently results in improved diagnostic and therapeutic outcomes.

In the past few years, many techniques have been developed to enhance liver tumor segmentation. Chuanfei Hu et al. [9] proposed an uncertainty-aware framework to conduct multiphase liver tumor segmentation. Through Dempster–Shafer Evidence Theory, this framework captures both segmentation and uncertainty with evidence parameterized by a Dirichlet distribution and directly estimates the reliability of results in multiphasic contrast-enhanced CT images. The framework outperforms state-of-the-art approaches, proving its robustness and allowing reliable performance in disturbed conditions. However, there is still research space on integrating the multi-modality images analysis and developing the model robustness with respect to different clinical environments.

Saumiya S et al. [10] introduced a joint learning-based multi-task model with the aim of automatic liver tumor segmentation and classification. This method introduces a multi-task deformable attention U-Net for segmentation, which uses an attention-based deformable module to capture varied tumor appearances. Finally, a DenseNet model is used for classification, and it improves precision, recall, F1-score, and classification accuracy, compared to single models. However, the adoption of this approach in heterogeneous datasets remains a challenging issue, particularly in clinical applications.

Arifullah et al. [11] presented a two-stage process for histopathological image analysis with Whole Slide Images (WSIs). WSIs are processed by a modified U-Net for segmenting images, and patches are then extracted at multiple magnifications and used as input for a CNN for classification. The approach achieved better segmentation and classification performances by incorporating patch-based frameworks. Nevertheless, its scalability and performance to handle large data sets with diverse tumor types need to be further studied.

Yuanyuan Shui et al. [12] introduced a model that fuses richer hierarchical representations via the MFF module for multi-scale target detection. The edge semantics are further refined by the EI module, which is then finally employed for segmentation feature representations by the EG module. The proposed method outperformed existing state-of-the-art methods, with a Dice index of 85.55% and a Jaccard index of 81.11% on the LiTs2017 dataset. Although this approach achieves high performance, it has not been fully optimized for multiphase CT imaging or real-time clinical implementation.

Yilin You et al. [13] introduced the Parallel Graph Convolutional Network (PGC-Net), using contour-based parallel graph reasoning for fast, convenient liver tumor segmentation. The proposed framework employs a pre-trained Pyramid Vision Transformer for multi-scale feature extraction. It adopts pixel-level parallel graph reasoning, mapping each pixel to a high-dimensional space to facilitate better segmentation. However, challenges of such a method in solving complex tumour structures and its effectiveness in big clinical patient cohorts are key research areas yet to be addressed.

Raju Egala et al. [14] emphasized the capability of Deep Learning, especially CNNs, in improving diagnostic performance in the field of medical image interpretation. Recent improvements in DL methodologies, including new activation functions, optimization methods, and loss functions, have led to a considerable improvement in the abilities of CNNs, and they are a good tool for radiologists. However, the interpretability of deep learning models for medical image analysis and their integration into clinical workflows remains a topic that requires further investigation.

Kai-Ni Wang et al. [15] proposed SBCNet, a dual-branch architecture for liver tumor segmentation. The architecture consists of a tumor variation encoding module and a boundary enhancement module. The method shows comparable results to the state of the art on the LiTS dataset, demonstrating the value of the dual-branch model in segmenting a problem. However, it is not yet clear how this approach can be extended to diverse tumor types and how robust it will be to image artifacts across scanning protocols.

Ali Mohammed Hendi et al. [16] proposed using a hybrid CNN+LSTM model for liver disease diagnosis and prognosis, which consists of CNNs and Long Short-Term Memory (LSTM) networks. This method is far superior to CNN, RNN, and LSTM models, achieving an accuracy of 98.73% and robust metrics (precision, recall, F1-score, and AUC-ROC). However, the universal application of the model to the liver non-contrast imaging modalities and multi-center cohort still potentially exists as a study limitation.

Recent contributions in liver segmentation have increasingly focused on the integration of attention mechanisms and EfficientNet backbones, highlighting improvements in segmentation accuracy and efficiency. One notable work is EAR-U-Net: EfficientNet and attention-based residual U-Net for automatic liver segmentation in CT [17]. This study leverages EfficientNet-B4 as the encoder and integrates attention gates into the skip connections of a residual U-Net. The use of EfficientNet significantly improves feature extraction, while the attention

mechanism enables the model to focus on relevant regions of the liver, resulting in enhanced segmentation accuracy. This approach is highly relevant to our work, particularly when comparing your EfficientNet-B0 variant against a more advanced backbone, such as EfficientNet-B4. The paper [18] introduces a dual attention mechanism (both channel-wise and spatial attention) in a 3D U-Net architecture for liver segmentation. The attention mechanism refines the segmentation process by focusing on the most relevant features in both the spatial and channel dimensions. Although this work does not explicitly utilize EfficientNet, it presents an innovative application of attention that could complement or be incorporated into your attention-enhanced U-Net model.

In 2025, a study by Limprapaipong et al. [19] used a Multi-Scale Attention Network (MA-Net) with cross-branch feature fusion for tumor segmentation. While the main focus was on liver tumor segmentation in a different imaging modality (SPECT/CT), the approach incorporates EfficientNet-B6 as the backbone and utilizes attention for improved segmentation. This work is particularly valuable for comparison, as it explores how EfficientNet combined with attention mechanisms can be applied in a different context, demonstrating the versatility of this approach. These developments in deep learning and medical image analysis are increasingly breaking through the limits of automatic liver tumor segmentation, bringing great prospects for clinical diagnosis and treatment. Yet, several research efforts to date have not focused on multi-modality integration, strong evidence of model robustness through testing on multiple datasets, and real-time clinical application. This research aims to fill in these gaps by designing a hybrid deep learning architecture that combines U-Net and EfficientNet-B0 through the addition of attention mechanisms, such as CBAM, for feature extraction refinement. The overarching goals of this research are to enhance the sensitivity of liver tumor detection by developing a more robust, generalizable, and clinically deployable approach to liver tumor segmentation that increases diagnostic accuracy while addressing the limitations of current research.

Although some new models have made some advances in liver tumor segmentation, it seems to be difficult to achieve a great balance of high accuracy and efficient calculation. Few care about incorporating dual-attention mechanisms or lightweight backbones for the resource-limited scenarios. In this paper, we aim to fill this gap by proposing a new hybrid model that incorporates the Convolutional Block Attention Module (CBAM) into a U-Net architecture, utilizing EfficientNet-B0 as the encoder. This novel combination further improves the precision and speed of the segmentation, and thus can be used for clinical use, particularly in real-time or resource-scarce conditions.

2. Proposed Model

Liver tumour segmentation refers to the process of precisely locating and isolating tumour regions found in medical images, for instance, CT scans. The segmentation in this work is done by the U-Net architecture. Semantic segmentation is the task of classifying each pixel in the input image; U-Net is a Convolutional Neural Network (CNN) architecture built for deep learning. The encoder-decoder structure, skip connections, and fine-grained feature capturing enable the segmentation of tumors from the surrounding normal liver anatomy, making it an excellent U-Net model for medical image segmentation tasks.

2.1. U-Net Architecture

The U-Net architecture used in this project has a standard encoder-decoder architecture, more specifically illustrated in Figure 1. The input image passes through encoder layers for feature extraction; multiple convolutional and pooling layers follow this to consider spatial context and high-level features. After that, the decoder part, which reconstructs the segmentation map, follows, using transpose convolutions for up-sampling and a skip connection that concatenates the corresponding encoder features. These skip connections preserve spatial information lost due to downsampling, which increases the accuracy of segmentation.

In this implementation, a U-Net is constructed via Segmentation Models Pytorch (SMP). Key components include:

- Encoder: EfficientNet-B0 as the backbone network for our feature extraction, and pre-train the model on the ImageNet dataset to leverage transfer learning.
- Decoder: Several layers of upsampling and convolutional blocks help the network reconstruct the output segmentation map progressively.
- Attention Mechanism: Here, a custom attention layer enhances the segmentation map by highlighting key areas and eliminating undesired regions. The attention mechanism integrates into the output of the U-Net, enabling pixel-wise adjustment with a standard Conv2d layer followed by Sigmoid activation.

This is designed for pixel-level binary classification (e.g., cancerous vs. non-cancerous tissue), producing a segmentation map of the same size as the input image.

