Received July 18, 2018, accepted September 22, 2018, date of publication October 11, 2018, date of current version November 8, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2873043

Artificial Intelligence Based Diagnosis for Cervical Lymph Node Malignancy Using the Point-Wise Gated Boltzmann Machine

QI ZHANG[1,2], YUE LIU[2], HONG HAN[3], JUN SHI[1], AND WENPING WANG[3]

[1]Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China
[2]The SMART (Smart Medicine and AI-Based Radiology Technology) Laboratory, Institute of Biomedical Engineering, Shanghai University, Shanghai 200444, China
[3]Department of Ultrasound, Zhongshan Hospital, Fudan University, Shanghai 200032, China

Corresponding author: Hong Han (hanhongzs@163.com) and Wenping Wang (puguang61@126.com)

This work was supported by the National Science Foundation of China under Grant 61671281, Grant 81627804, and Grant 61471231.

ABSTRACT This paper aims to build an artificial intelligence (AI) architecture for automated extraction of learned-from-data image features from contrast-enhanced ultrasound (CEUS) videos and to evaluate the AI architecture for classification between benign and malignant cervical lymph nodes. An AI architecture for CEUS feature extraction and classification was constructed by using the point-wise gated Boltzmann machine (PGBM). The PGBM consisted of task-relevant and task-irrelevant hidden units for both feature learning and feature selection, and the task-relevant units were connected to the support vector machine (SVM) to yield the likelihood for classification. The synthetic minority over-sampling technique was used to improve the classification ability for an unbalanced data set. Experimental evaluation was performed with the five-fold cross validation on a database of 127 lymph nodes (39 benign and 88 malignant) from 88 patients. The SVM likelihood exhibited a significant difference between benign and malignant cervical lymph nodes (0.74 ± 0.21 versus 0.33 ± 0.28, $p < 0.001$). On the test set, the accuracy, precision, sensitivity, specificity, and Youden's index of the AI architecture were 82.55%, 89.58%, 84.75%, 77.56%, and 62.32%, respectively. The AI architecture using the PGBM shows promising classification results, and it may be potentially used in clinical diagnosis for cervical lymph node malignancy.

INDEX TERMS Artificial intelligence, contrast-enhanced ultrasound, cervical lymph nodes, point-wise gated Boltzmann machine.

I. INTRODUCTION

Cervical lymphadenopathy often presents with head and neck malignancies such as thyroid cancer, lymphoma, and esophageal cancer [1]. Differentiating malignant from benign cervical lymph nodes is important for clinical management including staging, prognosis, and optimization of treatment process [1]. Contrast-enhanced ultrasound (CEUS) is an emerging technique for diagnosis of lymphadenopathy [2]–[7]. CEUS involves harmonic imaging and administration of ultrasound contrast agent intravenously to enhance the backscattering signal and visualize the perfusion of intranodal blood vessels [2], [4], [7].

Lymph node examination using CEUS provides a video lasting for a few minutes that may include thousands of sequential images. Visual interpretation of a CEUS video is subjective, tedious and time-consuming for a radiologist or oncologist [8], which also limits the accuracy of lymphadenopathy diagnosis. Therefore, it would be desirable to develop computer-aided diagnosis (CAD) approaches for more accurately, objectively and efficiently interpreting CEUS images and distinguishing malignant from benign cervical lymph nodes [9], [10].

Recent CAD systems for lymphadenopathy on ultrasound uses the statistical features (SFs), also called the human-crafted features in research areas of computer vision [11], [12]. The SFs include shape features and intensity statistics of a lymph node that are calculated on B-mode, Doppler sonography and elastography [12]–[15], as well as perfusion features that are derived from the time-intensity curve (TIC) analysis on CEUS [3]. The SFs are often extracted by depending on professional knowledge or human labor, and the choice of specific SFs largely affects the

diagnosis performance [16]. Artificial neural networks (ANNs) have gained a huge attention in various fields on account of recent advances in artificial intelligence (AI) technology [17]. ANNs involve applying connected units called artificial neurons to raw data (such as the pixel values in an image) to learn features (representations) of the data. Thus in an image classification scenario, ANNs allow a computer to be fed with raw pixel values and to automatically explore the learned-from-data features needed for end-to-end classification [17]. Recent development of CAD for breast tumors with ultrasound has suggested that the image features automatically induced from pixels by ANNs could outperform the SFs [11], [18]. The performance of ANNs on the classification of lymphadenopathy with CEUS is yet to be determined.

However, lymph node CEUS images contain speckle noise, artifacts including acoustic shadows, and other irrelevant patterns including the interference from neighboring tissues or large vessels such as carotid arteries. These irrelevant patterns may all hinder precise classification of lymph nodes. To deal with the irrelevant patterns is an extreme difficulty in building an ANN architecture that can robustly learn from complex CEUS video images. Hence the challenge is how to learn robust representations that can differentiate useful (i.e., task-relevant) patterns from large amounts of distracting (i.e., task-irrelevant) patterns [19].

