An Internet of Medical Things Based Liver Tumor Detection System using Semantic Segmentation

Mehr Yahya Durrani[1], Sadaf Yasmin[1], Seungmin Rho[2*]

[1] Department of Computer Science, COMSATS University Islamabad, Attock Campus, Pakistan
[2] Department of Industrial Security, Chung-Ang University, South Korea
mehryahya@ciit-attock.edu.pk, sadaf_yasmin@cuiatk.edu.pk, smrho@cau.ac.kr

Abstract

Internet of Medical Things (IOMT) based systems provide a framework for remote health monitoring. Liver tumors tend to have parallel intensities with neighboring lesions and may have an abnormal apparent form that directly depends on the stage, state, type, and luminosity setup. In this research, a segmentation model based on improved UNet has been deployed to segment the tumors by incorporating the side-by-side convolution layers based on Filter Response normalization layers (FRN) along with Threshold Linear Units (TLU). This combination of FRN along with TLU has a very strong impact on the performance of the model as the FRN layer operates on each batch sample and each response filter during training, and thus it eliminates the problem of batch dependence. Furthermore, we have also switched from the traditional up-sampling layers to fractionally strided convolutions in UNet which performs up-sampling of the required image with proper learning. Moreover, the tumors are directly segmented by the proposed framework from the given CT scan without any extraction of ROIs. To evaluate the performance of our proposed method, we use a publicly available 3DIRCADb dataset. The proposed technique has shown excellent results with 93.0% accuracy and 71.2% Jaccard score.

Keywords: IOMT, Liver tumor detection, Medical decision making, Semantic segmentation, U-Net

1 Introduction

Multimedia-based medical diagnostic systems provide efficient end-to-end solutions to assist physicians and brings revolution in smart industries [1-3]. Liver cancer is one of the most common cancers causing all over the world and falls in the sixth position according to statistics [4]. In the statistical figures given by WHO in 2017, deaths caused by liver cancer in Ethiopia were 0.16% out of the total number [5]. Hepatocellular carcinoma falls under one of the two types of liver cancer which are primary and secondary. HCC belongs to primary cancer and is accountable for 80% of deaths, causing the deaths of almost 700,000 people yearly [6]. The root cause of primary liver cancers is cirrhosis which comes into being by consuming alcohol, hepatitis B and C viruses, and liver disorder due to obesity [7]. Its diagnosis can be performed by making use of various imaging tests including ultrasound, MRI, and CT.

Among all these tests, CT or computed tomography is most widely used as it provides detailed cross-sectional images of the abdomen. By making use of abdominal CT, further processing is done to perform liver tumor segmentation [8]. However, there are some intensity similarities found between the tumor region and other neighboring lesions which makes tumor detection a difficult task to perform [9]. To address this issue, images are preferred to be enhanced for improved and accurate detection of a cancerous lesion. In a CT scan, a cancerous entity is detected by identifying the difference in pixel intensity of that region in the liver [10]. This difference is termed as hypodense if it is darker as compared to the surrounding healthy liver or termed as hyper-dense if it is brighter than the surrounding liver region. The traditional approach of manual segmentation is a time-taking task for a clinical setup [11]. The human liver consists of 150 slices in a given volume of CT and along with this reason, there is the irregular shape of the lesion, low-intensity contrast between tumor and neighboring tissue, the variation of liver shape as well as size among patients [12]. Keeping in consideration all these reasons, researchers have been working on CAD systems to segment the liver from the CT scan image as well as tumor. In current times, all the conventional approaches used for tumor extraction are not as effective as they should be.

Among all these tests, CT or computed tomography is most widely used as it provides detailed cross-sectional images of the abdomen. By making use of abdominal CT, further processing is done to perform liver tumor segmentation [8]. However, there are some intensity similarities found between the tumor region and other neighboring lesions which makes tumor detection a difficult task to perform [9]. To address this issue, images are preferred to be enhanced for improved and accurate detection of a cancerous lesion. In a CT scan, a cancerous entity is detected by identifying the difference in pixel intensity of that region in the liver [10]. This difference is termed as hypodense if it is darker as compared to the surrounding healthy liver or termed as hyper-dense if it is brighter than the surrounding liver region. The traditional approach of manual segmentation is a time-taking task for a clinical setup [11]. The human liver consists of 150 slices in a given volume of CT and along with this reason, there is the irregular shape of the lesion, low-intensity contrast between tumor and neighboring tissue, the variation of liver shape as well as size among patients [12]. Keeping in

*Corresponding Author: Seungmin Rho; E-mail: smrho@cau.ac.kr
DOI: 10.53106/160792642022032302015

consideration all these reasons, researchers have been working on CAD systems to segment the liver from the CT scan image as well as tumor. In current times, all the conventional approaches used for tumor extraction are not as effective as they should be.

Many of these approaches are manual or semi-automatic which implies their dependency on detectors such as edge detectors instead of considering pixels as images. [13]. Some examples of these methods include graphical models [14], atlas-based models [15], and deformable models [16]. Although these methods provide a good segmentation quality, but the parametric steps involved in these algorithms limit their use in real-time scenarios. A diverse range of machine learning techniques as well as deep learning techniques has been designed now for performing semi-automated or automated segmentation of tumors. In traditional machine learning techniques, the result of automatic segmentation is highly dependent upon careful engineering of features. The high value of sensitivity towards these constructed features makes the learning-based models unstable for all clinical scenarios. However, the most widely used method nowadays is artificial neural networks which further includes Recurrent Neural Networks (RNNs) [17] and Convolutional Neural Networks (CNN) [18]. They are used in several domains such as medical imaging [19-21], as well as in image segmentation [22-23]. Many researchers have made use of CNN as well as its extensions for liver tumor segmentation [24], brain tumor segmentation [25], and skin lesion segmentation [26].

More specifically, semantic segmentation conventionally assigns a category label for all the pixels found in an image and yields encouraging results as far as accuracy is concerned. There are various challenging aspects in medical imaging, including the shape of the lesion or a limited number of cases that are dealt with using semantic CNNs [27]. Moving on to the architecture of semantic CNNs it can be observed that such methods acquire semantic representations from input images [28]. These methods traditionally obtain rich contextual data by enhancing convolution layers. Kernel of convolution acquires specific contextual data via specific receptive locations [29]. A minute or outsized receptive field keeps up a correspondence to small or large scale features, in that order. On the other hand, convolution features pay attention to the region of interest for a particular receptive field and ignore contextual data. Research has shown that there are extremely fine boundaries among the lesions in medical images which ought to be having the same features [30]. This provides grounds for many issues and confusion in performing accurate segmentation. For this reason, it is mandatory to gain maximum contextual data dealing with the region of interest. To address this, various researchers have used patch-based techniques [31]. The key step to these techniques is to convert the medical image into little patches and to perform segmentation after that but these techniques are very costly and time-consuming Further, liver tumor detection first requires accurate identification of the liver, and then the tumor is identified. However, despite advancements in CAD systems, direct liver tumor detection remains an open research area and more automated end-to-end solutions are still required to be designed. Moreover, the fine and accurate segmentation of liver tumors assists in planning procedures of liver therapy and also resulted in classification response of more reliable diagnosis of liver tumors.

