Optik 125 (2014) 4057–4063

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound

Kaizhi Wu[a,c], Xi Chen[b], Mingyue Ding[a,*]

[a] Department of Biomedical Engineering, School of Life Science and Technology, Key Laboratory of Image Processing and Intelligent Control of Education Ministry of China, Huazhong University of Science and Technology, Wuhan 430074, China
[b] Institute for Pattern Recognition and Artificial Intelligence, Science and Technology on Multi-spectral Information Processing Laboratory, Huazhong University of Science and Technology, Wuhan 430074, China
[c] School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China

ARTICLE INFO

Article history:
Received 23 July 2013
Accepted 18 January 2014

Keywords:
Deep learning
Contrast-enhanced ultrasound
Focal liver lesions

ABSTRACT

Classification of liver masses is important to early diagnosis of patients. In this paper, a diagnostic system of liver disease classification based on contrast enhanced ultrasound (CEUS) imaging is proposed. In the proposed system, the dynamic CEUS videos of hepatic perfusion are firstly retrieved. Secondly, time intensity curves (TICs) are extracted from the dynamic CEUS videos using sparse non-negative matrix factorizations. Finally, deep learning is employed to classify benign and malignant focal liver lesions based on these TICs. Quantitative comparisons demonstrate that the proposed method outperforms the compared classification methods in accuracy, sensitivity and specificity.

1. Introduction

Primary liver cancer is the sixth most common cancer worldwide, and the third most common cause of death from cancer [1]. In order to increase the chances for survival by providing optimal treatments, early detection and accurate diagnosis of liver cancer is of utmost importance [2,3]. Biopsy is currently the golden standard for diagnosing cancer, but it is invasive, uncomfortable, and is not always a viable option depending on the location of the tumor [4–6]. Noninvasive diagnosis of focal liver lesions (FLLs) can be evaluated by using CEUS to determine the liver vascularization patterns in real-time, and thus, improve the diagnostic accuracy for the classification of FLLs [7].

Recently, many studies have investigated CEUS patterns of FLLs, establishing their typical behaviour in the arterial, portal and venous phases [8,9]. The normal liver is a highly vascular organ predominantly supplied by both hepatic artery (25%) and portal vein (75%) [10]. However, malignant focal liver lesions (i.e., hepatocellular carcinomas (HCCs), hypervascularity metastases) are supplied by the hepatic artery as well as tumor vessels. Therefore, the enhancement patterns of FLLs in the arterial and portal venous phases of CEUS can be used for characterizing FLLs [11]. Compared with healthy parenchyma, benign liver lesions are typically hyper-enhanced at all time, whereas malignant lesions usually present a hyper-enhanced pattern during the arterial phase and become hypoenhanced in the later portal-venous phase [12].

Time intensity curves (TICs) are a graphical illustrating representative contrast uptake kinetics represented in a CEUS investigation. Comparative TIC analysis between a tumoral region of interest (ROI) and parenchymal equivalent ROI could enhance the diagnostic accuracy of CEUS, thus establishing its role in liver cancer diagnosis [13,14]. Previous reports have shown that the analysis of TICs such as the area under the curve (AUC) and time to peak (TTP) have statistical significance between benign and malignant lesions of various types of tumors in the hemodynamic measurements [15–18].

Based on the TICs of CEUS, diagnostic systems had been developed to assist ultrasonographist in liver cancer processing to further improve the diagnostic accuracy. Casey et al. [18] extracted the TIC of each pixel within a ROI and used the measured TICs' parameters as the features to classify the benign and malignant tumors of rats by linear discriminant analysis. The problem of the method lies in that the extraction of TICs is susceptible to noise because the CEUS imaging signal is noisy due to many factors (i.e., speckle noise, fluctuations in the concentration of microbubbles) [18]. Streba et al. [19] extracted the TIC with the mean of the signal intensity within a manual drawing ROI surrounding the tumor and also used the measured TICs' parameters (i.e., AUC, TTP) as the features to classify liver tumors by artificial neural networks (ANN). The TICs obtained from ROI measurements may be composites of

* Corresponding author.
E-mail address: mingyueding.hust@gmail.com (M. Ding).

http://dx.doi.org/10.1016/j.ijleo.2014.01.114
0030-4026/

activities from different overlapping components in the selected ROI. It is the major disadvantage of ROI measurements method. Junji et al. [11] estimated TICs for an FLL by use of a series of the temporally averaged microflow imaging (MFI) images and employed a cascade of six independent ANN by use of extracted temporal and image features for classifying liver diseases.

The limitations of the above-mentioned methods were that the feature selection was determined empirically and always operator-dependent. Furthermore, the parameters setting are based on the experimental knowledge. To address the problem, we introduce deep learning [20] into the diagnostic system to classify the benign and malignant focal liver lesions. Deep learning is employed in this work because it is received much attention recently. It combines the feature extraction and recognition together perfectly. The feature extraction is implemented from low level to high level through unsupervised feature learning instead of being hand-designed [21]. Deep learning simulates the human brain to recognize objects through different layers' features.

Moreover, to overcome the subjectivity of TICs extracted with manual ROI selection and the impact of speckle noise, an automatical TICs extraction method is used. The TICs are extracted from the dynamic CEUS image sequences by Factor Analysis of Dynamic Structures (FADS) techniques. As far as we know, this is the first report combining TICs extracted automatically with deep learning to develop a diagnostic system.