Popular ANN methods including the convolutional neural network and autoencoder are inappropriate for overcoming the challenge, because they do not focus on distinguishing task-relevant and irrelevant patterns [11]. Instead, a newly developed ANN approach, the point-wise gated Boltzmann machine (PGBM), appears to be promising through introduction of a gating mechanism for estimating where task-relevant patterns occur [19]. Thus, the aim of this study is to construct an AI architecture by using PGBM for automated extraction of learned-from-data, task-relevant image features from CEUS videos, and to evaluate the AI architecture in classification between benign and malignant cervical lymph nodes.

II. MATERIALS AND METHODS

A. IMAGE ACQUISITION AND PRE-PROCESSING

This retrospective study was approved by our institutional review board and informed consent was obtained from all patients. Cervical lymph nodes suspected for malignancy were referred to CEUS examination at the Department of Ultrasound, Zhongshan Hospital, Fudan University. The CEUS video images of 127 lymph nodes from 88 patients (33 male and 55 female; 49.1 $\pm$ 15.4 years old; age range: 20–77 years) were collected. Malignant diagnosis was reached through pathology as the gold standard, while the benignancy was confirmed through pathology or negative follow-up. There were 39 benign and 88 malignant lymph nodes in the cohort.

The CEUS imaging was carried out with the MyLab Twice ultrasound system (Esaote SpA, Genoa, Italy) and a bolus injection of 1.5 mL of the contrast agent SonoVue (Bracco SpA, Milan, Italy). Each CEUS video was in a size of 800 $\times$ 555 with 256 gray scales, a resolution of 138.3 $\pm$ 24.3 pixel/cm, and a sampling frequency of 24.2 $\pm$ 3.4 frame/s.

After injection of contrast agent, the average gray level (AGL) within a lymph node on CEUS varies along time due to blood perfusion. In a CEUS video sequence, an elliptical region of interest was specified along the nodal border, and the AGL within the region was calculated to generate a TIC. From the TIC, the frame with the largest AGL called the peak frame was detected. A sequence of successive frames, which lasted several seconds and was centered on the peak frame, was automatically selected from the video images with the Butterworth band-pass and low-pass filtering. These selected frames were averaged to derive a temporal mean image (TMI). The TMI was downsampled to a fixed size of 32 $\times$ 54 and converted to a vector of 1728 dimensions as the input of the AI architecture. For details of the CEUS image pre-processing, readers can refer to our previous work [8], [20], [21].

B. PGBM-BASED AI ARCHITECTURE FOR FEATURE EXTRACTION AND CLASSIFICATION

A CEUS image of a cervical lymph node contained a large amount of irrelevant sensory patterns such as speckle noise, acoustic shadows, and interference from neighboring tissues or vessels, which were distracting to the diagnosis. Thus it was crucial to explore an automated algorithm to distinguish between relevant and irrelevant patterns. Here we used the PGBM to estimate where task-relevant patterns occurred on CEUS and thereby constructed an AI architecture for CEUS feature extraction and classification based on the PGBM.

As shown in Fig. 1, first, the CEUS video images were pre-processed to produce a TMI image, and the TMI image was converted to a vector as an input to the PGBM. Second, the PGBM introduced a gating mechanism by relying on the so-called ''switch units'', through which the PGBM performed feature selection both on learned high-level features (i.e., hidden units) and on raw features (i.e., visible units that were TMI image pixels) [11], [19]. Finally, based on the features selected by the PGBM, the support vector machine (SVM) was used to yield the likelihood of classification and to identify the benign and malignant cervical lymph nodes. Here, the SVM likelihood of classification was a posterior probability between 0 and 1 representing the possibility of a lymph node being classified into a benign one [11], [12]. For details of the PGBM technique, please refer to the pioneering technical work by Sohn et al. [19] and our recent application of PGBM to breast tumor sonoelastogrpahy [11].

C. COMBINATION OF PGBM AND SMOTE FOR UNBALANCED DATA

The ratio of benign and malignant sample size was 1:2.3 in our study, which means the dataset was unbalanced. The unbalance of data has a large influence on the

