



# Automatic Liver Segmentation from Multiphase CT Using Modified SegNet and ASPP Module

P. Vaidehi Nayantara<sup>1</sup> · Surekha Kamath<sup>1</sup> · Rajagopal Kadavigere<sup>2</sup> · Kanabagatte Nanjundappa Manjunath<sup>3</sup>

Received: 4 September 2023 / Accepted: 15 February 2024  
© The Author(s) 2024

## Abstract

Liver cancer is one of the dominant causes of cancer death worldwide. Computed Tomography (CT) is the commonly used imaging modality for diagnosing it. Computer-based liver cancer diagnosis systems can assist radiologists in image interpretation and improve diagnosis speed and accuracy. Since liver segmentation is crucial to such systems, researchers are relentlessly pursuing various segmentation approaches. A clinically viable computer-aided system requires examining multiphase CT images. However, most of the research focuses only on the portal venous phase. In this work, we developed an automatic and efficient Deep Learning (DL) method using SegNet, atrous spatial pyramid pooling module and leaky ReLU layers for liver segmentation from quadriphasic abdominal CT volumes. The proposed method was validated on two datasets, an internal institutional dataset consisting of multiphase CT and a public dataset of portal venous phase CT volumes. The Dice Coefficients (DC) obtained were greater than 96% for the latter dataset and the portal venous phase of the former. For arterial, delayed and plain CT phases of the former dataset, the DC achieved were 94.61%, 95.01% and 93.23%, respectively. Experiments showed that our model performed better than the other state-of-the-art DL models. Ablation studies have revealed that the proposed model leverages the strengths of all the three components that make it up. The promising performance of the proposed method suggests that it is appropriate for incorporation in hepatic cancer diagnosis systems.

**Keywords** Liver segmentation · SegNet · Computed tomography · Liver cancer · Computer-aided diagnosis · Atrous spatial pyramid pooling

## Introduction

Liver cancer is one of the most lethal cancers in the world. It was the third major cause of cancer mortality (approximately 830,000 deaths) in 2020 [1]. Computed Tomography (CT) is the most frequently used imaging technique for identifying hepatic cancer. Various Computer Aided Diagnosis (CAD<sub>x</sub>) solutions have been investigated to aid radiologists

in decision-making and increase diagnosis efficiency. Liver segmentation is the first and most critical stage of a CAD<sub>x</sub> system and is therefore decisive in determining the success of a diagnosis. However, liver delineation is difficult due to: (i) ambiguous boundaries with adjacent structures, (ii) large shape variability, (iii) the presence of organs with similar intensity in the vicinity, (iv) intensity variations and noise in the liver due to image acquisition and injection protocols [2] and (v) division of liver into right and left lobes.

Liver segmentation has been the subject of extensive research for over two decades. The earlier studies explored traditional segmentation methods, primarily level set, Fuzzy C-Means (FCM) and region growing. Xu et al. [3] presented a semiautomatic approach in which region growing was utilized for initial liver delineation and level set method for final refinement. Wang et al. [4] suggested a shape-intensity prior level set method using probabilistic atlas and probability map constraints. Eapen et al. [5] delineated the liver using a Bayesian probabilistic level set framework. The different swarm optimization techniques were explored in

✉ Surekha Kamath  
surekha.kamath@manipal.edu

<sup>1</sup> Department of Instrumentation and Control Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India

<sup>2</sup> Department of Radiodiagnosis and Imaging, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India

<sup>3</sup> Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India

[6–8]. Most of these methods required user intervention and were not very robust.

In recent years, the application of Deep Learning (DL) approaches for segmentation has risen rapidly. Liu et al. [9] delineated the liver using UNet and dense feature selection. Jeong et al. [10] incorporated long short-term memory network and attention mechanism into UNet. Sun et al. [11] proposed a UNet based architecture for overcoming the pitfalls in the skip connections and self attention mechanism for liver segmentation. In [12] a 3D version of UNet was developed by incorporating residual connections. Chung et al. [13] presented a Convolutional Neural Network (CNN) by combining auto-context and self-supervised sparse contour attention mechanisms. Ahmad et al. [14] employed a deep belief network for initial liver delineation and Chan-Vese active contour method for final refinement. Senthilvelan et al. [15] developed a cascaded CNN model consisting of V-Net for initial liver segmentation and H-DenseUNet for final refinement. Araújo et al. [16] cascaded multiple UNets for segmenting simple and complex cases. However, the computational cost was high. Fan et al. [17] presented a variant of UNet, in which the skip connections were modified to extract better features. They also introduced special modules to fuse high- and low-level features and to capture multiscale details. Xie et al. [18] combined dynamic adaptive pooling, residual modules and UNet to segment liver from CT data. Ahmad et al. [19] developed an efficient CNN initialized randomly with Gaussian weights for liver segmentation. Wei et al. [20] integrated generative adversarial network into mask region-based CNN to enhance liver segmentation results. Wang et al. [21] combined EfficientNetB4, attention gate and residual learning for liver delineation. Wu et al. [22] presented a DL model based on UNet by including pyramidal convolution and attention mechanisms. These works were automatic but only focused on extracting liver from the Portal Venous (PV) phase.

In clinical practice, generally, plain and contrast-enhanced CT images consisting of arterial, PV and delayed phase images are analyzed for tumor identification. Radiologists observe the enhancement patterns (generated by the contrast agent) in and around the tumor to diagnose them. Majority of the research on liver segmentation is centered on segmenting the liver solely from the PV phase. Very few authors have worked on multiple CT phases. For instance, Xu et al. [23] employed a network derived from UNet to segment liver from the triphasic CT data. The approach used by Rusko et al. [24] was based on region growing algorithm. They incorporated various pre and post processing operations using anatomical and multiphase information, to reduce over and under segmentation of liver. These studies required image registration, which is very time consuming.

A liver segmentation method feasible for a CAD<sub>x</sub> system must be automatic, accurate, robust and computationally

efficient; being effective for multiple CT phases would further add value to the method. This paper aims to accomplish such a method. We have developed a DL model from SegNet using two key components: Atrous Spatial Pyramid Pooling (ASPP) module and leaky Rectified Linear Unit (ReLU) layers. The ASPP module captured multiscale features without reducing the feature map resolution and the leaky ReLU layers improved the model's generalizability. Performance evaluation on a public dataset with challenging cases and our institutional dataset (consisting of multiphase CT volumes) yielded satisfactory results. Ablation studies justified the significance of the model components. Comparison with the state-of-the-art techniques indicated that our model was comparatively superior.

The rest of the paper is structured into the following sections: Section "[Materials and methods](#)": elaborates on the datasets employed and the proposed method. Section "[Experimental results](#)": presents the quantitative and qualitative results, and Section "[Discussion](#)": discusses the results obtained. Finally, Section "[Conclusion](#)", outlines the contributions and presents the future work.

## Materials and Methods

### Dataset Details

The training\validation and test datasets were prepared primarily from different databases. Hence, we have presented the details of the datasets in two separate subsections.

*Training and Validation* We collected 4994 diverse CT images from three databases (two public and one internal), namely, 3D-IRCADb [25], LiTS [26] and Kasturba Medical College (KMC), Manipal. They comprised liver of different shapes and intensity distributions, with/without tumor and abdominal images of lung, heart, intestine etc. that did not have liver. This data was then split into training and validation sets, such that over 78% of the total CT images was used for training and the remaining for validation. Thus, the training and validation datasets comprised 3930 and 1064 CT images, respectively. The data splitting was done manually so that the training and validation datasets contained difficult cases equally. We ensured that both the datasets have similar diversity and that there is no overflow of only certain type of images in either of the two datasets. This was important because only the training images are used by the model for learning features. If the training set would not have images of all the types mentioned above, it would be unable to learn relevant features required for identification of liver from the challenging images. The validation accuracies for two split ratios 78:22 and 90:10 was compared, and it was found that the former ratio gave slightly better results, hence it was considered in the work.

The number of cases and CT images considered from each of the databases, along with other relevant attributes, are detailed in Table 1. The images in the public datasets had a fixed resolution of  $512 \times 512$ , whereas the images in the KMC, Manipal dataset had differing resolutions. The 3D-IRCADb and KMC, Manipal datasets are in Digital Imaging and Communications in Medicine (DICOM) format, whereas the LiTS dataset is in Neuroimaging Informatics Technology Initiative (NIFTI) format. Only PV images were considered from the former two databases, whereas multiphase images were taken from the latter database. Since the training images were selected from different databases, reasonable diversity was introduced in the training data.

*Test set* Two datasets were used to test the proposed model: an internal institutional dataset (KMC, Manipal) and CHAOS, a public dataset [27]. The KMC, Manipal dataset consists of ten CT volumes, out of which six CT volumes have all four phases (plain, arterial, PV and delayed) and the remaining have three CT phases (plain, arterial, PV). This can be seen in Table 4. The dataset has cases with different

abnormalities, viz. metastasis, cysts and hepatocellular carcinoma.

The CHAOS dataset comprised twenty CT training datasets (labeled 1, 2, 5, 6, 8, 10, 14, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 in the database) acquired at PV phase. We chose this dataset for evaluation, firstly because it had CT images in DICOM format which is the standard format for medical images. Secondly, the images were acquired using three different scanners namely, Philips SecuraCT with 16 detectors, Philips Mx8000 CT with 64 detectors and Toshiba AquilionOne with 320 detectors. Hence, the robustness of the model could be assessed. Thirdly, because the dataset had challenging cases, as will be discussed in the next sections. It is to be noted that none of the test set images were part of the training set. An overview of the test datasets is provided in Table 2.

The ground truths for the training, validation and test data required for the KMC dataset were generated using the ITK-SNAP tool [28] under the guidance of an experienced radiologist (co-author) with over twenty years of expertise in medical imaging. The ground truths for the three external

**Table 1** Details of training and validation sets

Dataset specifications	Training set			Validation set		
	3D-IRCADb-01	LiTS	KMC, Manipal	3D-IRCADb-02	LiTS	KMC, Manipal
No. of 3D CT volumes (cases)	20	45	39	2	30	11
No. of images	2242	1502	186	195	651	218
Image resolution	$512 \times 512$	$512 \times 512$	From $464 \times 358$ to $955 \times 873$	$512 \times 512$	$512 \times 512$	From $412 \times 412$ to $664 \times 510$
x–y spacing (in mm) [minimum–maximum]	[0.56–0.87]	[0.60–1]	[0.48–1.13]	0.961	[0.56–1]	[0.68–0.92]
Slice thickness (in mm) [minimum–maximum]	[1.00–4.00]	[0.7–5]	5	[1.8–2.4]	[0.7–5]	[5–5.02]
Image file format	DICOM	NIFTI	DICOM	DICOM	NIFTI	DICOM
Imaging systems	—	—	Philips	—	—	Philips

**Table 2** Details of test datasets

Dataset specifications	CHAOS	KMC, Manipal			
		Plain	Arterial	Portal venous	Delayed
No. of 3D CT volumes	20	10			
Image resolution	$512 \times 512$	$367 \times 369$ to $666 \times 512$	$340 \times 370$ to $666 \times 512$	$348 \times 348$ to $666 \times 512$	$332 \times 332$ to $664 \times 510$
Slices in the volumes [minimum–maximum]	[77–105]	[78–106]	[83–101]	[85–105]	[41–51]
x/y voxel spacing (in mm) [minimum–maximum]	[0.7–0.8]	[0.69–1.18]	[0.68–1.15]	[0.68–1.15]	[0.69–0.98]
Slice thickness (in mm) [minimum–maximum]	[3–3.2]	[5–5.03]	[5–5.02]	[5–5.02]	5
Imaging systems	Philips and Toshiba	Philips			

datasets are available publicly in the database. The KMC dataset will be hereafter referred to as institutional dataset.

