

Classification of benign and malignant solid breast lesions on the ultrasound images based on the textural features: the importance of the perifocal lesion area

A.A. Kolchev¹, D.V. Pasyukov^{1,2,3}, I.A. Egoshin^{1,2}, I.V. Kliouchkin⁴, O.O. Pasyukova²

¹ Kazan (Volga region) Federal University, Ministry of Education and Science of Russian Federation, 420008, Kazan, Russia, Kremlevskaya St. 18;

² Mari State University, Ministry of Education and Science of Russian Federation, 424000, Yoshkar-Ola, Russia, Lenin square 1;

³ Kazan State Medical Academy - Branch Campus of the Federal State Budgetary Educational Institution of Further Professional Education «Russian Medical Academy of Continuous Professional Education», Ministry of Healthcare of the Russian Federation, 420012, Kazan, Russia, Butlerova St. 36;

⁴ Kazan Medical University, Ministry of Health of Russian Federation, 420012, Kazan, Russia, Butlerova St. 49

Abstract

The amount of ultrasound (US) breast exams continues to grow because of the wider endorsement of breast cancer screening programs. When a solid lesion is found during the US the primary task is to decide if it requires a biopsy. Therefore, our goal was to develop a noninvasive US grayscale image analysis for benign and malignant solid breast lesion differentiation. We used a dataset consisting of 105 ultrasound images with 50 benign and 55 malignant non-cystic lesions. Features were extracted from the source image, the image of the gradient module after applying the Sobel filter, and the image after the Laplace filter. Subsequently, eight gray-level co-occurrence matrices (GLCM) were constructed for each lesion, and 13 Haralick textural features were calculated for each GLCM. Additionally, we computed the differences in feature values at different spatial shifts and the differences in feature values between the inner and outer areas of the lesion. The LASSO method was employed to determine the most significant features for classification. Finally, the lesion classification was carried out by various methods. The use of LASSO regression for feature selection enabled us to identify the most significant features for classification. Out of the 13 features selected by the LASSO method, four described the perilesional tissue, two represented the inner area of the lesion and five described the image of the gradient module. The final model achieved a sensitivity of 98%, specificity of 96%, and accuracy of 97%. Considering the perilesional area, Haralick feature differences, and the image of the gradient module can provide crucial parameters for accurate classification of US images. Features with a low AUC index (less than 0.6 in our case) can also be important for improving the quality of classification.

Keywords: breast ultrasound, solid lesion, benign lesion, malignant lesion, classification, feature selection.

Citation: Kolchev AA, Pasyukov DV, Egoshin IA, Kliouchkin IV, Pasyukova OO. Classification of benign and malignant solid breast lesions on the ultrasound images based on the textural features: the importance of the perifocal lesion area. *Computer Optics* 2024; 48(1): 157-165. DOI: 10.18287/2412-6179-CO-1244.

Acknowledgements: The main results of sections "Materials and methods" and "Results" were obtained by D.V. Pasyukov and I.A. Egoshin with the support by Grant of Russian Science Foundation (Project 22-71-10070, <https://rscf.ru/en/project/22-71-10070/>). The authors are grateful to the Kazan Federal University Strategic Academic Leadership Program (PRIORITY-2030) for the technical feasibility of using hardware and software.

Introduction

Breast cancer (BC) represents an aggressive tumor characterized by high incidence and mortality rates. In 2018, breast cancer accounted for 11.6 % of all malignancies worldwide, sharing first place with lung cancer. At the same time, BC is responsible for 6.6 % of all cancer-related deaths, ranking fourth after lung, gastric, and liver cancer [1].

Nowadays mammography is the only screening approach that has shown improved survival rates [2]. How-

ever, up to 10 % of mammographic exams reveal suspicious findings, with varying malignancy rates ranging from 3 % to 94 % [3]. Biopsying all of these findings would not be a practical approach. As a result, ultrasound (US) is commonly used as a cost-effective and non-invasive modality to determine which lesions require further investigation.

The US image of the breast mass is complex and can be divided into the hypoechoic central and iso- or hyperechoic peripheral components. The central component corresponds to the tumoral tissue, often containing a sig-

nificant amount of fibrotic tissue. This hard tissue significantly increases the stiffness of the central tumor component and frequently causes acoustic shadowing due to US attenuation. Consequently, reliable quantitative analysis of the central component is challenging in most cases [4].

On the other hand, the peripheral part of the tumor contains cancerous cells, as well as components of desmoplastic reaction and inflammatory infiltration, forming a perifocal rim of varying echogenicity and width. This area does not significantly attenuate the US, and therefore can be utilized for quantitative analysis. On the contrary, benign breast tumors usually have no such peripheral component, or it has less width and higher homogeneity [5]. Therefore it looks mandatory to estimate the peripheral breast mass component when performing classification tasks.

