34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Semi-automatic Approach to the Segmentation of Liver Parenchyma from 3D CT Images with Extreme Learning Machine

W. Huang, Z.M. Tan, Z. Lin, G.-B. Huang, J. Zhou, C.K. Chui, Y. Su, and S. Chang

Abstract— **This paper presents a semi-automatic approach to segmentation of liver parenchyma from 3D computed tomography (CT) images. Specifically, liver segmentation is formalized as a pattern recognition problem, where a given voxel is to be assigned a correct label – either in a liver or a non-liver class. Each voxel is associated with a feature vector that describes image textures. Based on the generated features, an Extreme Learning Machine (ELM) classifier is employed to perform the voxel classification. Since preliminary voxel segmentation tends to be less accurate at the boundary, and there are other non-liver tissue voxels with similar texture characteristics as liver parenchyma, morphological smoothing and 3D level set refinement are applied to enhance the accuracy of segmentation. Our approach is validated on a set of CT data. The experiment shows that the proposed approach with ELM has the reasonably good performance for liver parenchyma segmentation. It demonstrates a comparable result in accuracy of classification but with a much faster training and classification speed compared with support vector machine (SVM).**

I. Introduction

Accurate liver segmentation is fundamentally important in subsequent steps such as surgical planning for tumor resection, transplantation and diagnosis of hepatic diseases. Currently automatic liver segmentation is still a challenging problem. The difficulties of obtaining precise delineation of liver boundary mainly arise from factors such as enormous shape variability of livers, the lack of clear boundary and poor image quality.

One of the popular techniques for liver segmentation is the application of active shape model (ASM) [1] where the basic idea is to fit a deformable shape model onto the image. It involves the construction of the shape model using principal component analysis (PCA) on a set of landmark points on the shape surfaces. However using shape model can at times be over-restrictive and may not capture all possible variations of livers. Also ASM requires an extensive database of liver shapes as the prerequisite for training. For example Kainmuller *et al.* [1] constructed their model based on as many as 112 liver shapes. Classification based methods [2-4] however only need a small amount of image data samples for training. Users just need an initialized region of interest (ROI) on CT scans to start generating training samples on-the-fly. Hence our classification framework is more appealing for applications where the training set is very scarce and no good shape prior model is available.

Also, many segmentation approaches in the literature rely heavily on proper initialization of liver boundary to function effectively. The level set framework of Lee *et al.* [5] necessitates estimated liver boundary to be initialized first by applying manually seeded region growing. Even for the ASM method, initial localization of liver location is a necessary step to ensure proper registration of the shape model [1]. To our best knowledge, there is currently no effective method for automatic or even semi-automatic liver segmentation with very few training CT data sets and with a fast training speed.

To address these problems, we propose the use of Extreme Learning Machine (ELM) [6] to segment liver voxels from non-liver ones in CT scans. Here, liver segmentation is formalized as a pattern recognition problem. Each voxel is associated with a feature vector which consists of a set of texture features, namely the mean and variance, Law's features [7] and sum-and-difference histograms [8]. Based on these features, an ELM classifier is learned for voxel classification, followed with morphological smoothing and 3D level set [9] refinement. The only assumption is that the liver is the biggest abdominal organ, which helps to locate the correct 3D liver segmentation region.

Our approach does not need any other prior knowledge of the exact shape model of the liver. Furthermore, our framework requires minimal human intervention and tuning. Unlike support vector machine (SVM) [10] or level set method [9], ELM is not very sensitive to selection of network parameters. If the number of hidden nodes is set adequately large, the choice of an activation function is not really important. Therefore significant amount of time can be saved in tuning the parameters, and the training and testing with ELM is very fast.

The organization of the paper is as follows. Section II briefly introduces ELM. Section III presents the proposed approach. Section IV summarizes and discusses the result. The paper is concluded in Section V.

II. Introduction of ELM

One of the recent advancements in computational intelligence is the introduction of ELM by Huang *et al.* [6]. The ELM is designed for single hidden layer feed forward networks (SLFN) and it addresses the speed limitation of traditional learning methods. In ELM, input weights and

*Work partially funded by BEP 1021480009 and JCOAG03-SG05-2009.

Weimin Huang and Jiayin Zhou are with the Institute for Infocomm Research, Singapore. (Phone: +65 6402516. fax: +65 64082000; e-mail: {wmhuang, jzhou}@i2r.a-star.edu.sg).

Zu Ming Tan, Zhiping Lin and Guang-Bin Huang are with the School of EEE, Nanyang Technological Univ, Singapore; email: ({tanz0062,ezplin,egbhuang}@ntu.edu.sg).

