








Article

Deep Learning Framework for Liver Segmentation from T_1 -Weighted MRI Images

Md. Sakib Abrar Hossain ^{1,2} , Sidra Gul ^{3,4}, Muhammad E. H. Chowdhury ^{2,*} , Muhammad Salman Khan ², Md. Shaheenur Islam Sumon ² , Enamul Haque Bhuiyan ⁵, Amith Khandakar ² , Maqsd Hossain ¹, Abdus Sadique ¹, Israa Al-Hashimi ⁶, Mohamed Arselene Ayari ⁷ , Sakib Mahmud ²  and Abdulrahman Alqahtani ^{8,9} 

¹ NSU Genome Research Institute (NGRI), North South University, Dhaka 1229, Bangladesh

² Department of Electrical Engineering, Qatar University, Doha 2713, Qatar

³ Department of Computer Systems Engineering, University of Engineering and Technology Peshawar, Peshawar 25000, Pakistan

⁴ Artificial Intelligence in Healthcare, IIPL, National Center of Artificial Intelligence, Peshawar 25000, Pakistan

⁵ Center for Magnetic Resonance Research, University of Illinois Chicago, Chicago, IL 60607, USA

⁶ Hamad Medical Corporation, Doha 3050, Qatar

⁷ Department of Civil Engineering, Qatar University, Doha 2713, Qatar

⁸ Department of Medical Equipment Technology, College of Applied, Medical Science, Majmaah University, Majmaah City 11952, Saudi Arabia

⁹ Department of Biomedical Technology, College of Applied Medical Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

* Correspondence: mchowdhury@qu.edu.qa

Abstract: The human liver exhibits variable characteristics and anatomical information, which is often ambiguous in radiological images. Machine learning can be of great assistance in automatically segmenting the liver in radiological images, which can be further processed for computer-aided diagnosis. Magnetic resonance imaging (MRI) is preferred by clinicians for liver pathology diagnosis over volumetric abdominal computerized tomography (CT) scans, due to their superior representation of soft tissues. The convenience of Hounsfield unit (HoU) based preprocessing in CT scans is not available in MRI, making automatic segmentation challenging for MR images. This study investigates multiple state-of-the-art segmentation networks for liver segmentation from volumetric MRI images. Here, T_1 -weighted (in-phase) scans are investigated using expert-labeled liver masks from a public dataset of 20 patients (647 MR slices) from the Combined Healthy Abdominal Organ Segmentation grant challenge (CHAOS). The reason for using T_1 -weighted images is that it demonstrates brighter fat content, thus providing enhanced images for the segmentation task. Twenty-four different state-of-the-art segmentation networks with varying depths of dense, residual, and inception encoder and decoder backbones were investigated for the task. A novel cascaded network is proposed to segment axial liver slices. The proposed framework outperforms existing approaches reported in the literature for the liver segmentation task (on the same test set) with a dice similarity coefficient (DSC) score and intersect over union (IoU) of 95.15% and 92.10%, respectively.

Keywords: deep learning; automated liver segmentation; MRI; diagnostic radiology; T_1 -weighted contrast



Citation: Hossain, M.S.A.; Gul, S.; Chowdhury, M.E.H.; Khan, M.S.; Sumon, M.S.I.; Bhuiyan, E.H.; Khandakar, A.; Hossain, M.; Sadique, A.; Al-Hashimi, I.; et al. Deep Learning Framework for Liver Segmentation from T_1 -Weighted MRI Images. *Sensors* **2023**, *23*, 8890. <https://doi.org/10.3390/s23218890>

Academic Editors: Dae-Ki Kang and Sukho Lee

Received: 23 June 2023

Revised: 8 August 2023

Accepted: 15 August 2023

Published: 1 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past decade, remarkable advancements in deep learning (DL) algorithms have led to a rapid transformation in the field of radiology. DL-aided diagnostics have achieved exceptional accuracy in detecting abnormalities in various domains such as ophthalmology, respiratory, and breast imaging. In some cases, multimodal DL solutions now exhibit accuracy levels comparable to expert radiologists. The high performance and clinically satisfactory outcomes achieved through computer-aided diagnostic radiology were previously considered inconceivable [1–3].

Semantic segmentation is a prerequisite for any DL-driven diagnostic task, as it allows the model to learn from the region of interest. Formerly, for semantic segmentation tasks in radiology/medical imaging, distinct mathematical models were implemented, but such approaches often lacked a generalized solution. Deep learning-based segmentation tasks outperform conventional mathematical modeling-based approaches. Segmentation has always helped in improving the performance of computer-aided diagnosis [4,5]. Q. Dou et al. present a unique 3D deeply supervised network (3D DSN) explicitly designed for liver segmentation from CT data [6]. The network incorporates deep supervision to enhance optimization and discrimination capabilities during the learning process, resulting in competitive segmentation results compared to state-of-the-art approaches, along with improved processing speeds. In another study, C. Chen et al. propose an innovative method for lung lesion segmentation in CT scans of COVID-19 patients. Their approach involves region-of-interest extraction and employs a 3D network with attention mechanisms to enhance segmentation accuracy [7]. Additionally, C. Chen et al. introduce a rapid and precise lung segmentation technique, utilizing the edge-weighted random walker algorithm with spatial and clustering information to achieve a heightened accuracy and reduced segmentation time [8]. Similarly, P. Hu et al. develop a liver segmentation framework by integrating a 3D convolutional neural network (CNN) with globally optimized surface evolution. Their approach demonstrates effective segmentation outcomes suitable for clinical applications [9]. Together, these contributions significantly enhance the field of automated organ segmentation, offering valuable insights for medical imaging research and clinical implementations.

However, such a segmentation task in an anatomical paradigm, i.e., the identification and delineation of an anatomical area or structure in magnetic resonance imaging (MRI), encounters a colossal amount of complexity. The complexity can be due to topology, spatial distance, location, relative motion, texture, geometrical structure, and other varying anatomical information. As a consequence, anatomical segmentation has always been a demanding task. In particular, compared to other anatomical structures, very few significant works can be found that focus on liver segmentation [10–12].

For any deep learning-based liver disease diagnosis system, precise automated liver segmentation is indispensable. However, similar to any anatomical segmentation task, it is extensively challenging. This is due to the fact that, compared with other abdominal organs, its anatomy can noticeably differ with patients and clinical conditions. Additionally, the liver's proximity to contiguous abdominal organs (the spleen and kidneys) generates substantial ambiguity [13,14].

However, recent research has demonstrated excellent results for deep neural network (DNN)-based liver segmentation tasks from volumetric abdominal computed tomography (CT) images. Tang et al. [15] achieved a dice similarity coefficient (DSC) of 98% in the liver segmentation task from a plain CT scan using a modified multiscaled convolutional neural network (CNN). Hu et al. [9] used a three-dimensional CNN for the same task and achieved a high performance of around a 97.25% dice similarity coefficient. These works utilized Hounsfield unit (HoU) scaling as a hyperparameter for image enhancement in the preprocessing stage [16]. The review by Xiang et al. [17] observed that, in terms of liver segmentation from magnetic resonance imaging (MRI) scans, high performance could not be achieved and also very little significant work exists in this domain. Owing to the absence of such homogeneous HoU-based image enhancement convenience, in terms of automated liver segmentation from volumetric abdominal MR scans, achieving a similarly high performance to CT images is challenging.

Moreover, MRI scans are extensively adopted by clinicians for liver pathology investigation, due to their superior contrast and spatial resolution for soft tissues compared to CT scans [18,19]. CT scans can provide solid anatomical information. On the contrary, MRI demonstrates high signal intensity in comparison to CT scans. As a result, both anatomical and physiological information can be derived from MRI scans. In particular, both CT and MRI can provide accurate anatomical information about the liver lesion or

haemangioma; however, MRI scans can provide a further important basis for screening benign or malignant types [20–22].

The above studies demonstrate the significance of liver segmentation, specifically from volumetric MRI scans, as this modality is favored by clinicians in relation to pathological diagnosis. In this regard, liver segmentation from MRI scans holds significant importance. A. Mostafa et al. investigated a whale optimization algorithm for liver segmentation from MRI scans [23]. A. Hänsch et al. studied multimodal training and three-dimensional CNN for the task [24]. X. Zhong et al. used deep action learning with a 3D UNet [25], and P. Pandey et al. investigated contrastive semisupervised learning for the liver segmentation task [26] on the CHAOS abdominal MRI dataset [27]. D. Mitta et al. implemented a weighted UNet with attention gates for the liver segmentation task [28] on the same dataset, and J. Hong et al. achieved a slightly better performance using a source-free unsupervised UNet [29]. X. Wang et al. investigated a bidirectional search of the neural net for the task [30]. Additionally, S. Mulay et al. used a geometric edge enhancement-based mask R-CNN [31]. The more recent work of L. Zbinden et al. achieved better performance than previous research for liver segmentation on the same testing set by implementing nnUNet on T_1 -weighted MRI slices [32].

In this research, we investigated 24 state-of-the-art segmentation networks for liver segmentation tasks from T_1 -weighted MR scans using a publicly available dataset, which annotated ground truths for the liver segmentation of 20 patients. The prospect of predicting a precise mask from T_1 -weighted MR scans is higher as fat (and protein) contents are brighter and more distinguishable in such a group. The investigation explores state-of-the-art segmentation networks, such as UNet, UNet++, and feature pyramid network (FPN) segmentation networks with varying dense encoder backbones, along with various image enhancement techniques in the preprocessing stage. The proposed cascaded network showed superior performance to many high-performance state-of-the-art approaches on the same test set. Finally, we developed a software prototype by deploying our proposed DL model in a cloud server for public usage. The cross-platform software is open source and can be accessed from <http://130.211.209.103/projects/the-big-mri-project-beta>, accessed on 9 August 2023.

The main contributions of the research are listed below.

- This research extensively investigates state-of-the-art approaches for precise liver segmentation from T_1 -weighted abdominal MR scans to facilitate clinicians with AI-driven assistance for liver pathology diagnosis;
- This research investigates the effects of multiple image enhancement techniques for automated liver segmentation tasks from MR scans;
- This research proposes a novel cascaded network for the liver segmentation task that demonstrated state-of-the-art performance compared to the literature;
- The proposed model was deployed in a cloud server for demonstration purposes so that clinicians can directly benefit from the results of this investigation.

