Liver and Tumour Segmentation Using Anchor Free Mechanism-Based Mask Region Convolutional Neural Network

(Liver and Tumour Segmentation)

Sangi Narasimhulu, Ch D V Subba Rao
Department of Computer Science and Engineering,
Sri Venkateswara University College of Engineering, Tirupathi, India

***Abstract*—An accurate liver tumour segmentation helps acquire the measurable biomarkers for decision support systems and Computer-Aided Diagnosis (CAD). However, most existing approaches fail to effectively segment tumours in the liver due to the overlapping of liver with any other organ in the image. To solve this problem, this research proposes Anchor Free with Masked Region-based Convolutional Neural Network (AFMRCNN) approach for segmenting liver tumours. The AF attains a precise localization of tumours by directly predicting the tumour location without relying on predefined anchor boxes. Standard datasets like LiTS and CHAOS are utilized to experiment with the efficiency of the proposed method. An EfficientNetB2 is performed to extract the most relevant features from the segmented data. The Deep Neural Network (DNN) is performed for the classification of liver tumours into binary classes by capturing intricate patterns and relationships in the data with the help of a non-linear activation function. The experimental results exhibit the proposed ARMRCNN method's commendable segmentation performance of 0.998 Dice Similarity Coefficient (DSC), as opposed to the existing methods, UoloNet and UNet++ + pre-activated multiscale Res2Net approach with Channel-wise Attention (PARCA) on the LiTS dataset.**

Keywords*—Anchor free; computer-aided diagnosis; deep neural network; EfficientNetB2; liver and tumor segmentation; masked region-based convolutional neural network

I. INTRODUCTION

Medical segmentation plays a vigorous role in Computer-Aided Diagnosis (CAD) by effectively improving the diagnostic performance and accuracy. This process enhances the precision of diagnosis, allowing more accurate identification and analysis of medical conditions [1], [2]. Globally, the liver disease is considered as the deadly disease and it is the predominant causes for liver cancer mortality. Liver tumour segmentation is important in the tumour phase and is a primary requirement for various radiological and surgical interventions including ablation therapy, liver transplant, etc. [3], [4]. There are various diagnosis tools like Computed Tomography (CT) scans, Magnetic Resonance Imaging (MRI), etc. which are the most extensively utilized techniques for the detection and diagnosis of hepatic cancer [5]. Among these, CT scans are majorly utilized to diagnose and provide highly detailed images of the body's internal structures and soft tissues. This high level details help accurately diagnosing various conditions. Different CAD solutions have been examined to aid radiologists in decision-making with diagnostic effectiveness. Liver segmentation is the most complex phase of CAD systems and hence plays a pivotal in identifying the success of diagnosis [6]. This process allows doctors to identify tumours with similar appearances in medical images and assists in developing tailored treatment pathways [7]. The reliability and precision of the segmentation approaches are significant for acquiring the clinically relevant boundary, as well as the volumetric calculations in the stage of liver tumour [8].

The structural data of shape, size and location are obtained from the segmented liver areas which offer a helpful understanding for disease assessment and treatment planning [9],[10]. Thus, introducing the automatic and accurate approaches for liver tumour segmentation has fascinated to an enhancing consideration with a crucial worth in clinical practice [11], [12]. Additionally, researchers have introduced various CAD and Deep Learning (DL) approaches to aid radiologists in understanding the CT images [13],[14]. Simultaneously, the CAD systems detect Regions of Interest (RoI) and provide the probability of these areas being specific types of lesions of either malignant or benign [15]. However, most of the existing approaches have failed to effectively segment tumours in the liver due to the overlapping of the liver with any other organ in the image. To overcome this problem, this research proposes the novel DL approach of Anchor Free with Masked Region-based Convolutional Neural Network (AFMRCNN) approach for segmenting liver tumours. The foremost contributions of this research are as follows:

- For liver tumour segmentation, this research proposes the AFMECNN approach. The AF directly detects the tumour location without relying on predefined anchor boxes, leading to the minimization of the existence of false positives and negatives.
- For the feature extraction process, the pre-trained architecture of EfficientNetB2 is utilized to extract the pertinent features. The EfficientNetB2 technique attains the most efficient extraction of high-quality features from images by employing effective scaling approaches.

- The Deep Neural Network (DNN) is performed to classify liver portions into two categories: Tumour and Non-tumour. The DNN attains a greater accuracy by capturing intricate patterns and relationships in data with the help of non-linear activation functions.

This research paper is further provided as follows: Section II displays the literature survey based on liver and tumour segmentation using DL approaches, Section III shows the proposed methodology, while Section IV demonstrates the results and discussion, and the conclusion of this research is given in Section V.

II. Literature Survey

In order to address the issue of liver tumour segmentation, various approaches were introduced by researchers. This section discusses the related works of the DL-based approaches for liver and tumour segmentation, along with their advantages and limitations. Table I demonstrates the advantages and limitations of the literature survey of Liver and Tumor Segmentation discussed in this research.

Zheng et al. [16] developed a UoloNet model to enhance the small target medical segmentation model for liver tumours using the LiTs17 dataset. The UoloNet model comprised three main modules namely, shared encoder module, object detection module, and mask generating module. The share encoder-decoder module generally implemented the extracting features of the image. The object detection module performed in a dual task mode with object detection and segmentation. The prediction module enhanced the segmentation accuracy by utilizing detection outputs to improve and highlight specific regions. Finally, the Intersection Over Union (IOU) metric was used in the traditional YOLO approach to increase the convergence speed of the network. However, the high amount of noise in the prediction module affected the identification of small labels in the image.

Huang et al. [17] implemented a Semi-supervised Double-cooperative Network (SD-Net) for liver tumour segmentation using a CT scan image. The SD-Net framework reduced the requirement of dense labelling in liver image segmentation. For the segmentation process, collaborative networks like V-net and 3D-ResVnet were utilized to transfer the labelled image from to the unlabelled image in the target domain. A dynamic Pseudo-label generation strategy was introduced to enhance the label qualities in unsupervised learning by collecting better-predicted masks as pseudo-labels from both network models. Nonetheless, the implemented SD-net model was memory intensive due to the high resolution of images, leading to limited scalability.