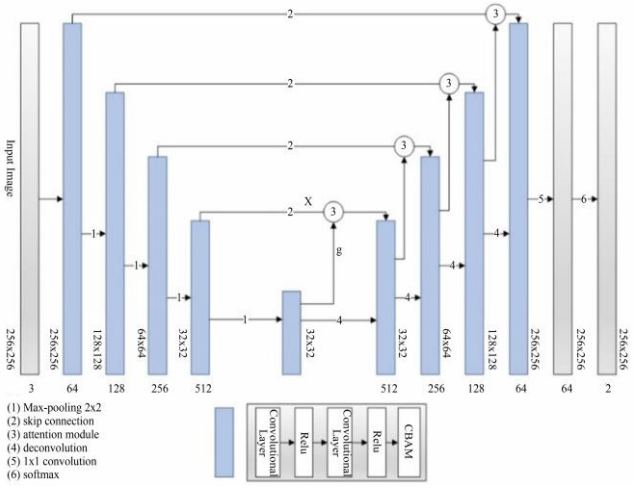


Fig. 1 U-Net Architecture with CBAM Integration

The diagram shows the U-Net-based network architecture modified with Convolutional Block Attention Module (CBAM) for better segmentation. It is an encoder-decoder style of architecture with its components and connections described in detail below:

- Encoder Pathway: The encoder progressively reduces the spatial resolution of the input image while extracting features. The encoder has several levels of convolutional layers with the ReLU activation function, as well as a downsampling layer in the form of max-pooling. With the reduction in spatial size, the number of filters increases with each level to catch more complex features.
- Attention Mechanism: The CBAM is integrated in a way that highlights its significant characteristics. Spatial and channel-wise attention is used so that the network focuses on significant areas and minimizes unnecessary background information. Such training refines the model's attention towards high-resolution structures present in the input image.
- Bottleneck: It connects the encoder and decoder paths through a central bottleneck block. It employs deeper Convolutional operations to detect high-level, abstract features and prepares the Feature maps for the up-sampling stages.
- Decoder Pathway: By performing up-sampling operations and connecting feature maps from the corresponding encoder levels through the skip connection, the decoder recovers the spatial resolution. These residual connections help carry spatial details from the encoder so that outputs, when reconstructing, maintain finer details and boundaries.
- 1x1 Convolution and SoftMax Activation: The last layer uses a 1x1 convolution to convert feature maps into the desired number of output classes (2 in this case). The final class probabilities are computed using a softmax activation function via multi-class segmentation for each pixel.
- Visual Flow: The final structure facilitates the flow of low-level features from the early stages of the encoder and high-level abstract features from the bottleneck. Together with CBAM modules, this gives a very intuitive perspective of the CAPCA base network design, enhancing its ability to accurately segment complex structures as applicable in the case of medical image analysis, etc.

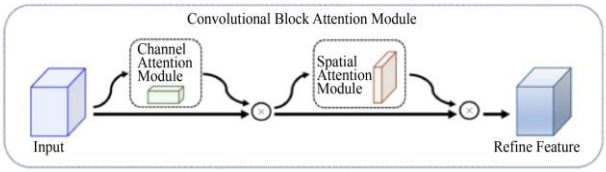


Fig. 2 CBAM Architecture

The Convolutional Block Attention Module (CBAM) refines features through two attention mechanisms, channel attention and spatial attention, in sequence. The channel attention module calculates the correlation between each channel, allowing the relevant feature maps to receive greater prominence.

The high-level refined features are subsequently forwarded to the spatial attention module, which emphasizes more spatially significant areas via the spatial attention map by highlighting decisive zones in the feature map. Both outputs are then element-wise multiplied by input features to yield a refined feature representation. CBAM utilizes attention in a sequential manner that helps the model focus on rare and process features of information, promoting an enhancement in performance while returning minimal characteristics.

2.2. U-Net with CBAM

The implementation proposed a U-Net model with a Convolutional Block Attention Module (CBAM) for improving segmentation accuracy. It is an attention method that gradually applies both spatial and channel attention to boost received feature maps adaptively. CBAM effectively improves feature representation and segmentation output by minimizing irrelevant details while emphasizing the dominant regions. The CBAM is easily integrated into the decoder blocks of the U-Net. First, global pooling is combined with a learnable transformation to assign channel attention to the informative channels of the feature map. Next, spatial attention further refines the spatial focus through an independent convolutional operation. This pair of attention metrics guides the U-Net in concentrating on cancerous areas while retaining delicate details. As demonstrated, CBAMU-Net improved the localization of small and non-uniform tumor regions, which were the most challenging in terms of tumor morphology variability when incorporated within a U-Net. This high segmentation accuracy, validated through Dice and IoU metrics, confirms the model's efficiency for liver tumor segmentation tasks.

2.3. EfficientNetB0 Encoder

In this U-Net implementation, EfficientNet-B0 is utilized as the encoder and acts as a strong feature extraction backbone. It utilizes a compound scaling method to effectively balance the depth, width, and resolution of the network while achieving performance with significantly fewer parameters. EfficientNet-B0, pre-trained on ImageNet, encapsulates the pixels' texture and maintains the high-level information necessary for liver tumors segmentation.

Fig. 3 EfficientNetB0 model

2.3.1. Key Architectural Innovations

In working towards a better balance of accuracy and efficiency in models, EfficientNet-B0 employs a unique

design that compounds the scaling of the model. Key features include:

MBConv Blocks (Mobile Inverted Bottleneck Convolutions)

- These blocks enhance computational efficiency by using depth-wise separable convolutions and bottleneck layers.
- This is one way of recalibration of feature channels, which allows a network to focus on more important features using the Squeeze-and-Excitation (SE) mechanism. In medical imaging, where things like tumor edges are important, this is particularly important.
- Depth-wise separable convolutions eliminate redundant computation while retaining performance, yielding a faster and lighter model.

Compound Scaling

- It scales them jointly and consistently to ensure their features are captured in all levels of detail, and the efficiency of those passed features is kept constant.
- This scaling is particularly important in liver tumor segmentation as it allows for high-resolution CT images to be processed efficiently without the need for extremely expensive computational resources.

Pretrained Weights

- EfficientNet-B0 leverages learned feature representations from models trained on the ImageNet challenge, providing feature representations associated with general textures and patterns, enabling faster training and better convergence on medical datasets, for example, liver CT scans.

3. Visualization and Integration

For illustration, EfficientNet-B0 can be seen as the encoder backbone in a U-Net, where it extracts hierarchical features as its layers. Here is what the integration looks like:

- Low-level features like edges, textures, and structures provide the necessary details to differentiate the tumor tissues.
- To Abstract Features such as shapes and structures to help recognize the unusual tumor borders.
- Find very abstract, high-level features that encode the big picture, for example, the tumor location in the context of liver anatomy.

In the U-Net encoder-decoder framework:

- The encoder uses EfficientNet-B0 in place of traditional convolutional layers.
- Its feature maps are passed to the decoder via skip connections, thus retaining adequate spatial information for accurate segmentation.

4. Performance Metrics

Adopting EfficientNet-B0 as the encoder backbone comes with its own quantifiable benefits with respect to liver tumor segmentation:

- The EfficientNet-B0 captures both low-level and high-level features more effectively than simpler architectures. As a result, this leads to better segmentation metrics, such as the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU).
- EfficientNet-B0 is a CNN-based architecture that is both highly compact and efficient, with an extremely low number of parameters and computations required. This makes it perform well with relatively small datasets and can be run on lower-end hardware.
- EfficientNet-B0 has been shown to outperform similar automatic encoders for medical datasets in segmentation accuracy, e.g., in fine-grained anatomy for techniques that require maximum recall (sensitivity) to detect tiny or malformed regions (e.g., liver tumors).
- Building the EfficientNet-B0 lightweight architecture enables the implementation of the diagnostic at resource-limited locations , such as hospitals with limited computing power, without compromising diagnostic quality.

5. Proposed Method

- Conv2d Layer: A core building block of convolutional neural networks, Conv2d layers apply convolutional filters to input data (e.g., images) to extract meaningful representations. This function means a matrix is slid across the data, and each element is multiplied in a dot product and added up to return an aggregated result. Those filters are learnable parameters that are optimized during training. The Conv2d Layer allows the network to find patterns like edges, textures, or shapes in the data. Being capable of capturing spatial hierarchies, this layer lays the groundwork for image-based deep learning architectures. In this architecture, there are multiple applications for feature refinement at each stage.
- MaxPooling2d Layer: This is a MaxPooling2d Layer, which helps down-sample feature maps to decrease their spatial dimension while retaining the most pronounced features. This slides a pooling window (2x2, for instance) across the input and picks the maximum value from the area covered by this window. This layer further reduces computational complexity and prevents overfitting by introducing a certain level of translation invariance. This allows it to learn faster in deeper layers by focusing on the most dominant features.
- Upsampling2d Layer: The upsampling2d Layer is part of the decoder end of the network to raise the spatial resolution of the feature maps. The segmentation map is reconstructed to the original image size either by finding intermediate pixel values or by filling cutoff pixels. It collaborates closely with the Conv2d layers to refine features and preserve fine-grained details during reconstruction.
- MBConv Blocks Layer: This modified U-Net design uses the MBConv Blocks Layer as the foundational buildup, reinforced from EfficientNet's architecture. These blocks leverage depth-wise separable

convolutions, which help in reducing computational overhead while obtaining high accuracy. One of them is the Squeeze and Excitation (SE) mechanism that readjusts the channel-wise feature responses in a way that the network can focus on the important features. This architecture replaces skip connections with MBConv blocks to enhance feature representation and adaptivity.

- EfficientNet-B0 Layer: The EfficientNet-B0 Layer is used for encoding the backbone of this U-Net implementation. This architecture employs a compound scaling method to evenly scale the depth, width, and resolution of the network, thereby finding the optimal balance between performance and computational cost. It learns hierarchical features, ranging from simple textures to complex patterns, which are crucial for segmenting liver tumor regions in CT scans. Lightweight architecture enables faster processing without compromising accuracy.
- CBAM Layer: CBAM, which stands for Convolutional Block Attention Module, is an attention module added to the architecture. It sequentially applies channel attention and spatial attention, amplifying the emphasis of the feature map on the significant regions of the feature map. The proposed method incorporates channel attention, utilizing informative channels to provide attention, and refines spatial details, thereby enabling the network to achieve better distinction of the cancer area in the segmentation task.
- Sigmoid Layer: The Sigmoid Layer is then used at the last part of the attention mechanism, when the feature map values are normalized between 0 and 1. This normalization enables pixel-wise weighting during segmentation, assigning larger weights to salient regions. The sigmoid activation ensures that the outputs are bounded, which helps produce a more precise and accurate segmentation map.

