Random Feature Subspace Ensemble Based Extreme Learning Machine for Liver Tumor Detection and Segmentation

Weimin Huang, Yongzhong Yang, Zhiping Lin, Guang-Bin Huang, Jiayin Zhou, Yuping Duan, Wei Xiong

Abstract—**This paper presents a new approach to detect and segment liver tumors. The detection and segmentation of liver tumors can be formulized as novelty detection or two-class classification problem. Each voxel is characterized by a rich feature vector, and a classifier using random feature subspace ensemble is trained to classify the voxels. Since Extreme Learning Machine (ELM) has advantages of very fast learning speed and good generalization ability, it is chosen to be the base classifier in the ensemble. Besides, majority voting is incorporated for fusion of classification results from the ensemble of base classifiers. In order to further increase testing accuracy, ELM autoencoder is implemented as a pre-training step. In automatic liver tumor detection, ELM is trained as a one-class classifier with only healthy liver samples, and the performance is compared with two-class ELM. In liver tumor segmentation, a semi-automatic approach is adopted by selecting samples in 3D space to train the classifier. The proposed method is tested and evaluated on a group of patients' CT data and experiment show promising results.**

I. INTRODUCTION

In recent years, multi-detector computed tomography (MDCT) has been applied extensively for liver disease diagnosis. However, a large amount of CT images need to be interpreted by radiologists for diagnosis, and such tasks are time-consuming and tedious. In order to conduct the task more efficiently, computer-aided analysis is introduced. In this paper, a liver tumor detection and segmentation using random feature subspace ensemble ELM is proposed.

CT-based liver tumor detection has been investigated by various researchers in the past. In the work of Zhang et al., liver edge and gray maps were extracted from multi-phase CT images and subsequently, tumors were derived from the subtraction of edge and gray maps as well as referring to the score from the spherical gray -level differentiation searching filter [1]. Militzer et al. utilized a probabilistic boosting tree to fully automatically classify liver voxel as either lesion or parenchyma. A robust tumor segmentation was performed simultaneous by an iterative classification and refinement scheme [2]. Pescia et al. [3] proposed the use of advanced non-linear machine learning techniques to determine the optimal features, as well as the hyperplane that uses these features to separate tumor voxels from healthy liver tissues.

For liver tumor segmentation, Zhou et al. has developed a semi-automatic scheme, which is composed of a two-class support vector machine (TSVM) classifier cum a propagational learning strategy for automated sampling, learning and further voxel classification among neighboring slices [4]. For the semi-automatic method reported by Freiman et al., it first classifies the liver voxels into tumor and healthy tissue with a TSVM engine from which a new set of high-quality seeds is generated. Over the 3D images, these seeds then conduct the propagation procedure [5]. Häme et al. reported a semi-automated method using non-parametric intensity distribution estimation and a hidden Markov measure field model, with application of a spherical shape prior. A post-processing operation was also utilized to remove the overflow to adjacent tissue [6]. In [7], Li et al proposed the semi-automatic tumor segmentation method using unified level set method to address the boundary leakage problem by integrating other object information in the object indication function. In a previous work [8], a kernel based ELM classifier is presented for tumor detection and segmentation. However as the one-class ELM is a simply mapping to a hyperplane, limited performance is achieved.

In this paper, we propose an ensemble based ELM to detect and segment liver tumors in CT images. An ELM autoencoder is also used to enhance the performance. A classifier ensemble comprises of learning algorithms and classifies new voxel by a majority voting approach. It has been shown in various cases [9] that classifier ensembles outperform individual classifiers. Random subspace ensembles for fMRI classification are studied by randomly selecting the voxels from fMRI volume to construct ensembles [10]. ELM has been recognized as a fast learning algorithm and it has been proved that kernel based ELM achieves superior performance than traditional ELM [11]. Liu and Wang [12] proposed Ensemble ELM for face recognition, where the data sampling is used to construct the ELMs. Here we propose a new ensemble based ELM by resampling the feature space to construct the classifiers for tumor detection and segmentation. The remaining of the paper is organized as follows. Section II briefs the kernel ELM and autoencoder. Section III presents Ensemble ELM by random feature subspace sampling for tumor detection and segmentation. Section IV discussed the implementation issue. Section V shows the experiment result and it is concluded in Section VI.

