TPCNN: Two-path convolutional neural network for tumor and liver segmentation in CT images using a novel encoding approach

Amirhossein Aghamohammadi [1], Ramin Ranjbarzadeh*[2], Fatemeh Naiemi [3], Marzieh Mogharrebi [4], Shadi Dorosti [5], Malika Bendechache [6]

[1] Institute of IR 4.0 (IIR4.0), Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia.
amho.mohammadi@gmail.com

[2] Department of Telecommunications Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran.
ranjbar.ramin24@gmail.com

[3] Department of Electrical and Computer Engineering, Faculty of Semnan Girls', Semnan Branch, Technical and Vocational University (TVU), Semnan, Iran.
daneshjo_naimi@yahoo.com

[4] Institute of IR 4.0 (IIR4.0), Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia.
mogharrebi.marzieh@gmail.com

[5] Department of Industrial Engineering, Urmia University of Technology (UUT), Urmia, Iran
Shadi.dorosti@gmail.com

[6] School of Computing, Faculty of Engineering and Computing, Dublin City University, Ireland.
malika.bendechache@dcu.ie

* corresponding author

Abstract

Automatic liver and tumour segmentation in CT images are crucial in numerous clinical applications, such as postoperative assessment, surgical planning, and pathological diagnosis of hepatic diseases. However, there are still a considerable number of difficulties to overcome due to the fuzzy boundary, irregular shapes, and complex tissues of the liver. In this paper, for liver and tumor segmentation and to overcome the mentioned challenges a simple but powerful strategy is presented based on a cascade convolutional neural network. At the first, the input image is normalized using the Z-Score algorithm. This normalized image provides more information about the boundary of tumor and liver. Also, the Local Direction of Gradient (LDOG) which is a novel encoding algorithm is proposed to demonstrate some key features inside the image. The proposed encoding image is highly effective in recognizing the border of liver, even in the regions close to the touching organs. Then, a cascade CNN structure for extracting both local and semi-global features is used which utilized the original image and two other obtained images as the input data. Rather than using a complex deep CNN model with a lot of hyperparameters, we employ a simple but effective model to decrease the train and testing time. Our technique outperforms the state-of-the-art works in terms of segmentation accuracy and efficiency.

Keywords: Image segmentation, Deep learning, Lesion detection, Liver segmentation, Convolutional Neural Network

1. Introduction

The liver is a meaty and large organ that plays a significant role in our digestive system and located on the right side of the belly. Cancerous growths in the liver (hepatic tumours) can pose significant risks to human life and almost always occurs after cirrhosis is present. The increasing occurrence of liver cancer draws attention to the treatment efficacy which relies mainly on its primary diagnosis. The liver is surely the primary probable candidate amongst possible cancers for metastases including bronchus, mouth, brain, bladder, lips, lymphoma, lung, prostate, stomach, cavernous nasal sinus, and appendix. (Yang et al. 2008; Ker et al. 2011). An increasing number of cancer-related deaths (13% of total deaths overall worldwide) are caused by hepatic tumours (Sethi, Saini & Singh 2016). Providing the prompt and precise recognition of cancer cells and the

proper localization of the tumours, are the core building block and play a vital role in the planning process for effective patient treatment and survival (Yang et al. 2008; Yamashita & Wang 2013; Frid-Adar et al. 2018).

If the cancer cells are successfully detected in the early phases of tumour growth, they can often be effectively treated by employing medical procedures or surgical resection chiefly of non-cancerous (benign) cells. If the size of the tumor is small and occupies a restricted section of the liver, this part of the organ can be removed by a surgeon only to stop the cancer growing and spreading (Ranjbarzadeh & Baseri Saadi 2020).

Computed tomography (CT) that is broadly utilized in hospitals can accurately perform early screening for cancer cells to recognize hepatic tumours precisely in the presence of some other organs, such as the heart and the stomach. However, due to intensity similarity in CT images, the precise distinction of the affected organic tissue (tumour) is certainly an arduous and difficult task, even for experienced doctors to make correct diagnosis. Also, as a detailed rule, the visual analysis of CT outputs is insufficient for accurate image interpretation (Wu et al. 2015; Xi et al. 2020). In recent years, many computer-aided diagnostic systems are employing that are mainly based on computer vision and artificial intelligence techniques to properly identify the border differences or even subtle shape between two objects. These procedures are highly reproducible and can enhance diagnostic accuracy by helping radiologists to be more accurate using a combination of various classification models with reasonable running time. Due to CT superior characteristics and capability of discrimination of soft tissues, this kind of imaging has found the highest diversity of applications such as image labeling or image segmentation (Ranjbarzadeh & Saadi 2020; Xi et al. 2020).

Finding and labeling objects inside the image is a quite important stage in numerous medical image analysis and is considered as image segmentation. This processing step can principally be employed either as a first or final processing stage. In the field of image processing, segmentation refers to the process of partitioning image into multiple segments or grouping together pixels based on similar attributes in neighbouring proximity. For instance, pixels in a particular part of the image can be considered as an analogous region based on various region criteria such as size, shape, density, colour, and texture characteristics (smooth or rough) (Chen & Pan 2018; Huang et al. 2018; LI, TSO & HE 2020).

In recent years, to overcome the segmentation difficulty of tumor and liver due to low

contrast, irregular shape and the fuzzy boundary between liver tissue and touching organs from CT images many highly sophisticated methods have been developed. These sophisticated frameworks can be classified into one of the three main categories on the basis of their characteristics, including semi-automatic strategies (Yang et al. 2014; Bakas et al. 2017; Kavur et al. 2020), interactive approaches (Baâzaoui et al. 2017; Chartrand et al. 2017; Zhou et al. 2018), and automatic frameworks (Li et al. 2013; Ranjbarzadeh & Baseri Saadi 2020). The problems of the over-segmentation and leakage were overcomed using a semi-automatic approach consisting of three steps in (Xu et al. 2020). Firstly, the binary images were obtained by a series of techniques from the input image. Next, some random seed points are selected on the binary image to obtain the primary shape of the liver tissue. Finally, a novel level set active contour method is used to refine the primary liver border. The semi-automatic and interactive models normally are used by various user directions or human-machine interaction (HMI), leading to increased radiologist's mistake in hospitals and contributing to rising healthcare costs. To address this, recently, fully automatic systems have been used in various applications to develop accuracy and steadily reduce the costs and time of diagnosis.

In the field of the tumour and liver analysis, current algorithms mainly can be split into two broad groups, including anti-learning and learning methods (Lu et al. 2017). The anti-learning methods regularly include (Luo, Li & Li 2014) the active contour (Guo, Schwartz & Zhao 2019; Xu et al. 2020), clustering (Cai 2019; Ranjbarzadeh & Baseri Saadi 2020), region growing (Lu et al. 2014; Zhou et al. 2018; Zeng et al. 2018; Liu et al. 2019), graph cut (Liao et al. 2016; Huang et al. 2018; Liu et al. 2019), and level set (Hoogi et al. 2017; Li et al. 2020) algorithms. Region growing approach selects the touching pixels with a high degree of similarity in intensity or variance value as the same object or area. The efficiency of this technique highly depends on the choice of the seed points. Graph cut is a powerful energy optimization (minimization) technique that characterizes the image to an undirected weighted graph. In this approach, prior knowledge or learning is unrequired and each pixel $p \in I$ in the digital image is displayed as a node inside the graph (p implies pixels and I indicates the image). Also, every edge connects two adjacent nodes (pixels), so that the weight of the edge specifies the criteria of the similarities between each pair (Luo, Li & Li 2014; Lu et al. 2017).

Cai (Cai 2019) for segmentation of the liver in a non-uniform background proposed a kernel space fuzzy clustering method. Firstly, a high-dimensional feature space (high-dimensional Hilbert

kernel space) has been gained using a mapping algorithm in the Euclidean space by using a kernel function. Euclidean distance between the clustering center c_j and the mean value of the sample $\bar{x_i}$ is demonstrated using a fuzzy index m. Also the influence of pixel information in the neighboring region on the current pixel clustering is described by parameter $\propto$, which is an important factor to define the fuzzy index m. Next, by combining the current pixel and all pixels around it and considering space information in the CT image, the linear weighted filtering image has been achieved. Finally, by using the Lagrange multiplier technique in kernel space of the fast fuzzy clustering, the two-dimensional histogram between the segmented pixel and its vicinity mean was implemented. Wang et al. (Wang et al. 2017) employed an approach based on a statistical shape model for solving the problem of similarity matching of image blocks. Their structure used the sparse coefficients and dictionary together to develop the segmentation efficiency of the a priori technique and the difficulty of the initialization of the deformation model. In the first step, to normalize the a priori shape models a generalized Procrustes analysis (GPA) is utilized. In the second step, all mark points are chosen from the corresponding vertexes. Next, using the a priori shape models and their corresponding marks the inquiry dictionary is created. In the fourth step, the sparse codes are calculated using the dictionary for mark points. In the next step, the sparse statistical shape model is created using the sparse codes. Finally, the boundary energy, sparse matching constraints, and intensity energy are employed to deform the statistical shape model. In (Ranjbarzadeh & Saadi 2020) a novel approach for removing the unwanted border of the liver was presented that is able to increase segmentation accuracy using a combination of the convex and concave points. The eight direction filter masks (Kirsch filter) were employed to increase the accuracy of the edge detection. They also applied a mean-shift clustering strategy to improve the local contrast. Moreover, by using these concave and concave points the problem of segmenting touching organs is successfully solved. Finally, by using the fuzzy c-means (FCM) clustering method the promising results were obtained.

A combination of the globally optimized surface evolution strategy and a 3D CNN is applied in (Hu et al. 2016) for segmenting the liver tissue in a CT slice automatically. In the first step, by learning the method of the initial surface, the probability map of the liver tissue was recognized. To increase the accuracy, all global and local information of the shape of the liver tissue was incorporated into a segmentation model by employing prior partitioning. As their network utilized both global and local features, the model can have recognized blurred boundaries effectively.

Although promising performance has been reported, their model is not well-designed and requires optimization of several auxiliary side networks. Due to the fact that the weighting of some networks for the overall loss function in the fuzzy borders cannot be an easy task, their model is not able to define the border between two touching organs effectively. A deep belief network (DBN) employed by (Ahmad et al. 2019) based on an automatic feature learning method for segmenting the liver. In their strategy a DBN for supervised fine tuning and unsupervised pre-training was trained. Although due to accurate feature learning their network is able to recognize the tumor area within the liver superbly, but it is not enough sensitive to the tumor on the boundary of a liver. The drawback of their structure is due to the low contrast, fuzzy boundaries, weak edges, and overlapping areas of touching organs at liver boundaries that caused misclassification. In (Budak et al. 2020) two deep encoder-decoder CNN (EDCNN) were implemented for segmenting of both the liver and tumors. The first EDCNN is responsible for detecting the border of the liver and the second EDCNN can detect the tumor area. Their model reaches acceptable results in terms of all evaluation metrics even at the presence of the fuzzy borders. The main drawbacks of their model are the inappropriate performance of the EDCNN on a quantity-limited dataset and the slow training speed which lead its applications encountered a lot of restrictions. Also, their proposed method was unable to obtain the acceptable performance against hepatic tumors.

In this study, a new pipeline to distinguish the exact border of the liver and tumors from medical CT images is introduced. It is generally considered to be challenging to extract the borders of the liver and lesions meticulously due to sharing similar intensity values across the liver. We first propose a new encoding method to increase the malleability of border detection with shape variation and extract the key local shape details more faithfully, particularly when a few samples of training images are available. Then another encoding algorithm (Z-Score normalization) is employed that improves the distinction capability of touching organs. Finally, a segmentation approach based on a two-path CNN structure is suggested which uses both local and semi-global features to accurate segmentation.

2. Material and Methods

This study structured as follows. In part 2.1, the Z-Score normalization strategy is explained. In part 2.2, a new encoding method (LDOG) is proposed. In part 2.3, the architecture of the convolutional neural network (CNN) is described. In part 2.4, we demonstrate our CNN model. The experiment and concluding remarks are explained in section 3. The workflow of our technique

is depicted in Fig. 1. As is illustrated in Fig. 1, our approach comprises of three steps; 1) Normalizing the original image to highlight key information, 2) Representing the original image in different form (encoding image) to emphasize of some key structures of the image, 3) Applying three different images to a two-path Convolutional Neural Network.