2. Materials and Methods

The brief methodology of the research is explained in Figure 1. The methodology will be discussed in detail in the following section.

2.1. Dataset

The dataset was collected from the Combined Healthy Abdominal Organ Segmentation (CHAOS) grant challenge [27]. The public portion of the CHAOS dataset includes computed tomography (CT) and magnetic resonance imaging (MRI) abdominal scans of 20 patients, in the Digital Imaging and Communications in Medicine (DICOM) format. The ground truths (GT) were provided from the source, which includes masks for the right kidney, left kidney, liver, and spleen. The ground truth masks were annotated by certified radiologists. All scans are of healthy patients. The MRI scans include T_1 -weighted in-phase and out-phase, along with T_2 -weighted scans, which are discussed in the next subsection.

Each T_1 scan includes 26 to 56 slices; for 20 patients the total number of T_1 -weighted slices is 647 [27].

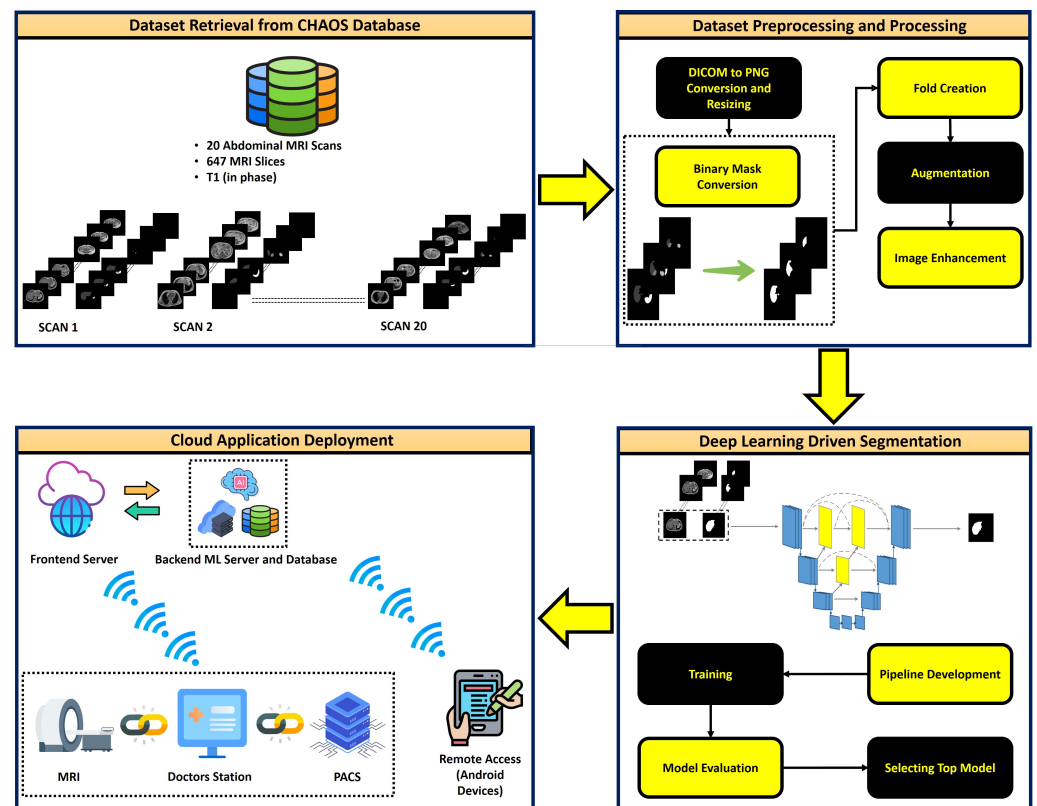


Figure 1. Flow diagram explaining methodology for automated liver segmentation from T_1 -weighted MRI scans.

2.2. Selecting Task-Specific Contrast Group

Among different contrast-enhanced groups (in-phase T_1 -weighted, out-phase T_1 -weighted, and T_2 -weighted), specific groups were chosen by analyzing the relevant abdominal anatomy and attributes of the available contrast-enhanced groups.

2.2.1. Relevant Abdominal Anatomy

The supplied masks (left kidney, right kidney, liver, and spleen) lie in close proximity to each other in the abdominal region, leading to a colossal amount of ambiguity in distinguishing any of the organs. The anatomy of these organs is briefly visualized in Figure 2 [33]. The superior part of the liver (left lobe) lies within the epigastric and left hypochondriac regions. It is in close proximity to the spleen and rests in front of the spleen in terms of the axial plane. The middle part of the liver resides above the umbilical region. The inferior part of the liver is just in front of the upper pole of the right kidney, which occupies the right lumbar region. The left kidney lies in the left lumbar region just below the spleen. Therefore, such close proximity generates an enormous amount of complexity and obscurity in automated abdominal organ segmentation tasks using machine learning.

2.2.2. T_1 - and T_2 -Weighted Images

The T_1 and T_2 parameter represents relaxation time for longitudinal (M_z) and transverse (M_t) magnetization components for each proton. T_1 is noted as the spin–spin relaxation phenomenon, and T_2 is noted as the spin–lattice relaxation phenomenon. When the

macroscopic magnetizing vector for each voxel is M_0 , then the relationship among the magnetization components and T_2 , T_1 is denoted as [34]

$$M_t(t) = M_0 \sin \alpha e^{-\frac{t}{T_2}} \quad (1)$$

$$M_z(t) = M_0 \cos \alpha e^{-\frac{t}{T_1}} + M_0(1 - e^{-\frac{t}{T_1}}) \quad (2)$$

where α denotes the flip angle, which represents a rotation in net magnetization. Characteristically, the T_1 tissue relaxation time is always larger than T_2 . The relaxation times vary broadly with tissue attributes and characteristics. These varying intervals can also be used to distinguish between healthy and abnormal tissues. Table 1 denotes T_1 and T_2 values for relevant abdominal tissues [35].

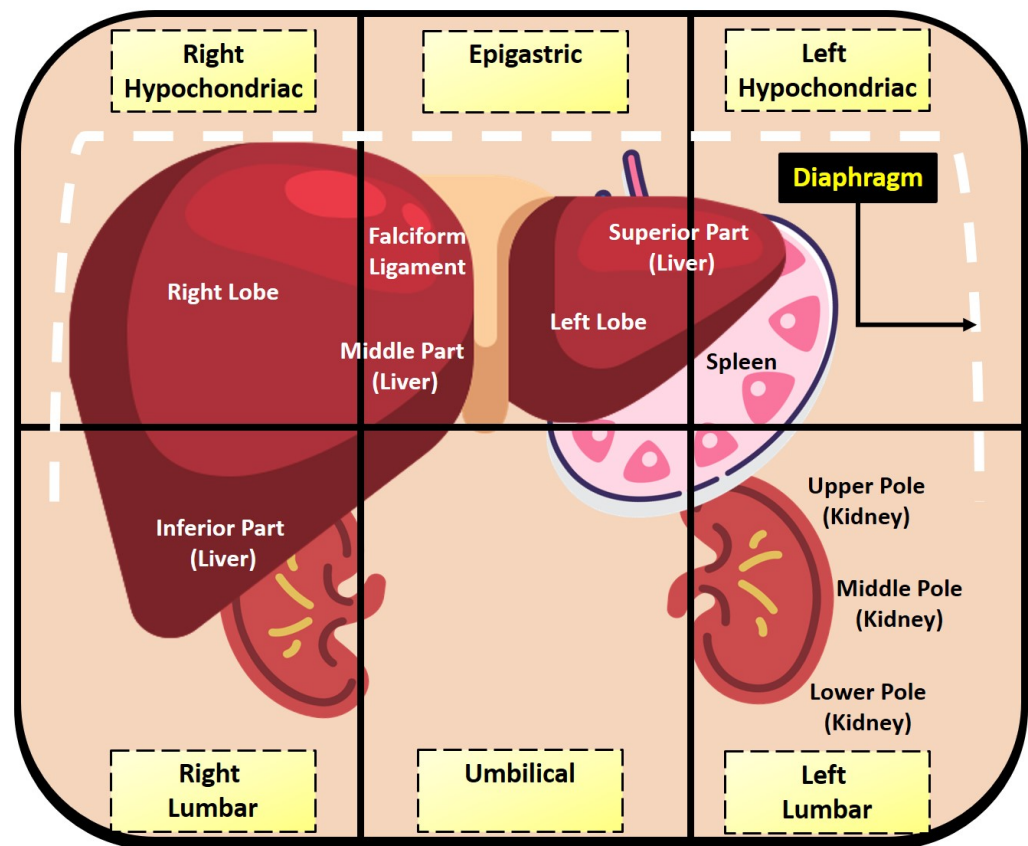


Figure 2. Superficial visualization of relevant abdominal anatomy for describing underlying ambiguity in the segmentation task.

Table 1. Average T_1 and T_2 relaxation time (msec) for 1.5 T and 3.0 T MRI scans.

Tissue	1.5 T		3.0 T	
	T_1 (msec)	T_2 (msec)	T_1 (msec)	T_2 (msec)
Kidney	966–1412	85–87	1142–1545	76–81
Liver	586	46	809	34
Spleen	1057	79	1328	79
Lipid	343	58	382	68

The relation among the image intensity of each voxel $I(x, y)$, the tissue density $\rho(x, y)$, the echo time (TE), and the repetition time (TR) can be denoted as

$$I(x, y) = \rho(x, y) \frac{(1 - e^{-\frac{TR}{T_1}}) \sin \alpha}{1 - e^{-\frac{TR}{T_1}} \cos \alpha} e^{-\frac{TE}{T_2}} \quad (3)$$

In Equation (3), α is optimized by following

$$\alpha_{Ernst} = \cos^{-1} e^{-\frac{TR}{T_1}} \quad (4)$$

when $TE \ll T_2$, and either $\alpha \sim \alpha_{Ernst}$ or $TR \sim T_1$, then the image is defined as T_1 -weighted. Moreover, the image is defined as T_2 -weighted when $TE > T_2$, and either $\alpha \ll \alpha_{Ernst}$ or $TR \gg T_1$ [36].

