Few-shot decision tree for diagnosis of ultrasound breast tumor using BI-RADS features

Article *in* Multimedia Tools and Applications · November 2018

DOI: 10.1007/s11042-018-6026-1

3 authors, including:

Qinghua Huang
Northwestern Polytechnical University
198 PUBLICATIONS **4,949** CITATIONS

SEE PROFILE

Few-shot Decision Tree for Diagnosis of Ultrasound Breast Tumor Using BI-RADS Features

Qinghua Huang[1,2,3], Fan Zhang[2], Xuelong Li[4]

[1]College of Information Engineering, Shenzhen University, Shenzhen 518060, China.

[2] School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China.

[3]School of Electronics and Information, and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072,Shaanxi, PR China.

[4]Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China

Abstract: This paper proposes an ultrasound breast tumor CAD system based on BI-RADS features scoring and decision tree algorithm. Because of the difficulty of biopsy label collection, the proposed system adopts a few-shot learning method. The SVM classifier is employed to preliminarily mark the unlabeled cases firstly. Then these unlabeled cases with the pseudo labels are combined with the few real-labeled cases to train the decision tree. To test the performance of the proposed method, 1255 ultrasound breast images were collected, and three well-experienced clinicians and three interns evaluated these images according to the BI-RADS scoring scheme. All of the images are transformed into vectors such that the algorithm can process. The experimental results show that the system performance improves significantly with the help of pseudo-labeled data. Compared to the decision tree trained by the real-labeled cases only, when the number of real-labeled cases was 40, the accuracy, specificity, sensitivity of the proposed system were increased by 4.05%, 0.09%, and 5.26%, respectively; the positive predictive value (PPV) and the negative predictive value (NVP) were increased by 0.41% and 9.77%, respectively. Meanwhile, the performance of the proposed method was the same as the method using sufficient samples. When the number of the labeled cases reached 100, the accuracy, specificity, sensitivity, PPV and NVP of the proposed method were 90.07%, 76.72%, 94.27%, 92.83%, and 81.36%, respectively. The results demonstrate that our method can efficiently distinguish the malignant tumor although the labeled data is not sufficient.

Keywords: Breast tumors, CAD system, Few-shot learning, BI-RADS, Decision tree

Corresponding email address: qhhuang@scut.edu.cn,

1 Introduction

The breast cancer is one of the largest menaces to the health of women. Thousands of women suffer and die for the breast cancer every year. According to the statistics of US women in 2013, there were around 232340 new breast cancer cases found, and up to 39620 women dead for breast cancer [1]. The reason for the breast cancer is uncertain at present. The major way to decrease the death rate of the breast cancer is the early detection [2].

Mammography and ultrasonography are two of the most commonly used technology in the early detection of the breast lesion. However, compared to ultrasonography, there are some limitations when applying mammography to examine the breast cancer. First, the accuracy of mammography examination is not satisfactory. According to [3], more than 65% women who take the mammography examinations suffer the unnecessary biopsies. Second, the performance of mammography examinations in the dense breast lesions is not adequately acceptable. It is not suitable for mammography to detect the breast lesion of young women [4]. Third, the radial ray in mammography brings additional health risk for patients. In contrast, the ultrasonography is safer and more accurate than mammography [5]. Ultrasonography is a convenient and economical scheme to detect breast lesion, which has been widely applied in the early tumor detection [6-9]. As an important application of ultrasonography, thousands of computer aided diagnostic (CAD) systems were proposed to help the radiologists and the clinicians improve the accuracy of diagnosis. Su et al. [10] suggested a breast tumor CAD system using the Affinity Propagation (AP) clustering and achieved a good performance. Zhou et al. [11] used the shearlet-based texture feature and support machine vector (SVM) to classify the breast tumor. The accuracy of classification reached 91%. Prabusankarlal et al. [12] assessed the performance of textual features and morphological features in the breast mass diagnosis. Their study showed that the accuracy of SVM classifier using gray level co-occurrence matrix (GLCM) feature obtained an accuracy of 90.83% and the accuracy of SVM classifier combining textual features and morphological features reached 95.83%. It can be seen that most of these previous breast tumor CAD systems are based on the textural features and morphological features. Although the method applying these features can have a fair performance, there are still some limitations in the clinical practices.

