

Received December 29, 2018, accepted January 22, 2019, date of publication February 1, 2019, date of current version February 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2896961

Deep Belief Network Modeling for Automatic Liver Segmentation

MUBASHIR AHMAD^{ID}¹, DANNI AI¹, GUIWANG XIE¹, SYED FURQAN QADRI²,
HONG SONG², YONG HUANG¹, YONGTIAN WANG¹, AND JIAN YANG^{ID}¹

¹Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

²School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China

Corresponding author: Jian Yang (jyang@bit.edu.cn)

This work was supported in part by the National Key R&D Program of China under Grant 2017YFC0112000, and in part by the National Science Foundation Program of China under Grant 81627803 and Grant 61771056.

ABSTRACT The liver segmentation in CT scan images is a significant step toward the development of a quantitative biomarker for computer-aided diagnosis. In this paper, we propose an automatic feature learning algorithm based on the deep belief network (DBN) for liver segmentation. The proposed method was based on training by a DBN for unsupervised pretraining and supervised fine tuning. The whole method of pretraining and fine tuning is known as DBN-DNN. In traditional machine learning algorithms, the pixel-by-pixel learning is a time-consuming task; therefore, we use blocks as a basic unit for feature learning to identify the liver, which saves memory and computational time. An automatic active contour method is applied to refine the liver in post-processing. The experiments on test images show that the proposed algorithm obtained satisfactory results on healthy and pathological liver CT images. Our algorithm achieved 94.80% Dice similarity coefficient on mixed (healthy and pathological) images while 91.83% on pathological liver images, which is better than those of the state-of-the-art methods.

INDEX TERMS Liver segmentation, deep learning, deep belief network, restricted Boltzmann machine.

I. INTRODUCTION

Liver segmentation is an important step toward different kinds of clinical applications, such as segmentation of a tumor, transplantation, and other liver diagnoses [1]. However, given the large variability in the liver shape, similar intensity in nearby organs, tissue linkage, low contrast, and overlapping of different organs are major challenging problems. In the last few years, a variety of methods have been proposed for liver segmentation from CT images. A review of different methods can be found in the literature [2], [3]. Another comparative study with different techniques of liver segmentation based on a publicly available dataset can be found in the reference [4]. Several methods have been proposed for liver segmentation which are, region growing [5], level-set [6], thresholding [7], and graph cut-based methods [8], have been proposed. Preventing segmentation from boundary leaks and under segmentation are the major challenges for gray level methods where organs have

The associate editor coordinating the review of this manuscript and approving it for publication was Po Yang.

similar intensity. Automatic methods are usually initialized with a certain threshold and morphological operations to deal with these issues [9]. However, semi-automatic methods with minimal user interaction and initialization can obtain better results. A semi-automatic method in which blood vessels are used to segment the liver from CT images was proposed [10]. Peng *et al.* [11], introduced a method, which is a combination of surface smoothness, regional appearance, and combined intensity methods, to deal with heterogeneous background and fuzzy boundaries [12]. Although these methods have promising results, user initialization and selection are the main drawbacks. Model-based methods, in which the typical shape of the liver is used to carry out the segmentation process, are robust. Some current model-based methods that rely on the statistical shape model (SSM) have been reported in the literature [13]. The small datasets are a real problem in SSM-based methods because of the large variation in the shape of a liver where each slice is different from the previous one; therefore, deformable models are used with a combination of SSM to overcome this problem [14] or merged in level-set method [15] and SSC [16] were used to design a better

model that can deal with complex variations in the liver shape. A local composition-based SSC method is proposed [17] to build a repository at a local level for each region and blood vessel-based initialization of a liver that can make the initial shape patient specific. Atlas-based methods [10], [18] are initiated by registering the single or multiple atlases with a target image and deforming the labelled image of the atlas by fusing the labels. These methods are difficult due to the large variation in the shape of the liver that heavily depends on the registration process. The complexity of atlas selection and computation was addressed in recent works [19].

In recent years, progress on deep-learning methods has been made in the field of image processing. Significant advances in clinical effects and medical imaging were acknowledged in [20]. There are three main primitives or building blocks of deep learning. These are Convolutional Neural Network (CNN), Auto Encoder (AE) and Deep Belief Network (DBN). These deep learning architectures already have been applied in many fields like bio-informatics, speech and audio recognition and computer vision [21]. An important benefit of various deep learning algorithms is to deal with unlabeled data. The most common architecture that can be trained in unsupervised fashion are Deep Belief Network (DBN) and Auto encoder (AE) [22]. DBN [29] model is a deep learning model that has achieved a popularity as a more successful implementation of a proficient learning method that stacks a simpler models which is known as Restricted Boltzmann Machine (RBM) [29]. A deep convolutional neural network is used for different tasks, such as image classification [23] and visual object recognition [24]. A study based on a deep belief network (DBN) is proposed in the literature for unsupervised feature learning, and the network was fine tuned to segment vertebrae from CT images [25], [43]. Skin segmentation was performed using a stacked autoencoder that represents high-level features of an image [26]. A stacked sparse autoencoder was proposed for unsupervised feature learning for pedestrian gender classification, and outstanding results were obtained [27]. Recent work on liver segmentation is presented in [28] where unsupervised features learned from stacked autoencoder and classification is performed to obtain the initial probability map. The background was removed by using a post-processing method. This work performed well on normal liver images, but unsatisfactory results were obtained on abnormal liver images.

In this work, we proposed a DBN-DNN model where training is completed in two steps: DBN unsupervised pretraining and supervised fine tuning. Pretraining relates to a highly proficient learning technique that stacks RBMs, which are individually trained layer by layer. The RBM parameters can be adjusted; for example, probability distribution can be made to fit the distribution of the training data. The pretraining of a DBN is unsupervised. The second step is supervised fine-tuning, in which all the layers, together with backpropagation neural network, are fine tuned to perform the classification task; this process is known as DBN-DNN. The sigmoid layer

at the end is composed of two units: background of a liver and foreground. We designed a DBN-DNN network with a limited number of layers and hidden nodes within each layer. This methodology reduced the training time and improved the overall accuracy. For noise removal, we utilized a fully automatic Chen–Vase (CV)-based 3D active contour method [30]. Next section demonstrates the methodology of our proposed framework.

II. METHODOLOGY

The framework of liver segmentation in CT images involves training and testing steps. In the training step, DBN-DNN is trained as an initial liver detector. In test images, the DBN-DNN model is used to detect the liver from CT scan image, and the volumetric CT image is refined by using CV based 3D Active contour method.

A. DATA PREPROCESSING

Preprocessing is a vital part of image segmentation, in which raw CT images are converted into a processed form to distinguish the liver features from the other abdominal parts of the human body. The Hounsfield units for each organ are different from those in the liver. A preprocessing step is manipulated in slice by slice manner and the Hounsfield unit is applied in window range $[-100, 400]$ to exclude irrelevant organs from CT images. Then, we enhanced the contrast of images by using histogram equalization and normalized by using zero mean and unit variance. To remove noise, we used a Gaussian filter, which can smoothen the image, and we utilized a sigma value of 0.2. The preprocessing step is given in Fig. 1.