In accordance with its definitions, fat (and protein) content in T_1 -weighted MRI scans is brighter. Owing to such characteristics, the liver is more distinguishable in T_1 -weighted MRI scans. Figure 3 shows sample T_1 - and T_2 -weighted slices of different axial views. It is clear from the figure that for T_1 -weighted in-phase scans, the liver is far more distinguishable (even in the slices where the liver is small) in the inferior part of the liver, and the superior part of the liver in the axial view. In the MRI slices where the liver is larger (i.e., the middle part of the liver), both in-phase and out-phase T_1 -weighted scans can be used. As T_1 -weighted out-phase scans represent out-of-phase protons, a darker boundary can be noticed around regions of varying intensities. As a result, unwanted artifacts are introduced in these slices.

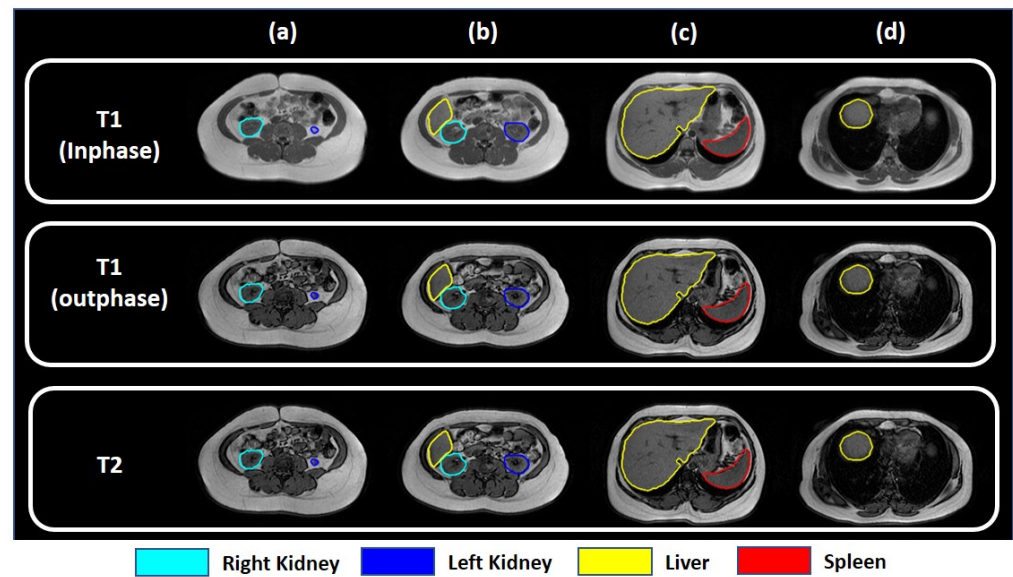


Figure 3. Visualization of MRI slices from (a) upper pole of the kidney, (b) inferior part of the liver, (c) middle part of the liver, and (d) superior part of the liver for different types of data available in the dataset.

Due to such attributes among different contrast-enhanced MRI scans, in-phase T_1 -weighted contrast-enhanced scans were selected for the liver segmentation task. Such groups provide initially enhanced images, which can contribute to boosting the performance of deep neural networks.

2.3. Dataset Preprocessing

Firstly, in-phase T_1 -weighted 647 DICOM slices were converted to PNG format in order to optimize the preprocessing and processing steps. In the ground truth mask, there are multiple organs (right kidney, left kidney, liver, and spleen) present. Binary

masks are generated for the liver alone. Each slice and GT mask pair is then resized to 256×256 dimensions from their original 512×512 dimensions. Reducing the size of the dataset offers notable benefits in terms of enhancing computational efficiency during the training process of segmentation networks.

In order to ensure the data are ready for the machine learning investigation, there are important steps that include fold creation from the preprocessed dataset. Fold creation involves dividing the data into the training set, validation set, and testing set for five folds. In order to avoid biases during training, it is important to make sure that the dataset is balanced; this is achieved by the augmentation of the training set. Finally, the authors investigated different image enhancement techniques for each of the created folds. Figure 4 represents techniques for fold creation and augmentation, which performed following the literature in [37–39]. Image enhancement techniques are demonstrated in Figure 5.

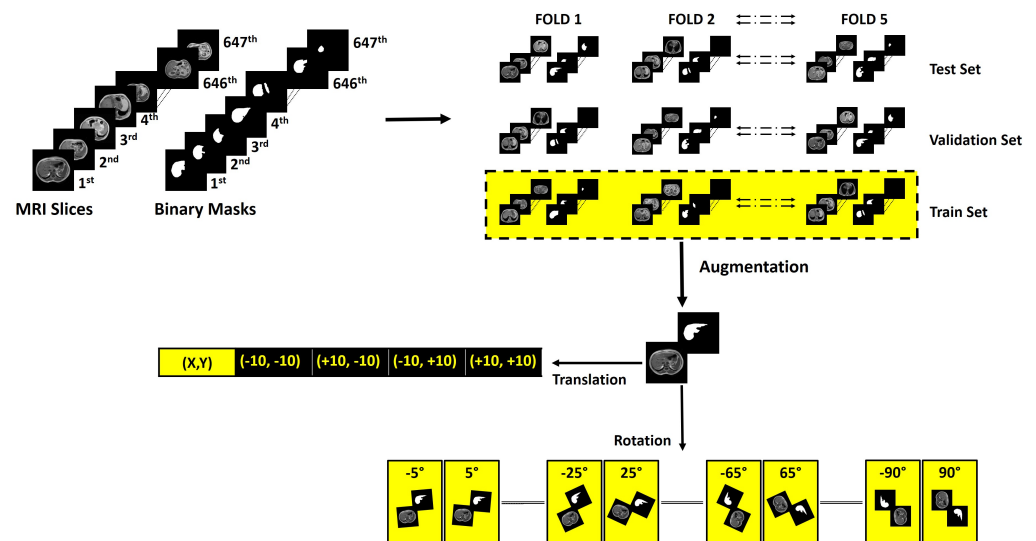


Figure 4. Flow diagram explaining the methodology for fold creation and augmentation in the training set.

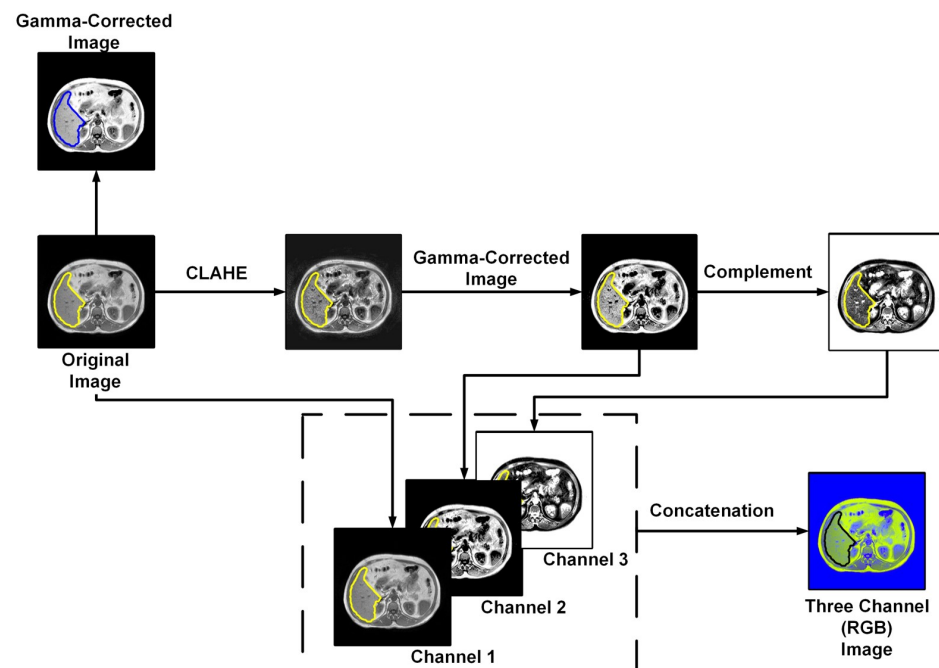


Figure 5. Visualization of image enhancement techniques.

2.3.1. Fold Creation

The methodology follows five-fold cross-validation techniques for validating the network performances. From the preprocessed dataset, five folds were created. In each fold training, validation, and testing set ratios were 70%, 10%, and 20%, which corresponds to 453, 65, and 129 DICOM slices, respectively. This was done to make sure that the performance metric represents the performance of the trained network on the complete dataset.

2.3.2. Augmentation

The training set for each fold was augmented using geometrical spatial transformation of coordinates (rotation and translation). Geometric spatial transformations represent a widely recognized and efficient technique for processing topographic imaging datasets [40,41].

The affine matrix for rotation $I_{rotation}$ and for translation $I_{translation}$ can be denoted as

$$I_{rotation} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$I_{translation} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad (6)$$

where the values of θ are defined by the set:

$$\theta = \{\pm 5^\circ, \pm 10^\circ, \pm 15^\circ, \pm 20^\circ, \dots, \pm 90^\circ\} \quad (7)$$

and the values (t_x, t_y) are defined by the set:

$$(t_x, t_y) = \{(-10, 10), (+10, -10), (-10, +10), (10, 10)\} \quad (8)$$

The validation and testing sets were not augmented. After augmentation, each training fold consisted of around 6700 slices. The validation set was used to avoid overfitting, which is a common problem in machine learning model development [42,43].