The first problem is “semantic gap” between human and the intelligent algorithms. The “semantic gap” is a long-standing obstacle on the way of pattern recognition applications [13-15]. Most of the previous CAD systems employed gray features and textural features to recognize the breast tumor. The gray feature and the textural feature are effective and classical in the area of pattern recognition. However, these features are difficult to understand by the human. For most of the clinicians, those CAD systems are “black boxes”. They cannot understand the reason of the diagnosis result given by the computer. The diagnosis rules are hidden for the doctors. It is hard for the doctor to map the gray feature and the textural feature into the clinical decisions. This problem leads to the distrust in the clinical practice and limits the applications of those previously proposed CAD systems.

The second problem is the image preprocessing. The image preprocessing is an important step in

improving the image quality and extracting the image features, e.g. the gray feature, the textural feature and the morphological feature. These preprocessing operations like denoising and saliency detection play an indispensable role in image classification [16]. The better the result of image preprocessing is, the better the classification accuracy will be. However, the result of preprocessing heavily depends on the parameters set by users, and it is hard to find universal parameters for all images. Therefore, most of the CAD systems can only recognize the breast tumor images from a specific ultrasonic machine. In clinical practice, however, the query image may be collected from different ultrasonic machines. The lack of data commonality limits the application of those previously reported ultrasound CAD systems.

The third problem is that the CAD systems were designed based on the sufficient training data. However, as for the breast ultrasound CAD system, the shortage of labeled data is often a problem. It is difficult to acquire the biopsy label timely, and not all of the patients are required to carry out the biopsy in the clinical practice. There are many ultrasound breast tumor images only with the Breast Imaging Reporting and Data System (BI-RADS) report. In the other hand, the process of biopsy usually takes more than two days. Many patients cannot receive their biopsy results immediately. As a result, the number of cases with the biopsy labels is often small, while the number of cases without the biopsy labels is relatively large in clinical practices. It is difficult for the CAD system to learn adequately smart diagnostic rules with few cases with the biopsy labels. Consequently, the Semi-Supervised Learning (SSL) method is a way to address this problem [17]. The method utilizes unlabeled samples to improve the performance of few-shot learning system. In 2013, Kim et al. [18] attempted to use the SSL to predict the breast cancer survivability. In 2015, Tu et al. [19] proposed the Posterior Distribution Learning (PDL) framework that used the unlabeled samples to improve the classification performance. In 2016, Liu et al. [20] employed the SSL method to predict the customer behavior. The success of the previous research has justified the effectiveness of this method. However, to the best of our knowledge, no success of the SSL on diagnosis of breast tumors in ultrasound images was reported. In this paper, a pseudo-labeled scheme based on the SVM is designed to address this problem.

It is worth noting that in clinical practices, doctors often use the BI-RADS report to guide the diagnosis. The BI-RADS is a standard characteristic description of breast tumor images proposed by American College of Radiology [21]. In clinical examinations, the radiologist is asked to allocate a BI-RADS category for each ultrasound breast tumor image according to the BI-RADS lexicon to evaluate the danger of the tumor. The features in the BI-RADS lexicon contain sufficient information which can be useful in diagnosis of breast masses. Shen et al. [22] attempted to extract the BI-RADS features from image automatically in 2007. The accuracy of the classifier employing the extracted features reached 91.70%. Zhang et al. [23] utilized Fourier Irregularity index to measure the irregularity of breast tumors standardized by BI-RADS lexicon and employed the index to classify the breast tumors. Shan et al. [24] compared the performance of different classifiers using the automatically extracted BI-RADS features, and the highest accuracy was 78.5%. However, automatically extracting BI-RADS features cannot guarantee the accuracy of classification because of the first and second problems mentioned above.

This paper proposes a new ultrasound CAD system using BI-RADS features to address the first two problems and designed a pseudo-labeled scheme to address the third problem. Firstly, a series of BI-RADS features are assessed by the clinicians in the process of image inspection. The breast image is then transformed into a vector, which is subsequently classified by the decision tree algorithm. In this way, the problem of "semantic gap" between the CAD system and clinicians can be overcome. Meanwhile, the scores of BI-RADS features marked by the clinician are independent of the image preprocessing, and hence the proposed method can be widely used with different ultrasound machines. To address the shortage of labeled data, we make use of the SVM to mark the unlabeled cases. Each unlabeled case is assigned with a pseudo label. The decision tree is thereafter trained on the mixed dataset including the pseudo-labeled cases and the cases with real labels. In this approach, the classification information hidden in the unlabeled data can be discovered, and the performance of decision tree can be improved.