We cropped the training images at a level that will not affect the liver and rotated the dataset according to our problem. We distributed each image into blocks; these blocks were used as input for our experiment. In the test images, we performed the same preprocessing steps, except cropping and rotation. Thus, the block was considered a basic unit for our experiment. Positive training blocks referred to the area or block from the liver, and we treated this as a foreground. However, negative blocks were the samples obtained from the background, as explained in Fig. 2. After preprocessing, we processed the datasets for further experimentation. We trained our DBN-DNN network on the balanced class of features concerning its background and foreground. We extracted 2,099,712 overlapping blocks from our training dataset, where half of the block represents the background and the remaining half belongs to the foreground. The size of each block is 23×23 pixels.

B. TRAINING THE DBN-DNN MODEL

DBN [29] network is the stacked network of multiple restricted Boltzmann machines (RBMs). Each RBM [29] consists of a hidden layer and a visible layer. The visible layer obtains the input data and transfers that data to the hidden layer. The state of the visible unit v and hidden unit h of RBM



FIGURE 1. The raw CT image is given as an input to the system. Fuzzy boundaries (left). Applied Hounsfield Unit window range [−100, 400] for liver (middle). Finally enhanced the contrast of an image (Right).

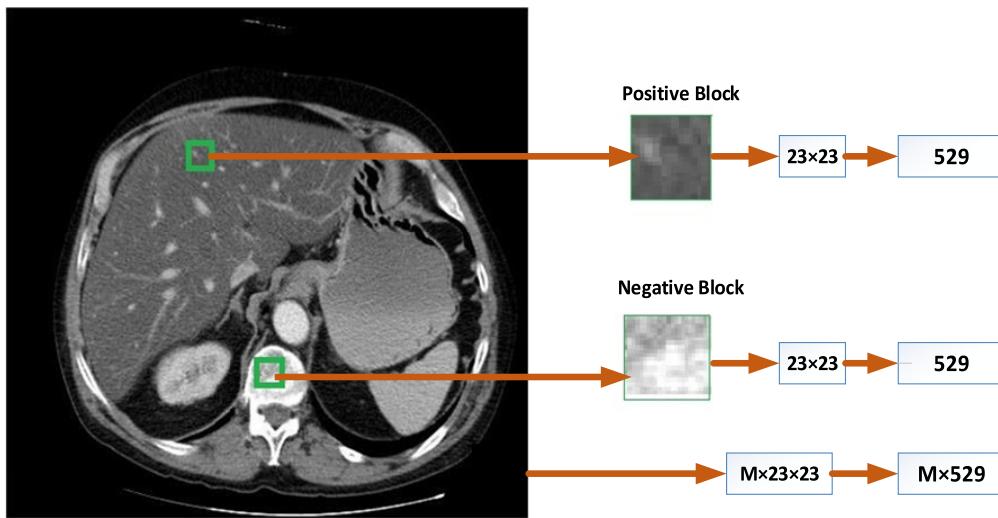


FIGURE 2. The positive block is a part of the liver where more pixels belongs to the liver. The negative block is the background area which is not a part of a liver. The total number of blocks in each image is M.

are represented by the following equations

$$v = (v_1, v_2, v_3, \dots, v_n)^T \quad (v_i \in \{0, 1\}) \quad (1)$$

$$h = (h_1, h_2, h_3, \dots, h_n)^T \quad (h_j \in \{0, 1\}) \quad (2)$$

The energy function of RBM is given by the following equation

$$E(v, h | \theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (3)$$

where n and m are the hidden and visible units, a_i and b_j represents the biases and w_{ij} is the weight between visible and hidden unit. The joint probability of visible unit v and hidden unit h is given by

$$p(v) = \frac{1}{z(\theta)} \sum_h \exp(-E(v, h | \theta)) \quad (4)$$

where,

$$p(h) = \frac{1}{z(\theta)} \sum_v \exp(-E(v, h | \theta)) \quad (5)$$

The visible and hidden units are autonomous, so the conditional probabilities of them are given in equation 6 and 7 respectively.

$$p(v|h) = \prod_i P(v_i|h) \quad (6)$$

$$p(h|v) = \prod_j P(h_j|v) \quad (7)$$

So there are no visible-visible, and hidden-hidden connections, the distribution of conditional probability of units represented in the following equation.

$$P(v_i = 1 | h) = \frac{1}{1 + \exp(-a_i - \sum_{j=1}^m w_{ij} h_j)} \quad (8)$$

$$P(h_j = 1 | v) = \frac{1}{1 + \exp(-b_j - \sum_{i=1}^n v_i w_{ij})} \quad (9)$$

The training process of RBM is described as follows. The visible unit is responsible for accepting the training samples to produce v_i and the hidden unit h_j is sampled similar to Eq. 8.

To repeat this process, the hidden and visible units are updated to produce the reconstructed states h'_j and v'_i .

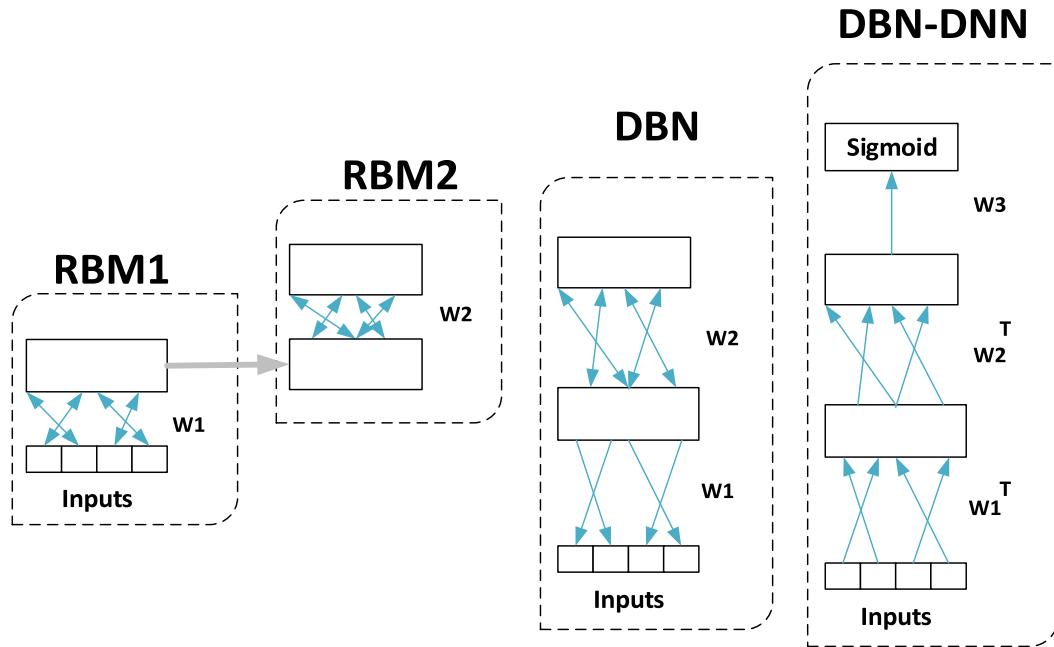


FIGURE 3. Pre-training of a deep neural network with DBN-DNN. Steps to create a DBN using greedy layer-wise pre-training and the whole network convert it into DBN-DNN. First Restricted Boltzmann machine with weights W_1 represents its hidden binary unit. After this step weights are frozen and the first output is generated and this output gives an input to the next RBM and weights W_2 are optimized to represent the correlation between its visible units. All these weights are combined to get a DBN. Then all the layers undergo fine-tuning to build a DNN.

