Advanced Hybrid Segmentation Model Leveraging AlexNet Architecture for Enhanced Liver Cancer Detection

Venkata Raja Sekhar Reddy Nagireddy[*][iD], Khaja Shareef Shaik[iD]

Department of Information Technology, MLR Institute of Technology, 500043 Hyderabad, India

[*] Correspondence: Venkata Raja Sekhar Reddy Nagireddy (rajasekhar.nv1@gmail.com)

Abstract: Liver cancer, one of the rapidly escalating forms of cancer, remains a principal cause of mortality globally. Its death rates can be attenuated through vigilant monitoring and early detection. This study aims to develop a sophisticated model to assist medical professionals in the classification of liver tumours using biopsy tissue images, thereby facilitating preliminary diagnosis. The study presents a novel, bio-inspired deep learning strategy purposed for augmenting liver cancer detection. The uniqueness of this approach rests in its two-fold contribution: Firstly, an innovative hybrid segmentation technique, integrating the SegNet network, UNet network, and Al-Biruni Earth Radius (BER) procedure, is introduced to extract liver lesions from Computed Tomography (CT) images. The algorithm initially applies the SegNet to isolate the liver from the abdominal image in a CT scan. Since hyperparameters significantly influence segmentation performance, the BER algorithm is hybridized with each network for optimal tuning. The method proposed herein is inspired by the pursuit of a common objective by swarm members. Al-Biruni's methodology for calculating Earth's radius sets the search space, extending beyond local solutions that require exploration. Secondly, a pre-trained AlexNet model is utilized for diagnosis, further enhancing the method's effectiveness. The proposed segmentation and classification algorithms have been compared with contemporary state-of-the-art techniques. The results demonstrated that in terms of specificity, F1-score, accuracy, and computational time, the proposed method outperforms its competitors, indicating its potential in advancing liver cancer detection.

Keywords: Liver cancer detection; Al-Biruni earth radius optimization; U-Net; SegNet; Computed tomography; AlexNet

1 Introduction

The liver, the largest internal organ in the human body, carries out an array of vital functions owing to its strategic location, vast size, and multifaceted capabilities. Unfortunately, it is frequently the site of one of the deadliest cancers globally, liver cancer [1]. Early detection is paramount for the prevention and management of liver cancer. Alas, the disease often remains asymptomatic in its nascent stages, rendering early detection challenging, despite its propensity for devastating liver cells [2].

Liver cancer can be broadly classified into two categories - primary and secondary. Primary liver cancers originate within the liver cells, with Hepatocellular Carcinoma (HCC) being the most prevalent type, constituting about 80% of initial liver malignancies [3]. Secondary liver cancers, often referred to as Metastases (MET), originate in other organs and subsequently metastasize to the liver. The treatment and presentation of these two types are markedly different, despite both having cancerous liver cells [4–6].

Liver tumours can be detected through a myriad of diagnostic techniques, including blood tests, biopsies, and imaging scans. Among these, medical image processing stands out as a non-invasive or minimally invasive method [7]. Non-invasive imaging technologies such as MRI, X-ray, ultrasound, and CT scans have been employed in the diagnosis of liver tumours [8]. These methods allow for tracking the progression of cancer throughout the body, ascertaining its initial location and size. CT scans are particularly efficacious for liver cancer detection, producing two-dimensional images of the organ from various angles [9]. However, the manual analysis of the copious data

generated by medical image processing is time-consuming. Therefore, the application of automated computer-aided diagnostic (CAD) approaches is essential for real-time, accurate tumour detection [10].

Traditional analysis of CT scans of the liver has been manual, a process that is expensive, time-consuming, and prone to errors. To rectify these issues and enhance liver cancer detection, several computational techniques have been proposed. However, these systems have been inadequate in identifying liver lesions due to the complexity of the liver and surrounding organs, small tumour sizes, and irregular tumour growth [11]. Hence, a novel approach is necessitated. Studies have shown that Convolutional Neural Networks (CNNs) can significantly enhance radiological diagnosis without the need to train the system to recognize specific radiological traits [12]. A fusion of the expertise of radiologists and the processing power of AI systems could dramatically improve the efficacy and safety of patient care.

This study proposes a completely bio-inspired approach for liver cancer detection using CT scans, which stands in contrast to contemporary systems based on either feature engineering methods or hybrid procedures. It integrates BER-optimized deep learning models for the analysis of lesions. The key contributions of this paper include:

• A comprehensive survey introducing current state-of-the-art procedures for diagnosing liver cancer and other malignancies.

• The proposal of a novel hybrid segmentation method using the Seg-Net network [13], the UNet [14], and BER for CT images. This approach employs the SegNet network to identify liver tissue in an abdominal CT image, and the UNet network to detect liver lesions within that tissue.

• The tuning of the deep learning network's hyperparameters, which significantly influence its segmentation performance, using BER optimization components. As a result, the proposed method can yield near-optimal segmentation results when applied to liver lesions, compared to best-in-class procedures.

The rest of the paper is structured as follows: Section 2 elaborates on the proposed methodology, while Section 4 presents the results of the trials and their discussion. The study concludes in Section 5 with a discussion on future research prospects.

2 Related Works

A substantial amount of research has been conducted in the field of liver cancer detection using advanced computational techniques. This section provides a review of the significant contributions made by various researchers.

Zhang et al. [15] developed an innovative FSVM+ based method that employs a feature SVM+ framework for the transfer learning problem. With a specific focus on increasing the gap between classes, the FSVM+ transformation matrix was utilized to minimize the encompassing data ball's dimension. They cleverly assigned appropriate weights to each Contrast-Enhanced Ultrasound (CEUS) image by calculating the maximum mean divergence. Their experimental outcomes, based on a dataset of bi-modal images from liver cancer patients, indicated an impressive improvement in the Computer Aided Design (CAD) model's performance.

Zhao et al. [16] proposed a pioneering machine learning methodology aimed at constructing an automated Hepatocellular Carcinoma staging system that incorporates a significantly larger number of clinical features than existing systems. Their approach, based on random survival trees, utilised B-splines to transform functions into vectors in a low-dimensional space, allowing for the grouping of similar patients into staging cohorts. The performance of their final staging system significantly surpassed the Barcelona Clinic Liver Cancer (BCLC) system in differentiating patients at diverse stages.