Chee Kong Chui is with the Dept of ME, National Univ of Singapore. (e-mail: mpecck@nus.edu.sg).

Yi Su is with the Inst. of High Performance Computing, Singapore (e-mail: suyi@ihpc.a-star.edu.sg).

Stephen Chang is with the National Univ Hospital, Singapore. (e-mail: cfscky@nus.edu.sg).

biases to the hidden nodes are randomly generated and the output weights to the output layer are analytically determined. It has been rigorously proven by Huang *et al.* [11] and others [12] that ELM has dramatically faster learning speed than all existing learning algorithms while offering generalization performance comparable to others such as SVM.

Consider a SLFN network with L hidden nodes and a training dataset consisting of N arbitrary samples $\{(\mathbf{x}_j, t_j)\}_{j=1}^{N}$ where $\mathbf{x}_j \in R^n$ is an input feature vector and $t_j \in R^m$ is the corresponding target vector. Given a sample $(\mathbf{x}_j, t_j)$ the SLFN output is written as

$$f_L(\mathbf{x}_j) = \sum_{i}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i), \qquad j = 1, \ldots, N. \tag{1}$$

Here $\mathbf{w}_i$ and b_i are the input weights and biases to the hidden layer respectively. The output weight β_i links the i-th hidden node to the output layer and $g()$ is the activation function of the hidden nodes. Equation (1) can be compactly represented as

$$\mathbf{H}\beta = \mathbf{T} \tag{2}$$

where

$$\mathbf{H}(\mathbf{w}_1, \ldots, \mathbf{w}_L, b_1, \ldots, b_L, \mathbf{x}_1, \ldots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \tag{3}$$

$$\beta = [\beta_1 \ldots \beta_L]^T \text{ and } T = [t_1 \ldots t_L]^T \tag{4}$$

After the random initialization of the input weights $\mathbf{w}_i$ and biases b_i, matrix $\mathbf{H}$ is essentially known and requires no further tuning. In practice since $L << N$ equality (2) does not hold. ELM determines the optimal solution to (2) in the least square sense by expressing the minimum norm least square solution $\hat{\beta}$ as

$$\hat{\beta} = \mathbf{H}^{\dagger}\mathbf{T} \tag{5}$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix $\mathbf{H}$. With (5), the SLFN can be constructed easily with the training samples.

III. Methodology

A. Pre-processing and ROI selection

A pre-processing step is necessary as the acquired images are often noisy and have different contrast ranges. For computational feasibility the transversal scans are down-sampled to a more manageable resolution of 256 x 256. Users are then requested to mark some liver and non-liver regions in the CT scans either by drawing a circle or using a free-hand delineation of the regions. The drawing needs not to be accurate. The purpose of having the ROI is two-fold: 1) to normalize the intensity range onto the [0,1] range and 2) to generate training samples for the ELM learning. This is the only stage of the framework where users need to manually intervene. Once the samples have been set up, the ensuing procedure is fully automated. After the intensity adjustment, a 2D anisotropic diffusion filter [13] is applied to noise suppression. The anisotropic diffusion filter is preferred over convolution filters because of its edge preserving property.

B. Feature Extraction

Proper selection of features is critical for ensuring high classification accuracy. Intensity feature (e.g. neighborhood mean and variance) is not an effective discriminant as liver parenchyma often shares similar intensity range with adjoining tissues. As liver parenchyma has slightly different textures, incorporation of higher order statistics to characterize such textures can improve the classification performance. Here we quantify texture patterns within a local neighborhood of a given voxel by four types of features, i.e. (i) neighborhood mean, (ii) neighborhood variance, (iii) Law's texture [7], and (iv) Unser's sum-and-difference histograms [8].

The Law's texture in general is capable of discerning a variety of patterns such as average intensity, edge, spot, wave and ripple. The method of sum-and-difference histograms is a way of modeling joint probability distribution between two voxels (random variables) separated at a certain displacement. From the histograms, texture descriptors such as correlation, entropy, contrast, homogeneity, cluster shade and so forth could be derived easily to characterize the statistical structure of the object appearance.

Resolution (or scale) at which the texture is defined is also equally important. In our work, we define the scales of the voxel neighborhood to be 3x3 (fine scale) and 5x5 (coarse scale) respectively for all the features. For computing local average and variance, the method of integral images proposed by Viola and Jones [14] is used to speed up feature computation. All the extracted features are then concatenated into a single feature vector and fed into the ELM as input. To normalize the features, we perform z-score normalization and use a hyperbolic tangent function to map the feature values into a fixed range of (-1, 1).