Previous studies have examined the area external to the visible hypoechoic mass and showed that it may enhance classification reliability. However as an external area they used fixed distance from the visible border (e.g. 5 mm [6], 40% of the hypoechoic lesion size [7], or 20 pixels [8]). At the same time, the area of the fixed width may incorrectly represent the perilesional tissue and include pixels unrelated to the mass. Moreover, certain works related to breast US image classification included cystic lesions which differ significantly from both benign and malignant solid lesions and can be visually distinguished [9].

The classification of benign and malignant lesions in US images relies on the observation that their structures differ [10–12]. The gray-level co-occurrence matrix (GLCM) provides comprehensive information about pixel relationships in an image. GLCM parameters include the number of gray levels and the distance between compared pixels in the image. The GLCM is calculated in four directions (0° , 45° , 90° , 135°) and serves as a source data for the Haralick textural features, commonly used for mass classification in US images [13–15]. Haralick proposed 14 features for each GLCM, which is considered more reasonable than using morphological features because the latter can significantly overlap in benign and malignant lesions [3].

It has been suggested that incorporating additional texture features obtained by applying Sobel or Laplace filters can improve the quality of lesion classification, similar to machine learning methods utilizing image pre-processing techniques like local binary patterns (LBP), histogram of oriented gradients (HOG), and GLCM [16]).

Therefore, our study aimed to estimate Haralick textural features of the segmented perilesional area of the hypoechoic central mass component for the classification of benign and malignant lesions in US images.

1. Materials and methods

To test the proposed approach we used the data obtained with the help of the following ultrasound systems: Medison SA8000SE, Siemens X150, Esaote MyLab C,

and 7.5–12 MHz probes. These systems allowed us to obtain digital 8-bit ultrasound images of the breast with the detected lesions (see fig. 1a, c). The dataset included 105 ultrasound images of cytologically and/or histologically proven benign, and 55 malignant non-cystic lesions (as displayed in tab. 1).

Tab. 1. Characteristics of the lesions included into the analysis

Feature	Size				Total
	≤10 mm	11–20 mm	21–50 mm	>50 mm	
BIRADS category (n=105)					
BIRADS 2	5	3	1	–	8
BIRADS 3	15	7	5	2	29
BIRADS 4	14	16	5	1	36
BIRADS 5	13	11	7	1	32
Malignant lesions (n=55)					
Breast cancer	28	16	4	2	50
Breast metastases	–	2	1	–	3
Lymphoma	–	1	1	–	2
Benign lesions (n=50)					
Fibroadenoma	3	5	1	–	9
Focal fibrosis	4	3	–	–	7
Intracystic papilloma	1	1	–	–	2
Sclerosing adenosis	7	2	–	–	9
Lipoma	–	6	3	–	9
Inflammatory infiltration	–	3	1	1	5
Gynecomastia	–	–	2	–	2
Phylloid tumor	–	1	1	–	2
Intramammary lymph node	1	4	–	–	5

On each image, the radiologist manually traced the contours of the inner hypoechoic area R_1 (as shown in fig. 1b, d – digit 1) of the lesion and its outer area R_2 (as depicted in fig. 1b, d – digit 2).

To differentiate between malignant and benign lesions, we estimated both inner and outer parts of the lesion. Typically, the outer neighborhood of the lesion exhibited distinctive textural features.

Image features. The GLCM was employed to compute the textural features of the ultrasound images. To construct the GLCM, the gray scale of ultrasound images was limited to 6 bits (the GLCM is $N \times N$, where $N=64$). The GLCM were constructed for four directions α ($\alpha=0^\circ$, 45° , 90° , 135°) and for two distance values between the compared pixels in the GLCM matrix: $d=1$ and $d=5$ (vertical or horizontal spatial shift 0.08 mm and 0.4 mm, respectively). Thus, a total of 8 matrices were determined for each region of interest (ROI).

Features were determined from three images: the source image (fig. 2a, d), the image of the gradient module after applying the Sobel filter (fig. 2b, e), and the image after applying the Laplace filter (fig. 2c, f).

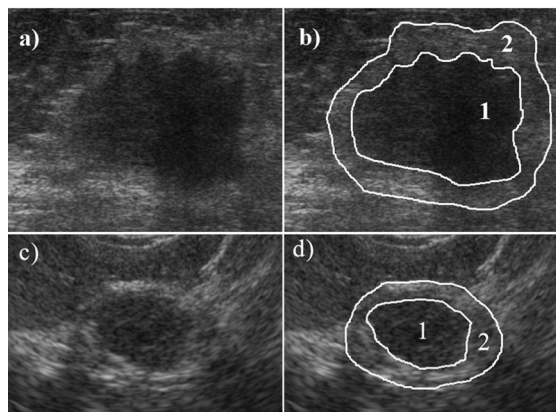


Fig. 1. (a) and (c) 8-bit ultrasound images of the breast carcinoma and benign solid lesion (fibroadenoma), respectively; (b) and (d) traced boundaries of the internal (digit 1) and external (digit 2) parts of the lesion

The Sobel filter was used to approximate the brightness gradient at each point of the image. This filter involved convolving two 3×3 kernels with the original image to calculate derivative approximations: one for horizontal changes and another for vertical changes. If A represents the matrix of the source image, and G_x and G_y are two images that at each point contain approximations of the vertical and horizontal derivatives, respectively, then the operation can be defined as:

$$G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * A \text{ and } G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} * A, \quad (1)$$

where * – convolution operation.