2.3.3. Image Enhancement

Image enhancement includes gamma correction for each fold. For each of the pixels $f(x, y)$, the gamma correction can be denoted as [44]

$$g(x, y) = 255 \left(\frac{f(x, y)}{255} \right)^{\frac{1}{\lambda}} \quad (9)$$

where $g(x, y)$ denotes the gamma corrected pixel value, and the value of λ is considered to be 0.5 in this study. And, for all $f(x, y) > 200$, $f(x, y)$ is considered to be 255 to enhance the targeted region. Another image enhancement technique called contrast-limited adaptive histogram equalization (CLAHE) was used in the three-channel (or RGB) image construction technique. If in a histogram k th, the intensity value is r_k , and the number of pixels with the r_k intensity value is n_k , then for an $M \times N$ dimensional image, the equalized histogram can be represented by

$$p(r_k) = \frac{n_k}{M \times N} \quad (10)$$

CLAHE is an adaptive histogram equalization technique that undergoes transformation over local regions. Here, a matrix of 8×8 dimension was used for local histogram equalization. The output histogram from the CLAHE transformed image follows the

Rayleigh distribution. Gamma correction was applied to the CLAHE-enhanced image, and finally, the image was complemented. The image complement $f^{-1}(x)$ can be expressed as

$$f^{-1}(x) = 255 - f(x) \quad (11)$$

The three-channel (or RGB) image was constructed by concatenating the original image, the gamma-corrected CLAHE enhanced image, and the complement of the gamma-corrected CLAHE enhanced image.

2.4. Deep Neural Networks

UNet-like architectures with pretrained deep dense, residual, and inception encoder backbones previously showed high performance in both classification and segmentation tasks for 2D chest X-rays [45]. UNet++ with deep dense blocks showed benchmark performance in segmenting lung content from volumetric CT scans [46]. These segmentation networks also performed well in solving complex problems such as detecting intracranial hemorrhages [47]. These studies inspired us to investigate these UNet-like architectures with pretrained encoder backbones for liver segmentation tasks from MR scans. UNet, UNet++, and feature pyramid network (FPN) segmentation networks were investigated with varying depths of dense, residual, and inception encoder backbones. The network architectures are shown in Figure 6. The encoder backbones were pretrained dense, residual, and inception blocks (marked in light orange). The decoder (light blue) uses transpose convolution blocks for upscaling the vector output from the bottleneck (marked in dark blue) output to construct the segmentation mask. The yellow blocks in UNet++ and FPN represent both concatenation and convolution blocks.

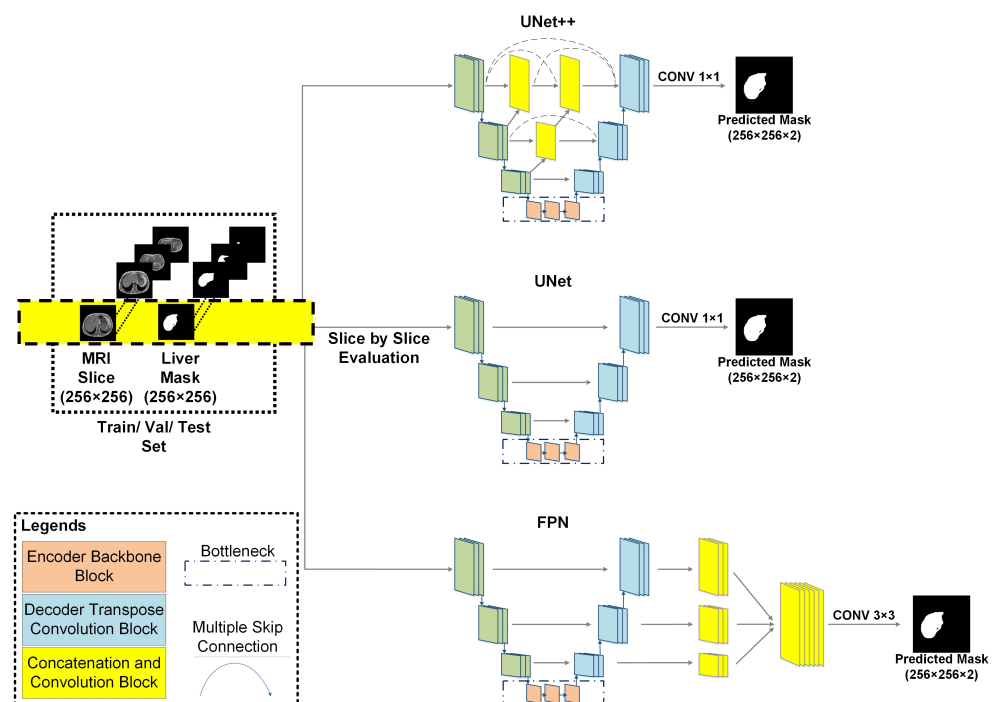


Figure 6. Network architectures of different segmentation networks and investigation frameworks. The varying depth of pretrained dense, residual, and inception encoder backbones were investigated for UNet++, UNet, and FPN segmentation network architectures.

2.4.1. UNet

UNet architecture consists of an encoder and a decoder part. The encoder part reduces the input image size in each of the convolutional blocks through max pooling. In the final encoder block, the two-dimensional original image matrix is reduced to a vector array.

The decoder part upscales the converted vector array in each block through convolutional blocks and upconvolution layers. Lastly, the skip connections among encoder–decoder blocks transfer weights for localizing the region of interest. These skip connections are similar to the attention mechanism [48].

2.4.2. UNet++

UNet++ is an extension of the UNet and wide UNet architecture. It utilizes the concept of deep supervision. UNet++ also introduces nested convolutional blocks inside each skip pathway, and such blocks enhance the quality of feature spaces that are passed to the decoder blocks [49,50].

2.4.3. Feature Pyramid Network (FPN)

In the FPN network, weight connections from the UNet decoder blocks are fed through skip connections to feature pyramid blocks. Further, the output from each feature pyramid block is fed into a single convolutional block. Finally, the output from the convolutional block is fed into a rectilinear unit (ReLU) activation layer for generating the predicted masks [51].

2.4.4. Pretrained Backbones

The concept of transfer learning is utilized to enhance the segmentation performance and reduce the training time. Several pretrained encoders (variants of dense, residual, and inception networks), which were trained on the ImageNet computer vision database [52], were used as the backbones. For each backbone variant, three varying depths were investigated. The variants of DenseNets were DenseNet201, DenseNet161, and DenseNet121 [52–54], while the variants of residual networks were ResNet152, ResNet50, and ResNet18 [55]. InceptionV4 and InceptionResNet were the variants of the inception backbones [56].

2.5. Experiments

Two major experiments were carried out in this study: (i) the generalized model and (ii) the specialized network for handling anatomical ambiguity.

2.5.1. Generalized Model

In this experiment, MRI slices with different liver sizes were used in training and evaluation and the model was generally not specific to any particular liver size. Then, the effects of image enhancement on the generalized model were investigated. A total of 24 networks (three architectures with eight backbones) were tested on three versions of MRI images (i.e., original, gamma-corrected, and 3-channel view) to segment the liver.

2.5.2. Specialized Network for Handling Anatomical Ambiguity

To enhance the performance of the segmentation network in segmenting the liver region from the MRI slices where the liver shape varies, multiple segmentation networks needed to be trained to segment the liver region reliably. A total of 90 slices from the inferior part of the liver and the upper right pole of the kidney were trained separately. Exact preprocessing and processing steps were followed for this set of slices, which was discussed previously. For this specific task, only varying depths of the ResNet encoder–decoder backbones with UNet++, UNet, and FPN were investigated as the ResNet showed better performance in the preliminary study. Three variants of ResNet and Inception-ResnetV2 with three architectures (a total of 12 experiments) were investigated specifically for the slices with small liver contents.

2.5.3. Cascaded Network

Since the liver size varies in the MRI volume, every single generalized model proposed in the literature fails to generalize. Therefore, we propose a cascaded model using a decision function to improve the performance of the segmentation network. The architecture of the

cascaded network is depicted in Figure 7. The volumetric MRI scan is fed into the network slice by slice. At first, a liver mask is predicted from the generalized network. From the first predicted mask, the number of predicted white pixels is calculated. The following equation is used to decide the potential shape of the liver mask in the slice under investigation, where k represents the white pixel count:

$$Liver_Content = \begin{cases} Absent, & \text{if } k = 0 \\ Small, & \text{if } 1 \leq k \leq 750 \\ large, & \text{if } k > 750 \end{cases} \quad (12)$$

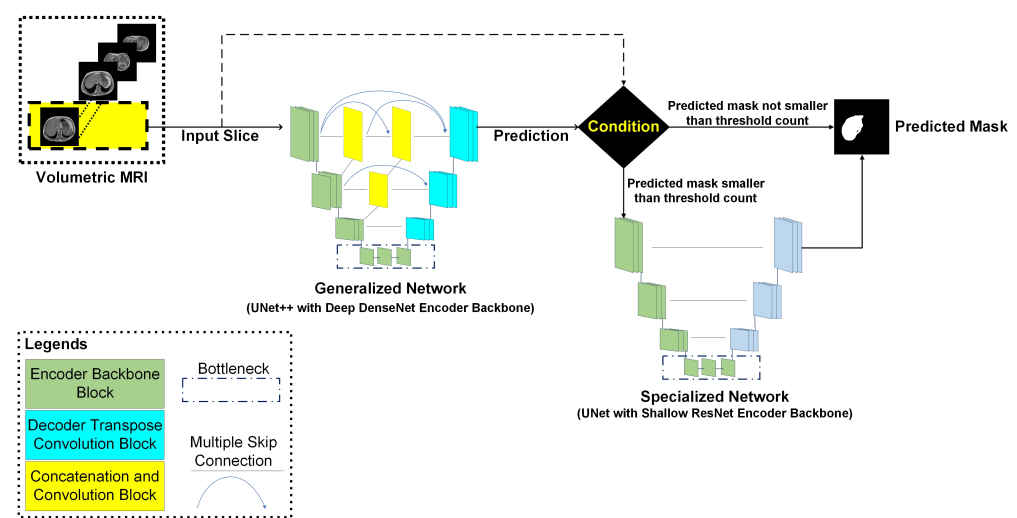


Figure 7. Cascaded network for handling anatomical ambiguity: the generalized network predicts an initial mask; if the pixel count for the predicted mask refers to a constrained or null liver content, then the input slice is fed into the specialized network.

If the number of white pixels is zero, there is no liver in the slice and so the mask is completely black. However, if the decision function identifies a number between 1–750, the slice is again fed into the specialized network for producing the final mask. However, if the number is higher than 750, the mask generated by the generalized model is used as the final liver mask.