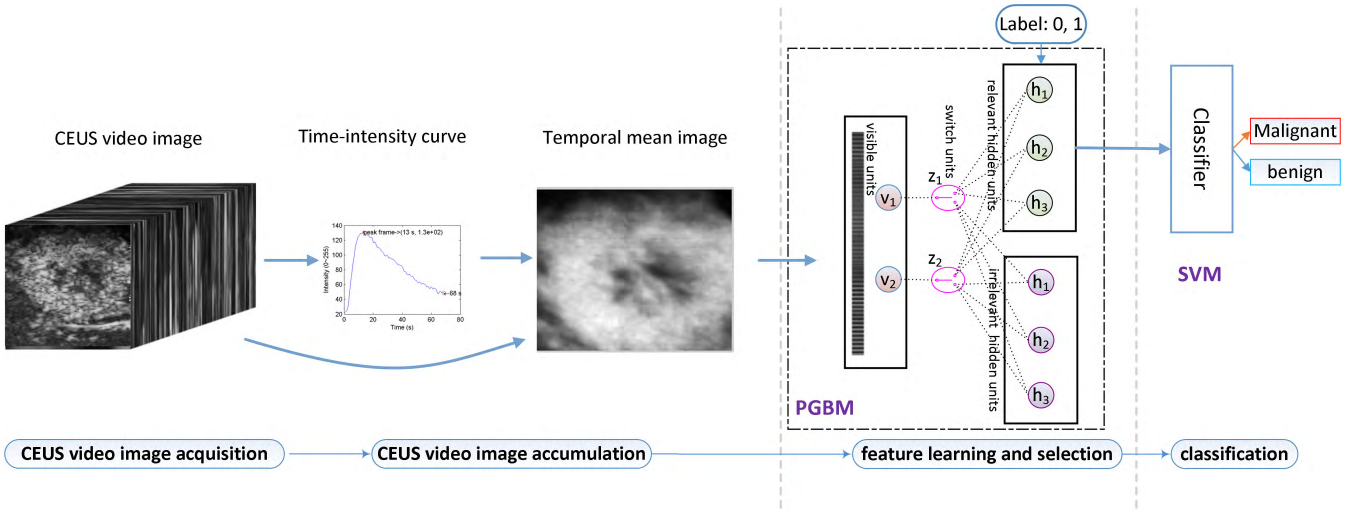


FIGURE 1. Schematic diagram of our artifically intelligent diagnosis architecture for lymphanopathy on contrast-enhanced ultrasound (CEUS) images based on the point-wise gated Boltzmann machine (PGBM).

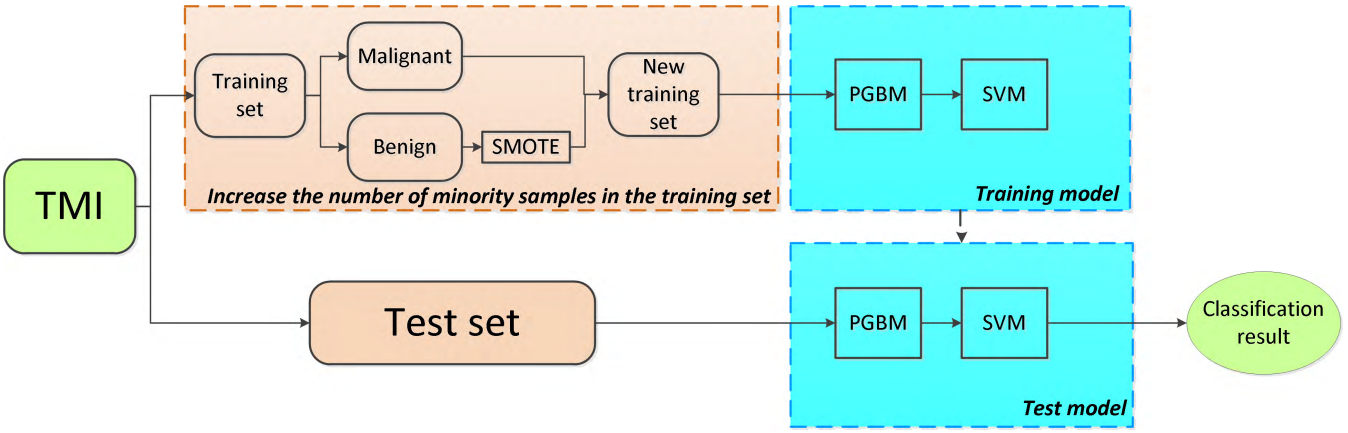


FIGURE 2. The AI architecture for diagnosis on an unbalanced dataset with the point-wise gated Boltzmann machine (PGBM) and the synthetic minority over-sampling technique (SMOTE).

classification performance of SVM [22], [23]. Therefore, we used the synthetic minority over-sampling technique (SMOTE) to reduce the class unbalance by creating synthetic minority class examples [22]. The basic principle of the SMOTE is to interpolate between two minority class samples, resulting in a new minority class sample.

The PGBM-based AI architecture for diagnosis on an unbalanced dataset is illustrated in Fig. 2, including the following steps:

1) The dataset was divided into a training set and a test set;

2) The SMOTE was used to increase the number of minority class samples in the training set to reduce the unbalance of data;

3) The PGBM and SVM were used on the training set for training the feature extraction and classification model;

4) The PGBM and SVM were used on the test set for validating the performance of the trained model.

D. CROSS VALIDATION AND PERFORMANCE EVALUATION

The proposed approach for AI based diagnosis, as well as the statistical analysis and validation, was implemented with MATLAB R2014a (The MathWorks, Natick, MA, USA).