II. RELATED WORKS

A. Extreme Learning Machine

Extreme Learning Machine, proposed by Huang et al., is a single-hidden-layer feed-forward neural network (SLFN). It

* Weimin Huang, Jiayin Zhou, Yuping Duan and Wei Xiong are with the Institute for Infocomm Research, 1 Fusionopolis Way, #21-01 Connexis, Singapore 138632 (+65 6408 2516; fax: +65 64082000; email: {wmhuang, jzhou, duany, wxiong}@i2r.a-star.edu.sg).

Yongzhong Yang, Zhiping Lin, Guang-Bin Huang are with the School of EEE, Nanyang Technological Univ, Singapore; email: ({yyang8, ezplin, egbhuang}@ntu.edu.sg).

has been shown that the learning speed is much faster than those of other learning algorithms such as SVM and other learning algorithms [11]. Another advantage of ELM is that the randomly generated hidden layer parameters $\{w_i, b_i\}$ are independent of the training data. The ELM algorithm maps input data from the input space to the L-dimensional hidden layer feature space. Later Kernel based ELM is introduced to enhance the robustness with a regularization coefficient [11]. The output function of the kernel ELM is written as:

$$f(\mathbf{x}) = [\mathbf{h}(\mathbf{x})]' \, \mathbf{H}' \, (\frac{\mathbf{I}}{\lambda} + \mathbf{H}\mathbf{H}')^{-1} \mathbf{T}$$

$$= \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_N) \end{bmatrix}' (\frac{\mathbf{I}}{\lambda} + \mathbf{\Omega}_{ELM})^{-1} \mathbf{T} \quad (1)$$

where $\mathbf{h}(\mathbf{x}) = [g(w_1 \times \mathbf{x} + b_1),..., g(w_L \times \mathbf{x} + b_L)]'$, $\mathbf{T} = [t_1,...t_N]'$ are the labels of training samples $\mathbf{x_1}$, …, $\mathbf{x_N}$, and $\mathbf{H} = [\mathbf{h}(\mathbf{x}_1), ... , \mathbf{h}(\mathbf{x}_N)]'$. Let $\mathbf{\Omega}_{ELM} = \mathbf{H}\mathbf{H}'$, $\mathbf{\Omega}_{ELMi,j} = \mathbf{h}(\mathbf{x}_i)\mathbf{h}(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$. In the kernel implementation, the feature mapping $\mathbf{h}(\mathbf{x})$ need not to be known to users, instead its corresponding kernel $K(\mathbf{u},\mathbf{v})$ can be computed. Here the Gaussian kernel is used, $K(\mathbf{u},\mathbf{v}) = \exp(-\gamma\|\mathbf{u}-\mathbf{v}\|^2)$.

B. Autoencoder Networks

Autoencoder is a neural network that encodes input $X \in R^d$ as a representation $H \in R^l$, and then reconstructs the input data X by setting the target output of to be the same as the input X. Usually back-propagation is used as the training algorithm in autoencoder networks. The work of [13] showed that autoencoder networks can be used as an unsupervised pre-training step which helps to achieve state of the art performance in classification problems, especially in deep architecture approaches. In this paper we adopt the ELM autoencoder for the preprocessing of the feature data.

III. LIVER TUMOR DETECTION AND SEGMENTATION

Here we present the Random Feature Subspace Ensemble ELM (RFSE-ELM) classifier for tumor detection and segmentation. The method mainly consists of five steps: pre-processing of the CT image, feature extraction, pre-training using ELM autoencoder, voxel classification by majority voting using results from ELM classifier ensemble, and post-processing to refine detection and segmentation results.