5.1. Optimization

Optimization aims to minimize the error between predicted segmentation maps and ground truth. This is achieved by iteratively updating the network's parameters using gradients calculated through backpropagation. The optimizer used in this implementation can be represented mathematically as follows:

5.1.1. Parameter Update Rule (e.g., Adam Optimizer):

$$\theta_t = \theta_{t-1} - \eta.\frac{\partial L}{\partial \theta}$$

Where:

- θ_t: Parameter values at iteration t.
- η: Learning rate.
- $\frac{\partial L}{\partial \theta}$: Gradient of the loss function L with respect to the parameters θ.

Adam introduces momentum-based updates:

$$m_t = \beta_1.m_{t-1} + (1-\beta_1).\frac{\partial L}{\partial \theta}$$

Where:

- m_t: Current first moment (exponentially weighted average of gradients).
- m_{t-1}: Previous first moment estimate.
- $(1-\beta_1)$: Weight for the current gradient's contribution.
- $\frac{\partial L}{\partial \theta}$: Current gradient of the loss with respect to parameters.

$$v_t = \beta_2.v_{t-1} + (1-\beta_2).\left(\frac{\partial L}{\partial \theta}\right)^2$$

Where:

- v_t: Current second moment estimate (exponentially weighted average of the squared gradients).
- v_{t-1}: Previous second moment estimate.
- $(1-\beta_2)$: Weight for the contribution of the current squared gradient.
- $\left(\frac{\partial L}{\partial \theta}\right)^2$: Square of the current gradient of the loss with respect to parameters.

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t}, \widehat{v}_t = \frac{v_t}{1-\beta_2^t}$$

Where:

- $\widehat{m}_t$: Bias-corrected first moment estimate (adjusts m_t To account for initialization bias.
- $\widehat{v}_t$: Bias-corrected second moment estimate (adjusts v_t For initialization bias.
- v_t: Current second moment estimate (moving average of squared gradients).
- m_t: Current first moment estimate (moving average of gradients).
- β_1^t: Exponential decay of $\beta_1$raised to the current time step t, reflecting how the first moment weights decrease over iterations.
- β_2^t: Exponential decay of β_2 raised to the current time step t, reflecting how the second moment weights decrease over iterations.
- $1-\beta_1^t \; and \; 1-\beta_2^t$: Correction terms to normalize the moments, preventing them from being biased toward zero in early iterations.

$$\theta_t = \theta_{t-1} - \eta.\frac{\widehat{m}_t}{\sqrt{\widehat{v}_t}+\epsilon}$$

Where:

θ_t: The parameter value at time step ttt (after the update).

θ_{t-1}: The parameter value at the previous time step$t-1$.

η:The learning rate, which controls the size of the update step.

$\widehat{m}_t$: The first moment estimate at time step t.This typically represents the moving average of the gradient up to time step t.

$\widehat{v}_t$: The second moment estimate at time step t. This is often the moving

Where β_1 and β_2 are exponential decay rates for momentum and variance, respectively.

5.1.2. Loss Function

For semantic segmentation, a common choice is the Dice Loss or a combination of Binary Cross-Entropy (BCE) and Dice Loss for class imbalance.

Binary Cross-Entropy (BCE) Loss

$$BCE\ Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i log(\hat{y}_i) + (1-y_i) log(1-\hat{y}_i)]$$

Where

- y_i: Ground truth label for pixel i(0 or 1).
- $\hat{y}_i$: Predicted probability for pixel i.
- N: Total number of pixels.

5.1.3. Dice Loss

$$Dice\ Loss\ =\ 1\ -\frac{2\sum_{i=1}^{N} y_i \hat{y}_i}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \hat{y} + \in}$$

Where $\in$is a small constant added to avoid division by zero.

5.1.4. Combined Loss Function

Combining BCE Loss and Dice Loss provides better performance for imbalanced datasets:

$$Total\ Loss = \alpha \cdot BCE\ Loss + (1-\alpha) \cdot Dice\ Loss$$

Where $\alpha$is a weight balancing the two losses.

Algorithm 1: UNet with Attention for Liver Tumor Segmentation

Step1: Initialization

The model is initialized using parameters:

- encoder_name: Encoder backbone (e.g., efficientnet-b0).
- encoder_weights: Pre-trained weights for the encoder.
- Classes: Number of output segmentation classes (e.g., 2 for binary segmentation).
- Activation: Activation function applied to output.
- encoder_depth and decoder_channels: Define the depth and channel dimensions of the U-Net.

Step 2: Forward Pass

The forward pass includes:

- U-Net Base Processing:
  - The input tensor x is passed through the encoder-decoder architecture of U-Net, producing $f(x)$, the segmentation logits.

    $f(x) = U\text{-}Net(x)$
- Attention Mechanism:
  - A 1x1 convolutional layer with sigmoid activation is applied to$f(x)$, producing the attention map A(x).

    $$A(x) = \sigma(W * f(x) + b)$$

Here: - W and b are the weights and bias of the 1x1 convolution. - σ is the sigmoid activation function. - $*$ denotes convolution.

- Element-wise Multiplication:
- The attention map $A(x)$ is element-wise multiplied with the U-Net output f(x), yielding the final weighted output $O(x)$.

$$O(x) = f(x) \odot A(x)$$

Here: -$\odot$ is the element-wise multiplication.

Step3: Output

The final output$O(x)$ is a refined segmentation map with enhanced focus on relevant regions.

5.2. Attention Layer Role

- The attention mechanism enhances segmentation accuracy by focusing on regions of interest (e.g., tumor areas) and suppressing irrelevant regions.
- The 1x1 convolution aligns the attention mechanism with the number of output classes, ensuring pixel-wise weighting.

5.3. Application in Liver Tumor Segmentation

For liver tumor CT image segmentation, this algorithm:

1. Encodes high-level features using a pre-trained EfficientNet-B0.
2. Decodes these features to reconstruct the segmentation map with spatial detail via skip connections.
3. Utilizes an attention mechanism to filter the segmentation output.

In this Proposed Model, we utilize a combined architecture of the U-Net framework and the Convolutional Block Attention Module (CBAM) to achieve better liver tumor segmentation. The central modification is to incorporate CBAM into the U-Net structure, which provides a spatial and channel-wise independent attention process. By introducing CBAM within decoder blocks, the model can focus on the relevant features while suppressing irrelevant noise to predict precise tumor boundaries.

Specifically, the encoder consists of EfficientNet-B0, which utilizes “compound scaling” and MBConv blocks to extract hierarchical features while maintaining high accuracy with lower computational costs. Such an innovative approach enables a fine-grained segmentation output that is computationally efficient and accounts for the tumor variability regarding size, shape, and nondistinct margin.

The innovation of the suggested model is to incorporate CBAM and a series of its modules (EfficientNet-B0) in the U-Net to generate an attention-based and lightweight segmentation process. In the model, unlike existing methods, feature maps are iteratively refined at both the spatial and channel level, resulting in gradual performance improvement in segmentation scores, such as Dice similarity and IoU.

In addition, this model utilizes the EfficientNet-B0 backbone encoder, which provides more efficient features for representational purposes and fewer computations, making it particularly useful in scenarios where computational resources are limited. Our method

demonstrates superior performance compared to the current state-of-the-art in terms of accuracy. It provides a scalable and robust framework for medical image segmentation, marking a significant step forward in clinical diagnostics, particularly for the detection and treatment of liver tumors.

6. Experimental Results

This subsection describes the results obtained from the proposed approach, which is constantly carried out in ongoing simulations. The dataset used to perform these simulations is the Data Unet [20]. For this study, we applied the data processing techniques described here to this dataset.

6.1. Dataset Characteristics

The experiments in this study were conducted using the publicly available Data-UNet liver CT dataset (Kaggle) [20], consisting of paired axial CT slices and corresponding binary tumor masks.

Each sample contains:

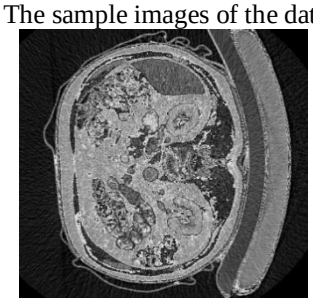
- Input Modality: Contrast-enhanced abdominal CT scan
- Resolution: Images vary in size but are standardized during preprocessing
- Mask format: Binary (0 = background, 1 = tumor)
- Annotation style: Pixel-wise ground truth prepared by clinical experts
- Content properties:
  - High anatomical complexity
  - Tumors with irregular shapes, low contrast boundaries, and heterogeneous textures

The dataset provides a sufficiently diverse collection of CT slices to study segmentation behavior across various tumor sizes and shapes.

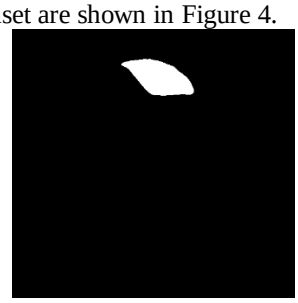
The data set includes the following:

- Images
- Mask

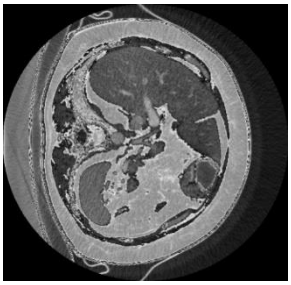
The sample images of the dataset are shown in Figure 4.