Yu et al. [18] presented a liver tumour segmentation network utilizing multi-phase CT images. To increase the importance of reciprocal data from various stages, Cross-Modal feature Guidance (CMG) and multi-feature fusion modules were developed in the model. The two modules were combined to effectively acquire the multi-phase features and improve the performance of liver tumour segmentation. The DL architecture was designed to exchange information between multiple phases and modalities accurately. However, the presented CMG model was constrained in its capability to analyse the complex details of lesions and generally focused only on segment lesions from inaccurate registration.

Kushnure et al. [19] implemented a new lightweight multi-level network through core architecture from the UNet++ network for automatic liver segmentation. The designed Pre-activated multiscale Res2Net approach with Channel-wise Attention (PARCA) block in a network model was able to extract coarse multi-scale features from various levels. The modification in the UNet++ network with a PARCA feature mechanism extracted more semantic and contextual data, thereby enhancing the performance of the decoder in the network. In addition, a customized loss function was implemented to maintain data imbalance and effectively consider complex samples in the collected data. Nevertheless, the UNet++ failed to effectively segment tumours in the liver that overlapped with other tissues.

Manjunath and Kwadiki [20] introduced a DL method for automatic liver tumour segmentation through CT scan images. For the liver tumour segmentation process, a 2D-modified ResUNet was designed to segment the affected regions in CT scan images. The 2D-modified ResUNet network was constructed by utilizing a parts encoder, the decoder, and the bridge. The encoder in the network encoded the resized image to a compact representation, then a pixel-wise fashion of an image was represented by the decoder, and finally, the bridge connected the paths of the encoder and decoder. Yet, the 2D-modified ResUNet network model was unable to identify and segment small tumours in the image.

TABLE I. Literature Survey of Liver and Tumour Segmentation

Author	Advantages	Limitations
Zheng et al. [16]	UoloNet's CNN component significantly model spatial relationships within CT images, helping it identify subtle variations in tissue texture and intensity of a small tumour.	A high level of noise in the segmented module impacted the accurate identification of small labels in the image.
Huang et al. [17]	SD-Net leveraged both labelled and unlabelled data by semi-supervised learning approach, enabled for more accurate segmentation of small amount of labelled data.	SD-net model was memory intensive because of its high resolution of images, resulted in limited scalability.
Yu et al. [18]	Cross-modal guidance mechanism enabled the network to significantly learn from balancing data over various stages, improved the segmentation performance.	CMG model was limited in its capability to determine the complex details of lesions and basically focused only on segment lesions from inaccurate registration.
Kushnure et al. [19]	Through the fusion of features from various levels, the network integrated the global and local information, enhanced the segmentation accuracy.	UNet++ was failed to significantly segment tumours in the liver that overlapped with other organs, resulted in poor performance.
Manjunath and Kwadiki [20]	The end-to-end approach removed the necessities of intermediate feature extraction steps and potentially enhanced the overall performance.	2D-modified ResUNet network model was unable to identify and segment small tumours in the image.

From this overall analysis, the limitations of the existing methods are identified as follows: the presence of noise in the prediction module, memory intensive, limited detail analysis and ineffective tumour segmentation. To solve those aforementioned problems, this research proposes the AFMRCNN approach for the effective segmentation of liver tumours. The ARMRCNN approach solves the memory-intensive problem by removing anchor-based complexity by considering effective sparse predictions.

III. Proposed Methodology

This research aims to overcome the segmentation in overlapping images. The AFMRCNN approach is proposed to effectively segment the liver tumour regions from the input images. The AF mechanism frequently utilizes key point-based predictions, provides more accurate boundary information and leads to enhanced segmentation performance. The DNN is used for the classification of segmented regions based on learned patterns by assigning the class labels to various regions in the image. The proposed method is comprised of five significant phases: data acquisition, pre-processing, liver tumour segmentation, feature extraction and classification. Fig. 1 depicts the pipeline of the proposed method.

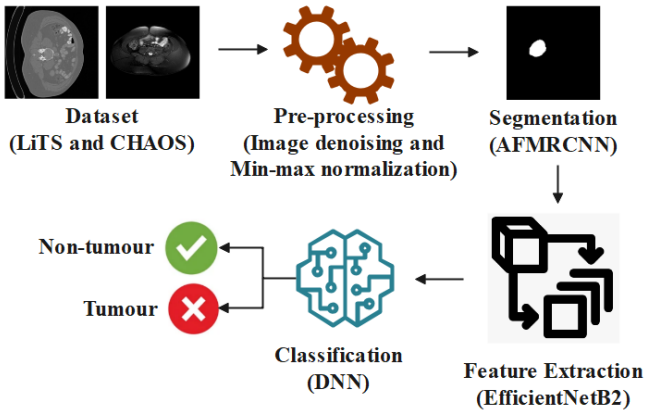


Fig. 1. Pipeline of the proposed method

A. Dataset Collection

This research estimates the efficacy of the proposed method with two publicly available standard liver segmentation datasets, Liver Tumour Segmentation (LiTS) dataset [21] and Combined (CT-MRI) Healthy Abdominal Organ Segmentation (CHAOS) dataset [22]. The detailed explanation of these datasets is discussed below.

1) LiTS dataset: The LiTS dataset involves 201-contrast enhanced 3D abdominal CT images, where 194 CT include lesions. This dataset is obtained from various scanners with scanning protocols from the clinical areas. The minimum and maximum number of axial slices in CT scans are 74 and 987, respectively. Fig. 2 depicts the sample images of the LiTS dataset.

2) CHAOS dataset: CHAOS involves 120 DICOM volumes obtained from binary MRI modalities: T1-DUAL and T2-SPIR. T1-Dual involves phases which are, in-phase and out-of-phase, each of 40 volumes. Every modality involves 20 labelled training data and 20 unlabelled test data. The labelled training data are utilized for multi-modal medical image segmentation tests and estimations. Fig. 3 depicts the sample images of the CHAOS dataset.