A. Pre-processing and feature extraction

The original CT images of different scans and patients may have different contrast. In implementation of tumor detection and segmentation, semi-automatic and supervised approach is used here so the training sample is available for normalization of the contrast window level and range around healthy liver intensity [14].

For each training voxel sample, as described in the method in [8], totally 22 features are generated in 3D volume to make them more representative.

B. ELM autoencoder pre-training

In [15], an ELM based autoencoder is proposed. Different from conventional autoencoders which apply back-propagation algorithm for training to obtain the identity

function, ELM is used as the training algorithm in the ELM autoencoder. In the ELM autoencoder, input data $X \in R^d$ is mapped to ELM feature space $H \in R^l$, then the original input X is reconstruected from the ELM feature space H through the reconstruction matrix $\beta \in R^{l,d}$, namely by $X = H\beta$. In this process, the reconstruction matrix β may have captured underlying input data information, and it is used for data training in a base ELM classifier. Therefore β^T is not randomly generated in the ELM, and such process is shown in Fig.1 below.
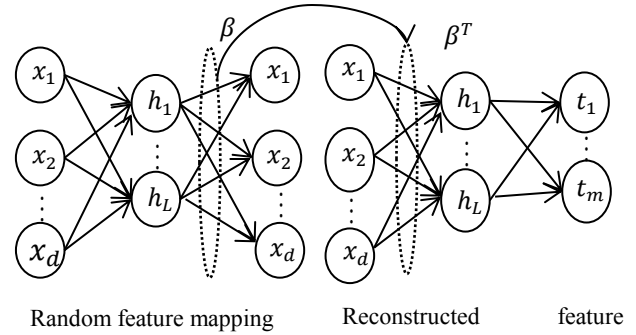

Figure 1. The β^T is used as the input weight of a base ELM classifier in ELM autoencoder.

Before training of each base ELM classifier, ELM auto-encoder pre-training is performed on each feature subset. Experiments have shown that such a process can suppress the noise and on average about 2% of segmentation performance enhancement is obtained. After that, each base ELM classifier is trained, and then a thresholding is performed to obtain classification results for the base classifiers. In the end, a majority voting strategy is applied on all base ELMs to obtain the final result.

C. ELM Classifier Ensemble Construction

Ensemble Learning (EL) has been used in many machine learning research, and there are a spectrum of EL algorithms developed such like boosting, multi-boost and random subspace. An effective EL system needs highly diversified base classifiers. Techniques like resampling, feature subspace partitioning are developed to fulfil this purpose. One of the most popular methods is random subspace ensemble. The idea is to sample from the feature set rather than using all features for each base classifier in the ensemble. Random subspace (RS) ensemble has been proved to outperform single classifiers and other most widely used classifier ensembles such as bagging, adaboost, random forest and rotation forest for functional magnetic resonance imaging (fMRI) classification [10]. In this paper, a random feature subspace ensemble is constructed using kernel based ELM as base classifiers. The reason is that ELM has the advantage of fast training speed and excellent generalization capability.

Different from [10], where the original voxels are taken as features, here the 22 extracted 3D image features are used for subspace sampling. Generally, small ensemble size and relatively medium feature subset could improve the performance [10]. Therefore, we have compared the size of ensembles using 5 to 30 base ELMs, where each ELM has 17 randomly sampled features (around 80% of the feature space

of 22 features). For each base ELM output, a pre-fixed threshold is used to obtain the classification result which can be varying when we computing the receiver operating characteristic (ROC) curve. A majority voting is followed to decide the final result.