1) GENERAL EVALUATION

First on the entire cohort, the independent two-sample t-test was employed to examine the difference of SVM likelihood between benign and malignant lymph nodes. Then the five-fold cross validation was used to verify the feature extraction and classification model. The entire dataset was randomly divided into five equal-sized subsets, four of which were used for training and the remaining one was used for testing. This process was repeated five times, and each time one different subset was used as a test set. The five-fold cross validation was again randomly repeated five times, taking the average values from 25 test sets as the final results.

In order to quantitatively evaluate the classification performance, the following indices were calculated:

TABLE 1. Results of eight models based on the point-wise gated Boltzmann machine (PGBM) for classification of cervical lymph nodes (unit: %).

	ACC	PRE	SEN	SPC	YI
PGBM	82.68±0.49	83.00±1.51	94.32±1.82	56.41±5.47	50.73±3.70
PGBM-SMOTE	82.55±1.40	89.58±1.88	84.75±2.76	77.56±4.92	62.32±3.61
PGBM-EC	72.07±1.94	74.48±0.38	88.18±0.91	31.79±2.05	19.98±1.14
PGBM-SMOTE-EC	77.80±0.32	82.72±0.15	85.91±0.91	59.49±1.03	45.40±0.12
PGBM-RBM	82.36±0.63	80.59±0.85	98.18±1.36	46.67±1.02	44.85±0.34
PGBM-RBM-SMOTE	74.49±0.39	93.03±2.59	68.41±1.67	88.21±5.02	56.61±3.35
PGBM-RBM-EC	65.04±2.20	72.31±0.83	80.23±3.18	30.77±0.00	10.99±3.18
PGBM-RBM-SMOTE-EC	62.99±6.22	79.86±3.07	62.05±8.99	65.13±4.76	27.17±9.63

accuracy (ACC), precision (PRE), sensitivity (SEN), specificity (SPE), and Youden's index (YI). In addition, we summarized the numbers of cases correctly and wrongly classified, and we performed χ^2-test to compare two classification models.

2) COMPARISON BETWEEN DIFFERENT MODELS

Our final AI architecture (i.e., model) for feature extraction and classification using both PGBM and SMOTE was named PGBM-SMOTE. The PGBM-SMOTE model was compared with the model only using PGBM but without SMOTE, namely the PGBM model. The PGBM-SMOTE and PGBM were both shallow ANNs which only had a single layer of networks. In our previous work, we proposed using a deep ANN model named the PGBM-RBM model, which had two layers of networks, for classification of breast tumors on sonoelastography [11]. Here, the PGBM-SMOTE model was also compared with the PGBM-RBM model and its variant combining with SMOTE (PGBM-RBM-SMOTE).

We further compared the features learned and selected by using our AI architecture with the SFs on classification of lymph nodes. The SFs included the TIC features, gray level statistical features, gray level co-occurrence matrix features, and binary image texture features [11], [20], [21]. The TIC features consisted of the peak intensity, enhanced intensity, mean transit time, time to peak, etc. [8], [20]. The gray level statistical features included the mean, median, standard deviation, entropy and several percentiles of the gray levels within a lymph node on the TMI [20]. The gray level co-occurrence matrix features included contrast, energy, entropy and homogeneity [21], [24]. The binary image texture features included center deviation degree, dispersion degree, etc [20]. In total, 312 SFs were calculated. Then we performed two schemes regarding feature selection of SFs: (a) the least absoulute shrinkage and selection operator (LASSO), and (b) t-test [11]. Thus, two SF models for feature extraction and classification were derived and named SF-LASSO and SF-T, respectively.

Adding an error cost (EC) in the SVM classifier may contribute to enhanced performance for unbalanced data [23]. For comparison purposes in this study, we set a larger EC for the minority class in order to reduce the influence of the unbalanced data. The PGBM model modified with adding an EC term was named PGBM-EC.

In total, there are 16 feature extraction and classification models for comparison, named as "XX-YY". Here, "XX" stands for PGBM, PGBM-RBM, SF-LASSO, or SF-T; "YY" stands for SMOTE, EC, both SMOTE and EC ("SMOTE-EC"), or none of SMOTE and EC (only "XX").

III. RESULTS

A. CHARACTERISTICS OF CERVICAL LYMPH NODES

Among 39 benign lymph nodes, 35 were reactive hyperplasia and 4 were lymph node tuberculosis. Among 88 malignant lymph nodes, 13 were lymphoma, and 57 were metastasis from thyroid cancer, 8 from lung cancer, 4 from esophageal cancer, 3 from breast cancer, 1 from nasopharyngeal cancer, 1 from ovarian cancer, and 1 from mesothelioma.

The SVM likelihood in benign cervical lymph nodes was significantly different from that in malignant nodes (0.74 ± 0.21 vs. 0.33 ± 0.28, $p < 0.001$).