In the pursuit of early disease prediction using Machine Learning (ML) methods, Dritsas and Trigka [17] evaluated various ML models and Ensemble techniques. Their results demonstrated that the Voting Classifier outperformed other models in predicting liver disease occurrence, achieving an Area Under the Curve (AUC) of 88.4%, accuracy of 80.1%, recall of 80.1%, and F-measure of 80.1%.

Deshmukh et al. [18] concentrated on the examination of images of two subtypes of cancer. Their framework, evaluated with 2871 images, achieved remarkable precision using a dual hybrid model. Their approach involved a result prioritizer that decided the most suitable model for image analysis based on the outputs of both neural networks. This deep learning system offered valuable insights into the defining features contributing to predictions.

Md et al. [19] employed a variety of data preparation techniques in their model, including accuracy with imputation to fill in gaps left by missing values. Their approach involved the application of the Log1p transformation on skewed columns, followed by normalization, and the use of feature selection techniques such as univariate analysis, feature importance, and correlation matrices. Their model, trained on enhanced preprocessed data using ensemble learning techniques such as Random Forest, Extra Tree, and Stacking, achieved a testing accuracy of 86.06%. This result outperformed those of prior models, thereby emphasizing the efficacy of their approach in identifying liver disease.

Huang et al. [20] exploited Raman spectroscopy to examine human hepatic tissue samples for in vitro cancer. The authors demonstrated the potential of Raman spectroscopy, coupled with cancer tissue identification, including subtype and stage, in differentiating carcinoma tissues from surrounding non-tumor tissues in a rapid, non-invasive,

and label-free manner. The potential of a portable Raman device for real-time intraoperative diagnosis of human liver cancer was also highlighted.

Further, the correct diagnosis and determination of a patient's survival time rely critically on effective treatment [21–23]. Histopathological images, often deemed the gold standard for diagnosing liver cancer, accurately represent various stages of the disease. Research into the intelligent categorization of histological images of liver cancer with differing degrees of differentiation can significantly benefit liver cancer patients, despite the existing challenges associated with this categorization.

The aforementioned studies demonstrate the extensive application of machine learning and computational techniques in the detection and diagnosis of liver cancer. This paper seeks to build upon these contributions by proposing a bio-inspired approach for liver cancer detection using CT scans and BER-optimized deep learning models.

3 Methodology

In this section, we delve into the databases leveraged by our proposed model. A comprehensive examination of the proposed methodology is also presented. Figure 1 provides a high-level schematic representation of the proposed method's architecture.
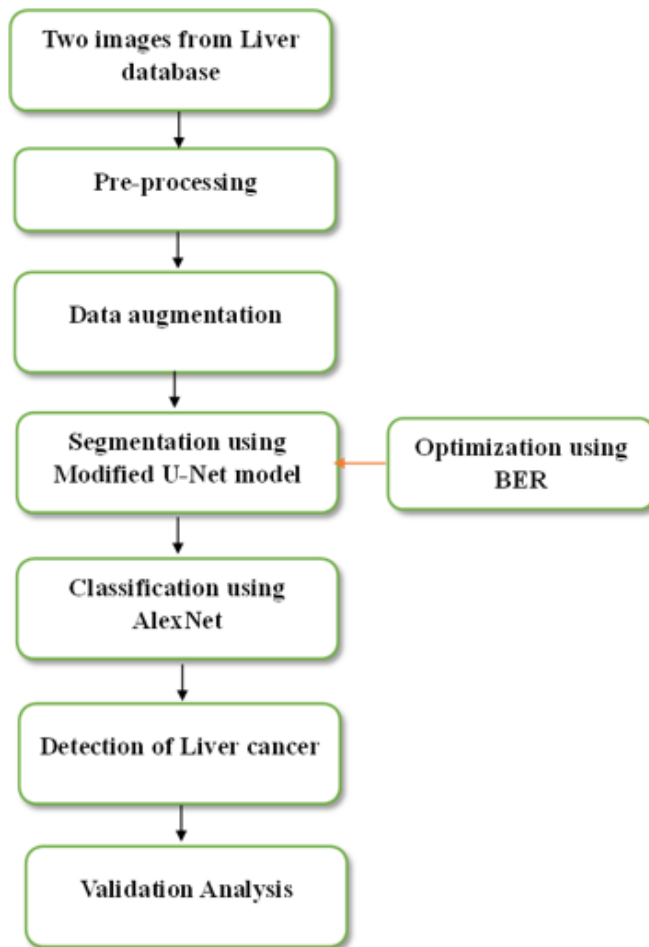
Figure 1. Working flow of the research work

3.1 Dataset Description

3.1.1 LiTS17 dataset

Our proposed approach was evaluated using the LiTS17 dataset [24], specifically designed for the tumor lesion problem. This dataset provides a training set comprised of 130 CT scans.

3.1.2 3D-IRCADb-01 dataset

The 3D-IRCADb-01 dataset incorporates 3D CT images from twenty patients, with liver tumors present in 75% of these cases. The typical liver densities range from 40 to 135. Each image is a square, with a resolution of 512 by

512 pixels. Segmentation was performed on the DICOM images with labels.

In this investigation, 10% of all CT scan images were reserved for testing purposes. This test set was comprised of 26 randomly selected images, chosen using Python packages. Additionally, 10% of the remaining training set was partitioned for validation, leaving 80% for subsequent training. Consequently, the final distribution for the study comprised 80% training, 10% validation, and 10% testing from the remaining images post-validation.

3.2 CT Image Preprocessing

The proposed method is designed for the segmentation of liver tumor images for clinical application. The process involves the utilization of data augmentation, preprocessing, and a Convolutional Neural Network (CNN) to diagnose tumors in the liver and its surrounding organs. During the preprocessing phase of CT scans, focus is placed predominantly on the liver, with surrounding organs disregarded.

Next, the image undergoes histogram equalization, a process aimed at enhancing image contrast. This procedure is followed by data amplification, employing insights derived from the augmentation steps to train the necessary invariant features.

Every medical image analysis system deploys image preprocessing to enhance the quality of the initial input image. This can involve a variety of techniques, including but not limited to noise reduction and enhancement methods. The necessity of preprocessing derives from its impact on subsequent processes, which depend heavily on image quality for defining blocks and feature extraction.