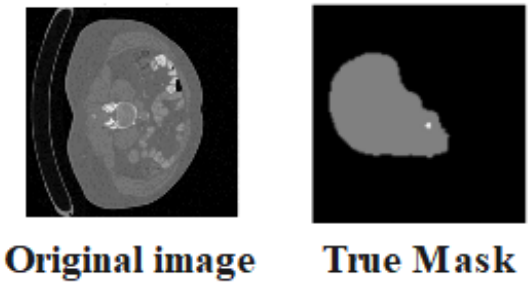


Fig. 2. Sample images of the LiTS dataset.

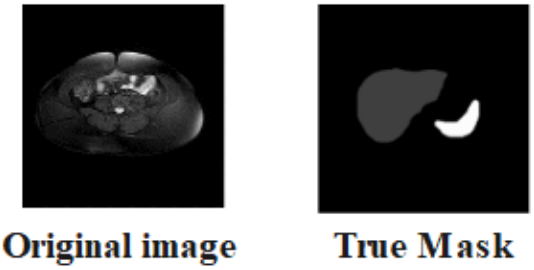


Fig. 3. Sample images of the CHAOS dataset.

Then, the collected datasets are portioned into two categories: 80% of the data are utilized for training and the remaining data for testing. Table II represents the sample number of datasets. The collected dataset is then provided for the pre-processing step.

TABLE II. Sample Number of Datasets

Dataset	Tumour	Non-Tumour
LiTS	21478	42160
CHAOS	6217	8205

B. Pre-processing

After the data collection, pre-processing is performed to modify the input data for achieving the desired results of the DL approach. The collected data contains problems of high resolution and unwanted noise which leads to inaccurate segmentation results. Hence, this research utilizes pre-processing techniques of image denoising and min-max normalization. In the image denoising process, the Gabor filtering approach [23] is utilized and is realized by performing convolution on the Gaussian function with trigonometric functions. By choosing the suitable Gabor function, various scales and directional features are detected from the collected data. This allows the utilization of Gabor filtering in image denoising and edge detection applications. By using image denoising, the Gaussian noise is removed as it produces random variations in pixel values, which obscure the significant features and structures in medical images. This denoising supports to enhance the precision of segmentation algorithms.

After denoising the image, the normalization is performed to equally balance the input data to enhance the segmentation performance because the collected data involves various units and scales. The min-max normalization [12] is a technique that is used to standardize and measure data in the direction to be taken out to the comparable range and magnitude. Then, the pre-processed data are provided to the segmentation process to effectively localize the affected regions.

C. Segmentation

The pre-processed images are fed as input to the segmentation technique to effectively segment the liver tumour regions for classification. Segmentation is a technique generally used in image processing to segment into multiple parts or regions based on the characteristics of pixels in the image. To effectively segment the tumour regions, the AFMRCNN approach is proposed. The detailed explanation of this proposed method is described in the following section. Fig. 4 illustrates the segmented images of LiTS and CHAOS datasets.

1) Anchor-free with mask region-based convolutional neural network: As compared with the other segmentation approaches, the MRCNN approach produces high-quality pixel-to-pixel masks for every case. It performs pixel-level segmentation operations and obtains better target-positioning performance in liver tumour segmentation. The MRCNN [24] approach utilizes a two-stage network model. Initially, the Region Proposal Network (RPN) generates predictions on Regions of Interest (RoI). Then, the Fully Convolutional Network (FCN) processes these RoIs to predict the binary mask, bounding box offsets, and categories for every RoIs. The network model of the MRCNN approach involves three significant types of backbone network, pixel (mask) prediction and alignment of RoI.

In the process of liver tumour segmentation, this research designs the Anchor Free with RPN (AF-RPN) approach to acquire better localization of the liver tumours. The AF-RPN architecture is an FCN to achieve the sharing computation with the binary class classification network. To acquire the tumour regions, the AF-RPN uses a sliding window mechanism applied to feature maps developed through a fusion module. This network performs 3×3 spatial window on an input feature map, and extracts the features through every sliding window. It is mapped to 512-dimensional feature vector through 3×3 kernel convolutional operation via 512 channels. These channel features are forwarded to the two parallel fully convolutional networks. One branch provides likelihoods for classification, demonstrating whether or not the regions are objects, while the other branch provides the coordinates of the boxes for localization. These are processed through the 1×1 convolutional layer and eventually, all output proposals are performed through the Non-Maximum Suppression (NMS) approaches to eliminate irrelevant proposals. Here, the Receptive Field (RF) is developed and anchor-free boxes are applied within the proposed region network to acquire parameterization for position boxes. The RF is introduced by sliding the window across fusion feature maps in CNN.

The RF for pixel is described as the rectangle region in the input data. Particularly for $k \times k$ sliding window with centroid (x, y), the RF is obtained with (x_r, y_r, w_r, h_r), where, $x_r = v \times x$ and $y_r = v \times y$ and $w_r = h_r = k \times x$ where v represents an upsampling factor of the scaling coefficient from feature map to the input data. At every feature map location, there are k anchor-free boxes in the convolutional approach, but the proposed approach involves the individual RF. Providing the convolutional feature map of size $w \times h$, obtaining the $w \times h$ RF samples in total leads to k times, not more than the present AF mechanism. In this research, the RF is introduced as the position boxes to acquire parameterized coordinates of the ground truth bounding box $t* \left(t* = \{t_x*, t_y*, t_w*, t_h*\}\right)$ and predicted box $\left(t = \{t_x, t_y, t_w, t_h\}\right)$, as formulated in Eq. (1) and (2).

$$t* = \frac{(x* - x_r)}{w_r}, t_y* = \frac{(y* - y_r)}{h_r}, t_w* = \frac{log(w*/w_r)}{t_h}, t_h* = log(h*/h_r) \quad (1)$$

$$t_x = \frac{(x - x_r)}{w_r}, t_y = \frac{(y - y_r)}{h_r}, t_w = log(w/w_r), t_h = log(h* /h_r) \quad (2)$$

Where, $x*$ and $y*$ represent the centre coordinates of ground truth box; $w*$, $h*$ denote the width and height, x, y, w and h illustrates the complements of predicted box, x_r, y_r, w_r and h_r represent the RF box which is observed as the regression from the RF to its closest ground truth box. By eliminating the anchor-based predictions, this approach simplifies the prediction process, which supports the more effective maintenance of memory when processing high-resolution images. Then, the segmented results are provided to the feature extraction process.