B. CLASSIFICATION RESULTS OF PGBM-BASED MODELS

Table 1 shows classification results of eight PGBM-based models on the test sets. According to the YI-value, the PGBM-SMOTE performed best among eight models (62.32%), and its ACC, PRE, SEN, and SPE were 82.55%, 89.58%, 84.75%, and 77.56%. From Table 1, we can also see that the four single-layer PGBM models (i.e., the shallow networks) were all better than the corresponding two-layer PGBM-RBM model (i.e., the deep networks) in terms of ACC and YI.

C. CLASSIFICATION RESULTS OF STATISTICAL FEATURES

The classification results of SF models on the test sets are enumerated in Table 2. The SF-LASSO-SMOTE achieved

TABLE 2. Results of eight models based on statistical features for classification of cervical lymph nodes (unit: %).

	ACC	PRE	SEN	SPC	YI
SF-T	68.50±3.73	73.33±1.96	85.68±4.47	29.74±5.28	15.42±7.47
SF-T-SMOTE	65.51±2.51	73.02±2.56	80.00±4.10	32.82±10.56	12.82±8.39
SF-T-EC	69.13±2.01	78.97±1.99	75.68±2.65	54.36±5.71	30.04±5.33
SF-T-SMOTE-EC	66.77±2.08	72.35±1.62	84.32±1.33	27.18±5.52	11.50±5.92
SF-LASSO	73.23±2.17	77.62±2.10	86.36±2.87	43.59±7.25	29.95±6.41
SF-LASSO-SMOTE	73.53±0.81	78.73±1.31	83.77±1.54	48.21±4.70	31.98±3.51
SF-LASSO-EC	68.03±3.36	78.06±2.85	75.00±2.24	52.31±6.99	27.31±8.56
SF-LASSO-SMOTE-EC	71.97±2.52	78.27±1.93	82.50±3.01	48.21±5.71	30.71±6.05

the best performance among eight models in terms of YI (32.98%). Its ACC, PRE, SEN, and SPE were 73.53%, 78.73%, 84.77%, and 48.21%. The SF-LASSO-SMOTE was inferior to the PGBM-SMOTE. From the experimental results, we can also see that the SF-LASSO models were generally better than the SF-T models. In addition, the SEN and SPC of the SF models had a large difference even they used SMOTE or EC methods for compensation of the data unbalance.

D. COMPARISON BETWEEN NUMBERS OF CASES CORRECTLY OR WRONGLY DIAGNOSED

Table 3 gives the numbers of cases, in one experiment of five-fold cross validation, correctly or wrongly diagnosed by three typical models, namely the PGBM-SMOTE, PGBM and SF-LASSO-SMOTE. There were 78 cases correctly classified by all of the three models; the typical examples are shown in Fig. 3a. There were 16 cases correctly classified by both the PGBM-SMOTE and PGBM but misclassified by the SF-LASSO-SMOTE (Fig. 3b). In addition, there were 7 cases correctly classified by the PGBM-SMOTE but misclassified by both the PGBM and SF-LASSO-SMOTE (Fig. 3c). The PGBM-SMOTE correctly classified more cases than SF-LASSO-SMOTE (104 correct: 23 wrong vs. 93:34, $p = 0.098$).

PGBM-SMOTE and PGBM had similar classification results (104:23 vs. 107:20, $p = 0.616$). Table 4 further gives the numbers of cases correctly or wrongly diagnosed by the PGBM-SMOTE and the PGBM. In benignancy, there were 10 cases correctly classified by the PGBM-SMOTE but misclassified by PGBM. Meanwhile, there was only 1 case correctly classified by the PGBM but misclassified by PGBM-SMOTE. In malignancy, there were 12 cases correctly classified by PGBM but misclassified by PGBM-SMOTE. Meanwhile the number of cases correctly classified by PGBM-SMOTE but misclassified by PGBM was 0. The classification results of the PGBM on the benign and malignant subsets were very different (22:17 vs. 85:3, $p < 0.001$). In contrast, the results of PGBM-SMOTE were relatively balanced between benignancy and malignancy (31:8 vs. 73:15, $p = 0.641$). These detailed results on benign and malignant subsets showed that compared with the PGBM, the PGBM-SMOTE increases the numbers of correctly classified in benignancy but decreases them in malignancy. It indicated that the PGBM-SMOTE elevated the specificity while sacrificed in the sensitivity, and thus made the sensitivity and specificity balanced for such an unbalanced dataset consisting of much more malignant nodes than benign.

TABLE 3. Numbers of cases correctly or wrongly diagnosed with three feature extraction and classification methods.

PGBM-SMOTE	PGBM	SF-LASSO-SMOTE	Case No.
Correct	Correct	Correct	78
Correct	Correct	Wrong	16
Correct	Wrong	Correct	3
Correct	Wrong	Wrong	7
Wrong	Correct	Correct	6
Wrong	Correct	Wrong	7
Wrong	Wrong	Correct	6
Wrong	Wrong	Wrong	4

IV. DISCUSSION

The main contribution of this study is proposing an end-to-end AI architecture for learning and selecting image features from CEUS videos and for automatic classification of cervical lymph nodes. The AI architecture has elevated classification ability for the unbalanced dataset compared with the traditional SF methods. In future clinical diagnosis, our AI architecture could be used as a valuable tool for distinguishing malignant and benign cervical lymph nodes.

