Semantic segmentation and PSO based method for segmenting liver and lesion from CT images

P Vaidehi Nayantara, Surekha Kamath, Manjunath KN and Rajagopal Kadavigere

Abstract—The liver is a vital organ of the human body and hepatic cancer is one of the major causes of cancer deaths. Early and rapid diagnosis can reduce the mortality rate. It can be achieved through computerized cancer diagnosis and surgery planning systems. Segmentation plays a major role in these systems. This work evaluated the efficacy of the SegNet model in liver and particle swarm optimization-based clustering technique in liver lesion segmentation. Over 2400 CT images were used for training the deep learning network and ten CT datasets for validating the algorithm. The segmentation results were satisfactory. The values for Dice Coefficient and volumetric overlap error achieved were 0.940 ± 0.022 and 0.112 ± 0.038, respectively for liver and the results for lesion delineation were 0.4629 ± 0.287 and 0.6986 ± 0.203, respectively. The proposed method is effective for liver segmentation. However, lesion segmentation needs to be further improved for better accuracy.

Keywords—Liver lesion segmentation; Computed Tomography; Semantic segmentation; SegNet; Particle swarm optimization-based clustering; Hounsfield Unit

I. INTRODUCTION

The liver performs an array of vital functions like protein synthesis for blood plasma, bile production for digestion, blood purification, storage and release of nutrients, regulating blood clotting, etc [1]. Hence it is essential to keep the liver healthy and protect it from various diseases like fatty liver disease, cirrhosis, hepatitis and liver cancer. In this paper, the discussion is limited to liver cancer. Hepatocellular carcinoma (HCC) and metastatic liver cancer are the most commonly encountered liver cancers. Computer-aided detection and diagnosis systems (CAD) and computer-aided surgery are computer-aided systems that can assist the radiologist in the early and accurate diagnosis and treatment of liver cancer. Liver and liver lesion segmentations play a crucial role in the development of computer-aided systems. Former is essential for surgical resurrection and latter is the critical first step for accurate diagnosis of liver cancer in CAD systems. Over the years, Computed Tomography (CT) has been widely employed for diagnosing liver cancer due to its low cost, short acquisition time and ease of use. In this work, the liver and liver lesions are segmented from the portal venous phase of abdominal CT images using Deep Learning (DL) and Particle Swarm Optimization (PSO) based clustering techniques.

Liver lesions are segmented in many ways, one of the most trusted methods is manual contouring by experienced radiologists [2]–[4]. But it is not feasible when the number of cases to be diagnosed is very high. The standard methods used for segmentation are region growing, level set, active contour model and graph cut. Nayak et al. [5] performed liver segmentation using the region growing algorithm where the seed point was randomly selected from the Region Of Interest (ROI) defined by the user for the first slice. For the remaining slices, multiple seed points were randomly chosen from the segmented liver region of the previous slice. Out of the liver masks corresponding to the different seed points, the one with the highest dice coefficient with the liver mask of previous slice was retained. Other works with this method were performed in [6] and [7]. The main limitation of this technique is that it is sensitive to seed point selection. Yang et al. [8], proposed an algorithm in which multiple seed points were accepted from the user for obtaining the initial liver contour using the fast-marching level-set method, then a threshold-based level-set method with initial liver contour as input segmented the actual liver region. In [9], the liver was segmented using FCM integrated with Grey Wolf Optimization to deal with local minima convergence and speed optimized FCM performed liver lesion segmentation. The algorithm requires the number of clusters as input and the execution is computationally complex. Some authors have also combined two or more standard algorithms to achieve better results. Xu et al. [10], combined region growing with a region-based active contour model with a new signed pressure function for liver segmentation.

Some of the factors that make segmentation of liver challenging are ambiguous boundaries with surrounding structures and the presence of large lesions that change the topology of the liver [11]. Many of the conventional segmentation methods mentioned above rely on the intensity of the pixels and do not give accurate results when the desired ROI is not homogeneous. Recently, a lot of DL models were used for liver and liver lesion segmentation. The DL techniques extract relevant features from the input CT images to perform segmentation and give better results than the conventional methods [12], [13]. In most of the papers published, the researchers mostly worked on the segmentation of JPEG images. Digital Imaging
and Communications in Medicine (DICOM) images were hardly considered. DICOM images contain more intensity value range than JPEG images; hence they are more suitable for CAD systems. In our work, we have used DICOM images and performed semantic segmentation. The main advantage of the DL-based segmentation methods is that they are automatic and yield highly accurate results. Although, they require considerable training time, high-performance processors and a large volume of labeled datasets. The benefits outweigh these pitfalls. We have employed a pre-trained SegNet model for liver segmentation and PSO-based clustering for liver lesion segmentation. To the best of our knowledge, this is the only work that uses PSO for liver lesion segmentation.