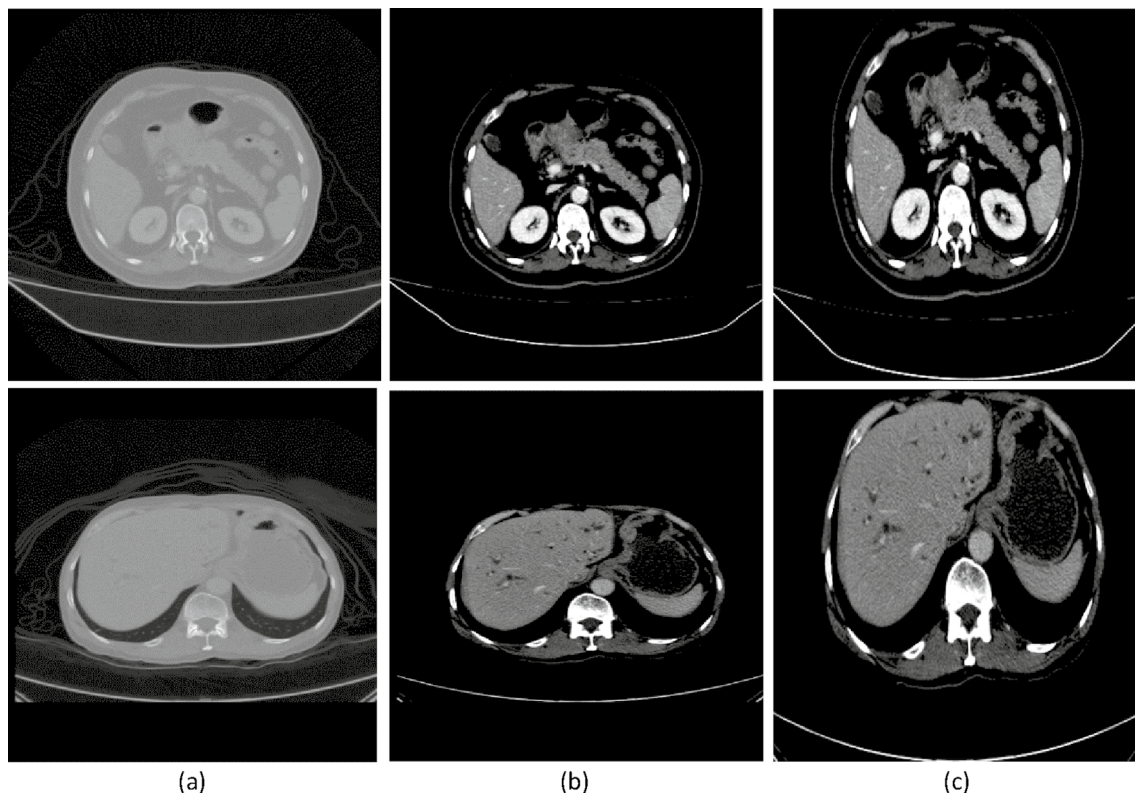
## Preprocessing

The following operations were performed on the training/validation datasets before training: (i) the pixel intensities are first converted to Hounsfield units using a linear transformation described in the DICOM documentation, (ii) the images are converted to unsigned 8-bit integer format (Fig. 1b), (iii) the background pixels in the upper region of the image are removed through cropping to reduce the unwanted areas and magnify the abdominal region, (iv) the images are resampled to the dimension  $384 \times 384 \times 3$  pixels to satisfy the RGB input requirement of the DL model (Fig. 1c). The image dimension was chosen as a tradeoff between image quality and training time. Training images of higher dimensions gave better results at the cost of higher training time; and lower dimension images reduced the training time, however, produced inferior results. The preprocessed training and validation images were saved in .mat format. For the test sets all the preprocessing operations except step (iii) were performed. It is to be noted that all the above mentioned preprocessing steps were automated.

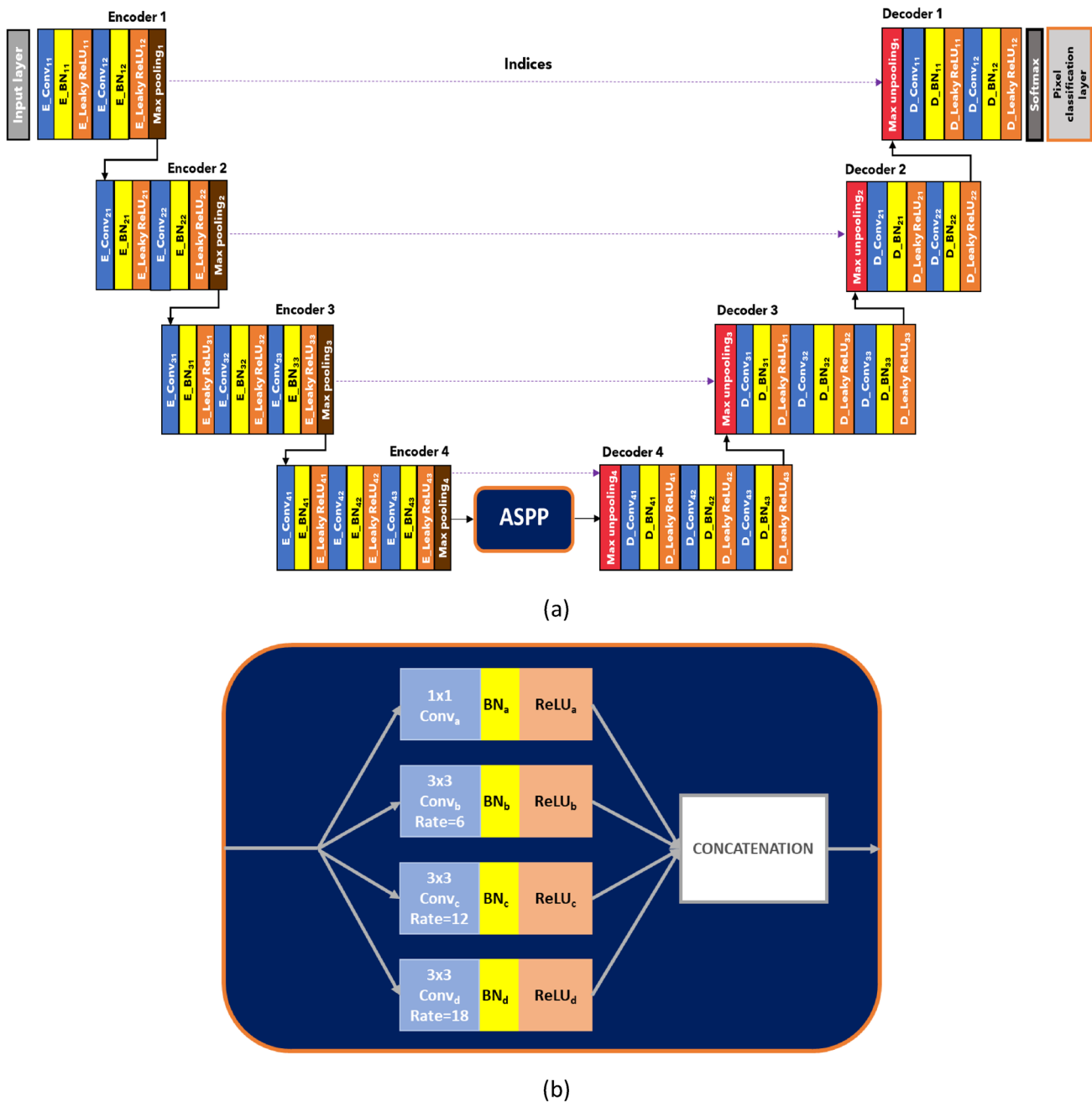
During training, augmentation techniques like scaling and translation were applied on the training and validation datasets on-the-fly. Hence only the diversity in the data was increased and the dataset size remained the same. The scaling factor was randomly selected from the range 60–100% (horizontally) and 40–100% (vertically). For translation, the value was randomly chosen between  $[-25, 25]$  pixels (horizontally) and  $[-5, 5]$  pixels (vertically). The ranges chosen ensured that the abdominal region remained sufficiently intact in the image, post augmentation.

## Proposed Framework

The proposed SegNet based framework for liver segmentation consists of three parts: an encoder, an ASPP module and a decoder (Fig. 2). The model has four encoder-decoder pairs, consisting of convolution (conv.), Batch Normalization (BN) and leaky ReLU layers. The max pooling layer at the decoder performs non-linear upsampling of the input feature maps using pooling indices derived at the corresponding encoder's max-pooling layer. The encoder layers are initialized with weights from the VGG-16 network [29] trained on the ImageNet database. The kernel size for the convolutional and max pooling layers are  $3 \times 3$  and  $2 \times 2$ , respectively; the stride considered are 1 and 2, respectively, for the two layers.



**Fig. 1** Overview of the preprocessing steps. **a** Raw CT axial slice **b** CT slice in unsigned 8-bit integer format after rescaling **c** CT slice after cropping and resizing



**Fig. 2** Proposed framework. **a** Liver segmentation module consisting of four encoder-decoder blocks and ASPP module. **b** Detailed structure of the ASPP module

Cross entropy was the loss function employed during training. Table 3 gives the details of the model architecture.

To overcome the dying ReLU problem sometimes encountered in the ReLU layers, we employed the leaky ReLU layers [30] in the encoder and decoder blocks. A leaky ReLU is an activation function that multiplies any input less than zero by a fixed scalar. It is mathematically defined as follows,

$$f(x) = \begin{cases} scale * x, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (1)$$

where  $x$  is the input value and  $scale$  is the scalar, chosen as 0.01 in this work.