The rest of this paper is organized as follows: the related works are introduced in Section 2. Section 3 describes the data acquisition and pre-processing, sparse non-negative matrix factorizations, the deep learning classifier and the framework of the classification system. In Section 4, we present the classification results and discuss the results obtained in the experiment. Finally, we give the conclusion in Section 5.

2. Related works

The TICs extracted from dynamic CEUS image sequences can be used to detect the aberrant functionality of tumor vasculature. Factor Analysis of Dynamic Structures (FADS) [22,23] is a technique used for the extraction of TICs from a series of dynamic images. The technique allows homogenous physiological structures with different temporal characteristics to be identified. Recently, FADS has been investigated in hepatic perfusion studies based on CEUS imaging [24–27]. However, one of the major drawbacks of FADS is that the solution is not mathematically unique when only non-negativity constraints are used. In order to guarantee the solution corresponding to the physiological truth uniquely, we use a sparse non-negative matrix factorizations as presented in our previous work [27] to extract the TICs.

Due to the dual blood supply from the hepatic artery and portal vein, the balance between arterial and portal blood supply is an indication of the type of lesion [26]. From a hemodynamic perspective, benign and malignant lesions in the liver differ in their respective needs in arterial blood supply [12,28]. Furthermore, most FLLs show unique enhancement patterns in the first two phases [18,29]. Therefore, in the study, we extracted the TICs of arterial and portal vein phases.

3. Materials and methods

3.1. Data acquisition and pre-processing

Ultrasound examinations were performed by an experienced ultrasonographist using a Philips iU22 equipped with a C5-1 transducer (Philips Medical Systems, Bothel, WA) and contrast specific imaging (CSI). Initially, a B-mode scan was performed to identify the best approach to the lesion. Thereafter, a bolus of 1.5–2.4 ml of Sonovue (Bracco, Milan, Italy) was injected intravenously through a cubital vein, followed by flush of NaCl 0.9% 5 ml in bolus. Real-time side by side contrast-enhanced mode continuous video clip with a mechanical index of less than 0.20 were acquired at a frame rate of 8-15 fps.

The study population comprised 22 patients with 26 lesions who underwent CEUS in Huazhong University of Science and Technology affiliated Wuhan Union Hospital between March 2012 and May 2013. Positive diagnosis was reached through a combination of other imagistic methods (CT and CE-MRI), liver biopsy in uncertain cases or followup for a minimum period of sixth months. All the cases consisted of 6 hepatocellular carcinomas (HCCs), 10 cavernous hemangiomas (CHs), 4 liver abscesses, 3 metastases (METASs), and 3 localized fat sparings (LFSs). The patients' ages ranged from 18 to 73 years (mean, 43.5 ± 9.9 years), 12 case of male and 10 case of female. The average size of the tumor was 21.2 ± 13.8 mm (size range, 10.0–56.3 mm) for benign lesions and 23.0 ± 10.3 mm (size range, 9.0–32.2 mm) for malignant tumors.

To minimize the impact of breathing motion on TICs extraction and improve the accuracy of the classification system, an image correction technique as presented in our previous work [30] that combining of template matching and frame selection was applied to compensate respiratory motion throughout each CEUS video. Respiratory motion compensation for the free-breathing data is an obligatory pre-processing step before the TICs extraction.

3.2. Sparse non-negative matrix factorizations

To extract the TICs from the dynamic CEUS image sequences, we introduce a sparse non-negative matrix factorizations (SNMF) [27,31] which uses the sparseness of each pixel in all coefficient images as a degree of the amount of mixing. The sparseness degree function is the ℓ_1-norm of each pixel. The objective function is defined by:

$$\min_{C,F} \quad \frac{1}{2}\left\|A - CF\right\|_F^2 + \alpha\sum_{i=1}^{N}\left\|C_i\right\|_1^2, \quad s.t. \quad C, F \geq 0 \tag{1}$$

where the size of matrix A is $N \times M$, N is the number of pixels in the image and M is the number of dynamic images. Matrix C and matrix F are the coefficients image and the TICs which defined in FADS model. C_i is the ith row vector of C.

In order to correct for the nonuniqueness of the solution in optimization problem (1), we use the Frobenius norm of F to constraint it. The final objective function becomes:

$$\min_{C,F} \quad \frac{1}{2}\left\|A - CF\right\|_F^2 + \alpha\sum_{i=1}^{N}\left\|C_i\right\|_1^2 + \beta\left\|F\right\|_F^2, \quad s.t. \quad C, F \geq 0 \tag{2}$$

where $\beta > 0$ is a parameter to suppress $\| F \|_F^2$, and $\alpha > 0$ is a regularization parameter to balance the trade-off between the accuracy of approximation and sparseness of C.

The objective function is solved by the alternating non-negativity-constrained least squares (ANLS) algorithm referred in [31]. The sparse NMF (SNMF) algorithm begins with the initialization of F with non-negative values. Then, it iterates the following ANLS until it is convergence:

$$\min_{C}\left\|C(F \quad \sqrt{\alpha}\boldsymbol{e}_{K\times 1}) - (A \quad \mathbf{0}_{K\times 1})\right\|_F^2, \quad s.t. \quad C \geq 0 \tag{3}$$

where $\boldsymbol{e}_{\boldsymbol{K\times 1}} \in \mathrm{R}^{\boldsymbol{K\times 1}}$ is a column vector with all components equal to one, $\mathbf{0}_{K\times 1} \in \mathbb{R}^{K\times 1}$ is a zero vector, and

$$\min_{F}\left\|F^T(C^T \quad \sqrt{\beta}\boldsymbol{I}_K) - (A^T \quad \mathbf{0}_{M\times K})\right\|_F^2, \quad s.t. \quad F \geq 0 \tag{4}$$