The sets of large, small, and absent liver contents are created on the basis of the topographic visualization of the abdominal anatomy, which was described earlier in Section 2.2.1. The liver content is maximum in the axial views from the middle part of the liver. Moreover, the liver content is medium and constrained in the axial views from the superior part of the liver and the inferior part of the liver, respectively. In the axial view from the upper part of the kidney, the liver content is absent. In this perspective, the set of large liver content is constructed with the axial views from the middle part and superior part of the liver. The axial views from the inferior part of the liver are represented in the set of small liver content. Lastly, the set of absent liver content is formed by the axial views from the upper part of the right kidney. The threshold values are then determined by analyzing the pixel counts in each of the sets.

Generally, the axial views from the superior part and inferior part of the liver have significant liver content and the liver area can be comfortably segmented. However, ambiguity arises for the axial views from the inferior part of the liver and the upper part of the right kidney, as the liver portion is significantly constrained. In such a perspective, segmentation performance may improve if slices from these two complicated axial views are handled with a separate network, which is only trained with such cases. Thus, such a cascaded approach was investigated.

2.6. Loss Function

Binary cross-entropy (BCE) loss is typically used for classification tasks. As any semantic segmentation task can be considered as a classification task at the pixel level, this loss is also effective for segmentation. BCE loss can be expressed by [57,58]

$$Loss_{(BCE)} = \frac{1}{N} \sum_{i=0}^{N-1} -(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (13)$$

The dice coefficient is used to calculate the similarity index between ground truth and predicted masks for segmentation tasks. Dice loss is a region-based loss function and it is introduced in [59]. Dice loss can be expressed by

$$Loss_{(DICE)} = 1 - \frac{\sum_{i=0}^{N-1} y_i \hat{y}_i}{\sum_{i=0}^{N-1} y_i^2 + \sum_{i=0}^{N-1} \hat{y}_i^2 + \epsilon} \quad (14)$$

In Equations (13) and (14), N represents the total number of pixels, y_i represents the i th pixel in the ground truth mask, and \hat{y}_i represents the i th pixel

Initially, both the mentioned loss functions were investigated to find the optimum solution. However, the detailed investigation was carried out with the BCE loss, as it demonstrated superior performance over dice loss in the initial investigation.

2.7. Training Parameters

In order to conduct a uniform comparison among the network performances, it was indispensable to use the same training parameters for all the networks. All the training was conducted in an NVIDIA Tesla P100-PCIE graphics processing unit (GPU) with 16 gigabytes (GB) of memory. The initial learning rate was set to 0.0001 with a learning factor LR of 0.02. If the validation loss did not show significant changes in 10 epochs, the learning rate was reduced by $\frac{1}{LR}$. The maximum epoch number was set to be 100 for each fold, but if the validation loss was constant for 20 epochs, the training was terminated. Z-score normalization was used, which uses the standard deviation and mean of the raw MRI slices for normalizing each image. For optimizing the process of gradient descent, in each of the epochs, the ADAM optimization algorithm was used as it showed superior performance over stochastic gradient descent (SGD) in the initial investigation [60,61].

2.8. Evaluation Metrics

For evaluating the performance of each investigated network, the accuracy, dice similarity coefficient (DSC) (i.e., F1-score), and intersection of union (IoU) were computed. For each network, the average metrics for each ground truth mask and predicted mask pairs were calculated. Accuracy, DSC, and IoU can be expressed by

$$Accuracy = \frac{TP + TN}{2 \times TP + TN + FP + FN} \quad (15)$$

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (16)$$

$$IoU = \frac{TP}{2 \times TP + FP + FN} \quad (17)$$

In Equations (15), (16), and (17), TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative, respectively.

2.9. Cloud Deployment

A cloud-based application for real-time liver segmentation from MRI images was deployed. The deep learning model was deployed in the cloud back-end server, which runs on an 8-core, 32 GB Memory Apache Linux instance hired from Google Cloud Perform (GCP). The back-end server was connected to a SQL database for storing the MRI images. The application is cross-platform compatible and users can access the application anytime via a web browser from any edge device. The cloud-based application can be remotely connected with a Picture Archiving System (PACS) for assisting radiologists in liver pathology investigations. In order to provide more convenient remote access for clinicians, an Android application was also developed. Figure 1 superficially describes the cloud application. To ensure the robustness of the segmentation network, an automated self-learning scheduler was implemented in the back-end server following the concept discussed [62]. The scheduler automatically retrains the deployed model with the incoming new data provided by the user, and such an approach boosts the network's performance on unseen real-world data.

3. Results and Discussion

The results from different investigations are discussed in this section. Later, a performance comparison is presented, which compares the efficiency of the proposed approach with the reported high-performance techniques in the literature for liver segmentation on the same MRI test set. In the following section, the performance of the generalized model, specialized model, and cascaded models are presented.

3.1. Generalized Model

The Table 2 summarizes the network performance for segmenting the liver region in the MR slices using a generalized model. Deep networks showed superiority in performance in comparison to shallow networks. On the nonenhanced images (original image), UNet++ with dense backbones showed the top performance. UNet++ with a DenseNet201 backbone showed the best performance with a DSC and IoU of 94.3% and 91.00%, respectively. On the nonenhanced images, UNet with different DenseNet backbones exhibited a similar performance.

For the three-channel image set, networks with DenseNet backbones demonstrated slightly better performance compared to other networks. Among the variants of the DenseNet model, DenseNet161 performed the best for this specific image set. Both FPN and UNet with DenseNet161 backbones achieved a DSC of over 93%. Among the investigated image enhancement techniques, the gamma-enhanced image set performed the worst.

The liver content is maximum in the slices showing the middle part of the liver and also significantly larger in the superior part of the liver. For these two specific types of slices, all of the DenseNet backbones showed excellent performances. Figure 8a shows the predicted liver masks for the slices from the middle part of the liver. The figure shows that all of the networks can segment the liver region accurately.

3.2. Effects of Image Enhancement for Generalized Model

It can be observed that the network performances were slightly decreased when image enhancement techniques were implemented (Table 2). This is due to the ambiguity that arises from the slices where the liver content varies widely. Though image enhancement was very effective for the slices where the liver portion was significant, the performance dropped when the liver size was minimum in the slice of investigation. Figure 8b shows such a sample liver slice with the ground truth mask and the masks predicted by different models.

Due to this ambiguity, finding a generalized image enhancement technique for such a complex and varying anatomy is very challenging.

3.3. Limitation of the Generalized Model

In the case of the slices of the middle part of the liver, all ResNet and inception backbones demonstrated satisfactory performance. Figure 9 shows the predicted masks from top-performing networks for the middle part of the liver (large liver content), inferior part of the liver (small liver content), and upper pole of the kidney (no liver content) for the original T_1 -weighted images. Figure 9 shows that the generalized model performed well for the slices with large liver content and for the slices where the liver was absent. However, when the liver content was small, the generalized model struggled to locate the liver area precisely.

Table 2. Summary of the investigated network performances from the generalized approach. UNet++ with DenseNet201 encoder exhibited the best performance.

Networks		Original			Three Channel			Gamma Corrected		
Architecture	Backbone	Acc. (%)	IoU (%)	DSC (%)	Acc. (%)	IoU (%)	DSC (%)	Acc. (%)	IoU (%)	DSC (%)
UNet++	DenseNet201	99.73	91.00	94.30	99.60	88.95	92.35	99.42	89.28	91.91
	DenseNet161	99.68	89.78	93.06	99.71	89.60	92.95	99.66	87.00	90.30
	DenseNet121	99.43	89.17	92.57	99.66	90.08	93.40	99.56	87.58	90.92
	ResNet152	99.70	89.79	93.13	99.67	87.97	91.34	99.70	89.00	92.33
	ResNet50	99.70	90.42	93.81	99.68	89.54	92.98	99.66	88.13	91.64
	ResNet18	99.70	89.73	93.08	99.70	89.63	93.01	99.63	84.55	88.50
	Inception-resnet-v2	99.71	89.16	92.57	99.65	88.31	92.10	99.70	89.60	91.98
	inception-v4	99.70	87.98	91.29	99.68	89.23	92.62	99.70	89.79	92.16
UNet	DenseNet201	99.76	89.98	93.22	99.72	88.77	92.13	99.78	88.74	92.18
	DenseNet161	99.57	90.48	93.84	99.43	90.08	93.60	99.45	87.58	90.92
	DenseNet121	99.43	89.88	93.27	99.66	89.48	92.90	99.64	87.04	90.31
	ResNet152	99.69	89.46	92.97	99.67	88.66	92.25	99.67	88.91	92.35
	ResNet50	99.68	87.48	90.93	99.66	85.49	89.01	99.68	88.79	92.36
	ResNet18	99.67	88.16	91.83	99.67	86.83	90.38	99.68	88.77	92.31
	Inception-resnet-v2	99.66	87.68	91.41	99.68	88.20	91.80	99.70	87.81	91.32
	inception-v4	99.68	88.64	92.34	99.70	90.68	93.47	99.62	87.89	91.71
FPN	DenseNet201	99.65	89.45	92.83	99.36	89.50	92.97	99.47	88.33	91.87
	DenseNet161	99.70	89.38	92.77	99.66	89.32	93.00	99.53	88.11	91.85
	DenseNet121	99.47	87.52	91.08	99.47	89.39	92.94	99.71	86.91	90.49
	ResNet152	99.66	88.49	92.08	99.67	88.90	92.59	99.68	87.85	91.46
	ResNet50	99.69	89.01	92.52	99.65	88.15	91.88	99.66	88.76	92.40
	ResNet18	99.68	88.33	91.91	99.66	88.10	91.95	99.67	88.58	92.33
	Inception-resnet-v2	99.61	87.03	91.46	99.67	88.17	92.08	99.65	88.52	92.39
	inception-v4	99.62	85.52	90.85	99.66	88.64	92.55	99.66	88.64	92.55

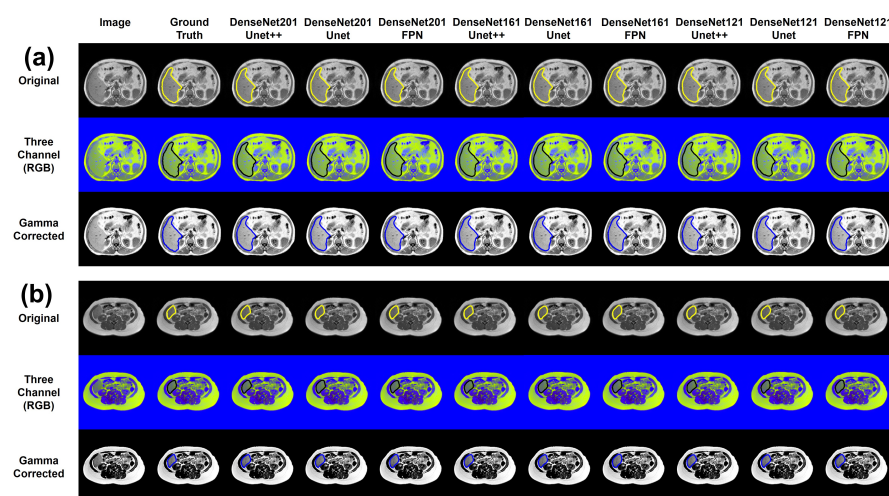


Figure 8. Visualization of the predicted masks from the networks with DenseNet backbones for a sample axial slice showing the middle part of the liver (a) and the inferior part of the liver (b).