In this research work, we have used an improved U-Net-based semantic segmentation method for liver tumor segmentation. Previously, this has been done by many of the researchers. However, their approaches involved the process of segmenting out the liver first from the CT scan and then tumor segmentation is performed. Tumor segmentation is rather a challenging job as compared to liver segmentation. Here, the proposed technique performs tumor segmentation directly. We have incorporated side-by-side convolutions based on FRN layers and TLU units to the exiting UNet architecture to improve the performance. Due to pre-activations and weights of filter, these FRN layers eliminate the scaling problem and the same relative importance is ensured to all weight matrices. For performance evaluation, a 3dircadb dataset is used and compare our approach with the existing techniques which have been used for tumor segmentation in the recent past. This comparison shows that our proposed model exhibits encouraging results. The proposed study has the following contributions:

- We propose an end to end fully automated method for liver tumor detection
- The proposed model can detect liver tumors directly from the CT scans without extracting liver ROIs
- The proposed method integrates side-by-side convolutions along with the FRN layers and TLU units to improve U-Net performance

The rest of the paper is organized as: Section II gives an account of elaborated literature in our specified domain. Section III of this research article discusses our proposed methodology and its details. Section IV presents the results and discussion. Section V of this research paper deals with the conclusion.

2 Related Work

There exist numerous research studies for both liver and liver tumor segmentation. In the context of traditional approaches of segmentation, Li et al. [32] proposed the new model based on a unified level set algorithm. This method integrates the prior information, region competition, and gradient of an image to segment the tumors from a given CT scan. Stawiaski et al. [33] also uses the traditional method of Markov random field and minimal surfaces. In this work, the watershed transform graph of region adjacency is applied with these models to mark tumors in CT scans. Zhang et al. [34] performs tumor segmentation from CT scan by using an interactive method. In this work, the tumor segmentation starts from the preprocessing steps which include contrast enhancement of liver followed by segmentation. The seed points are selected by a user and give as an input support vector machine (SVM) classifier for training and tumors are extracted in the next step. The segmentation results generated by SVM are further refined by applying the morphological operations. All these methods perform accurate segmentation but require a lot of preprocessing steps and extra operations to refine the results.

The liver and tumor segmentation using deep learning-based techniques are also well researched by different researchers. Christ et al. [35] presented an approach in which they segmented liver and lesions with cascaded deep neural networks and 3D conditional random fields. They have made use of two cascaded U-Net models for both liver and tumor in their research work. After that, outputs obtained from these

models were passed as inputs to 3D conditional random fields (CRF) for improvement. In their work, the whole process is divided into three stages i-e in the first stage liver is segmented and passed as a region of interest to the second model for tumor segmentation. Later on, the 3D CRF is used to further improve the performance. The dice similarity coefficient achieved is 0.943 in liver segmentation. However, they excluded their tumor segmentation results in their research study.

Subsequently, Sun et al. [36] proposed liver tumor segmentation using a multichannel fully convolutional network (FCN) from contrast-enhanced multiphase CT scan images. The Single channel of FCN consists of 8 convolutional layers, 3 subsampling layers, 3 deconvolution layers, and two feature fusion layers. The convolutional layers carry out convolutional operations over the output of the preceding layer with diverse kernel sizes and obtain features from an image by sustaining spatial correlations. The sub-sampling layer minimizes the size of an image without impacting its resolution. After that, the deconvolution layer and fusion layers have been exploited for up sampling and fusion process correspondingly. They have used three FCN layers for feature extraction out of three diverse phases of CT scan images. All FCN channels went through an independent training phase and their features were also combined to achieve improved segmentation. They used various performance measures to evaluate the results. They had got a volumetric overlap error (VOE) of 8.1±4.5. Afterward, Chlebus et al. [5] also used the U-Net model for the segmentation of liver tumors. They revised the already implemented U-Net model by further adding up the dropout layer. Furthermore, they also added short skip connections for parameter updates as well as for the escalating speed of the model in the phase of training. After that, the output from this model had been gone through the post-processing phase using 3D connected component, a shape-based processing method, and classified again using a random forest classifier for improved results. They got an average DSC of 0.58 in their segmentation result. However, the shape-based process did not seem to be much efficient as tumors are found in various shapes and sizes.

Furthermore, the most recent works in the liver tumor include the work of Lei et al. [37]. They proposed a segmentation approach using GIU-Net which incorporates improved U-Net along with a graph cut algorithm. In their proposed model, they enhanced the depth of structure for better semantic segmentation results and make the skip connections be from the pooling layers output, unlike the original UNet which concatenates the up-sampling layer output with the output of the corresponding convolutional layer, to reduce the lost information and called this new UNet structure IU-Net. In the next step, they merged this structure with the graph cut technique and as a result, provided their latest segmentation approach. At first, they segmented out the liver from a liver CT sequence with the help of their improved U-Net model. In the next step, they improved their result by making use of graph cuts. For the evaluation of results generated by their proposed algorithm, they used various evaluation measures. By calculating the DSC, the value was 0.9505. Similarly, Li et al. [38] and similar work proposed by others. They exploited a convolutional neural network for the detection of HCC similar to the work of Christ et al. In this research, it can be seen that they performed two major activities which include segmentation of liver and its pathology, HCC using a fully convolutional neural network, and classification of HCC into massive classes [39]. During the segmentation phase, they were completely relying on the FCN-8s structure. This model comprises four max-pooling layers along with two skip structures for concatenation of the last two outputs of the max-pooling layer with the corresponding up-sampling layer. The proposed model had further two skip connections which were there for the concatenation of remaining outputs of the max-pooling layer with the corresponding up-sampling layer. This was important for boosting the number of features being used in predicting the output. Their proposed model comprises two major parts. To begin with, the first part of the model has 13 convolution layers of standard VGG-16 model that performed convolution operation with a kernel size of 3 x 3 and ReLU as activation function and four max-pooling layers. In the second part of their structure up-sampling using deconvolution was performed on the up-sampling layer and the output was fused with the corresponding output from the first half of the model on fusion layers. The accuracy could have been 0.994 via 100 epochs training but due to noise blotches in the output, the algorithm had a dissatisfying performance during the segmentation process.

Moreover, Budak et al. [29] designed two cascaded encoder-decoder-based networks to achieve accurate segmentation of tumors present in the liver along with the liver itself. They put forward an EDCNN algorithm that is made up of two symmetric parts of encoder and decoder. Both parts have ten convolutional layers with batch normalization and ReLU activation followed by a max-pooling layer. In every convolutional layer, they exploited a total of 64 filters, by which they minimized the number of required parameters. For segmentation, they made use of two cascaded deep neural networks, one for liver and one for tumor. The output of the first network is passed as input to the next network. The average DSC value of 0.9522 and 0.634 was achieved on liver and tumor segmentation in that order.

In the presented literature, most of the research works segment the tumor from liver ROIs. Here we slightly changed the approach and segment the tumors directly from CT scan with FRN and TLU units based UNet architecture. The proposed methodology is shown in Figure 1.

3 Material and Methods

An improved UNet based semantic segmentation method is proposed and presented in this section.

3.1 Dataset Extraction

Images used for this thesis are collected from publicly datasets, 3Dircadb01 (3D Image Reconstruction for Comparison of Algorithm Database). The 3DIRCADb dataset provides a wide range of data that has been exploited for this research work. The dataset we have used incorporates huge diversity of medical images including tumor complexity with up to 20 venous phases enhanced CT volumes from different European hospitals with numerous CT scanners. The dataset has assorted 3D CT scans of 10 women and 10 men diagnosed with a tumor in 75% of cases. Every folder of this dataset has 4 sub-folders which are “PATIENT_DICOM”,

"LABELLED_DICOM", "MASKS_DICOM", and "MESHES_VTK". Patient images are found to be in DICOM format and its corresponding label images include images with a region of interest in DICOM format. The dataset of 3D-ircadb01 comprised up of twenty folders. For each patient, the actual image along with tumor annotation is given. The details are given in Table 1.