2 Methods

There are four steps in the proposed CAD method. In the first step, the breast image is evaluated and transformed into a vector by clinicians according to BI-RADS feature scoring scheme. In the second step, SVM is applied to mark the unlabeled data weakly, and the pseudo-labeled data set is built. In the third step, the decision tree is trained with the pseudo-labeled data set and real-labeled data set. Finally, the trained decision tree is applied to classify the query cases.

2.1 BI-RADS feature scoring scheme

In proposed system, a manually scoring scheme is designed to quantify the BI-RADS features. As shown in Table 1, all of the attributes in the scheme are designed by experienced clinicians according to the ACR BI-RADS–US Lexicon Classification Form [25]. Given a breast ultrasound image, each BI-RADS feature is assigned with a score ranging from 0 to 5. A score associated with a BI-RADS feature reflected the subjective visual assessment by the clinician.

Table 1 The feature scoring standard scheme

Feature	0	1	2	3	4	5
Shape	Oval	Round	Irregular	N/A	N/A	N/A
Orientation	Parallel	Not parallel	N/A	N/A	N/A	N/A
Margin integrality	Circumscribed	Not circumscribed	N/A	N/A	N/A	N/A
Margin ambiguity	Distinct	Indistinct	N/A	N/A	N/A	N/A
Angular	Absent	Present	N/A	N/A	N/A	N/A
Microlobulated	Absent	Present	N/A	N/A	N/A	N/A
Spiculated	Absent	Present	N/A	N/A	N/A	N/A
Echo pattern	Anechoic	Hyperechoic	Isoechoic	Heterogeneous	Complex cystic &	Hypoechoic

					solid	
Posterior feature	Enhancement	None	Combined pattern	Shadowing	N/A	N/A
Calcification in mass	Absent	Coarse calcification	Microcalcifi-cation	N/A	N/A	N/A
Architectural distortion	Absent	Present	N/A	N/A	N/A	N/A
Ducts changes	Normal	Cystic expansion	Object found in ducts	N/A	N/A	N/A
Skin thickening	Absent	Present	N/A	N/A	N/A	N/A
Skin retraction	Absent	Present	N/A	N/A	N/A	N/A
Edema	Absent	Present	N/A	N/A	N/A	N/A
Vascularity	Absent	Vessels in rim	Internal vascularity	N/A	N/A	N/A
Elasticity	Soft	Intermediate	Hard	N/A	N/A	N/A
Cyst type	Simple cyst	Clustered micro cysts	Complicated cyst	N/A	N/A	N/A
Skin mass	Absent	Present	N/A	N/A	N/A	N/A
Foreign body	Absent	Present	N/A	N/A	N/A	N/A
Lymph nodes-intramammary	Absent	Present	N/A	N/A	N/A	N/A
Lymph nodes-axillary	Absent	Reactive	Metastatic	N/A	N/A	N/A
Vascular abnormalities	Absent	Present	N/A	N/A	N/A	N/A
Postsurgical fluid collection	Unoperated	Absent	Present	N/A	N/A	N/A
Fat necrosis	Absent	Present	N/A	N/A	N/A	N/A

2.2 Pseudo-labeled scheme based on SVM

When there are few cases with real biopsy labels, we employed the SVM to mark the unlabeled samples. The SVM is a powerful classifier when the training data is little. We used the labeled dataset as the training set to build the SVM decision function as below,

$$f(x)=\sum_{i=1}^{n} y_i \alpha_i K(x_i, x)+b \tag{1}$$

where, $x_i \in R^d$, is the training vectors, and $y_i \in \{-1,1\}^n$ is the label of each training case. α_i is the Lagrange multiplier, and n is the total number of training vectors. $K(\bullet)$ is the kernel function, and b is an independent term. We adopt the Radial Basis Function (RBF) kernel function defined as:

$$K(x_i, x)=\exp(-\gamma \|x-x_i\|^2) \tag{2}$$

where, γ is defined as:

$$\gamma=\frac{1}{d} \tag{3}$$

The decision function $f(x)$ reflects the distance between input sample and the decision hyperplane. In general, the sample which is farther to the decision hyperplane is more likely to be classified correctly. From this perspective, we designed the pseudo-labeled scheme as follows:

$$y(x) = \begin{cases} 1, f(x) > \lambda_p \\ -1, f(x) < \lambda_n \end{cases} \quad (4)$$

where, λ_p is the threshold for the sample which is above the decision hyperplane, and λ_n is the threshold for the sample which is located below the decision hyperplane. Only the samples whose distances to the decision hyperplane are larger than the threshold can be regarded as the convinced pseudo-labeled samples. In this approach, the unlabeled data is weakly marked by the SVM classifier. In the experiment, we define the median of the distance from above samples to the decision hyperplane as λ_p, and the median of the distance from below samples to the hyperplane as λ_n.