The ASPP module performs four convolutions in parallel, one with  $1 \times 1$  filter and the remaining are atrous convolutions (with dilation rates of 6, 12, 18) with  $3 \times 3$  filters (Fig. 2b). Each of these layers is followed by BN and ReLU

**Table 3** Detailed architecture of the proposed model

Components	Layers	Feature map size
Encoder 1	Input	$384 \times 384 \times 3$
	E_Conv <sub>11</sub> + E_BN <sub>11</sub> + E_Leaky ReLU <sub>11</sub>	$384 \times 384 \times 64$
	E_Conv <sub>12</sub> + E_BN <sub>12</sub> + E_Leaky ReLU <sub>12</sub>	$384 \times 384 \times 64$
Encoder 2	Max Pooling <sub>1</sub>	$192 \times 192 \times 64$
	E_Conv <sub>21</sub> + E_BN <sub>21</sub> + E_Leaky ReLU <sub>21</sub>	$192 \times 192 \times 128$
Encoder 3	E_Conv <sub>22</sub> + E_BN <sub>22</sub> + E_Leaky ReLU <sub>12</sub>	$192 \times 192 \times 128$
	Max Pooling <sub>2</sub>	$96 \times 96 \times 128$
	E_Conv <sub>31</sub> + E_BN <sub>31</sub> + E_Leaky ReLU <sub>31</sub>	$96 \times 96 \times 256$
	E_Conv <sub>32</sub> + E_BN <sub>32</sub> + E_Leaky ReLU <sub>32</sub>	$96 \times 96 \times 256$
	E_Conv <sub>33</sub> + E_BN <sub>33</sub> + E_Leaky ReLU <sub>33</sub>	$96 \times 96 \times 256$
Encoder 4	Max Pooling <sub>3</sub>	$48 \times 48 \times 256$
	E_Conv <sub>41</sub> + E_BN <sub>41</sub> + E_Leaky ReLU <sub>41</sub>	$48 \times 48 \times 512$
	E_Conv <sub>42</sub> + E_BN <sub>42</sub> + E_Leaky ReLU <sub>42</sub>	$48 \times 48 \times 512$
	E_Conv <sub>43</sub> + E_BN <sub>43</sub> + E_Leaky ReLU <sub>43</sub>	$48 \times 48 \times 512$
ASPP block	Max Pooling <sub>4</sub>	$24 \times 24 \times 512$
	Conv <sub>a</sub> Conv <sub>b</sub> Conv <sub>c</sub> Conv <sub>d</sub> + + + + BN <sub>a</sub> BN <sub>b</sub> BN <sub>c</sub> BN <sub>d</sub> + + + + ReLU <sub>a</sub> ReLU <sub>b</sub> ReLU <sub>c</sub> ReLU <sub>d</sub>	$24 \times 24 \times 128$
	Concatenation	$24 \times 24 \times 512$
	Max Unpooling <sub>4</sub>	$48 \times 48 \times 512$
Decoder 4	D_Conv <sub>43</sub> + D_BN <sub>43</sub> + D_Leaky ReLU <sub>43</sub>	$48 \times 48 \times 512$
	D_Conv <sub>42</sub> + D_BN <sub>42</sub> + D_Leaky ReLU <sub>42</sub>	$48 \times 48 \times 512$
	D_Conv <sub>41</sub> + D_BN <sub>41</sub> + D_Leaky ReLU <sub>41</sub>	$48 \times 48 \times 256$
	Max Unpooling <sub>3</sub>	$96 \times 96 \times 256$
Decoder 3	D_Conv <sub>33</sub> + D_BN <sub>33</sub> + D_Leaky ReLU <sub>33</sub>	$96 \times 96 \times 256$
	D_Conv <sub>32</sub> + D_BN <sub>32</sub> + D_Leaky ReLU <sub>32</sub>	$96 \times 96 \times 256$
	D_Conv <sub>31</sub> + D_BN <sub>31</sub> + D_Leaky ReLU <sub>31</sub>	$96 \times 96 \times 128$
Decoder 2	Max Unpooling <sub>2</sub>	$192 \times 192 \times 128$
	D_Conv <sub>22</sub> + D_BN <sub>22</sub> + D_Leaky ReLU <sub>22</sub>	$192 \times 192 \times 128$
	D_Conv <sub>21</sub> + D_BN <sub>21</sub> + D_Leaky ReLU <sub>21</sub>	$192 \times 192 \times 64$

**Table 3** (continued)

Components	Layers	Feature map size
Decoder 1	Max Unpooling <sub>1</sub>	$384 \times 384 \times 64$
	D_Conv <sub>12</sub> + D_BN <sub>12</sub> + D_Leaky ReLU <sub>12</sub>	$384 \times 384 \times 64$
	D_Conv <sub>11</sub> + D_BN <sub>11</sub> + D_Leaky ReLU <sub>11</sub>	$384 \times 384 \times 2$
	<i>SOFTMAX</i>	$384 \times 384 \times 2$
	<i>Pixel Classification layer</i>	$384 \times 384 \times 2$

layers; finally, a concatenation layer fuses the four outputs. The multiple receptive fields incorporated in the ASPP module aid in visualizing the abdominal CT image from different scales [31]. Unlike SegNet, which further downsamples the feature map (to dimension  $12 \times 12$ ), resulting in information loss, the proposed model retrieves useful multiscale context information from the input feature map (of resolution  $24 \times 24$  pixels) after the fourth encoder (shown in Table 3).

By increasing the receptive field of the convolution filters more context information can be extracted that can aid in identifying the liver region more precisely. Different receptive fields enable the model to learn different features pertaining to the region of interest, like neighboring structures, location and size with respect to other organs and many inherent patterns that are beyond human perception. However, increasing the receptive field makes the model more complex in terms of learnable parameters (weights and biases); thus, increasing the network complexity and training time. These are two very crucial factors for training DL models and hence should be used judiciously. A solution to this issue is to use multiple atrous convolutions instead of standard convolutions. Atrous convolutions with dilation rates of 6, 12 and 18 for a  $3 \times 3$  kernel expands the receptive field to  $13 \times 13$ ,  $25 \times 25$  and  $37 \times 37$ , respectively. Hence, the same pixel can be viewed with respect to 168, 624 and 1368 surrounding pixels, thus extracting more information without increasing the learnable parameters. The ASPP module was incorporated to improve the liver segmentation accuracy without increasing the learnable parameters.

### Evaluation Metrics

Five standard metrics namely, Dice Coefficient (DC), Jaccard Index (JI), Matthews's Correlation Coefficient (MCC), Absolute Volume Difference (AVD) and Average symmetric Surface Distance (ASD) evaluated the segmentation accuracy. DC and JI compute the percentage of overlap between the segmented and ground truth volumes [32]; MCC measures the quality of a binary classification. It is appropriate even when the pixels in the two classes are unbalanced [33]. AVD computes the absolute difference

between the segmented and ground truth volumes and ASD gives the average distance between the surfaces of the two volumes in mm. Higher values indicate better performance for DC, JI and MCC; the opposite is true for AVD and ASD.

### Experimental Results

The programs were implemented in MATLAB R2021a and the DL models were trained on a server with NVIDIA T4 GPU with 16 GB memory. The trained models were evaluated with the test sets on a laptop with Intel Core i7-10750H processor, 16 GB RAM (DDR4) and Windows 10 operating system.

### Parameter Setting

The model was trained with Stochastic Gradient Descent with Momentum (SGDM) optimizer with momentum and minibatch size of 0.7 and 2, respectively. The initial learning rate was 0.1. It was lowered by a factor of 0.1 every 50 epochs. The proposed model was trained for eight different epochs viz. 50, 54, 58, 60, 90, 110, 130 and 150 to find the optimal epoch. The DC for the PV phase of the Institutional dataset was computed (Fig. 3). These results were analyzed and since the 150th epoch gave the highest DC, we finalized our epoch as 150.

### Liver Segmentation Results

Tables 4 and 5 show the liver segmentation results for the two test sets. From Table 4, we can see that the best DC achieved was 97.01% (PV phase) and the poorest DC was 87.3% (plain phase). The average DC values obtained for the four CT phases are 96.12% (PV), 94.61% (arterial), 95.01%

(delayed) and 93.23% (plain). For the CHAOS dataset, DC greater than 97% was achieved for majority of the cases. The average DC accomplished was 96.69%.

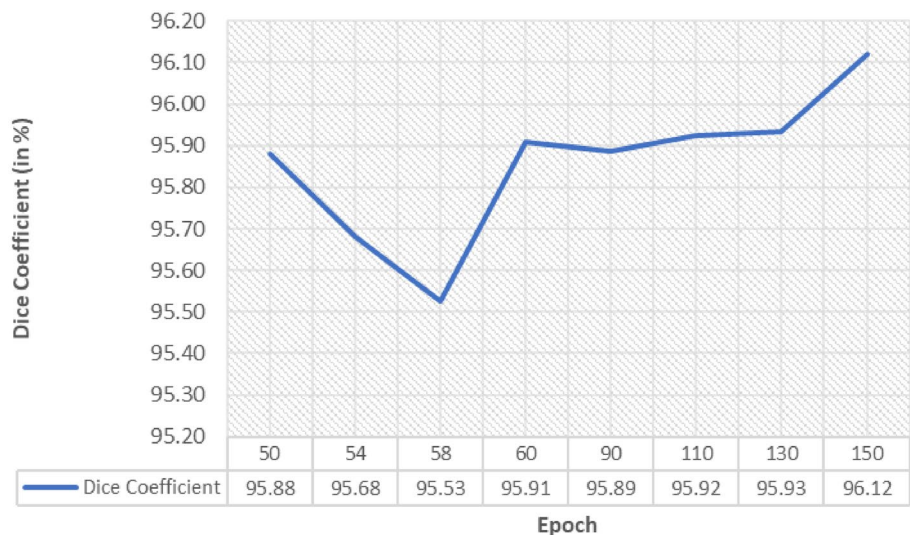
The liver segmentation results for some of the cases are illustrated in Figs. 4 and 5 for the two datasets. The first row of Fig. 4 shows a case that is grainy, the liver has an unusual shape and a large peripheral tumor. Although the contrast between the liver and adjacent structures like rib muscles is lower in the arterial and plain CT phases, the liver is segmented well. The second row shows a case where most part of the liver contains a heterogenous tumor and the contrast between the liver and heart is very low in all the phases except arterial. In both cases, the liver is segmented quite accurately in all phases. Slight discrepancies in the contouring of liver boundary and Inferior Vena Cava (IVC) mainly created differences in the ground truths and predicted masks in Figs. 4 and 5.

The training and validation accuracy and loss curves for the proposed model are shown in Fig. 6. The validation curves exhibited large fluctuations until the fifty-second epoch, after which the variations were reduced.

### Ablation Study

An ablation study was conducted to validate the necessity of the different components of the proposed model. Three DL models were studied: Model 1 (Original SegNet with five encoder-decoder pairs), Model 2 (SegNet with four encoder-decoder pairs), Model 3 (Model 2 with ASPP block inserted between encoder 4 and decoder 4). Tables 6 and 7 summarize the results obtained for the institutional and CHAOS datasets, respectively. Table 8 illustrates the network complexity for the different models.