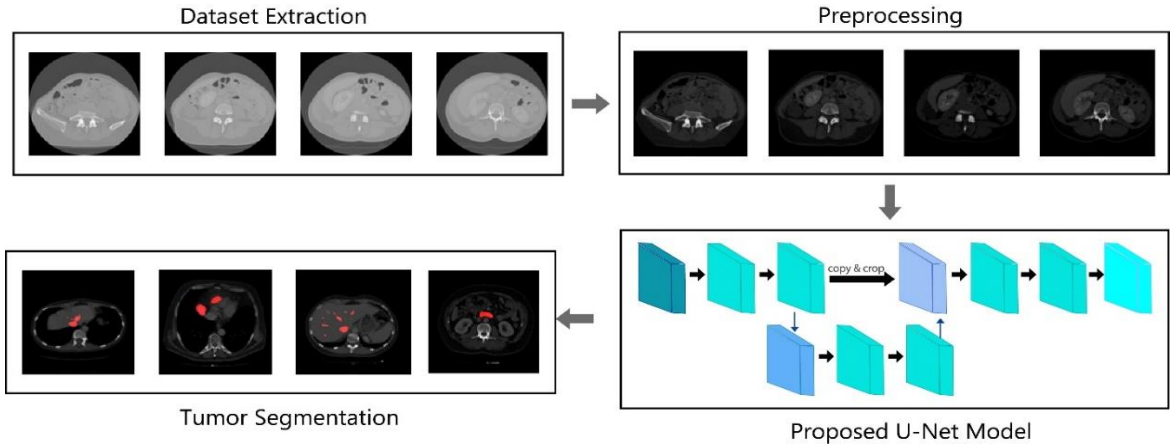

Figure 1. The proposed methodology for fully automated liver tumor segmentation

Table 1. Samples from the dataset

S. No	Gender	YOB	Liver size (cm)	Image Size (Pixels)	Liver Average Density	Voxel Size (mm)	Segmentation Drawbacks
1	F	1944	18,3	512	111	0,57	Stomach,
			15,1	512		0,57	pancreas,
			14.1	129		1,6	duodenum
2	F	1987	20,1	512	84	0,78	Pancreas,
			16,9	512		0,78	duodenum
			15,7	172		1,6	
3	M	1956	16,7	512	108	0,62	Artifact due to
			14,9	512		0,62	metal
			15,2	200		1,25	
4	M	1942	16,9	512	107	0,74	Heart
			12,0	512		0,74	
			17,2	91		2	
5	M	1957	19,8	512	69	0,78	Diaphragm,
			16,8	512		0,78	duodenum
			19,1	139		1,6	

3.2 Preprocessing

Medical imaging datasets tend to have unnecessary objects which do not fall in the region of interest in most cases. Usually, the data to be used for further experimentation needs preprocessing so that images become clearer and are processed easily. For this reason, the step of data preprocessing is mandatory, and it is performed over raw data to enhance the information that is required. Various techniques are normally exploited for this purpose. In this research, we have performed image enhancement to achieve good quality images that are further used for processing. Mainly, we tried to improve the contrast by applying a contrast stretching approach which has enhanced the raw image to a certain extent. Figure 2 shows sample raw images extracted from the dataset.

3.3 Segmentation

The detailed analysis of literature brings us to our proposed technique, and we have used U-Net architecture to achieve accurate segmentation. Segmentation refers to segregating the image into various parts. In this research, we have directly segmented out the liver tumor and for this purpose, we exploited U-Net architecture along with some modifications.

3.4 U-Net Architecture

U-Net architecture is a widely used framework that has emerged in the area of deep learning over the years. The U-Net framework was first developed by Ronneberger et al. [40] in the year 2015. Since then, many researchers have exploited this architecture in the field of deep learning. We continue to use a similar model which comprises the contracting path, expansive path, and bottleneck. However, we do introduce some alterations and modify the traditional U-Net architecture accordingly. Figure 3 shows the standard U-Net model. Our proposed U-Net model along with modifications is explained below in detail.