During normalization and scaling processes, the image values are adjusted and the range is narrowed, facilitating the precision of the classifier. Noise reduction improves the performance of image processing operations such as edge detection, segmentation, and compression. Contemporary medical imaging can advantageously employ either a spatial domain or a spectral domain method for noise reduction.

Mean filtering is a method in which each pixel is replaced by the average of its neighbors, resulting in a softened and blurred image. The adaptive mean filtering method maintains edges and features by utilizing local image statistics, such as mean, variance, and correlation. By changing the original value to the local mean, noise is significantly reduced.

This filter learns local image characteristics and assists in the removal of noise from specific regions. The statistic filter produces less blur than the mean filter while retaining edge sharpness. The values are estimated using a maximum a posteriori filter and an unobserved signal to maximize the Bayes theorem.

Noise in medical imaging can also be reduced using the Curvelet transform, an indexed frame multi-scale transformation that allows for scale, position, and element indexing. Histogram equalization is employed in image processing to improve the visibility of organs. This process enables more precise segmentation of the liver tumor.

Volumes of tumors and liver masks are stored separately for each CT imaging slice dataset, contributing to a comprehensive and detailed data repository for the study.

3.2.1 Data augmentation

The application of data augmentation techniques allows the envisaged system to encounter a broader diversity of tumors. Figure 2 showcases several instances of images that have been enhanced through this method. In this strategy, images with new features—derived from the original set but subjected to existing procedures—are incorporated into both the training and testing datasets. This supplemental information enables the assessment of the constructed models' capacity to identify images that have undergone rotation or magnification. If the models can recognize shapes in a specific orientation—or indeed any orientation—they retain these shapes, thereby generating accurate results.

## 3.3 Proposed Hybrid Algorithm for Liver Tumor Segmentation

Various architectures are leveraged by Convolutional Neural Networks (CNN) to address classification problems. Recently, semantic segmentation has seen a notable increase in the usage of SegNet and UNet. However, the accuracy of the segmentation is intrinsically linked to the network hyperparameters. To achieve near-optimal segmentation results, adjustment of these hyperparameters is required. Selecting the appropriate hyperparameters necessitates the use of state-of-the-art optimization techniques. Such algorithms are capable of efficiently searching the space of potential solutions on a global scale. Our proposed hybrid method, termed SegNet-UNet-BER [23], amalgamates the advantages of SegNet and UNet deep learning architectures with the benefits of BER optimization to maximize efficiency.

3.3.1 Liver segmentation from the CT image using network

Abdominal CT scans encompass more than just the liver, necessitating the extraction of the liver as a crucial step in obtaining an accurate cancer diagnosis. In this context, a Convolutional Neural Network (CNN) utilizing a SegNet architecture is employed. This architecture has been effectively utilized for semantic segmentation tasks at the pixel level. The SegNet architecture is predicated on an encoder-decoder mechanism that culminates in a final layer.

The encoder constituent of SegNet is constructed from a sequence of stacked layers, with a max-pooling layer separating the convolutional layers within each cluster. This is demonstrated by the first 13 convolutional layers of the VGG16 architecture. In these layers, input from a filter bank undergoes convolution to generate the required number of feature maps. Subsequently, the resulting feature maps are subjected to batch normalization. The pixel-wise operation, the Rectified Linear Unit (ReLU) procedure, is then enacted, with the output being $max(0, k)$. A max-pooling layer is utilized here to achieve downsampling by a factor of 2, as evidenced by the use of a 2x2 window and a stride of 2. SegNet relies extensively on max-pooling to attain translational invariance. However, the loss of boundary information during segmentation presents a challenge with this approach. Prior to the implementation of the max-pooling operation, the feature maps of the encoder are indexed to retain boundary information. In practice, the location of the highest value pixel in the feature map window is recorded.

The SegNet decoder mirrors the layer structure of the encoder but operates in reverse. The input maps are initially upsampled, resulting in a sparse feature map using the remembered max-pooling indices. The decoder's filter banks then convolve to generate a dense feature map. Batch normalization is applied subsequent to the convolution, as is the case in the encoder. In the final stage of the decoder, the pixels are activated with a softmax activation function before being relayed to the output layer. The desired segmentation is achieved by assigning each pixel to a specific category.
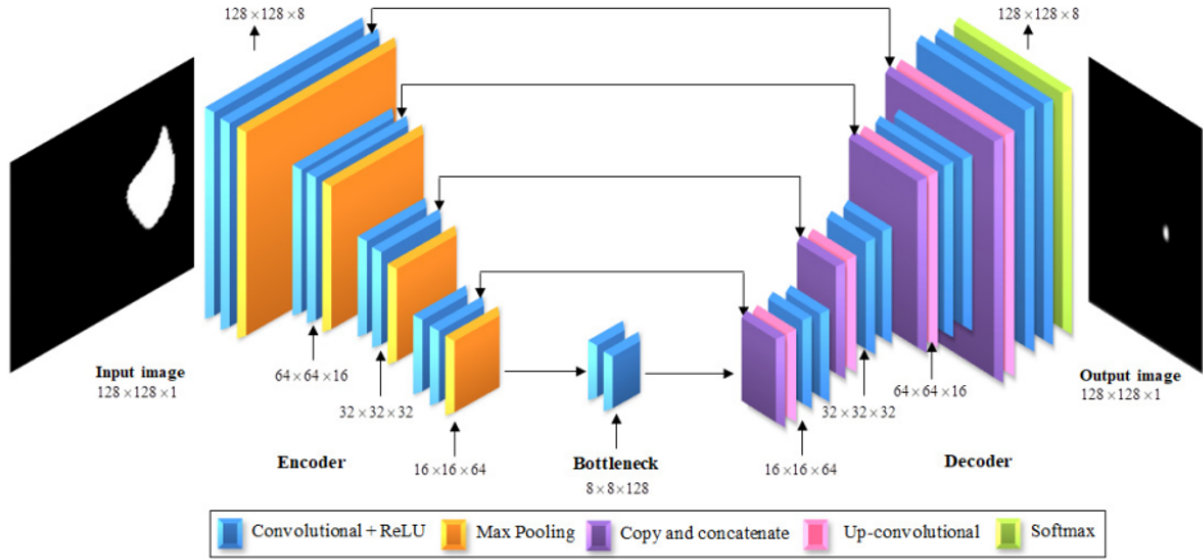
Figure 2. Illustration of the SegNet-UNet-BER method, using UNet for liver lesion segmentation in CT images, with BER optimizing the segmentation results

3.3.2 Lesion segmentation from the liver tissue utilizing the UNet system

The isolation of lesions within liver tissue is a critical initial step for further examination. For this purpose, the UNet architecture is utilized. This framework has demonstrated effectiveness when applied to medical images. The input layer, as displayed in Figure 2, accepts liver images in a 128×128×1 format. The UNet architecture is composed of three main components.