The paper is organized as follows: Section II describes the SegNet model and PSO-based clustering technique in the current context. The details regarding the methodology adopted are given in section III. The experimental settings, results and discussion are presented in section IV and section V concludes the paper.

II. THEORETICAL BACKGROUND

A. SegNet model

SegNet is a Deep Convolutional Neural Network (DCNN) used for semantic segmentation. It consists of an encoder, a corresponding decoder and a pixel-wise classification layer (Fig. 1). It was introduced by Badrinarayanan et al. [14] for road scene segmentation. The encoder has convolutional layers similar to the VGG-16 network [15]. It consists of one or more convolutional layers with batch normalization and a Rectified Linear Unit (ReLU) layer, then non-overlapping max pooling and sub-sampling layers. At the decoder, upsampling and convolution are performed. Each decoder upsamples the low-resolution feature maps using the respective max-pooling indices produced by the encoder. This improves boundary delineation and reduces the total number of trainable parameters in the decoders. Finally, a softmax classifier layer predicts the category of each pixel, for instance, liver or background.

Fig. 1. SegNet architecture [14]

B. Particle Swarm Optimization

PSO is an evolutionary and population-based optimization algorithm that simulates the movement of bird flocks, where a flock of birds randomly searches for food in a region where only one piece of food exists. The birds do not know where the food is, but they know how far the food is in each time step. They find food by following the bird which is nearest to it [16].

The objective of the PSO algorithm is to find the optimal solution for complex problems. The algorithm maintains a set of potential solutions to the optimization problem called particles. Each particle possesses three details:

\( x_i : \) Current position in \( N_d \) dimensional space
\( v_i : \) Current velocity
\( y_i : \) Personal best position

The particles move through the \( N_d \) dimensional space by adjusting the position towards their current best position and the best position in the neighborhood called global best (gbest) using the following equations:

\[
\begin{align*}
    v_{i+1} &= w \cdot v_i + c_1 \cdot r_1 \cdot (y_{i+1} - x_i) + c_2 \cdot r_2 \cdot (g_{best} - x_i) \quad (1) \\
    x_{i+1} &= x_i + v_{i+1} \quad (2)
\end{align*}
\]

In Eq.1, \( w \) represents the inertia weight, \( c_1 \) and \( c_2 \) are the acceleration constants and \( r_1 \) and \( r_2 \) are sampled from a uniform distribution.

The best position of the particle is calculated using Eq. 3,

\[
y_i^{f+1} = \begin{cases} 
    y_i^f & \text{if} \quad f(x_i^{f+1}) \geq f(y_i^f) \\
    x_i^{f+1} & \text{if} \quad f(x_i^{f+1}) < f(y_i^f) 
\end{cases} \quad (3)
\]

In the proposed algorithm, the fitness or objective function that needs to be minimized is given by,

\[
J_c = \frac{\sum_{j=1}^{N_c} \sum_{z_{p \in C_{i,j}}} d(z_p, m_j)}{N_c} \quad (4)
\]

where \( N_c \) is the number of clusters, \( |C_{i,j}| \) is the number of data vectors belonging to cluster \( C_{i,j} \). \( z_p \) is the vector of the input data belonging to cluster \( C_{i,j} \) and \( m_j \) is the jth centroid of the ith particle in the cluster \( C_{i,j} \). The parameter values that gave the best results are \( i = 1 \) particle, \( N_c = 3 \) clusters corresponding to background, liver and liver lesion, \( N_d = 1 \), \( w = 0.6 \), \( c_1 = 2.5 \), \( c_2 = 2.5 \) and \( t_{max} = 60 \) (maximum number of iterations required to minimize the objective function). The algorithm is described below:

1) For each particle, initialize the centroids for the \( N_c \) clusters randomly.

2) For \( t = 1 \) to \( t_{max} \), for each particle \( i \) do
   a) For each pixel \( z_p \) in the input image,
      i) Compute the Euclidean distance \( d(z_p, m_j) \) \forall cluster centroids \( C_{i,j} \).
      ii) Assign \( z_p \) to cluster \( C_{i,j} \) such that \( d(z_p, m_j) = min_{k \in \{1,2,3, \ldots, N_c\}} \{d(z_p, m_j)\} \).
      iii) Compute the fitness using Eq. 4.
   b) Update \( g_{best} \) and \( y_i \).
   c) Update the cluster centroids using Eq. 1 and 2.
III. OVERVIEW OF THE PROPOSED METHOD

The flow diagram of the proposed algorithm which is based on SegNet and PSO clustering methods, along with the intermediate results is shown in Fig. 2. Each stage is explained in detail in the following subsections.