The equation updates for w_{ij} is given as

$$\Delta w_{ij} = \eta(\langle v_i h_i \rangle - \langle v'_i h'_i \rangle) \quad (10)$$

where $\eta \in (0, 1)$ is the learning rate, and $\langle \cdot \rangle$ denotes the mean over training data. The construction of DBN is based on RBM, where the output of the lower layer of RBM is used as input for the next RBM layer. The visible layer (input layer) and the first hidden layer build the first RBM (RBM1). The first and second hidden layers construct the second RBM (RBM2). The same process is repeated until the last hidden layer in the network. In our case, we only had two hidden layers in our network; hence, RBM1 and RBM2 are the units for construction of the DBN. The sigmoid layer is added to diagnose the liver in the CT images. Sigmoid is used as a binary classifier, and we refer to the whole method as DBN-DNN. When the unsupervised training is completed successfully layer by layer, the fine-tuning process is started and accomplished by using backpropagation neural networks. Supervised fine tuning is a method that further trains the network to reduce and recover the training error of the DBN. In unsupervised DBN training process, only one RBM is considered at a time. The backpropagation training process considers all the DBN layers concurrently. The training procedure of DBN-DNN model is given in Fig. 3.

C. POST-PROCESSING FOR LIVER SEGMENTATION

Initial segmentation is achieved through DBN-DNN-based method. Some limitations exist in the initial probability map, especially in the CT images that exhibit a tumor inside the

liver that leaves holes. These holes are filled by morphological operations. The median filter is used to smoothen the segmented liver. We obtained the initial results from the DBN-DNN model and then passed these findings to CV-based 3D active contour method because initial segmentation revealed slight noise in the beginning and ending slices of the CT image in some cases. To refine the liver, we constructed a CV based 3D active contour method. For this purpose, we automatically defined a 2D mask for initial seed slice to initialize the algorithm. We determined the liver slice with the most pixels in the whole image and set it as an initial seed point for this method. These methods have advantages of smooth boundaries and uncertain definition by the gradient. The second advantage is that it does not need smoothing of images, thereby proficiently processing images with noise. In our case, we applied this method to 3D image where data are noisy from initial segmentation.

III. EXPERIMENTATION

A. DATASETS

MICCAI-Silver07 and 3Dircadb01 CT datasets were used for the experiments. Silver07 is a publicly available dataset containing 20 training samples with labels. Each sample varies by the number of slices from 64 to 512 while slice thickness and pixel spacing varied from 0.5mm to 5.0mm and 0.54mm to 0.87mm respectively. This dataset is available on the organizer of MICCAI-Silver07 website [31]. The 3Dircadb01 dataset is also publicly available 20 CT images with target [32]. It is highly complex due to pathologies and

a large number of variations in it. In 3Dircadb01 dataset having 15 patients and 120 lesions which make the segmentation task very difficult. Slice thickness and pixel spacing varied from 1mm to 5.0mm and 0.55mm to 0.95mm respectively. Both datasets have the axial dimension of 512×512 . We selected 30 CT scan images for the training of the DBN-DNN model and 10 CT scan images for testing. These datasets have normal and abnormal (hemangioma, hepatoma, cysts, and so on) images. This work was performed in MATLAB 2017a with Intel Core i7, 3.60 GHz CPU, and 32 GB of RAM. Experimental testing of CT images consisted of normal and pathological, in which former images displayed different types of tumors and cavities.

B. PARAMETRIC SETTINGS

After preprocessing, we set up the DBN-DNN model for the experiment as follows: all biases and weights were initially set to zero. Different algorithms on several datasets show that random search experimentation is more efficient for the optimization of hyper-parameters [33]. In our DBN-DNN model, we performed random search experiments to identify the range of hyper-parameters and then selected its values randomly until we get the best performance. Random search found better network training and required less computational time. Based on hyper-parameter values, random experimentation is more efficient because all hyper-parameters are not equally important for tuning. During pretraining, the learning rate was set at 0.052 and momentum at 0.9 with 70 epochs to train RBM layers. Each RBM was trained with contrastive divergence in a layer-wise greedy manner. The input 23×23 block size was flattened to obtain 529 neurons, which was the input vector of our RBM1, and the total number of samples was M. Fig. 2 demonstrates the whole input criteria. To train the first RBM, the whole dataset $M \times 529$ was transformed through the first RBM, resulting in a new RBM of size $M \times 100$. The second RBM was trained on the output of the first RBM. The feedforward neural network was initialized where weights and biases were used in pretraining with three layers of sizes 100-42-2, and the last two neurons were the output of our model. The hidden layers were 100 and 42 units. In this study, our model was trained using a stochastic gradient descent (SGD) algorithm during pretraining. Fig. 4 shows the pretraining visualization results of the first RBM layer. We set up the learning rate at 0.0061, momentum at 0.9, and weight decay at 0.00005 using backpropagation. Weight decay was used to prevent the network from overfitting. Sigmoid was used as an activation function where the mini-batch size of 512 was used for training. The output of our model was the foreground and background of the liver in CT images. Our model took 5 hours to complete the pre-training and 43 hours took to complete the fine-tuning by the backpropagation algorithm. In total, our entire approach took 48 hours for the training. In this study, our DBN-DNN model was based on MATLAB library “Deep learning Toolbox.” [34]. From the weights in Fig. 4, we can see that RBM learned on a large degree blobs detectors

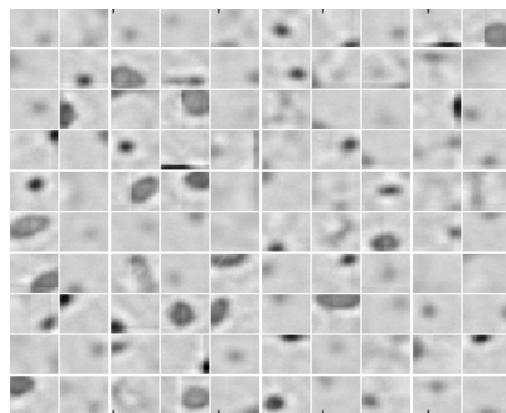


FIGURE 4. Visualization of the weights of RB.

and strokes. Some detectors are less meaningful where they either learned more or no learning at all.