In the downsampling pathway, a max-pooling layer with a 2×2 window size and a stride value of 2 succeeds two convolutional layers. The input liver image undergoes two convolutions with a 3×3 filter, followed by activation with a ReLU filter. The output image retains the dimensions of the initial image, hence the padding value remains unchanged. Commencing with a value of eight in the convolution layer of the first group, the number of filters is incremented by a factor of two for each subsequent layer until the fifth layer.

Subsequently, upsampling is performed by halving the sample size of each group's feature maps. Within the UNet design, features from the layer are amalgamated into a single concatenation layer with the equivalent number of the group. This is followed by a ReLU activation function and then a pair of layers with a 3×3 convolutional filter. This series of layers is replicated from the sixth group through to the ninth. The tenth and final layer comprises a convolutional layer with a 1×1 filter and eight feature channels. In total, this architecture encompasses 27 layers: 18 convolutional layers paired with ReLU levels, 4 pooling layers, 4 up-convolution layers, and a solitary softmax layer.

3.3.3 Optimization of segmentation presentation using the BER procedure

The BER method initially generates solution vectors. All feasible hyperparameter optimization values are included in each created vector. These numbers are used as inputs to the SegNet network's training procedure. Eq. (1) is used to calculate the fitness value of the BER's output hyperparameter vectors. To do this, we compare the predicted image P to the ground truth image G and calculate the contour matching score (C score). Therefore, improving the F1-score, precision, and recall will determine the best approach for liver the abdominal CT scan. Three different masses w_f, w_p, and w_r were recall, correspondingly.

$$\text{Fitness } = C - \text{score}(P, G) \tag{1}$$

$$C - \text{score}(P, G) = w_f(F1 - \text{ score }) + w_p(\text{ Precision }) + w_r(\text{ Recall }) \tag{2}$$

The precision and recall are totalled as shadows,

$$\text{Precision } = \frac{\text{True Positive}}{\text{True Positive } + \text{ False Positive}} \tag{3}$$

$$\text{Recall } = \frac{\text{True Positive}}{\text{True Positive } + \text{ False Negative}} \tag{4}$$

(1) Basic concepts and formulation

The goal of optimization procedures is to locate the best answer to a problem with a given set of restrictions. In BER, a vector can be used to represent a single member of the population, $\vec{S} = \{S_1, S_2, \ldots; S_d\} \in R_d$, where, d is the size and S_i is a parameter or feature in the optimization issue. The proposed method uses a fitness function f to assess an individual's performance up to point. The following stages of the optimization process are used to explore populations for a certain start with a collection of randomly chosen people (solutions). The fitness function, size are necessary for the optimization process to start.

(2) Exploration-exploitation balance

To better balance the demands of exploitation and exploration, the proposed strategy divides the population into subgroups and dynamically adapts the composition of each group. The population is divided into two groups at the outset of the process for exploration and exploitation. The exploration group makes up 70% of the population, whilst the misuse group makes up just 30% of the population. In order to increase the fitness values of people in each group, the number of persons in task is initially set at 30% and increased during the optimisation rounds to reach other hand, the exploratory group's membership falls from 70% to 30% during the course of the iterations. Thanks to this technique, the average level of people's fitness can be raised more clearly. Additionally, the elitism approach is employed if no better solution is discovered, preserving the process' leading solution to guarantee the population's convergence of the optimisation process. If utilising the BER optimisation process, another separate can be formed by performing the mutation procedure.

(3) Exploration-exploitation balance

• Exploration operation

In addition to finding interesting locations in the search space, exploration is in charge of moving away from local optima stagnation and towards the ideal answer, as will be detailed below.

• Heading towards the best solution

This method is used by the exploration group member to look for promising areas near its current location in the search space. This is done by repeatedly looking among nearby viable options for a better choice in terms of fitness value. The BER research uses the subsequent equations to achieve this.

$$\text{r} = h\frac{\cos(x)}{1 - \cos(x)} \tag{5}$$

$$\vec{D} = \vec{r_1} \cdot (\vec{S}(t) - 1) \tag{6}$$

$$\vec{S}(t + 1) = \vec{S}(t) + \vec{D} \cdot (2\vec{r_2} - 1) \tag{7}$$

where, $S(t)$ is the solution vector at iteration, 0×180, r_1 and r_2 are coefficient vectors, and their values are given by Eq. (7). h is a random number selected from the range [0, 2], where, D is the diameter of the search agent's search region for gifted areas.

• Exploitation operation

The task of improving already implemented solutions falls to the exploitation team. The BER determines the best individual by calculating the fitness values of all participants at each cycle. To achieve exploitation, the BER uses two distinct strategies, which are described in more depth in the sections following.

• Heading towards the best solution

The subsequent reckonings are used to change the search agent to the best key:

$$\vec{S}(t + 1) = r^2(\vec{S}(t) + \vec{D}) \tag{8}$$

$$\vec{D} = \vec{r_3}(\vec{L}(t) - \vec{S}(t)) \tag{9}$$

where, $\vec{r_3}$ is a chance vector intended using Eq. (7) that controls the drive ladders towards the best solution, $\vec{S}(t)$ is the repetition t, $\vec{L}$ is the finest solution course, and $\vec{D}$ refers to the distance vector.