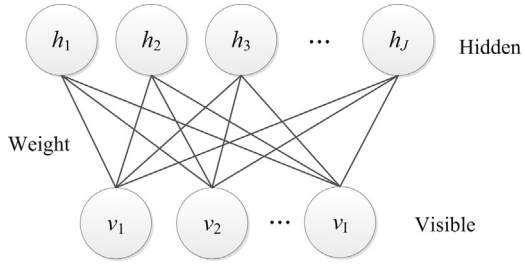

Fig. 1. A RBM with I visible units and J hidden units.

where $\boldsymbol{I_K}$ is an identity matrix of size $K \times K$ and $\mathbf{0}_{M\times K}$ is a zero matrix of size $M \times K$.

Our algorithm can be described below:

Algorithm 1. TICs extraction using SNMF

Input: load a dynamic CEUS image sequence A
Initialization: F^0 is initialized by a non-negative random matrix. C^0 is initialized to be inf. N is the max iteration times. ε is the stopping threshold. K is determined by using orthogonal singular value decomposition of A
Output: C^*, F^*
Method:
1. Begin $i = 0$
2. While $i < N$ and $\| A - C^i F^i \|_F > \varepsilon$
   - Using NLS algorithm to solve the subproblem: $C^{i+1} = \underset{C}{\arg\min} \left\| C(F \quad \sqrt{\alpha}\boldsymbol{e}_{K\times 1}) - (A \quad \mathbf{0}_{K\times 1}) \right\|_F^2, \quad s.t. \quad C \geq 0$
   - Using NLS algorithm to solve the subproblem: $F^{i+1} = \underset{F}{\arg\min} \left\| F^T(C^T \quad \sqrt{\beta}\boldsymbol{I}_K) - (A^T \quad \mathbf{0}_{M\times K}) \right\|_F^2, \quad s.t. \quad F \geq 0$
   - $i \leftarrow i + 1$
3. $C^*(k, :) = F^i(k, :) / \max\left(C^i(:, k)\right), k = 1, \cdots, K$
4. $F^*(:, k) = C^i(:, k) * \max\left(C^i(:, k)\right), k = 1, \cdots, K$

3.3. Deep learning

Deep learning was firstly introduced by Hinton et al. [20,32] for a class of deep probabilistic generative models called Deep Belief Networks (DBNs). Generative models provide a joint probability distribution over observable data and labels, facilitating the estimation of both P(Observation | Label) and a P(Label | Observation), while discriminative models are limited to the latter, P(Label | Observation) [33]. DBNs are composed of several layers of restricted Boltzmann machines (RBM), a type of neural network [21,34–36]. As shown in Fig. 1, these networks are "restricted" to a two-layer architecture in which the visible, binary stochastic units v are connected to hidden binary stochastic units h, where connections are formed between the layers (units within a layer are not connected). The hidden units are trained to capture higher-order data correlations that are observed at the visible units.

As proposed in [35], a joint configuration, (v, h) of the visible and hidden units has an energy given by:

$$E(v, h) = -\sum_{i\in visible} a_i v_i - \sum_{j\in hidden} b_j h_j - \sum_{i,j} v_i h_j w_{ij} \tag{5}$$

where v_i and h_j are the binary states of visible unit i and hidden unit j, a_i and b_j are their biases and w_{ij} is the weight between them. The network assigns a probability to every possible pair of a visible and a hidden vector via this energy function:

$$p(v, h) = \frac{1}{Z} e^{-E(v,h)} \tag{6}$$

where the "partition function", Z, is given by summing over all possible pairs of visible and hidden vectors:

$$Z = \sum_{v,h} e^{-E(v,h)} \tag{7}$$

The probability that the network assigns to a visible vector, v, is given by summing over all possible hidden vectors:

$$p(v) = \frac{1}{Z} \sum_{h} e^{-E(v,h)} \tag{8}$$

The derivative of the log probability of a training vector with respect to a weight is:

$$\frac{\partial \log p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model} \tag{9}$$

where the angle brackets are used to denote expectations under the distribution specified by the subscript followed. $\langle v_i h_j \rangle_{data}$ denotes an expectation with respect to the data distribution and $\langle v_i h_j \rangle_{model}$ is an expectation with respect to the distribution defined by the model. This leads to a simple learning rule for performing stochastic steepest ascent in the log probability of the training data:

$$\Delta w_{ij} = \epsilon(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}) \tag{10}$$

where ϵ is a learning rate.

Because there are no direct connections between hidden units in an RBM, it is easy to get an unbiased sample of $\langle v_i h_j \rangle_{data}$. Given a randomly selected training data, v, the binary state, h_j, of each hidden unit, j, is set to 1 with the probability:

$$p(h_j = 1|v) = \sigma\left(b_j + \sum_i v_i w_{ij}\right) \tag{11}$$

where $\sigma(x)$ is the logistic sigmoid function $\sigma(x) = 1/(1 + exp(-x))$. $v_i h_j$ is then an unbiased sample.

Because there are no direct connections between visible units in an RBM, it is also easy to get an unbiased sample of the state of a visible unit, *given a hidden vector*:

$$p(v_i = 1|h) = \sigma\left(a_i + \sum_j h_j w_{ij}\right) \tag{12}$$

However, it is much more difficult to get an unbiased sample of $\langle v_i h_j \rangle_{model}$. In practice learning, it is done by following an approximation to the gradient of a different objective function, called the "contrastive divergence" [37].