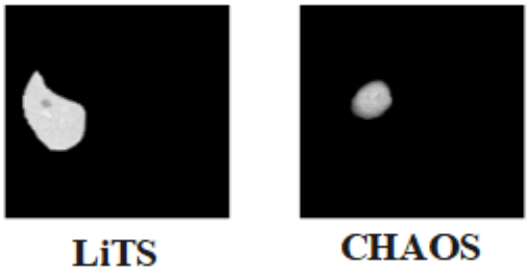


Fig. 4. Segmented images of LiTS and CHAOS datasets.

D. Feature Extraction

The segmented portions of the liver tumour are fed to the feature extraction process. Feature extraction is a process of identifying and extracting relevant features from the data, further used in classification and prediction tasks. Extracting the important features of shape, edge, texture and intensity information support the distinguishing between normal liver tissue and tumour tissue. This leads to more accurate segmentation results and outlines tumours from the surrounding healthy tissue. In this research, the pre-trained architecture of EfficientNetB2 is performed to extract the relevant (texture and shape) features. The detailed explanations of this architecture are discussed below.

1) EfficientNetB2: As compared to other pre-trained architectures, the EfficientB2 utilizes a compound scaling approach to balance the network resolution, depth and width. This process effectively scales dimensions, leading to the most effective utilization of floating-point operations and parameters. The network depth, resolution and width are reliably scaled through the Efficient Net logically by the utilization of compound coefficient. The EfficientNetB2 [25] involves arranging the model through global max pooling and the dropout layer to solve the overfitting problem and succeed through the dense layer for the binary classification. This architecture complies with the suitable loss as well as optimization functions and performs the callbacks for effective training. In EfficientNetB2, the CNN network involves various layers designed for the image processing tasks. With the output shape, the EfficientNetB2 incorporates the 7×7 spatial dimensions with 1408 channels. Ensuing the convolutional layers, the Global Average Pooling (GAP) operation minimizes the spatial dimensions while recollecting the most salient features, resulting in the size of (None, 1408). Then, the dropout layer is performed to alleviate overfitting by arbitrarily deactivating the neurons in training. Eventually, the dense layer with 1 unit is used for the extraction process with 1409 parameters. The extracted texture and shape features using EfficientNetB2 provide detailed information about tumour regions. This information supports classification by helping differentiate tumour regions from normal liver tissue and other structures. Then, the extracted 1408 features from the GAP layer are provided for the classification process.

E. Classification

After feature extraction, the features are fed as input to the liver tumour classification. In this research, the DNN approach is utilized for the classification of tumours into two types as tumour and non-tumour. The detailed description of the DNN is discussed below.

1) Deep neural network: The DNN [26] is inspired by the biological nervous network. The DNN is a robust promising approach in modelling the mechanical materials behaviour because of its powerful nonlinear mapping capability. The DNN approach comprises three significant layers of input, activation and Fully Connected (FC) layer which are utilized for the classification tasks.

A neural network requires an activation function for the final prediction of liver tumours. Rectified Linear Unit (ReLU) is the default activation function; it extends the nonlinearity to the network which provides output 0 for negative values and similar values for non-negative values. Also, Sigmoid is an activation function which is appropriate for binary classification with the output within the range of 0 to 1. The sigmoid identifies the multinomial probability distribution with two classes. The dense layer is the Neural Network (NN) layer, in which every neuron in the layer obtains an input from whole neurons of its previous layer. It utilizes the operation function to map each input with output.

In the classification process, the DNN effectively classifies the liver tumour into binary classes with the help of various layers. Hence, DNN is more robust to variations in input data like differences in tumour size, shape, and appearance, as opposed to the traditional classifiers which require effective tuning of parameters. The classified result are implanted to estimate the effectiveness of the model, and the detailed explanation result is represented in the following section.

IV. Experimental Results

The segmentation effectiveness of the liver tumour is estimated experimentally based on two standard datasets. The experiments of the proposed AFMRCNN approach are implemented on Python 3.10.12 with the system configuration of Windows 10 (64 bit) OS, intel i5 processor and 8GB RAM. The model effectiveness is estimated through various segmentation performance metrics of Volumetric Overlap Error (VOE), Dice Similarity Coefficient (DSC), Relative Volume Difference (RVD) and Intersection over Union (IoU). The classification metrices of accuracy, precision, specificity, sensitivity/recall, and F1-score are used to estimate AFMRCNN performance. The mathematical expressions of these assessment metrics are formulated in Eq. (3) to (10).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{5}$$

$$Specificity = \frac{TN}{TN+FP} \tag{6}$$

$$F1-score = \frac{2TP}{2TP+FP+FN} \tag{7}$$

$$DSC = 2\frac{|A \cap B|}{|A|+|B|} \tag{8}$$

$$VOE = 1 - \frac{|A \cap B|}{|A \cup B|} \tag{9}$$

$$RVD = \frac{|B|-|A|}{|A|} \tag{10}$$

Where, TN is True Negative, TP is True Positive, FN is False Negative and FP is False positive, while A and B represent the binary masks. Table III represents the hyper parameter settings of the proposed AFMRCNN approach.

TABLE III. Hyperparameter Settings of the Proposed AFMRCNN Approach

Hyper parameters	Values
Optimizer	Adam
Learning Rate	0.0001
Loss Function	Binary cross entropy
Batch size	32
Activation function	Sigmoid
No. of Epochs	20

A. Quantitative and Qualitative Analysis

The achievements of the proposed AFMRCNN method is estimated against the existing methods on the LiTS and CHAOS datasets. Table III demonstrates the analysis of the segmentation results. Table IV represents an analysis of the feature extraction results with an analysis of the classification results.