Our AI architecture was constructed based on the PGBM. The switch units in the PGBM allowed the model to estimate for each TMI of CEUS where the task-relevant patterns occurred and to make only those relevant visible units

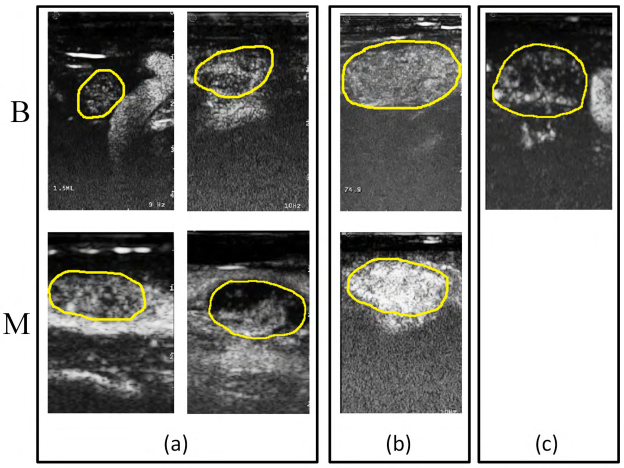


FIGURE 3. Typtical samples of benign (B, top) and maligant (M, bottom) cervical lymph nodes that were correctly classified with our PGBM-SMOTE model. The lymph nodes shown in (a) were also correctly classified with the PGBM and the SF-LASSO-SMOTE, the nodes shown in (b) were correctly classified with the PGBM but misclassified with SF-LASSO-SMOTE, and the nodes shown in (c) were misclassified with both PGBM and SF-LASSO-SMOTE.

TABLE 4. The numbers of cases correctly and wrongly classified in benign and malignant lymph nodes with the PGBM-SMOTE and the PGBM models.

	PGBM-SMOTE	PGBM	Case No.
Benign	Correct	Correct	21
	Correct	Wrong	10
	Wrong	Correct	1
	Wrong	Wrong	7
Malignant	Correct	Correct	73
	Correct	Wrong	0
	Wrong	Correct	12
	Wrong	Wrong	3

(pixels) to contribute to the final classification. The architecture ignored the task-irrelevant portion of the visible units, and thus it performed dynamic feature selection, namely choosing a variable subset of the raw features depending on adaptive interpretation of an individual image.

From Tables 1 and 2, we can see that the SMOTE and EC were both helpful for improving the classification performance of the unbalanced data, but the combination of the two techniques did not perform well. In addition, for the PGBM models, the SMOTE performed better than EC, while for the SF models, the EC seemed to yield better results than SMOTE. This phenomenon might indicate that the SMOTE and EC were suitable for different feature spaces. The optimal solution for classification of unbalanced data still needs to be investigated in the future.

It can be seen from Tables 1 and 2 that three PGBM-based models, namely PGBM, PGBM-SMOTE and PGBM-SMOTE-EC, achieved better results than the corresponding SF-based models in terms of all five indices. For instance, compared with the SF-LASSO-SMOTE, the PGBM-SMOTE increased the ACC, PRE, SEN, SPC, and YI by 9.02%, 10.85%, 0.98%, 29.35% and 30.34% respectively. These results indicated that the task-relevant component in the PGBM could capture the intrinsic differences in a CEUS video between benign and malignant nodes and thus improve the classification performance, while the SF models were disturbed by the task-irrelevant patterns in the video. In addition, the SF generation required prior expert knowledge of cervical lymph nodes, which complicated the diagnosis process. Our AI architecture needed little prior knowledge, and hence it could be more proper and convenient for future clinical diagnosis.

The single-layer PGBM models were better than the two-layer PGBM-RBM models (Table 2), especially when SMOTE was applied, implying that the shallow networks outperformed the deep networks (i.e., deep learning) for cervical lymph node classification on CEUS. Since the medical image datasets are usually in a small size, deep networks on such small datasets may tend to overfit the data. A recent study on music genre classification also demonstrated that shallow networks were better suited for small datasets than deep networks [25]. In contrast, our previous work on classification of breast tumors with sonoelastography showed superior performance of deep networks to shallow networks [11]. Therefore, whether deep or shallow models are appropriate for a small medical dataset still needs to be further investigated and the answer may depend on specific applications.

In the PGBM-based models, the classification results with different numbers of hidden units were similar. For instance, when we varied the unit number from 10 to 500 in an interval of 10, the classification accuracy only changed by 4.0%. It showed the robustness of the proposed architecture to this parameter. Finally we set the number as 80 to achieve best performance, where the numbers of relevant and irrelevant units were both 40.