• Investigating area around best solution

The area surrounding the best answer (leader) is the most talented. As a result, search close to the greatest option in the hopes of discovering a better one. The BER uses the subsequent equation to carry out this way.

$$\vec{S}(t + 1) = r(\overrightarrow{S*}(t) + \vec{k}) \tag{10}$$

$$\vec{k} = Z + \frac{2 \times t^2}{N^2} \tag{11}$$

where, $\overrightarrow{S*}$ refers to the best answer, where t is the iteration sum and N is the total sum of iterations, and z is a random number between [0, 1].

• Mutation operation

The BER uses the mutation as one more study method. The genetic operator responsible for maintaining and generating population variety. It may be seen as a local, probabilistic random disturbance of one or more distinct components. By aiding in the avoidance of local optima, it helps prevent early convergence; such a change in the search field serves as a launching pad for another intriguing subject. The strong exploration capability of the BER is affected by the mutation, in fact.

$$\vec{S}(t + 1) = \vec{k} * z^2 - h\frac{\cos(x)}{1 - \cos(x)} \tag{12}$$

• Selection of the best solution

The BER chooses the best choice for the next iteration in order to ensure the quality of the solutions identified. Although the elitism technique increases algorithm efficiency, it also increases the risk of multimodal functions convergent too soon. It is interesting that the BER offers exceptional exploration capabilities by using a mutation strategy and searching close to members of the exploration group. The BER can prevent early convergence because to its excellent exploration capabilities. Iterations, population size, mutation rate, and other input parameters are initially given to the BER. The people are then split up into the exploration group and the exploitation group by the BER. The BER method continuously controls the number of participants in each group while iteratively seeking the optimal solution. To complete their responsibilities, each group chooses one of two methods. To guarantee variety and high exploration, the BER randomly ranks findings after each cycle. A solution from the exploration group in one iteration can, for instance, join the exploitation group in the next. The elitist mindset of the BER aids in maintaining the leader over the iterations. Table 1 lists the hyper-parameters that the proposed model picked.

3.4 Classification Using Pre-trained AlexNet Model

The AlexNet model is suggested as the dependable foundation of the suggested model. We chose AlexNet over other pre-trained models since we want to work with a straightforward model and test performances without sacrificing memory and testing time. AlexNet employs two distinct ideas. Both models have roughly the same

amounts of layers, neurons, and filters of the same size. To address the issue with AlexNet, a better framework is presented by adding batch normalisation functions:

1. ReLU's insurmountable barrier, which can vanish and never act on point.

2. Adjust the local response normalisation (LRN) exercised in the conventional AlexNet's normalisation effect. Since batch normalisation (BN) is trainable but LRN is not, using the latter produces more encouraging results than using LRN.

The AlexNet model's repetition of the layers is due to the efficient use of the GPU for convolution and all other processing during CNN training to speed up training. Because AlexNet is distributed across two parallel GPUs, the model's processing speed is accelerated and model training time is reduced.

Table 1. Optimized hyperparameters for lesions using SegNet and UNet, correspondingly

Hyperparameter	Optimized Values	
	SegNet	**UNet**
Early learning rate	0.01	0.05
Shuffle	Once	Every epoch
Minibatch size	11	16
Momentum	0.9	0.9
Supreme epochs	30	150
12 regularization	0.0006019928	0.0004795852

3.4.1 Proposed CNN architecture

On many datasets, the AlexNet perfect can yield accuracy capacities. However, removing any one of the convolutional layers must significantly reduce AlexNet's performance. Five convolutional layers, three fully linked layers, and eight successive layers make up the original AlexNet model network structure. AlexNet is one of the greatest options for forging because of its intricate nature. Except for the final completely connected layer, which uses the softmax function, other layers employ a max-out activation function. The following are the main contributions:

1. Replace RELU with function across all AlexNet layers.

2. Swap out LRN for batch normalisation.

3. The projected construction will be employed as a classifier for the output result to notice the bogus result.

As the input, an RBG image of 227 by 227 must be utilised. Due to the significant overfitting that would have occurred without this picture size, AlexNet would have had to use much lower network layers. If an RGB picture doesn't already exist, the supplied image is converted to one. If not, the input image's dimensions will be modified to 227×227. Convolution and max-pooling with BN are performed by 96 separate 11×11-sized filters in the first convolution layer. An input image with the dimensions 227×227×3 might be subjected to a convolution:

• There are (227×227×96) output nerve cells in L, one for each of the 227×227 input "pixels" and for each of the 96 yield maps.

• There are 96 kernels total, (11×11×3) per filter (the input size processed done the kernel), and (11×11×3) weights in total.

• It is possible to connect to (227×227×3×11×11×96) connections. For each of the (227×227×96) output units, a single filter procedure the (11×11×3) input values.

Through the following four convolution layers, this is repeated in a similar manner. Every layer has distinct input and filter sizes, as well as the matching numbers of maps.

On the connected layers that were used as the discriminant and trained using dropout followed by model.

A specific aspect will be briefly explained and will centre on the key terminology used in the model in order to aid families in understanding and explaining the intended work:

1. Max Pooling (MXP): To reduce the dimension, the suggested model uses a max-pooling strategy that retains just the extreme value in the filter.

2. Dropout: Using a predetermined probability, this approach turns off individual nodes. For the suggested AlexNet model, we kept the dropout rate at 50%. The 50% dropout rate was chosen since it will provide the model the most regularisation. This is so because a loss function that follows a distribution is minimised using the dropout.

3. Softmax Activation Function: Given a neuron, each neuron's softmax value results in an output as shown in Eq. (13).

$$y_j = \frac{e^{x_j}}{\sum_{k=1}^{k=K} e^{x_k}} \tag{13}$$

where, K is the length of x, j stands for the output components, so j=1, 2..., K, and x is the layer.

4. Max-out Function: The projected model substitutes for ReLU because it is known to hasten the junction of big datasets. Eq. (14) can be used to represent the max-out function as follows.

$$\max\left(w_1^T x + b_1, w_2^T x + b_2\right) \tag{14}$$

where, the input vector x, the matrix w, and the bias b are present. A well-known learning activation function is max-out. Known as a maxed-out special variant, ReLU. ReLU is an easy-to-train and simple-to-implement piecewise linear function. The model can be trained more quickly thanks to ReLU. As a function has none of the drawbacks (dying ReLU) and all the advantages of a ReLU.