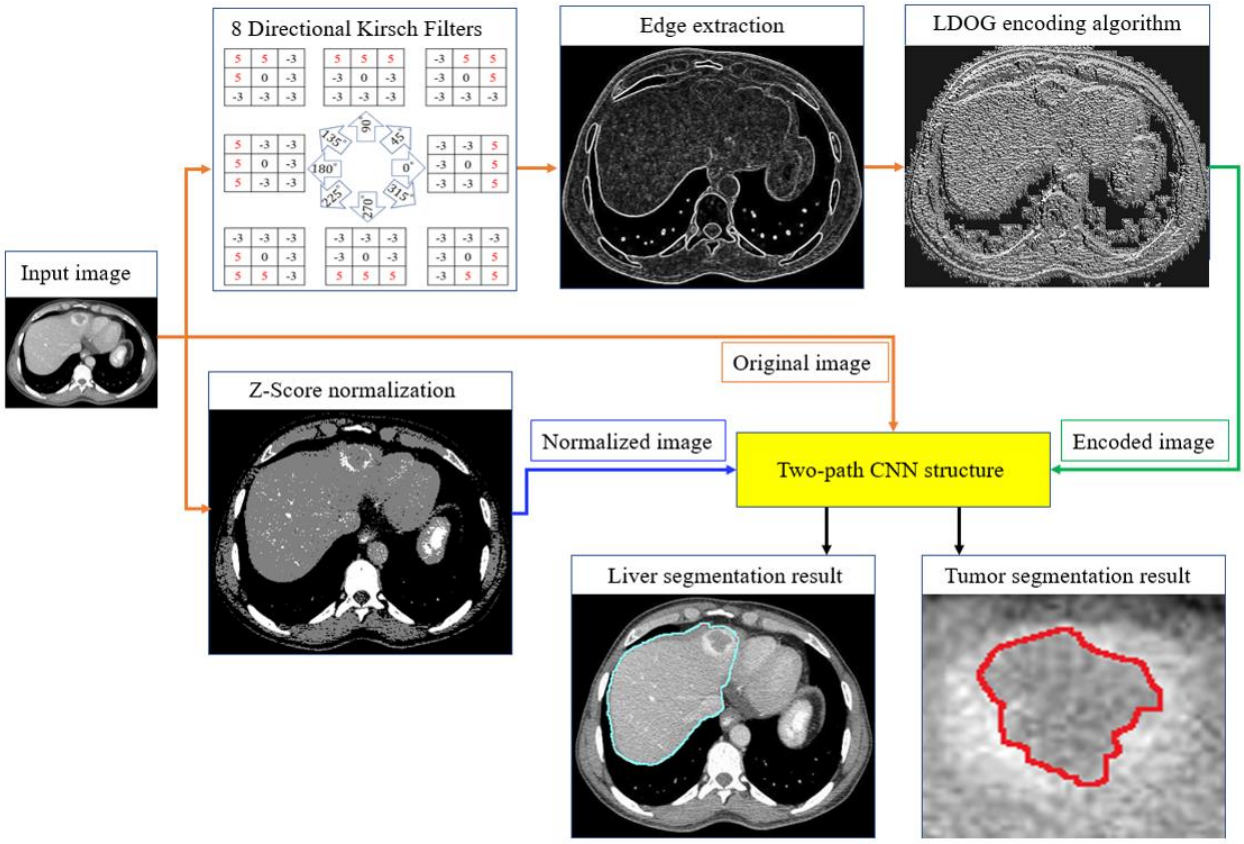

Fig. 1. Schematic of our framework for extracting the borders of the liver and lesion (tumor).

2.1 Z-Score normalization

As mentioned in (Willner et al. 2015; Bae, Lee & Hong 2020) when we are dealing with a CT image of a liver, due to the presence of the noise, there is a deviation in the Hounsfield units about a mean that leads to high variance in intensity in each image (Hounsfield units is a quantitative scale for describing radio density). These important noises in CT images can be categorized into three sources: 1) electronic noise. 2) noise of the reconstruction process. 3) stochastic noise.

The main source of the noise in these images is stochastic noise that can be reduced in the imaging process by increases the number of photons. However, in received images from any clinic

or hospital, still there is a significant amount of noise which has to be diminished before the segmentation step. So, to overcome this issue, a normalization approach is vital to be applied so that all the non-zero pixels inside the image have zero mean and unit variance (Jain, Shukla & Wadhvani 2018; Ranjbarzadeh et al. 2021). Equation. 1 demonstrates how to accomplish the Z-Score normalization (Robitaille et al. 2012).

$$Z = (x - \mu)/\sigma \qquad (1)$$

where σ and μ represent the standard deviation and the mean of the intensity of non-zero pixels, respectively. Also, x depicts the current pixel intensity.

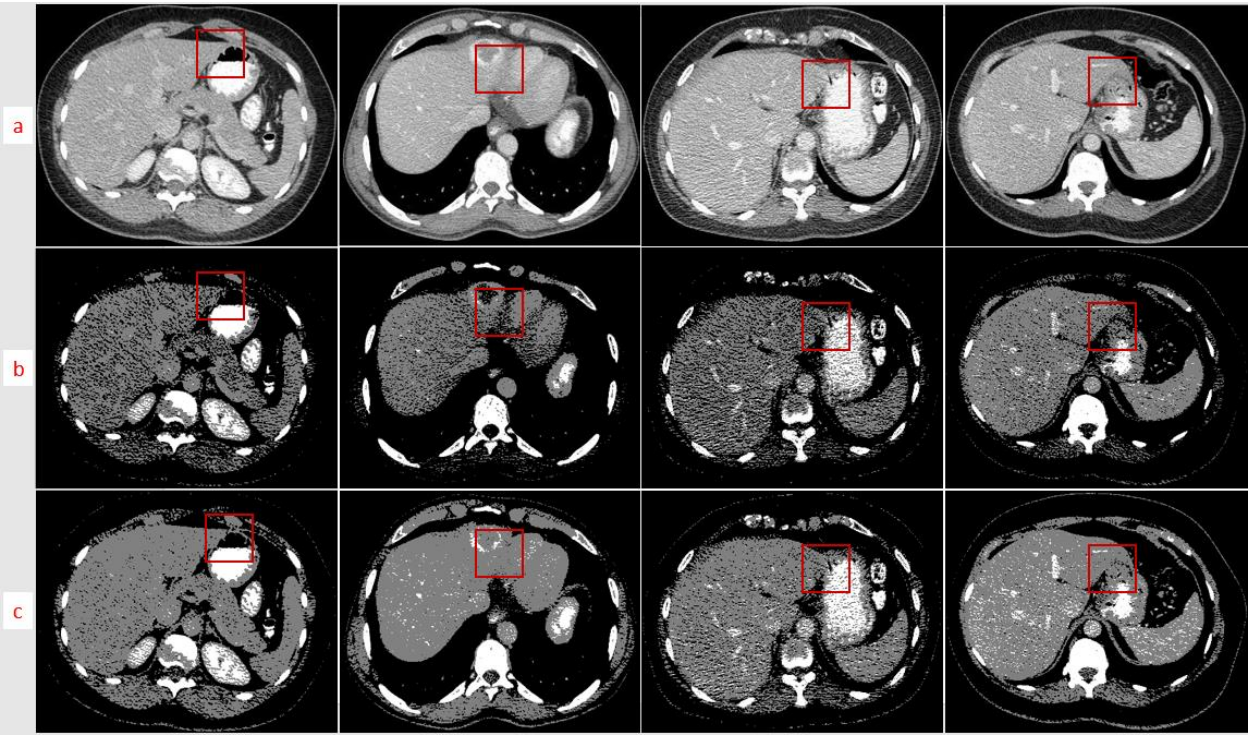

Fig. 2. An illustration of applying Z-Score normalization. There are four images from our dataset which indicate the difficulty of recognizing the liver border in the touching organs areas. a) Original images. b) Z-Score normalization by using non-zero pixels. c) Z-Score normalization by using all pixels.

By doing this method along with other following approaches, we are trying to obtain more accuracy of the final tumor and liver segmentation. The result of normalization is shown in Fig. 2.

In Fig. 2, the first row indicates the original CT images, the second and third rows are demonstrated the Z-Score output by using only non-zero pixels and all pixels, respectively. In each column, the original image and the results of normalization of that image using different initialization are shown. Initialization means which kind of pixels inside the image should participate in the normalization procedure. When we are using all pixels inside the image, the output image is much smoother than another method, but also finding the border of touching organs is much difficult (most of the time, it is impossible.). Using only the non-zero pixels lead to obtaining a gap between organs, so that border recognition can be easier. We indicate the key regions using the red windows (Fig. 2) that clarify the difference between the two methods visually. As is clearly shown in Fig. 2(b), in contrast to Fig. 2(c), the white elliptical objects (masses arising in the ribs) can be easily detected and many complicated borders are easily recognized.

2.2 LDOG algorithm

Texture analysis endeavors to explore the characterization of a surface such as colors, contrast, and shapes. In the encoding procedure of texture, the local descriptors are employed to convert the image into a new representation based on a pre-defined coding algorithm or code-book of visual patterns. As discussed in (Di Cataldo & Ficarra 2017; Song et al. 2018; Pourasad, Ranjbarzadeh & Mardani 2021), numerous types of descriptors can be applied to denote biological textures. Local Binary Pattern (LBP) and local ternary patterns (LTP) operators are the simple and efficient texture analyzer approaches that depend on the change of the intensity of surrounding neighbor pixels in clockwise or counter-clockwise to encode the low-level information of the curve, line, spot, edges and other local features in the image and consider the result as a binary number (Liu et al. 2016; Di Cataldo & Ficarra 2017; Karimi, Ranjbarzadeh Kondrood & Alizadeh 2017; Rakshit, Nath & Kisku 2018; Sotoodeh, Moosavi & Boostani 2019; Ghazouani & Barhoumi 2020). Due to the more stability of the gradient rather than gray-level intensity, in recent studies, pixel's gradient magnitude-based approaches such as local directional number patterns (LDN) and local directional patterns (LDP) have been gained much attention.

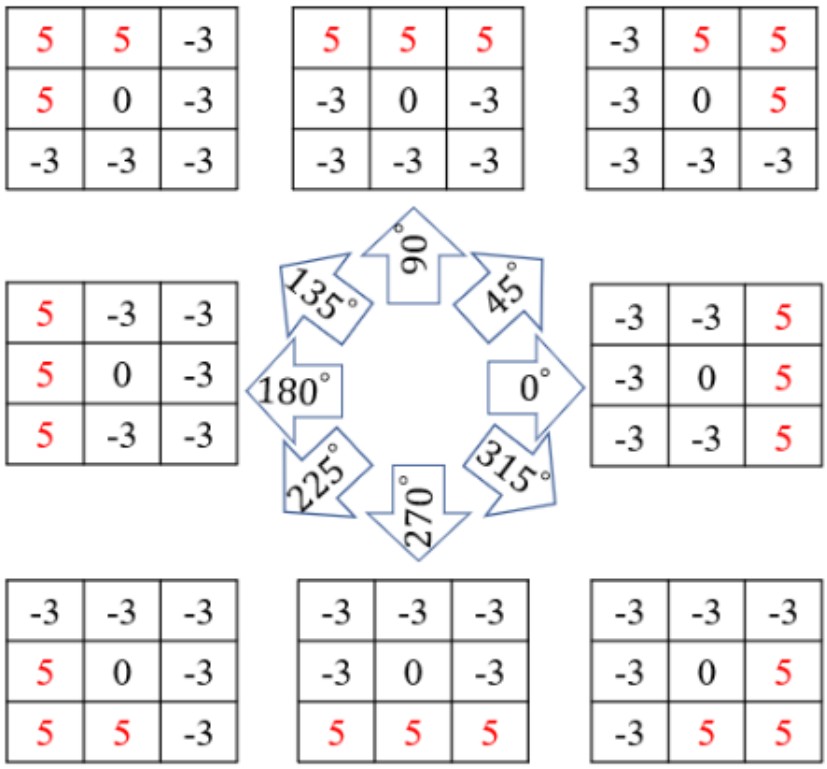

Fig. 3. Kirsch non-linear edge detector masks in eight directions (Ranjbarzadeh & Saadi 2020). The red color is used to emphasize the direction of the filter.