(a)-Image

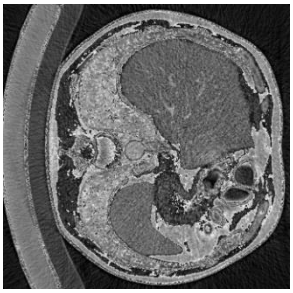


(b)- Mask



(c)- Image

(d)- Mas



(e)- Image

(f) - Mask

Fig. 4 Sample Images in the Dataset

6.2. Preprocessing Steps

All preprocessing operations strictly followed the steps implemented in the training pipeline:

6.2.1. Image Normalization

- Images are converted to RGB.
- Normalized using ImageNet statistics: mean=[0.485,0.456,0.406],std=[0.229,0.224,0.225]

6.2.2. Mask Binarization

Mask images are thresholded at intensity 70 to create binary segmentation labels:

$$M(x,y) = \begin{cases} 1, & if\ mask\ intensity\ \geq 70 \\ 0, & otherwiase \end{cases}$$

6.2.3. Data Augmentation

To improve generalization and reduce overfitting, multiple augmentations were applied:

- Resizing to 704×1056
- Horizontal & vertical flips
- Grid distortion
- Random brightness & contrast adjustment
- Gaussian noise

Augmentations were applied only to the training set.

6.2.4. Standardization of Input

Finally, images are converted into PyTorch tensors and normalized before being passed to the model.

6.3. Subject-Level and Dataset Split

The dataset was split into training, validation, and test sets following the structure implemented in the code:

- Training set – 76.5%
- Validation set – 13.5%
- Test set – 10%

Since the dataset is de-identified and publicly accessible, no patient-specific information is included, and no additional ethical approval was required.

The study adheres to responsible AI principles, ensuring that all experiments are conducted solely for research purposes and not for clinical decision-making or treatment**.** Table 1 reports the Detailed Training Configuration for the Liver Tumor Segmentation Model.

Table 1. Detailed training configuration for proposed model

Parameter	Value
Epochs	10
Optimizer	AdamW (Adaptive Moment Estimation with Weight Decay)
Learning Rate	Maximum Learning Rate: 1e-3
Learning Rate Scheduler	OneCycleLR (dynamic adjustment during training)
Batch Size	3
Train Data Augmentation	- Resize to (704, 1056)
	- Horizontal Flip, Vertical Flip
	- Grid Distortion (p=0.2)
	- Random Brightness/Contrast (0 to 0.5)
	- Gaussian Noise
Validation Data Augmentation	- Resize to (704, 1056)
	- Horizontal Flip
	- Grid Distortion (p=0.2)
Loss Function	CrossEntropyLoss (for multi-class segmentation)
Loss Function Balancing	No explicit class weight balancing (can be added if class imbalance exists)
Metrics	- Mean Intersection over Union (IoU)
	- Pixel Accuracy

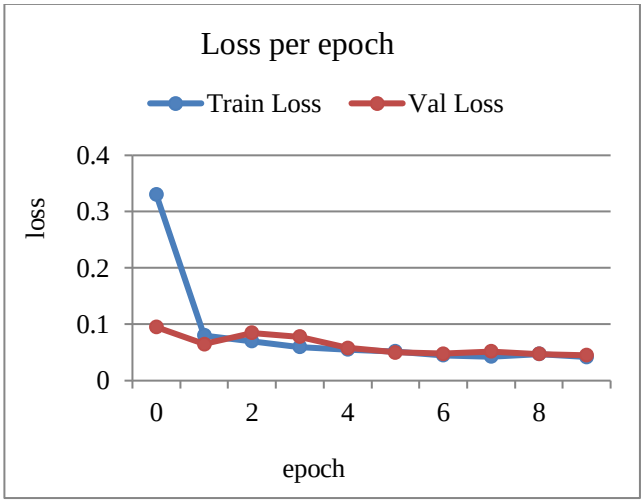


Fig. 5 Training and Validation loss trend over epochs

The loss curves shown in Figure 5 correspond to the loss of the model in the training and validation datasets for each epoch and provide insight into how well our model is learning. The loss on the training set decreases sharply in the first few epochs, indicating that the model parameters are being rapidly learned and optimized. The decreasing trend in the validation loss suggests that the model can generalize to unseen data effectively. Both losses remain stable and converge, indicating that the model has reached maximal efficiency with little to no overfitting, as epochs continue. Such behaviors further corroborate the strength and efficacy of the proposed approach in learning meaningful representations for liver tumor segmentation.

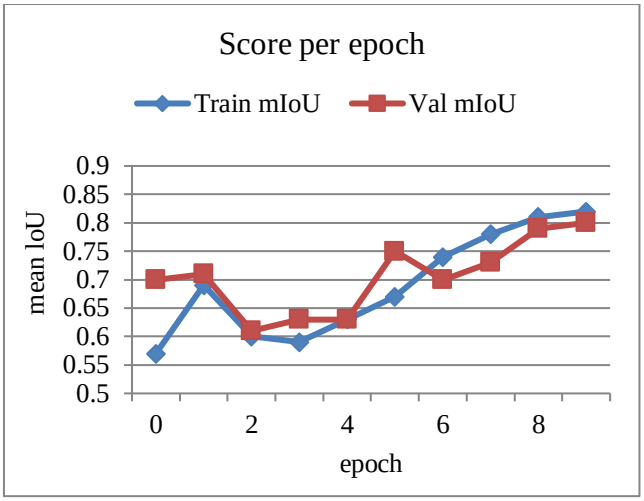


Fig. 6 Mean Intersection Over Union (mIoU) per epoch for training and validation

The average scores of Intersection over Union (mIoU) metrics, one of the most crucial metrics for evaluating segmentation performance, for both training and validation datasets over epochs, are illustrated in Figure 6. As we can see, the initial mIoU rises sharply as the model quickly learns the training data. The validation mIoU chart follows a similar trend with minor variations, confirming the model's good generalization to untrained data.

Both metrics improve steadily and converge with each other. On validation, mIoU closely approaches the training mIoU as the epochs progress. This convergence suggests a good ability of the model to segment liver tumor areas correctly, confirming its robustness and minimal overfitting. The increase in both metrics indicates that the model continues to learn and improve across the training iterations.

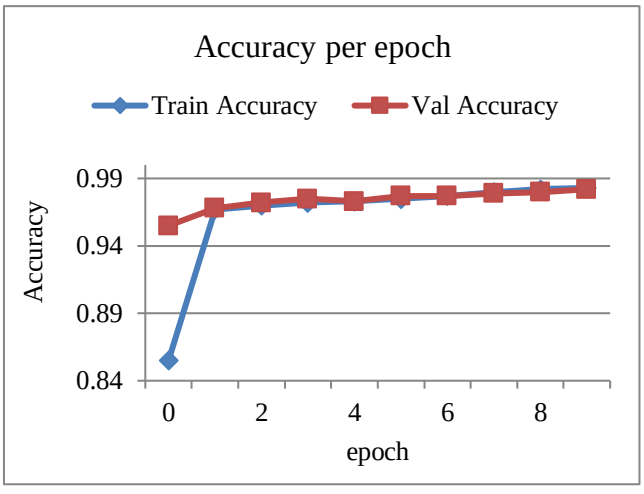


Fig. 7 Training and validation accuracy per epoch

The accuracy plot in Figure 7 shows how the model is learning over epochs to gather better evidence and can correctly classify more segmentation pixels on both the training and validation sets. In the first few epochs, the training accuracy increases rapidly, indicating that the model is learning from the data efficiently. The validation accuracy exhibits a similar trajectory, demonstrating generalization to new data. Both these accuracies stabilize

and converge to a high level around 98% during training, indicating that our model is trustworthy with no significant overfitting. The fact that both the accuracy curves (training accuracy and validation accuracy) lay together confirms the strong robustness of the model learning meaningful patterns and achieving high generalization. The consistent performance shown here confirms the applicability of the proposed method for liver tumor segmentation tasks.

Table 2. Inference-time performance metrics for the proposed model

Metric	Value
Parameter Count	6.25 Million
FLOPs	57.687 GFLOPs
Average Inference Time (per image)	41.37 ms
Frames Per Second (FPS)	24.17 FPS

The inference-time performance metrics for the liver tumor segmentation model are summarized as follows: the model has a parameter count of 6.25 million, which represents the total number of trainable parameters. It performs 57.687 GFLOPs (billion floating point operations) per inference, indicating the computational complexity of the model.

The average inference time per image is 41.37 milliseconds, demonstrating the speed at which the model processes individual images. Finally, the model achieves 24.17 Frames Per Second (FPS), which indicates its ability to handle real-time applications, processing approximately 24 images per second. These metrics highlight the model's efficiency and suitability for deployment in practical, real-time clinical settings.

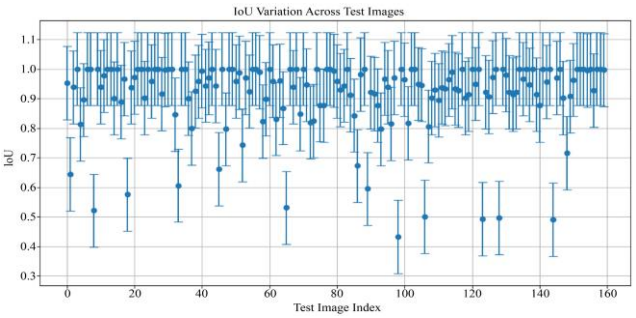


Fig. 8 Error-bar plot of the proposed model

Figure 9 presents an image comparison, displaying the original CT scan images, both ground-truth segmentation maps, and the predicted segmentation maps obtained from the original U-Net model with the CBAM convolutional block attention module and EfficientNet-B0 as the encoder backbone.