We trained a backpropagation neural network where the criterion for error measurement adopted was mean squared error (MSE). During the experiment, our training error was 0.020 and the validation error was 0.022 with $10,000 = 10^4$ epochs; no convergence was recorded during training and validation. We trained our DBN-DNN models with the combination of different activation functions and output layers. Fig. 5 shows the training and validation errors of our model on different parameters. A learning rate of 0.0061 yielded the best training and validation results with sigmoid layer as a classifier. The most important finding in our model was that the low learning rate with sigmoid output provided robust results. The top right portion of Fig. 5-b shows a low error rate, which was helpful to efficiently segment the liver from the CT images. Figs. 5-c and 5-d, with the softmax output layer, exhibited a large difference in training and validation error, and no smoothness was observed. Initial segmentation was performed using the DBN-DNN model. After segmentation of each 2D CT scan slice, the 5×5 2D median filter was applied to smoothen the liver in each segmented 2D image.

C. LIVER REFINEMENT

We acquired all the 2D slices to stack them into 3D form and then applied CV-based 3D active contour method to refine segmentation. The selection of initial seed slice was automatic in our work. We obtained satisfactory results of segmentation from our DBN-DNN model. Therefore, in the beginning and ending slices, our algorithm mis-segmented the small parts of other organs, which created noise in volumetric CT images. These errors were automatically removed by CV-based 3D active contour method and considered tiny errors. The detection of initial seed selection for the active contour method was based on most liver pixel slices in the whole volumetric image.

The red circle in Fig. 6 (large liver slice) points out the initial seed slice of the 3Dircadb01 dataset where beginning

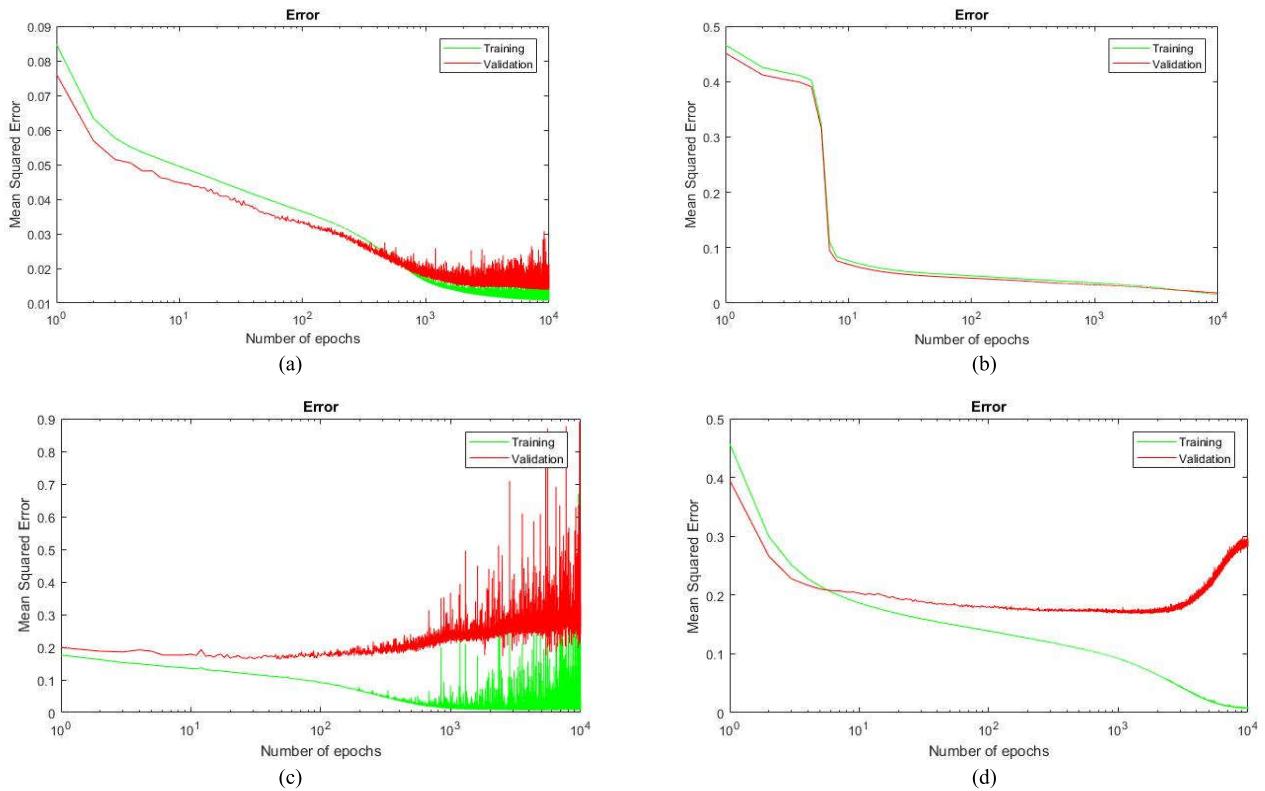


FIGURE 5. 5-a shows the mean squared error on training and validation datasets with sigmoid as an activation function and sigmoid (a binary classifier) as an output layer. The learning rate is 0.0061. Low training and validation errors. The best results of our model recorded on these parameters. 5-b: shows the mean squared error on training and validation datasets with sigmoid as an activation function and sigmoid (a binary classifier) as an output layer. The learning rate is 0.0061. Low training and validation errors. The best results of our model recorded on these parameters. 5-c: Bad training and validation. Shows the mean squared error on training and validation datasets with sigmoid as an activation function and softmax as an output layer. The learning rate is 0.061. Many variations in training and validation error on each epoch. 5-d: Bad validation. Shows the mean squared error on training and validation datasets with sigmoid as an activation function and softmax as an output layer. The learning rate is 0.0061.

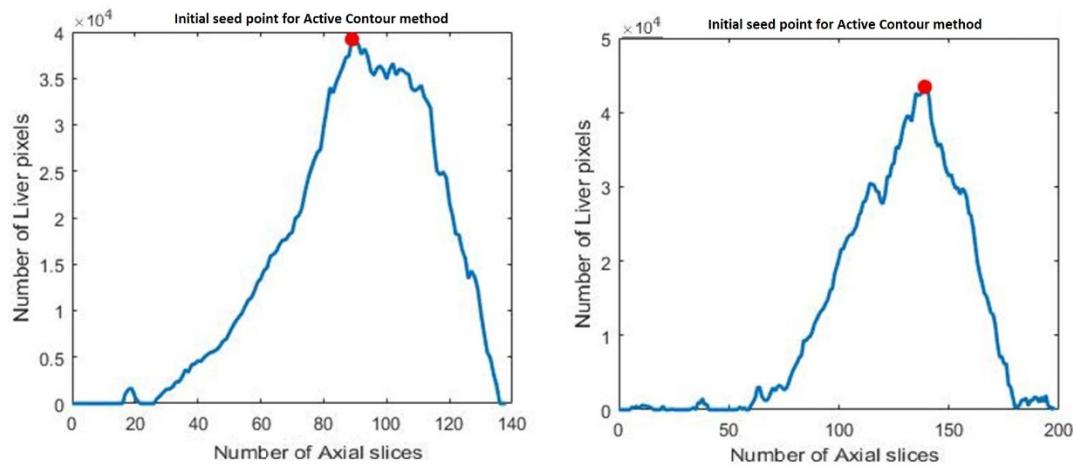


FIGURE 6. Automatic detection of the larger liver slice for initial seed point for active contour method (Two CT images from our testing dataset).

slices displayed noise in both the CT scan images; in second case (right), few ending slices showed noise. To detect the large liver in the CT scan image which is the initial seed slice for 3D active contour method. From the segmentation of liver with the DBN-DNN model, we obtained tiny errors

around the surface of the liver; this surface was not connected to the liver area where we applied the CV-based 3D active contour model to remove such errors. For this purpose, we initialized the active contour model automatically based on the initial seed slice with 300 iterations to run the algorithm.