2.3 Decision tree for tumor classification

The proposed system uses the decision tree as the classifier to distinguish the breast tumor. The decision tree is a powerful algorithm for classification. The rule learned by a decision tree is easy to understand. Figure 1 shows an example of a decision tree used in the tumor classification in this study. From the root node to the leave node, every node is assigned with a BI-RADS feature. In the leave node, the category of the testing sample is given. There are many decision tree algorithms proposed in the past decades, e.g. ID3 [26], C4.5 [27], C5.0 [28], and CART [29]. In this study, we use the Classification and Regression Tree (CART) algorithm.

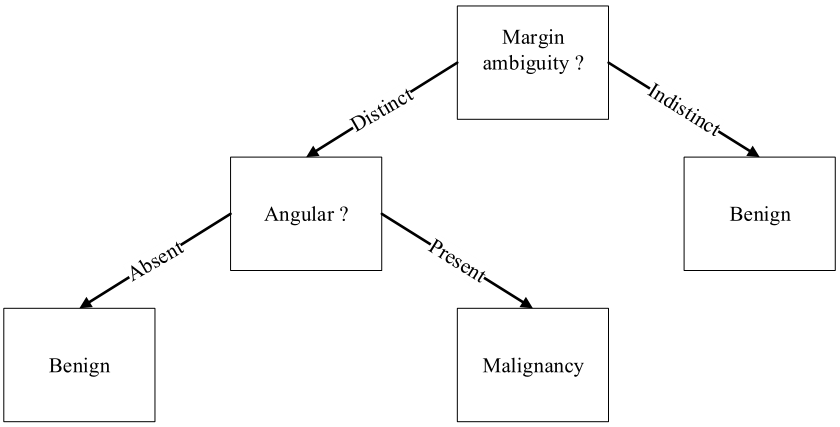

Fig. 1. An example of a decision tree for classifying tumors. Explanatory variables are margin ambiguity and angular.

2.4 Diagnostic system

Figure 2 shows the flow chart of the proposed CAD system. The ultrasound image is scored by the clinicians firstly. Then, the unlabeled sample is weakly marked with a pseudo label by the SVM pseudo-labeled scheme. Later, the training dataset which combines the real-labeled data and pseudo-labeled data is employed to train the decision tree. When the query image is input, it is

transformed to a 25×1 vector (each element corresponds to a BI-RADS feature) by the clinician, and then the vector is classified by the trained decision tree.