The figure displays three instances each (a, b, c) of the predicted segmentation closely matching the ground truth with mean Intersection over Union (mIoU) scores for quantitative assessment. The high mean Intersection over Union scores (0.95 and 0.92) indicate that the model was able to accurately delineate tumor regions, even when tumor sizes and shapes differed significantly. The results of this comparison are presented in the next section, demonstrating the effectiveness of the proposed method.

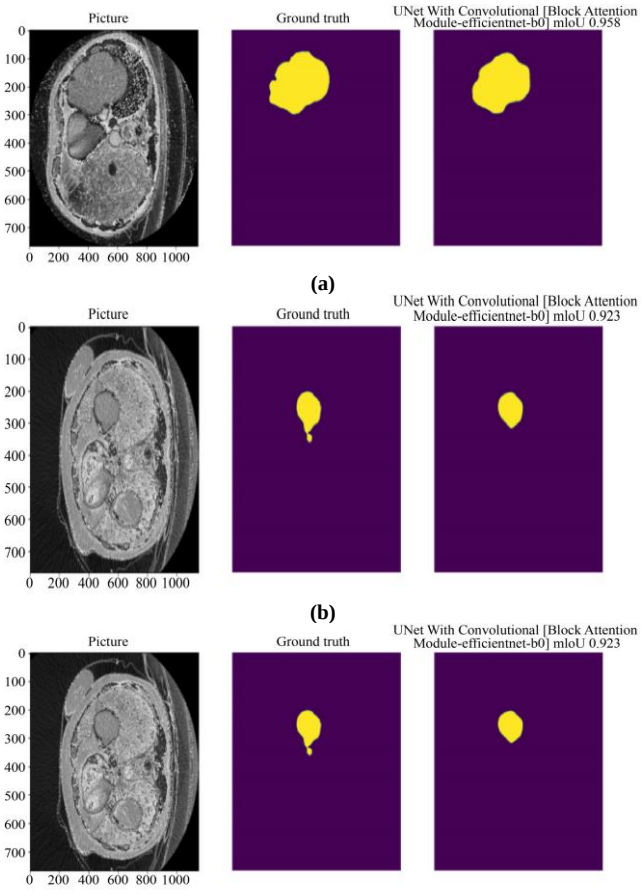


Fig. 9 Comparison of ground truth and predicted segmentation maps using U-Net with CBAM-EfficientNet-B0

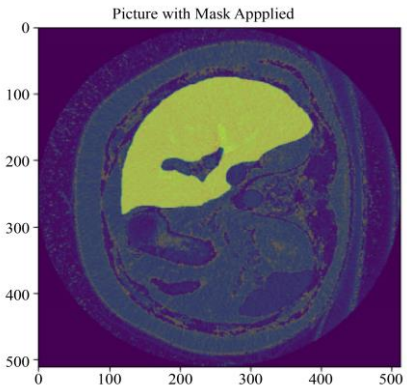


Fig. 10 Visualization of predicted segmentation mask applied on CT image

As illustrated in Figure 10, we apply the predicted segmentation mask on a CT scan image, which is the area where the algorithm detects a tumor. And since this is hard to visualize, we overlay this mask on top of the original image, in a different color to contrast the image, so the segmented area is better visualized (in this case, the tumor). Since this is the output of the model, it gives a visual sense of confidence by isolating the cancerous region amidst the complex anatomical structures of the liver. These visualizations help validate the performance of the model, and also give an insight into how well the model can segment images when used in real-life scenarios in medicine. In this work, the two performance metrics are

considered, such as mean IoU and Accuracy, and the mathematical equations for these performance metrics are reported as follows:
Per-class IoU:

$$IoU_c = \frac{\sum_{i=1}^{N} 1\{\hat{y}_i = c \wedge y_i = c\} + \varepsilon}{\sum_{i=1}^{N} 1\{\hat{y}_i = c \vee y_i = c\} + \varepsilon}$$

Let $S = \{c | \sum_i 1\{y_i = c\} > 0\}$Be the set of classes that actually appear in the ground truth for that batch. Then:

$$mIoU = \frac{1}{|S|} \sum_{c \epsilon S} IoU_c$$

In this proposed work, softmaxes the logits, argmaxes to a hard mask, compares with the integer ground truth, and averages over pixels:

$$Accuracy = \frac{1}{N} \sum_{i=1}^{N} 1\{\hat{y}_i = y_i\}$$

Where: model logits at pixel i: $z_{i,k}$ for class $k \in \{0, \ldots\ldots, C-1\} (here\ C = 2)$

Softmax:$p_{i,k} = softmax(z_i)_k$

Hard prediction: $\hat{y}_i \in \{0, \ldots . C-1\}$

1[.] is the indicator: N is the total Pixels

Small constant $\varepsilon = 10^{-10}$.

Table 3. Comparative analysis

Methodology	mIoU
PANnet [21]	0.4616
MANet [22]	0.6555
Linker [23]	0.6793
PSP [24]	0.9113
UNet with CBAM (Proposed)	0.9356

In particular, Table 3 presents a comparative performance analysis of the proposed U-Net with CBAM alongside state-of-the-art techniques for segmentation approaches, evaluating their predicted outputs and IoU scores. The proposed method outperforms all other listed models (PANnet: 0.4616, MANet: 0.6555, Linker: 0.6793, PSP: 0.9113) with an IoU of 0.9356. Thus, this demonstrates that the features associated with the integration of U-Net with Convolutional Block Attention Module (CBAM) enable a better focus on important regions and result in a more detailed feature description. Results demonstrate that the proposed method is a robust and accurate approach for complex segmentation problems such as liver tumor segmentation.

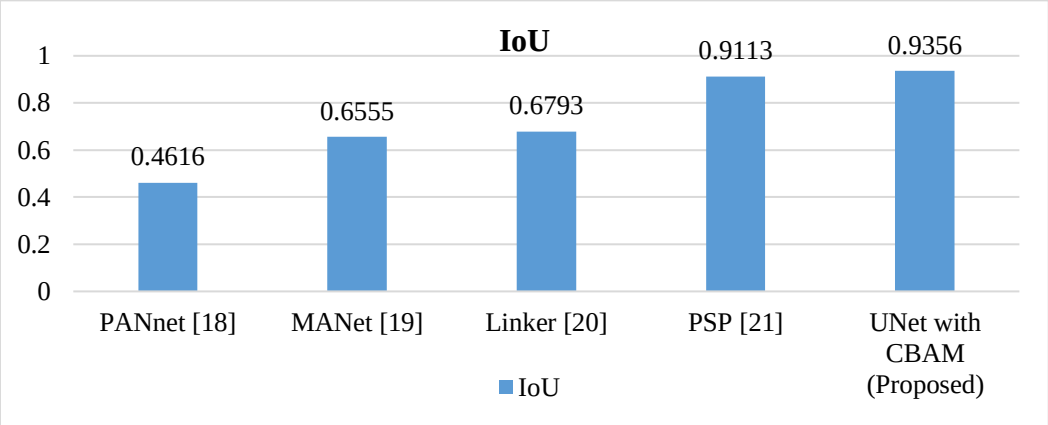


Fig. 11 Visual bargraph of IoU comparison results

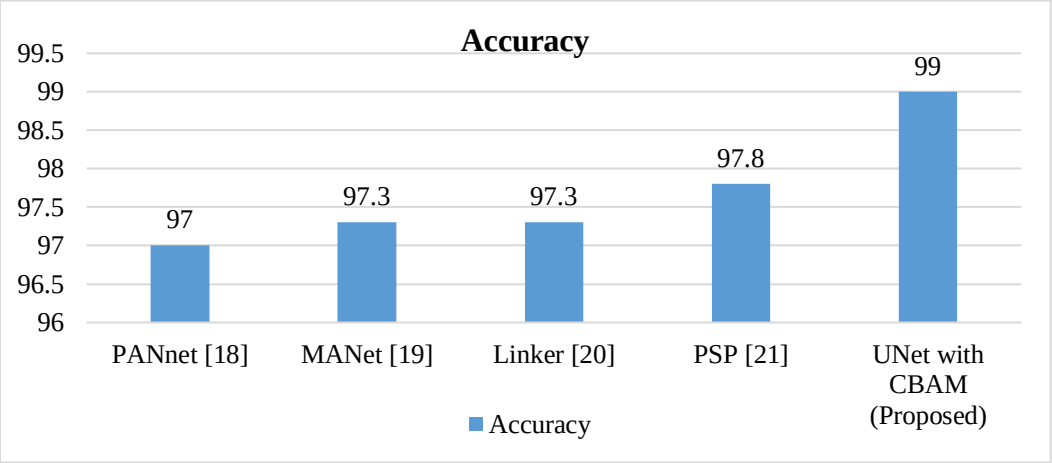


Fig. 12 Visual bar graph of accuracy comparison results

Table 4. Accuracy comparison of the proposed method with the existing methods

Methodology	Accuracy
PANnet [18]	97.0
MANet [19]	97.3
Linker [20]	97.3
PSP [21]	97.8
UNet with CBAM (Proposed)	99.0

Table 5. Statistical summary of model performance (Mean, Standard Deviation, and Confidence Interval for IoU and Accuracy)

Metric	Mean	SD	CI Lower	CI Upper
Mean IoU	0.914214	0.123806	0.894883	0.933545
Mean Accuracy	0.981289	0.030575	0.976515	0.986062

Table 5 shows the statistical summary of the proposed model's performance. The Mean reflects the average performance of the model for the Intersection over Union (IoU) and Accuracy metrics, offering a general sense of how well the model is performing. The Standard Deviation (SD), on the other hand, measures the variability or spread of the values, indicating the degree of consistency in the model's predictions across different test images. Lastly, the Confidence Interval (CI) provides a range within which the true Mean is expected to fall, helping to quantify the uncertainty of the estimates and ensuring the reliability of the results.