FIGURE 7. Tinny error regions (misclassified regions) on the left. Refined the liver by CV-based 3D active contour method in right.

Fig. 7 shows the tiny error region in a 3D volumetric image, which was efficiently refined by our post-processing method.

IV. RESULTS AND DISCUSSION

Before we further examine our results, some common methods for statistical performance measurement are presented. True Positive (TP) denotes all the pixels related to the liver, True Negative (TN) shows all the pixels related to the background. False Negative (FN) is the liver pixels that does not classify accurately and false positive (FP) denotes the pixels which belong to the background but not classified as a background. Dice similarity coefficient (DSC) [35] represents the overall performance of the segmentation correctly where each pixel in the ROI is mapped individually with the original labels. The higher number represents the good performance. The following equation represents the DSC.

$$DSC(\%) = 100 \times \left(\frac{2TP}{2TP + FP + FN} \right) \quad (11)$$

Jaccard similarity coefficient (JSC) [36] is a common statistical measurement method to comparing the similarity of the two sets, higher value represents a best performance.

$$JSC(\%) = 100 \times \left(\frac{TP}{TP + FP + FN} \right) \quad (12)$$

Symmetric volume difference (SVD) provides a symmetric measure of the difference in segmented image and the ground truth image as calculated in [37].

$$SVD(\%) = 100 \times (1 - DSC) \quad (13)$$

Relative volume difference (RVD) is a statistical method used to measure the size between segmented image (S) and ground truth image (T), as calculated in [4]. The negative value of RVD represents under-segmentation and a positive value represents over-segmentation. S represents the segmentation and T represents the ground truth. RVD is given in the following equation.

$$RVD(\%) = 100 \times \left(\left(\frac{|S| - |T|}{|T|} \right) \right) \quad (14)$$

TABLE 1. Statistical performance of our approach in comparison to expert radiologist segmentation on MICCAI-Sliver07 dataset.

Testing Images	VOE%	RVD %	SVD %	JSC %	DSC %
#1	4.16	1.36	5.98	88.71	94.02
#2	5.40	1.76	4.60	91.20	95.40
#3	4.55	2.93	4.80	90.84	95.20
#4	4.18	1.79	5.70	89.21	94.30
#5	3.26	-1.40	4.90	90.66	95.10
Mean	4.31	1.28	5.19	90.12	94.80
SD	0.77	1.61	0.60	1.09	0.60

TABLE 2. Statistical performance of our approach in comparison to expert radiologist segmentation on 3Dircadb01 dataset.

Testing Images	VOE%	RVD %	SVD %	JSC %	DSC %
#1	4.10	1.30	6.97	86.97	93.03
#2	6.91	4.13	8.99	83.50	91.01
#3	5.96	-1.80	7.87	85.41	92.13
#4	7.20	10.30	6.88	87.13	93.12
#5	6.32	14.05	10.10	81.65	89.90
Mean	6.09	5.59	8.16	84.93	91.83
SD	1.21	6.49	1.37	2.34	1.37

The volumetric overlap error (VOE) is based on the division of intersection of segmented image (S) and ground truth (T) by the total number in the union. A 0 score means there is no error and 100 means there is no similarity between T and S.

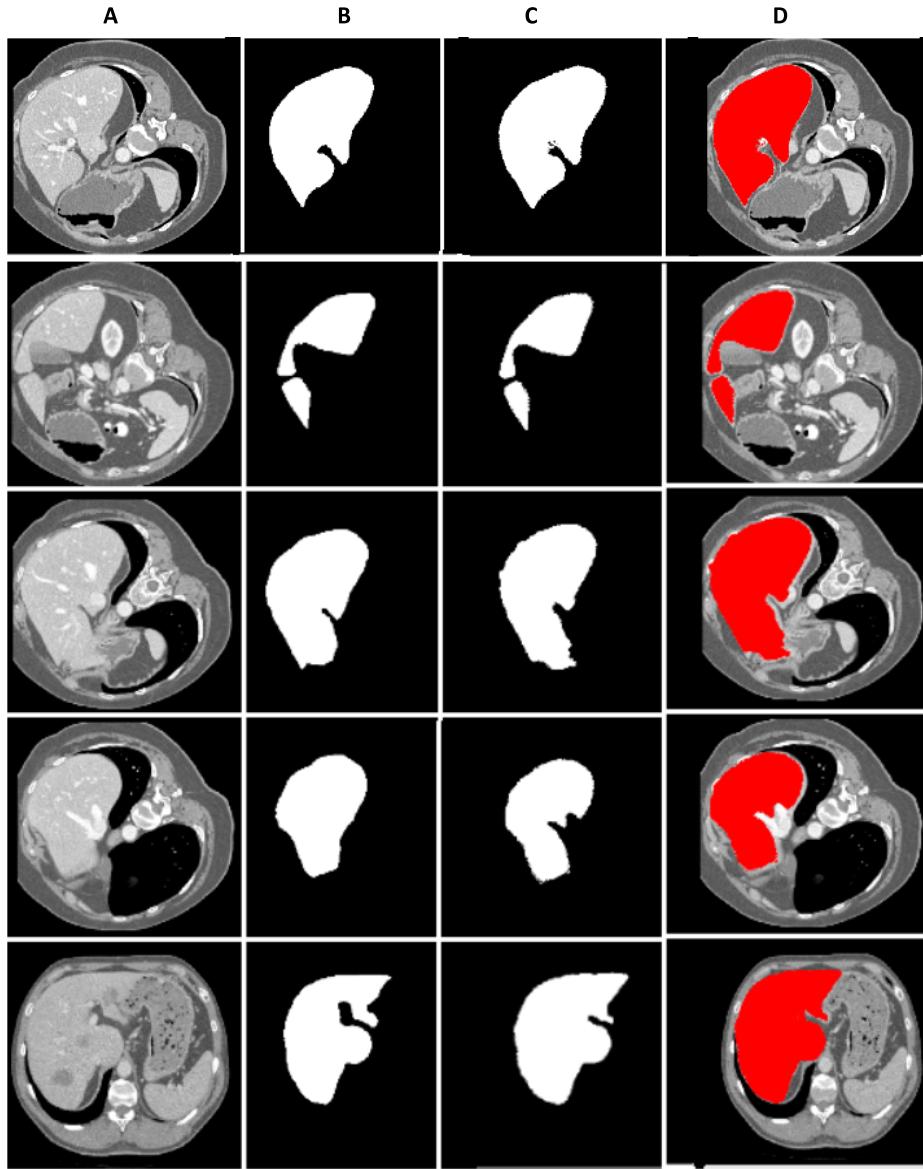


FIGURE 8. Axial slice segmentation view of proposed DBN-DNN model with different angles. (A) First Column is the contrast-enhanced CT image. (B) The second column is the original label of the image. (C) Third column is the segmentation of the liver by the proposed DBN-DNN method. (D) Fourth column is the overlapping of the segmented liver on CT image.