A more detailed picture is shown in Table 3. The table illustrates the fold-wise slice distribution and the observed DSC for each of the different groups of liver shapes for the top-performing UNet++ model with the DenseNet201 backbone. Though the percentage of slices from the middle part of the liver was the minimum for all the folds, the DSC value for the best-performing model was still over 95% for each fold. Slices with medium liver content occurred the most and the DSC value for each fold was around 95%. The network also efficiently handled slices where the liver content is absent.

However, the model performance was greatly reduced for the slices where liver content was small. It is worth mentioning that the number of such slices in the training set was also insignificant. Our hypothesis was that handling such slices by a separate model may improve the overall segmentation performance, which is explored in this study.

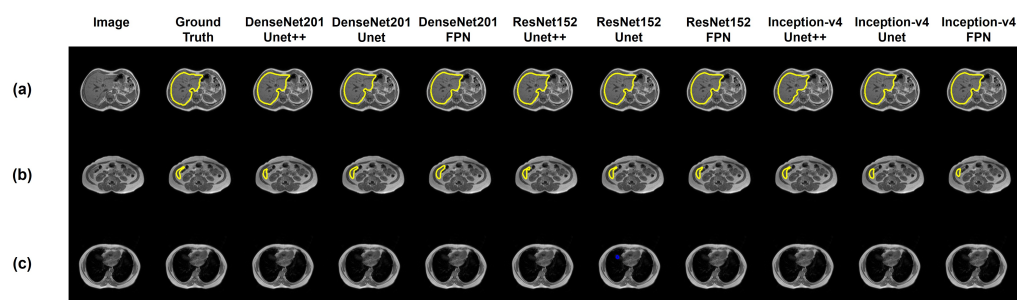


Figure 9. Visualization of the predicted masks from selected networks for a sample slice in (a) middle part of the liver (large liver content), (b) inferior part of the liver (small liver content), and (c) upper pole of the kidney (no liver content).

Table 3. Distribution of slices of distinct axial views in train set and test set, along with the observed DSC. Slices depicting axial view from the inferior part of the liver holds a constrained liver content and exhibits anatomical ambiguity.

Fold No	Middle Part of Liver (Liver Content: Large)			Superior Part of Liver (Liver Content: Medium)			Inferior Part of Liver (Liver Content: Small)			Upper Part of Kidney (Liver Content: Absent)		
	Train Set Slice %	Test Set Slice %	DSC (%)	Train Set Slice %	Test Set Slice %	DSC (%)	Train Set Slice %	Test Set Slice %	DSC (%)	Train Set Slice %	Test Set Slice %	DSC (%)
1	7.74%	19.38%	97.03%	45.63%	34.89%	95.33%	12.71%	16.28%	81.95%	34.12%	29.46%	100.00%
2	7.73%	11.63%	95.75%	44.36%	41.86%	95.11%	13.25%	11.63%	82.64%	34.66%	34.88%	95.55%
3	6.69%	17.83%	96.17%	43.43%	41.86%	95.20%	11.79%	12.40%	78.13%	35.85 %	30.23%	97.43%
4	7.67%	10.85%	95.70%	43.26%	42.63%	95.23%	14.90%	9.30%	80.20%	34.18%	37.21%	97.91%
5	6.86%	16.79%	97.35%	44.25%	38.17%	93.90%	14.15%	9.16%	82.46%	34.75%	35.88%	94.78%

3.4. Specialized Network for Handling Anatomical Ambiguity

Table 4 illustrates the performance of the specialized models trained with different architectures and different backbones in comparison to the best-performing generalized model in segmenting the MR slices with a small liver content.

The UNet with ResNet18 encoder backbone showed superior performance over the other investigated networks, with an IoU and DSC of 77.00% and 86.22%, respectively. For UNet++, the Inception-resnet-V2 encoder backbone showed better performance over the varying depths of ResNet backbones with an IoU and DSC of 75.58% and 84.03%, respectively. The shallow ResNet18 backbone performed better for the FPN architecture over other pretrained encoder backbones with an IoU and DSC of 71.20% and 82.04%, respectively. Each of the top-performing encoder backbones for UNet, UNet++, and FPN performed better than the top-performing generalized network for this task, which demonstrated an IoU and DSC of 70.74% and 80.88%, respectively. Lastly, the shallow networks performed better compared to deep networks for this specific task, as the liver content in the slice was small.

3.5. Cascaded Network

Figure 10 shows the predicted masks from the proposed cascaded network. It can be observed that such an approach enhances the mask quality for slices with a small liver content. Combining both the generalized and specialized network enhances the performance of the network for segmenting the liver region. Table 5 summarizes the performance metrics for the generalized network and the cascaded network. Cascading both the networks improves the overall DSC score (from 94.3 to 95.15%).

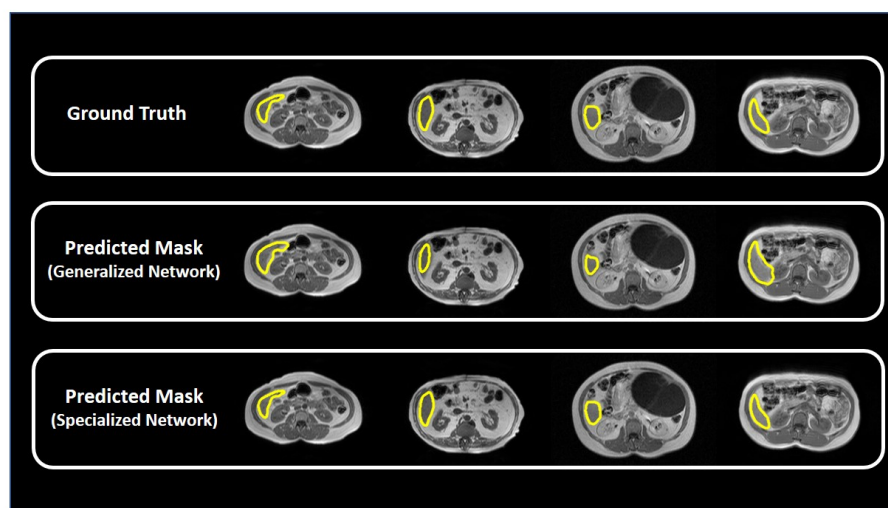


Figure 10. Comparison of the predicted masks from the generalized and specialized networks for sample MR slices with small liver content.

Table 4. Summary of the investigated network performance for the slices with small liver content using different specialized models and the best-performing generalized model. UNet with ResNet18 backbone showed improved performance for the task.

Networks		Metrics (Specialized Network)			Metrics (Best-Performing Generalized Network)		
Architecture	Backbone	Acc. (%)	IoU (%)	DSC (%)	Acc. (%)	IoU (%)	DSC (%)
UNet	ResNet18	99.64	77.00	86.22	99.76	70.74	80.88
	ResNet50	99.81	72.06	80.94			
	ResNet152	99.78	70.00	79.73			
	Inception-resnet-v2	99.70	72.73	81.72			
UNet++	ResNet18	99.78	71.71	78.38	99.76	70.74	80.88
	ResNet50	99.76	71.62	80.96			
	ResNet152	99.80	71.89	81.02			
	Inception-resnet-v2	99.78	75.58	84.03			
FPN	ResNet18	99.80	71.20	82.04	99.76	70.74	80.88
	ResNet50	99.77	69.75	79.77			
	ResNet152	99.78	71.86	81.20			
	Inception-resnet-v2	99.80	70.92	80.41			

Table 5. Performance metrics for the best-performing generalized and cascaded network. Here, the results of the cascaded network are marked in gray.

Experiments	Acc. (%)	IoU (%)	DSC (%)
Generalized Network	99.73%	91.00%	94.30%
Cascaded Network	99.70%	92.10%	95.15%

3.6. Discussion

The performance comparison of our proposed framework with all of the existing high-performance networks (using the same test for evaluation) is summarized in Table 6. X. Zhong et al. [25] investigated deep action learning for abdominal organ segmentation tasks from volumetric MRI images. Their proposed network demonstrated superiority

over 3D UNet in terms of overall performance, and achieved a DSC of 80.6% for the liver segmentation task. P. Pandey et al. [26] explored a contrastive semisupervised approach for the same task, and it achieved a DSC of 85.9%. The proposed method generates patches for each slice, which enhances the feature space. Mitta et al. [28] achieved a DSC of 88.12% on the test set by using W-Net with attention gates. J. Hong et al. [29] and X. Wang et al. [30] used source-free unsupervised learning and bidirectional searching for the segmentation task, respectively. By using geometric edge enhancement, S. Mulay et al. [31] boosted the performance of the mask R-CNN for the liver segmentation task on this test set. L. Zbinden et al. [32] achieved a DSC of 93.60% by implementing nnUNet on T_1 -weighted MRI slices.