An approximation for the modulus of the gradient G (fig. 2b, e) was obtained by combining of these derivatives:

$$G = \sqrt{G_x^2 + G_y^2}. \quad (2)$$

The Laplace filter approximates the Laplacian of the image brightness function at each point and is given by the kernel:

$$D_{xy}^2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (3)$$

The result of the filtering is the convolution of this kernel with the source image A (fig. 2c, f):

$$L = D_{xy}^2 * A. \quad (4)$$

For each ultrasound image we obtained three images: A , G and L . Two ROI were selected on each of these images: the inner region R_1 of the lesion and the outer region R_2 of the lesion.

As mentioned above, eight GLCM were constructed for each ROI, and 13 Haralick textural features were determined for each GLCM. Therefore a total of $3 \times 2 \times 8 \times 13 = 624$ features were obtained for the one ultrasound image.

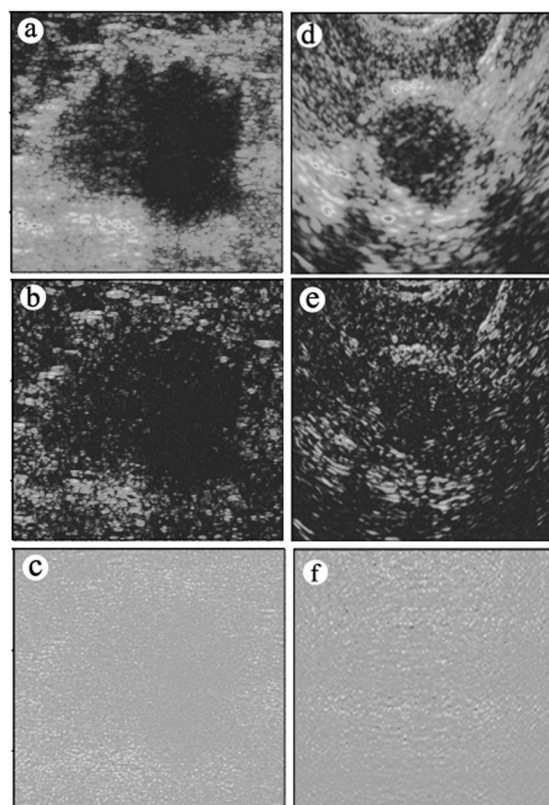


Fig. 2. (a) and (d) source images of the breast carcinoma and benign solid lesion (based on the images showed on fig. 1a and 1c, respectively); (b) and (e) images of the gradient modulus after the Sobel filter, respectively; (c) and (f) images after the Laplace filter, respectively

The paper tested the assumption that not only the features themselves are important for lesion classification, but also the differences between the features at different spatial shift values d , as well as the differences in feature values between the inner and outer areas of the lesion.

Let $V(A, G, L, R_1, d_1)$ is the Haralick feature vector of four GLCM (0°, 45°, 90° and 135°) matrices constructed with $d=1$ for the inner area R_1 of three images A , G and L (totally $4 \times 13 \times 3 = 156$ features). Similarly, let $V(A, G, L, R_2, d_2)$ be the Haralick feature vector of four GLCM matrices constructed with $d=5$ for the outer area R_2 of three images A , G and L . Additional features considered were the difference in feature values:

$$\begin{aligned} V(A, G, L, R_1, d_{15}) &= V(A, G, L, R_1, d_1) - V(A, G, L, R_1, d_5) \\ V(A, G, L, R_2, d_{15}) &= V(A, G, L, R_2, d_1) - V(A, G, L, R_2, d_5) \end{aligned} \quad (5)$$

$$V(A, G, L, R_{12}, d_{15}) = V(A, G, L, R_1, d_{15}) - V(A, G, L, R_2, d_{15}).$$

Two additional features were the average distance between the contours and the standard deviation distance between the contours. The total number of the features considered for the classification was $624 + 3 \times 156 + 2 = 1094$ features.

Various machine learning methods are commonly used for image classification using textural features, including ensemble classifier AdaBoost [17–18], support vector machine [16, 19–20], artificial neural network [19], random forest classifier [19, 21], naive Bayes classi-

fier [16, 19], K – nearest neighbor [16, 19], and logistic regression [22 – 23].

The quality of the image feature classification was analyzed using the following methods: linear discriminant analysis, Gaussian naive Bayes, linear supported vector machine (SVM), nearest neighbor method, bagged trees. We evaluated the Accuracy (Ac), Sensitivity (Se), and Specificity (Sp) for each method when classifying benign and malignant lesions.