A. Liver segmentation

The input CT images in DICOM format are initially processed using various operations to make them suitable for training the DL model. The pixel intensities in the input CT images are first converted to Hounsfield Units (HU) then, they are mapped to grayscale values (in the range \([0, L-1]\), where \(L = 2^8\)) by assigning the pixels that have values outside 0-255 range to either 0 or 255 and retaining other pixels as they are. This step was required since training the SegNet model with the image obtained in the previous step did not give satisfactory results. In the next step, the dimensions of the abdominal CT images are changed for optimal training and according to the requirements of the SegNet model. Each grayscale image is converted to RGB image by assigning the same intensity value to R, G and B channels. Then, the images are resized to the resolution 380 × 380. The hyperparameters used for training the SegNet are given in Table I. In an attempt to generalize the SegNet model the training dataset was augmented. The augmentation strategies incorporated include rotation between -45° to 45°, flipping (both horizontal and vertical) and scaling between 50-400%.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Epochs</td>
<td>110</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001</td>
</tr>
<tr>
<td>Hyperparameter tuning algorithm</td>
<td>Stochastic gradient descent with momentum</td>
</tr>
<tr>
<td>Mini batch size</td>
<td>2</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.9</td>
</tr>
</tbody>
</table>

The algorithm for liver segmentation is briefly summarized below:

1) Convert the abdominal CT images’ CT numbers in DICOM format to HUs.
2) Retain the pixels that have values within the range, where and assign all other pixels to 0 or 255.
3) Retain the CT images to the dimension.
4) Train the SegNet model with pre-trained weights from VGG-16 network. The training parameters are given in Table I.
5) Use the fine-tuned SegNet model obtained in Step 4 to segment the liver from the abdominal CT datasets (dataset I) given as input.

B. Liver lesion segmentation

After liver segmentation, the hepatic lesion is segmented by PSO-based clustering and morphological operations on the segmented liver. The algorithm for the same is given below:

1) Apply median operator on the segmented liver volume to remove high-frequency components (mottle).
2) Perform PSO clustering segmentation on the segmented liver volume to obtain the three cluster centroids corresponding to background, liver and liver lesion.
3) Assign each voxel to the nearest centroid.
4) Remove the connected components with total no. of pixels less than 70 as it was observed that most of the lesions were larger than this size.
5) Perform morphological operations like hole filling, erosion and dilation (with disk-shaped structuring element of radius 2) to refine the segmented lesion.

IV. RESULTS AND DISCUSSION

In this section, the details of the algorithm implementation, the qualitative and quantitative results are discussed.

A. Dataset Description and Experimental Settings

The SegNet model was trained with 2428 CT images and their corresponding segmented liver masks obtained from 3D-IRCADb database [17] and Kasturba Medical College and hospital, Manipal. From the latter source only 100 images were used during training, the remaining were from the former. Ten different patient datasets from the latter source were used for testing the algorithm. Each dataset comprised around thirty consecutive slices of liver.

Experiments were performed on a core i7-10750H processor with 16GB RAM using MATLAB R2021a programming platform. The DL model was trained on a server with NVIDIA GeForce GTX 1650 Ti with 4 GB GPU memory.

B. Liver and lesion segmentation results

The evaluation criterion, qualitative and quantitative analysis of the results are discussed here. For qualitative analysis, the opinion of an expert radiologist was considered and for quantitative analysis Dice Coefficient (DC), Volumetric Overlap Error (VOE), Absolute Volume Difference (AVD) and Root Mean Square Symmetric Surface Distance (RMSD) were applied. The proposed algorithm’s output is compared
with the ground truth obtained using ITK-SNAP tool [18], [19] under the guidance of an experienced radiologist for computing the metrics. DC indicates the amount of overlap between two masks (predicted and ground truth). Its value can lie between 0 (no overlap) and 1 (perfect overlap). AVD gives the difference in the predicted and ground truth volumes in percentage. Here, 0% indicates perfect segmentation and 100% stands for worst case. RMSD computes the root value of the squared average symmetric surface distance between the border voxels of predicted and ground truth volumes [20]. It is measured in mm and 0 mm represents perfect segmentation.

The quantitative liver and lesion segmentation results achieved are summarized in Tables II and III, respectively. These are the results obtained for ten patient datasets from Kasturbha hospital each comprising around 30 slices of liver.