In this study, we represent a new encoding technique based on both gradient magnitude and gray-level intensity to create the illumination-invariant representations that are much useful for textural analysis. The first step for encoding the image in our study is obtaining all significant edges. It is accomplished by using the compass Kirsch filter. These non-linear edge detector kernels detect the edges and the maximum value generated by one of the eight filters that are rotated by 45° in the eight directions (Fig. 3) characterizes the edge direction (Li, Sang & Gao 2016; Luo et al. 2016; R. & Chandra 2016; Uddin et al. 2017; Punarselvam & Suresh 2019; Ranjbarzadeh, Saadi & Amirabadi 2020). An example of applying the non-linear edge detector to the CT images is illustrated in Fig. 4. In Fig. 4, there are three images from our dataset which indicate the difficulty of recognizing the liver and tumor borders due to the irregular shapes and complex tissues. In each column, the original image and the results of the edge extracting and encoding image are shown.

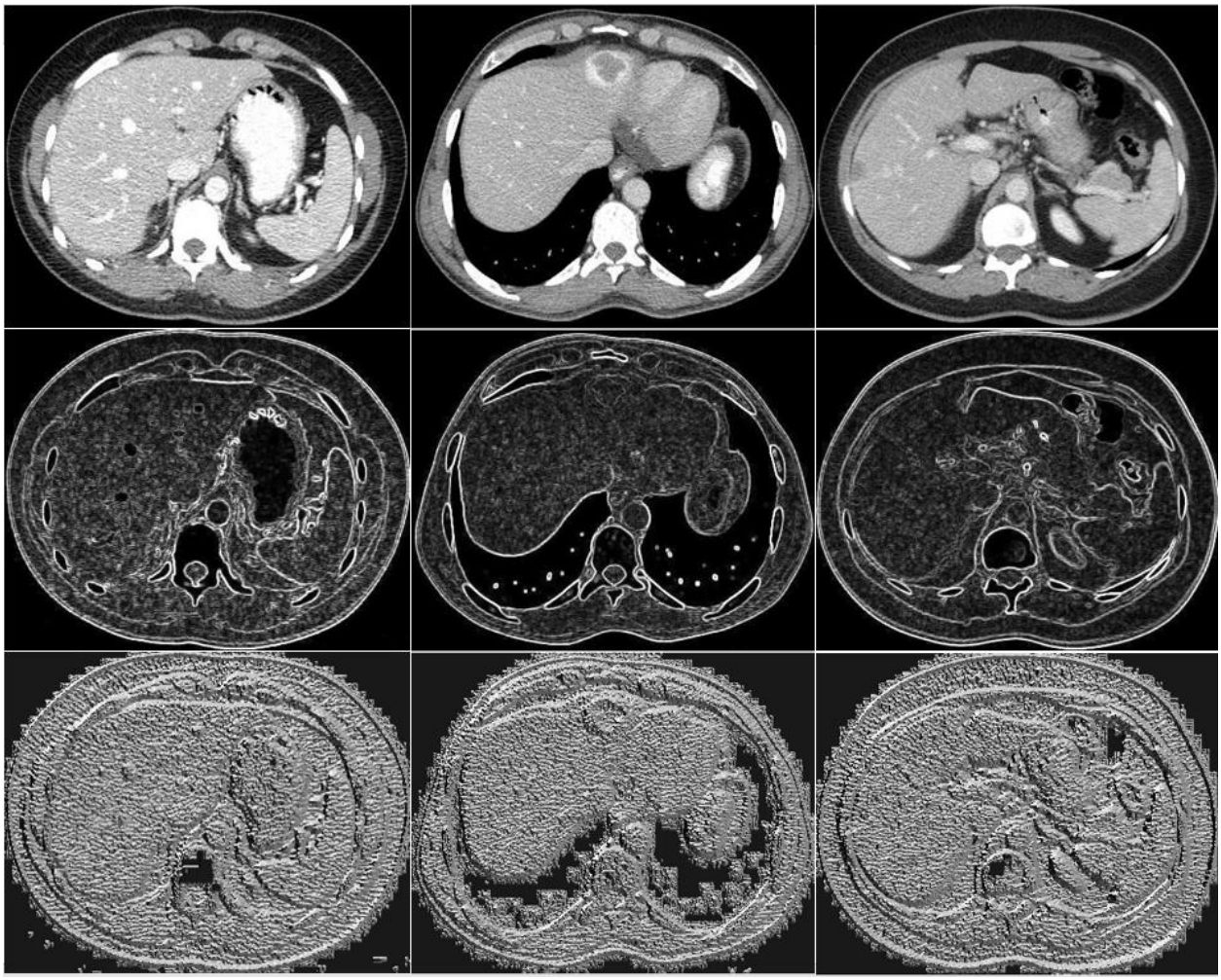

Fig. 4. The result of encoding the images. The first row indicates the original images. The second row depicts edge detection using the eight directional Kirsch filters. The third row illustrates the results of the LDOG method. By applying the LDOG encoding method more information about the structure of the image (textural information) can be extracted.

The Local Direction of Gradient (LDOG) is represented to produce the illumination-invariant representation of a CT image. This strategy computes a local demonstration of texture and is based on the selection of two vicinity group surrounding each pixel. For a better understanding of the algorithm, we clarify it visually with different possible colors in Fig. 5. Also, Fig. 6 demonstrates a simplified representation of the pseudo-code of the LDOG algorithm. This part causes a significant increase in exact segmentation, especially in tumor borders. Due to its discriminative capability and insensitivity to noise, the LDOG encoding texture can be utilized in numerous

applications in the computer vision field.

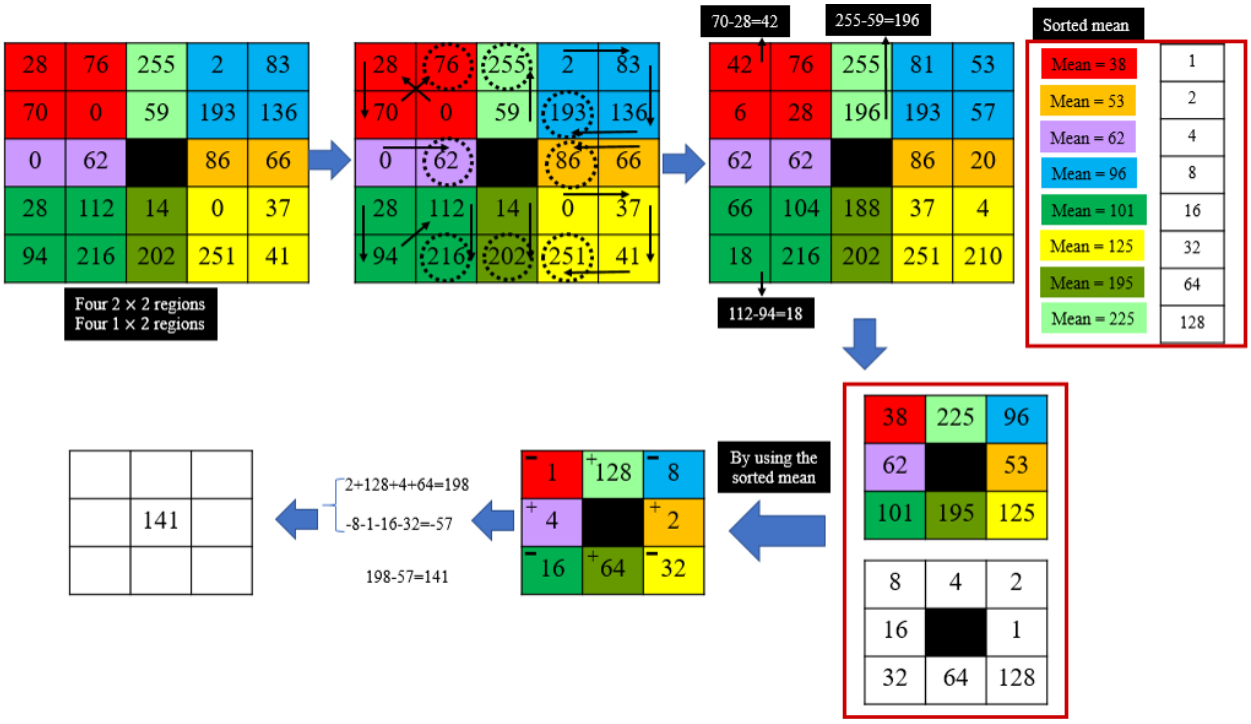

Fig. 5. The procedure of encoding an image using LDOG approach.

For encoding image based on the LDOG algorithm, firstly, we need to select a 5 × 5 window from the original image, which the location of the center of the window demonstrates the location of the calculated value inside the encoded image. In the second step, as depicted in Fig. 5, all 25 pixels inside the window are divided to four 2 × 2 and four 1 × 2 distinct patches. Then, all values inside each patch are sorted in descent order. In the next step, every sorted value subtracted from adjacent smaller value to compute the gradient value between them. These gradient value are replaced with smaller value, but maximum value inside of the patch still remains intact. In the next step, mean values of all eight patches are computed and are sorted in ascending order. Next, all eight sorted mean values are located in a 3 × 3 template patch based on their original location inside the 5 × 5 window. Each value inside the in the 3 × 3 is labeled and replaced with the sorted number. Then, the plus sign is applied to N, S, W, and E directions and others receive the minus sign. Finally, all mean values with their own sign are added together and the final value is generated. This final value represents the encoded value in the same location of the original image.

Proposed LDOG method

Input: A 5 × 5 patch
Output: One value

1. Read the new patch with overlapping.
2. Segment all pixels to four 2 × 2 and four 1 × 2 groups.
3. Calculate the gradient of each two values with minimum value difference inside each group. The maximum value inside each group still remains intact.
4. The mean value of each group is calculated.
5. All eight mean values are sorted.
6. An empty template 3 × 3 patch is considered that center pixel is ignored. Then values of 1, 2, 4, 8, 16, 32, 64 and 128 are written inside it in the clockwise direction.
7. All eight sorted mean values are located in the template patch in the clockwise direction and replace with one of eight values in step 6.
8. The plus sign is applied to N, S, W, and E directions and others receive the minus sign.
9. All mean values with their own sign are aggregated and final value is generated.

Fig. 6. Pseudo code of Local Direction of Gradient strategy.

2.3 Convolutional Neural Network Design

Pattern recognition is the science of distinguishing patterns by computers and is closely related to machine learning and artificial intelligence for countless applications like biomedical and biological imaging. In today's pattern recognition applications and methods, the convolutional neural network (CNN) structures represent a huge breakthrough in image analyzing. The CNN structures largely exploit the texture content and can be found at the core of everything from remote sensing to automated tumor segmentation (Mahmood et al. 2017; Ullah et al. 2018; de Assis Neto et al. 2020; Cecotti et al. 2020).

This neuron-based network, that has a grid-like topology (1D grid for time series data and 2D or 3D grid for image data) enables us to extract characteristics efficiently from the image

content by passing through a series of convolution layers (Affonso et al. 2017; Golrizkhatami & Acan 2018; Dang et al. 2019). This grid-like structure comprises some trainable weights and biases and is utilized for feature extraction, prediction, and classification. These trainable weights can be defined randomly at the beginning. The core building block or main part of any convolutional neural network is described as the convolutional layer that calculates the dot product between a set of learnable filters (two-dimensional arrays) and input data (image) (Chen et al. 2018; Özyurt et al. 2019). Routinely, the first convolution layers are located at the beginning of the CNN architecture and play a key role as the prior layer for extracting features from an input image (Bengio 2012; Mahmood et al. 2017). To control the size of the feature maps, pad the input matrix with zeros (zero paddings) in the convolving process that can be chosen.

Basically, the Spatial and Temporal dependencies are able to successfully capture by the convolutional layers (Petersen, Rodrigues & Pereira 2019). The multiplying operation that performs a dot product (algebraic operation) the values of the kernel with any receptive field and related pixels on the depth of the input can be considered as the convolutional operation (Dolz, Desrosiers & Ben Ayed 2018). The kernel with the arbitrary size of the receptive field is a matrix of numbers that requires to have as depth as the input image. For instance, the input image and filter are dimensions of $150 \times 150 \times 3$ and $\sim \times \sim \times 3$, respectively. Where $\sim$ is an odd number. In contrast to the non-convolutional neural network, in the convolutional layers (include stacks of 2D or 3D filters), every component of the kernel array is convolved with the input 2D or 3D data (gray-scale or color image) more than once which takes up 80% to 90% of the execution time (Mahmood et al. 2017).

Note that the final output image is much smaller than the size of the input image. The reduction in the dimension of the image depends on the kernel (filter) size applied for the convolution process and also the dimension and shape of the strides. In the procedure of convolving a filter, the stride represents the number of pixels that a center of the kernel is dislocated while iterating through the input image (Torres et al. 2018).