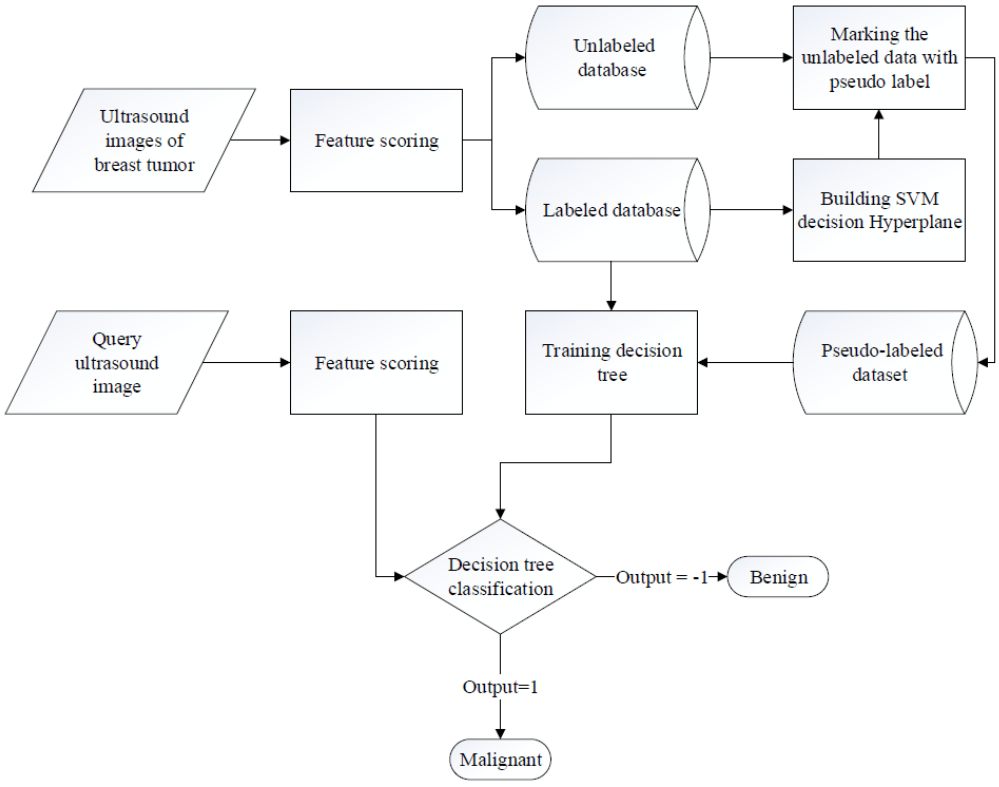

Fig. 2. The **f**lowchart of the proposed CAD system.

3 Experiments and results

3.1 Data source

The data sources from the Cancer Center of Sun Yat-sen University. A total of 1255 breast ultrasound images were collected with the help of clinicians from the Cancer Center of Sun Yat-sen University. All of the samples in the dataset have been marked with the real biopsy label. There are 964 malignant samples and 291 benign samples in the dataset. Three well-experienced clinicians and three interns were involved to evaluate the samples. Each sample was transformed into a 25×1 vector such that the computer can deal with it.

3.2 Data filtering

Although there are 1255 cases in the dataset, not all of them are appropriate for the experiment. It is unnecessary for the case which belongs to BI-RADS category 0, 1 or 6 to be classified by our system. The clinical diagnosis for those cases is sufficiently convinced. After removing those cases falling into BI-RADS category 0, 1 and 6, there were 922 malignant cases and 286 benign cases in the dataset.

Table 2 shows the number of samples grouped into different BI-RADS categories in our dataset.

Table 2 The statistic of samples belonging to different BI-RADS categories

BI-RADS categories	Malignant	Benign	Total
2	1	9	10
3	8	118	126
4	406	154	560
5	507	5	512

3.3 Feature selection

In the feature scoring, we found that some features were noneffective in the decision tree algorithm. Decision tree algorithm divides the dataset according to the sample value of each feature. But there are some features under which the sample values are almost the same. These features cannot improve the performance of classification and moreover increase the computational load. We calculated the variance of the sample values for all BI-RADS features and ranked these features from high to low. Table 3 shows the variance of different features. In this study, the features whose variances were less than 0.01 were abandoned. After the feature selection, there were 19 features left in the dataset.

Table 3 The variance value of all features

Rank	Feature	Variance
1	Calcification in mass	1.5612
2	Posterior feature	1.2800
3	Echo pattern	0.9297
4	Vascularity	0.6817
5	Shape	0.5501
6	Spiculated	0.2466
7	Orientation	0.2311
8	Angular	0.2273
9	Microlobulated	0.2125
10	Architectural distortion	0.1883
11	Margin ambiguity	0.1370
12	Margin integrality	0.1285
13	Elasticity	0.0998
14	Ducts changes	0.0877
15	Skin thickening	0.0848
16	Skin retraction	0.0435
17	Edema	0.0351
18	Skin mass	0.0156

19	Cyst type	0.0147
20	Postsurgical fluid collection	0.0049
21	Vascular abnormalities	0.0017
22	Lymph nodes-intramammary	0.0008
23	Lymph nodes-axillary	0.0000
24	Foreign body	0.0000
25	Fat necrosis	0.0000

3.4 Evaluation method

In the experiment, although all of the gathered data has been labeled by biopsy, to evaluate the proposed method, the labeled data was divided into two subsets. One of them is taken as the labeled dataset which includes 300 samples evenly covering the BI-RADS categories 2-5, and the other subset was assumed to be unlabeled dataset which includes 908 samples. We evaluated the proposed method with the different numbers of training samples. The number of training samples was set to be 40, 60, 80, and 100, respectively. We carried out the experiment for 30 runs with each training sample volume. In each run, the training samples were randomly selected from labeled dataset to compose the training set for building up the SVM decision hyperplane. From the rest of labeled samples, 200 samples were randomly selected to compose the testing dataset. The sample in unlabeled dataset was marked by the pseudo-labeled scheme to produce the pseudo-labeled dataset. The size of pseudo-labeled dataset was not constant, ranging from 450 to 460 in our experiments, which depended on the position of the decision hyperplane. In each run, five indices i.e. the accuracy, specificity, sensitivity, PPV and NVP, respectively, were calculated. The final experiment result was the average value of these indices over the 30 runs. This experiment was carried on Intel(R) Core(TM) i7-4790 CPU 3.60GHz with Microsoft Windows 10 operating system.