We use morphological opening followed by closing to remove small misclassification errors, especially those near the liver boundary. And in tumor detection clustering is performed to connect adjacent detections to merge the false positive tumor regions. The whole process is shown below in the Fig.2.
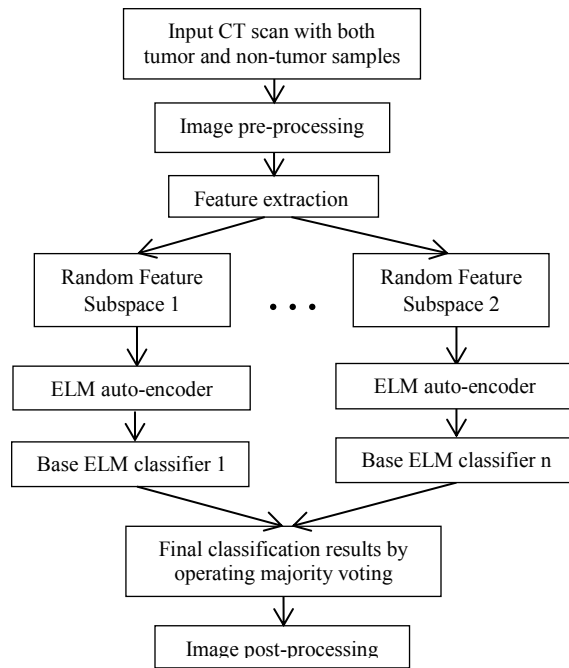
Figure 2. Block diagram of the RFSE- ELM.

IV. EXPERIMENT RESULTS AND DISCUSSION

A. Tumor detection

Tumor detection is performed using both one-class and two-class ensemble based ELM. CT data are available for 20 tumors from 7 patients. Cross-validation is adopted in the following way. All patients' CT data are used for training excluding one patient, and then automatic tumor detection is applied to the patient's CT data. In two-class ELM, training data set contains both tumor and non-tumor samples, whereas only healthy samples are used for training in one-class ELM.

It is proposed by Huang et al. [8] that ELM can also be used for one-class novelty detection. It is applied for tumor detection using a learnt kernel based ELM. However, the performance can be further enhanced by the RFSE ELM proposed here. Firstly, ELM auto-encoder is used as a pre-training step to better represent features, and then the ensemble based ELM is used to take advantage of decision making from majority voting. Sensitivity (S) and false positive (FP) error is used to measure the performance of tumor detection. The detection performance is shown in Fig. 3. Base classifiers are constructed by randomly sampling the feature space, and the ensemble size ranging from 5 to 30 ELM classifiers. The one-class detection method in [8] is used as the baseline for comparison. For clarity, RFSE-ELMs with 5, 15 and 25 classifiers are presented.

From the results, it can be observed that basically all the classifier ensembles outperform or have comparable performance as the baseline, and the best result is obtained with ensemble of 15 classifiers. The general trend is that the detection performance improves with increasing ensemble size until it hits 15, and then the performance degrades. This pattern of performances is presented in other literature as well [8][10].
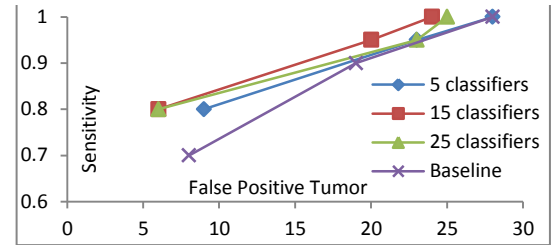
Figure 3. ROC curves for tumor detection by one-class ensembles.

For two-class ensemble based ELM, the performances of ensembles are also better than the baseline in [8]. As the ensemble size changes, a similar trend of performance as one-class detection appears. The detection performances are shown below in Fig. 4.
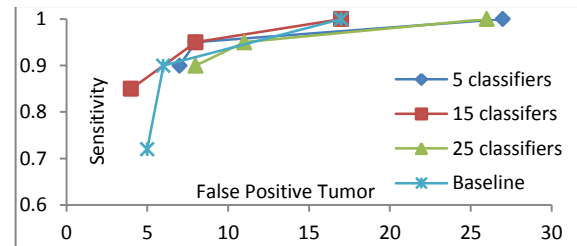
Figure 4. ROC curves for tumor detection by two-class ensembles.