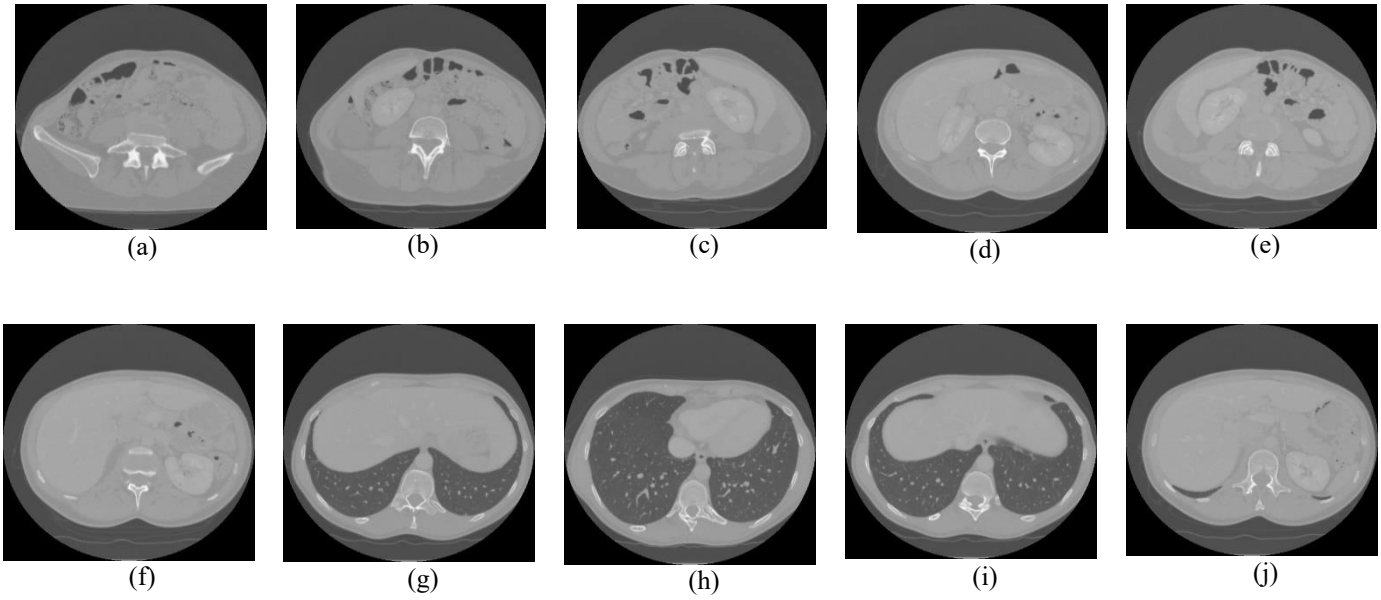

Figure 2. (a, b, c, d, e, f, g, h, i, j) Raw images from 3DIRCADb dataset

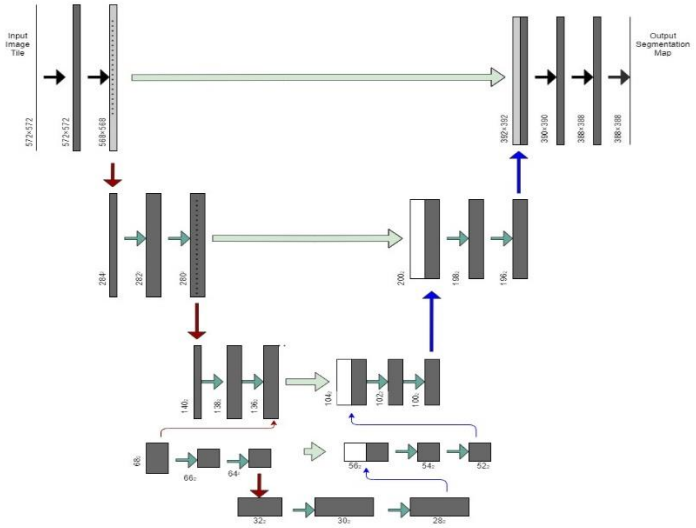

Figure 3. Diagrammatic representation of U-Net model

3.4.1 Encoder

The U-Net model is comprised mainly and a total of three parts. The first one is called an encoder and the second one is called a bottleneck and the third one is called a decoder. The encoder has 4 major blocks, every block comprising of three convolution layers operating on input with 1*1, 3*3, and 5*5 filters. Along with this, FRN layers are incorporated along with TLU units [41]. An FRN is the filter response normalization layer and exhibits the joint combination of both normalization and non-linear activation. This layer operates on each batch sample and each response filter during training, and thus it eliminates the problem of batch dependence. This layer mainly consists of two components. First is that for each response of the filter all the values are normalized independently and then dividing the resulting values by the square root of their un-centered second moment. With the help of this operation, there is no need to perform the mean operation. Later on, the second component Thresholded Linear Unit (TLU). This unit is parametrized by rectification threshold that is to be learned and hence it is as pointwise activations. The combination of FRN along with TLU has a very strong impact on the performance of the model. A schematic structure of FRN layers-based convolution block is shown in Figure 4. For the mathematical description, consider the two-dimensional feature vectors which after passing from the convolution layer produce the 4D-tensor of shape $[B, W, H, C]$. Here the B represents the size of the mini-batch and W and H are the dimension of the input and C denotes the total filters used in the layer of convolution. Let $x = X_{b,,,,,,c} \in \mathbb{R}^N$ where N is the tensor result from the convolution layer and considered as responses of the filter. So for every sample batch point b^{th} and c^{th} filter, and let the mean squared norm value of x is $v^2 = \sum_i x_i^2 / N$. Then the FRN layer is described as in equation (1):

$$\hat{x} = \frac{x}{\sqrt{v^2 + \varepsilon}} \quad \textbf{(1)}$$

In the above equation, the invalid operation of division by zero is handled by a small constant ε. Moreover, the TLU in the FRN is computed by equation (2):

$$z = \max(y, T) \quad \textbf{(2)}$$

After this, we have used the down-sampling layer known as a max-pool layer in which the total size of the window is 2*2. Moreover, the parameter of stride is having a value of 2. Max-pooling layer is present after every convolutional block over the down-sampling path and usually, a dropout layer of 0.05 is introduced after each max-pooling layer. To solve and handle the problems of deep neural networks such as overfitting in the network, there is a dropping of the information carried by the different sets of neurons in the model. The first layer which is termed as convolution layer in the model consists of 16 kernels matrices to be applied to the image. Similarly, in the upcoming layers, the number of kernels is ultimately increased such as 32, 64, and 128. This path of U-Net is designed to extract a different set of features from the given image. From this process, the model is such a capacity that it identifies that what the image represents. This process is completed by incorporating a different number of layers which includes both the pooling and convolution but with loss of spatial information.

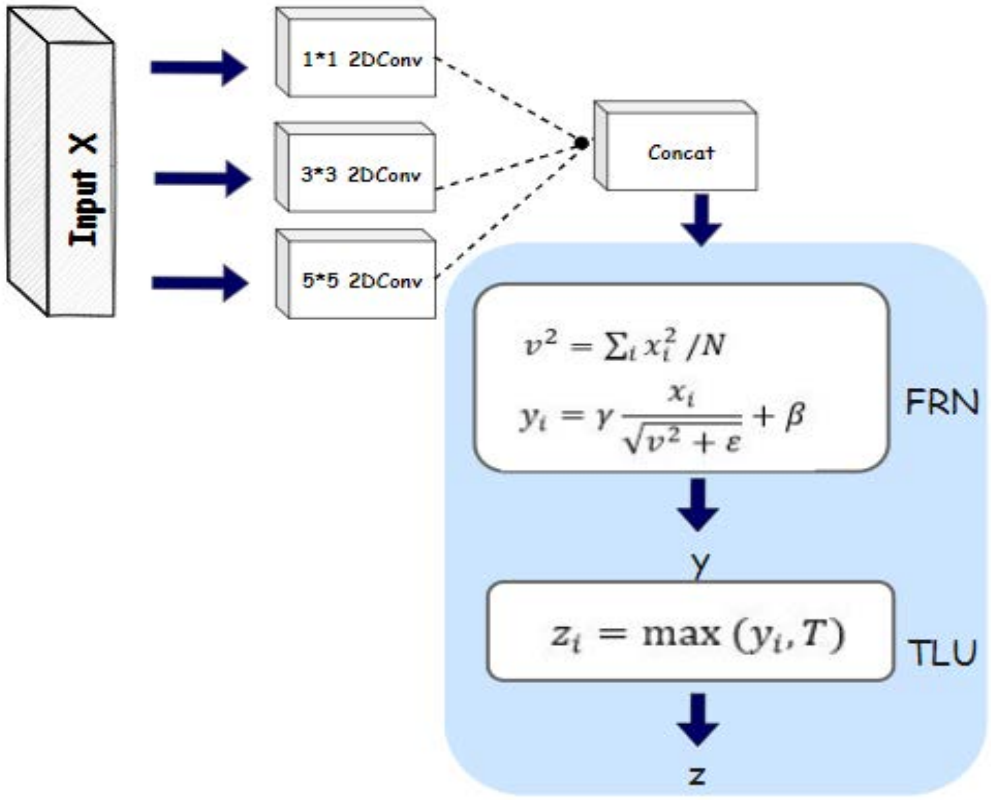

Figure 4. Diagrammatic representation of proposed FRN layers based convo block

3.4.2 Bottleneck path

Among the encoder and decoder path follows the bottleneck part. The bottleneck path has three side-by-side layers of the convolution on input along with weight kernel sizes of 1*1,3*3 and 5*5 followed by FRN and TLU units respectively.

3.4.3 Decoder Path

Structurally comprising of 4 blocks, the decoder path is also called the expanding and up-sampling path. All of the 4 blocks have a deconvolution layer with filter size 3*3 and stride 2. For up-sampling, the deconvolution layer here used is named as transposed convolution layer. The Conv2DTranspose is complicated as compared to traditional up-sampling. This layer performed an invert the process of traditional convolution over the given input. When we input the given image then this layers up samples the provided input with the help of proper learning of weights. On the other, when the same input is given to the traditional and simple up-sampling layer only doubles that is duplicating columns and rows the dimensions of a given input image. The other term that is used for transposed convolutions in the existing literature is fractionally strided convolutions. First, suppose if we have a kernel w that has some suitable weights which is going to perform convolution over the image with a parameter of stride is 1. Moreover, consider that the process does not involve the padding so the inputs that we provided and outputs that are produced are n from unrolled vectors. In the given scenario, the resulting matrix of convolution is denoted by the sparse matrix C in which the kernel elements in both x and y directions are represented by wi, j. This sparse matrix makes the computation of backward pass faster and easier.

Furthermore, the backpropagation of weights is done only when we perform the multiplication of transpose matrix and loss. More specifically, the calculation of both passes in a network such as forward and backward is done only when the multiplication is performed between transpose matrix and sparse matrix. So, in the case of transposed convolutions, these passed are computed as if we perform the process of multiplication on the given C and $(C^T)^T$ respectively. After deploying these transposed convolutions, we add another operation called concatenation. The responsibility of this operation is to combine the outputs of the encoder and decoder. After this, a side-by-side convolution layer is also employed in which kernel size is defined as 1*1,3*3, and 5*5. Then the output is passed as an input to the FRN layer and TLU Units. The objective of this path is to restore the information that we lose during down-sampling. The context information of tumors is integrated with the localization information of tumors. This process is completed by incorporating the skip connections utilizing the concatenation operation that we defined earlier. This involves combining the outputs produced and generated by the encoder with the decoder-produced outputs.