Our proposed cascaded framework outperforms all of these existing high-performance techniques by a large margin with a DSC of 95.15%. As discussed previously, the size of the liver content in an arbitrary MRI slice depends on its axial view source. Any generalized segmentation network can perform comparatively better when the liver content is significant in the given MRI slice (axial view from the middle part of the liver). On the contrary, the network faces ambiguity when the liver content is reduced for the given MRI slice (axial view from the upper pole of the kidney, the inferior pole of the kidney, and the superior part of the liver). As a result, the network performance drops significantly for these specific groups of slices where the liver content is small. This specific cause for reduced segmentation performance is overlooked in all of the previous studies. Our proposed framework separately handles this specific group of slices with a small liver content, which generates ambiguity through a specialized network, thus enhancing the overall segmentation performance.

Table 6. Comparison of the proposed method (marked in gray) with existing studies that used the same testing set.

Authors	Methodology and Approach	Metric (DSC)
X. Zhong et al. [25]	Deep action learning with 3D UNet	80.60 ± 5.30%
P. Pandey et al. [26]	Contrastive Semi Supervised Learning Approach with UNet	85.90%
D. Mitta et al. [28]	W-Net with attention gates	88.12%
J. Hong et al. [29]	Source Free Unsupervised UNet	88.40%
X. Wang et al. [30]	Bidirectional Searching Neural Net	89.80%
S. Mulay et al. [31]	Mask R-CNN	80.00%
	Geometric Edge Enhancement based Mask R-CNN	91.00%
L. Zbinden et al. [32]	nnUNet	93.60%
Proposed	Cascaded Network for Handling Anatomical Ambiguity	95.15%

4. Conclusions

Abdominal organ segmentation is a challenging task due to the complexity of the anatomy of the abdominal area and the close proximity of multiple organs. Ambiguity in the segmentation of the liver arises due to the variance in its anatomical shape in the MRI volume. The MRI modality is favored by clinicians for liver pathology diagnosis. However, automated liver segmentation from MRI scans is a demanding task. In this research, we proposed a novel cascaded network for liver segmentation from T_1 -weighted MR images. The proposed network treats each axial view distinctly and achieved a DSC of 95.15% on the publicly available CHAOS MRI dataset. Such an approach can also be investigated for other abdominal organ segmentation tasks, such as those involving the kidneys and spleen. The proposed network was also deployed as an open-source application in a cloud server for demonstration purposes. This application can later be integrated with PACS for clinical usage. Lastly, we also investigated the effects of different image enhancement techniques for liver segmentation tasks from MR scans.

Author Contributions: Conceptualization, M.S.A.H., M.E.H.C., M.S.K. and M.H.; methodology, M.S.A.H., M.E.H.C., S.G., M.S.K., E.H.B., M.H. and A.S.; software, M.S.I.S., M.S.A.H., S.G., M.A.A., A.A. and S.M.; validation, M.S.A.H., A.K., I.A.-H. and E.H.B.; formal analysis, M.S.A.H.; investigation, M.S.A.H., E.H.B., M.H., I.A.-H. and M.E.H.C.; resources, M.S.K. and A.K.; data curation, M.S.A.H., S.G., A.S., S.M. and I.A.-H.; writing—original draft preparation, M.S.A.H., S.G. and S.M.; writing—review and editing, all authors; visualization, M.S.A.H., M.S.I.S., A.K., M.A.A. and A.A.;

supervision, M.E.H.C., M.S.K., A.A., M.A.A. and M.H.; project administration, M.E.H.C., M.S.K., A.A. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Qatar University High Impact grant QUHI-CENG-23/24-216 and student grant QUST-1-CENG-2023-796 and is also supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/2023/R/1444). The statements made herein are solely the responsibility of the authors. The open-access publication cost is covered by the Qatar National Library.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This article is based on previously conducted studies and does not contain any studies with human participants or animals. The dataset used in this research is open access and can be accessed from <https://chaos.grand-challenge.org/Data/>, accessed on 9 August 2023. Kavuar et al. [27] followed standard protocol for creating this open-access dataset.

Acknowledgments: The authors would like to thank Kavuar et al. [27] for sharing the dataset publicly which make the research possible.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Aggarwal, R.; Sounderajah, V.; Martin, G.; Ting, D.S.; Karthikesalingam, A.; King, D.; Ashrafi, H.; Darzi, A. Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *NPJ Digit. Med.* **2021**, *4*, 65. [CrossRef] [PubMed]
- Barragán-Montero, A.; Javaid, U.; Valdés, G.; Nguyen, D.; Desbordes, P.; Macq, B.; Willems, S.; Vandewinckele, L.; Holmström, M.; Löfman, F.; et al. Artificial intelligence and machine learning for medical imaging: A technology review. *Phys. Med.* **2021**, *83*, 242–256. [CrossRef]
- Chen, Z.; Song, Y.; Chang, T.-H.; Wan, X. Generating Radiology Reports via Memory-Driven Transformer. *arXiv* **2020**, arXiv:2010.16056.
- Tahir, A.M.; Chowdhury, M.E.; Khandakar, A.; Rahman, T.; Qiblawey, Y.; Khurshid, U.; Kiranyaz, S.; Ibtehaz, N.; Rahman, M.S.; Al-Maadeed, S.; et al. COVID-19 infection localization and severity grading from chest X-ray images. *Comput. Biol. Med.* **2021**, *139*, 105002. [CrossRef] [PubMed]
- Abbas, T.O.; AbdelMoniem, M.; Khalil, I.; Hossain, M.S.A.; Chowdhury, M.E. Deep learning based automated quantification of urethral plate characteristics using the plate objective scoring tool (POST). *arXiv* **2023**, arXiv:2209.13848.
- Dou, Q.; Chen, H.; Jin, Y.; Yu, L.; Qin, J.; Heng, P.A. 3D Deeply Supervised Network for Automatic Liver Segmentation from CT Volumes. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: 19th International Conference, Athens, Greece, 17–21 October 2016; Proceedings, Part II 19; Springer: Cham, Switzerland, 2016. [CrossRef]
- Chen, C.; Zhou, K.; Zha, M.; Qu, X.; Guo, X.; Chen, H.; Wang, Z.; Xiao, R. An effective deep neural network for lung lesions segmentation from COVID-19 CT images. *IEEE Trans. Ind. Inform.* **2021**, *17*, 6528–6538. [CrossRef]
- Chen, C.; Xiao, R.; Zhang, T.; Lu, Y.; Guo, X.; Wang, J.; Chen, H.; Wang, Z. Pathological lung segmentation in chest CT images based on improved random walker. *Comput. Methods Programs Biomed.* **2021**, *200*, 105864. [CrossRef]
- Hu, P.; Wu, F.; Peng, J.; Liang, P.; Kong, D. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. *Phys. Med. Biol.* **2016**, *61*, 8676. [CrossRef]
- Liu, X.; Song, L.; Liu, S.; Zhang, Y. A review of deep-learning-based medical image segmentation methods. *Sustainability* **2021**, *13*, 1224. [CrossRef]
- Yang, R.; Yu, Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Front. Oncol.* **2021**, *11*, 638182. [CrossRef]
- Liu, L.; Wolterink, J.M.; Brune, C.; Veldhuis, R.N.J. Anatomy-aided deep learning for medical image segmentation: A review. *Phys. Med. Biol.* **2021**, *66*, 11TR01. [CrossRef] [PubMed]
- Peng, J.; Wang, Y.; Kong, D. Liver segmentation with constrained convex variational model. *Pattern Recognit. Lett.* **2014**, *43*, 81–88. [CrossRef]
- Liu, Z.; Song, Y.Q.; Sheng, V.S.; Wang, L.; Jiang, R.; Zhang, X.; Yuan, D. Liver CT sequence segmentation based with improved U-Net and graph cut. *Expert Syst. Appl.* **2019**, *126*, 54–63. [CrossRef]
- Tang, X.; Jafarholi Rangraz, E.; Coudyzer, W.; Bertels, J.; Robben, D.; Schramm, G.; Deckers, W.; Maleux, G.; Baete, K.; Verslype, C.; et al. Whole liver segmentation based on deep learning and manual adjustment for clinical use in SIRT. *Eur. J. Nucl. Med. Mol. Imag.* **2020**, *47*, 2742–2752. [CrossRef]
- Kim, K.; Chun, J. A new hyper parameter of hounsfield unit range in liver segmentation. *J. Internet Comput. Serv.* **2020**, *21*, 103–111. [CrossRef]
- Xiang, K.; Jiang, B.; Shang, D. The overview of the deep learning integrated into the medical imaging of liver: A review. *Hepatol. Int.* **2021**, *15*, 868–880. [CrossRef]