3.5 Evaluation of Pseudo-labeled scheme

To evaluate the effectiveness of the pseudo-labeled scheme, we compared the performance of the decision tree trained using the labeled data only with the performance of the decision tree trained using the pseudo-labeled data and labeled data. The result is shown in Table 4. It can be found that after applying the pseudo labeled and labeled data to train the decision tree, the accuracy of classification was improved by more than 2.45%. Meanwhile, the NPV was increased by more than 6.61%, and the sensitivity was improved by more than 2.91%. However, the improvement of PPV and specificity is not significant. Table 5 shows the changes of the five indices with various quantities of labeled training samples. It is seen that the increased accuracy, NPV and sensitivity in percentage were reduced gradually as the number of labeled training samples was increased; meanwhile the changes of PPV and specificity grow slightly. The reason is that the accuracy, NPV and sensitivity of the decision tree which was only trained on the labeled data grew faster than those of the decision tree trained on both

labeled data and pseudo-labeled data as the size of labeled samples increased. Instead, the PPV and specificity of the decision tree trained on labeled data were quite stable.

From the Table 5, the increased accuracy, NPV and sensitivity in percentage were 4.05%, 9.97% and 5.26%, respectively, when the number of training labeled samples is 40. As for the CAD system, the improvement of the accuracy, NPV and sensitivity are generally more important than those of the PPV and specificity in clinical practices. High NPV and high sensitivity can help avoid the case that a malignant case is misclassified into the benign class. Finally, it can be concluded that the performance of classification improves significantly with the help of pseudo-labeled data.

Table 4 The performance of the system in the different volume of labeled training samples. N is the number of labeled training samples.

N	Training dataset	Accuracy	PPV	NPV	Specificity	Sensitivity
40	Pseudo-labeled data + Labeled data	88.70%	92.66%	77.80%	75.96%	92.64%
	Labeled data	84.65%	92.25%	68.04%	75.87%	87.38%
60	Pseudo-labeled data + Labeled data	89.68%	92.97%	80.21%	76.73%	93.66%
	Labeled data	87.17%	92.69%	72.36%	76.54%	90.48%
80	Pseudo-labeled data + Labeled data	89.78%	93.12%	80.18%	77.62%	93.64%
	Labeled data	87.20%	92.69%	72.98%	77.04%	90.44%
100	Pseudo-labeled data + Labeled data	90.07%	92.83%	81.36%	76.72%	94.27%
	Labeled data	87.62%	92.33%	74.75%	75.73%	91.36%

Table 5 The changes of the performance in various quantities of labeled training samples. N is the number of labeled training samples.

N	40	60	80	100
Accuracy	4.05%	2.52%	2.58%	2.45%
PPV	0.41%	0.29%	0.42%	0.50%
NPV	9.77%	7.85%	7.20%	6.61%
Specificity	0.09%	0.20%	0.58%	0.99%
Sensitivity	5.26%	3.19%	3.20%	2.91%

3.6 Performance comparison with other methods

We compared the proposed CAD system with two traditional methods [22][24], which automatically extracted the BI-RADS features from the ultrasound images and trained classifiers for breast tumor diagnoses. In our method, we used 100 labeled samples as the training dataset to learn the SVM decision

hyperplane for producing the pseudo-labeled dataset and took the remaining 200 labeled samples as the testing dataset to evaluate the diagnostic performance of decision tree. The decision tree was trained using the pseudo-labeled dataset and 100 labeled samples in the training dataset. We carried out the experiment for 30 runs. In each run, the training dataset and the testing dataset were selected randomly from the labeled dataset. Five indices i.e. the accuracy, specificity, sensitivity, PPV and NVP, respectively, were calculated in each run. The final experiment result was the average value of these indices over the 30 runs. The Table 6 shows the comparison results. It is observed that the performance of proposed system outperformed the method mentioned in [24]. Compared to the method mentioned in [22], the PPV and the sensitivity of the proposed system were higher. However, our system could not achieve significant improvement in comparison with the methods in [22] and [24]. It is because our system was trained using few labeled data; meanwhile the methods in [22] and [24] are trained by sufficient labeled data. Consequently, the advantage of the proposed method can be obviously found. It is worth noting that the NPV of proposed system was lower than that of the method in [22], and the specificity of proposed system was also not satisfactory. The reason is that the number of benign cases is much smaller than the number of malignant cases in our dataset. The imbalance of dataset will damage the performance of decision tree algorithm [30].