### TABLE II
LIVER SEGMENTATION RESULTS

<table>
<thead>
<tr>
<th>Dataset No.</th>
<th>DC</th>
<th>VOE</th>
<th>AVD</th>
<th>RMSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.956</td>
<td>0.884</td>
<td>3.23</td>
<td>2.992</td>
</tr>
<tr>
<td>2</td>
<td>0.947</td>
<td>0.102</td>
<td>3.69</td>
<td>3.560</td>
</tr>
<tr>
<td>3</td>
<td>0.933</td>
<td>0.125</td>
<td>0.66</td>
<td>0.988</td>
</tr>
<tr>
<td>4</td>
<td>0.899</td>
<td>0.184</td>
<td>14.4</td>
<td>6.647</td>
</tr>
<tr>
<td>5</td>
<td>0.952</td>
<td>0.091</td>
<td>1.25</td>
<td>2.398</td>
</tr>
<tr>
<td>6</td>
<td>0.955</td>
<td>0.087</td>
<td>1.79</td>
<td>3.362</td>
</tr>
<tr>
<td>7</td>
<td>0.903</td>
<td>0.177</td>
<td>10.2</td>
<td>6.534</td>
</tr>
<tr>
<td>8</td>
<td>0.950</td>
<td>0.095</td>
<td>0.28</td>
<td>2.365</td>
</tr>
<tr>
<td>9</td>
<td>0.952</td>
<td>0.092</td>
<td>1.96</td>
<td>2.989</td>
</tr>
<tr>
<td>10</td>
<td>0.956</td>
<td>0.085</td>
<td>0.98</td>
<td>15.181</td>
</tr>
<tr>
<td>Average</td>
<td>0.940</td>
<td>0.112</td>
<td>3.84</td>
<td>4.702</td>
</tr>
</tbody>
</table>

The average values of the liver segmentation results achieved are DC = 0.940 ± 0.022, VOE = 0.112 ± 0.038, AVD = 3.84 ± 4.688% and RMSD = 4.702 ± 4.089 mm. It can be seen in Table II that the DC value is over 95% for majority of the datasets. The results show that the SegNet model is effective in liver segmentation. Moreover, it also proves the robustness of the model as the training was mostly done using a different dataset, as mentioned earlier. The conventional segmentation algorithms such as region growing perform segmentation based on the intensity values of the voxels; hence cannot give accurate liver segmentation results due to the challenges posed by abdominal CT images. The liver and its adjacent organs mostly share vague boundaries and have similar intensities. Large peripheral hepatic lesions further complicate liver segmentation as the portion of liver where these lesions are present usually gets excluded from the segmented liver image. DL models learn to segment by extracting relevant features from the training images. They do not rely solely on the intensity values. Hence, they have a greater capacity to segment abdominal CT images more accurately in spite of the complexities mentioned above. Moreover, they do not require any user input; therefore, the output is reproducible.

The evaluation metrics for liver lesion segmentation are DC = 0.4629 ± 0.287, VOE = 0.6986 ± 0.203, AVD = 1568.518 ± 2953.172 % and RMSD = 21.4777 ± 33.583 mm. The metric values illustrated in Table III show that the algorithm based on PSO was only fairly successful in delineating the liver lesion. The main reasons for the low accuracy were that it was unable to segment hypervascular and isodense lesions (Dataset Nos. 8 and 9). The algorithm was unable to detect the lesions for the last dataset (Dataset No. 10); hence we have not included its results in the table. In the qualitative analysis of the results, false positives were observed which has also reduced the segmentation accuracy. Since promising results were achieved for some of the datasets (hypodense lesions), further research may help improve segmentation accuracy. Fig. 3 shows sample results of liver and lesion segmentation for some of the input images. The segmented liver region is marked in blue and lesion region is marked in red for better understanding. Fig. 4 shows a false positive captured by the liver lesion segmentation algorithm for one of the CT slices.

Table IV shows the comparison between the results of the proposed liver segmentation algorithm and other works. Our algorithm has attained better DC than Danilov and Yurova [21]. We can see that the results obtained by Peng et al. [22] and Chen et al. [23] are slightly better. However, their method is semiautomatic and requires user intervention, whereas our method is fully automatic.

The time for training the SegNet model was around 44.5 hours. The algorithm took 108.2 secs to fetch the CT dataset from the source folder. The liver and liver lesion segmentation algorithms took 124 secs and 191.1 secs, respectively. More powerful graphics processing units can reduce the high training time.
V. CONCLUSION

The proposed work discusses liver and lesion segmentation methods for abdominal CT images in DICOM format. The liver is segmented using SegNet and liver lesion using PSO-based clustering technique. SegNet has given good liver segmentation results; hence it can be used for liver delineation in computer-aided systems. However, the lesion segmentation algorithm gave satisfactory results for only hypodense lesions. It was unable to segment certain hypervascular lesions. Further, false positives have reduced the segmentation accuracy. We would try to improve the lesion segmentation algorithm for false positive reduction and better accuracy in the future. Work also needs to be done to enhance the contrast between the liver and lesions.

ACKNOWLEDGMENT

We would like to thank KSTePS, DST, Government of Karnataka, India for their support. We are also grateful to Manipal Institute of Technology, MAHE, Manipal for providing the facilities to carry out the research and Kasturba Medical College and hospital, Manipal for providing the images.

REFERENCES