In Table IV, the performance analysis of the segmentation results is presented on the LiTS and CHAOS datasets. The success of the proposed AFMRCNN method is estimated and compared with conventional segmentation approaches like UNet, CNN and MRCNN approaches. When compared to these conventional approaches, the AF in the MRCNN approach that depends on the predicting key points, leading to more accurate localization of liver tumours. This approach specifically aids when dealing with irregular shapes and sizes of tumours, leading to enhanced segmentation outcomes. In LiTS dataset, the proposed AFMRCNN approach attains a better DSC of 0.978, VOE of 14.75, RVD of 11.92 and IoU of 0.98. In the CHAOS dataset, the proposed AFMRCNN approach attains a superior DSC of 0.988, VOE of 6.52, RVD of 5.32 and IoU of 0.98.

In Table V, the performance analysis of the feature extraction results is demonstrated based on the LiTS and CHAOS datasets. The efficiency of the EfficientNetB2 approach is estimated and compared with existing feature extraction approaches of ResNet, VGG19 and InceptionNet. As opposed to these approaches, EfficientNetB2 architecture permits the model to extract relevant and detailed features, capturing complex patterns and structures in the collected data with the help of the compound scaling method. In the LiTS dataset, the performed EfficientNetB2 approach attains a better accuracy of 0.997, precision of 0.998, recall of 0.978, specificity of 0.995 and F1-score of 0.987. Also on the CHAOS dataset, the EfficientNetB2 attains a better accuracy of 0.996, precision of 0.997 precision, recall of 0.988, specificity of 0.996 and F1-score of 0.992.

In Table VI, the performance analysis of the classification results is demonstrated based on the LiTS and CHAOS datasets. The effectiveness of the DNN approach is estimated and compared with the existing feature extraction methods of CNN, Simple Neural Network (SNN) and Feedforward Neural Network (FNN). When compared to these approaches, the DNN minimizes the loss of important features by allowing end-to-end learning. In the LiTS dataset, the performed DNN approach attains a superior accuracy of 0.997, precision of 0.998, recall of 0.978, specificity of 0.995 and F1-score of 0.987. Additionally on the CHAOS dataset, the DNN attains a superior accuracy of 0.996, precision of 0.997, recall of 0.988, specificity of 0.996 and F1-score of 0.992.

TABLE IV. ANALYSIS OF SEGMENTATION RESULTS

Dataset	Method	DSC	VOE	RVD	IoU
LiTS	UNet	0.874	19.80	15.45	0.90
	CNN	0.896	18.90	14.75	0.91
	MRCNN	0.912	17.65	13.80	0.93
	AFMRCNN	0.978	14.75	11.92	0.98
CHAOS	UNet	0.853	8.45	6.75	0.85
	CNN	0.875	8.60	7.85	0.86
	MRCNN	0.893	7.95	6.45	0.88
	AFMRCNN	0.988	6.52	5.32	0.98

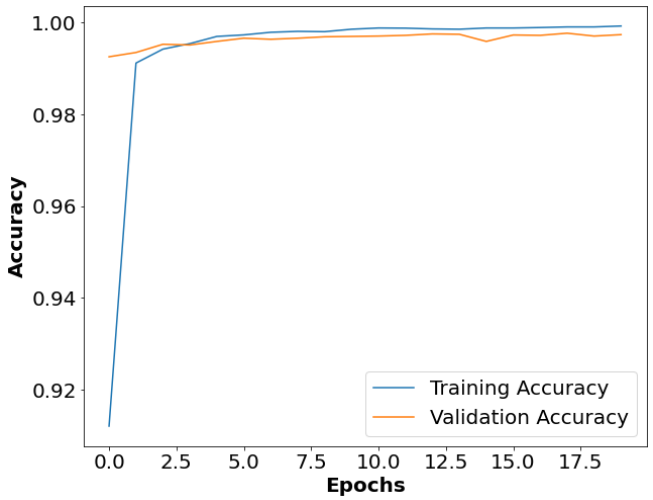
TABLE V. ANALYSIS OF FEATURE EXTRACTION RESULTS

Dataset	Method	Accuracy	Precision	Recall	Specificity	F1-score
LiTS	ResNet	0.952	0.941	0.964	0.948	0.952
	VGG19	0.961	0.955	0.972	0.960	0.961
	InceptionNet	0.844	0.832	0.857	0.841	0.844
	EfficientNetB2	0.997	0.998	0.978	0.995	0.987
CHAOS	ResNet	0.923	0.912	0.934	0.921	0.923
	VGG19	0.938	0.929	0.945	0.937	0.938
	InceptionNet	0.819	0.806	0.832	0.818	0.821
	EfficientNetB2	0.996	0.997	0.988	0.996	0.992

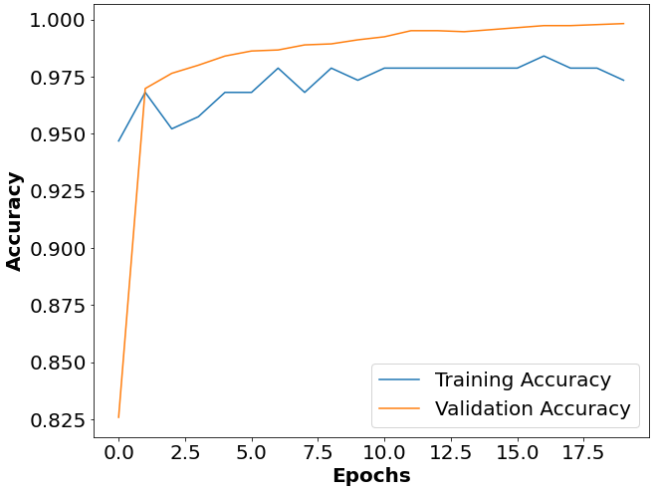
TABLE VI. ANALYSIS OF CLASSIFICATION RESULTS

Dataset	Method	Accuracy	Precision	Recall	Specificity	F1-score
LiTS	CNN	0.945	0.952	0.938	0.964	0.945
	SNN	0.889	0.895	0.874	0.870	0.884
	FNN	0.927	0.933	0.956	0.922	0.945
	DNN	0.997	0.998	0.978	0.995	0.987
CHAOS	CNN	0.918	0.921	0.905	0.934	0.913
	SNN	0.853	0.855	0.848	0.863	0.851
	FNN	0.896	0.901	0.878	0.869	0.889
	DNN	0.996	0.997	0.988	0.996	0.992

1) Accuracy function: Fig. 5(a) explains the training and validation of the accuracy for the AFMRCNN approach based on the LiTS dataset. Fig. 5(b) explains the training and validation of the accuracy for the AFMRCNN approach based on the CHAOS dataset. The accuracy of the validation is attained through the model on the 20th epoch. The training data is continuously partitioned into smaller portions, which are utilized to update the parameter of the model. The number of epochs describes the total number of iterations that the model works on the training data. Accuracy allows for the easy comparison among the models, representing a commendable model's performance, considering the classes balanced.