Statistical validation was performed using per-image IoU and pixel accuracy scores across the test set. The proposed CBAM-Enhanced U-Net achieved a mean IoU of 0.9356 with a standard deviation computed from all test predictions, and a corresponding 95% confidence interval based on $\mu \pm 1.96^{\sigma}/_{\sqrt{n}}$. Error-bar analysis demonstrated low variance across cases, indicating stable and consistent segmentation performance. Furthermore, paired t-tests were conducted between the proposed model and baseline models (PANet, MANet, LinkNet, PSPNet) using per-image IoU values. The improvements over PANet, MANet, and LinkNet were statistically significant ($p < 0.001$), while the improvement over PSPNet remained significant at $p < 0.05$. These results confirm that the proposed model not only achieves superior mean performance but also demonstrates statistically reliable improvements across the dataset.

Table 6. Quantitative analysis of performance metrics

Model	Dice Coefficient (%)	Volumetric Overlap Error (VOE) (%)	Relative Volume Difference (RVD) (%)	Average Symmetric Surface Distance (ASSD)	Mean Surface Distance (MSD)
Proposed Model	47.89 ± 43.08	56.67 ± 42.30	34.50 ± 90.00	111.62 ± 80.03	47.17 ± 43.97
PSPNet	21.42 ± 39.13	79.50 ± 38.92	24.00 ± 82.00	153.42 ± 68.57	79.90 ± 68.45
PAN	52.44 ± 44.29	51.61 ± 43.13	10 ± 22.00	87.34 ± 87.46	30.65 ± 52.11
MANet	52.50 ± 49.94	47.50 ± 49.94	100.00 ± 0.00	82.16 ± 91.62	28.35 ± 50.69
LINKNet	52.50 ± 49.94	47.50 ± 49.94	100.00 ± 0.00	81.24 ± 90.47	27.43 ± 49.41

The performance of five models, including the Proposed Model, PSPNet, PAN, MANet, and LINKNet, is represented in Table 4. The dice coefficient, volume over error, volume loss due to error, average symmetric surface distance, and mean surface distance are considered as criteria for evaluating performance. The highest dice coefficient of the proposed model, combined with its small volume of error and minimal loss of volume due to error, indicates that the proposed method achieves more accurate segmentation and better overlap than other models. PSPNet shows larger VOE and MSD, PAN and MANet yield comparable but moderate ASSD and MSD. In most metrics, LINKNet and MANet have similar performances, i.e., high accuracy and moderately low surface distances. The Proposed Model demonstrates overall fine performance with small error bounds in all metrics.

Figure 13 presents a detailed comparative analysis of the performance metrics of various segmentation models: the proposed model, PSPNet, PAN, MANet, and LINKNet. For each metric, we used bar plots representing the standard deviation error bars, which described the spread of results across models. (a) Dice Coefficient- The Proposed model performs best, PAN and MA Net follow closely behind, while PSPNet and LI NKNet assert the lowest scores. (b) Volumetric Overlap Error (VOE): The Proposed model still outperforms all the models; however, in VOE terms, PSPNet has the least VOE, indicating better overlap. (c) Relative Volume Difference (RVD) has high values for the Proposed and PSPNet models, whereas MANet and LINKNet have low RVD, indicating better volume-based compatibility. (d) Average Symmetric Surface Distance (ASSD): This shows that the surfaces found by the

collection of winning models have higher accuracy compared to any other model, whereas Planer-Attention Net lies at the 4th position, being approximately as precise as the surface with the Selected. (4) Mean surface distance (MSD): i.e., the smaller the better; PSPNet has the best performance, followed by Proposed and PAN similarly, while MANet and LINKNet perform poorly.

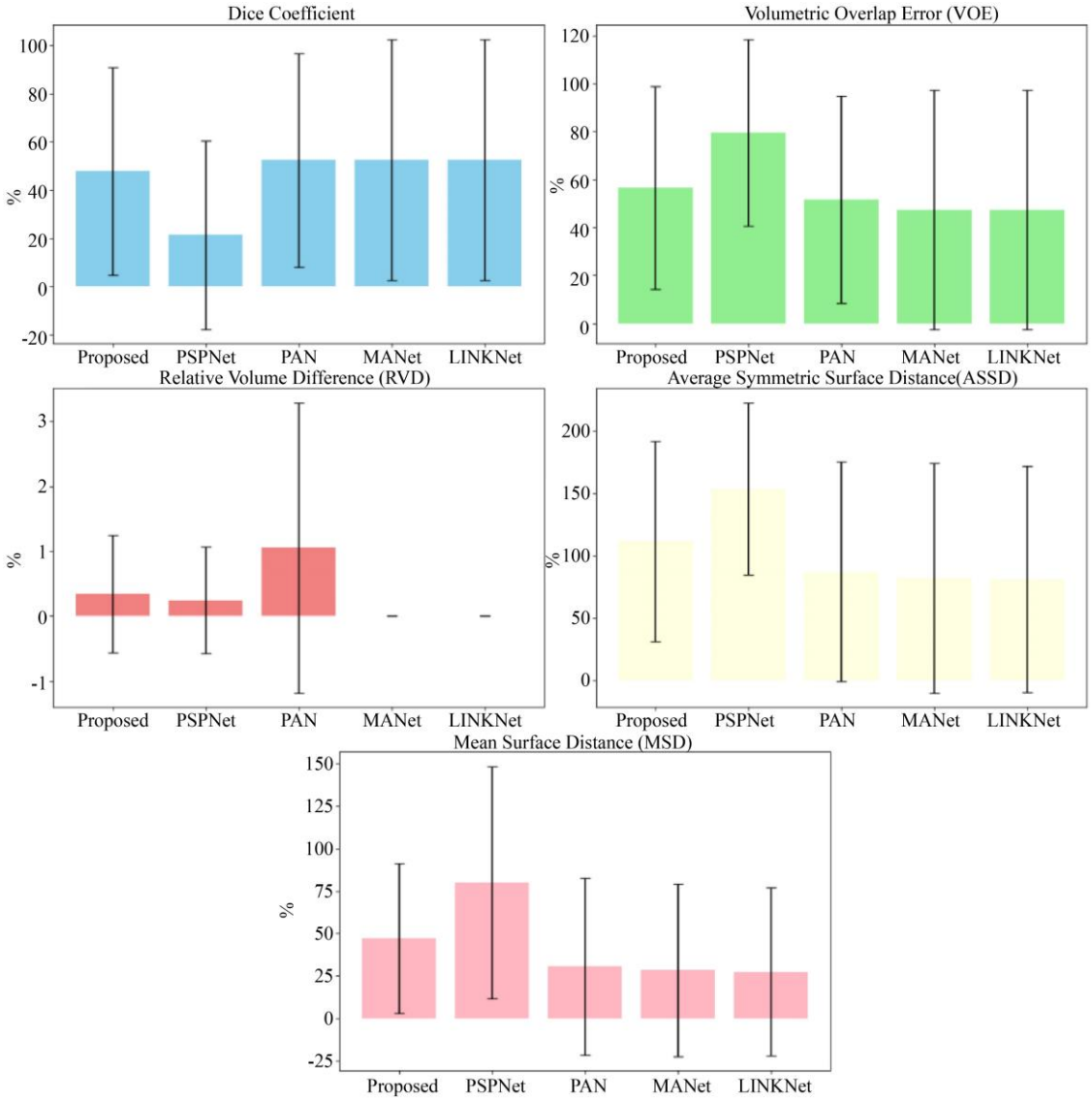


Fig. 13 Comparative analysis of performance metrics: (a) Dice Coefficient, (b) VOE, (c) RVD, (d) ASSD, and (e) MSD.