The following equation represents the VOE as in [38].

$$\text{VOE}(\%) = 100 \times \left(1 - \left(\frac{|S \cap T|}{|S \cup T|} \right) \right) \quad (15)$$

The results were presented using different statistical performance measurement methods in the following section. We categorized our testing datasets into two categories: pathological and mixed. Former datasets comprised normal and pathological CT images, whereas the pathological dataset only included abnormal images. Tables 1 and 2 represent different performance measurements of our DBN-DNN model on testing images as compared with the radiologist's segmentation. We selected 10 CT scan images for testing (5 from the MICCAI-Sliver07 training dataset and 5 from the 3dircadb01 dataset), which were not used to train the

DBN-DNN model. The segmentation results achieved from the DBN-DNN method are presented in Fig. 8. From a statistical performance viewpoint, our model was the most accurate segmentation method developed and tested on the liver that achieved comparable results with the semi-automatic and automatic methods.

In the case of the Sliver07 dataset, the proposed model obtained 94.80% DSC and VOE of 4.31. The average RVD, SVD, and Jaccard similarity coefficients were 1.28%, 5.19%, and 90.12%, respectively. VOE and DSC showed that our model was consistent on the SLiver07 dataset when we discussed both the overlapping errors. The standard deviations for VOE and DSC were 0.77 and 0.60, respectively. The standard deviations of RVD, SVD, and Jaccard similarity coefficients were 1.61, 0.60, and 1.09, respectively. The mean

TABLE 3. Comparative results of our method with some well-known methods.

Method	Dataset type	Mean VOE %	Mean RVD %	Mean DSC %
[39] Automatic	Pathological	--	2.78 ± 0.39	92
[40] Automatic	Pathological	5.95 ± 2.56	7.49 ± 9.63	91 ± 2
[41] Automatic	Mixed	6.1 ± 2.1	-2.9 ± 2.9	--
[15] Automatic	Mixed	6.44 ± 0.6	1.53 ± 1.7	--
[42] Automatic	Mixed	--	--	94 ± 3
[40] Automatic	Mixed	4.47 ± 1.7	2.38 ± 2.61	94 ± 1
Proposed method	Mixed	4.31	1.28	94.80 ± 1
Proposed method	Pathological	6.09	5.59	91.83 ± 1

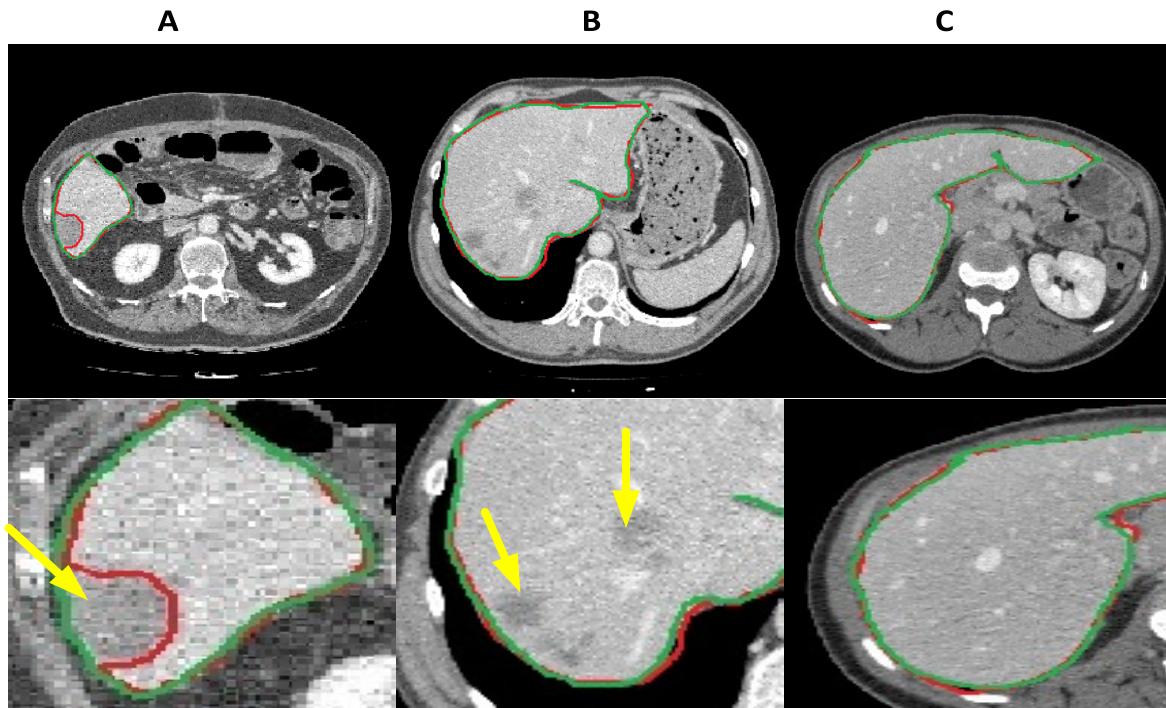


FIGURE 9. Green line represents the original label and red line represents our segmentation. The yellow arrow points the tumor in the liver. (A) Represents the image having a tumor on the boundary of the liver which is miss-segmented by our method. (B) Represents the CT 2D axial slice having the tumor inside the liver which is segmented efficiently. (C) Represents the CT 2D axial slice is a normal liver which segmented efficiently by our method.

Jaccard similarity coefficient was 90.12%. Both overlapping and size-based comparison methods achieved robust results on the SLiver07 dataset (Table 1).

Therefore, the 3Dircadb dataset is highly complex due to the different types of tumors within the images and on the boundary of the liver. The chances of mis-segmentation in this dataset are higher than those in SLiver07. The average DSC and VOE were 91.83% and 6.09%, respectively. The proposed algorithm was less sensitive against the tumors on the boundary of CT images. The standard deviations for VOE and DSC were 1.21% and 1.37%, respectively. RVD and SVD were 5.59% and 8.16%, respectively. The mean Jaccard similarity coefficient was 84.93%, which was low because of the irregularity of the 3Dircadb01 dataset. The standard

deviations of RVD, SVD, and Jaccard similarity coefficient were 6.49, 1.37, and 2.34, respectively. These observations revealed a large variation in each image of the 3Dircadb01 dataset. Table 2 presents the results of the 3Dircadb01 dataset.