Fig. 3 Dice coefficient for portal venous phase of institutional dataset



**Table 4** Quantitative results for internal institutional dataset

CT phase	Metrics	1	2	3	4	5	6	7	8	9	10	AVG
Portal venous	DC	95.39	96.76	94.57	97.01	95.51	96.83	95.48	96.28	96.41	96.95	<b>96.12</b>
	JI	91.18	93.73	89.70	94.19	91.41	93.86	91.36	92.82	93.06	94.09	<b>92.54</b>
	MCC	95.32	96.73	94.58	96.96	95.41	96.75	95.40	96.20	96.35	96.90	<b>96.06</b>
	AVD	1.64	2.41	8.89	0.49	0.16	1.23	2.38	1.80	0.18	0.60	<b>1.98</b>
	ASD	1.11	1.86	2.37	0.21	0.32	1.20	2.19	1.21	2.25	0.34	<b>1.31</b>
Arterial	DC	94.58	95.78	92.78	95.19	94.91	96.66	89.75	95.82	94.83	95.78	<b>94.61</b>
	JI	89.72	91.90	86.54	90.82	90.32	93.53	81.41	91.97	90.17	91.90	<b>89.83</b>
	MCC	94.52	95.72	92.86	95.12	94.82	96.58	89.62	95.75	94.76	95.73	<b>94.55</b>
	AVD	2.85	0.93	12.24	1.67	0.66	1.13	4.18	4.52	2.14	2.23	<b>3.25</b>
	ASD	3.38	5.16	2.33	0.44	0.35	0.92	6.88	1.23	1.76	0.33	<b>2.28</b>
Delayed	DC	–	–	93.55	–	95.31	–	94.54	96.23	95.07	95.36	<b>95.01</b>
	JI	–	–	87.88	–	91.04	–	89.64	92.74	90.61	91.14	<b>90.51</b>
	MCC	–	–	93.42	–	95.07	–	94.24	96.07	94.92	95.30	<b>94.84</b>
	AVD	–	–	10.80	–	0.06	–	0.71	2.89	0.69	0.04	<b>2.53</b>
	ASD	–	–	0.78	–	0.54	–	1.35	0.42	0.58	0.6	<b>0.71</b>
Plain	DC	93.49	95.60	87.30	93.08	93.47	94.49	89.58	95.23	94.03	96.01	<b>93.23</b>
	JI	87.78	91.57	77.46	87.06	87.73	89.56	81.13	90.90	88.73	92.34	<b>87.43</b>
	MCC	93.41	95.53	87.73	92.98	93.35	94.35	89.57	95.15	93.94	95.96	<b>93.20</b>
	AVD	3.29	0.84	19.78	2.66	0.16	2.85	10.95	3.24	1.71	2.36	<b>4.78</b>
	ASD	10.07	1.2	7.33	0.74	0.60	5.72	4.70	2.52	3.69	0.42	<b>3.70</b>

**Table 5** Quantitative results for CHAOS dataset

Dataset No	DC	JI	MCC	AVD	ASD
1	97.15	94.45	96.90	2.49	0.89
2	97.22	94.58	97.06	1.82	0.81
5	97	94.17	96.74	3.78	1.06
6	96.02	92.35	95.55	1.91	1.1
8	95.51	91.4	95.24	2.41	1.34
10	97.18	94.52	96.97	2.92	0.86
14	93.4	87.62	92.84	2.87	1.81
16	97.16	94.49	96.94	1.38	0.97
18	96.52	93.28	96.28	6.38	1.1
19	95.41	91.22	95.04	5.88	1.65
21	97.77	95.63	97.66	1.93	0.96
22	97.34	94.81	97.16	0.19	0.76
23	97.74	95.58	97.61	2.56	0.74
24	97.83	95.76	97.70	1.81	0.61
25	95.6	91.58	95.31	0.98	1.83
26	97.51	95.14	97.35	3.24	0.86
27	97.59	95.29	97.44	0.30	0.73
28	96.90	93.99	96.67	0.01	4.43
29	96.29	92.85	96.14	4.41	0.86
30	96.63	93.48	96.48	0.07	0.95
AVG	<b>96.69</b>	<b>93.61</b>	<b>96.45</b>	<b>2.37</b>	<b>1.22</b>

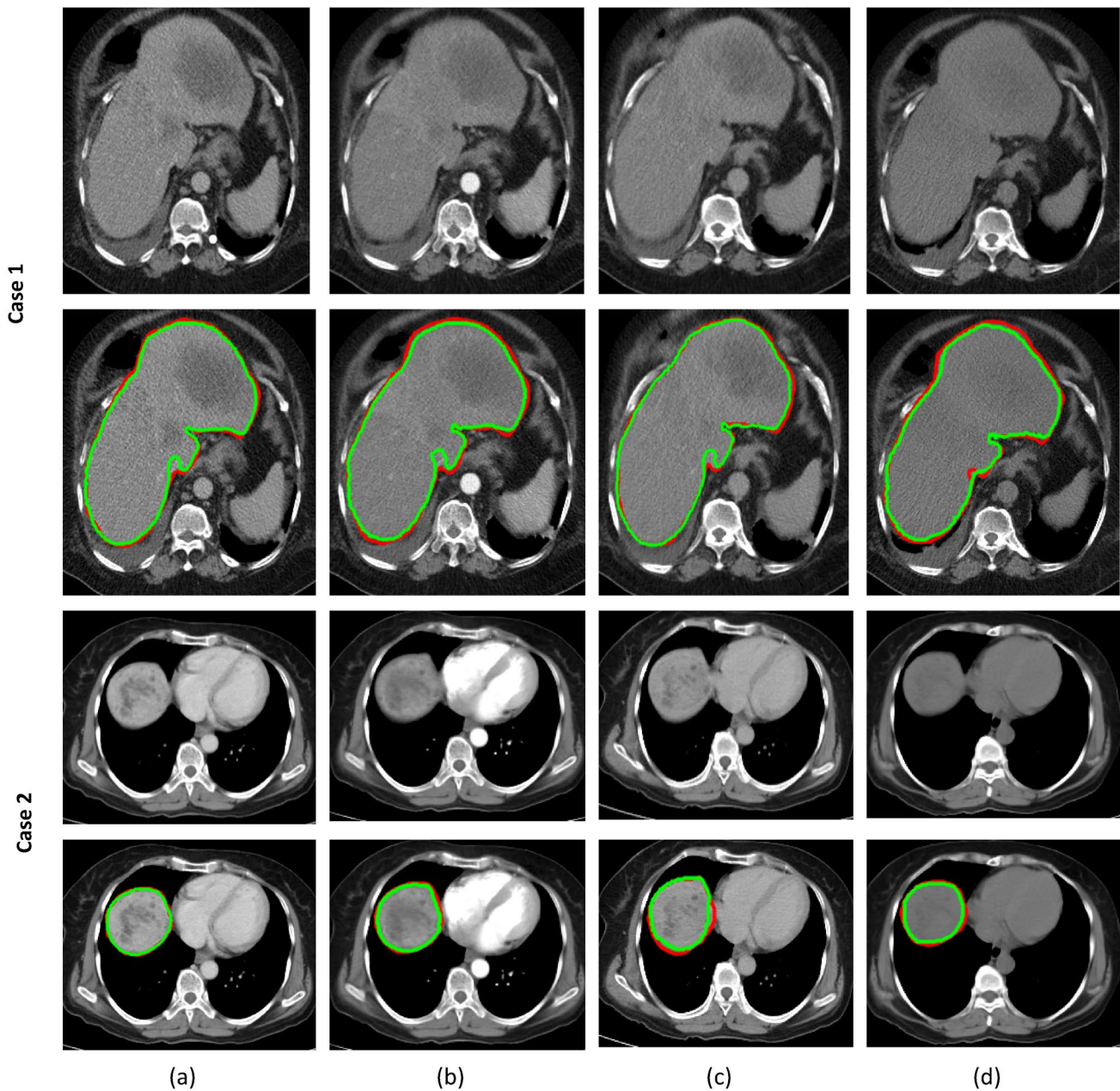
The average values are highlighted in bold

*Comparison with Model 1* It can be observed from Table 6 that the proposed model has outperformed the original SegNet (model 1) for all the CT phases and all the metrics, barring AVD for arterial phase and ASD for plain phase (institutional test set). Table 7 shows that for the CHAOS dataset, the proposed model has performed better for all metrics. Table 8 infers that the proposed model is also superior with respect to number of learnable parameters and training time by approximately 42% and 5 h, respectively. These results show that replacing the fifth encoder-decoder pair in the original SegNet with ASPP block and employing leaky ReLU layers, have improved the liver segmentation accuracy; and also reduced the network parameters and training time. Hence, we can say that the proposed model is superior to SegNet (model 1).

*Comparison with Model 2* In order to examine the usefulness of the fifth encoder-decoder block in model 1, we removed the same and developed model 2. The number of learnable parameters and the training time were reduced considerably (Table 8). However, the segmentation results in the tables clearly show that model 1 is better than model 2 for most of the phases. Thus, the fifth encoder-decoder block is critical for the SegNet.

Compared to the proposed model, model 2 required 1.9 million lesser learnable parameters. However, the proposed model gave better segmentation results. For the PV, arterial, delayed and plain CT phases of the institutional test set the increase in DC was by 1.45%, 3.05%, 0.5% and