2. Results

K -fold cross-validation was used to estimate the classification quality of various machine learning models. This method involves randomly dividing the data into k non-overlapping blocks of approximately equal size. Each block is then treated as a validation set, while the remaining $k-1$ blocks are used as a training set. The model was trained on $k-1$ blocks and evaluated on the validation block. This process is repeated k times, resulting in k scores. The average of these scores represents the final score of the model.

In this study, the classification quality was assessed for each model using $k=3, 4$, and 5 , and the results were averaged.

Tab. 2 shows the classification results of the feature vectors $V(A, G, L, R_1, d_1)$ (where features were determined for the inner area R_1 , and GLCM was constructed with $d=1$) using the different models: model 1 – linear discriminant analysis, model 2 – naive Bayes, model 3 – SVM, model 4 – nearest neighbor method, model 5 – bagged trees. The table displays the classification quality values for various k values, the mean value, and the estimated standard deviation (SD).

Tab. 2. Classification results for $V(A, G, L, R_1, d_1)$. Note: the best values of accuracy, sensitivity and specificity are given in bold

Model	Classification quality, %					
	Parameter	$k=3$	$k=4$	$k=5$	Mean	SD
1	Accuracy	73	69	75	72	1.4
	Sensitivity	71	60	73	68	3.3
	Specificity	76	78	78	77	0.5
2	Accuracy	69	70	68	69	0.5
	Sensitivity	71	69	62	67	2.2
	Specificity	66	70	76	71	2.4
3	Accuracy	77	78	80	78	0.7
	Sensitivity	75	75	80	77	1.4
	Specificity	80	82	80	81	0.5
4	Accuracy	64	67	72	68	1.9
	Sensitivity	67	73	76	72	2.2
	Specificity	60	60	67	62	1.9
5	Accuracy	74	72	77	74	1.2
	Sensitivity	82	71	80	78	2.8
	Specificity	66	74	73	71	2.1

Model 3 (linear SVM) achieves the highest average accuracy and specificity, while model 5 exhibits the highest sensitivity. However, it's important to note that model 5 also has a significantly higher standard deviation.

Tab. 3 shows the classification results for the following feature vectors of the ultrasound images: 1) $V(A, G, L, R_1, d_5)$, 2) $V(A, G, L, R_2, d_1)$, 3) $V(A, G, L, R_2, d_5)$, 4) $V(A, G, L, R_1, d_{15})$, 5) $V(A, G, L, R_2, d_{15})$, 6) $V(A, G, L, R_{12}, d_{15})$. The average classification quality results for $k=3, 4$, and 5 are given for each feature vector.

By comparing the average classification quality results for the feature vectors $V(A, G, L, R_1, d_1)$ and $V(A, G, L, R_1, d_5)$, as well as $V(A, G, L, R_2, d_1)$ and $V(A, G, L, R_2, d_5)$, we observed that using textural features obtained with $d=5$ (spatial shift of 0.4 mm) yields higher sensitivity and specificity compared to $d=1$ (spatial shift of 0.08 mm). Additionally, the textural features extracted from the outer area R_2 of the lesion exhibit higher sensitivity and specificity values than those from the inner area R_1 of the lesion. Notably, the best classification quality was achieved when utilizing the differences between textural features, specifically feature vectors $V(A, G, L, R_1, d_{15})$, $V(A, G, L, R_2, d_{15})$, $V(A, G, L, R_{12}, d_{15})$. Model 3 (the linear SVM model) consistently provided the highest accuracy and specificity among all considered feature vectors.

The vector $V(A, G, L, R_{12}, d_{15})$, which considers the differences in feature values between the outer and inner areas, yielded the highest sensitivity value in classifying the ultrasound images. It's worth mentioning that besides model 3, models 2 and 4 also exhibit high-quality classification indicators for the feature vector $V(A, G, L, R_{12}, d_{15})$.

Previously we demonstrated that the LASSO feature selection method can enhance the quality of the lesion classification in US images [24].

The LASSO is commonly used when dealing with datasets containing numerous variables to exclude some of them and understand how the important features (i.e., those selected by the LASSO as significant) impact the model's performance.

The mathematical equation of the LASSO model can be defined as follows:

$$\arg \min \left(\sum_{i=1}^n (y_i - \sum_{j=1}^m \beta_j x_{ij})^2 + \lambda |\beta| \right),$$

where x_{ij} is the value of the independent variable (1094 considered features for 105 US images), y_i is the value of the dependent variable (-1 for benign and 1 for malignant lesion), λ is the penalty parameter ($\lambda \geq 0$), β_j is the regression coefficient, $n = 105$, $m = 1094$ [25].

In this case, a compromise is achieved between the regression error and the dimension of the used feature space. It is expressed as the sum of the absolute values of the coefficients $|\beta|$. During the minimization process, some coefficients become zero, which determines the selection of informative features.

Since the best results were obtained using the feature difference vectors, a feature vector was created that included all the features of the vectors $V(A, G, L, R_1, d_{15})$, $V(A, G, L, R_2, d_{15})$, and $V(A, G, L, R_{12}, d_{15})$. Additionally, two more features were added: the average distance between the contours and the SD distance between the con-

tours. The total number of features analyzed using the LASSO method was $3 \times 156 + 2 = 470$ features.