5. A CNN that homogenises inputs for each mini-batch is trained using batch normalisation. As a result, the learning process is required to train CNNs is drastically reduced. The advantages of BN (ICS) and quickening network training. Before the output is passed to the activation function in BN, the following processing is done on the output:

i. Set the entire batch B to have a mean of 0 and a variance of 1.
• Compute the mean of the whole batch yield: μ_B
• Compute the alteration of the whole batch output: σ_B^2
• Regularize the batch by deducting the mean and in-between by the alteration.
ii. Suggest two training limits (γ: for scaling and β: for instable).
iii. Smear the scaled and shifted regularized batch to a purpose.

By generating μ_B and σ_B^2 for a channel, batch normalisation normalises inputs x_i before creating the normalised activation as in Eq. (15).

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \tag{15}$$

where, μ_B is used to increase stability if the mini-batch's variation is small.

In the training process, mean and variance are calculated across the entire training dataset. By the conclusion of network training, the calculated mean and variance are preserved as properties. This contrasts with Local Response Normalization (LRN), which is a non-trainable layer. LRN primarily square-normalizes a feature map within a given neighbourhood. LRN reduces uniformly high activations within these neighbourhoods, leading to an enhancement of contrast in the feature map. This process is based on the principle of lateral inhibition, which involves achieving local maximum contrast. It's essential to note that LRN does not provide a regularization effect, whereas Batch Normalization (BN) does. The differences between LRN and BN are elaborated on in Table 2.

Table 2. Difference among local response normalization

| Normalization Category	Trainable	Regularization	# of Trainable Limits
LRN	No	No	0
BN	Yes	Yes	2

4 Results and Discussion

The existing procedures are considered and applied with our datasets and then, results are averaged in Table 3 and Table 4.

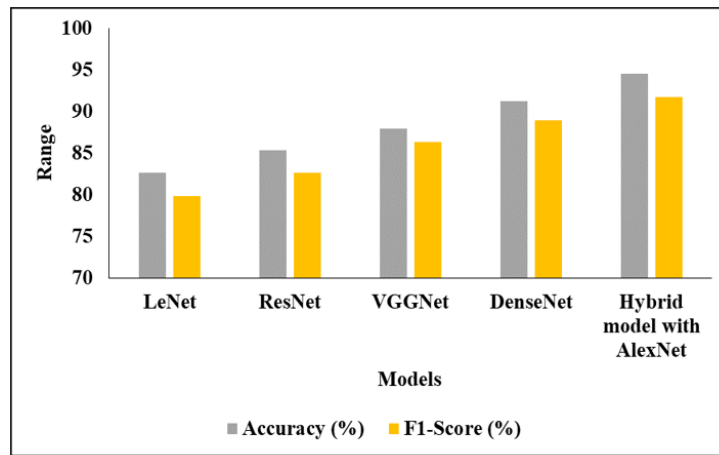**Table 3.** Analysis of projected model on LiTS17 dataset

Model	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
LeNet	81.40	78.41	82.68	79.88
ResNet	83.57	81.71	85.38	82.63
VGGNet	87.14	85.62	87.90	86.37
DenseNet	88.81	89.11	91.23	88.96
Hybrid Model with AlexNet	**90.78**	**92.62**	**94.53**	**91.69**

Table 3 illustrates an analysis of the projected model on the LiTS17 Dataset. The evaluation considers various models, and their respective performance metrics are as follows: LeNet Model: This model achieved a precision rate of 81.40%, a recall value of 78.41%, an accuracy value of 82.68%, and an F1 score of 79.88%. ResNet Model: The ResNet model performed better, with a precision rate of 83.57%, a recall value of 81.71%, an accuracy value

of 85.38%, and an F1 score of 82.63%. VGGNet Model: This model demonstrated further improvement, achieving a precision rate of 87.14%, a recall value of 85.62%, an accuracy value of 87.90%, and an F1 score of 86.37%. DenseNet Model: The DenseNet model surpassed the others so far, with a precision rate of 88.81%, a recall value of 89.11%, an accuracy value of 91.23%, and an F1 score of 88.96%. Hybrid Model with AlexNet: The Hybrid model, when combined with AlexNet, outperformed all the previous models. It reached a precision rate of 90.78%, a recall value of 92.62%, an accuracy value of 94.53%, and an F1 score of 91.69%. This comparative analysis shows that the proposed Hybrid model, in combination with AlexNet, yields superior results compared to the other models (See Figure 3 and Figure 4).
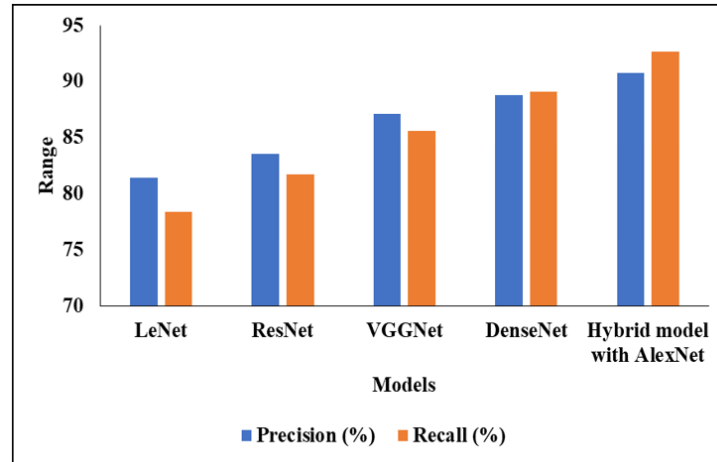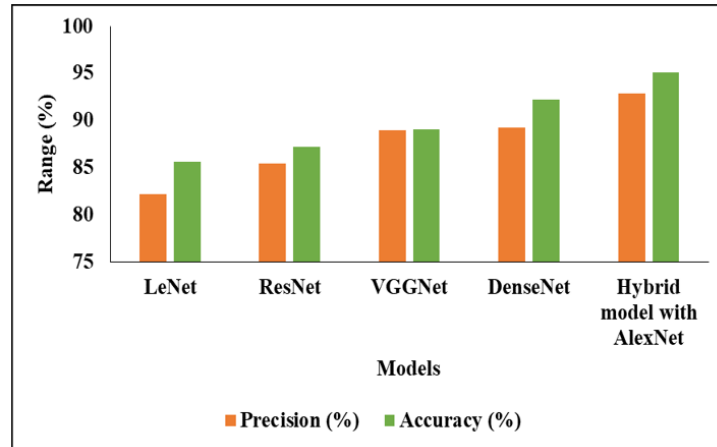