C. Classifier architecture

There are only a few user-specified parameters to tweak in ELM learning, i.e. the number of hidden nodes L and the type of activation function $g()$ as defined in equation (1). It is observed that ELM is insensitive to the choice of activation function with adequately high number of hidden nodes, which is smaller than that of the support vectors used in a direct SVM learning. Fig.1 shows the performance with different number of hidden nodes and activation functions used for ELM training.

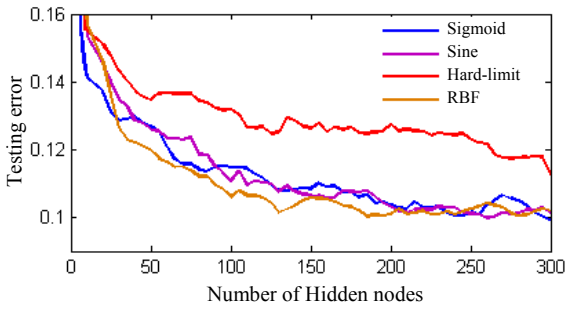
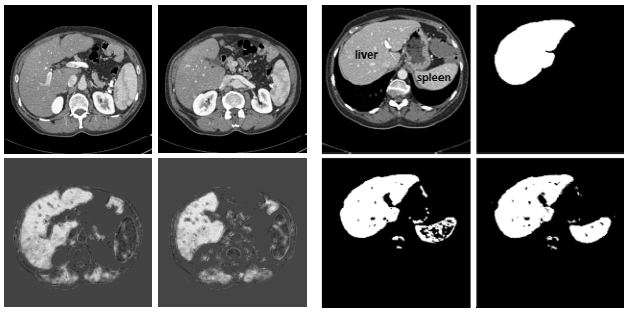


Figure 1. Testing error vs. the number of hidden nodes with various types of activation functions

As shown in Fig.1, the sigmoid, sine and radial basis functions (RBF) exhibit comparable performance when the size of hidden layer is set sufficiently large. In this paper we select 150 hidden nodes and the sigmoid function for ELM.

D. Binary Classification

The ELM classifier is deployed to segment the whole CT volume of each patient, slice-by-slice in the transversal view. A threshold (sign) function is applied to the ELM output to obtain the binary segmentation. Fig. 2(a) illustrates the ELM output (saliency map) where the brightness indicates the likelihood of a voxel belonging to the liver, while Fig. 2(b) compares ELM with SVM, which shows it is able to achieve very similar segmentation result as SVM.



(a) ELM output (b) Binary segmentation

Figure 2. (a) ELM output: Original CT image (top row) and the ELM output (bottom row). (b) Binary segmentation: Original CT image (top left), Ground truth (top right), SVM segmentation (bottom left) and ELM segmentation (bottom right).

E. Post-processing

With ELM classification, the segmentation could still contain both the liver and some falsely classified non-liver voxels. To remove these unwanted tissues, we employ 3D morphological operations (Fig. 3) to post-process the volume. To remove cylindrical structures like inferior vena cava (IVC), we fit an ellipse to each 2D connected component on the transversal view. If the eccentricity of a component is below a predefined cutoff value, the 2D component is discarded since it is too circular to be a liver component. A 5-*mm* sized spherical element is used to morphologically open and smooth the binary volume. From anatomical perspective we know that liver is the largest organ in human body. Incorporating this prior knowledge, the largest 3D connected component is identified as the liver volume.

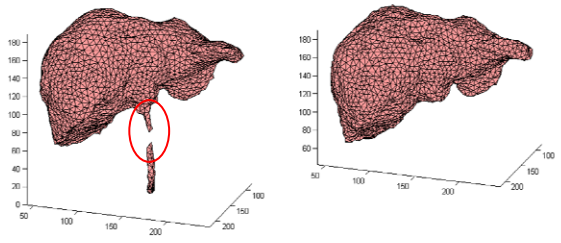


Figure 3. (Left) The liver is connected to the part of inferior vena cava, circled in the ellipse. (Right) Result after applying opening, removing cylindrical structures and extracting the largest connected component.

The opening procedure tends to produce a slightly under-segmented liver volume. To remedy this problem, a 3D distance regularized geodesic level set [9] is used to refine the liver boundary using the previous segmented result. Illustrated in Fig. 4 is an example of the initialization (cyan) from the ELM and morphologic smoothing and the result after level-set refinement (yellow). Since the initial ELM segmentation is very close to the true boundary, the level set method is relatively insensitive to parameters setting (e.g. curvature weightage and weightage of area term) which is an added advantage to our framework.