Tab. 3. Classification results for the six feature vectors. Note: the highest accuracy, sensitivity and specificity values for the assessed classification methods are highlighted in bold

Model	Classification quality, %						
	Parameter	Feature vector					
		$V(A, G, L, R_1, d_5)$	$V(A, G, L, R_2, d_1)$	$V(A, G, L, R_2, d_5)$	$V(A, G, L, R_1, d_5)$	$V(A, G, L, R_2, d_5)$	$V(A, G, L, R_{12}, d_{15})$
1	Accuracy	76	75	78	82	84	80
	Sensitivity	77	77	82	85	85	85
	Specificity	75	73	73	79	82	75
2	Accuracy	70	75	79	85	83	92
	Sensitivity	78	76	85	86	91	94
	Specificity	62	73	71	83	73	91
3	Accuracy	83	76	90	92	92	92
	Sensitivity	83	84	93	94	94	96
	Specificity	82	68	86	90	91	89
4	Accuracy	66	68	83	85	85	91
	Sensitivity	74	73	87	88	88	94
	Specificity	56	63	79	82	81	87
5	Accuracy	72	73	80	86	88	86
	Sensitivity	78	80	82	88	87	87
	Specificity	65	65	77	82	89	85

The LASSO method identified 14 significant features as follows:

- Two features were selected from the feature vector of the inner area $V(A, G, L, R_1, d_{15})$ for the image of the gradient modulus G (fig. 2b, e) are the total variance for $\alpha=0^\circ$ (feature 1) and the total mean for $\alpha=135^\circ$ (feature 2).

- Four features from the feature vector $V(A, G, L, R_2, d_{15})$ of the outer area. These features include two textural features of the source image A : Haralick correlation for $\alpha=90^\circ$ (feature 3) and total entropy for $\alpha=135^\circ$ (feature 4). Additionally, two textural features of the image gradient module G were selected: Haralick correlation and the first information measure of correlation for $\alpha=0^\circ$ (features 5 and 6, respectively).

- Seven features were obtained from the feature difference vector of the inner and outer areas $V(A, G, L, R_{12}, d_{15})$. These features include six textural features of the source image A : total entropy and second information measure of correlation for $\alpha=0^\circ$; total entropy and Haralick correlation for $\alpha=45^\circ$; and the first information measure of correlation for $\alpha=90^\circ$ and $\alpha=135^\circ$ (features 7, 8, 9, 10, 11, 12, respectively). Furthermore, one textural feature was selected for the image of the gradient modulus G : the second information measure of correlation for $\alpha=0^\circ$ (feature 13).

If $P(i, j)$ is an element of the GLCM, then the selected textural features of Haralick can be represented by the following equations:

- Haralick correlation:

$$f_1 = \frac{\sum_{i=1}^N \sum_{j=1}^N (i \cdot j) P(i, j) - \mu_i \mu_j}{\sigma_i \sigma_j}, \tag{6}$$

where

$$\begin{aligned} \mu_i &= \sum_{i=1}^N i \cdot p_i, \mu_j = \sum_{j=1}^N j \cdot p_j, p_i = \\ &= \sum_{j=1}^N P(i, j), p_j = \sum_{i=1}^N P(i, j), \\ \sigma_i^2 &= \sum_{i=1}^N (i - \mu_i)^2 p_i, \sigma_j^2 = \sum_{j=1}^N (j - \mu_j)^2 p_j; \end{aligned}$$

- total mean:

$$f_2 = \sum_{k=0}^{2N-2} k \cdot P_{x+y}(k), \tag{7}$$

where

$$P_{x+y}(k) = \sum_{j=1}^N \sum_{i=1}^N P(i, j), k = i + j, k = 2, 3, \dots, 2N - 2;$$

- total variance:

$$f_3 = \sum_{k=0}^{2N-2} (k - f_2)^2 \cdot P_{x+y}(k); \tag{8}$$

- total entropy:

$$f_4 = - \sum_{k=0}^{2N-2} P_{x+y}(k) \cdot \log(P_{x+y}(k)); \tag{9}$$

- first information measure of correlation:

$$f_5 = \frac{HXY - HXY1}{\max(HX, HY)}, \tag{10}$$

where

$$HX = - \sum_{i=1}^N p_i \cdot \log p_i, HY = - \sum_{j=1}^N p_j \cdot \log p_j,$$

$$HXY = - \sum_{i=1}^N \sum_{j=1}^N P(i, j) \cdot \log(P(i, j)),$$

$$HXY1 = - \sum_{i=1}^N \sum_{j=1}^N P(i, j) \cdot \log(p_i \cdot p_j);$$

- second information measure of correlation:

$$f_6 = \sqrt{1 - \exp(-2(HXY2 - HXY))}, \tag{11}$$

where

$$HXY2 = - \sum_{i=1}^N \sum_{j=1}^N (p_i \cdot p_j) \cdot \log(p_i \cdot p_j).$$

- Additionally, the standard deviation of the distance between the outer and inner contours was determined as a significant feature (feature 14).