3.4.4 Training Details and Hyper-parameters

The proposed model designed with the U-Net framework is trained on the dataset of 3DIRCADb for tumor segmentation with a split ratio of 80-20. The data of images and their ground mask of highlighted tumors is given as input to further initiate the process of training. The model weights are updated with the help of an optimizer called "Adam" optimizer. The learning rate specified for the trained model is 0.01. The RMSprop is combined with the Adam optimizer. It also involves the new term called momentum. The stochastic gradient is also utilized with it. To update the weights associated with each neuron of the model is given by in equation (3):

$$W_t = W_{t-1} - \eta \frac{\widehat{m}t}{\sqrt{\widehat{v}t}+\epsilon} \quad \textbf{(3)}$$

In equation (7), W shows weights associated with each neuron, and step size is indicated by η. The value of given step size η contributes larger towards the iteration. The computation of values for the terms of $\widehat{m}_t$ and $\hat{v}_t$ are done by equation (4):

$$\widehat{m}t = \frac{m_t}{1-\beta_1^t} \; and \; \hat{v}t = \frac{v_t}{1-\beta_2^t} \quad \textbf{(4)}$$

In equation (8), β_1 and β_2 have employed the default values that are 0.9 and 0.999 respectively. They are usually considered hyper-parameters. Moreover, during training of the model, we have used the loss function of binary cross-entropy which effectively computes the loss over predicted values generated by the model, and actual masks annotated by radiologists. It is mentioned below:

$$BCE = \frac{-1}{N}\sum_{i=1}^{N} y_i * \log\big(P(y_i)\big) + (1-y_i) * \log\big(1-p(y_i)\big) \quad \textbf{(5)}$$

In equation (9), the term BCE means binary cross-entropy y_i shows the class generated or predicted by the model while $P\ (y_i)$ denotes probability. This probability means that how much it is certain that a given is a pixel of the image belongs to either class of liver tumor or other backgrounds. A total of 150 epochs is set to train the network in which the value of batch size is set to 16 and an input image dimension is 256*256.

4 Results and Discussion

4.1 Performance Metrics

4.1.1 Dice Similarity Coefficient (DSC)

Dice similarity coefficient (DSC) is defined as the total size of overlap between two masks and then the resulting term is divided by the size of objects present in the two binary masks. Its value is calculated by the following mentioned equation:

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP}+\text{FP}+\text{FN}} \quad \textbf{(6)}$$

4.1.2 Jaccard Similarity Coefficient (JSC)

Jaccard Similarity Coefficient (JSC) is defined as intersection ratio among two which includes both binary masks of actual and predicted with the union term. It is defined by equation (7):

$$\text{JSC} = \frac{\text{TP}}{\text{TP}+\text{FP}+\text{FN}} \quad \textbf{(7)}$$

4.1.3 Accuracy

Accuracy [42-47] shows the total number of pixels that are present in the given input image and hence correctly classified and can be calculated using the following equation:

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \quad \textbf{(8)}$$

4.1.4 Symmetric Volume Difference (SVD)

SVD gives the difference found in segmented images with ground truth. When SVD is approximated as zero, it denotes encouraging segmentation results, and it is defined in equation (9):

$$\text{SVD} = 1 - \text{DSC} \quad \textbf{(9)}$$

In the above equations (6-9), TP represents true positive, TN represents a true negative, FN shows false-negative, and FP denotes false positive.

4.2 Results

A deep learning-based segmentation algorithm is used for tumor segmentation directly from abdominal CT scan images in this research work. The proposed technique is based on U-Net architecture and segmentation of the tumor is performed with significant results. The dataset which has been used for this research work is 3DIRCADb which provides a diverse range of tumor images. Our proposed model has been trained over this dataset and the images were preprocessed to achieve clarity. We have performed an enhancement operation and the contrast of images has been improved as shown in Figure 5.

Once the contrast of images gets improved, it becomes easy to process that data further for segmentation purposes. After applying the enhancement operation on enhanced images of our dataset, we have further passed the input to our proposed model which is based on the architecture of U-Net. The result is analyzed as that the tumor is segmented out directly from the CT scan image achieving 93% accuracy. It shows that our proposed technique outperformed other existing techniques that have been used for this purpose. The results happened to be highly encouraging when other evaluation measures have been observed. It is worth mentioning that our proposed technique achieved an average dice score of 78.96%, Jaccard similarity score of 71.2%, and SVD with a value of 0.21. These scores are reported by taking an average of scores of the model by running it three different times as presented in Table 2. The resultant output images are given below in Figure 8. Column (a) shows the CT scan images, column (b) shows the original tumor annotations column(c) shows the overlay images with actual annotations of tumors, column (d) shows tumor segmented image by the model, and column (e) shows red-colored part represents the segmented tumor region in output overlay images. Moreover, there are some tumors which model finds difficult to correctly segment as shown in Figure 8 row 3. It is due to the heterogeneous nature of tumors in terms of shape and sizes, and it is very difficult to overcome all of these variations. The proposed model still lacks and is limited to overcome these shape variations.

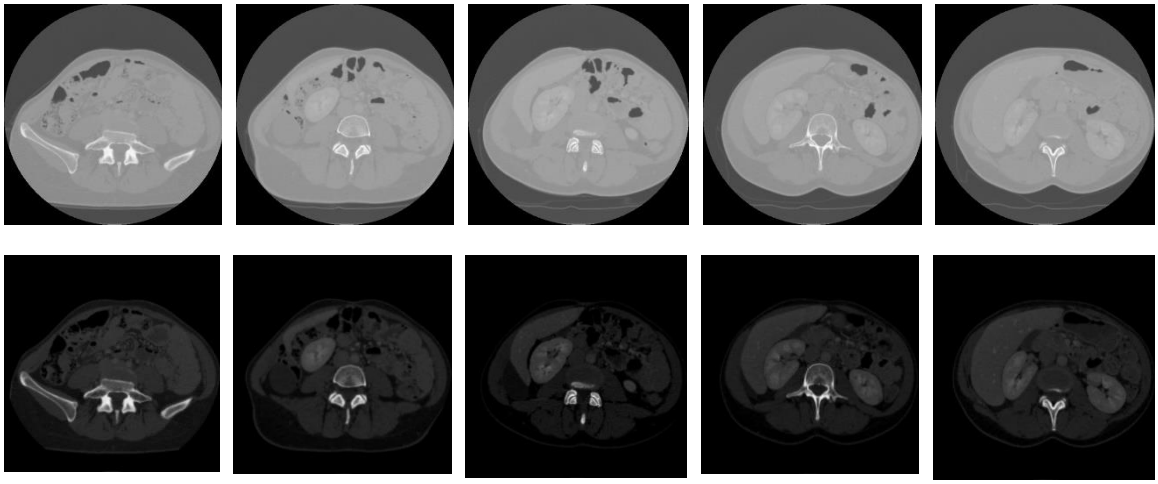

Figure 5. (Left to Right, Top to Bottom) Raw and corresponding enhanced images

Table 2. Comparative results of proposed model and Unet

Architecture	Dice Score	Jaccard	Accuracy	SVD
U-Net [40]	67.5%	56.0%	92%	0.33
1st Run	79.3%	71.4%	93.13%	0.21
2nd Run	79.0%	71.2%	93%	0.21
3rd Run	78.6%	71%	93.1%	0.22
Proposed Method (Average)	**78.96%**	71.2%	93.0%	0.21

To achieve validation, we also perform a comparative analysis of results with the original U-Net architecture. With U-Net, the value of dice score achieved is 67.5%. Moreover, the other scores which include the Jaccard, SVD, and accuracy are 56.0%, 0.33, and 92% respectively. It is perceived that side-by-side convolution based on FRN layers and TLU units brings a rise in performance. The up-sampling transpose convolution layers also play a significant role in architecture. Moreover, in standard U-Net there are only two convolution layers in convo block but here we applied side-by-side convolutions with different matrix sizes of kernels on the input images, and later on, the results of all convolution layers are combined and pass them to FRN layers and TLU Units. Furthermore, training graphs of accuracy and loss are also plotted for both standard U-Net and the proposed model. The graph plotted for accuracy shows how close the obtained result is to a particular target. For instance, Epoch is plotted on the x-axis, and Accuracy is plotted on the y-axis as shown in Figure 6. In the same way, a loss can also be determined as shown in Figure 7 below shows plotted values of model loss. According to Figure 7, Epoch is plotted on the x-axis, and Loss is plotted on the y-axis.