There are several limitations and directions for future work. First, multiple modalities, including conventional ultrasound, sonoelastography, magnetic resonance imaging, computed tomography and positron emission tomography, are expected to be incorporated in the AI architecture for better differentiation of lymphadenopathy. Second, this study focuses on binary classification of benign and malignant lymph nodes. There are many sub-classes in benign and malignant nodes, and the detailed classification of nodal sub-classes is helpful for optimal therapeutic planning and deserves future investigation. Third, in addition to the differential diagnosis of cervical lymph nodes, the proposed AI architecture is expected to be easily applied to the diagnosis of other diseases with little modification but retraining on samples of the target diseases, and then a new network will be constructed for the new application [11]. Fourth, the AI architecture for diagnosis will be upgraded for evaluation of treatment response in a future study, for example, for timely identification of head

and neck cancer patients at risk of residual malignant lymph nodes after chemo-radiotherapy [24].

In conclusion, we construct an AI architecture by using the PGBM and SMOTE for differentiating malignant from benign cervical lymph nodes on CEUS. The experimental results show that the shallow PGBM-SMOTE model has better classification performance than the SF models and other PGBM-based models including the deep models. Our AI architecture could be potentially valuable for the future diagnosis and clinical management of lymphadenopathy.

REFERENCES

[1] C. P. Giacomini, R. B. Jeffrey, and L. K. Shin, "Ultrasonographic evaluation of malignant and normal cervical lymph nodes," *Seminars Ultrasound, CT MRI*, vol. 34, no. 3, pp. 236–247, Jun. 2013.

[2] R. Slaisova, K. Benda, J. Jarkovsky, H. Petrasova, P. Szturz, and V. Valek, "Contrast-enhanced ultrasonography compared to gray-scale and power Doppler in the diagnosis of peripheral lymphadenopathy," *Eur. J. Radiol.*, vol. 82, no. 4, pp. 693–698, Apr. 2013.

[3] L. Poanta, O. Serban, I. Pascu, S. Pop, M. Cosgarea, and D. Fodor, "The place of CEUS in distinguishing benign from malignant cervical lymph nodes: A prospective study," *Med. Ultrasonography*, vol. 16, no. 1, pp. 7–14, Mar. 2014.

[4] X.-W. Cui, C. Jenssen, A. Saftoiu, A. Ignee, and C. F. Dietrich, "New ultrasound techniques for lymph node evaluation," *World J. Gastroenterology*, vol. 19, no. 30, pp. 4850–4860, Aug. 2013.

[5] L. Rubaltelli, V. Beltrame, A. Tregnaghi, E. Scagliori, A. C. Frigo, and R. Stramare, "Contrast-enhanced ultrasound for characterizing lymph nodes with focal cortical thickening in patients with cutaneous melanoma," *Amer. J. Roentgenology*, vol. 196, no. 1, pp. W8–W12, Jan. 2011.

[6] J.-B. Liu *et al.*, "Contrast-enhanced sonography for detection of secondary lymph nodes in a melanoma tumor animal model," *J. Ultrasound Med.*, vol. 33, no. 6, pp. 939–947, Jun. 2014.

[7] T. Sato *et al.*, "Monitoring of blood vessel density using contrast-enhanced high frequency ultrasound may facilitate early diagnosis of lymph node metastasis," *J. Cancer*, vol. 8, no. 5, pp. 704–715, 2017.

[8] Q. Zhang *et al.*, "Evaluating pathologic response of breast cancer to neoadjuvant chemotherapy with computer-extracted features from contrast-enhanced ultrasound videos," *Phys. Medica*, vol. 39, pp. 156–163, Jul. 2017.

[9] Q. Huang, X. Huang, L. Liu, Y. Lin, X. Long, and X. Li, "A case-oriented Web-based training system for breast cancer diagnosis," *Comput. Methods Programs Biomed.*, vol. 156, pp. 73–83, Mar. 2018.

[10] X. Feng, X. Guo, and Q. Huang, "Systematic evaluation on speckle suppression methods in examination of ultrasound breast images," *Appl. Sci.*, vol. 7, no. 1, p. 37, 2017.

[11] Q. Zhang *et al.*, "Deep learning based classification of breast tumors with shear-wave elastography," *Ultrasonics*, vol. 72, pp. 150–157, Dec. 2016.

[12] Q. Zhang, J. Suo, W. Chang, J. Shi, and M. Chen, "Dual-modal computer-assisted evaluation of axillary lymph node metastasis in breast cancer patients on both real-time elastography and B-mode ultrasound," *Eur. J. Radiol.*, vol. 95, pp. 66–74, Oct. 2017.

[13] J. Zhang, Y. Wang, B. Yu, X. Shi, and Y. Zhang, "Application of computer-aided diagnosis to the sonographic evaluation of cervical lymph nodes," *Ultrason. Imag.*, vol. 38, no. 2, pp. 159–171, Mar. 2016.