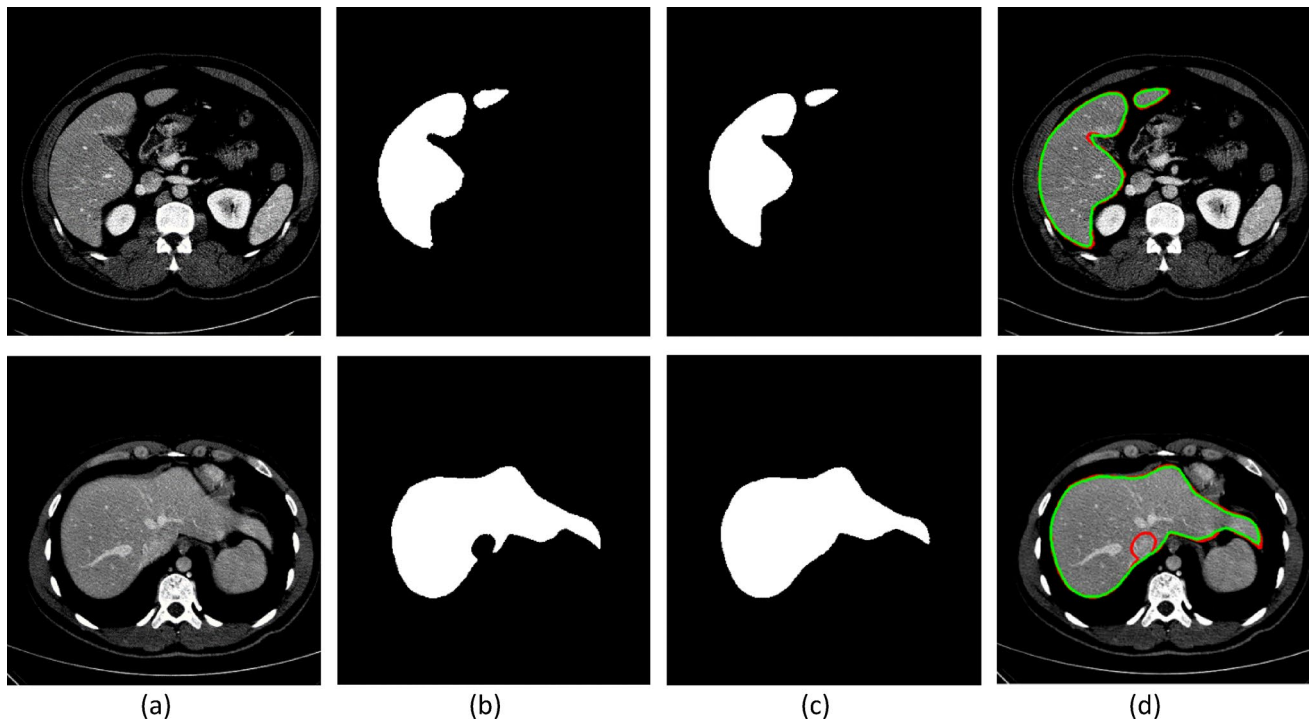


**Fig. 4** Liver segmentation results for multiphase CT images. **a** Portal venous. **b** Arterial. **c** Delayed. **d** Plain (Ground truth: red, predicted: green) (Note: The images have been cropped for better visualization of the liver and the contours)

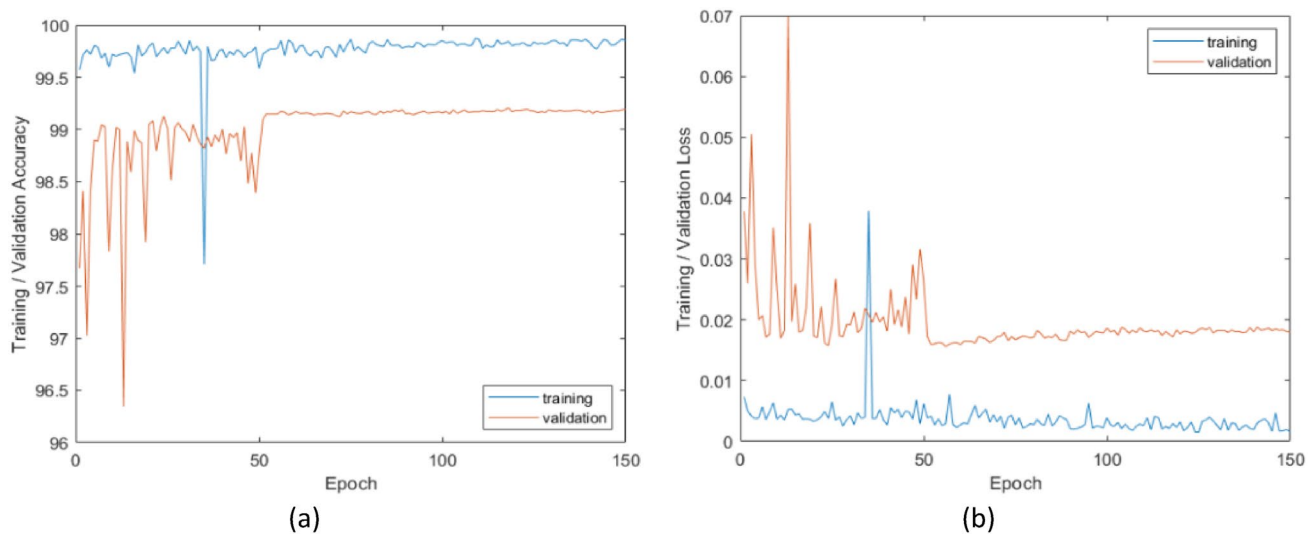
5.02%, respectively; the improvement in JI was by 2.46%, 4.95%, 0.9% and 7.98% for the four phases. The ASD values were better for the proposed model by 7.27 mm, 8.49 mm, 0.41 mm and 4.37 mm for the four phases in the same order. The improvement in MCC was by 1.4%, 2.9%, 0.52% and 4.89%; and AVD was by 2.73%, 6.42%, 0.91% and 6.64%, respectively. For the CHAOS dataset, the improvement was by 0.46%, 0.78%, 0.47%, 0.86% and 1.26 mm for DC, JI, MCC, AVD and ASD, respectively. These results indicate that by removing the fifth encoder-decoder pair, although

there is some reduction in the learnable parameters, there is deterioration in the outcomes of segmentation. Thus, a four encoder-decoder SegNet is not as effective as our proposed model for liver segmentation.

*Comparison with Model 3* In an attempt to get high accuracy along with a reduced number of learnable parameters, we investigated model 3, in which an ASPP module was inserted after the fourth encoder block. Model 3 outperformed model 2 in accuracy; however, it required 3.5 h more for training. Compared to model 1, model 3 gave



**Fig. 5** Liver segmentation results for CHAOS dataset. (a) Axial CT slice (b) Ground truth mask (c) Predicted liver mask (d) Contours marked on CT image (Ground truth: red, predicted: green)



**Fig. 6** Learning curves of the proposed model **a** accuracy **b** loss

better results barring ASD for arterial phase (Table 6). For the CHAOS dataset, model 1 gave better results than model 3 for all metrics except ASD (Table 7). However, the latter model took 1.5 h less than the former for training.

The proposed model has performed better than model 3 for all metrics except AVD for the PV phase. For the arterial phase, ASD was better by 1.15 mm (Table 6). For

the CHAOS dataset, the proposed model has given better results for all metrics. In addition, although both the models had the same number of learnable parameters (17.1 M), the proposed model needed lesser training time. It is noted from Table 6 that model 3 performed slightly better than the proposed model for some of the phases in the institutional dataset. Nevertheless, the better results for the PV phase of

**Table 6** Quantitative results of ablation study on the institutional test dataset

CT phase	Metrics	Model 1	Model 2	Model 3	Proposed
Portal venous	DC	95.39	94.67	95.97	<b>96.12</b>
	JI	91.22	90.08	92.25	<b>92.54</b>
	MCC	95.33	94.66	95.90	<b>96.06</b>
	AVD	2.65	4.71	<b>1.72</b>	1.98
	ASD	2.75	8.58	2.46	<b>1.31</b>
Arterial	DC	93.66	91.56	<b>94.63</b>	94.61
	JI	88.25	84.88	<b>89.84</b>	89.83
	MCC	93.59	91.65	<b>94.56</b>	94.55
	AVD	3.10	9.67	<b>3.05</b>	3.25
	ASD	2.77	10.77	3.43	<b>2.28</b>
Delayed	DC	93.81	94.51	<b>95.01</b>	<b>95.01</b>
	JI	88.38	89.61	<b>90.51</b>	<b>90.51</b>
	MCC	93.61	94.32	94.83	<b>94.84</b>
	AVD	4.88	3.44	<b>2.50</b>	2.53
	ASD	1.04	1.12	<b>0.63</b>	0.71
Plain	DC	92.43	88.21	<b>93.30</b>	93.23
	JI	86.10	79.45	<b>87.49</b>	87.43
	MCC	92.38	88.31	<b>93.24</b>	93.20
	AVD	5.56	11.42	4.85	<b>4.78</b>
	ASD	2.86	8.07	<b>2.84</b>	3.70

The best results are highlighted in bold

**Table 7** Quantitative results of ablation study on the CHAOS dataset

Model	DC	JI	MCC	AVD	ASD
Model 1	96.62	93.48	96.38	2.38	1.67
Model 2	96.23	92.83	95.98	3.23	2.48
Model 3	96.33	92.99	96.07	2.81	1.31
Proposed	<b>96.69</b>	<b>93.61</b>	<b>96.45</b>	<b>2.37</b>	<b>1.22</b>

The best results are highlighted in bold

**Table 8** Comparison of network complexity

DL model	# learnable parameters	# layers	Approximate training time (in hours)
Model 1	29.4 M	91	57
Model 2	<b>15.2 M</b>	<b>71</b>	<b>52</b>
Model 3	17.1 M	84	55.5
Proposed	17.1 M	84	<b>52</b>

The best results are highlighted in bold

M million

the institutional dataset and the challenging CHAOS dataset; and the lesser training time required make the proposed model superior to model 3.

**Table 9** Comparison of other DL models on the internal test dataset

CT phase	Metrics	SegNet	UNet	DeepLab v3+	Proposed
Portal venous	DC	95.39	94.52	95.42	<b>96.12</b>
	JI	91.22	89.73	91.26	<b>92.54</b>
	MCC	95.33	94.54	95.36	<b>96.06</b>
	AVD	2.65	4.80	3.14	<b>1.98</b>
	ASD	2.75	6.01	4.29	<b>1.31</b>
Arterial	DC	93.66	92.53	94.05	<b>94.61</b>
	JI	88.25	86.21	88.81	<b>89.83</b>
	MCC	93.59	92.48	93.98	<b>94.55</b>
	AVD	<b>3.10</b>	6.67	3.86	3.25
	ASD	2.77	5.37	3.89	<b>2.28</b>
Delayed	DC	93.81	93.41	94.35	<b>95.01</b>
	JI	88.38	87.71	89.32	<b>90.51</b>
	MCC	93.61	93.20	94.16	<b>94.84</b>
	AVD	4.88	5.72	4.16	<b>2.53</b>
	ASD	1.04	0.94	1.05	<b>0.71</b>
Plain	DC	92.43	89.07	92.45	<b>93.23</b>
	JI	86.10	80.56	86.08	<b>87.43</b>
	MCC	92.38	88.94	92.43	<b>93.20</b>
	AVD	5.56	13.58	6.96	<b>4.78</b>
	ASD	2.86	2.97	<b>2.80</b>	3.70

The best results are highlighted in bold

**Table 10** Comparison of other DL models on the CHAOS dataset

DL Model	DC	JI	MCC	AVD	ASD
SegNet	96.62	93.48	96.38	2.38	1.67
UNet	96.19	92.72	95.90	3.07	1.52
DeepLab v3+	93.22	89.66	93.44	6.23	1.99
Proposed	<b>96.69</b>	<b>93.61</b>	<b>96.45</b>	<b>2.37</b>	<b>1.22</b>

The best results are highlighted in bold

The above results emphasize that replacing the fifth encoder-decoder pair with the ASPP block and incorporating leaky ReLU instead of ReLU layers has enhanced the performance of the original SegNet. The improvement is in terms of accuracy, computational complexity, training time and model generalizability.