Table 6. The performance comparison of different systems.

Method	Accuracy	PPV	NPV	Specificity	Sensitivity
[22]	91.70%	84.62%	95.40%	92.22%	90.59%
[24]	78.50%	82.50%	74.70%	82.00%	75.30%
Proposed method	90.07%	92.83%	81.36%	76.72%	94.27%

4 Discussion and Conclusions

The proposed system utilizes the information generated in the manual BI-RADS category evaluation to classify the breast tumor. To address the difficulty of biopsy label collection, the proposed system adopted a few-shot learning method. The SVM classifier was used to weakly mark the unlabeled case, and these unlabeled cases were assigned pseudo labels according to their distance to the SVM decision hyperplane. The decision tree was trained by the pseudo-labeled data and real-labeled data. In this approach, the information of unlabeled data was dug out. According to the experiment results, the pseudo-labeled scheme improved the performance of classification immensely.

The accuracy of breast early diagnosis is significant [31]. The evaluation result shows that the accuracy of proposed system is not second to the previous systems. In particular, the PPV and sensitivity are higher than the previous system. The specificity and NPV of the proposed system are not satisfactory because of the imbalance of dataset. Nevertheless, we expect that these indices will be improved when more benign cases can be added in our further study.

Compared with the traditional ultrasonography CAD systems that adopted the low-level features

e.g. texture and gray, the proposed system is much more comprehensible for physicians. All of the features employed in the proposed system are designed by experienced clinicians according to ACR BI-RADS–US Lexicon Classification Form. Meanwhile, the diagnostic rules generated by decision tree are much easier to understand than those generated by other machine learning algorithms, e.g. the SVM and Artificial Neural Network (ANN). Clinicians can easily check the decision path if the diagnostic results given by decision tree are different to theirs. In addition to the improvement of the transparency, the proposed system is compatible for the different ultrasound machines. The image preprocessing is avoided in the proposed system. The BI-RADS features are evaluated by doctors directly. The proposed method is not limited by the scope of ultrasonic examinations. We expect that it can be more widely applied in various CAD applications in hospital.

Acknowledgements

This work was partially supported by National Natural Science Foundation of China (Nos. 61372007 and 61571193), Guangzhou Key Lab of Body Data Science (No.201605030011), Natural Science Foundation of Guangdong Province, China (Nos. 2017A030312006 and 2015A030313210), Fundamental Research Funds for the Central Universities (NO.2015ZM138), Project of Science and Technology Department of Guangdong province (2016A010101021, 2016A010101022 and 2016A010101023), and Guangdong Provincial Science and Technology Program - International Collaborative Projects (No. 2014A050503020).

References:

1. DeSantis, C., Ma, J., Bryan, L., Jemal, A.: Breast cancer statistics, 2013. CA: a cancer journal for clinicians, **64**(1), 52-62 (2014)
2. Aminikhanghahi, S., Shin, S., Wang, W., Jeon, S.I., Son, S.H.: A new fuzzy Gaussian mixture model (FGMM) based algorithm for mammography tumor image classification. Multimedia Tools & Applications, 1-15 (2016)
3. Jesneck, J.L., Lo, J.Y., Baker, J.A.: Breast mass lesions: Computer-aided diagnosis models with mammographic and sonographic descriptors 1. RADIOLOGY, **244**(2), 390-398 (2007)
4. Huang, Q., Luo, Y., Zhang, Q.: Breast ultrasound image segmentation: a survey. INT J COMPUT ASS RAD, 1-15 (2017)
5. Drukker, K., Giger, M.L., Horsch, K., Kupinski, M.A., Vyborny, C.J., Mendelson, E.B.: Computerized lesion detection on breast ultrasound. MED PHYS, **29**(7), 1438-1446 (2002)
6. M, N., V, D., M, H., M, W., F, D., H, M., S, W., US, A.: Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. BMC CANCER, **9**(1), 335 (2009)
7. Luo, Y., Liu, L., Huang, Q., Li, X.: A Novel Segmentation Approach Combining Region-and Edge-Based Information for Ultrasound Images. BIOMED RES INT, **2017** (2017)
8. Feng, X., Guo, X., Huang, Q.: Systematic Evaluation on Speckle Suppression Methods in Examination of Ultrasound Breast Images. Applied Sciences, **7**(1), 37 (2016)
9. Huang, Q., Yang, F., Liu, L., Li, X.: Automatic segmentation of breast lesions for interaction in