It should be noted that the features corresponding to the image after the Laplace filter (fig. 2c, f) were not identified as significant.

The classification results of the 14 feature vectors selected by the LASSO method are presented in tab. 4. The table displays the classification performance at various values of k , along with the average values. It can be observed that for these 14 features, the model 3 (SVM) achieved the best classification results: accuracy of 97%, sensitivity of 98%, and specificity of 95%. The application of the LASSO method for ultrasound image feature selection has significantly enhanced the quality of the classification.

Tab. 4. Classification results for the LASSO feature vectors. Note: the highest accuracy, sensitivity and specificity values for the assessed classification methods are highlighted in bold

Model	Classification quality, %				
	Parameter	$k=3$	$k=4$	$k=5$	Mean
1	Accuracy	92	91	92	92
	Sensitivity	89	92	91	91
	Specificity	95	88	93	92
2	Accuracy	91	93	93	92
	Sensitivity	89	91	92	91
	Specificity	93	95	93	94
3	Accuracy	99	96	97	97
	Sensitivity	98	98	98	98
	Specificity	100	93	95	96
4	Accuracy	98	93	94	95
	Sensitivity	96	96	92	95
	Specificity	100	88	95	94
5	Accuracy	84	90	85	86
	Sensitivity	89	92	83	88
	Specificity	79	86	88	84

Tab. 5. Values of correlation coefficients between the features. Note: the highest correlation coefficients are highlighted in gray

Feature	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	0.32	0.13	0.22	0.17	0.53	0.36	0.36	0.52	0.5	0.17	0.08	0.3	0.31
2		1	0.6	0.2	0.1	0.44	0.55	0.5	0.45	0.58	0.04	0.13	0.39	0.52
3			1	0.16	0.21	0.24	0.23	0.33	0.34	0.32	0.04	0.02	0.22	0.28
4				1	0.01	0.3	0.11	0.41	0.22	0.12	0.04	0	0.53	0.19
5					1	0.08	0.15	0	0.33	0.24	0.09	0.11	0.03	0.07
6						1	0.46	0.46	0.4	0.6	0.1	0.03	0.37	0.64
7							1	0.52	0.56	0.85	0.03	0.04	0.31	0.41
8								1	0.25	0.5	0.26	0.27	0.78	0.28
9									1	0.72	0.05	0.04	0.23	0.3
10										1	0.11	0.07	0.28	0.56
11											1	0.9	0.37	0.36
12												1	0.43	0.28
13													1	0.13
14														1
AUC	0.82	0.84	0.79	0.74	0.55	0.86	0.83	0.84	0.75	0.86	0.53	0.52	0.80	0.83
P value	4.1E-9	4.3E-9	1.2E-5	2.4E-6	0.12	3.2E-11	8.7E-09	4.1E-10	5.7E-06	9.9E-11	0.0092	0.028	2.3E-08	6.8E-10

The statistical significance of the differences in average values between benign and malignant lesions for each feature was evaluated using Student's t-test (last row in tab. 5) [22]. The quality of classification when utilizing individual features was assessed using the AUC indicator (area under the ROC curve) (penultimate row in tab. 5).

Features 5, 11, and 12 demonstrate the lowest AUC value (close to 0.5) and the highest P value. However,

Fig. 3 shows the normalized distributions of all features for benign and malignant breast lesions.

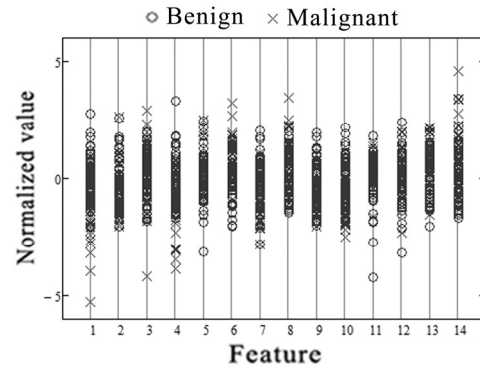


Fig. 3. The normalized distributions of all features

Fig. 3 shows that the values of the individual features of benign lesions overlap with those of malignant lesions. In contrast to the findings in a previous work [13], there are no features that exhibit a clear separation between the two classes.

Tab. 5 shows the absolute values of the correlation coefficients r for the selected 14 features. It can be observed that most of the features exhibit weak correlations. Only in four cases out of 91 correlation coefficients (highlighted in gray cells) the absolute value of the correlation coefficient is greater than 0.7 (excluding the correlation of a feature with itself). Additionally, 74 correlation coefficients have values less than 0.5.

excluding these features from the set of features for classification leads to a deterioration in classification results. Specifically, the specificity of model 3 decreases from 95% to 91% for $k=5$.