We validated our model on mixed and pathological images, and we found that our model had limitations on pathological images where the tumor was on the boundary of the liver. On normal CT images, DBN-DNN performed well. The DSC of the proposed method on mixed images was high and comparable with sophisticated methods published recently. Statistical performance on Sliver07 CT images revealed a combination of pathological and normal CT images. The mean DSC was 94.80%, which was comparable with recently published methods [40], [42] in the literature. For this

observation, our model performed well, especially on normal CT images. The 3dircadb01 CT images constrained by lesions made automatic segmentation challenging. The proposed method obtained a DSC of 91.83%, whereas another method [39] obtained 92% on the pathological liver using the semi-automatic approach on the private dataset. Thus, our method was less sensitive to the pathological liver. RVD and VOE are important validation techniques to calculate and determine the liver volume than other segmentation methods. The mean RVD and VOE noted on the Silver'07 dataset were 1.28% and 4.31%, respectively, whereas those on the 3Dircadb01 dataset were 5.59% and 6.09%, respectively. Table 3 represents the comparative results with other well-known techniques.

The DBN-DNN model can detect the tumor region within the liver very well due to accurate feature learning, but it was less sensitive to the tumor on the boundary of a liver. Fig. 9 shows the segmentation results on normal and abnormal livers. The limitation of DBN-DNN model on pathological images at the liver boundary is due to the weak edges, low contrast and almost the same intensities of the background pixels that caused misclassification. Our model can't recognize the boundaries of the liver very well where a tumor lies. We deal with two class problems in our model, non-liver (background) and a liver (foreground). In our future work, we will try to solve this problem by using three classes, background, liver and one more class from the boundary region. The addition of third class can solve this limitation which can learn the blocks from boundaries and liver tumor on the boundary region can be easily discriminated.

In our future work, we will focus on the limitations that were identified in our work and use the newest python libraries (Pytorch, Tensorflow, etc.) with GPU support.

V. CONCLUSION

In this paper, we propose a 3D liver segmentation model using DBN. The algorithm can overcome problems caused by the tumor within the liver. DBN-DNN is used to segment the liver from CT scan images. We applied an automatic CV-based 3D active contour method to refine the liver, which improved the DSC and volumetric error. Without any human interaction, the algorithm learned the robust representation of the liver among other abdominal organs. The overall DSC of our model was 94.80% on mixed CT images and 91.83% on pathological CT images; these values are better than the comparing state-of-the-art methods.

REFERENCES

- [1] Y. Chen *et al.*, "Artifact suppressed dictionary learning for low-dose CT image processing," *IEEE Trans. Med. Imag.*, vol. 33, no. 12, pp. 2271–2292, Dec. 2014.
- [2] Y. Zhao, Y. Liu, X. Wu, S. P. Harding, and Y. Zheng, "Retinal vessel segmentation: An efficient graph cut approach with retinex and local phase," *PLoS one*, vol. 10, no. 4, 2015, Art. no. e0122332.
- [3] P. Campadelli, E. Casiraghi, and A. Esposito, "Liver segmentation from computed tomography scans: A survey and a new algorithm," *Artif. Intell. Med.*, vol. 45, nos. 2–3, pp. 185–196, 2009.
- [4] T. Heimann *et al.*, "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1251–1265, Aug. 2009.
- [5] L. Ruskó, G. Bekes, G. Németh, and M. Fidrich, "Fully automatic liver segmentation for contrast-enhanced CT images," in *Proc. MICCAI Workshop 3D Segmentation Clinic, A Grand Challenge*, vol. 2, 2007, pp. 143–150.
- [6] D. A. B. Oliveira, R. Q. Feitosa, and M. M. Correia, "Segmentation of liver, its vessels and lesions from CT images for surgical planning," *Biomed. Eng. Online*, vol. 10, p. 30, Apr. 2011.
- [7] K.-S. Seo, H.-B. Kim, T. Park, P.-K. Kim, and J.-A. Park, "Automatic liver segmentation of contrast enhanced CT images based on histogram processing," in *Proc. Int. Conf. Natural Comput.*, 2005, pp. 1027–1030.
- [8] A. Afifi and T. Nakaguchi, "Liver segmentation approach using graph cuts and iteratively estimated shape and intensity constrains," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2012, pp. 395–403.
- [9] X. Song, M. Cheng, B. Wang, S. Huang, X. Huang, and J. Yang, "Adaptive fast marching method for automatic liver segmentation from CT images," *Med. Phys.*, vol. 40, no. 9, 2013, Art. no. 091917.
- [10] M. G. Linguraru, J. K. Sandberg, Z. Li, F. Shah, and R. M. Summers, "Automated segmentation and quantification of liver and spleen from CT images using normalized probabilistic atlases and enhancement estimation," *Med. Phys.*, vol. 37, pp. 771–783, Feb. 2010.
- [11] J. Peng, F. Dong, Y. Chen, and D. Kong, "A region-appearance-based adaptive variational model for 3D liver segmentation," *Med. Phys.*, vol. 41, Apr. 2014, Art. no. 043502.
- [12] J. Peng, Y. Wang, and D. Kong, "Liver segmentation with constrained convex variational model," *Pattern Recognit. Lett.*, vol. 43, pp. 81–88, Jul. 2014.
- [13] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Med. Image Anal.*, vol. 13, no. 4, pp. 543–563, 2009.
- [14] H. Ling, S. K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and D. Comaniciu, "Hierarchical, learning-based automatic liver segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA, Jun. 2008, pp. 1–8.
- [15] S. D. S. Al-Shaihkh, M. Y. Yang, and B. Rosenhahn. (2015). "Automatic 3D liver segmentation using sparse representation of global and local image information via level set formulation." [Online]. Available: <https://arxiv.org/abs/1508.01521>
- [16] G. Wang, S. Zhang, H. Xie, D. N. Metaxas, and L. Gu, "A homotopy-based sparse representation for fast and accurate shape prior modeling in liver surgical planning," *Med. Image Anal.*, vol. 19, no. 1, pp. 176–186, 2015.
- [17] C. Shi, Y. Cheng, F. Liu, Y. Wang, J. Bai, and S. Tamura, "A hierarchical local region-based sparse shape composition for liver segmentation in CT scans," *Pattern Recognit.*, vol. 50, pp. 88–106, Feb. 2016.
- [18] Z. Xu *et al.*, "Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning," *Med. Image Anal.*, vol. 24, no. 1, pp. 18–27, 2015.
- [19] C. Platero and M. C. Tobar, "A multitalas segmentation using graph cuts with applications to liver segmentation in CT scans," *Comput. Math. Methods Med.*, vol. 2014, 2014, Art. no. 182909. doi: [10.1155/2014/182909](https://doi.org/10.1155/2014/182909).
- [20] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.
- [21] T. M. Hassan, M. Elmogy, and E.-S. Sallam, "Diagnosis of focal liver diseases based on deep learning technique for ultrasound images," *Arabian J. Sci. Eng.*, vol. 42, pp. 3127–3140, Aug. 2017.
- [22] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
- [23] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2013, pp. 411–418.
- [24] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2013, pp. 246–253.
- [25] S. F. Qadri, M. Ahmad, D. Ai, J. Yang, and Y. Wang, "Deep belief network based vertebra segmentation for CT images," in *Proc. Chin. Conf. Image Graph. Technol.*, 2018, pp. 536–545.
- [26] Y. Lei, W. Yuan, H. Wang, Y. Wenhui, and W. Bo, "A skin segmentation algorithm based on stacked autoencoders," *IEEE Trans. Multimedia*, vol. 19, no. 4, pp. 740–749, Apr. 2017.