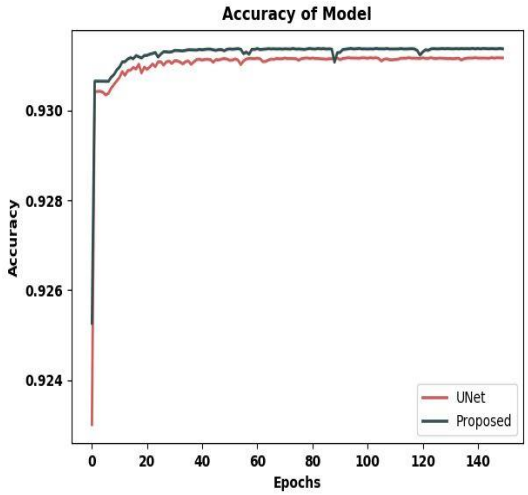

Figure 6. Graph plotted to determine model accuracy

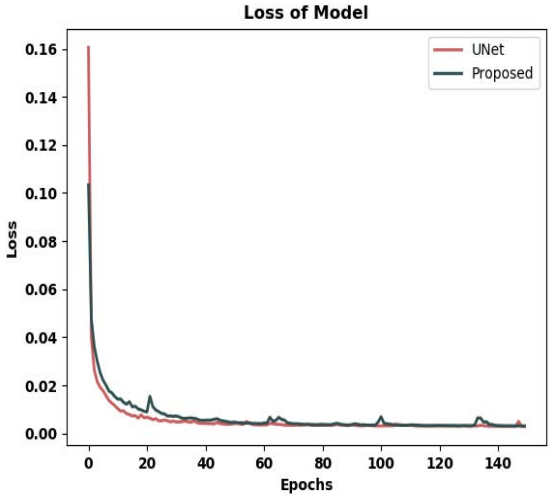

Figure 7. Graph plotted to determine model loss

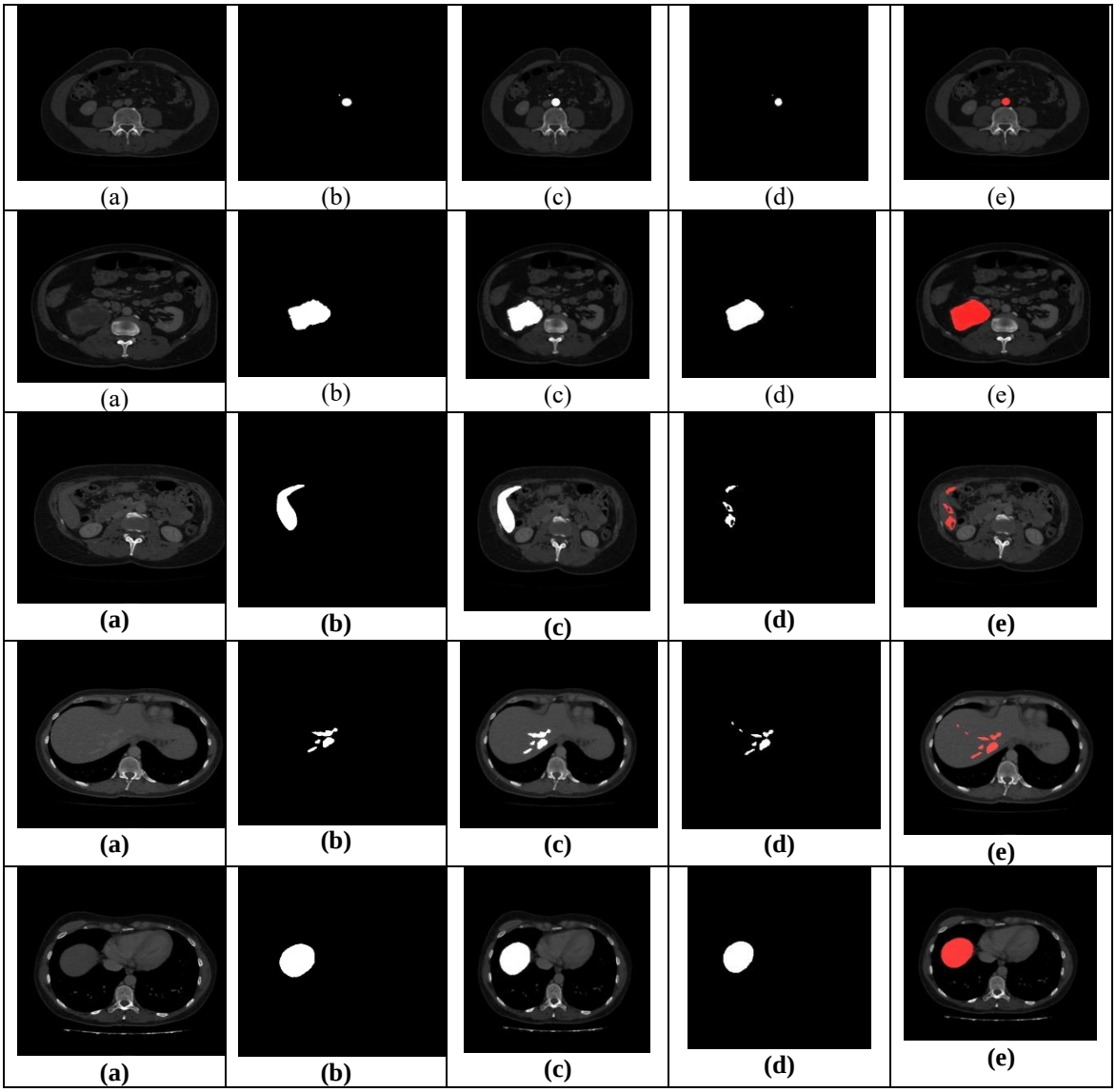

Figure 8. (a) Original enhanced image (b) Actual mask (c) Actual overlay image (d) Predicted mask (e) Predicted overlay

In this section, we compare the results of our proposed model along with the results acquired from some previous techniques and validate our results. Christ et al. [48] segmented the tumor from the liver in their research work. They performed their experimentations over images obtained from the 3DIRCADb dataset by using cascaded FCNs and achieved a dice score of 56%. Later on, Budak et al. [29] improved the previous dice score percentage and obtained

63% with encoder-decoder-based CNNs (EDCNN). They have also segmented out the liver tumor from the CT scan and used the 3DIRCADb dataset for the experimental setup. Although their results were significantly improved there was more room left to achieve finer segmentation results. Therefore, Ayalew et al. [49] have also performed segmentation of tumors. In their work, they combined the images of two datasets namely LITs and 3DIRCADd and achieved the 63% dice score. If a comparison is to be made among all the above-mentioned techniques, our model shows some improvements. Table 3 provides a comparative analysis of the proposed technique along with other approaches. To add more, it is also evident from the comparative analysis that the most significant dice score has been presented by a technique we have put forward in this research work. This validates our technique being used in the current study and proves it to be efficient enough for tumor segmentation. It subsequently assists in achieving accurate liver tumor segmentation results.