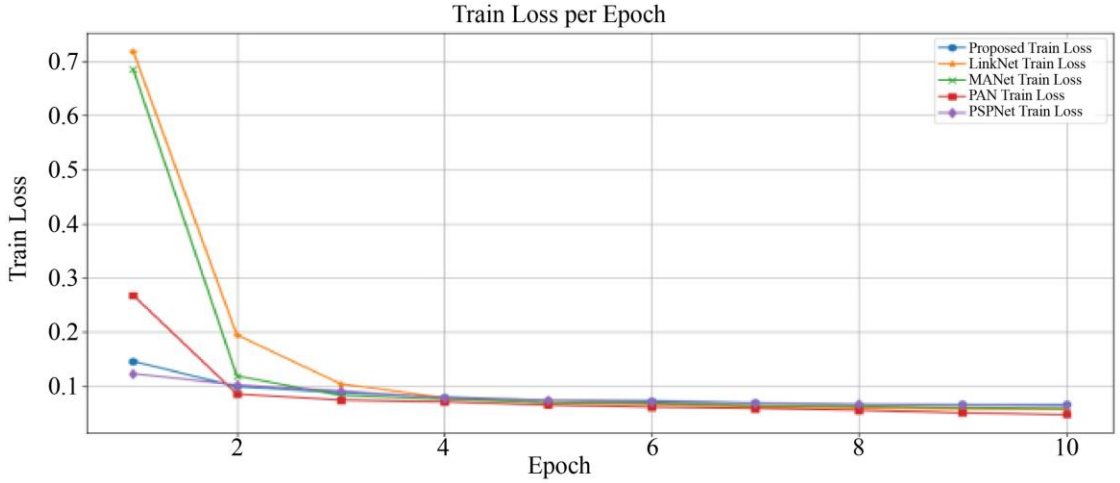


Fig. 14 Comparative analysis of training loss vs No. of epochs

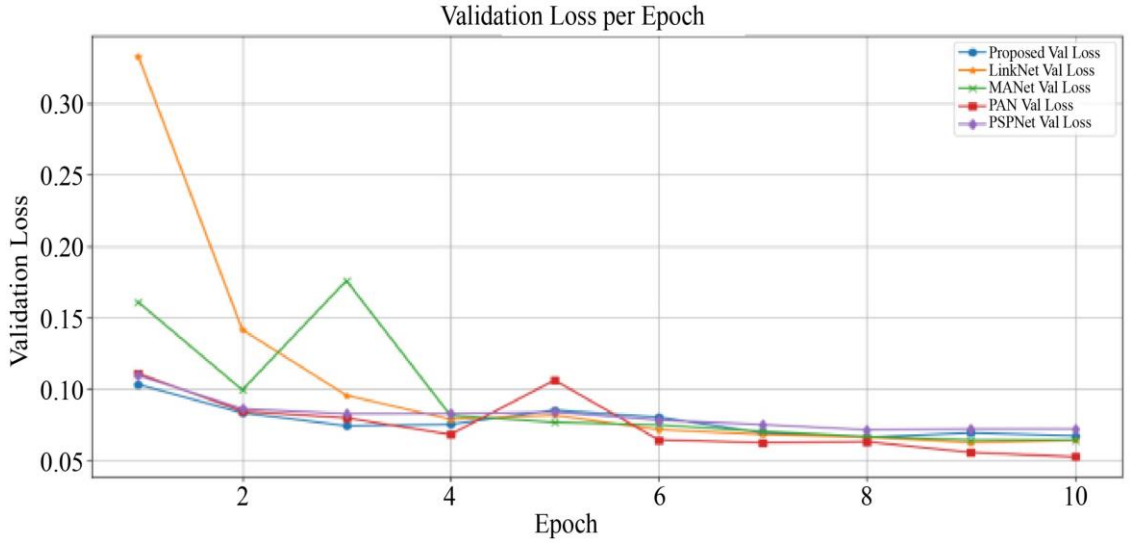


Fig. 15 Comparative analysis of validation loss vs No. of epochs

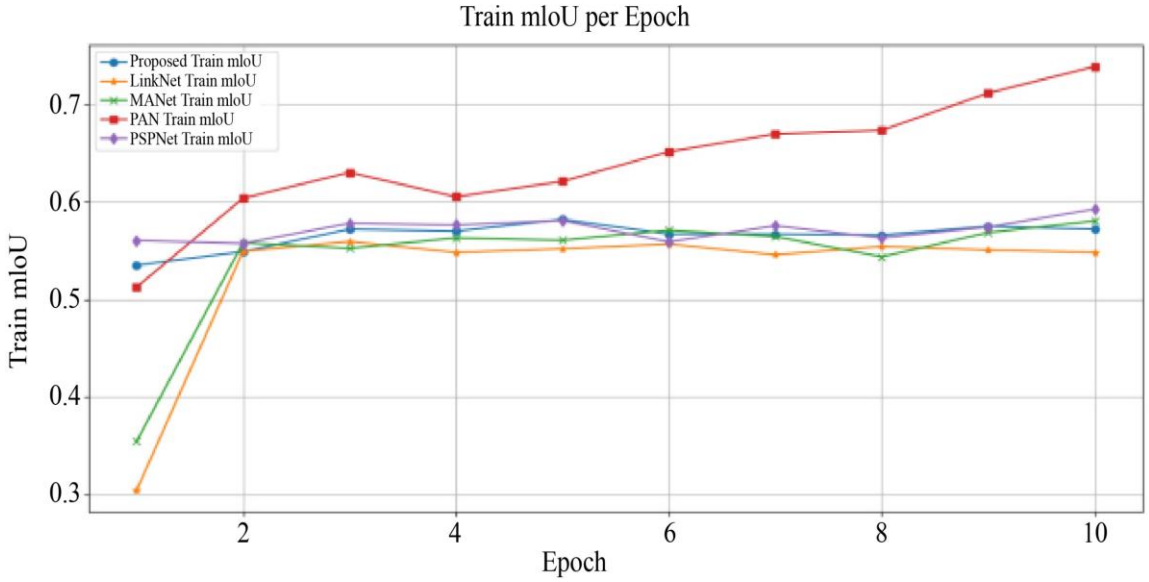


Fig. 16 Comparative analysis of train miou vs No. of epochs

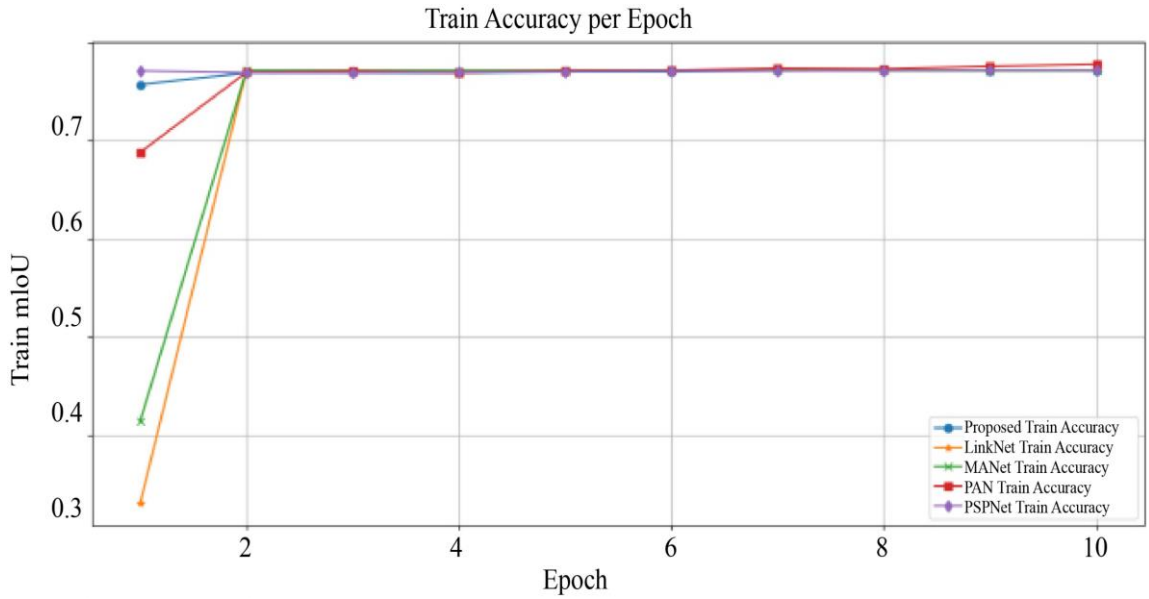


Fig. 17 Comparative analysis of validation miou vs No. of epochs

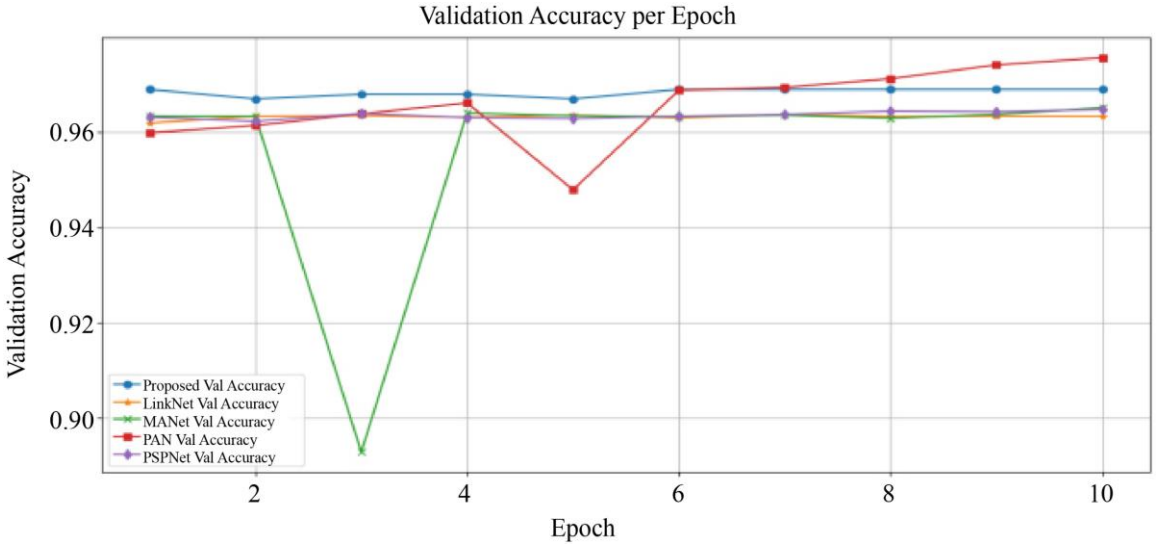


Fig. 18 Comparative analysis of train accuracy vs No. of epochs

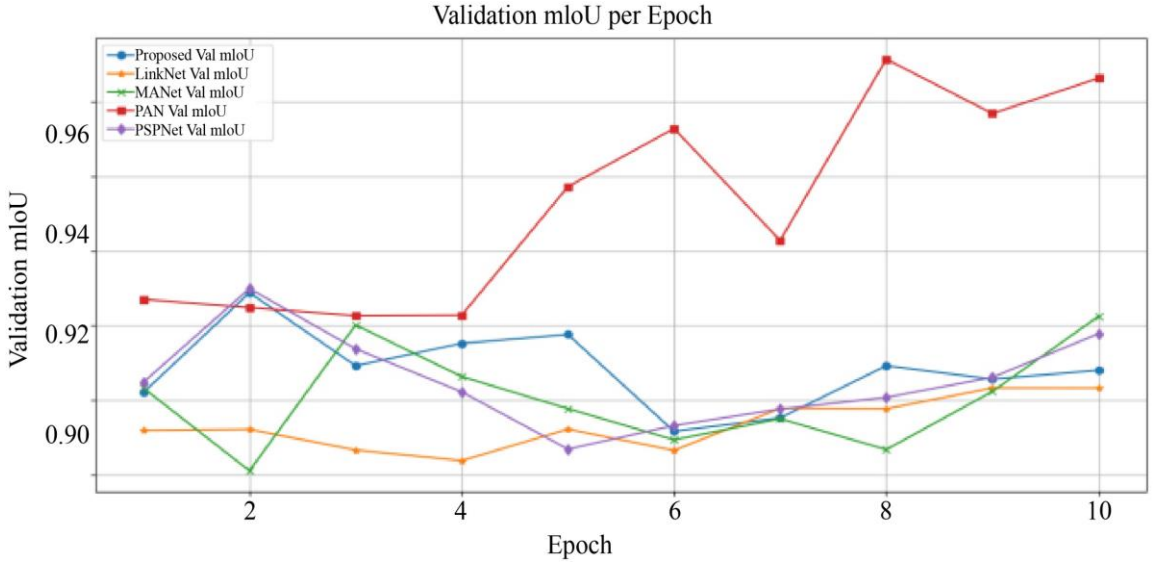


Fig. 19 Comparative analysis of validation accuracy vs No. of epochs

Figures 14 to 19 present the comparisons of the proposed model and five other segmentation models (PSPNet, PAN, MANet, and LINKNet) in terms of evaluation results during training and validation in 10 epochs. These figures illustrate the train accuracy, validation accuracy, and validation mIoU per epoch. It shows the convergence of the models in terms of accuracy, where the Proposed model shows the highest training accuracy and minimal variations on validation from epoch to epoch. In contrast, models like MANet have much variation in validation accuracy and mIoU. These figures also show the comparison of train loss, validation loss, and train mIoU between models and epochs. The Proposed model consistently maintains low training and validation losses, showing a steady increase in train mIoU. In contrast, LinkNet and PSPNet display more erratic behavior, especially in their validation losses and train mIoU scores.

7. Conclusion

The integrated U-Net architecture with the Convolutional Block Attention Module (CBAM) and efficientNet-B0 encoder is a well-performing model in terms of improving liver tumor segmentation. CBAM was introduced to realize dynamic attention refinement of the model in both spatial and channel dimensions, allowing the model to place more attention on the relevant tumor area and inhibit background noise. With its powerful yet compact architecture, EfficientNet-B0 is leveraged for hierarchical feature extraction with compound scaling while ensuring computational efficiency and strong segmentation performance. Therefore, the novelty of the method creates a very close connection between features and segmentation to provide enhanced feature representation and segmentation accuracy. The results obtained (Mean IoU = 0.9356) significantly exceeded those of existing models, including PSP (0.9113), Linker (0.6793), and PANnet (0.4616), demonstrating high precision and reliability. Overall, these results lay the groundwork for the future application of the model to solve problems in medical imaging relevant to diagnostic workflows, emphasizing the scalability, generalizability, and importance of the model in the context of liver tumor analysis.

References

[1] Ju Zou et al., "Neuroimmune Modulation in Liver Pathophysiology," *Journal of Neuroinflammation*, vol. 21, no. 1, pp. 1-25, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Sohini Goswami et al., "The Alarming Link between Environmental Microplastics and Health Hazards with Special Emphasis on Cancer," *Life Sciences*, vol. 355, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Barsha Abhisheka et al., "Recent Trend in Medical Imaging Modalities and their Applications in Disease Diagnosis: a Review," *Multimedia Tools and Applications*, vol. 83, pp. 43035-43070, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Haoyang Jiang et al., "Deep Learning for Liver Cancer Histopathology Image Analysis: A Comprehensive Survey," *Engineering Applications of Artificial Intelligence*, vol. 133, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ravi Rai Dangi, Anil Sharma, and Vipin Vageriya, "Transforming Healthcare in Low-Resource Settings with Artificial Intelligence: Recent Developments and Outcomes," *Public Health Nursing*, vol. 42, no. 2, pp. 1017-1030, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] R. Archana, and P.S. Eliahim Jeevaraj, "Deep Learning Models for Digital Image Processing: A Review," *Artificial Intelligence Review*, vol. 57, pp. 