- [27] M. Raza, M. Sharif, M. Yasmin, M. A. Khan, T. Saba, and S. L. Fernandes, "Appearance based pedestrians' gender recognition by employing stacked auto encoders in deep learning," *Future Gener. Comput. Syst.*, vol. 88, pp. 28–39, Nov. 2018.
- [28] M. Ahmad, J. Yang, D. Ai, S. F. Qadri, and Y. Wang, "Deep-stacked auto encoder for liver segmentation," in *Proc. Chin. Conf. Image Graph. Technol.*, 2017, pp. 243–251.
- [29] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [30] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [31] B. Van Ginneken, T. Heimann, and M. Styner, "3D segmentation in the clinic: A grand challenge," in *Proc. 10th Intel Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Brisbane, QLD, Australia, Oct./Nov. 2007, pp. 7–15.
- [32] (2016). *IRCAD*. [Online]. Available: <http://www.ircad.fr/research/3dircadb>
- [33] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, Feb. 2012.
- [34] R. B. Palm, "Prediction as a candidate for learning deep hierarchical models of data," Tech. Univ. Denmark, Kongens Lyngby, Denmark, 2012, vol. 5. [Online]. Available: <http://www.imm.dtu.dk/English.aspx>
- [35] L. R. Dice, "Measures of the amount of ecological association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [36] P. Jaccard, "The distribution of the flora in the alpine zone," *New Phytologist*, vol. 11, pp. 37–50, Feb. 1912.
- [37] V. Yeghiazaryan and I. Voiculescu, "An overview of current evaluation methods used in medical image segmentation," Dept. Comput. Sci., Univ. Oxford, Oxford, U.K., Tech. Rep. CS-RR-15-08, 2015.
- [38] Q. Zhang, Y. Fan, J. Wan, and Y. Liu, "An efficient and clinical-oriented 3D liver segmentation method," *IEEE Access*, vol. 5, pp. 18737–18744, 2017.
- [39] M. Goryawala, S. Gulec, R. Bhatt, A. J. McGoron, and M. Adjouadi, "A low-interaction automatic 3D liver segmentation method using computed tomography for selective internal radiation therapy," *BioMed Res. Int.*, vol. 2014, 2014, Art. no. 198015. doi: [10.1155/2014/198015](https://doi.org/10.1155/2014/198015).
- [40] M. Moghbel, S. Mashohor, R. Mahmud, and M. I. B. Saripan, "Automatic liver segmentation on computed tomography using random walkers for treatment planning," *EXCLI J.*, vol. 15, pp. 500–517, Aug. 2016. doi: [10.17179/excli2016-473](https://doi.org/10.17179/excli2016-473).
- [41] D. Kainmüller, T. Lange, and H. Lamiecker, "Shape constrained automatic segmentation of the liver based on a heuristic intensity model," in *Proc. MICCAI Workshop 3D Segmentation Clinic, A Grand Challenge*, 2007, pp. 109–116.
- [42] X. Wang, J. Yang, D. Ai, Y. Zheng, S. Tang, and Y. Wang, "Adaptive mesh expansion model (AMEM) for liver segmentation from CT image," *PLoS one*, vol. 10, no. 3, 2015, Art. no. e0118064.
- [43] S. F. Qadri et al., "Automatic deep feature learning via patch-based deep belief network for vertebrae segmentation in CT images," *Appl. Sci.*, vol. 9, no. 1, p. 69, 2019.



MUBASHIR AHMAD received the BCS degree from Hazara University, Mansehra, Pakistan, in 2007, and the MSCS degree from IQRA University, Islamabad, Pakistan, in 2012. He is currently pursuing the Ph.D. degree in optical engineering with the Beijing Institute of Technology, China. From 2012 to 2014, he was a Php Programmer with Microcom IT Ltd. His research interests include deep learning, pattern recognition, and medical imaging. He received the Excellent Research Paper Award at EI Conference, Beijing, China, in 2017. He is also a recipient of the Chinese Government Scholarship for his Ph.D. studies.



DANNI AI received the B.S. and M.S. degrees from Xi'an Jiaotong University, in 2005 and 2008, respectively, and the Ph.D. degree from Ritsumeikan University, Japan, in 2011. She is currently an Associate Professor with the School of Optics and Photonics, Beijing Institute of Technology. Her research interests include medical image analysis, surgical navigation, virtual reality, and augmented reality.



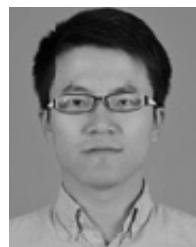
GUIWANG XIE received the B.S. degree in the measurement and control technology and instrument from the Beijing Institute of Technology, China, in 2016, where he is currently pursuing the master's degree with the School of Optics and Photonics. His research interests include medical image processing and augmented reality.



SYED FURQAN QADRI received the M.Sc. (CS) and M.S. (CS) degrees from the University of Agriculture Faisalabad, in 2011 and 2013, respectively. He is currently pursuing the Ph.D. degree in computer science and technology with the Beijing Institute of Technology, China. He has worked as a Lecturer with the CS Department, Government College University Faisalabad (Layyah Campus). His current research interest areas are medical Image processing, computer vision, deep learning, and information and communication technology.



HONG SONG received the Ph.D. degree in computer science from the Beijing Institute of Technology, China, in 2004, where she is currently an Associate Professor with the School of Software. She focuses her research interests on medical image processing, augmented reality, and computer vision.



YONG HUANG received the Ph.D. degree from the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA, in 2013. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, Johns Hopkins University, from 2013 to 2014. He is currently an Associate Professor with the School of Optics and Photonics, Beijing Institute of Technology, China. His research interests include biomedical imaging, image processing, and high-performance computation.



YONGTIAN WANG received the Ph.D. degree from the Department of Physics, University of Reading, U.K., in 1986. He has been with the Beijing Institute of Technology, since 1988. He is currently a Professor of both the School of Computer Science and School of Optoelectronics, Beijing Institute of Technology. His research interests include computer science, computer vision, medical image processing, and image-guided surgery.



JIAN YANG received the Ph.D. degree in optical engineering from the School of Optics and Photonics, Beijing Institute of Technology, Beijing, China, in 2007. He held a postdoctoral position with Children's Hospital Toronto, Canada, from 2007 to 2009. He has been with the Beijing Institute of Technology, since 2009, where he is currently a Full Professor with the School of Optics and Photonics. His research interest includes medical image processing, computer vision, and artificial intelligence.