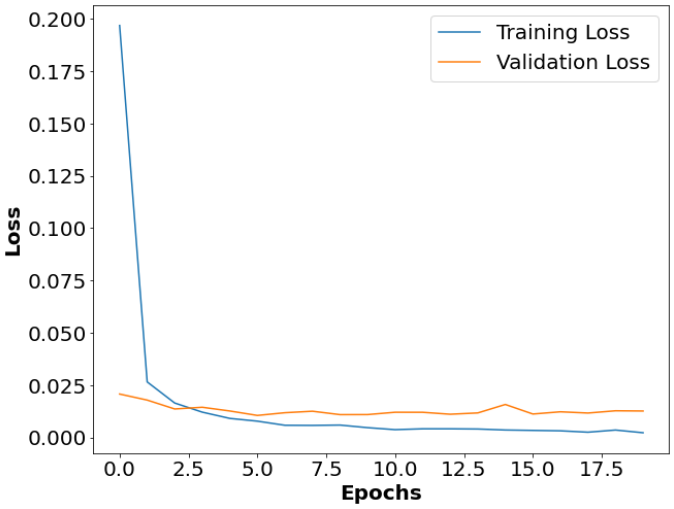
(a)



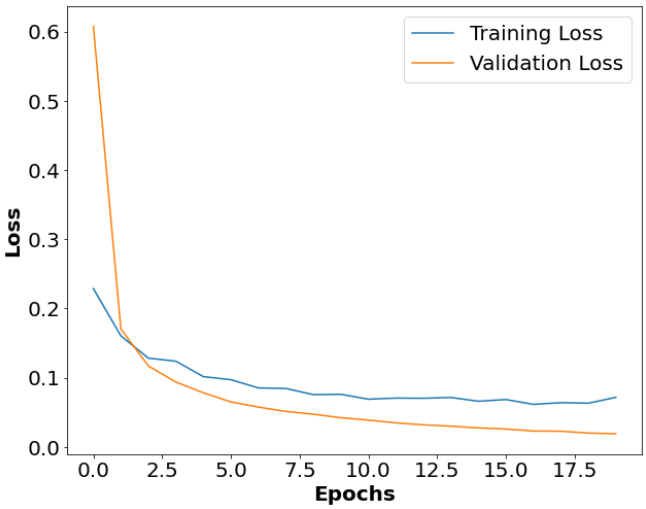
(b)

Fig. 5. Accuracy function: (a) LiTS and (b) CHAOS.

2) Loss function: Fig. 6(a) explains the training and validation of the loss function for the AFMRCNN on LiTS dataset. Fig. 6(b) explains the training and validation of the loss function for the AFMRCNN on CHAOS dataset. The minimum loss of the validation is attained through the model on the 17th epoch, while the training continues until the 18th epoch. However, after the 17th epoch, the loss was further not minimized. At every epoch, the model often changes its internal parameters by focusing on the input data when compared with the target labels to reduce the loss. This loss function indicates that the AFMRCNN has learned to simplify well to the validation data up to this point. This is a significant indicator of the performance of the AFMRCNN on the given data.



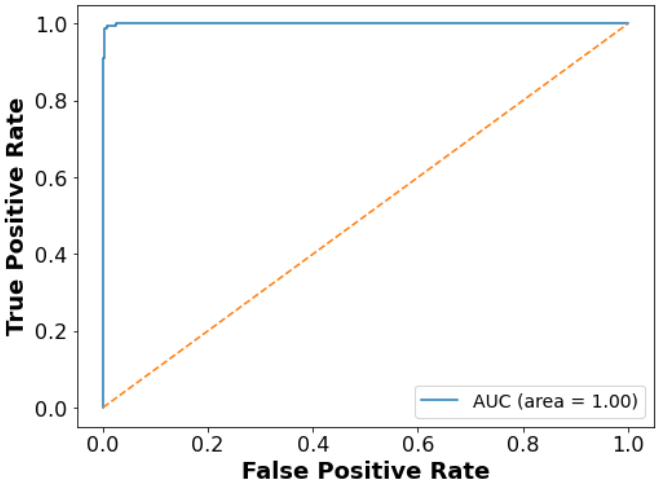
(a)



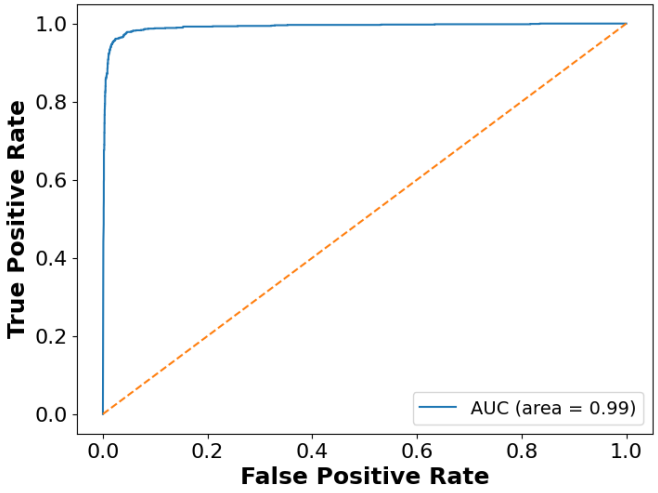
(b)

Fig. 6. Loss function: (a) LiTS and (b) CHAOS.