The output of the convolution layer defines as the input for an activation layer (Ting, Tan & Sim 2019). To diminish the effect of the vanishing gradient difficulty, an activation function is employed for each feature map which leads to improving computational effectiveness by inducing sparsity (Morabito et al. 2018). Moreover, by applying the smaller size of the kernels comparison to the input image, the number of connections between the output and the input layer

is diminished and sparse connectivity can be achieved (Mahmood et al. 2017).

As in object recognition, there are no matters what the size or location of the object is, only the spatial variance need to be extracted. To accomplish this, a downsampling layer (subsampling layer) is used. By reducing the size of each activation map (feature map), the efficiency of feature extraction is increased. Also, since the number of pixels in each feature map generated in the previous layer (in both column and row) is decreased, it leads to a decrease in model computing time and control overfitting (Zhong et al. 2019; Ranjbarzadeh et al. 2021).

An appropriate strategy for dimensionality reduction of feature maps needs to be robust to alter the high-frequency information (vital information) and preserves significant features (Bengio 2012). This reduction happens by using a filter with a predefined size that moves across the extracted feature map taking the average or the maximum of the adjacent values selected by the filter that called average pooling and max pooling, respectively. In this paper, the max-pooling layer is utilized that partitions the feature map into a set of areas that don't overlap and then selects the maximum illumination value inside each region. The max pooling technique also exerts as a noise suppression approach (Yin et al. 2016; Doğantekin et al. 2019).

Moreover, a Fully-Connected layer (FC) can be applied for gaining more high-level features in an input image. Each node inside the FC layer with its learnable corresponding weight multiplies each input vector and outputs the sum of the nodes are totalized to a learnable bias before performing an activation function operation (Shen et al. 2019; Singh et al. 2020).

Training a CNN needs learning weights and biases for each layer so that a cost function is minimized. The minimization of a cost function is accomplished iteratively using a gradient descent strategy which involves the calculation of fractional derivatives of the cost function. Moreover, computing of this cost function is accomplished by the backpropagation algorithm (Husain, Dellen & Torras 2017).

For the training step, while a CNN structure is used, it is essential to have a relatively big dataset. When we are working with a small dataset, it is extremely straightforward and easy for the model to be specialized according to its application area and based on a few sets of rules (to be less intelligent). So there are two main approaches to overcome this issue. Firstly, a transform learning method can be used to bring some trained weights and biases into our algorithm rather than randomly chosen them at the first (He et al. 2016; Khatami et al. 2018; Salaken et al. 2019; Efimova, Shalamov & Filchenkov 2020). Secondly, the diversification and number of training

examples are artificially boosted (Dvornik, Mairal & Schmid 2019). We utilized the affine transformations, blurring, contrast changes, hue/saturation changes, and random intensity variation approaches of data augmentation in this paper.

2.4 Proposed Convolutional Neural Network model

As mentioned in the previous section, CNNs are more capable to extract significant information from an input image. So, in this work, we designed a novel pipeline based on the combination of local and global features to classify each pixel inside the image into three classes; tumor border, liver border, or other tissues. Also, to improve the segmentation accuracy, the three input images including original image, Z-Score generated image, and LDOG generated image were used. Using these extra two images as the input of the network cause the network can be learned faster with a high degree of the accuracy. The flowchart of this complex strategy is shown in Fig. 7.

When we are dealing with the CNN models which are having millions of parameters, the best-suited feature maps are produced based on the best possible probability to calculate class probability. Although many CNN architectures have been proposed for liver and lesion (tumor) segmentation in prior studies, none of them has focused on combining and integrating the textural encoding approach and CNN. Our strategy comprises of multiple input feature maps corresponding to a different image and related textural features (extracted by LDOG). It means the pixels inside each patch of three images (3D patches) are applied together into the network. Since diverse images or textural features definitely comprise complementary and detailed information, our experimental results indicate that this complex multi-input technique is effective in enhancing the value of the evaluation indexes.

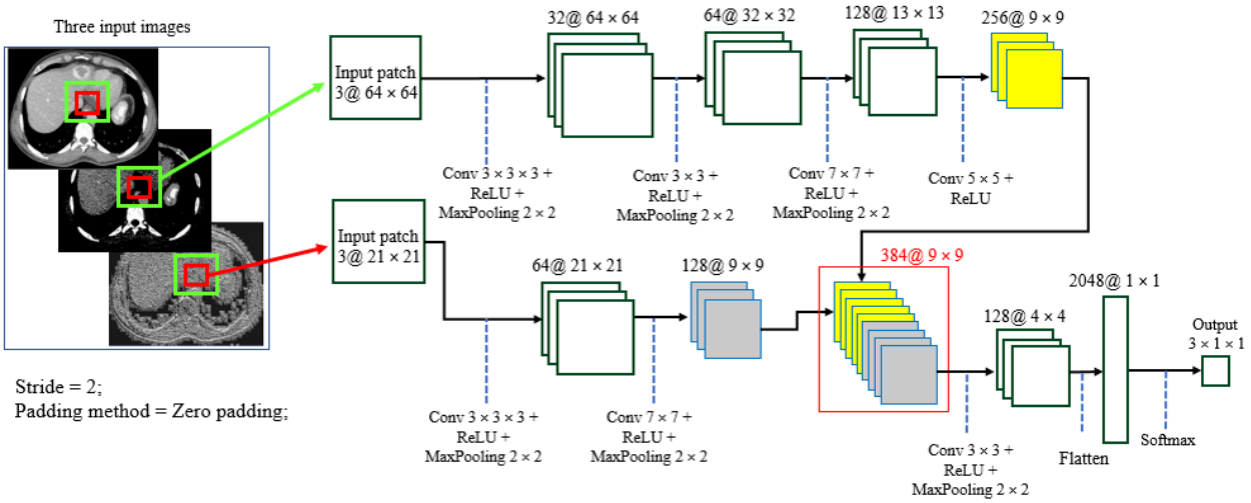

Fig. 7. Our implemented two-path way CNN pipeline.

As is clearly demonstrated in Fig. 7, our CNN architecture is based on a two-path feature extraction approach (Cascading). In some complex textural images such as CT images, there are some borders of touching objects (organs) that cannot be segmented properly without the knowledge of neighbor features in a little further of the pixel location. So, in this study, we are considering both local and semi-global features (information) around each target pixel to categorize the target pixel. This means that we select a 21×21×3 patch with the central location of the current pixel as a local window and also a 64×64×3 patch with the central location of the current pixel as a semi-global window. Where the number of three represents the three different input images. The function of the first convolutional layers is detecting the low-level features such as curves points, and edges. Moreover, the higher-level features such as ears, legs, and face are identified in the deeper layer of the model (Zhong et al. 2019; Zhang, Wu & Li 2020; Ranjbarzadeh et al. 2021).

The semi-global patches give us more information about the similar touching tissues to draw a segmentation line between them which leads to increasing the dice score significantly. Moreover, the result of the algorithm for tumor segmentation highly depends on information extracted from the semi-global windows. In Table 1, we demonstrate the effect of applying both local and semi-global windows in the final result of the method. The best accuracy related to the best size of the patches is illustrated in bold.

Table 1. Evaluating the accuracy of the proposed liver and tumor segmentation approach by applying different dimension of the patches to the final result of the method. The best outcomes are shown in the bold form.

Size of the local patch	**Size of the semi-global patch**	**Liver segmentation (DICE %)**	**Tumor segmentation (DICE %)**
11×11× 3	30×30× 3	40 %	63 %
11×11× 3	40×40× 3	45 %	66 %
11×11× 3	50×50× 3	37 %	70 %
11×11× 3	60×60× 3	42 %	60 %
11×11× 3	64×64× 3	41 %	60 %
11×11× 3	70×70× 3	44 %	55 %
11×11× 3	80×80× 3	38 %	53 %
15×15× 3	30×30× 3	76 %	71 %
15×15× 3	40×40× 3	78 %	75 %
15×15× 3	50×50× 3	80 %	76 %
15×15× 3	60×60× 3	82 %	76 %
15×15× 3	64×64× 3	81 %	74 %
15×15× 3	70×70× 3	77 %	70 %
15×15× 3	80×80× 3	77 %	69 %
21×21× 3	30×30× 3	85 %	87 %
21×21× 3	40×40× 3	87 %	87 %
21×21× 3	50×50× 3	88 %	90 %
21×21× 3	60×60× 3	89 %	92 %
21×21× 3	**64×64× 3**	**92 %**	**95 %**
21×21× 3	70×70× 3	90 %	93 %
21×21× 3	80×80× 3	88 %	90 %
25×25× 3	30×30× 3	90 %	94 %
25×25× 3	40×40× 3	86 %	90 %
25×25× 3	50×50× 3	85 %	85 %
25×25× 3	60×60× 3	85 %	84 %
25×25× 3	64×64× 3	84 %	82 %
25×25× 3	70×70× 3	82 %	80 %
25×25× 3	80×80× 3	83 %	81 %

The local extracted patch has a dimension of 21×21×3, which 3 indicates the number of the input images. Then these 3D patches are convolved to the 64 kernels with the receptive field of 3×3×3 in order to produce the 64 feature maps with 2D dimensions. In the next step, the number and size of filters are changed to 128 and 7×7, respectively.

Unlike the path related to the local features, in the feature extraction procedure of the semi-global patch, there are four convolutional layers with increasing the number of the kernels. After extracting both semi-global and local features in two distinct paths, all features need to be concatenated. In this step, there are 384 feature maps with a size of 9×9 which are fed to the next convolutional layer. Next, all generated 2D feature maps (128 feature maps) are transformed into a 1D feature vector with a size of 2048×1 features. Finally, using a Softmax layer all features all labeled to one of three possible classes (1 represents the liver border, 2 indicates the tumor border, and 3 shows a non-important pixel).

Our network was learned through stochastic gradient descent (Wahab, Khan & Lee 2017) for minimizing the value of this cross-entropy loss (cost function) in Eq. (3), that maximize the final accuracy using calculating the amount of discrepancy between actual (ground-truth labeled images) and predicted (estimated) output for liver and tumour segmentation. The output layer three logistic units were employed in the output layer, to generate the probabilities of the given sample pattern relating to either of the three predefined outputs. The backpropagation method with respect to the design parameters was utilized to calculate the derivative of the objective function.

$$loss_i = -log\left(\frac{e^{U_K}}{\sum_{d=1}^{O} e^{U_d}}\right) \qquad (3)$$

where $loss_i$ indicates the loss for training sample i, U_K shows the un-normalized production value for the ground-truth of the predefined class K and can be attained by multiplying the outputs from the former FC neurons with the parameters of the corresponding logistic unit. To obtain a normalized value (score) for each class between 0 and 2, the denominator aggregates the scores for all the logistic units O. As in this study there are only three final neurons at the end of the model, in the above equation O is equal to three, which normalizes the output score and is considered as a probability score. It means each pixel is categorized into one of three classes. Also, to reduce the overfitting effect by controlling the fitting process, a dropout layer (Srivastava et al. 2014) (before the FC layer) with a 35% dropout probability, was incorporated into the proposed CNN model.

3. Experiments

3.1 Datasets and Evaluation metrics

Our novel TPCNN model and seven techniques (Concave and Convex Points (CCP) (Ranjbarzadeh & Saadi 2020), Cascaded deep convolutional encoder-decoder neural networks (EDCNN) (Budak et al. 2020), Laplacian Mesh Optimization (LMO) (Chartrand et al. 2017), Graph Cuts (GC) (Liao et al. 2017), Deep Belief Network (DBN-DNN) (Ahmad et al. 2019), Adaptive Scale-Kernel Fuzzy Clustering Model (FCM) (Cai 2019), and multiphase contrast-enhanced FCN (MC-FCN) (Sun et al. 2017)) were evaluated in different scenarios on non-public CT data sets to evaluate the validity, reliability, and efficiency of experiments.

In this study, we used the dataset in (Ranjbarzadeh & Saadi 2020), but the whole samples have been increased to 1000 patient cases (most of them include slices with a vague border and touching organs border.), who underwent CAD screening at some hospitals in Iran. This dataset contains over 20,000 scan slices from around 1000 patients each having a resolution of 512×512 pixels. The dataset is divided into 10000 benign and 10000 malignant samples. To detect the border of the liver and lesion tissues accurately, two experienced specialists manually segmented the borders in all slices in a slice-by-slice manner. It should be mentioned that, by using a data augmentation approach, all samples have been increased up to 100,000 scan slices which 70 percent for training, 20 percent for the test, and the rest are used for the validation process.