### Comparison with Other DL Models

A comparative analysis with other widely used semantic segmentation networks like UNet, DeepLab v3+ and SegNet was performed and the results are reported in Table 9, 10 and 11. The proposed model outperformed all the DL models in all CT phases except for AVD in arterial phase and ASD in plain phase.

UNet gave the poor results for most CT phases (Table 9). The increase in DC for our model compared to UNet was by

**Table 11** Comparison of network complexity

DL model	# parameters	# layers	Training time (in hours)
SegNet	29.4 M	91	57
UNet	124.3 M	<b>70</b>	<b>48.5</b>
DeepLab v3 +	43.9 M	206	65
Proposed	<b>17.1 M</b>	84	52

The best results are highlighted in bold

1.6%, 2.08%, 1.6% and 4.16% for the PV, arterial, delayed and plain phases, respectively. The improvement in JI was by 2.81%, 3.62%, 2.8% and 6.87%; and MCC was better by 1.53%, 2.06%, 1.64% and 4.26% for the four phases. The ASD metrics improved by 4.7 mm, 3.09 mm and 0.23 mm for PV, arterial and delayed phases. For the CHAOS dataset, the proposed model was better by 0.5%, 0.89%, 0.55%, 0.71% and 0.3 mm for DC, JI, MCC, AVD and ASD, respectively (Table 10).

When compared to DeepLab v3 +, the DC of our model was better by 0.7%, 0.56%, 0.66% and 0.78% for PV, arterial, delayed and plain CT phases, respectively. The improvement in JI was by 1.28%, 1.02%, 1.19% and 1.35% for the four phases. The ASD metrics were better by 2.98 mm, 1.61 mm, 0.34 mm for PV, arterial and delayed phases, respectively. For the CHAOS dataset, proposed model gave the best results and DeepLab v3 + gave the poorest results (Table 10). The improvement in DC, JI, MCC, AVD and ASD were by 3.47%, 3.95%, 3.01%, 3.87% and 0.77 mm, respectively for the proposed model. The comparison of the proposed model with SegNet has already been discussed in the previous subsection.

From Table 11 it can be observed that the proposed model required the least number of learnable parameters (17,115,718 ~ 17.1 million). The learnable parameters constitute weights and biases from the convolutional layers; and offset and scale from the batch normalization layer. The other layers do not have any learnable parameters. The details of the parameters in these layers are given in Table 12. Compared to SegNet, UNet and DeepLab v3 + models, our proposed model requires approximately 42%, 86% and 61% lesser learnable parameters. The minimum training time required was 48.5 h for UNet, our model required 3.5 h more; however our segmentation results were better. Compared to the remaining models, our model was superior in terms of learnable parameters and segmentation accuracy. Hence, we conclude that our model is the best considering all the aspects.

It was observed that all the DL models segmented the simple cases equally well. However, in the majority of the challenging cases, the proposed model outperformed the

other models. The liver segmentation results for some of the unusual cases from the CHAOS dataset are presented in Fig. 7. Here, first column, shows the results for case 14, where the liver has varying intensity and nonuniform texture due to contrast injection. SegNet and UNet exhibited over and under segmentation; and DeepLab v3 + gave abysmal results for all the slices in the volume. However, apart from the inclusion of IVC, our model delineated the liver quite precisely. For case 6, the shape of liver is atypical and the boundary with spleen is vague. It is apparent from Fig. 7c (second column) that the proposed model has quite successfully segmented the liver compared to the other models. For cases 23 and 28, the other models incorrectly segmented parts of the spleen as liver, producing False Positives (FP). The proposed model has segmented only the liver regions, including both lobes for case 28. In case 25, an unusual liver shape is delineated best by the proposed model. The above analysis implies that our model is a clear improvement compared to the well-known DL models.

Figure 8 shows the segmentation results obtained for healthy (third and fourth columns) and unhealthy liver (first and second columns). In the first and second columns the tumors are present near the border of the liver. The liver has however been segmented well in both cases. The third column shows liver with two lobes, that has been segmented accurately. The fourth column shows a healthy liver with heart and other structures with similar intensity in the vicinity, also segmented precisely by the proposed model.

## Discussion

The observations made based on our study are outlined in this section. It is seen that most of the literature on semantic segmentation of the liver has focused on UNet. SegNet based architectures are very rarely used. This trend may be because UNet was initially developed for medical image understanding and segmentation, whereas SegNet was primarily used for road scene segmentation. Our findings suggest that SegNet and the proposed SegNet-based model delineate the liver more accurately than UNet for all the CT phases.

The ablation studies have highlighted that integrating the ASPP scheme into the four encoder-decoder SegNet model improves the segmentation accuracy. Although the DeepLab v3 + network also uses the scheme, our investigations reveal that the proposed model is more efficient, effective and robust (Tables 9, 10 and 11). The proposed model requires 61% fewer learnable parameters and comparatively lesser training time.

The ablation studies indicate that leaky ReLU layers in the encoder and decoder sections have made the model more robust. The liver segmentation results for case 14

**Table 12** Layer wise details of the learnable parameters of the proposed model

Block Name	Layer name	Filter	Stride	Learnable parameters				Learnable parameters
				For convolutional layers		For BN layer		
				Weights	Bias	Offset	Scale	
Encoder 1	E_Conv <sub>11</sub>	3 × 3, 64	1	3 × 3 × 3 × 64	1 × 1 × 64	–	–	1792
	E_BN <sub>11</sub>	–	–	–	–	1 × 1 × 64	1 × 1 × 64	128
	E_Conv <sub>12</sub>	3 × 3, 64	1	3 × 3 × 64 × 64	1 × 1 × 64	–	–	36,928
	E_BN <sub>12</sub>	–	–	–	–	1 × 1 × 64	1 × 1 × 64	128
Encoder 2	E_Conv <sub>21</sub>	3 × 3, 128	1	3 × 3 × 64 × 128	1 × 1 × 128	–	–	73,856
	E_BN <sub>21</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	E_Conv <sub>22</sub>	3 × 3, 128	1	3 × 3 × 128 × 128	1 × 1 × 128	–	–	147,584
	E_BN <sub>22</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	E_Conv <sub>31</sub>	3 × 3, 256	1	3 × 3 × 128 × 256	1 × 1 × 256	–	–	295,168
	E_BN <sub>31</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	E_Conv <sub>32</sub>	3 × 3, 256	1	3 × 3 × 256 × 256	1 × 1 × 256	–	–	590,080
	E_BN <sub>32</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	E_Conv <sub>33</sub>	3 × 3, 256	1	3 × 3 × 256 × 256	1 × 1 × 256	–	–	590,080
	E_BN <sub>33</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	E_Conv <sub>41</sub>	3 × 3, 512	1	3 × 3 × 256 × 512	1 × 1 × 512	–	–	1,180,160
	E_BN <sub>41</sub>	–	–	–	–	1 × 1 × 512	1 × 1 × 512	1024
	E_Conv <sub>42</sub>	3 × 3, 512	1	3 × 3 × 512 × 512	1 × 1 × 512	–	–	2,359,808
	E_BN <sub>42</sub>	–	–	–	–	1 × 1 × 512	1 × 1 × 512	1024
	E_Conv <sub>43</sub>	3 × 3, 512	1	3 × 3 × 512 × 512	1 × 1 × 512	–	–	2,359,808
	E_BN <sub>43</sub>	–	–	–	–	1 × 1 × 512	1 × 1 × 512	1024
	Conv <sub>a</sub>	1 × 1, 128	1	1 × 1 × 512 × 128	1 × 1 × 128	–	–	65,664
ASPP	BN <sub>a</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	Conv <sub>b</sub>	3 × 3, 128, dilation rate = 6	1	3 × 3 × 512 × 128	1 × 1 × 128	–	–	589,952
	BN <sub>b</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	Conv <sub>c</sub>	3 × 3, 128, dilation rate = 12	1	3 × 3 × 512 × 128	1 × 1 × 128	–	–	589,952
	BN <sub>c</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	Conv <sub>d</sub>	3 × 3, 128, dilation rate = 18	1	3 × 3 × 512 × 128	1 × 1 × 128	–	–	589,952
	BN <sub>d</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	D_Conv <sub>43</sub>	3 × 3, 512	1	3 × 3 × 512 × 512	1 × 1 × 512	–	–	2,359,808
Decoder 4	D_BN <sub>43</sub>	–	–	–	–	1 × 1 × 512	1 × 1 × 512	1024
	D_Conv <sub>42</sub>	3 × 3, 512	1	3 × 3 × 512 × 512	1 × 1 × 512	–	–	2,359,808
	D_BN <sub>42</sub>	–	–	–	–	1 × 1 × 512	1 × 1 × 512	1024
	D_Conv <sub>41</sub>	3 × 3, 256	1	3 × 3 × 512 × 256	1 × 1 × 256	–	–	1,179,904
	D_BN <sub>41</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	D_Conv <sub>33</sub>	3 × 3, 256	1	3 × 3 × 256 × 256	1 × 1 × 256	–	–	590,080
Decoder 3	D_BN <sub>33</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	D_Conv <sub>32</sub>	3 × 3, 256	1	3 × 3 × 256 × 256	1 × 1 × 256	–	–	590,080
	D_BN <sub>32</sub>	–	–	–	–	1 × 1 × 256	1 × 1 × 256	512
	D_Conv <sub>31</sub>	3 × 3, 128	1	3 × 3 × 256 × 128	1 × 1 × 128	–	–	295,040
	D_BN <sub>31</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	D_Conv <sub>22</sub>	3 × 3, 128	1	3 × 3 × 128 × 128	1 × 1 × 128	–	–	147,584
Decoder 2	D_BN <sub>22</sub>	–	–	–	–	1 × 1 × 128	1 × 1 × 128	256
	D_Conv <sub>21</sub>	3 × 3, 64	1	3 × 3 × 128 × 64	1 × 1 × 64	–	–	73,792
	D_BN <sub>21</sub>	–	–	–	–	1 × 1 × 64	1 × 1 × 64	128

Table 12 (continued)

Block Name	Layer name	Filter	Stride	Learnable parameters				Learnable parameters
				For convolutional layers		For BN layer		
				Weights	Bias	Offset	Scale	
Decoder 1	D_Conv <sub>12</sub>	3×3, 64	1	3×3×64×64	1×1×64	–	–	36,928
	D_BN <sub>12</sub>	–				1×1×64	1×1×64	128
	D_Conv <sub>11</sub>	3×3, 2	1	3×3×64×2	1×1×2	–	–	1154
	D_BN <sub>11</sub>	–				1×1×2	1×1×2	4
Total								<b>17,115,718</b>

The best result is highlighted in bold

(with hyperdense liver) of the CHAOS dataset is depicted in Fig. 9. The results illustrate that model 3 (with ReLU layers) could not segment the liver as effectively as the proposed model (Fig. 8 (c) and (d)). Although both the models gave similar results for most cases, the results obtained for case 14 of the CHAOS dataset illustrate that the proposed network is more robust.