Table 3. Comparative analysis with existing work

Authors	Technique used	Dice Score	Year
Christ et al. [48]	Tumor segmentation from liver	56%	2017
Budak et al. [29]	Tumor segmentation from liver	63.4%	2019
Ayalew et al. [49]	Tumor directly segmented from CT scan	63%	2020
Proposed Method	**umor directly segmented U-Net+ FRN+ TLU Architecture**	**78.96%**	2021

5. Conclusion

This research work addressed the issue of liver tumor detection on a different scale for IOMT based health monitoring systems. Our proposed technique focuses on tumor segmentation directly from CT scans instead of detection of tumors from liver ROIs which is extremely challenging in terms of medical imaging. As a solution, we have proposed a model that exploits improved U-Net architecture in which side-by-side convolutions are operated on input images along with FRN layers and TLU units to perform segmentation. The proposed work showed encouraging results and also enhanced the performance of liver tumor segmentation. Furthermore, to evaluate our model, we have also performed a comparative analysis among our proposed techniques with the existing state-of-the-art techniques, and the proposed model achieves good results but there is still some further research are required in terms of performance. In the future, we will investigate the performance of a model on more strenuous tumors available in other challenge datasets as well as validating the model with advanced data augmentation methods such as Generative Adversarial Networks. Moreover, a feedback unit is also integrated with the framework of IOMT in which diagnosis results from real-time patients are monitored and recorded along with feedback of these results which is further used to optimize the model.

Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC Information Technology Research Center) support program (IITP-2021-2018-0-01799) supervised by the IITP Institute for Information & communications Technology Planning & Evaluation) and also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2019R1F1A1060668).

References

[1] J. C.-W. Lin, G. Srivastava, Y. Zhang, Y. Djenouri, M. Aloqaily, Privacy-preserving multiobjective sanitization model in 6G IoT environments, *IEEE Internet of things journal*, Vol. 8, No. 7, pp. 5340-5349, April, 2021.

[2] J. C.-W. Lin, Y. Djenouri, G. Srivastava, Efficient closed high-utility pattern fusion model in large-scale databases, *Information Fusion*, Vol. 76, pp. 122-132, December, 2021.

[3] L. Jiang, J. Shen, S. Ji, Y. Dong, T. Miao, Hash Forest Structure Assisted Bi-auditing Protocol with Multiuser Modification in E-health Systems, *Journal of Internet Technology*, Vol. 22, No. 4, pp. 923-934, July, 2021.

[4] A. M. Anter, A. E. Hassenian, CT liver tumor segmentation hybrid approach using neutrosophic sets, fast fuzzy c-means and adaptive watershed algorithm, *Artificial Intelligence in Medicine*, Vol. 97, pp. 105-117, June, 2019.

[5] G. Chlebus, A. Schenk, J. H. Moltz, B. van Ginneken, H. K. Hahn, H. Meine, Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing, *Scientific Reports*, Vol. 8, No. 1, Article No. 15497, October, 2018.

[6] P. Campadelli, E. Casiraghi, A. Esposito, Liver segmentation from computed tomography scans: A survey and a new algorithm, *Artificial Intelligence in*

Medicine, Vol. 45, No. 2-3, pp. 185-196, February-March, 2009.

[7] W. Huang, N. Li, Z. Lin, G. Huang, W. Zong, J. Zhou, Y. Duan, Liver tumor detection and segmentation using kernel-based extreme learning machine, in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013, pp. 3662-3665.

[8] W. Huang, Z. M. Tan, Z. Lin, G. Huang, J. Zhou, C. K. Chui, Y. Su, S. Chang, A semi-automatic approach to the segmentation of liver parenchyma from 3D CT images with Extreme Learning Machine, in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 2012, pp. 3752-3755.

[9] T.-N. Le, P. T. Bao, H. T. Huynh, Liver Tumor Segmentation from MR Images Using 3D Fast Marching Algorithm and Single Hidden Layer Feedforward Neural Network, *BioMed Research International*, Vol. 2016, Article No. 3219068, August, 2016.

[10] W. Li, F. Jia, Q. Hu, Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks, *Journal of Computer and Communications*, Vol. 3, No. 11, pp. 146-151, November, 2015.

[11] Y. Masuda, T. Tateyama, W. Xiong, J. Zhou, M. Wakamiya, S. Kanasaki, A. Furukawa, Y. W. Chen, Liver tumor detection in CT images by adaptive contrast enhancement and the EM/MPM algorithm, in *2011 18th IEEE International Conference on Image Processing*, Brussels, Belgium, 2011, pp. 1421-1424.

[12] D. Smeets, D. Loeckx, B. Stijnen, B. De Dobbelaer, D. Vandermeulen, P. Suetens, Semi-automatic level set segmentation of liver tumors combining a spiral-scanning technique with supervised fuzzy pixel classification, *Medical image analysis*, Vol. 14, No. 1, pp. 13-20, February, 2010.

[13] X. Zhang, J. Tian, K. Deng, Y. Wu, X. Li, Automatic Liver Segmentation Using a Statistical Shape Model With Optimal Surface Detection, *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 10, pp. 2622-2626, October, 2010.

[14] Q. Luo, W. Qin, T. Wen, J. Gu, N. Gaio, S. Chen, L. Li, Y. Xie, Segmentation of abdomen MR images using kernel graph cuts with shape priors, *Biomedical engineering online*, Vol. 12, pp. 1-19, December, 2013.

[15] D. Li, L. Liu, J. Chen, H. Li, Y. Yin, B. Ibragimov, L. Xing, Augmenting atlas-based liver segmentation for radiotherapy treatment planning by incorporating image features proximal to the atlas contours, *Physics in Medicine & Biology*, Vol. 62, No. 1, pp. 272-288, January, 2017.

[16] G. Chartrand, T. Cresson, R. Chav, A. Gotra, A. Tang, J. A. De Guise, Liver segmentation on CT and MR using Laplacian mesh optimization, *IEEE Transactions on Biomedical Engineering*, Vol. 64, No. 9, pp. 2110-2121, September, 2017.

[17] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, C.-M. Chen, LSTM Network for Transportation Mode Detection, *Journal of Internet Technology*, Vol. 22. No. 4, pp. 891-902, July, 2021.

[18] J. C.-W. Lin, Y. Shao, Y. Djenouri, U. Yun, ASRNN: a recurrent neural network with an attention model for sequence labeling, *Knowledge-Based Systems*, Vol. 212, Article No. 106548, January, 2021.

[19] A. Belhadi, Y. Djenouri, V. G. Diaz, E. H. Houssein, J. C. W. Lin, Hybrid intelligent framework for automated medical learning, *Expert Systems*, pp. 1-4, May, 2021.

[20] Q. Nie, Y.-B. Zou, J. C.-W. Lin, Feature extraction for medical ct images of sports tear injury, *Mobile Networks and Applications*, Vol. 26, No. 1 pp. 404-414, February, 2021.

[21] J. M.-T. Wu, G. Srivastava, J. C.-W. Lin, Q. Teng, A Multi-Threshold Ant Colony System-based Sanitization Model in Shared Medical Environments, *ACM Transactions on Internet Technology (TOIT)*, Vol. 21, No. 2, pp. 1-26, June, 2021.

[22] M. Bukhari, K. B. Bajwa, S. Gillani, M. Maqsood, M. Y. Durrani, I. Mehmood, H. Ugail, S. Rho, An efficient gait recognition method for known and unknown covariate conditions, *IEEE Access*, Vol. 9, pp. 6465-6477, December, 2020.

[23] M. Maqsood, S. Yasmin, I. Mehmood, M. Bukhari, M. Kim, An Efficient DA-Net Architecture for Lung Nodule Segmentation, *Mathematics*, Vol. 9, No. 13, Article No. 1457, July, 2021.