18. Elbanna, K.Y.; Kielar, A.Z. Computed Tomography Versus Magnetic Resonance Imaging for Hepatic Lesion Characterization/Diagnosis. *Clin. Liver Dis.* **2021**, *17*, 159. [\[CrossRef\]](#)
19. Coenegrachts, K. Magnetic resonance imaging of the liver: New imaging strategies for evaluating focal liver lesions. *World J. Radiol.* **2009**, *1*, 72. [\[CrossRef\]](#)
20. Caseiro-Alves, F.; Brito, J.; Araujo, A.E.; Belo-Soares, P.; Rodrigues, H.; Cipriano, A.; Sousa, D.; Mathieu, D. Liver haemangioma: Common and uncommon findings and how to improve the differential diagnosis. *Eur. Radiol.* **2007**, *17*, 1544–1554. [\[CrossRef\]](#)
21. Wang, G.; Zhu, S.; Li, X. Comparison of values of CT and MRI imaging in the diagnosis of hepatocellular carcinoma and analysis of prognostic factors. *Oncol. Lett.* **2019**, *17*, 1184–1188. [\[CrossRef\]](#)
22. Gibbs, J.F.; Litwin, A.M.; Kahlenberg, M.S. Contemporary management of benign liver tumors. *Surg. Clin. N. Am.* **2004**, *84*, 463–480. [\[CrossRef\]](#)
23. Mostafa, A.; Hassanien, A.E.; Houseni, M.; Hefny, H. Liver segmentation in MRI images based on whale optimization algorithm. *Multimed. Tools Appl.* **2017**, *76*, 24931–24954. [\[CrossRef\]](#)
24. Hänsch, A.; Chlebus, G.; Meine, H.; Thielke, F.; Kock, F.; Paulus, T.; Abolmaali, N.; Schenk, A. Improving automatic liver tumor segmentation in late-phase MRI using multi-model training and 3D convolutional neural networks. *Sci. Rep.* **2022**, *12*, 12262. [\[CrossRef\]](#)
25. Zhong, X.; Amrehn, M.; Ravikumar, N.; Chen, S.; Strobel, N.; Birkhold, A.; Kowarschik, M.; Fahrig, R.; Maier, A. Deep action learning enables robust 3D segmentation of body organs in various CT and MRI images. *Sci. Rep.* **2021**, *11*, 3311. [\[CrossRef\]](#)
26. Pandey, P.; Pai, A.; Bhatt, N.; Das, P.; Makharia, G.; Ap, P. Contrastive semi-supervised learning for 2D medical image segmentation. *arXiv* **2021**. arXiv:2106.06801.
27. Kavur, A.E.; Gezer, N.S.; Barış, M.; Aslan, S.; Conze, P.H.; Groza, V.; Pham, D.D.; Chatterjee, S.; Ernst, P.; Özkan, S.; et al. CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation. *Med. Image Anal.* **2021**, *69*, 101950. [\[CrossRef\]](#)
28. Mitta, D.; Chatterjee, S.; Speck, O.; Nürnberger, A. Upgraded w-net with attention gates and its application in unsupervised 3d liver segmentation. *arXiv* **2020**, arXiv:2011.10654.
29. Hong, J.; Zhang, Y.D.; Chen, W. Source-free unsupervised domain adaptation for cross-modality abdominal multi-organ segmentation. *Knowl. Based Syst.* **2022**, *250*, 109155. [\[CrossRef\]](#)
30. Wang, X.; Xiang, T.; Zhang, C.; Song, Y.; Liu, D.; Huang, H.; Cai, W. Bix-Nas: Searching Efficient Bi-Directional Architecture for Medical Image Segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, 27 September–1 October 2021; Proceedings, Part I 24; Springer International Publishing: Cham, Switzerland, 2021; pp. 229–238. [\[CrossRef\]](#)
31. Mulay, S.; Deepika, G.; Jeevakala, S.; Ram, K.; Sivaprakasam, M. Liver Segmentation from Multimodal Images Using HED-Mask R-CNN. In Proceedings of the Multiscale Multimodal Medical Imaging: First International Workshop, MMMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, 13 October 2019; Proceedings 1; Springer International Publishing: Cham, Switzerland, 2019; pp. 68–75. [\[CrossRef\]](#)
32. Zbinden, L.; Catucci, D.; Suter, Y.; Berzigotti, A.; Ebner, L.; Christe, A.; Obmann, V.C.; Schnitzman, R.; Huber, A.T. Convolutional neural network for automated segmentation of the liver and its vessels on non-contrast T1 vibe Dixon acquisitions. *Sci. Rep.* **2022**, *12*, 22059. [\[CrossRef\]](#)
33. Netter, F.H. *Section 4: Atlas of Human Anatomy*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2003; pp. 239–338. [\[CrossRef\]](#)
34. Suetens, P. Chapter 4—Magnetic Resonance Imaging. In *Fundamentals of Medical Imaging*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017; pp. 64–104. [\[CrossRef\]](#)
35. de Bazelaire, C.M.; Duhamel, G.D.; Rofsky, N.M.; Alsop, D.C. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. *Radiology* **2004**, *230*, 652–659. [\[CrossRef\]](#)
36. Dimakis, N. Chapter 5—Magnetic Resonance Imaging (MRI). In *Introduction to Medical Imaging—Physics, Engineering and Clinical Applications*, 1st ed.; Cambridge University Press: Cambridge, UK, 2011; pp. 204–273. [\[CrossRef\]](#)
37. Rahman, T.; Chowdhury, M.E.; Khandakar, A.; Mahbub, Z.B.; Hossain, M.S.A.; Alhatou, A.; Abdalla, E.; Muthiyil, S.; Islam, K.F.; Kashem, S.B.A.; et al. BIO-CXRNET: A robust multimodal stacking machine learning technique for mortality risk prediction of COVID-19 patients using chest X-ray images and clinical data. *Neural Comput. Appl.* **2023**, *35*, 17461–17483. [\[CrossRef\]](#)
38. Hossain, S.A.; Rahman, M.A.; Chakrabarty, A.; Rashid, M.A.; Kuwana, A.; Kobayashi, H. Emotional State Classification from MUSIC-Based Features of Multichannel EEG Signals. *Bioengineering* **2023**, *10*, 99. [\[CrossRef\]](#)
39. Hossain, S.A.; Rahman, M.A.; Chakrabarty, A. MUSIC Model Based Neural Information Processing for Emotion Recognition from Multichannel EEG Signal. In Proceedings of the 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 26–27 August 2021; pp. 955–960. [\[CrossRef\]](#)
40. Nalepa, J.; Marcinkiewicz, M.; Kawulok, M. Data augmentation for brain-tumor segmentation: A review. *Front. Comput. Neurosci.* **2019**, *13*, 83. [\[CrossRef\]](#)
41. Safdar, M.F.; Alkobaisi, S.S.; Zahra, F.T. A Comparative Analysis of Data Augmentation Approaches for Magnetic Resonance Imaging (MRI) Scan Images of Brain Tumor. *Acta Inform. Med.* **2020**, *28*, 29–36. [\[CrossRef\]](#)
42. Islam, K.R.; Kumar, J.; Tan, T.L.; Reaz, M.B.I.; Rahman, T.; Khandakar, A.; Abbas, T.; Hossain, M.S.A.; Zughaier, S.M.; Chowdhury, M.E.H. Prognostic Model of ICU Admission Risk in Patients with COVID-19 Infection Using Machine Learning. *Diagnostics* **2022**, *12*, 2144. [\[CrossRef\]](#)

43. Mahmud, S.; Ibtehaz, N.; Khandakar, A.; Rahman, M.S.; Gonzales, A.J.; Rahman, T.; Hossain, M.S.; Hossain, M.S.A.; Faisal, M.A.A.; Abir, F.F.; et al. NABNet: A Nested Attention-guided BiConvLSTM network for a robust prediction of Blood Pressure components from reconstructed Arterial Blood Pressure waveforms using PPG and ECG signals. *Biomed. Signal Process. Control* **2023**, *79*, 104247. [\[CrossRef\]](#)
44. TRahman, T.; Khandakar, A.; Qiblawey, Y.; Tahir, A.; Kiranyaz, S.; Kashem, S.B.A.; Islam, M.T.; Al Maadeed, S.; Zughaier, S.M.; Khan, M.S.; et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Comput. Biol. Med.* **2021**, *132*, 104319. [\[CrossRef\]](#)
45. Rahman, T.; Khandakar, A.; Kadir, M.A.; Islam, K.R.; Islam, K.F.; Mazhar, R.; Hamid, T.; Islam, M.T.; Kashem, S.; Mahbub, Z.B.; et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *IEEE Access* **2020**, *8*, 191586–191601. [\[CrossRef\]](#)
46. Qiblawey, Y.; Tahir, A.; Chowdhury, M.E.; Khandakar, A.; Kiranyaz, S.; Rahman, T.; Ibtehaz, N.; Mahmud, S.; Maadeed, S.A.; Musharavati, F.; et al. Detection and severity classification of COVID-19 in CT images using deep learning. *Diagnostics* **2021**, *11*, 893. [\[CrossRef\]](#)
47. Khan, M.M.; Chowdhury, M.E.H.; Arefin, A.S.M.S.; Podder, K.K.; Hossain, M.S.A.; Alqahtani, A.; Murugappan, M.; Khandakar, A.; Mushtak, A.; Nahiduzzaman, M. A Deep Learning-Based Automatic Segmentation and 3D Visualization Technique for Intracranial Hemorrhage Detection Using Computed Tomography Images. *Diagnostics* **2023**, *13*, 2537. [\[CrossRef\]](#)
48. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. [\[CrossRef\]](#)
49. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA ML-CDS)*; Springer: Cham, Switzerland, 2018. [\[CrossRef\]](#)
50. Lee, C.-Y.; Xie, S.; Gallagher, P.; Zhang, Z.; Tu, Z. Deeply-Supervised Nets. In Proceedings of the 18th International Conference on Artificial Intelligence and Statistics, San Diego, CA, USA, 9–12 May 2015; pp. 562–570.
51. Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 936–944. [\[CrossRef\]](#)
52. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
53. Wang, S.; Zhang, Y. DenseNet-201-Based Deep Neural Network with Composite Learning Factor and Precomputation for Multiple Sclerosis Classification. *ACM Trans. Multimed. Comput. Commun. Appl.* **2020**, *16*, 3341095. [\[CrossRef\]](#)
54. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A Survey on Deep Transfer Learning. In *Artificial Neural Networks and Machine Learning—ICANN 2018*; Springer International Publishing: Cham, Switzerland, 2018; pp. 270–279.
55. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [\[CrossRef\]](#)
56. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. February. Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; IEEE: Piscataway, NJ, USA, 2017.
57. Jadon, S. A Survey of Loss Functions for Semantic Segmentation. In Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar, Chile, 27–29 October 2020; pp. 1–7. [\[CrossRef\]](#)
58. Yi-de, M.; Qing, L.; Zhi-Bai, Q. Automated Image Segmentation Using Improved PCNN Model Based on Cross-Entropy. In Proceedings of the 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, China, 20–22 October 2004; pp. 743–746. [\[CrossRef\]](#)
59. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Jorge Cardoso, M. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Cham, Switzerland, 2017; pp. 240–248. [\[CrossRef\]](#)
60. Diederik, P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
61. Ketkar, N. Stochastic Gradient Descent. In *Deep Learning with Python*; Apress: New York, NY, USA, 2017; pp. 113–132.
62. Ravandi, B.; Papapanagiotou, I. A Self-Learning Scheduling in Cloud Software Defined Block Storage. In Proceedings of the 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), Honolulu, HI, USA, 25–30 June 2017; pp. 415–422. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.