This starts by setting the states of the visible units to a training vector. Then the binary states of the hidden units are all computed in parallel using Eq. (11). Once binary states have been chosen for the hidden units, a "reconstruction" is produced by setting each v_i to 1 with a probability given by Eq. (12). The change of weight is given by:

$$\Delta w_{ij} = \epsilon\left(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recon}\right) \tag{13}$$

A DBN with l layers models the joint distribution between observed vector x and l hidden layers h^k as follows:

$$P(x, h^1, \cdots, h^l) = \left(\prod_{k=1}^{l-2} P(h^k|h^{k+1})\right) P(h^{l-1}, h^l) \tag{14}$$

where $x = h^0$, $P(h^k|h^{k+1})$ is a conditional distribution for visible hidden units in an RBM associated with level k of the DBN, and $P(h^{l-1}, h^l)$ is the visible-hidden joint distribution in the top-level RBM. Fig. 2 illustrates a graphical depiction of a 3-layer DBN. DBN is a generative model (generative path, with bold arrows) and a means

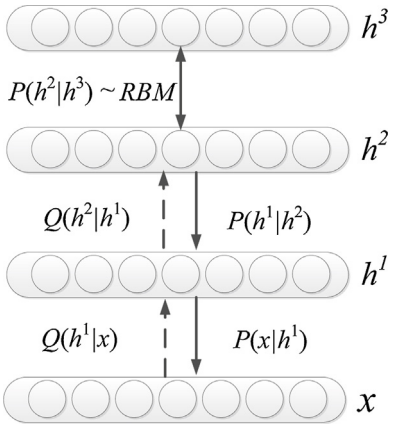

Fig. 2. A graphical depiction of a 3-layer DBN.

to extract multiple levels of representation of the input (recognition path, with dashed arrows). The top two layers h^2 and h^3 form an RBM (for their joint distribution). The lower layers form a directed graphical model (sigmoid belief net $h^2 \to h^1 \to x$) and the prior for the penultimate layer h^2 is provided by the top-level RBM.

The DBN can be trained in a purely unsupervised way, with the greedy layer-wise procedure in which each added layer is trained as an RBM by contrastive divergence. The DBN training processing can be described as Algorithm 2.

Algorithm 2. Training the classification model

Input: TICs extracted from dynamic CEUS image sequences
Definition: $\hat{p}$ is the input training distribution for the network; ϵ is a learning rate for the stochastic gradient descent in contrastive divergence; L is the number of layers to train; $n = (n_1, \cdots, n_L)$ is the number of hidden units in each layer; W^i is the weight matrix for level i, for i from 1 to L; b^i is the bias vector for level i, for i from 0 to L; $Q(h^k, h^{k+1})$ is defined as the kth RBM; $Q(h^k, h^{k+1})$ is an approximation of $P(h^k, h^{k+1})$.
Output: the classification model
Method:
• initialize $b^0 = 0$
for $l = 1$ to L do
  • initialize $W^i = 0$, $b^i = 0$
  while not stopping criterion do
    • sample $h^0 = x$ from $\hat{p}$
    for $k = 1$ to $l - 1$ do
      • sample h^k from $Q(h^k|h^{k+1})$
    end for
    • RBMupdate(h^{l-1}, ϵ, W^l, b^l, b^{l-1}) thus providing $Q(h^l|h^{l-1})$ for future use
  end while
end for

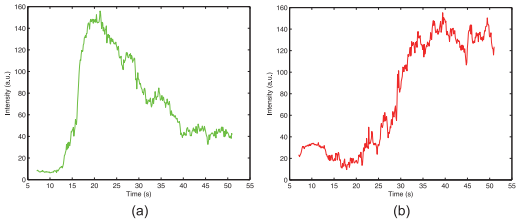

Fig. 4. TICs extracted from one of the malignant lesions. (a) and (b) show the TICs corresponding to the arterial and portal vein phases, respectively.

There are several studies demonstrating the effectiveness of deep learning in a variety of application domains, such as MNIST handwriting challenge [38], face detection [39,40], speech recognition and detection [41], general object recognition [42], natural language processing [43], and meat spoilage markers [44].

3.4. The framework of the classification system

The block diagram of the proposed framework is shown in Fig. 3. At the training phase, the acquired liver CEUS videos in the training set are preprocessed through respiratory motion compensation. The TICs of the CEUS videos are extracted by the ultrasonographist. The ground truths of whether the CEUS videos belong to benign or malignant cases are labeled. During training phase, the TICs set are used as inputs to deep learning classifier in order to train a model for classification. The obtained classification model is the output at the training phase. At the testing phase, the test CEUS videos are also preprocessed as training phase and TICs are extracted by sparse non-negative matrix factorizations. Subsequently, at the testing phase, the classification model is used to determine the class of the case.