To prove the efficiency and robustness of the proposed technique in terms of different six performance measures, the result of our method and all baseline models were investigated by comparing the result of each them with its corresponding reference image (ground-truth). The significant accuracy/reliability of our structure was evaluated using the relative volume difference (RVD), root mean square symmetric surface distance (RMS SSD), volume overlap error (VOE), average surface distance (ASD), maximum surface distance (MSD), and dice similarity (DICE) (Heimann et al. 2009; Vinícius dos Santos Ferreira et al. 2018; Ranjbarzadeh & Saadi 2020). For RVD, ASD, RMS SSD, and VOE, a zero value represents a complete match with the reference image (ground-truth). The DICE, RVD, ASD, RMS SSD, MSD, and VOE measures can be formulized as follows (Ranjbarzadeh & Saadi 2020; Tang et al. 2020):

$$
\begin{cases}
DICE(M_{seg}, M_{gnd}) = \left(2 \times \dfrac{M_{seg} \cap M_{gnd}}{M_{seg} + M_{gnd}}\right) \times 100\% \;\; or \;\; \left(2 \times \dfrac{TP}{2TP + FP + FN}\right) \times 100\% \\
VOE(M_{seg}, M_{gnd}) = \left(1 - \dfrac{M_{seg} \cap M_{gnd}}{M_{seg} \cup M_{gnd}}\right) \times 100\% \\
RVD(M_{seg}, M_{gnd}) = \left(\dfrac{M_{seg} - M_{gnd}}{M_{gnd}}\right) \times 100\% \\
ASD = \dfrac{1}{\left|B_{M_{seg}}\right| + \left|B_{M_{gnd}}\right|} \times \left(\sum_{x \epsilon B_{M_{seg}}} d\left(x, B_{M_{gnd}}\right) + \sum_{y \epsilon B_{M_{gnd}}} d\left(y, B_{M_{seg}}\right)\right) \\
RMS\ SSD\ (M_{seg}, M_{gnd}) = \left(\sqrt{\dfrac{1}{\left|B_{M_{seg}}\right| + \left|B_{M_{gnd}}\right|}} \times \sqrt{\sum_{x \epsilon B_{M_{seg}}} d^2\left(x, B_{M_{gnd}}\right) + \sum_{y \epsilon B_{M_{gnd}}} d^2\left(y, B_{M_{seg}}\right)}\right) \\
MSD\ (M_{seg}, M_{gnd}) = \max\left\{max\left(d\left(x, B_{M_{gnd}}\right)\right), max\left(d\left(y, B_{M_{seg}}\right)\right)\right\}
\end{cases}
\tag{4}
$$

where $\boldsymbol{M_{seg}}$ and $\boldsymbol{M_{gnd}}$ denote the proposed method segmentation result and reference labeled image, respectively, and $\boldsymbol{B_{M_{seg}}}$ and $\boldsymbol{B_{M_{gnd}}}$ illustrate the borders of the proposed method segmented result and ground-truth image, respectively. When we apply Boolean data, the TP, FP, and FN represent true positive, false positive, and false negative, respectively. The dice similarity coefficient (DSC) equals one for a perfect segmentation. MSD can measure the distance between each segmentation pixel from its corresponding pixel in the ground-truth border. RVD is a statistical approach applied to measure the different sizes between segmented results and ground truth images. The observed positive value shows over-segmentation and a negative value of RVD indicates under-segmentation. So, the best-observed value is a zero value which represents the segmented volume is equal to the reference labeled image.

3.2 Experimental Results and Discussions

Our algorithm was implemented in Matlab 2019b, and the experiments were run on Intel I7-7600@3.4 GHz. Also, the computer equipped with a GEFORCE GTX1070 Ti GPU, windows 10 operating system and 16 Gigabytes of RAM. The outcomes of our segmentation model from the CT images are evaluated according to the ground truths and described in Tables 2 and 3. Slices with an uneven density, irregular texture, large liver region, not well-defined border (blurred or unclear margin), and at least one attached component have the most part of the train and test sample in the dataset to learn and evaluate the proposed architecture.

To exemplify the importance of using the combination of the Z-Score, LDOG, and CNN algorithms, Figs. 8 and 9 show the results of the proposed technique (denoted by solid colour line)

on a few slices exhibiting ambiguous boundaries, liver, irregular texture, heterogeneous appearances, intensity inhomogeneity, and various livers and tumour shape. It can clearly be seen that due to intensity inhomogeneity in the input image, the lesion and liver areas were not suitably identified when the EDCNN (Budak et al. 2020) and DBN-DNN (Ahmad et al. 2019) strategies were utilized.

As demonstrated in Fig. 8, segmentation by applying the DBN-DNN (Ahmad et al. 2019) technique indicates the least match with the reference labeled image when similar intensity values are present close to the margins of other body parts. The outline of the final liver segmented volume acquired using our structure is similar to the corresponding reference labeled image (ground-truth image). As it is clearly demonstrated the CCP (Ranjbarzadeh & Saadi 2020) seems to perform better than the DBN-DNN (Ahmad et al. 2019) and EDCNN (Budak et al. 2020) in segmenting liver border between two touching organs, whereas our method is much better not only in detecting such borders but also it can recognize objects with irregular shapes and complex tissues more accurately. Also, the DBN-DNN (Ahmad et al. 2019) and EDCNN (Budak et al. 2020) techniques under-segment the long and thin areas and over-segment similar objects with equivalent illumination especially the touching organs. Moreover, such algorithms are prone to boundary leakage, especially on the blurred liver and tumour boundaries. The proposed segmentation algorithm also has higher accuracy and has not significant boundary leakage, under-segmentation, or over-segmentation, predominantly in specific regions with touching objects (organs). As is shown, the use of the Z-Score algorithm in our algorithm leads to contrast enhancement of the liver border to obtain better performance in both liver and tumour segmentation.

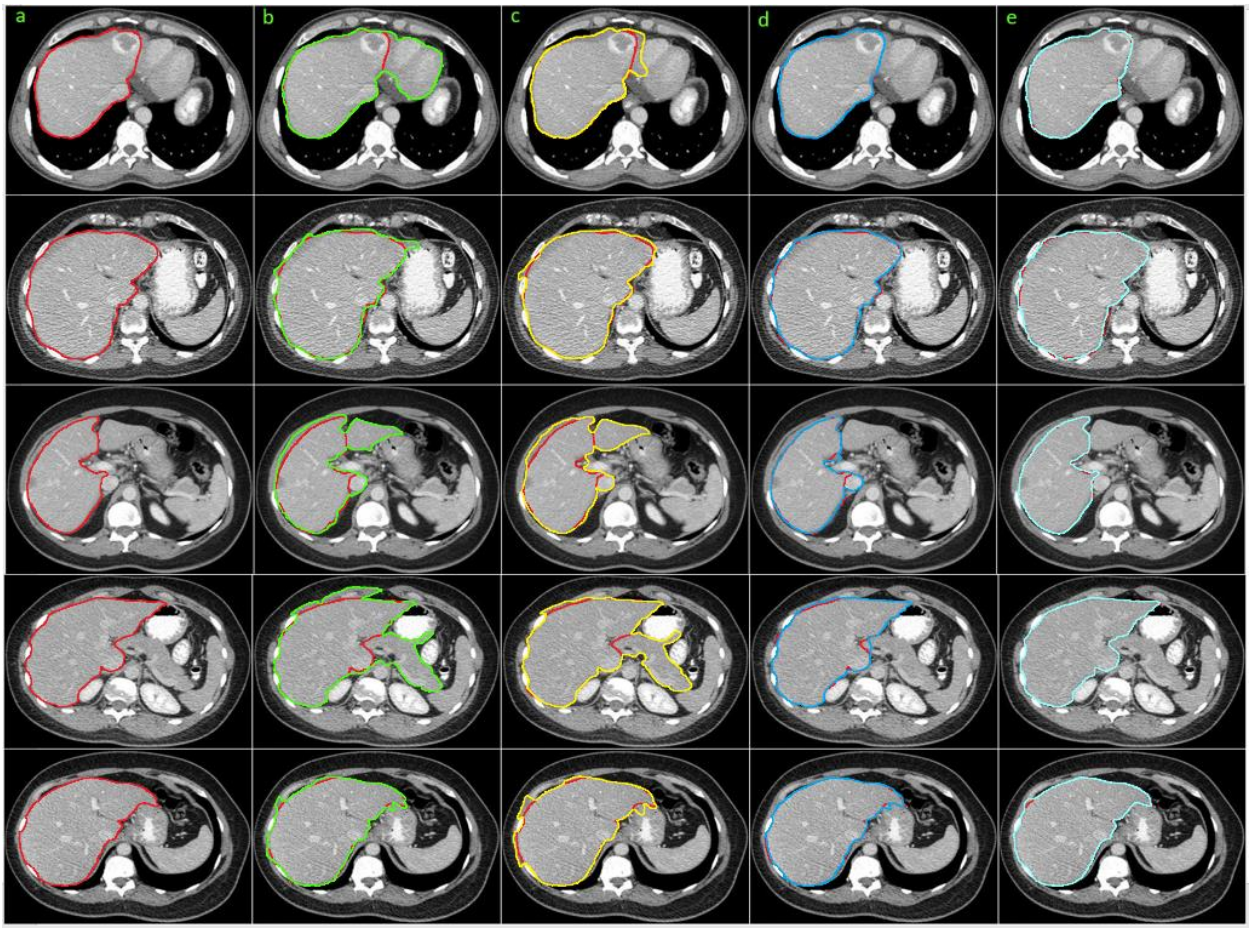

Fig. 8. Comparisons between four different algorithms for liver segmentation. (a) Input CT images. Segmentation based on the (b) Deep Belief Network (DBN-DNN) (Ahmad et al. 2019) (c) EDCNN (Budak et al. 2020), (d) Concave and Convex Points (CCP) (Ranjbarzadeh & Saadi 2020), and (e) Ours methods. The red contours illustrate the ground-truth border.

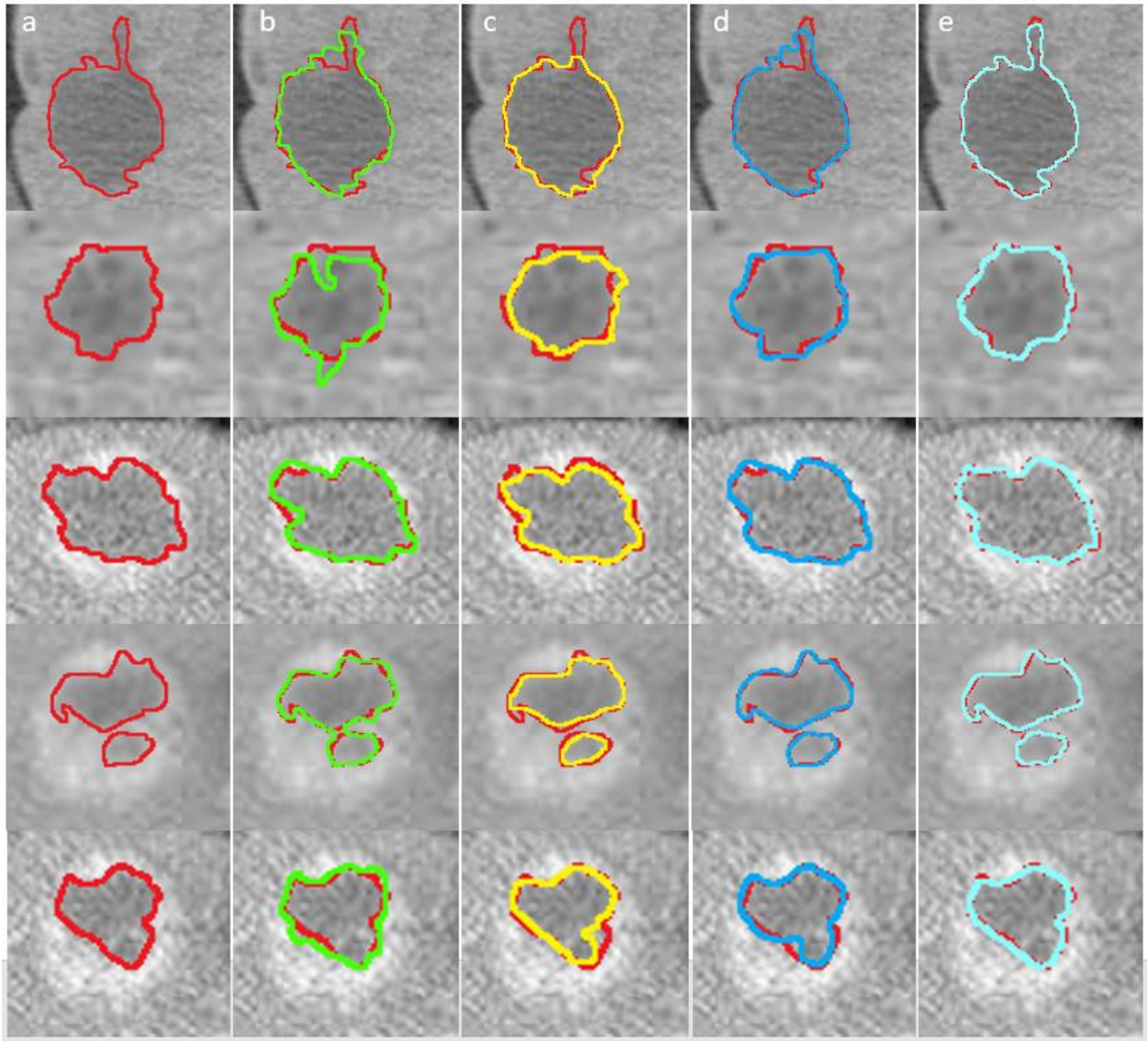

Fig. 9. Comparisons between four different algorithms for tumour segmentation. (a) Input CT images. Segmentation based on the (b) Deep Belief Network (DBN-DNN) (Ahmad et al. 2019) (c) EDCNN (Budak et al. 2020), (d) Concave and Convex Points (CCP) (Ranjbarzadeh & Saadi 2020), and (e) Ours methods. The red contours illustrate the ground-truth border.