From Fig.3 and Fig.4, it can be concluded that the ensemble based ELM improves the performances for both one-class and two-class detection. Two-class detection still performs better than one-class, but this is totally reasonable. Two types of training data will better characterize liver tumors, and this leads to better performance. However, one-class detection will have its edge when tumors having differently characteristics from existing ones appear. The operation of majority voting helps rule out those healthy voxels which resembles tumors. These voxels are normally liver tumor regions and vessels since their intensities differ significantly from healthy tissue.

B. Tumor segmentation

Fully automatic tumor detection is very time-consuming, and segmentation results are always not satisfying. Some human intervention will greatly enhance the performance, and hence a semi-automatic tumor segmentation approach with the assistance of a region of interest (ROI) is adopted instead in this paper.

To measure the segmentation performance, the metrics proposed in [16] are used: Volume Overlapped (VO),

Volume Difference (VD), Average Symmetric Surface Distance (ASD) and Root Mean Square Symmetric Surface Distance (RMSD) and Maximum Surface Distance (MSD).

The ensemble size ranges from 5 to 30 ELM as well. It is found that for one-class segmentation, the performance increases with ensemble size till 15 base classifiers are reached, and then the performance decreases. However, for two-class segmentation, different ensembles have comparable performances and there is no apparent trend. Similarly, the method in [8] is used as baseline for comparison. Table I summarizes the segmentation results over all the tumors for one-class and two-class ELM ensembles of 15 base classifiers, and they are compared with the baseline.

TABLE I. METRICS OF TUMOR SEGMENTATION

	Metrics	VO (%)	VD (%)	ASD (mm)	RMSD (mm)	MSD (mm)
Two class	mean	74.75	11.89	1.03	1.28	4.77
	max	86.41	36.52	1.94	2.58	7.00
	min	36.62	0.83	0.70	0.77	1.80
One-class	mean	68.82	14.12	1.65	2.11	7.14
	max	95.76	43.47	9.23	12.24	36.73
	min	15.57	1.44	0.59	0.66	1.53
Baseline	mean	67.15	14.16	2.27	2.47	8.46
	max	85.72	50.39	11.92	12.35	20.62
	min	20.49	1.36	0.80	1.03	3.50

From the table, it can be seen that the mean VO of tumor segmentation for two class ensemble ELM is 74.75%. This significantly outperforms the baseline by 7.6 % for mean VO. In addition, for VD, ASD, RMSD and MSD, the ELM classifier ensemble all performs better than the baseline. Although the baseline results are based on two class kernel based ELM, the one class ensemble based ELM turns to outperform it. Nevertheless, it is noted that differences between max and min values are significant for one class ensemble ELM. This implies that the method is still not stable enough, and two-class segmentation is more reliable. Fig. 5 shows some segmentation results (green) using two-class ELM overlapped with ground truth (red) of tumors. The over-segmented regions are basically those voxels at liver-vessel and inter-lobe gaps.
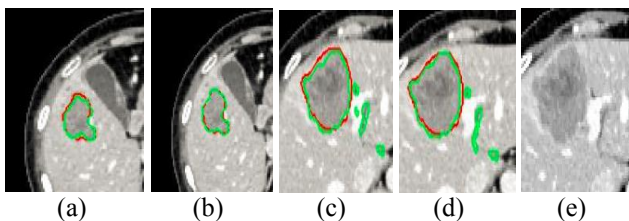

(a)　　(b)　　(c)　　(d)　　(e)

Fig. 5. (a) and (b) show good segmentation results by one-class and two-class ELM respectively. (c) and (d) are segmentation results with errors on (e). The over-segmented regions are in liver-vessel and inter-lobe gaps.