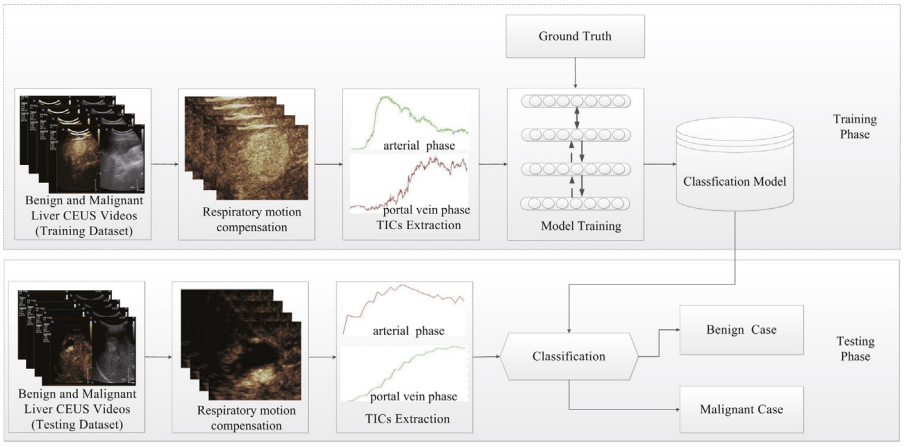

Fig. 3. Block diagram of the proposed framework for focal liver lesions classification; the blocks inside the dash shaded rectangular box represent the training phase, and the blocks within the real line box represent the testing phase.

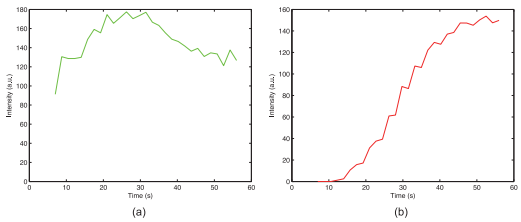

Fig. 5. TICs extracted from one of the benign lesions. (a) and (b) show the TICs corresponding to the arterial and portal vein phases, respectively.

4. Results and discussions

4.1. TICs extracted by sparse non-negative matrix factorizations

Fig. 4 shows the TICs extracted from one of the malignant lesions (HCC cases). Fig. 5 shows the TICs extracted from one of the benign lesions (hemangioma cases). It is easy to see that benign and malignant lesions can be differentiated by the wash-in (arterial phase) and wash-out phase (portal vein phases).

4.2. Classification results

To demonstrate the effectiveness of the classification method in the framework, we do some comparison experiments. Firstly, we determine the experiment parameters of the training system by comparison the classification result of the training data with different numbers of hidden layer and hidden unit. The experimental method with tenfold cross-validation testing was employed. In tenfold cross-validation testing protocol, the training data is divided into 10 subsets and the processing is repeated 10 times. The results are averaged. Then, we use the training data to train a model for classification and test the model using testing data. To evaluate the classification performance, comparison experiments are made among the proposed method and other methods such as linear discriminant analysis (LDA) [45], K-nearest neighbors (KNN) [46], support vector machine (SVM) [47] and back propagation net (BPN) [37]. For quantitative evaluation, classification performance is measured by the accuracy, sensitivity and specificity [48].

In the proposed method, the learning rate is fixed at 0.1, the max epoch is set to 100, the momentum for smoothness is set to 0.5, and the weight decay factor is set to 2e−4. For SVM, the Kernel function is linear function. For KNN, the number of neighbors is set to 10, the Euclidean distance is used. For BPN, the Feed-Forward network is used, the learning rate is fixed at 0.01, the max epoch is set to 500, and the training target error is set to 0.01.

Tables 1–3 list the accuracy, sensitivity and specificity values of the proposed method on different hidden unit for the number of hidden layer with 1, 2 and 3, respectively. It is easy to see that with the increasing hidden unit at different hidden layer, there are not differences in accuracy values. The value of sensitivity and specificity has some differences, but the difference is little. It demonstrates that the proposed method is robust to the parameters.

Table 4 lists the accuracy, sensitivity and specificity values of the compared methods. Fig. 6 shows the visual comparisons among the accuracy, sensitivity and specificity values using the evaluated methods. Obviously, the proposed method outperforms the compared methods in terms of accuracy, sensitivity and specificity values.

Table 1
Experiment results of the proposed method with one hidden layer on different hidden units (mean± std).

Units	Accuracy	Sensitivity	Specificity
5	87.95 ± 4.02	89.16 ± 7.85	87.74 ± 5.09
10	87.95 ± 3.56	89.59 ± 7.16	87.75 ± 4.87
20	87.95 ± 4.42	88.70 ± 7.79	87.58 ± 5.09
30	87.95 ± 4.29	88.56 ± 7.20	87.27 ± 4.75
40	87.95 ± 6.07	89.79 ± 13.19	87.51 ± 6.74
50	87.95 ± 4.02	89.56 ± 5.25	87.41 ± 5.90
60	87.95 ± 5.36	89.20 ± 13.10	87.51 ± 5.10
70	87.95 ± 4.55	90.44 ± 9.30	87.55 ± 4.92
80	87.95 ± 4.68	89.82 ± 6.51	87.36 ± 6.15
90	87.95 ± 5.67	89.48 ± 10.18	87.76 ± 7.31
100	87.95 ± 4.02	89.00 ± 8.47	87.43 ± 4.07
200	87.95 ± 6.59	89.33 ± 7.19	87.81 ± 7.54

Table 2
Experiment results of the proposed method with two hidden layer on different hidden units (mean± std).