Fig. 9 indicates the lesion boundary that was detected using the proposed pipeline and the DBN-DNN and EDCNN techniques. Considering the heterogeneous textures, blur boundary, and different sizes of the tumour, it is more evident that the proposed structure suitably finds the tumour border, which displays its robust performance on countless lesion outlines.

The implemented model illustrates that a higher segmentation robustness can be gained when other encodings of the input image are provided (representation of information inside the image in another way), meaning more key features are available to the distinction between classes. Due to the combination of the Z-Score and LDOG encoding algorithms, the proposed pipeline reaches an acceptable result with respect to all the assessment indexes. The segmentation accuracy and performance of the proposed model was assessed based on the dice metric, as demonstrated in Figs. 10 and 11 for detecting the border of liver and tumour, respectively. It is worth to mention that the obtained averages of the dice score for the segmented region of the tumour and liver altered from 90% to 95% and 85% to 92% respectively. This illustrates that the intensity values, texture, and shape of the target organs can define the segmentation accuracy of the tumour and liver.

Considering the Dice value from the above results and the Figs. 10 and 11, we conclude that for liver and tumour segmentation at the presence of similar intensity values across the liver, the use of the proposed algorithm shows more promising results over the others.

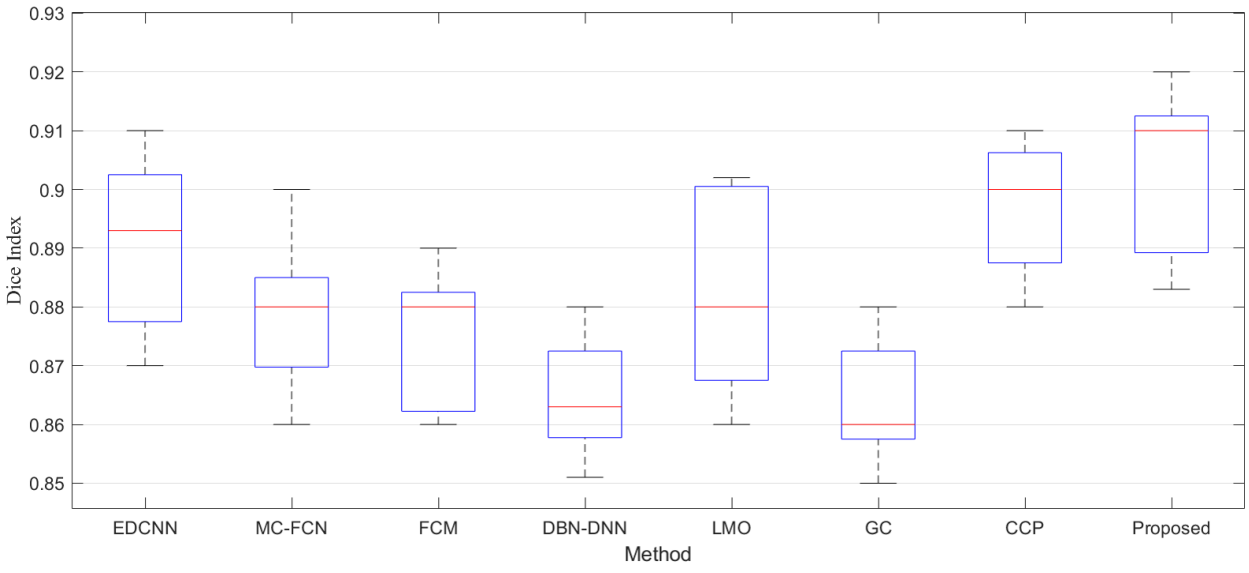

Fig. 10. The result of the algorithm performance evaluation in term of the dice scores for eight techniques applied for automated liver segmentation.

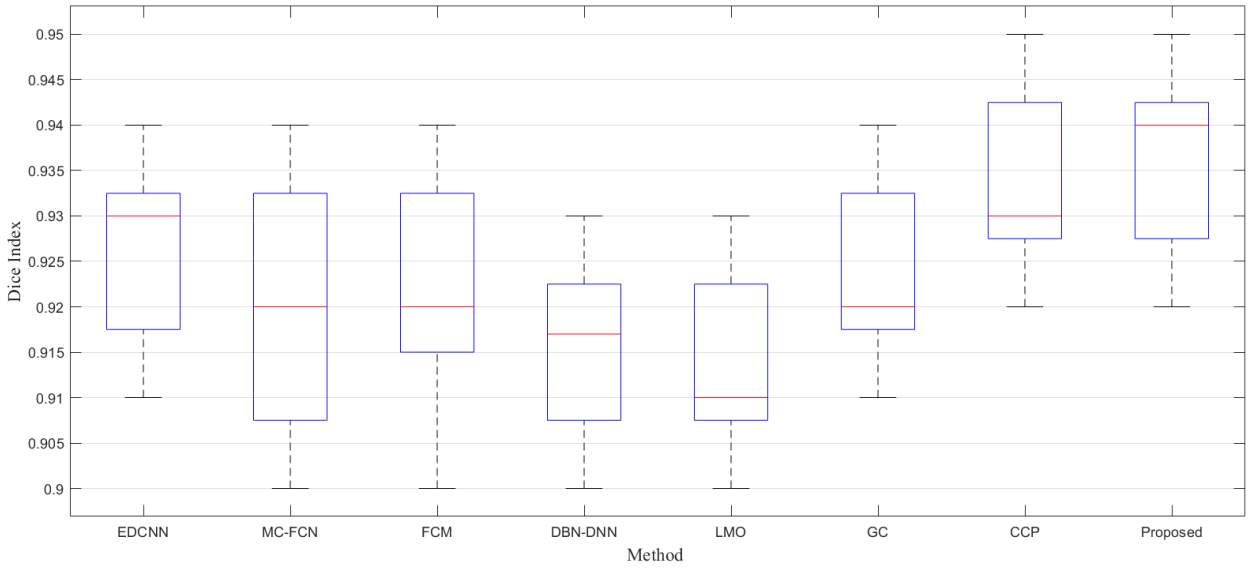

Fig. 11. The result of the algorithm performance evaluation in term of the dice scores for eight techniques applied for automated tumor segmentation.

Tables 2 and 3 appraise our automated segmentation model with the outcomes from seven recently published techniques in the field of tumor and liver segmentation.

For each index in Tables 2 and 3, for a quantitative comparison the highest ASD, VOE, RVD, MSD, and RMS values are highlighted in bold. The results of every five measurements are represented by standard deviation and mean of our dataset. Our automated detection algorithm for recognizing the border of the liver reaches a smaller mean in all five indexes. The mean VOE is meaningfully altered between all investigated techniques, while the RMS displays the lowest variance. There is a minimum difference between obtained RVD values using the LMO and GC algorithm for liver segmentation, whilst this is valid for the GC and MC-FCN techniques in tumor segmentation. The mean ASD of the CCP structure is partly similar to the proposed algorithm; however, this approach has a significantly different RVD. The highest VOE, MSD, and ASD for liver and tumor segmentation were obtained using DBN-DNN method. Also, this model gets the highest under-segment result among all evaluated algorithms. Both the LMO and DBN-DNN techniques indicate a large standard deviation in the RVD; however, a significant standard deviation in MSD value is observed in FCM and DBN-DNN methods. A lower standard deviation for a quantitative comparison can be realized in the MSD and RVD values when using the CCP, EDCNN and MC-FCN techniques, respectively. The RVD metric for LMO, FCM, DBN-DNN,

and CCP approaches are less than zero. In addition, the mean VOE and MSD of the technique utilized by DBN-DNN and GC for liver segmentation were noticeably higher as compared to our outcomes.

The segmentation accuracy is significantly higher for our approach and CCP compared with DBN-DNN and FCM, for abnormal shapes. There is no meaningful difference in dissimilar intensity. This indicates that using two more encoding images along with the original image is adversely effecting boundary detection in the regions with massive and similar densities. This could perhaps be due to representing key features in the densest areas; however, further work would be required to investigate this in more detail. Based on the information given above, the DBN-DNN method demonstrated the poorest performance among all the eight methods. Tables 2 and 3 exhibit the superiority of our model in terms of all five measures.

Table 2. Quantitative comparison of liver segmentation results using the proposed strategy and seven recently published techniques. The obtained values are based on relative volume difference (RVD), root mean square symmetric (RMS) surface distance, Volume overlap error (VOE), maximum surface distance (MSD), average surface distance (ASD).

Technique	RMS (mm)	VOE (%)	ASD (mm)	RVD (%)	MSD (mm)
LMO (Chartrand et al. 2017)	3.8±0.7	12±6.1	3.7±0.6	-8.5±7.3	14.7±3.5
GC (Liao et al. 2017)	4.2±0.5	16.4±6.3	5.8±0.4	8.8±6.2	20.6±4.1
MC-FCN (Sun et al. 2017)	3.8±0.6	11.9±5.2	3.7±0.4	4.7±5.3	14.9±3.2
DBN-DNN (Ahmad et al. 2019)	4.4±0.1	19.5±7.5	6.5±0.1	-12.3±7.7	22.1±5.3
FCM (Cai 2019)	4.2±0.2	13.3±7.2	4.4±0.3	-10.5±4.8	19.8±6.7
EDCNN (Budak et al. 2020)	3.2±0.4	10.5±4.4	2.4±0.2	7.3±5.7	13.4±3.5
CCP (Ranjbarzadeh & Saadi 2020)	2.6±0.1	3.8±2.7	1.6±0.4	-2.1±3.5	8.7±4.1
Proposed	**2.3±0.1**	**2.2±0.1**	**1.3±0.6**	**1.9±3.3**	**6.3±5.1**

Table 3. Quantitative comparison of tumor segmentation results using the proposed strategy and seven recently published techniques. The obtained values are based on relative volume difference (RVD), root mean square symmetric (RMS) surface distance, Volume overlap error (VOE), maximum surface distance (MSD), average surface distance (ASD).

Technique	RMS (mm)	VOE (%)	ASD (mm)	RVD (%)	MSD (mm)
LMO (Chartrand et al. 2017)	5.1±0.1	18.4±4.9	5.9±0.2	7.3±6.7	16.2±4.8
GC (Liao et al. 2017)	4.3±0.5	14.2±4.7	3.9±0.3	-5.5±6.8	12.3±4.6
MC-FCN (Sun et al. 2017)	4.7±0.4	17.5±6.2	5.4±0.4	-5.9±7.3	17.6±5.4
DBN-DNN (Ahmad et al. 2019)	5.3±0.4	19.1±7.5	6.3±0.2	-7.1±8.7	19.5±6.2
FCM (Cai 2019)	4.5±0.3	15.1±3.8	4.7±0.5	6.5±3.8	14.7±5.1
EDCNN (Budak et al. 2020)	4.2±0.3	12.6±4.7	3.1±0.1	-5.1±6.7	11.1±5.8
CCP (Ranjbarzadeh & Saadi 2020)	3.9±0.2	11.2±3.7	2.2±0.3	4.2±3.2	9.4±4.1
Proposed	**3.7±0.4**	**9.5±0.9**	**1.8±0.5**	**3.4±4.1**	**7.2±4.4**

.