V. CONCLUSION

We have proposed Random Feature Subspace Ensemble based ELM (RFSE-ELM) in this paper. Particularly we proposed the one-class RFSE-ELM classifier that demonstrates the promising and much better performance than the baseline classifier in tumor detection and segmentation. The one-class RFSE-ELM shows that it is possible to detect tumors even there is no tumor data for training. The advantage is that it could serve as a preliminary detection scheme, particularly in the case some unknown or new tumor is not well trained or represented by a two-class classifier. In two-class tumor segmentation, the new method achieved better performance as well. In this work we focus on the feature ensemble, and compared with the baseline classifiers. Due to the semi-automatic scheme, we do not have much data for training. It will be interesting in the future to study the data ensemble for tumor detection and segmentation.

REFERENCES

[1] X. Zhang, T. Furukawa, X. Zhou, et al., "Detection of metastatic liver tumor in multi-phase CT images by using a spherical gray-level differentiation searching filter," Proc. SPIE: Medical Imaging 2011, Vol. 7963, pp. 79632k1-8, Feb. 2011.

[2] A. Militzer, T. Hager, F. Jager, et al., "Automatic detection and segmentation of focal liver lesions in contrast enhanced CT images," Proc. International Conference on Pattern Recognition (ICPR), pp. 2524-2527, Aug. 2010.

[3] D. Pescia, N. Paragios, S. Chemouny, "Automatic detection of liver tumors," 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 672-67,.2008.

[4] J. Zhou et al. "Semi-automatic Segmentation of 3D Liver Tumors from CT Scans Using Voxel Classification and Propagational Learning," MICCAI workshop on 3D segmentation in the clinic: a grand challenge II. 2008.

[5] M. Freiman, O. Cooper, D. Lischinski, and L. Joskowicz, "Liver tumors segmentation from CTA images using voxels classification and affinity constraint propagation," Int. J. Comput. Assist. Radiol. Surg., vol. 6, no. 2, pp. 247-255, Mar. 2011.

[6] Y. Häme and M. Pollari, "Semi-automatic liver tumor segmentation with hidden Markov measure field model and non-parametric distribution estimation," Med. Image Anal. vol. 16, no. 1, pp. 140-149, June 2012.

[7] B. N. Li, C. K. Chui, S. Chang, S. H. Ong, "A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced CT images," Expert Systems with Applications, Vol. 39, pp.9661–9668, 2012.

[8] W. Huang, N. Li, Z. Lin, G.-B. Huang, W. Zong, J. Zhou and Y. Duan, "Liver tumor detection and segmentation using kernel-based extreme learning machine," IEEE Conf on Engineering in Medicine and Biology Society (EMBC), vol., no., pp.3662-3665, 3-7 July 2013.

[9] L. Rokach, "Ensemble-based classifiers," in Artificial Intelligence Review, vol. 33, no. 1-2, pp. 1–39, 2010.

[10] L. Kuncheva, J. Rodriguez, C. Plumpton, D. Linden and S. Johnston, "Random Subspace Ensembles for fMRI Classification," Medical Imaging, IEEE Transactions on, vol.29, no.2, pp.531-542, Feb. 2010.

[11] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme Learning Machines:A Survey," International Journal of Machine Leaning and Cybernetics, 2(2), pp. 107-122, 2011.

[12] N. Liu and H. Wang, "Ensemble Based Extreme Learning Machine," IEEE Signal Processing Letters, Vol. 17, No. 8, 2010, pp.754-757.

[13] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, Samy Bengio, "Why does unsupervised pre-training help deep learning?," Journal of Machine Learning Research, Vol. 11, pp.625-660, Feb. 2010.

[14] D. Smeets, D. Loeckx, B. Stijnen, et al., "Semi-automatic level set segmentation of liver tumors combining a spiral-scanning technique with supervised fuzzy pixel classification," Med. Image Anal. Vol. 14, No. 1, pp. 13-20, Feb. 2010.

[15] L. L. C. Kasun, H. Zhou, and G.- B. Huang, C. M. Vong, "Representational Learning with ELMs for Big Data," IEEE Intelligent Systems, 11/12, 2013, pp.3-6.

[16] 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.