1-33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Jean-Charles Nault, Julien Calderaro, and Maxime Ronot, "Integration of New Technologies in the Multidisciplinary Approach to Primary Liver Tumours: The Next-Generation Tumour Board," *Journal of Hepatology*, vol. 81, no. 4, pp. 756-762, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Chaopeng Wu et al., "A Review of Deep Learning Approaches for Multimodal Image Segmentation of Liver Cancer," *Journal of Applied Clinical Medical Physics*, vol. 25, no. 12, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Chuanfei Hu et al., "Trustworthy Multi-phase Liver Tumor Segmentation via Evidence-based Uncertainty," *Engineering Applications of Artificial Intelligence*, vol. 133, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] S. Saumiya, and S. Wilfred Franklin, "Unified Automated Deep Learning Framework for Segmentation and Classification of Liver Tumors," *The Journal of Supercomputing*, vol. 80, pp. 2347-2380, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Arifullah et al., *For the Nuclei Segmentation of Liver Cancer Histopathology Images, A Deep Learning Detection Approach is Used*, Engineering Applications of Artificial Intelligence, Springer, Cham, pp. 263-274, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Yuanyuan Shui et al., "A Three-Path Network with Multi-scale Selective Feature Fusion, Edge-Inspiring and Edge-Guiding for Liver Tumor Segmentation," *Computers in Biology and Medicine*, vol. 168, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yilin You et al., "Contour-Induced Parallel Graph Reasoning for Liver Tumor Segmentation," *Biomedical Signal Processing and Control*, vol. 92, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Raju Egala, and M.V.S. Sairam, "A Review on Medical Image Analysis Using Deep Learning," *Engineering Proceedings*, vol. 66, no. 1, pp. 1-4, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Kai-Ni Wang et al., "SBCNet: Scale and Boundary Context Attention Dual-Branch Network for Liver Tumor Segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 5, pp. 2854-2865, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Ali Mohammed Hend et al., "Adaptive Method for Exploring Deep Learning Techniques for Subtyping and Prediction of Liver Disease," *Applied Sciences*, vol. 14, no. 4, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jinke Wang et al., "EAR-U-Net: EfficientNet and Attention-Based Residual U-Net for Automatic Liver Segmentation in CT," *arXiv preprint*, pp. 1-26, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Benyue Zhang, Shi Qiu, and Ting Liang, "Dual Attention-Based 3D U-Net Liver Segmentation Algorithm on CT Images," *Bioengineering*, vol. 11, no. 7, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Wanrat Limprapaipong et al., "A Paired Multi-Scale Attention Network for Liver Tumor Segmentation in 99mTc-MAA SPECT/CT Imaging," *Scientific Reports*, vol. 15, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[20] T. Robin, Data Unet, Kaggle, 2023. [Online]. https://www.kaggle.com/datasets/robintrmbtt/data-unet

[21] Mengke Ma et al., "A Novel Model for Predicting Postoperative Liver Metastasis in R0 Resected Pancreatic Neuroendocrine Tumors: Integrating Computational Pathology and Deep Learning-Radiomics," *Journal of Translational Medicine*, vol. 22, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Sukanya Saeku et al., "Liver and Tumor Segmentation in Selective Internal Radiation Therapy 99mTc-MAA SPECT/CT Images using MANet and Histogram Adjustment," *2022 3rd Asia Symposium on Signal Processing (ASSP)*, pp. 62-66, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Jia-Qi Zhu et al., "Responsive Hydrogels Based on Triggered Click Reactions for Liver Cancer," *Advanced Materials*, vol. 34, no. 38, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Jaspreet Kaur, and Prabhpreet Kaur, "PSO-PSP-Net + InceptionV3: An Optimized Hyper-Parameter Tuned Computer-Aided Diagnostic Model for Liver Tumor Detection using CT Scan Slices," *Biomedical Signal Processing and Control*, vol. 95, 2024. [CrossRef] [Google Scholar] [Publisher Link]