ultrasonic computer-aided diagnosis. INFORM SCIENCES, **314**, 293-310 (2015)
10. Su, Y., Wang, Y.: Computer-Aided Classification of Breast Tumors Using the Affinity Propagation Clustering Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on, 2010. IEEE, p 1-4. (2010)
11. Zhou, S., Shi, J., Zhu, J., Cai, Y., Wang, R.: Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image. BIOMED SIGNAL PROCES, **8**(6), 688-696 (2013)
12. Prabusankarlal, K.M., Thirumoorthy, P., Manavalan, R.: Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound. Human-centric Computing and Information Sciences, **5**(1), 1-17 (2015)
13. Han, J., Chen, C., Shao, L., Hu, X., Han, J., Liu, T.: Learning computational models of video memorability from fMRI brain imaging. IEEE T CYBERNETICS, **45**(8), 1692-1703 (2015)
14. Tang, J., Shao, L., Li, X.: Efficient dictionary learning for visual categorization. COMPUT VIS IMAGE UND, **124**, 91-98 (2014)
15. Lu, X., Li, X., Mou, L.: Semi-supervised multitask learning for scene recognition. IEEE T CYBERNETICS, **45**(9), 1967-1976 (2015)
16. Zhang, D., Han, J., Han, J., Shao, L.: Cosaliency detection based on intrasaliency prior transfer and deep intersaliency mining. IEEE T NEUR NET LEAR, **27**(6), 1163-1176 (2016)
17. Zhang, R., Nie, F., Li, X.: Semi-supervised classification via both label and side information Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on, 2017. IEEE, p 2417-2421. (2017)
18. Kim, J.: Breast cancer survivability prediction using labeled, unlabeled, and pseudo-labeled patient data. Journal of American Medical Informatics Association Jamia, **20**(4), 613 (2013)
19. Tu, E., Yang, J., Kasabov, N., Zhang, Y.: Posterior Distribution Learning (PDL): A novel supervised learning framework using unlabeled samples to improve classification performance. NEUROCOMPUTING, **157**, 173-186 (2015)
20. Liu, J., Gu, Y., Kamijo, S.: Customer behavior classification using surveillance camera for marketing. Multimedia Tools & Applications, 1-28 (2016)
21. Committee, A.C.O.R., Radiology, A.C.O.: Breast imaging reporting and data system. American College of Radiology. (1998)
22. Shen, W., Chang, R., Moon, W.K., Chou, Y., Huang, C.: Breast ultrasound computer-aided diagnosis using BI-RADS features. ACAD RADIOL, **14**(8), 928-939 (2007)
23. Zhang, G., Wang, W., Shin, S., Hruska, C.B., Son, S.H.: Fourier irregularity index: A new approach to measure tumor mass irregularity in breast mammogram images. Multimedia Tools & Applications, **74**(11), 3783-3798 (2015)
24. Shan, J., Alam, S.K., Garra, B., Zhang, Y., Ahmed, T.: Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound in medicine & biology, **42**(4), 980-988 (2016)
25. American College of Radiology: ACR BI-RADS – US Lexicon Classification Form. (2013)
26. Quinlan, J.R.: Induction of decision trees. MACH LEARN, **1**(1), 81-106 (1986)
27. Quinlan, J.R.: C4. 5: programs for machine learning. Elsevier. (2014)
28. Quinlan, R.: C5. 0: An informal tutorial. Rule Quest. (1998)
29. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and regression trees. CRC press. (1984)

30. Estabrooks, A., Jo, T., Japkowicz, N.: A multiple resampling method for learning from imbalanced data sets. COMPUT INTELL-US, **20**(1), 18-36 (2004)
31. American Cancer Society: Breast Cancer Facts & Figures 2011-2012. (2013)