The results of the liver and tumour segmentation in this study in terms of the VOE, RMS, MSD, RVD, and ASD metrics, as a quantitative comparison between our approach and seven recently published techniques, demonstrated in Tables 2 and 3.

It was found that recognition of the liver border was not significantly different between MC-FCN and LMO methods; however, while there are not touching organs the detection of these borders was significantly better with MC-FCN compared with LMO method. Also, the MC-FCN method relies on the premise that objects (tumors or liver) have distinguished borders, which is often not the case, leading to unacceptable boundary leakage.

Table 3 displays the results of tumour boundary recognition in terms of the RMS, VOE, RVD, ASD, and MSD measures. The RVD metric for DBN-DNN, GC, EDCNN, and MC-FCN methods is less than zero. Besides that, the DBN-DNN method obtains the highest mean score of RMS. The RMS values show that the CCP and our approaches produced the best outcomes among the eight models. Our structure and DBN-DNN approach have significantly different MSDs (Lu et al. 2014). In contrast, the mean value achieved utilizing the three different input image proposed pipeline is

644 considered similar to that of the CCP approach. Considering the ASD, the lowest and highest mean
645 values belong to our strategy and the DBN-DNN model, respectively.
646 Tables 2 and 3 illustrate that the proposed algorithm segments wide and erratic tumours
647 (ranging from a few millimeters wide to several centimeters), asymmetrical tumour and liver
648 shapes, complex tissues, and various paradigms where the lesion is in the vicinity of tissues with
649 analogous densities and global distribution inside the liver. In most of the segmentation models
650 that merely rely on measuring the energy, shape, density, location, illumination, and entropy could
651 fail when the tumours, liver, and other touching organs have a similar intensity and solidity. Under
652 such circumstances, applying more distinguishable features from a different kind of images can
653 improve the segmentation process and played a major role in recognizing different areas associated
654 with the above-mentioned difficulties. The two-path convolutional neural network could
655 potentially be more advantageous when analyzing dissimilar CT images with wide-ranging liver
656 sizes and blurred tumour or liver boundaries. The procedure proposed herein provides an improved
657 classification in terms of simplicity, time consumption, and segmentation accuracy as compared
658 to previously proposed approaches.

659 **4. Conclusions**

660 In this paper, a novel and robust architecture was proposed that incorporates the three input
661 images rather than one CT image, to automatically recognize the border of the liver and tumours
662 in abdominal CT images. This approach first applied a normalization method (Z-Score) to obtain
663 a more distinguishable liver border using the original image. Then, a new method (LDOG) was
664 implemented for encoding images to extract a more significant image. Finally, using the original
665 image and two other mentioned images, a new two-path CNN structure was trained.
666 The suggested novel architecture for liver and tumor segmentation was appraised on a dataset
667 containing 1000 patient cases which include over 20,000 scan slice. Our significant findings infer
668 that the proposed complex pipeline gained the following: 1) indicated accurate segmentation result
669 when the liver tissue consists of the sharing extended border with ambiguous touching organs, 2)
670 was suitably robust as illustrated by the insignificant standard deviations for all investigation
671 measures, and 3) accomplished well in intricate cases with several different types of the liver
672 tumours, which had a shadow from fat or fibrous tissue, amoeboid shapes, an abnormally large
673 amount of scar tissue, and similar densities as the surrounding tissues.
674 Our structure addresses the problem of failing in recognition of a vague border with an aim

to provide better results. Moreover, the suggested approach does not need any additional parameter to feed into the software apart from the input image to produce the segmented regions. Also, our finding represents that the border of tumours is always difficult to recognize on the east side of the liver, where there is a vague boundary due to the presence of the other similar intensity values organs. The proposed method illustrated that for obtaining a good segmentation result, it is essential to provide more representations of the image (using some encoding methods) before starting to extract key features from the image.

Acknowledgement

The author Malika Bendechache is supported, in part, by Science Foundation Ireland (SFI) under the grants No. 13/RC/2094_P2 (Lero) and 13/RC/2106_P2 (ADAPT).

References

Affonso, C., Rossi, A.L.D., Vieira, F.H.A. & de Carvalho, A.C.P. de L.F. 2017. Deep learning for biological image classification. *Expert Systems with Applications* 85: 114–122.

Ahmad, M., Ai, D., Xie, G., Qadri, S.F., Song, H., Huang, Y., Wang, Y. & Yang, J. 2019. Deep Belief Network Modeling for Automatic Liver Segmentation. *IEEE Access* 7: 20585–20595.

de Assis Neto, S.R., Santos, G.L., da Silva Rocha, E., Bendechache, M., Rosati, P., Lynn, T. & Takako Endo, P. 2020. Detecting Human Activities Based on a Multimodal Sensor Data Set Using a Bidirectional Long Short-Term Memory Model: A Case Study. *Studies in Systems, Decision and Control*, hlm. 31–51. Springer.

Baâzaoui, A., Barhoumi, W., Ahmed, A. & Zagrouba, E. 2017. Semi-Automated Segmentation of Single and Multiple Tumors in Liver CT Images Using Entropy-Based Fuzzy Region Growing. *IRBM* 38(2): 98–108.

Bae, Y.-K., Lee, J.-W. & Hong, S. 2020. Effects of image distortion and Hounsfield unit variations on radiation treatment plans: An extended field-of-view reconstruction in a large bore CT scanner.

Bakas, S., Chatzimichail, K., Hunter, G., Labbé, B., Sidhu, P.S. & Makris, D. 2017. Fast semi-automatic segmentation of focal liver lesions in contrast-enhanced ultrasound, based on a

probabilistic model. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization* 5(5): 329–338.

Bengio, Y. 2012. Practical Recommendations for Gradient-Based Training of Deep Architectures. hlm. 437–478. Springer, Berlin, Heidelberg.

Budak, Ü., Guo, Y., Tanyildizi, E. & Şengür, A. 2020. Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation. *Medical Hypotheses* 134: 109431.

Cai, J. 2019. Segmentation and Diagnosis of Liver Carcinoma Based on Adaptive Scale-Kernel Fuzzy Clustering Model for CT Images. *Journal of Medical Systems* 43(11): 1–11.

Di Cataldo, S. & Ficarra, E. 2017. Mining textural knowledge in biological images: Applications, methods and trends. Elsevier B.V.

Cecotti, H., Rivera, A., Farhadloo, M. & Pedroza, M.A. 2020. Grape detection with convolutional neural networks. *Expert Systems with Applications* 159: 113588.

Chartrand, G., Cresson, T., Chav, R., Gotra, A., Tang, A. & De Guise, J.A. 2017. Liver Segmentation on CT and MR Using Laplacian Mesh Optimization. *IEEE Transactions on Biomedical Engineering* 64(9): 2110–2121.

Chen, J., Liu, Z., Wang, H., Nunez, A. & Han, Z. 2018. Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. *IEEE Transactions on Instrumentation and Measurement* 67(2): 257–269.

Chen, X. & Pan, L. 2018. A Survey of Graph Cuts/Graph Search Based Medical Image Segmentation. *IEEE Reviews in Biomedical Engineering* 11: 112–124.

Dang, L.M., Hassan, S.I., Im, S. & Moon, H. 2019. Face image manipulation detection based on a convolutional neural network. *Expert Systems with Applications* 129: 156–168.

Doğantekin, A., Özyurt, F., Avcı, E. & Koç, M. 2019. A novel approach for liver image classification: PH-C-ELM. *Measurement: Journal of the International Measurement Confederation* 137: 332–338.

Dolz, J., Desrosiers, C. & Ben Ayed, I. 2018. 3D fully convolutional networks for subcortical

segmentation in MRI: A large-scale study. *NeuroImage* 170: 456–470.

Dvornik, N., Mairal, J. & Schmid, C. 2019. On the Importance of Visual Context for Data Augmentation in Scene Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Efimova, V., Shalamov, V. & Filchenkov, A. 2020. Synthetic dataset generation for text recognition with generative adversarial networks. *Twelfth International Conference on Machine Vision (ICMV 2019)*, hlm. 62. SPIE.

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. & Greenspan, H. 2018. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 321: 321–331.

Ghazouani, H. & Barhoumi, W. 2020. Genetic programming-based learning of texture classification descriptors from Local Edge Signature. *Expert Systems with Applications* 161: 113667.

Golrizkhatami, Z. & Acan, A. 2018. ECG classification using three-level fusion of different feature descriptors. *Expert Systems with Applications* 114: 54–64.

Guo, X., Schwartz, L.H. & Zhao, B. 2019. Automatic liver segmentation by integrating fully convolutional networks into active contour models. *Medical Physics* 46(10): 4455–4469.

He, T., Huang, W., Qiao, Y. & Yao, J. 2016. Text-Attentional Convolutional Neural Network for Scene Text Detection. *IEEE Transactions on Image Processing* 25(6): 2529–2541.

Heimann, T., Van Ginneken, B., Styner, M.A., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., Bello, F., Binnig, G., Bischof, H., Bornik, A., Cashman, P.M.M., Chi, Y., Córdova, A., Dawant, B.M., Fidrich, M., Furst, J.D., Furukawa, D., Grenacher, L., Hornegger, J., Kainmüller, D., Kitney, R.I., Kobatake, H., Lamecker, H., Lange, T., Lee, J., Lennon, B., Li, R., Li, S., Meinzer, H.P., Németh, G., Raicu, D.S., Rau, A.M., Van Rikxoort, E.M., Rousson, M., Ruskó, L., Saddi, K.A., Schmidt, G., Seghers, D., Shimizu, A., Slagmolen, P., Sorantin, E., Soza, G., Susomboon, R., Waite, J.M., Wimmer, A. & Wolf, I. 2009. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Transactions on Medical Imaging* 28(8): 1251–1265.

Hoogi, A., Beaulieu, C.F., Cunha, G.M., Heba, E., Sirlin, C.B., Napel, S. & Rubin, D.L. 2017. Adaptive local window for level set segmentation of CT and MRI liver lesions. *Medical Image Analysis* 37: 46–55.

Hu, P., Wu, F., Peng, J., Liang, P. & Kong, D. 2016. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. *Physics in Medicine and Biology* 61(24): 8676–8698.

Huang, Q., Ding, H., Wang, X. & Wang, G. 2018. Fully automatic liver segmentation in CT images using modified graph cuts and feature detection. *Computers in Biology and Medicine* 95: 198–208.

Husain, F., Dellen, B. & Torras, C. 2017. Scene Understanding Using Deep Learning. *Handbook of Neural Computation*, hlm. 373–382. Elsevier Inc.

Jain, S., Shukla, S. & Wadhvani, R. 2018. Dynamic selection of normalization techniques using data complexity measures. *Expert Systems with Applications* 106: 252–262.

Karimi, N., Ranjbarzadeh Kondrood, R. & Alizadeh, T. 2017. An intelligent system for quality measurement of Golden Bleached raisins using two comparative machine learning algorithms. *Measurement: Journal of the International Measurement Confederation* 107: 68–76.

Kavur, A.E., Gezer, N.S., Barış, M., Şahin, Y., Özkan, S., Baydar, B., Yüksel, U., Kılıkçıer, Ç., Olut, Ş., Akar, G.B., Ünal, G., Dicle, O. & Selver, M.A. 2020. Comparison of semi-automatic and deep learning-based automatic methods for liver segmentation in living liver transplant donors. *Diagnostic and Interventional Radiology* 26(1): 11–21.

Ker, C.G., Kuo, K.K., Chang, W.T., Chen, J.S., Lee, K. Ter, Yang, S.F., Wu, C.C. & Chai, C.Y. 2011. Clinical significance of hepatic cancer stem cells. No longer published by Elsevier.

Khatami, A., Babaie, M., Tizhoosh, H.R., Khosravi, A., Nguyen, T. & Nahavandi, S. 2018. A sequential search-space shrinking using CNN transfer learning and a Radon projection pool for medical image retrieval. *Expert Systems with Applications* 100: 224–233.

Li, C., Wang, X., Eberl, S., Fulham, M., Yin, Y., Chen, J. & Feng, D.D. 2013. A likelihood and local constraint level set model for liver tumor segmentation from CT volumes. *IEEE*

Transactions on Biomedical Engineering 60(10): 2967–2977.