3) ROC Curve: Fig. 7(a) depicts the confusion matrix for AFMRCNN with an affiliation among True Positive Rate (TPR) and False Positive Rate (FPR) based on LiTS dataset. Fig. 7(b) depicts the confusion matrix for AFMRCNN with an affiliation among TPR and FPR based on CHAOS dataset. ROC is used to depict the graphical evaluation of the binary classification. The ROC curve is widely used to evaluate the performance of classifiers. In this process, the area under the ROC curve is fixed at 0.72. Estimating the ROC curve with TPR and TNR supports understanding the trade-offs between sensitivity and specificity. This supports selecting the optimal balance between correctly identifying positives and minimizing false positives, leading to a commendable AFMRCNN performance.



(a)



(b)

Fig. 7. ROC curve: (a) LiTS and (b) CHAOS.

B. Comparative Analysis

The efficiency of the proposed method is compared with the existing methods based on the LiTS and CHAOS datasets in this section. The existing methods like UoloNet [16], CMG [18] and UNet+++ PARCA [19] are compared and estimated with the proposed AFMRCNN approach. AF mechanism significantly minimizes the chances of missing small or irregularly shaped tumours which are not aligned well with predefined anchors, leading to better segmentation results. In classification tasks, the DNN is utilized for classifying liver tumours with the help of the sigmoid activation function in the final layer to produce a probability distribution over different classes. This process allowing for straightforward interpretation of the DNN confidence in every class and enabling decision-making, resulting in a superior classification of tumour and non-tumour. Table VII represents the comparative analysis on the LiTS and CHAOS datasets.

C. Discussion

The advantages of the proposed AFMRCNN and limitations of the existing works are discussed. The limitations of the previous works: In UoloNet [16], a high level of noise in the segmented module impacted the accurate identification of small labels in the image. SD-net model [17] was memory intensive because of its high resolution of images, resulted in limited scalability. CMG model [18] was limited in its capability to determine the complex details of lesions and basically focused only on segment lesions from inaccurate registration. UNet++ [19] was failed to significantly segment tumours in the liver that overlapped with other organs, resulted in poor performance. To overcome these problems, the AFMRCNN mechanism enhances the detection of small objects like tumors through reducing the dependency on predefined anchor boxes, which often introduce noise. The anchor-free approach enables the network to focus more precisely on the actual regions of interest, enhancing the small label segmentation accuracy. The anchor-free design enables the model to learn object boundaries dynamically, resulted in better segmentation.

TABLE VII. COMPARATIVE ANALYSIS USING LITS AND CHAOS DATASET

Dataset	Method	Recall	DSC	VOE	RVD
LiTS	UoloNet [16]	0.821	0.462	NA	NA
	UNet++ + PARCA [19]	0.964	0.963	0.057	0.015
	Proposed AFMRCNN	0.978	0.978	0.147	0.119
CHAOS	CMG [18]	0.927	0.928	NA	0.061
	UNet++ + PARCA [19]	0.912	0.951	0.096	0.079
	Proposed AFMRCNN	0.988	0.988	0.652	0.532

This research demonstrates that the binary classification with CNN is effectively utilized for semantic segmentation in medical diagnosis, particularly for liver tumour segmentation. Moreover, this research demonstrates that the segmentation has improved the AFMRCNN effectiveness by solving the individual network bias. This research proposes an end-to-end automatic liver segmentation by using the AFMRCNN approach. This approach conducts multi-level features and AF mechanism for decontaminating the features. ARMRCNN approach improves the quality of segmentation masks by considering them more accurately on RoI. This results in a better description of tumour boundaries, which is complex for treatment planning and assessment. Hence, the proposed AFMRCNN approach attains a better DSC of 0.978 and 0.988, VOE of 14.75 and 6.52, RVD of 11.92 and 5.32 and IoU of 0.98 and 0.98 on LiTS and CHAOS datasets. The results of this research demonstrate that the utilization of AFMRCNN enhances both accuracy and performance of medical diagnostics in the segmentation of liver tumours. The development in the expertise has the latent to improve patients' outcomes, while enabling the development of personalized treatment plans.

V. CONCLUSION

Utilizing the CT images for diagnosing the liver tumours is the most popular technique because it provides high-resolution images with fine anatomical details. Consistent liver tumour segmentation requires accurate image processing across different steps and inaccuracies in this process because of the subsequent segmentation tasks. Hence, this research proposes the AFMRCNN approach to address the problem of image overlapping for the precise segmentation of the liver tumour.

The AF mechanism attains greater segmentation accuracy, particularly for irregularly shaped liver tumours through the prediction of the presence and boundaries of tumours directly. The DNN approach is performed to classify the liver tumours into binary classes through the Adam optimizers and adjust the learning rate dynamically to enhance the convergence. The experimental results exhibit the proposed AFMRCNN approach's commendable DSC of 0.147 on the LiTS dataset as compared to the existing methods, UoloNet and UNet++ + PARCA. However, UoloNet and UNet++ + PARCA attain a minimum DSC of 0.462 and 0.963 on the LiTS dataset. Future work will involve the hybrid DL approach for enhancing the overall model results during the segmenting of the liver tumour.

REFERENCES

[1] T. Lei, R. Sun, X. Du, H. Fu, C. Zhang, and A. K. Nandi, "SGU-Net: Shape-guided ultralight network for abdominal image segmentation," IEEE J. Biomed. Health. Inf, Vol. 27, pp.1431-1442, March 2023.

[2] S. Fogarollo, R. Bale, and M. Harders, "Towards liver segmentation in the wild via contrastive distillation. Int. J. Comput. Assisted Radiol. Surg, Vol. 18, pp.1143-1149, May 2023.

[3] S. Bogoi, A. Udrea, "A Lightweight Deep Learning Approach for Liver Segmentation," Mathematics 2023, Vol.11, p. 95, December 2022.

[4] C. Hu, T. Xia, Y. Cui, Q. Zou, Y. Wang, W. Xiao, S. Ju, and X. Li, "Trustworthy multi-phase liver tumour segmentation via evidence-based uncertainty," Eng. Appl. Artif. Intell, Vol. 133, p.108289, July 2024.