Our model was trained on CT images from three databases namely, 3D-IRCADb, LiTS and our institutional database. We tested our model on two test sets: (a) CHAOS dataset and (b) our institutional dataset and achieved satisfactory results with both. It is to be noted that no images from the CHAOS database were included in the training set and that the test data from the institutional database were separate from the training images from the same database. Since the images in the CHAOS dataset were acquired using different scanners as mentioned earlier, good results for this dataset proves the robustness of our model. Moreover, it was effective in segmenting the liver from multiple CT phases (plain, arterial, PV and delayed) although it was mainly trained on PV images. Hence, our model has the potential to be integrated into a CAD<sub>x</sub> system.

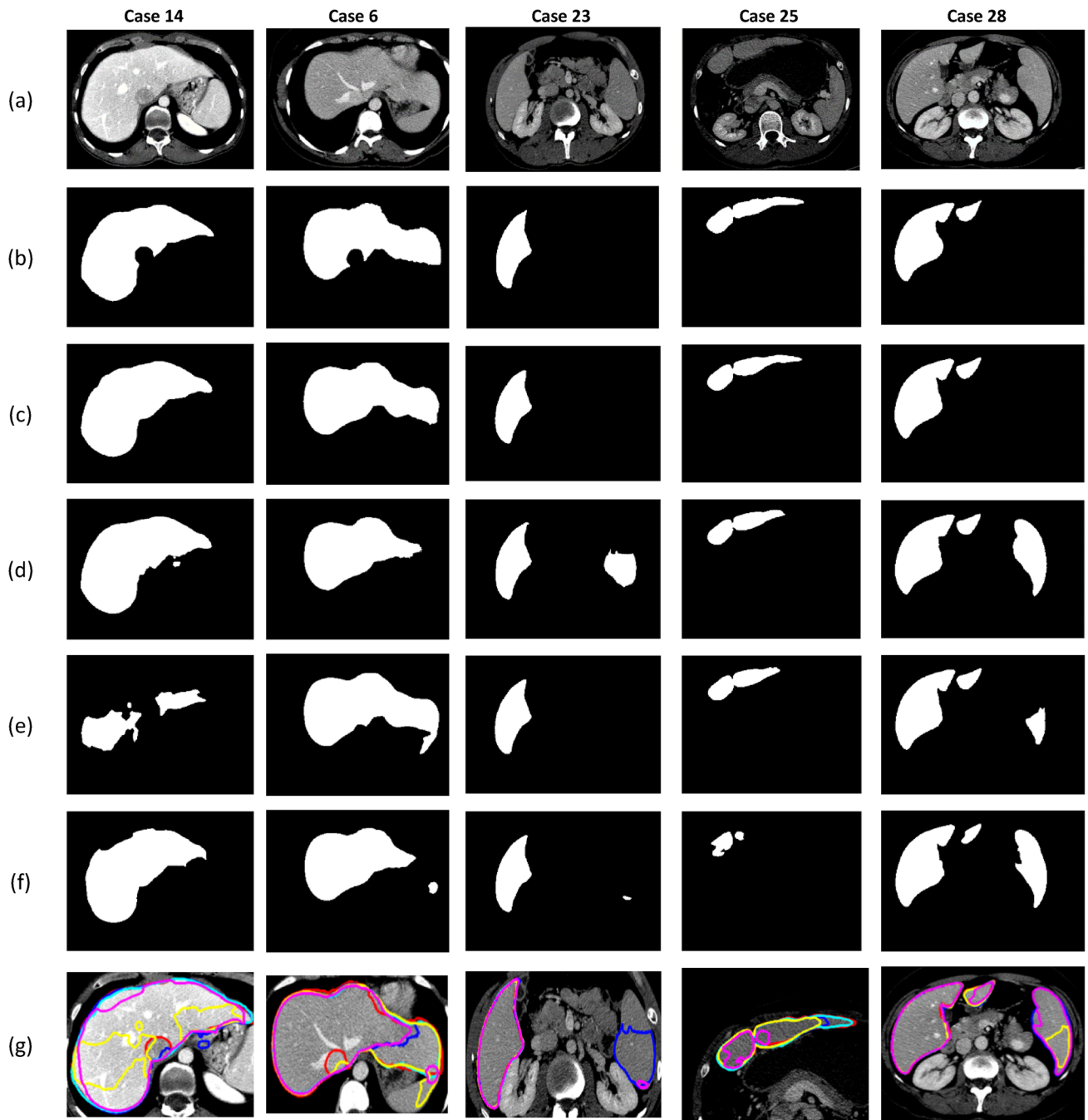
Table 13 shows a comparison of the proposed method with other recent works that employed the CHAOS dataset. The metrics compared were DC and JI, the former metric was specified in all the works whereas the latter was specified only in few works. The proposed has given better results compared to these works. Mulay et al. [34] presented a method based on Holistically-nested Edge Detection and Region-Convolutional Neural Network for liver segmentation. Their approach required the images to be enhanced through adaptive histogram equalization and sigmoid function. The DC value obtained by their method was 94%. Lei et al. [35] proposed a U-shaped network that employed improved pooling operation and skip connections and achieved DC = 95.58%. Khan et al. [36] integrated UNet, residual networks, dilated convolutions and a new loss function to segment liver and reported DC and JI of 95.49% and 89.13%, respectively. Wu et al. [22] developed a CNN based

on UNet, multiscale processing and attention mechanism and obtained a DC of 96.12%. The DC was only slightly less (0.57%) compared to the proposed method (DC = 96.69%), however the JI was lower by 0.93%. Moreover, their work required 24.97 million parameters while the proposed method utilized only around 17 million parameters. Since our model has performed better compared to the other works, in terms of DC, JI and other parameters we can conclude that our model is superior.

To sum up, the advantages of the proposed model compared to the other architectures are that (i) it delineates the liver more precisely from all the CT phases, (ii) it is more robust as complex and uncommon cases, especially in the CHAOS dataset (liver with unconventional shapes, heterogeneous and hyperdense intensity distribution) were comparatively segmented in a better manner. Moreover, the model was trained mainly on the PV phase images and tested on all four CT phases, (iii) it produces fewer FPs that can adversely affect the diagnosis made by the CAD<sub>x</sub> systems, (iv) it is better than other state-of-the-art methods that employed the same dataset (CHAOS) and (v) it is simple to implement as it is built from existing components. The key limitation of the model is that it has not identified the IVC in many cases. Another shortcoming is its sensitivity to the CT image format. Although it was trained using CT images in DICOM and NIFTI formats, the algorithm works better on the former and gives inferior results for CT images in the latter format.

## Conclusion

This study developed a DL model for liver segmentation from multiphase abdominal CT volumes. The network was trained on CT images from different databases and tested on two diverse datasets, firstly an institutional multiphase CT dataset and secondly on a public dataset. The experimental results of a comparative study indicate that the proposed model is superior to some of the

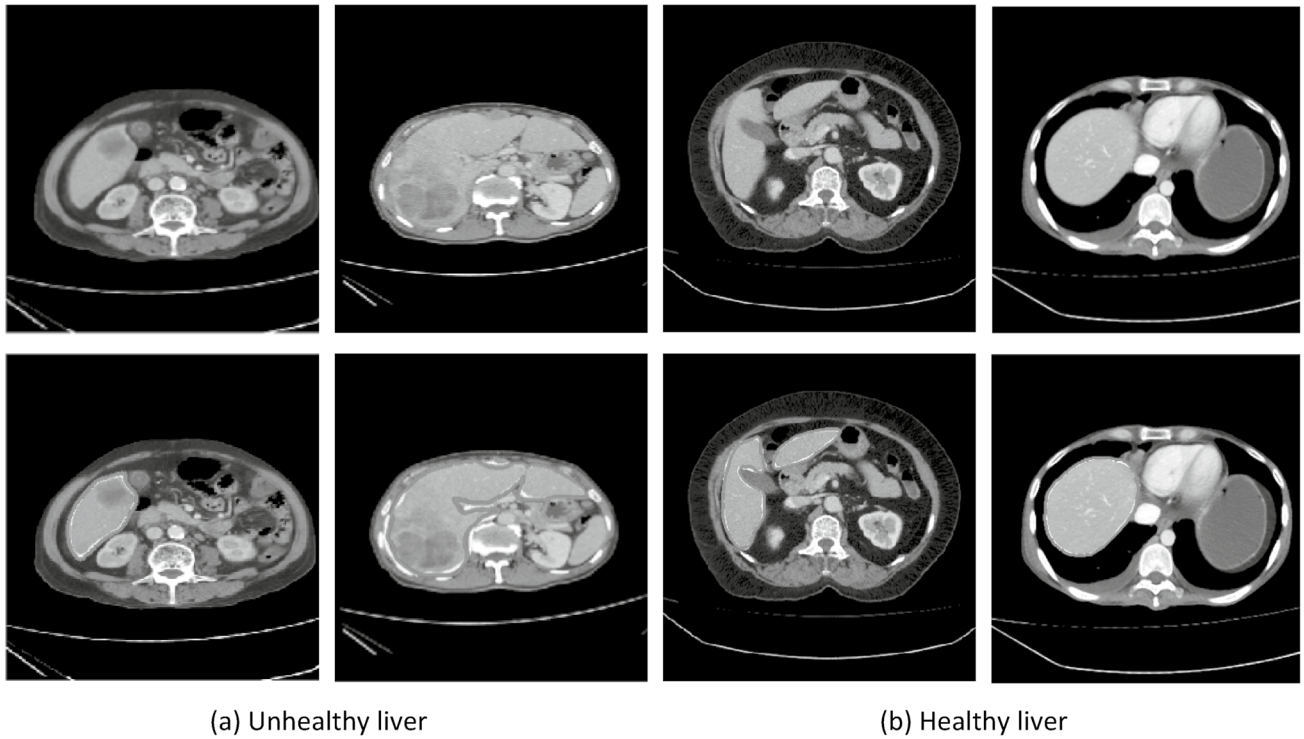


**Fig. 7** Comparison of liver segmentation results of different DL models. **a** Input axial CT image **b** ground truth **c** proposed method **d** SegNet **e** DeepLab v3+ **f** UNet. **g** Contours marked on CT image

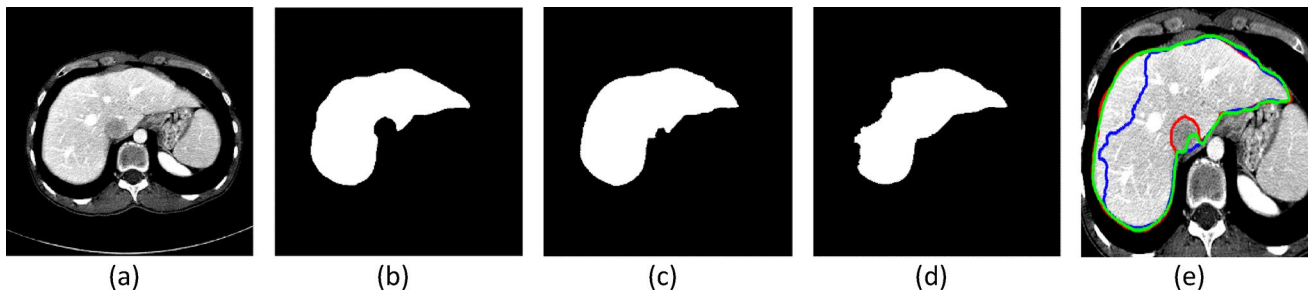
(Ground truth: Red, Proposed: Cyan, SegNet: Blue, DeepLab v3+:Yellow, UNet: Magenta)

commonly employed DL models. It has performed well in terms of accuracy, learnable parameters and training time. Hence, we believe that our liver segmentation algorithm is suitable for incorporation into a CAD<sub>x</sub> system. The future

work includes constructing a hepatic CAD<sub>x</sub> system for differentiating between normal and abnormal liver and diagnosing liver cancer.