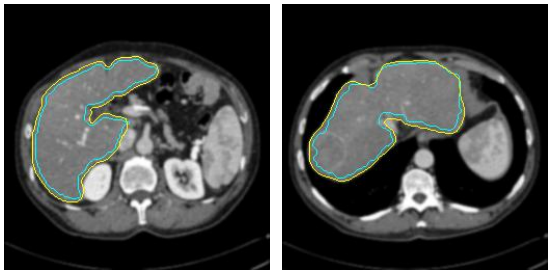


Figure 4. Examples of liver segmentation boundary before (cyan) and after (yellow) level set evolution.

IV. Results and Discussion

We validate our segmentation approach on MICCAI'07 dataset. The segmentation is done with MATLAB running on a standard 3 GHz PC with 2 GB of RAM. The time to segment one slice is about 40 seconds. Visual results of our ELM segmentation are presented in Fig. 5.

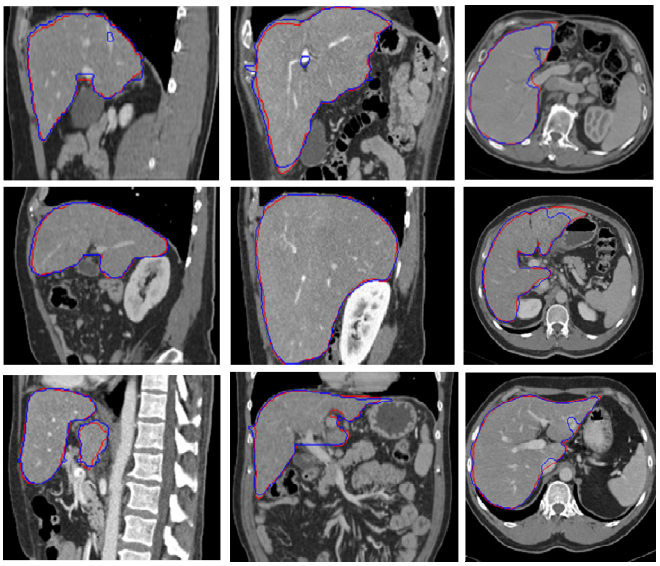


Figure 5. Comparison of our segmentations (blue lines) and ground truth (red lines) in sagittal, coronal and transversal view respectively.

The segmentation result is quantitatively evaluated using the following five metrics [15]: volume overlap (VO), volume difference (VD), average symmetric surface distance (in *mm*), Root Mean Square (RMS) symmetric surface distance (in *mm*) and maximum surface distance (in *mm*). There are a total of 20 datasets in MICCAI'07 liver database. However due to the speed limitation of SVM, we have only evaluated SVM results on 5 datasets. Table I compares the results from SVM and ELM segmentation. It shows the accuracy of ELM is comparable to that of SVM segmentation but at a much faster training speed. As the trained SVM has more than 4000 support vectors, compared to only 150 hidden nodes of ELM, the classification time of ELM is only around 10% of that of SVM based approach.

Table I. Comparison of ELM and SVM segmentation results

SVM						
Dataset	VO (%)	VD (%)	Avg. Dist.(mm)	RMS Dist. (mm)	Max. Dist. (mm)	Training time (s)
1	88.3	-0.2	2.2	3.6	21.0	2109.1
2	91.5	-0.8	1.6	1.8	20.4	2591.2
3	90.7	-3.1	1.7	2.4	15.3	2353.7
4	91.2	-5.4	1.6	2.1	14.5	1538.2
5	92.2	3.6	1.8	2.4	20.5	2292.2
ELM						
Dataset	VO (%)	VD (%)	Avg. Dist.(mm)	RMS Dist. (mm)	Max. Dist. (mm)	Training time (s)
1	90.6	-1.0	1.7	2.2	16.8	104.9
2	91.7	-2.6	1.6	2.1	15.6	106.6
3	91.7	-1.5	1.7	2.2	17.6	109.3
4	92.3	-3.7	1.5	1.9	14.3	90.5
5	91.1	-3.6	1.6	2.5	23.2	89.7

Here we compare SVM (code downloadable from [10]) and ELM classifiers in terms of their sensitivities to user-specified parameters. There are two parameters for SVM with radial basis function (RBF) kernel: cost parameter C and kernel parameter γ. It is rarely known beforehand which combination of (C, γ) yields the best performance for a given application. The SVM has to be tuned iteratively using multiple combinations of (C, γ). As illustrated in Fig. 6, SVM is sensitive to the choice of (C, γ). This shows that ELM is better than SVM since it requires fewer tunings and is less sensitive to the user-specified parameters (see Fig. 1).