[5] Z. Xia, M. Liao, S. Di, Y. Zhao, W. Liang, and N. N. Xiong, "Automatic liver segmentation from CT volumes based on multi-view information fusion and condition random fields," Opt. Laser Technol, Vol. 179, p.111298, December 2024.

[6] P. V. Nayantara, S. Kamath, R. Kadavigere, and K. N. Manjunath, "Automatic Liver Segmentation from Multiphase CT Using Modified SegNet and ASPP Module," SN Comput. Sci, Vol. 5, p.377, March 2024.

[7] M. Y. Ansari, A. Abdalla, M. Y. Ansari, M. I. Ansari, B. Malluhi, S. Mohanty, S. Mishra, S. S. Singh, J. Abinahed, A. Al-Ansari, S. Balakrishnan, and S. P. Dakua, "Practical utility of liver segmentation methods in clinical surgeries and interventions," BMC Med. Imaging, Vol. 22, p.97, May 2022.

[8] A. E. Kavur, L. I. Kuncheva, and M. A. Selver, "Basic ensembles of vanilla-style deep learning models improve liver segmentation from CT images," In Convolutional Neural Networks for Medical Image Processing Applications, (pp. 52-74), 2022. CRC Press.

[9] L. Jiang, J. Ou, R. Liu, Y. Zou, T. Xie, H. Xiao, and T. Bai, "RMAU-Net: residual multi-scale attention U-Net for liver and tumor segmentation in CT images," Comput. Biol. Med, Vol. 158, p.106838, May 2023.

[10] X. Liu, K. Ono, and R. Bise, "A data augmentation approach that ensures the reliability of foregrounds in medical image segmentation," Image Vision Comput. Vol. 147, p.105056, July 2024. https://doi.org/10.1016/j.imavis.2024.105056

[11] E. Othman, M. Mahmoud, H. Dhahri, H. Abdulkader, A. Mahmood, M. Ibrahim, "Automatic Detection of Liver Cancer Using Hybrid Pre-Trained Models," Sensors 2022, Vol. 22, pp. 5429, july 2022.

[12] A.M. Hendi, M. A. Hossain, N. A. Majrashi, S. Limkar, B. M. Elamin, M. Rahman, "Adaptive Method for Exploring Deep Learning Techniques for Subtyping and Prediction of Liver Disease," Appl. Sci. 2024, Vol. 14, p. 1488, Febrauary 2024.

[13] D. Popescu, A. Stanciulescu, M. D. Pomohaci, L. Ichim, "Decision Support System for Liver Lesion Segmentation Based on Advanced Convolutional Neural Network Architectures," Bioengineering 2022, Vol. 9, p. 467, September 2022.

[14] Y. Chen, C. Zheng, F. Hu, T. Zhou, L. Feng, G. Xu, Z. Yi, and X. Zhang, "Efficient two-step liver and tumour segmentation on abdominal CT via deep learning and a conditional random field," Comput. Biol. Med,Vol. 150, p.106076, November 2022.

[15] H. L. Elghazy, and M. W. Fakhr, "Dual-and triple-stream RESUNET/UNET architectures for multi-modal liver segmentation," IET Image Proc, Vol. 17, pp.1224-1235, March 2023.

[16] K. Zhang, L. Zhang, and H. Pan, "UoloNet: based on multi-tasking enhanced small target medical segmentation model," Artif. Intell. Rev, Vol. 57, p. 31, Febrauary 2024.

[17] S. Huang, J. Luo, Y. Ou, Wangjun shen, Y. Pang, X. Nie, and G. Zhang, "Sd-net: A semi-supervised double-cooperative network for liver segmentation from computed tomography (CT) images," J. Cancer Res. Clin. Oncol, Vol. 150, p. 79, Febrauary 2024.

[18] W. Yu, M. Wang, Y. Zhang, and L. Zhao, "Reciprocal cross-modal guidance for liver lesion segmentation from multiple phases under incomplete overlap," Biomed. Signal Process. Contro, Vol. 88, p.105561, Febrauary 2024.

[19] D. T. Kushnure, S. Tyagi, and S. N. Talbar, "LiM-Net: Lightweight multi-level multiscale network with deep residual learning for automatic liver segmentation in CT images," Biomed. Signal Process. Control Vol. 80, p.104305, Febrauary 2023.

[20] R.V. Manjunath, and K. Kwadiki, "Automatic liver and tumour segmentation from CT images using Deep learning algorithm," Results Control Optim. Vol. 6, p.100087, March 2022. https://doi.org/10.1016/j.rico.2021.100087

[21] LiTs dataset link: https://www.kaggle.com/datasets/andrewmvd/liver-tumor-segmentation. (Accessed on 30/07/2024)

[22] CHAOS dataset link: https://www.kaggle.com/datasets/anhoangvo/chaos-t1-and-t2. (Accessed on 30/07/2024)

[23] V. Dakshayani, G. R. Locharla, P. Pławiak, V. Datti, C. Karri, "Design of a Gabor Filter-Based Image Denoising Hardware Model," Electronics 2022, Vol. 11, p. 1063, March 2022.

[24] M. Zhou, J. Wang, B. Li, "ARG-Mask RCNN: An Infrared Insulator Fault-Detection Network Based on Improved Mask RCNN," Sensors 2022, Vol. 22, p. 4720, june 2022.

[25] A. M. J. Z. Rahman, M. Gupta, S. Aarathi, T. R. Mahesh, V. V. Kumar, S. Y. Kumaran, and S. Guluwadi, "Advanced AI-driven approach for enhanced brain tumor detection from MRI images utilizing EfficientNetB2 with equalization and homomorphic filtering," BMC Med. Inf. Decis. Making, Vol. 24, p.113, April 2024. https://doi.org/10.1186/s12911-024-02519-x

[26] X. Ding, X. Hou, M. Xia, Y. Ismail, and J. Ye, "Predictions of macroscopic mechanical properties and microscopic cracks of unidirectional fibre-reinforced polymer composites using deep neural network (DNN)," Compos. Struct, Vol. 302, p.116248, December 2022. https://doi.org/10.1016/j.compstruct.2022.116248.