Figure 3. Analysis of various models

Figure 4. Comparison performance of projected model

**Table 4.** Analysis of projected model on 3D-IRCADb-01 dataset

Model	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
LeNet	82.22	80.10	85.63	81.15
ResNet	85.41	82.30	87.18	83.83
VGGNet	88.98	87.80	89.08	88.39
DenseNet	89.20	91.78	92.25	90.47
Hybrid Model with AlexNet	**92.87**	**94.32**	**95.05**	**93.59**

Table 4 presents an examination of the projected performance on the 3D-IRCADb-01 Dataset. This analysis assesses several distinct models. The LeNet model achieved a precision rate of 82.22%, a recall value of 80.10%, an accuracy value of 85.63%, and an F1-score of 81.15%. The ResNet model demonstrated a higher precision rate of

85.41%, a recall value of 82.30%, an accuracy value of 87.18%, and an F1-score of 83.83%. The VGGNet model showed further improvement, with a precision rate of 88.98%, a recall value of 87.80%, an accuracy value of 89.08%, and an F1-score of 88.39%. The DenseNet model surpassed the others, achieving a precision rate of 89.20%, a recall value of 91.78%, an accuracy value of 92.25%, and an F1-score of 90.47%. Finally, the Hybrid model, coupled with the AlexNet model, outperformed all the previous models. It reached a precision rate of 92.87%, a recall value of 94.32%, an accuracy value of 95.05%, and an F1-score of 93.59%. These results illustrate the comparative performance metrics of different models when applied to the 3D-IRCADb-01 Dataset (See Figure 5 and Figure 6).

Figure 5. Validation analysis of different models
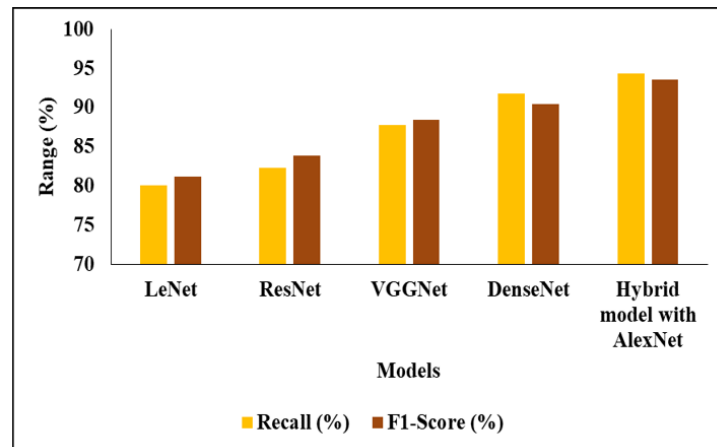
Figure 6. Graphical representation of various techniques

Table 5. Time required to obtain results using the proposed method on various hardware platforms

Platform	Time Obligatory to Get Result [s]
CPU, i3 processor, 8GB RAM	0.293
CPU, i5	0.193
CPU, I7	0.191
GPU, Nvidia K80	0.0026

Table 5 depicts the time duration required to yield results using the proposed system architecture across various hardware platforms. This study evaluates the performance of the system on distinct platforms. The platform equipped with a CPU, specifically an i3 processor with 8GB RAM, took 0.293 seconds to produce the result. Subsequently, a platform with a CPU, an i5 processor, and 8GB RAM required 0.193 seconds. Additionally, a platform with a CPU, an i7 processor, and 8GB RAM necessitated a time of 0.191 seconds. The results obtained from the i5 and i7 processor-based platforms are nearly identical, suggesting similar performance levels for these CPUs. Lastly, a

platform employing a GPU, namely the Nvidia K80, exhibited a significantly quicker result time of 0.0026 seconds. This stark contrast in result times underscores the superior performance of the GPU platform in this particular application.

5 Conclusion

In this investigation, a unique methodology for liver lesion identification was introduced, predicated on the hybridization of multiple models alongside the BER optimization algorithm. This novel method, termed SegNet-UNet-BER, was proposed for the extraction of liver lesions from CT images. It represents a synthesis of SegNet, UNet, and the BER process. The BER algorithm was employed to refine the deep learning architectures, aiming to maximize the efficacy of liver lesion segmentation. In contrast to preceding attempts in liver cancer detection which utilized various approaches, the suggested methodology harnesses the AlexNet architecture of CNN as both a feature extractor and a classifier. Two openly accessible datasets were scrutinized in this research.

The potential of the proposed SegNet-UNet-BER algorithm in the segmentation and classification of CT images was first evaluated by contrasting its performance with existing state-of-the-art segmentation techniques. Future investigations into liver cancer detection will aim to incorporate multiple modalities, including ultrasound and CT imaging, to formulate a multimodal diagnostic method underpinned by deep learning. By leveraging the advantages offered by medical imaging, this approach is hypothesized to bolster diagnostic confidence. The significance of our findings in the context of liver lesion detection and categorization establishes a compelling case for further exploration and refinement of this methodology.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] S. S. Kherde and A. A. Gurjar, "Detection of abnormalities in liver using image processing techniques," *Int. J. Creat. Res. Thoughts.*, vol. 9, no. 6, pp. 238–242, 2021.

[2] V. Hemalatha and C. Sundar, "Automatic liver cancer detection in abdominal liver images using soft optimization techniques," *J. Ambient. Intell. Humaniz. Comput.*, vol. 12, pp. 4765–4774, 2021. https://doi.org/10.1007/s12652-020-01885-4

[3] V. Gavini and G. R. J. Lakshmi, "CT image denoising model using image segmentation for image quality enhancement for liver tumor detection using CNN," *Trait. Signal*, vol. 39, no. 5, pp. 1807–1814, 2022. https://doi.org/10.18280/ts.390540

[4] A. Krishan and D. Mittal, "Ensembled liver cancer detection and classification using CT images," *J. Eng. Med.*, vol. 235, no. 2, pp. 232–244, 2020. https://doi.org/10.1177/0954411920971888

[5] H. V. S. L. Inamanamelluri, V. R. Pulipati, N. C. Pradhan, P. Chintamaneni, M. Manur, and R. Vatambeti, "Classification of a new-born infant's jaundice symptoms using a binary spring search algorithm with machine learning," *Rev. Intell. Artif.*, vol. 37, no. 2, pp. 257–265, 2023. https://doi.org/10.18280/ria.370202

[6] G. Mahalaxmi, T. Tirupal, and S. Shanawaz, "Liver cancer detection using various image segmentation approaches: A review," *IUP J. Telecomm.*, vol. 13, no. 3, pp. 48–61, 2021.