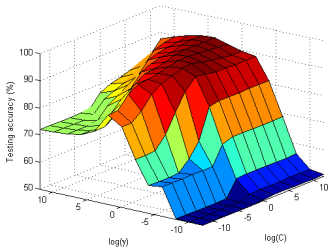


Figure 6. Performance of SVM with RBF kernel trained on multiple (C, γ) combinations (with dataset 2 as an example).

Currently, our method is only capable of segmenting healthy liver parenchyma, with/without small tumors, accurately. It may not be able to handle datasets with large liver tumors. Setting the pathological datasets (2 sets) aside, our overall ELM segmentation performance (on 18 sets) is given in Table II.

Table II. Overall ELM segmentation results

	mean	**std.**	**max.**	**min.**
VO (%)	90.3	1.7	92.3	86.8
VD (%)	-0.7	3.6	5.9	-6.2
Avg. dist. (mm)	1.7	0.3	2.4	1.3
RMS dist. (mm)	2.7	1.0	5.8	1.7
Max. dist. (mm)	22.9	7.5	38.0	14.3

V. CONCLUSION

We summarize our contributions as follows. In this paper, we have proposed an approach using ELM to segment liver voxels in CT scans. We also examined briefly the performance of ELM compared with that of SVM. To deal with the occasional leakage in ELM segmentation, we applied morphologic operation and level set for the post-processing to enhance the segmentation. In our approach, user interaction is only required to expedite the preparation of the training dataset. The remaining classification procedures are fully automated. Even without the shape prior model, the proposed approach performs reasonably well and it is easy to implement.

We have experimentally demonstrated that our method in liver segmentation and show advantageous over SVM in terms of training speed. Compared to existing techniques, our framework is robust to user-specified parameters. In addition, it can operate on a limited amount of training test cases and does not require any prior knowledge about the location, orientation or shape of the liver organ.

As an ongoing work, we are experimenting with using multiple ELM classifiers to segment the healthy parenchyma and pathological tissues (tumors) jointly.

REFERENCES

[1] D. Kainmüller, T. Lange, and H. Lamecker, "Shape constrained automatic segmentation of the liver based on a heuristic intensity model," in *Proc. MICCAI Workshop 3-D Segmentation Clinic: A Grand Challenge*, pp. 109–116, 2007.

[2] E. van Rikxoort, Y. Arzhaeva, and B. van Ginneken, "Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching," in Proc. MICCAI Workshop 3-D Segmentation.Clinic: A Grand Challenge, pp. 101–108, 2007.

[3] R. Susomboon, D. Raicu, J. Furst, and D. Channin, "Automatic single-organ segmentation in computed tomography images,", in Proc. *6th International Conference on Data Mining*, pp. 1081-1086, 2006.

[4] O. Gambino *et al.*, "Automatic volumetric liver segmentation using texture based region growing," in *Int. Conf. of Complex, Intelligent and Software Intensive Systems (CISIS),* pp. 146-152, 2010.

[5] J. Lee, et al, "Efficient liver segmentation exploiting level-set speed images with 2.5D shape propagation," in Proc. MICCAI Workshop 3-D Segmentat. Clinic: A Gand Challenge, pp. 189–196, 2007.

[6] G.-B. Huang, Q.-Y. Zhu and C.-K. Siew, "Extreme learning machine: theory and applications", *Neurocomputing*, vol. 70, pp. 489-501, 2006.

[7] K. Laws, "Textured Image Segmentation", Ph.D. Dissertation, University of Southern California, January 1980.

[8] M. Unser, "Sum and Difference Histograms for Texture Classification," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. PAMI-8, pp. 118-125, 1986.

[9] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," IEEE Trans. Image Processing, vol. 19, no. 12, 2010.

[10] C-C. Chang and C-J. Lin, "LIBSVM : A library for support vector machines", ACM Transactions on Intelligent Systems and Technology, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[11] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: a survey*," International Journal of Machine Leaning and Cybernetics*, pp. 107-122, vol. 2, no. 2, 2011.

[12] C. Pan, P. Dong and Y. Hyouck, "Leukocyte Image Segmentation by Visual Attention and Extreme Learning Machine," *Neural Computing and* Applications, pp. 1-11, 2011.

[13] P. Perona, and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 629-639, 1990.

[14] P. Viola and M. Jones, "Robust Real-time Object Detection", International Journal of Computer Vision, vol 57, no 2, pp. 137-154, 2004.

[15] T. Heimann *et al.*, "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Trans. Medical Imaging*, vol. 28, pp. 1251-1265, 2009.