3. Discussion

For our experiments, we decided to create our own image dataset. The reason for this choice is that open

source datasets often include images with a gray-level histogram shift towards the brighter end. This type of distortion can hinder the assessment of low-intensity gray levels. It is important to note that this issue stems from the diverse image presets available on different ultrasound machines. To ensure unbiased representation of gray levels and enable future processing of the images, it was necessary to carefully select images with accurate gray level representations.

It is indeed true that directed image pre-processing techniques can enhance the classification of breast lesions by highlighting relevant features [26]. In the analysis of image structure, filters such as Sobel and Laplace are commonly used to accentuate the borders of brightness differences, making object contours more distinct. Therefore, we propose that applying Sobel or Laplace filtering in combination with textural features can improve the quality of breast lesion classifications.

What concerns the features, it is well-known that the morphological ones are insensitive for breast lesion classification. As an alternative approach, we focused on textural features extracted from the key outer image area [22, 27]. Haralick textural features have been widely used in classifying breast ultrasound images in several studies [6, 13, 28]. We utilized the first 13 Haralick features, excluding the feature that requires eigenvalue calculations from the GLCM.

Our results indicate that both the individual features themselves and their differences calculated under different conditions play a significant role in classification. Comparing feature differences at various spatial distance d values yielded better classification performance than using fixed d values.

In our study, we demonstrated that, in line with the theoretical concept and some previous works [6, 29], the perilesional area has a significant importance for solid breast lesion classification. In our experiment, out of 13 features selected by the LASSO method, four ones described the perilesional tissue and only two represented the inner area of the lesion. Moreover, a majority of the selected features characterized the difference between the inner and outer areas of the lesion.

Another crucial aspect is the selection of the lesion region to be included in the processing. It is well-known that the upper 120° sector of the lesion is relatively free from US artifacts. On the other hand, the side and lower sectors often contain numerous unpredictable artifacts that may potentially diminish the quality of classification [30]. Taking this into consideration, we solely utilized the upper sector of the lesion for feature calculation in order to improve the classification accuracy.

Besides the textural features of the source image we also incorporated the textural features of the brightness gradient module image. Out of the 13 textural features selected by the LASSO method, eight represent the source US image and five describe the image of the gradient module. This highlights the importance of analyzing not

only the original US images but also the results of their filtering for effective lesion classification.

It should also be noted that the three of the selected features have AUC values below 0.6, which typically suggests limited utility for classification purposes. However, in our experiments these features still influenced the classification quality, possibly due to the inclusion of non-standard images in our training set.

Recently, deep learning has gained popularity in biomedical image analysis [31–32]. Unlike traditional machine learning methods, deep learning approaches are independent of the feature extraction step. Prior to the advent of deep learning, the feature extraction was performed manually and required domain experts knowledge. In contrast, deep learning relies on neural networks that automatically detect effective feature representations through the nonlinear transformation of primitive data features, such as word vectors and picture pixels [33].

However, deep neural networks require an extensive set of the previously labeled and specially prepared digitized data for training, which can be a problem in medical applications. Moreover, the training of the deep learning models from scratch requires tens of thousands of images to avoid overfitting.

Using data augmentation methods such as MixUp, the Mosaic method or spatial transformation of source images (flipping, rotation, padding) [34–36] allows to increase the number of images significantly. However, the resulting dataset is still not ideal for training deep learning models from scratch.

Simultaneously, merging of several public US databases can also be challenging due to the lack of standardization among them.

Often, to address the issue of overfitting, transfer learning methods are employed. These methods utilize pre-trained weights obtained from large open datasets such as PASCAL VOC, ImageNet, MS COCO, and others [37–39]. In transfer learning the last few layers of the network are usually replaced with new ones and initialized with random weights. The remaining layers can either be frozen (i.e. unmodifiable) or left learnable. However, since pretrained weights are usually trained on non-medical data like cars, buses, bicycles etc., it is necessary to adopt an appropriate methodology for transfer learning or domain adaptation. Furthermore, the presence of non-standard medical cases, which are often encountered in clinical practice, can also impact the performance of deep learning networks. Therefore, a comprehensive evaluation of deep learning approaches for medical purposes requires a unified database consisting of tens of thousands of images annotated by specialists.

Conclusions

In the classification of breast lesion ultrasound images, the differences in features between the internal and external lesion areas are more significant than those within the lesion itself. Given the wide range of shapes and

structures observed in breast lesions, features with an AUC parameter close to 0.5 should not be automatically excluded from the analysis. These features can contribute to improving the classification accuracy, especially for rare lesion types. Furthermore, incorporating filtered versions of the images, such as applying Sobel and Laplace filters, can enhance the classification quality, emphasizing the importance of considering both the original and processed images.