Units	Accuracy	Sensitivity	Specificity
5	87.95 ± 4.16	89.39 ± 16.07	87.44 ± 5.45
10	87.95 ± 3.56	89.90 ± 8.06	87.56 ± 4.08
20	87.95 ± 6.86	89.07 ± 8.35	87.30 ± 8.71
30	87.95 ± 5.47	88.30 ± 14.78	87.41 ± 5.05
40	87.95 ± 4.68	89.93 ± 9.53	87.37 ± 5.08
50	87.95 ± 4.92	88.58 ± 9.90	87.70 ± 5.40
60	87.95 ± 4.92	88.37 ± 12.24	87.36 ± 6.87
70	87.95 ± 5.77	89.41 ± 9.34	87.37 ± 7.13
80	87.95 ± 3.87	88.53 ± 6.24	87.69 ± 4.02
90	87.95 ± 3.40	89.23 ± 8.98	87.53 ± 3.11
100	87.95 ± 3.72	89.60 ± 6.19	87.64 ± 4.03
200	87.95 ± 3.40	89.29 ± 9.16	87.50 ± 3.72

Table 3
Experiment results of the proposed method with three hidden layer on different hidden units (mean± std).

Units	Accuracy	Sensitivity	Specificity
5	87.95 ± 4.29	88.96 ± 7.32	87.59 ± 5.28
10	87.95 ± 5.25	88.64 ± 10.39	87.46 ± 4.23
20	87.95 ± 5.47	88.47 ± 7.94	87.30 ± 8.71
30	87.95 ± 5.67	88.87 ± 11.43	87.42 ± 6.33
40	87.95 ± 6.34	89.59 ± 7.68	87.43 ± 6.50
50	87.95 ± 3.56	89.28 ± 6.65	87.86 ± 8.91
60	87.95 ± 6.07	89.29 ± 9.49	87.26 ± 4.33
70	87.95 ± 5.67	89.65 ± 6.50	87.66 ± 6.35
80	87.95 ± 4.55	89.49 ± 6.84	87.79 ± 6.75
90	87.95 ± 3.56	88.15 ± 8.50	87.48 ± 5.36
100	87.95 ± 4.92	89.43 ± 7.51	87.74 ± 6.99
200	87.95 ± 3.22	89.64 ± 9.25	87.42 ± 4.50

4.3. Discussions

We have evaluated deep learning for the design of a classifier between malignant and benign focal liver lesions from contrast enhanced ultrasound time-series data. There are a number of motivations for selecting deep learning as a classification mechanism. Deep learning has been largely applied to many classification systems and have been shown to perform well [38–44]. Deep learning architecture expresses its full potential when dealing with highly varying functions. It allows unsupervised pre-training to achieve good generalization performance in practice. Deep learning aims at learning feature hierarchies with features from higher levels of the hierarchy formed by the composition of lower level features.

Table 4
Experiment results of each learning algorithm on CEUS data.

Algorithm	Accuracy	Sensitivity	Specificity
LDA	81.82	66.67	81.25
KNN	83.33	80.00	82.45
SVM	77.27	66.67	81.25
BPN	82.58	81.86	86.61
Proposed	**86.36**	**83.33**	**87.50**

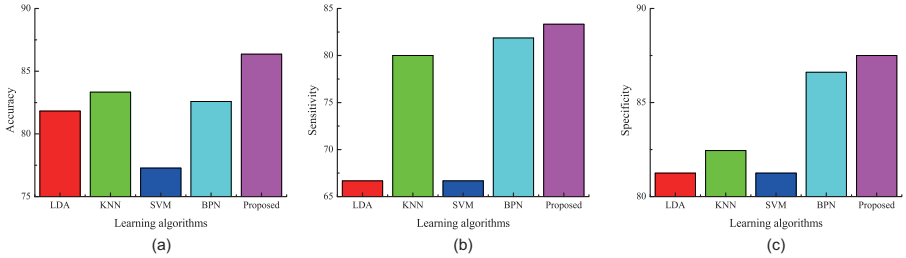

Fig. 6. Experiment results of several learning algorithms on CEUS data. (a) Accuracy. (b) Sensitivity. (c) Specificity.

We perform deep learning on the TICs of the CEUS videos automatically extracted by sparse non-negative matrix factorizations. The accuracy, sensitivity and specificity values are higher than other compared methods.

In this study, we extracted the TICs of the arterial and the portal venous phase of CEUS videos because the appropriate enhancement phase for distinguishing FLLs depended on the type of disease. However, the quality of the CEUS videos depended on the patients' condition. Therefore, dosage of contrast agent is an important variable required to extract accurate enhancement curves for both the hepatic artery and portal vein. This problem was addressed by the choice of the dosage and mechanical index to minimize these effects. Different patients have differing ideal doses which is the best chosen with experience. A similar issue related to image data analysis of CEUS. The video images provided by ultrasound systems are log-compressed data. The decompression function is unknown. This raw envelope data is typically unavailable. Thus, future work should operate on raw, unprocessed data.

Patient motion during acquisitions of different CEUS series images may introduce inaccuracies in the discrimination of lesion properties. To this end, an image registration method was applied to the dynamic series to correct such motion before TICs were extracted.

Nevertheless, the dataset is composed of a limited number of lesions and types of liver diseases. For example, we did not include focal nodular hyperplasia (FNH) because the number of FNH cases was small. However, we believe that the diagnostic framework can be used to classify other types of liver diseases which is our next research direction in the future.

5. Conclusion

A deep learning based classification for contrast enhanced ultrasound imaging is proposed in the paper. Distinctively, the proposed method trains the classification model through deep learning based on the TICs extracted from CEUS videos. Extensive testing has demonstrated that the proposed method outperforms the compared methods in accuracy, sensitivity and specificity.