[14] J. Lam *et al.*, "A comparison of the diagnostic accuracy and reliability of subjective grading and computer-aided assessment of intranodal vascularity in differentiating metastatic and reactive cervical lymphadenopathy," *Ultraschall Med.*, vol. 37, no. 1, pp. 63–67, Feb. 2016.

[15] A. Chmielewski, P. Dufort, and A. M. Scaranelo, "A computerized system to assess axillary lymph node malignancy from sonographic images," *Ultrasound Med. Biol.*, vol. 41, no. 10, pp. 2690–2699, Oct. 2015.

[16] M. A. Maraci, R. Napolitano, A. Papageorghiou, and J. A. Noble, "Object classification in an ultrasound video using LP-SIFT features," in *Proc. Int. MICCAI Workshop Med. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 71–81.

[17] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[18] S. Han *et al.*, "A deep learning framework for supporting the classification of breast lesions in ultrasound images," *Phys. Med. Biol.*, vol. 62, no. 19, p. 7714, 2017.

[19] K. Sohn, G. Zhou, C. Lee, and H. Lee, "Learning and selecting features jointly with point-wise gated Boltzmann machines," in *Proc. The 30th Int. Conf. Mach. Learn.*, 2013, pp. 1–9.

[20] Q. Zhang *et al.*, "Spatio-temporal quantification of carotid plaque neovascularization on contrast enhanced ultrasound: Correlation with visual grading and histopathology," *Eur. J. Vascular Endovascular Surg.*, vol. 50, no. 3, pp. 289–296, Sep. 2015.

[21] Q. Zhang, C. Li, H. Han, L. Yang, Y. Wang, and W. Wang, "Computer-aided quantification of contrast agent spatial distribution within atherosclerotic plaque in contrast-enhanced ultrasound image sequences," *Biomed. Signal Process. Control*, vol. 13, pp. 50–61, Sep. 2014.

[22] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[23] M. A. H. Farquad and I. Bose, "Preprocessing unbalanced data using support vector machine," *Decis. Support Syst.*, vol. 53, no. 1, pp. 226–233, Apr. 2012.

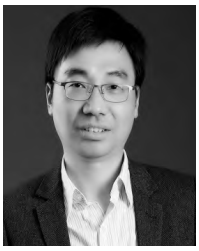
[24] E. Scalco, S. Marzi, G. Sanguineti, A. Vidiri, and G. Rizzo, "Characterization of cervical lymph-nodes using a multi-parametric and multi-modal approach for an early prediction of tumor response to chemo-radiotherapy," *Phys. Medica*, vol. 32, no. 12, pp. 1672–1680, Dec. 2016.

[25] A. Schindler, T. Lidy, and A. Rauber, "Comparing shallow versus deep neural network architectures for automatic music genre classification," in *Proc. 9th Forum Media Technol.*, Sankt Pölten, Austria, 2016, pp. 17–21.

QI ZHANG received the B.S. degree in electronic engineering and the Ph.D. degree in biomedical engineering from Fudan University, China, in 2005 and 2010, respectively. From 2008 to 2009, he was a Visiting Ph.D. Student with the Department of Biomedical Engineering, Duke University, USA, for one year. He joined the Institute of Biomedical Engineering, Shanghai University, China, in 2010, where he has been an Associate Professor since 2013. He has published over 70 articles. His research interests include biomedical image analysis and artificially intelligent diagnosis and treatment.

YUE LIU received the bachelor's degree from the Shanghai Institute of Technology in 2010. He is currently pursuing the master's degree with Shanghai University. His research interests include image processing and machine learning.

HONG HAN received the B.S. and Ph.D. degrees in imaging medicine both from Zhongshan Hospital, Fudan University, China, in 2002 and 2009, respectively. In 2012, she was a Visiting Scholar at the Department of Radiology, Keck Hospital, University of Southern California, USA. Since 2014, she has been an Associate Professor. She has published over 20 articles. Her research interests include diagnosis of lesions in abdominal organs and small organs with advanced technology of ultrasonography and diagnosis of vascular complications in transplanted liver with contrast enhanced ultrasonography.



JUN SHI received the B.S. and Ph.D. degrees from the Department of Electronic Engineering and Information Science, University of Science and Technology of China, in 2000 and 2005, respectively. In 2005, he joined the School of Communication and Information Engineering, Shanghai University, China, where he has been a Professor since 2015. From 2011 to 2012, he was a Visiting Scholar with the University of North Carolina at Chapel Hill. His current research interests include machine learning in medical imaging.

WENPING WANG is currently the Director of the Department of Ultrasound, Zhongshan Hospital, Fudan University. He is also the Deputy Director with the Institute of Medicine and Engineering, Fudan University. He has conducted the clinical and the basic research work in tumors for over 20 years. He has made contributions to early diagnosis and differential diagnosis in tumors with kinds of advanced ultrasonic technologies. He has published nearly 100 papers as the first author or corresponding author. He has also published over 10 books on various ultrasonic research fields.