**Fig. 8** Segmentation results for unhealthy and healthy liver (Red and yellow contours indicate ground truth and predicted output, respectively)



**Fig. 9** Comparison of liver segmentation results of model 3 and the proposed model for case 14. **a** Input axial slice. **b** Ground truth. **c** Proposed model. **d** Model 3. **e** Liver contours marked on CT image (Ground truth: Red, Proposed: Green, Model 3: Blue)

**Table 13** Comparison of the proposed method with other works that have used CHAOS dataset

Reference	Method	DC	Jl
Mulay et al. [34]	HED-Mask R-CNN	94%	–
Lei et al. [35]	Improved UNet with weighted pooling operation & full-scale skip connection	95.58%	–
Khan et al. [36]	Residual UNet with the dilated convolution	95.49%	89.13%
Wu et al. [22]	Pyramid convolutional attention fusion network	96.12%	92.68%
Proposed	SegNet with ASPP block and leaky ReLU layers	96.69%	93.61%

*HED*Holistically-nested Edge Detection, *R-CNN* Region-Convolutional Neural Network

**Acknowledgements** This work was supported by Karnataka Science and Technology Promotion Society, Department of Science and Technology, Government of Karnataka, India (No. DST/KSTePS/Ph.D.

Fellowship/ENG-02: 2019-20/419/19). We are grateful to Manipal Institute of Technology, Manipal Academy of Higher Education,



Manipal for providing the facilities to carry out the research and Kasturba Medical College, Manipal for providing the images.

**Funding** Open access funding provided by Manipal Academy of Higher Education, Manipal.

**Data Availability** The datasets cannot be made openly available due to institutional protocol.

## Declarations

**Conflict of Interest** The authors have no relevant financial or non-financial interests to disclose.

**Ethics Approval** We have conducted a retrospective study and the anonymized data was collected after obtaining approval from the Institutional Ethics Committee of Kasturba Medical College and Kasturba Hospital (IEC:606/2019).

**Consent to Participate** Written informed consent for participation was not required for this study in accordance with the institutional requirements.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Cancer. 2022. <https://www.who.int/news-room/fact-sheets/detail/cancer>. Accessed 28 Jun 2022.
2. Moghbel M, Mashohor S, Mahmud R, Bin Saripan MI. Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography. *Artif Intell Rev*. 2018;50(4):497–537. <https://doi.org/10.1007/s10462-017-9550-x>.
3. Xu L, Zhu Y, Zhang Y, Yang H. Liver segmentation based on region growing and level set active contour model with new signed pressure force function. *Optik (Stuttg)*. 2020;202(July):2019. <https://doi.org/10.1016/j.ijleo.2019.163705>.
4. Wang J, Cheng Y, Guo C, Wang Y, Tamura S. Shape–intensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal CT images. *Int J Comput Assist Radiol Surg*. 2016;11(5):817–26. <https://doi.org/10.1007/s11548-015-1332-9>.
5. Eapen M, Korah R, Geetha G. Computerized liver segmentation from CT images using probabilistic level set approach. *Arab J Sci Eng*. 2016;41(3):921–34. <https://doi.org/10.1007/s13369-015-1871-y>.
6. Mostafa A, Houssein EH, Houseni M, Hassanien AE, Hefny H. Evaluating swarm optimization algorithms for segmentation of liver images BT. In: Hassanien AE, Oliva DA, editors. *Advances in soft computing and machine learning in image processing*. Cham: Springer International Publishing; 2018. p. 41–62.
7. Mostafa A, Fouad A, Elfattah MA, Hassanien AE, Hefny H. Artificial bee colony based segmentation for CT liver images BT. In: Dey N, Bhateja V, Hassanien AE, editors. *Medical imaging in clinical applications: algorithmic and computer-based approaches*. Cham: Springer International Publishing; 2016. p. 409–30.
8. Sayed GI, Hassanien AE. Abdominal CT Liver Parenchyma Segmentation Based on Particle Swarm Optimization BT. In: *The 1st International Conference on advanced intelligent system and informatics (AISII2015)*, November 28–30, 2015, Beni Suef, Egypt,” 2016; p. 219–228.
9. Liu Z, et al. Automatic liver segmentation from abdominal CT volumes using improved convolution neural networks. *Multimed Syst*. 2021;27(1):111–24.
10. Jeong JG, Choi S, Kim YJ, Lee W-S, Kim KG. Deep 3D attention CLSTM U-Net based automated liver segmentation and volumetry for the liver transplantation in abdominal CT volumes. *Sci Rep*. 2022;12(1):6370. <https://doi.org/10.1038/s41598-022-09978-0>.
11. Sun J, Hui Z, Tang C, Wu X. Liver segmentation based on complementary features U-Net. *Vis Comput*. 2022;39:1–12.
12. Czipczer V, Manno-Kovacs A. Adaptable volumetric liver segmentation model for CT images using region-based features and convolutional neural network. *Neurocomputing*. 2022;505:388–401.
13. Chung M, Lee J, Park S, Lee CE, Lee J, Shin YG. Liver segmentation in abdominal CT images via auto-context neural network and self-supervised contour attention. *Artif Intell Med*. 2021;113(November 2020):102023. <https://doi.org/10.1016/j.artmed.2021.102023>.
14. Ahmad M, et al. Deep belief network modeling for automatic liver segmentation. *IEEE Access*. 2019;7:20585–95. <https://doi.org/10.1109/ACCESS.2019.2896961>.
15. Senthilvelan J, Jamshidi N. A pipeline for automated deep learning liver segmentation (PADLLS) from contrast enhanced CT exams. *Sci Rep*. 2022;12(1):1–11.
16. Araújo JDL, et al. Liver segmentation from computed tomography images using cascade deep learning. *Comput Biol Med*. 2022;140: 105095. <https://doi.org/10.1016/j.compbio.2021.105095>.
17. Fan T, Wang G, Wang X, Li Y, Wang H. MSN-Net: a multi-scale context nested U-Net for liver segmentation. *Signal, Image Video Process*. 2021;15(6):1089–97. <https://doi.org/10.1007/s11760-020-01835-9>.
18. Xie X, et al. Dynamic adaptive residual network for liver CT image segmentation. *Comput Electr Eng*. 2021;91: 107024. <https://doi.org/10.1016/j.compeleceng.2021.107024>.
19. Ahmad M, et al. A lightweight convolutional neural network model for liver segmentation in medical diagnosis. *Comput Intell Neurosci*. 2022;2022:7954333. <https://doi.org/10.1155/2022/7954333>.
20. Wei X, Chen X, Lai C, Zhu Y, Yang H, Du Y. Automatic liver segmentation in CT images with enhanced GAN and mask region-based CNN architectures. *Biomed Res Int*. 2021;2021:9956983. <https://doi.org/10.1155/2021/9956983>.
21. Wang J, Zhang X, Lv P, Wang H, Cheng Y. Automatic liver segmentation using EfficientNet and Attention-based residual U-Net in CT. *J Digit Imaging*. 2022;35:1–15.
22. Wu Y, Wang G, Wang Z, Wang H. PCAF-Net: A liver segmentation network based on deep learning. *IET Image Process*. 2021. <https://doi.org/10.1049/ipr2.12346>.
23. Xu H, Wang B, Xue W, Zhang Y, Zhong C, Chen Y, Leng J. Automatic segmentation of liver CT image based on dense pyramid network. In: *Multiscale multimodal medical imaging: first international workshop, MMMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China: Springer International Publishing*; 2019. p. 10–16.

24. Ruskó L, Bekes G, Fidrich M. Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images. *Med Image Anal.* 2009;13(6):871–82. <https://doi.org/10.1016/j.media.2009.07.009>.
25. Soler L et al. 3D image reconstruction for comparison of algorithm database: a patient-specific anatomical and medical image database. IRCAD, Strasbourg, France, Tech. Rep, 2010.
26. Bilic P et al. The liver tumor segmentation benchmark (lits). *arXiv Prepr. arXiv1901.04056*, 2019.
27. Kavur AE, et al. CHAOS Challenge-combined (CT-MR) healthy abdominal organ segmentation. *Med Image Anal.* 2021;69: 101950. <https://doi.org/10.1016/j.media.2020.101950>.
28. Yushkevich PA, Gao Y, Gerig G. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016; p. 3342–3345.
29. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd Int. Conf. Learn. Represent. ICLR 2015-Conf. Track Proc., 2015; p. 1–14.
30. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. *Proc icml.* 2013;30(1):3.
31. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell.* 2018;40(4):834–48. <https://doi.org/10.1109/TPAMI.2017.2699184>.
32. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging.* 2015;15:29. <https://doi.org/10.1186/s12880-015-0068-x>.
33. Oliveira WS, Teixeira JV, Ren TI, Cavalcanti GDC, Sijbers J. Unsupervised retinal vessel segmentation using combined filters. *PLoS ONE.* 2016;11(2): e0149943.
34. Mulay S, Deepika G, Jeevakala S, Ram K, Sivaprakasam M. Liver segmentation from multimodal images using HED-Mask R-CNN. In: Multiscale multimodal medical imaging. First International Workshop, MMMI 2019, Held in Conjunction with MICCAI 2019. Shenzhen, China: Springer International Publishing; 2020. p. 68–75.
35. Lei J, Lei T, Zhao W, Xue M, Du X, Nandi AK. Rethinking Pooling Operation for Liver and Liver-Tumor Segmentations. *Front Sig Proc.* 2022;1:808050
36. Khan RA, Luo Y, Wu F-X. RMS-UNet: residual multi-scale UNet for liver and lesion segmentation. *Artif Intell Med.* 2022;124: 102231. <https://doi.org/10.1016/j.artmed.2021.102231>.

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