Acknowledgements

The work is supported by the National 973 Project (Grant no. 2011CB933103) and the Project of the National 12th-Five Year Research Program of China (Grant no. 2012BAI13B02). The authors would like to thank Dr. Tianwei Yan and Ziming Zhang from Huazhong University of Science and Technology affiliated Wuhan Union Hospital, for their assistance with data acquisition and insightful comments.

References

[1] J. Ferlay, H.R. Shin, F. Bray, D. Forman, C. Mathers, D.M. Parkin, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int. J. Cancer 127 (2010) 2893–2917.
[2] G.T. Budd, M. Cristofanilli, M.J. Ellis, A. Stopeck, E. Borden, M.C. Miller, J. Matera, M. Repollet, G.V. Doyle, L.W.M.M. Terstappen, D.F. Hayes, Circulating tumor cells versus imaging-predicting overall survival in metastatic breast cancer, Clin. Cancer Res. 12 (2006) 6403–6409.
[3] R.A. Smith, V. Cokkinides, D. Brooks, D. Saslow, O.W. Brawley, Cancer screening in the United States, 2010: a review of current American Cancer Society guidelines and issues in cancer screening, CA. Cancer J. Clin. 60 (2010) 99–119.
[4] K. Yasufuku, T. Fujisawa, Staging and diagnosis of non-small lung cancer: invasive modalities, Respirology 12 (2007) 173–183.
[5] U. Veronesi, G. Viale, G. Paganelli, S. Zurrida, A. Luini, V. Galimberti, P. Veronesi, M. Intra, P. Maisonneuve, F. Zucca, G. Gatti, G. Mazzarol, C.D. Cicco, D. Vezzoli, Sentinel lymph node biopsy in breast cancer: ten-year results of a randomized controlled study, Ann. Surg. 251 (2010) 595–600.
[6] L.A. Eskew, R.L. Bare, D.L. McCullough, Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate, J. Urol. 157 (1997) 199–202.
[7] E. Quaia, F. Calliada, M. Bertolotto, S. Rossi, L. Garioni, L. Rosa, R. Pozzi-Mucelli, Characterization of focal liver lesions with contrast specific US modes and a sulfur hexafluoride-filled microbubble contrast agent: diagnostic performance and confidence, Radiology 232 (2004) 420–430.
[8] M. Claudon, D. Cosgrove, et al., Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver – update 2012, Ultraschall Med. 34 (2013) 11–29.
[9] V. Salvatore, A. Borghi, E. Sagrini, M. Galassi, A. Gianstefani, L. Bolondi, F. Piscaglia, Quantification of enhancement of focal liver lesions during contrast-enhanced ultrasound (CEUS). Analysis of ten selected frames is more simple but as reliable as the analysis of the entire loop for most parameters, Eur. J. Radiol. 81 (2012) 709–713.
[10] P.E. Sijens, Perfusion magnetic resonance imaging of the liver, World J. Gastroenterol. 16 (2010) 1598–1609.
[11] S. Junji, K. Sugimoto, F. Moriyasu, N. Kamiyama, K. Doi, Computer-aided diagnosis for the classification of focal liver lesions by use of contrast-enhanced ultrasonography, Med. Phys. 35 (2008) 1734–1746.
[12] N.G. Rognin, M. Arditi, L. Mercier, P.J. Frinking, M. Schneider, G. Perrenoud, A. Anaye, J.Y. Meuwly, F. Tranquart, Parametric imaging for characterizing focal liver lesions in contrast-enhanced ultrasound, IEEE Trans. Ultrason. Ferroelectr. Freq. Control. 57 (2010) 2503–2511.
[13] A. Ignee, M. Jedrejczyk, G. Schuessler, W. Jakubowski, C.F. Dietrich, Quantitative contrast enhanced ultrasound of the liver for time intensity curves – reliability and potential sources of errors, Eur. J. Radiol. 73 (2010) 153–158.
[14] R.S. Goertz, T. Bernatik, D. Strobel, E.G. Hahn, T. Haendl, Software-based quantification of contrast-enhanced ultrasound in focal liver lesions-a feasibility study, Eur. J. Radiol. 75 (2010) 22–26.
[15] X.Q. Dong, Y. Shen, L.W. Xu, C.M. Xu, W. Bi, X.M. Wang, Contrast-enhanced ultrasound for detection and diagnosis of renal clear cell carcinoma, Chin. Med. J. Engl. 122 (2009) 1179–1183.
[16] H. Marret, S. Sauget, B. Giraudeau, M. Brewer, J. Ranger-Moore, G. Body, F. Tranquart, Contrast-enhanced sonography helps in discrimination of benign from malignant adnexal masses, J. Ultrasound Med. 23 (2004) 1629–1639.
[17] M. Mitterberger, A. Pelzer, D. Colleselli, G. Bartsch, H. Strasser, L. Pallwein, F. Aigner, J. Gradl, F. Frauscher, Contrast-enhanced ultrasound for diagnosis of prostate cancer and kidney lesions, Eur. J. Radiol. 64 (2007) 231–238.
[18] N.T. Casey, K. Yuko, V.B. Christopher, F.M. Robert, C.K. Andrew, Automating tumor classification with pixel-by-pixel contrast-enhanced ultrasound perfusion kinetics, J. Vac. Sci. Technol. B 30 (2012) 02C103.
[19] C.T. Streba, M. Ionescu, D.I. Gheonea, L. Sandulescu, T. Ciurea, A. Saftoiu, C.C. Vere, I. Rogoveanu, Contrast-enhanced ultrasonography parameters in neural network diagnosis of liver tumors, World J. Gastroenterol. 18 (2012) 4427–4434.
[20] G.E. Hinton, S. Osindero, Y. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (2006) 1527–1554.
[21] D. Yu, L. Deng, Deep learning and its applications to signal and information processing, IEEE Signal Process. Mag. 28 (2011) 145–154.
[22] D.C. Barber, The use of principal components in the quantitative analysis of gamma dynamic studies, Phys. Med. Biol. 25 (1980) 283–292.
[23] R.D. Paola, J.P. Bazin, F. Aubury, A. Aurengo, F. Cavailloles, Y. Herry, E. Kahn, Handling of dynamic sequences in nuclear medicine, IEEE Trans. Nucl. Sci. 43 (1982) 1310–1321.