Li, J., Sang, N. & Gao, C. 2016. LEDTD: Local edge direction and texture descriptor for face recognition. *Signal Processing: Image Communication* 41: 40–45.

LI, S., TSO, G.K.F. & HE, K. 2020. Bottleneck feature supervised U-Net for pixel-wise liver and tumor segmentation. *Expert Systems with Applications* 145: 113131.

Li, Y., Zhao, Y. qian, Zhang, F., Liao, M., Yu, L. li, Chen, B. fan & Wang, Y. jin. 2020. Liver segmentation from abdominal CT volumes based on level set and sparse shape composition. *Computer Methods and Programs in Biomedicine* 195: 105533.

Liao, M., Zhao, Y. qian, Liu, X. yao, Zeng, Y. zhan, Zou, B. ji, Wang, X. fang & Shih, F.Y. 2017. Automatic liver segmentation from abdominal CT volumes using graph cuts and border marching. *Computer Methods and Programs in Biomedicine* 143: 1–12.

Liao, M., Zhao, Y. qian, Wang, W., Zeng, Y. zhan, Yang, Q., Shih, F.Y. & Zou, B. ji. 2016. Efficient liver segmentation in CT images based on graph cuts and bottleneck detection. *Physica Medica* 32(11): 1383–1396.

Liu, L., Lao, S., Fieguth, P.W., Guo, Y., Wang, X. & Pietikäinen, M. 2016. Median Robust Extended Local Binary Pattern for Texture Classification. *IEEE Transactions on Image Processing* 25(3): 1368–1381.

Liu, Z., Song, Y.Q., Sheng, V.S., Wang, L., Jiang, R., Zhang, X. & Yuan, D. 2019. Liver CT sequence segmentation based with improved U-Net and graph cut. *Expert Systems with Applications* 126: 54–63.

Lu, F., Wu, F., Hu, P., Peng, Z. & Kong, D. 2017. Automatic 3D liver location and segmentation via convolutional neural network and graph cut. *International Journal of Computer Assisted Radiology and Surgery* 12(2): 171–182.

Lu, X., Wu, J., Ren, X., Zhang, B. & Li, Y. 2014. The study and application of the improved region growing algorithm for liver segmentation. *Optik* 125(9): 2142–2147.

Luo, S., Li, X. & Li, J. 2014. Review on the Methods of Automatic Liver Segmentation from Abdominal Images. *Journal of Computer and Communications* 02(02): 1–7.

Luo, Y.T., Zhao, L.Y., Zhang, B., Jia, W., Xue, F., Lu, J.T., Zhu, Y.H. & Xu, B.Q. 2016. Local line directional pattern for palmprint recognition. *Pattern Recognition* 50: 26–44.

Mahmood, A., Bennamoun, M., An, S., Sohel, F., Boussaid, F., Hovey, R., Kendrick, G. & Fisher, R.B. 2017. Deep Learning for Coral Classification. *Handbook of Neural Computation*, hlm. 383–401. Elsevier Inc.

Morabito, F.C., Campolo, M., Ieracitano, C. & Mammone, N. 2018. Deep learning approaches to electrophysiological multivariate time-series analysis. *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, hlm. 219–243. Elsevier.

Özyurt, F., Tuncer, T., Avci, E., Koç, M. & Serhatlioğlu, İ. 2019. A Novel Liver Image Classification Method Using Perceptual Hash-Based Convolutional Neural Network. *Arabian Journal for Science and Engineering* 44(4): 3173–3182.

Petersen, N.C., Rodrigues, F. & Pereira, F.C. 2019. Multi-output bus travel time prediction with convolutional LSTM neural network. *Expert Systems with Applications* 120: 426–435.

Pourasad, Y., Ranjbarzadeh, R. & Mardani, A. 2021. A New Algorithm for Digital Image Encryption Based on Chaos Theory. *Entropy* 23(3): 341.

Punarselvam, E. & Suresh, P. 2019. Non-Linear Filtering Technique Used for Testing the Human Lumbar Spine FEA Model. *Journal of Medical Systems* 43(2): 1–13.

R., S.P. & Chandra, C.M. 2016. Dimensionality reduced local directional pattern (DR-LDP) for face recognition. *Expert Systems with Applications* 63: 66–73.

Rakshit, R.D., Nath, S.C. & Kisku, D.R. 2018. Face identification using some novel local descriptors under the influence of facial complexities. *Expert Systems with Applications* 92: 82–94.

Ranjbarzadeh, R. & Baseri Saadi, S. 2020. Corrigendum to “Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering” (Measurement (2020) 150, (S0263224119309522), (10.1016/j.measurement.2019.107086)). *Measurement: Journal of the International Measurement Confederation* 151.

Ranjbarzadeh, R., Jafarzadeh Ghoushchi, S., Bendechache, M., Amirabadi, A., Ab Rahman, M.N., Baseri Saadi, S., Aghamohammadi, A. & Kooshki Forooshani, M. 2021. Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images. *BioMed Research International* 2021: 1–16.

Ranjbarzadeh, R. & Saadi, S.B. 2020. Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering. *Measurement: Journal of the International Measurement Confederation* 150.

Ranjbarzadeh, R., Saadi, S.B. & Amirabadi, A. 2020. LNPSS: SAR image despeckling based on local and non-local features using patch shape selection and edges linking. *Measurement: Journal of the International Measurement Confederation* 164.

Robitaille, N., Mouiha, A., Crépeault, B., Valdivia, F. & Duchesne, S. 2012. Tissue-based MRI intensity standardization: Application to multicentric datasets. *International Journal of Biomedical Imaging* 2012.

Salaken, S.M., Khosravi, A., Nguyen, T. & Nahavandi, S. 2019. Seeded transfer learning for regression problems with deep learning. *Expert Systems with Applications* 115: 565–577.

Sethi, G., Saini, B.S. & Singh, D. 2016. Segmentation of cancerous regions in liver using an edge-based and phase congruent region enhancement method. *Computers and Electrical Engineering* 53: 244–262.

Shen, S., Han, S.X., Aberle, D.R., Bui, A.A. & Hsu, W. 2019. An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. *Expert Systems with Applications* 128: 84–95.

Singh, V.K., Rashwan, H.A., Romani, S., Akram, F., Pandey, N., Sarker, M.M.K., Saleh, A., Arenas, M., Arquez, M., Puig, D. & Torrents-Barrena, J. 2020. Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. *Expert Systems with Applications* 139: 112855.

Song, W., Zhang, Y., Liu, F., Chai, Z., Ding, F., Qian, X. & Park, S.C. 2018. Taking advantage of multi-regions-based diagonal texture structure descriptor for image retrieval. *Expert Systems with Applications* 96: 347–357.

868 Sotoodeh, M., Moosavi, M.R. & Boostani, R. 2019. A novel adaptive LBP-based descriptor for
869 color image retrieval. *Expert Systems with Applications* 127: 342–352.

870 Srivastava, N., Hinton, G., Krizhevsky, A., Ilya, S. & Salakhutdinov, R. 2014. Dropout: A
871 Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning*
872 *Research 15* 15(1): 1929–1958.

873 Sun, C., Guo, S., Zhang, H., Li, J., Chen, M., Ma, S., Jin, L., Liu, X., Li, X. & Qian, X. 2017.
874 Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images
875 based on FCNs. *Artificial Intelligence in Medicine* 83: 58–66.

876 Tang, W., Zou, D., Yang, S., Shi, J., Dan, J. & Song, G. 2020. A two-stage approach for
877 automatic liver segmentation with Faster R-CNN and DeepLab. *Neural Computing and*
878 *Applications* 32(11): 6769–6778.

879 Ting, F.F., Tan, Y.J. & Sim, K.S. 2019. Convolutional neural network improvement for breast
880 cancer classification. *Expert Systems with Applications* 120: 103–115.

881 Torres, A.D., Yan, H., Aboutalebi, A.H., Das, A., Duan, L. & Rad, P. 2018. Patient facial
882 emotion recognition and sentiment analysis using secure cloud with hardware acceleration.
883 *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering*
884 *Applications*, hlm. 61–89. Elsevier.

885 Uddin, M.Z., Hassan, M.M., Almogren, A., Zuair, M., Fortino, G. & Torresen, J. 2017. A facial
886 expression recognition system using robust face features from depth videos and deep
887 learning. *Computers and Electrical Engineering* 63: 114–125.

888 Ullah, I., Hussain, M., Qazi, E. ul H. & Aboalsamh, H. 2018. An automated system for epilepsy
889 detection using EEG brain signals based on deep learning approach. *Expert Systems with*
890 *Applications* 107: 61–71.

891 Vinícius dos Santos Ferreira, M., Oseas de Carvalho Filho, A., Dalília de Sousa, A., Corrêa
892 Silva, A. & Gattass, M. 2018. Convolutional neural network and texture descriptor-based
893 automatic detection and diagnosis of glaucoma. *Expert Systems with Applications* 110: 250–
894 263.

895 Wahab, N., Khan, A. & Lee, Y.S. 2017. Two-phase deep convolutional neural network for

reducing class skewness in histopathological images based breast cancer detection. *Computers in Biology and Medicine* 85: 86–97.

Wang, X., Zheng, Y., Gan, L., Wang, X., Sang, X., Kong, X. & Zhao, J. 2017. Liver segmentation from CT images using a sparse priori statistical shape model (SP-SSM). *PLOS ONE* 12(10): e0185249.

Willner, M., Fior, G., Marschner, M., Birnbacher, L., Schock, J., Braun, C., Fingerle, A.A., Noël, P.B., Rummeny, E.J., Pfeiffer, F. & Herzen, J. 2015. Phase-Contrast Hounsfield Units of Fixated and Non-Fixated Soft-Tissue Samples. *PLOS ONE* 10(8): e0137016.

Wu, S., Tu, R., Nan, R., Liu, G., Cui, X. & Liang, X. 2015. Impact of variations in fatty liver on sonographic detection of focal hepatic lesions originally identified by CT. *Ultrasonography* 35(1): 39–46.

Xi, X.F., Wang, L., Sheng, V.S., Cui, Z., Fu, B. & Hu, F. 2020. Cascade U-ResNets for Simultaneous Liver and Lesion Segmentation. *IEEE Access* 8: 68944–68952.

Xu, L., Zhu, Y., Zhang, Y. & Yang, H. 2020. Liver segmentation based on region growing and level set active contour model with new signed pressure force function. *Optik* 202: 163705.

Yamashita, T. & Wang, X.W. 2013. Cancer stem cells in the development of liver cancer. American Society for Clinical Investigation.

Yang, X., Yu, H.C., Choi, Y., Lee, W., Wang, B., Yang, J., Hwang, H., Kim, J.H., Song, J., Cho, B.H. & You, H. 2014. A hybrid semi-automatic method for liver segmentation based on level-set methods using multiple seed points. *Computer Methods and Programs in Biomedicine* 113(1): 69–79.

Yang, Z.F., Ho, D.W., Ng, M.N., Lau, C.K., Yu, W.C., Ngai, P., Chu, P.W.K., Lam, C.T., Poon, R.T.P. & Fan, S.T. 2008. Significance of CD90+ Cancer Stem Cells in Human Liver Cancer. *Cancer Cell* 13(2): 153–166.

Yin, W., Schütze, H., Xiang, B. & Zhou, B. 2016. ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs. *Transactions of the Association for Computational Linguistics* 4: 259–272.

Zeng, Y. zhan, Liao, S. hui, Tang, P., Zhao, Y. qian, Liao, M., Chen, Y. & Liang, Y. xiong. 2018. Automatic liver vessel segmentation using 3D region growing and hybrid active contour model. *Computers in Biology and Medicine* 97: 63–73.

Zhang, X., Wu, F. & Li, Z. 2020. Application of Convolutional Neural Network to Traditional Data. *Expert Systems with Applications*: 114185.

Zhong, J., Liu, Z., Han, Z., Han, Y. & Zhang, W. 2019. A CNN-Based Defect Inspection Method for Catenary Split Pins in High-Speed Railway. *IEEE Transactions on Instrumentation and Measurement* 68(8): 2849–2860.

Zhou, Z., Xue-Chang, Z., Si-Ming, Z., Hua-Fei, X. & Yue-Ding, S. 2018. Semi-automatic liver segmentation in CT images through intensity separation and region growing. *Procedia Computer Science*, hlm. 220–225. Elsevier B.V.