[7] S. Thulasidass, D. V. Soundari, S. Chinnaparaj, and R. Naveen, "Liver tumor diagnosis by using hybrid watershed segmentation method," *Materials Today: Proceedings*, vol. 37, pp. 2848–2857, 2021. https://doi.org/10.1016/j.matpr.2020.08.660

[8] M. Sato, R. Tateishi, Y. Yatomi, and K. Koike, "Artificial intelligence in the diagnosis and management of hepatocellular carcinoma," *J. Gastroenterol. Hepatol.*, vol. 36, no. 3, pp. 551–560, 2021. https://doi.org/10.1111/jgh.15413

[9] Y. Ikeda, K. Doma, Y. Mekada, and S. Nawano, "Lesion image generation using conditional GAN for metastatic liver cancer detection," *J. Eng. Med.*, vol. 9, no. 1, pp. 27–30, 2021. https://doi.org/10.18178/joig.9.1.27-30

[10] M. Maqsood, M. Bukhari, Z. Ali, S. Gillani, I. Mehmood, S. Rho, and Y. Jung, "A residual-learning-based multi-scale parallel-convolutions-assisted efficient CAD system for liver tumor detection," *Mathematics*, vol. 9, no. 10, pp. 1–15, 2021.

[11] Y. A. Ayalew, K. A. Fante, and M. A. Mohammed, "Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method," *BMC Biomed. Eng.*, vol. 3, 2021. https://doi.org/10.1186/s42490-021-00050-y

[12] A. Kaur, A. P. S. Chauhan, and A. K. Aggarwal, "An automated slice sorting technique for multi-slice computed tomography liver cancer images using convolutional network," *Expert Syst. Appl.*, vol. 186, p. 115686, 2021. https://doi.org/10.1016/j.eswa.2021.115686

[13] A. Rehman, M. A. Butt, and M. Zaman, "Liver lesion segmentation using deep learning models," *Acad. Trans. AI Mach. Learn.*, vol. 1, no. 1, pp. 61–67, 2022. https://doi.org/10.56578/ataiml010108

[14] N. Nanda, P. Kakkar, and S. Nagpal, "Computer-aided segmentation of liver lesions in CT scans using cascaded convolutional neural networks and genetically optimised classifier," *Arab. J. Sci. Eng.*, vol. 44, pp. 4049–4062, 2019. https://doi.org/10.1007/s13369-019-03735-8

[15] H. L. Zhang, L. H. Guo, J. Wang, S. H. Ying, and J. Shi, "Multi-view feature transformation based SVM+ for computer-aided diagnosis of liver cancers with ultrasound images," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 3, pp. 1512–1523, 2023. https://doi.org/10.1109/jbhi.2022.3233717

[16] Z. Y. Zhao, Y. C. Tian, Z. Yuan, P. Zhao, F. Xia, and S. Yu, "A machine learning method for improving liver cancer staging," *J. Biomed. Inform.*, vol. 137, p. 104266, 2023. https://doi.org/10.1016/j.jbi.2022.104266

[17] E. Dritsas and M. Trigka, "Supervised machine learning models for liver disease risk prediction," *Computers*, vol. 12, no. 1, p. 19, 2023. https://doi.org/10.3390/computers12010019

[18] S. P. Deshmukh, D. Choudhari, S. Amalraj, and P. N. Matte, "Hybrid deep learning method for detection of liver cancer," *Comp. Assist. Methods Eng. Sci.*, vol. 30, no. 2, pp. 151–165, 2023. https://doi.org/10.24423/CAMES.463

[19] A. Q. Md, S. Kulkarni, C. J. Joshua, T. Vaichole, S. Mohan, and C. Iwendi, "Enhanced preprocessing approach using ensemble machine learning algorithms for detecting liver disease," *Biomedicines*, vol. 11, no. 2, p. 581, 2023. https://doi.org/10.3390/biomedicines11020581

[20] L. P. Huang, H. W. Sun, L. B. Sun, K. Q. Shi, Y. Z. Chen, X. Q. Ren, Y. C. Ge, D. F. Jiang, X. H. Liu, W. Knoll, Q. W. Zhang, and Y. Wang, "Rapid, label-free histopathological diagnosis of liver cancer based on Raman spectroscopy and deep learning," *Nat. Commun.*, vol. 14, no. 48, 2023. https://doi.org/10.1038/s41467-022-35696-2

[21] K. Alawneh, H. Alquran, M. Alsalatie, W. A. Mustafa, Y. Al-Issa, A. Alqudah, and A. Badarneh, "Livernet: Diagnosis of liver tumors in human CT images," *Appl. Sci.*, vol. 12, no. 11, p. 5501, 2022. https://doi.org/10.3390/app12115501

[22] A. Kaur, L. Kaur, and A. Singh, "Ga-UNet: Unet-based framework for segmentation of 2D and 3D medical images applicable on heterogeneous datasets," *Neural Comput. Appl.*, vol. 33, pp. 14 991–15 025, 2021. https://doi.org/10.1007/s00521-021-06134-z

[23] C. Chen, C. Chen, M. R. Ma, X. J. Ma, X. Y. Lv, X. G. Dong, Z. W. Yan, M. Zhu, and J. J. Chen, "Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism," *BMC Med. Inform. Decis. Mak.*, vol. 22, 2022. https://doi.org/10.1186/s12911-022-01919-1

[24] R. M. Ghoniem, "A novel bio-inspired deep learning approach for liver cancer diagnosis," *Information*, vol. 11, no. 2, p. 80, 2020. https://doi.org/10.3390/info11020080