[24] H. Seo, C. Huang, M. Bassenne, R. Xiao, L. Xing, Modified U-Net (mU-Net) With Incorporation of Object-Dependent High Level Features for Improved Liver and Liver-Tumor Segmentation in CT Images, *IEEE Transactions on Medical Imaging*, Vol. 39, No. 5, pp. 1316-1325, May, 2020.

[25] S. Pereira, A. Pinto, V. Alves, C. A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images, *IEEE transactions on medical imaging*, Vol. 35, No. 5, pp. 1240-1251, May, 2016.

[26] Y. Yuan, M. Chao, Y.-C. Lo, Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance, *IEEE transactions on medical imaging*, Vol. 36, No. 9, pp. 1876-1886, September, 2017.

[27] H. Jiang, T. Shi, Z. Bai, L. Huang, AHCNet: An Application of Attention Mechanism and Hybrid Connection for Liver Tumor Segmentation in CT Volumes, *IEEE Access*, Vol. 7, pp. 24898-24909, February, 2019.

[28] L. Meng, Y. Tian, S. Bu, Liver tumor segmentation based on 3D convolutional neural network with dual scale, *Journal of Applied Clinical Medical Physics*, Vol. 21, No. 1, pp. 144-157, January, 2020.

[29] Ü. Budak, Y. Guo, E. Tanyildizi, A. Şengür, Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation, *Medical Hypotheses*, Vol. 134, Article No. 109431, January, 2020.

[30] A. Qayyum, A. Lalande, F. Meriaudeau, Automatic Segmentation of Tumors and Affected Organs in the Abdomen Using a 3D Hybrid Model for Computed Tomography Imaging, *Computers in Biology and Medicine*, Vol. 127, Article No. 104097, December, 2020.

[31] H. Khan, P. M. Shah, M. A. Shah, S. u. Islam, J. J. P. C. Rodrigues, Cascading handcrafted features and Convolutional Neural Network for IoT-enabled brain tumor segmentation, *Computer Communications*, Vol. 153, pp. 196-207, March, 2020.

[32] B. N. Li, C. K. Chui, S. Chang, S. H. Ong, A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced CT images, *Expert Systems with Applications*, Vol. 39, No. 10, pp. 9661-9668, August, 2012.

[33] J. Stawiaski, E. Decenciere, F. Bidault, Interactive liver tumor segmentation using graph-cuts and watershed, in *Workshop on 3D segmentation in the clinic: a grand challenge II. Liver tumor segmentation challenge. MICCAI*, New York, USA, 2008, pp. 1-12.

[34] X. Zhang, J. Tian, D. Xiang, X. Li, K. Deng, Interactive liver tumor segmentation from ct scans using support vector classification with watershed, in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, MA, USA, 2011, pp. 6005-6008.

[35] P. F. Christ, M. E. A. Elshaer, F. Ettlinger, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, M. Armbruster, F. Hofmann, M. D'Anastasi, W. H. Sommer, S.-A. Ahmadi, B. H. Menze, Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 2016, pp. 415-423.

[36] D. Sun, J. Lu, Z. Luo, L. Zhang, P. Liu, Z. Chen, Competitive electrochemical platform for ultrasensitive cytosensing of liver cancer cells by using nanotetrahedra structure with rolling circle amplification, *Biosensors and Bioelectronics*, Vol. 120, pp. 8-14, November, 2018.

[37] L. Lei, F. Xi, S. Chen, Z. Liu, Iterated graph cut method for automatic and accurate segmentation of finger-vein images, *Applied Intelligence*, Vol. 51, No. 2, pp. 673-689, February, 2021.

[38] Y. Li, Y.-Q. Zhao, F. Zhang, M. Liao, L.-L. Yu, B.-F. Chen, Y.-J. Wang, Liver segmentation from abdominal CT volumes based on level set and sparse shape composition, *Computer methods and programs in biomedicine*, Vol. 195, Article No. 105533, October, 2020.

[39] P. Bilic, P. F. Christ, E. Vorontsov, G. Chlebus, H. Chen, Q. Dou, C.-W. Fu, X. Han, P.-A. Heng, J. Hesser, *The liver tumor segmentation benchmark (lits)*, arXiv preprint arXiv:1901.04056, January, 2019.

[40] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical image computing and computer-assisted intervention*, Munich, Germany, 2015, pp. 234-241.

[41] S. Singh, S. Krishnan, Filter response normalization layer: Eliminating batch dependence in the training of deep neural networks, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, USA, 2020, pp. 11237-11246.

[42] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, K. Muhammad, A local and global event sentiment based efficient stock exchange forecasting using deep learning, *International Journal of Information Management*, Vol. 50, pp. 432-451, February, 2020.

[43] M. Yasir, M. Y. Durrani, S. Afzal, M. Maqsood, F. Aadil, I. Mehmood, S. Rho, An intelligent event-sentiment-based daily foreign exchange rate forecasting system, *Applied Sciences*, Vol. 9, No. 15, Article No. 2980, August, 2019.

[44] F. Jabeen, M. Maqsood, M. A. Ghazanfar, F. Aadil, S. Khan, M. F. Khan, I. Mehmood, An IoT based efficient hybrid recommender system for cardiovascular disease, *Peer-to-Peer Networking and Applications*, Vol. 12, No. 5, pp. 1263-1276, September, 2019.

[45] S. Afzal, M. Maqsood, F. Nazir, U. Khan, F. Aadil, K. M. Awan, I. Mehmood, O.-Y. Song, A data augmentation-based framework to handle class imbalance problem for Alzheimer's stage detection, *IEEE Access*, Vol. 7, pp. 115528-115539, August, 2019.

[46] A. Kalsoom, M. Maqsood, M. A. Ghazanfar, F. Aadil, S. Rho, A dimensionality reduction-based efficient software fault prediction using Fisher linear discriminant analysis (FLDA), *The Journal of Supercomputing*, Vol. 74, No. 9, pp. 4568-4602, September, 2018.

[47] J. C.-W. Lin, Y. Djenouri, G. Srivastava, U. Yun, P. Fournier-Viger, A predictive GA-based model for closed high-utility itemset mining, *Applied Soft Computing*, Vol. 108, Article No. 107422, September, 2021.

[48] P. F. Christ, F. Ettlinger, F. Grün, M. E. A. Elshaera, J. Lipkova, S. Schlecht, F. Ahmaddy, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, F. Hofmann, M. D. Anastasi, S.-A. Ahmadi, G. Kaissis, J. Holch, W. Sommer, R. Braren, V. Heinemann, B. Menze, *Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks*, arXiv preprint arXiv:1702.05970, February, 2017.

[49] Y. A. Ayalew, K. A. Fante, M. Aliy, Deep Learning Based Liver Cancer Segmentation from Computed Tomography Images, pp. 1-33, August, 2020.

Biographies

Mehr Yahya Durrani is serving as an Assistant Professor at COMSATS University Islamabad, Attock Campus, Pakistan. He graduated in the field of information Technology and later received master's degree in computer science. Currently, he is pursuing Doctorate in the field of Pattern Recognition.

Sadaf Yasmin is currently working as Assistant Professor at COMSATS University Islamabad, Attock Campus, Pakistan. She has completed her PhD from Capital University of Science and Technology, Islamabad. Her research interests include network protocol design, delay tolerant networks, future network architectures, CCN and Internet of Things approaches.

Seungmin Rho is currently a Faculty of the Department of Industrial Security, chung-Ang University, South Korea. His current research interests include database, big data analysis, music retrieval, multimedia systems, machine learning, knowledge management, and computational intelligence. He has published more than 180 papers in refereed journals and conference proceedings in these areas. He has been involved in more than 20 conferences and workshops as various chairs and more than 30 conferences/workshops as a program committee member.