[24] G. Renault, F. Tranquart, V. Perlbarg, A. Bleuzen, A. Herment, F. Frouin, A posteriori respiratory gating in contrast ultrasound for assessment of hepatic perfusion, Phys. Med. Biol. 50 (2005) 4465–4480.
[25] G.J. Lueck, T.K. Kim, P.N. Burns, A.L. Martel, Hepatic perfusion imaging using factor analysis of contrast enhanced ultrasound, IEEE Trans. Med. Imaging 27 (2008) 1449–1457.
[26] J. Zhang, M.Y. Ding, F. Meng, X.M. Zhang, Quantitative evaluation of two-factor analysis applied to hepatic perfusion study using contrast-enhanced ultrasound, IEEE Trans. Biomed. Eng. 60 (2013) 1449–1457.
[27] X. Chen, K.Z. Wu, M.Y. Ding, N. Sang, Sparse non-negative matrix factorizations for ultrasound factor analysis, Optik (2013), http://dx.doi.org/10.1016/j.ijleo.2013.04.048.
[28] K.V. Ramnarine, E. Leen, K. Oppo, W.J. Angerson, C.S. McArdle, Contrast-enhanced Doppler perfusion index: Clinical and experimental evaluation, J. Ultrasound Med. 21 (2002) 1121–1129.
[29] T.V. Bartolotta, A. Taibbi, M. Midiri, R. Lagalla, Focal liver lesions: contrast-enhanced ultrasound, Abdom. Imaging 34 (2009) 193–209.
[30] J. Zhang, M.Y. Ding, F. Meng, M. Yuchi, X.M. Zhang, Respiratory motion correction in free-breathing ultrasound image sequence for quantification of hepatic perfusion, Med. Phys. 38 (2011) 4737–4748.
[31] H. Kim, H. Park, Sparse non-negativematrix factorization via alternating nonnegativity-constrained least squares for microarray data analysis, Bioinformatics 23 (2007) 1495–1502.
[32] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (2006) 504–507.
[33] I. Arel, D.C. Rose, T.P. Karnowski, Deep machine learning – a new frontier in artificial intelligence research, IEEE Comput. Intell. Mag. 5 (2010) 13–18.
[34] Y. Bengio, Learning deep architectures for AI, Found. Trends Mach. Learn. 2 (2009) 1–271.
[35] G.E. Hinton, A practical guide to training restricted boltzmann machines, Lect. Notes Comput. Sci. 7700 (2012) 599–619.
[36] R. Salakhutdinov, I. Murray, On the quantitative analysis of deep belief networks, in: Proceedings of the 25th International Conference on Machine Learning (ICML'08), 2008, pp. 872–879.
[37] G.E. Hinton, Training products of experts by minimizing contrastive divergence, Neural Comput. 14 (2002) 1711–1800.
[38] The MNIST Database of Handwritten Digits [Online], Available: http://yann.lecun.com/exdb/mnist/
[39] B. Kwolek, Face detection using convolutional neural networks and Gabor filters, Lect. Notes Comput. Sci. 3696 (2005) 551–556.
[40] M. Osadchy, Y. LeCun, M. Miller, Synergistic face detection and pose estimation with energy-based models, J. Mach. Learn. Res. 8 (2007) 1197–1215.
[41] S. Sukittanon, A.C. Surendran, J.C. Platt, C.J.C. Burges, Convolutional networks for speech detection, Interspeech (2004) 1077–1080.
[42] F.J. Huang, Y. LeCun, Large-scale learning with SVM and convolutional nets for generic object categorization, in: Proc. Computer Vision and Pattern Recognition Conf. (CVPR'06), 2006.
[43] H. Lee, Y. Largman, P. Pham, A. Ng, Unsupervised feature learning for audio classification using convolutional deep belief networks, in: Advances in Neural Information Processing Systems, 22 (NIPS'09), 2009.
[44] M. Längkvist, S. Coradeschi, A. Loutfi, J.B.B. Rayappan, Fast classification of meat spoilage markers using nanostructured ZnO thin films and unsupervised feature learning, Sensors 13 (2013) 1578–1592.
[45] E.L. Shafey, C. McCool, R. Wallace, S. Marcel, A scalable formulation of probabilistic linear discriminant analysis: applied to face recognition, IEEE Trans. Pattern. Anal. Mach. Intell. 35 (2013) 1788–1794.
[46] S.Z. Erdogan, T.T. Bilgin, A data mining approach for fall detection by using k-nearest neighbour algorithm on wireless sensor network data, IET Commun. 6 (2012) 3281–3287.
[47] C.C. Chang, C.J. Lin, LIBSVM: A Library for Support Vector Machines, 2001, Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm
[48] N. Lavrac, Selected techniques for data mining in medicine, Artif. Intell. Med. 16 (1999) 3–23.