References

- [1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin* 2018; 68: 394-424. DOI: 10.3322/caac.21492.
- [2] Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, Straif K. Breast-cancer screening – viewpoint of the IARC Working Group. *N Engl J Med* 2015; 372: 2353-2358. DOI: 10.1056/NEJMs1504363.
- [3] Raza S, Goldkamp AL, Chikarmane SA, Birdwell RL. US of breast masses categorized as BI-RADS 3, 4, and 5: Pictorial review of factors influencing clinical management 1. *Radiographics* 2010; 30(5): 1199-1213. DOI: 10.1148/rg.305095144.
- [4] Guo R, Lu G, Qin B, Fei B. Ultrasound imaging technologies for breast cancer detection and management: A review. *Ultrasound Med Biol* 2018; 44(1): 37-70. DOI: 10.1016/j.ultrasmedbio.2017.09.012.
- [5] Krizmanich-Conniff KM, Paramagul C, Patterson SK, Helvie MA, Roubidoux MA, Myles JD, Jiang K, Sabel M. Triple receptor-negative breast cancer: Imaging and clinical characteristics. *AJR Am J Roentgenol* 2012; 199(2): 458-464. DOI: 10.2214/AJR.10.6096.
- [6] Klimonda Z, Karwat P, Dobruch-Sobczak K, Piotrkowska-Wroblewska H, Litniewski J. Breast-lesions characterization using Quantitative Ultrasound features of peritumoral tissue. *Sci Rep* 2019; 9: 7963. DOI: 10.1038/s41598-019-44376-z.
- [7] Bahareh B, Hamze R, Tehrani Ali KZ, Hassan R. Deep classification of breast cancer in ultrasound images: more classes, better results with multi-task learning. *Proc SPIE* 2021; 11602: 116020S. DOI: 10.1117/12.2581930.
- [8] Nemat H, Fehri H, Ahmadinejad N, Frangi AF, Gooya A. Classification of breast lesions in ultrasonography using sparse logistic regression and morphology-based texture features. *Med Phys* 2018; 45(9): 4112-4124. DOI: 10.1002/mp.13082.
- [9] Telagarapu P, Poonguzhali S. Analysis of contourlet texture feature extraction to classify the benign and malignant tumors from breast ultrasound images. *Int J Eng Technol* 2014; 6(1): 293-305.
- [10] Badawy SM, Mohamed AE-NA, Hefnawy AA, Zidan HE, GadAllah MT, El-Banby GM. Automatic semantic segmentation of breast tumors in ultrasound images based on combining fuzzy logic and deep learning-A feasibility study. *PLoS ONE* 2021; 16: e0251899. DOI: 10.1371/journal.pone.0251899.
- [11] Jabeen K., Khan MA, Alhaisoni M, Tariq U, Zhang Y-D, Hamza A, Mickus A, Damasevicius R. Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors* 2022; 22(3): 807. DOI: 10.3390/s22030807.
- [12] Gonzalez-Luna FA, Hernandez-Lopez J, Gomez-Flores W. A performance evaluation of machine learning techniques for breast ultrasound classification. 2019 16th Int Conf on Electrical Engineering, Computing Science and Automatic Control (CCE) 2019; 1-5. DOI: 10.1109/ICEEE.2019.8884547.
- [13] Hsu SM, Kuo WH, Kuo FC, Liao YY. Breast tumor classification using different features of quantitative ultrasound parametric images. *Int J CARS* 2019; 14: 623-633. DOI: 10.1007/s11548-018-01908-8.
- [14] Wei M, Wu X, Zhu J, Liu P, Luo Y, Zheng L, Du Y. Multi-feature fusion for ultrasound breast image classification of benign and malignant. *IEEE 4th Int Conf on Image, Vision and Computing (ICIVC)* 2019; 474-478. DOI: 10.1109/ICIVC47709.2019.8980898.
- [15] Prabhakar T, Poonguzhali S. Assessment of texture feature extraction to classify the benign and malignant lesions from breast ultrasound images. In Book: Dash S, Naidu P, Bayindir R, Das S, eds. *Artificial intelligence and evolutionary computations in engineering systems. Proceedings of ICAIECES 2017*. Singapore: Springer Nature Singapore Pte Ltd; 2018: 725-732. DOI: 10.1007/978-981-10-7868-2_69.
- [16] Wei M, Du Y, Wu X, Su Q, Zhu J, Zheng L, Lv G, Zhuang J. A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images. *Comput Math Methods Med* 2020; 2020: 5894010. DOI: 10.1155/2020/5894010.
- [17] Moustafa AF, Cary TW, Sultan LR, Schultz SM, Conant EF, Venkatesh SS, Sehgal CM. Color Doppler ultrasound improves machine learning diagnosis of breast cancer. *Diagnostics* 2020; 10(9): 631. DOI: 10.3390/diagnostics10090631.
- [18] Mishra AK, Roy P, Bandyopadhyay S, Das SK. Breast ultrasound tumour classification: A machine learning-radiomics based approach. *Expert Syst* 2021; 38: e12713. DOI: 10.1111/exsy.12713.
- [19] Faust O, Acharya UR, Meiburger KM, Molinari F, Koh JEW, Yeong CH, Kongmebol P, Ng KH. Comparative assessment of texture features for the identification of cancer in ultrasound images: A review. *Biocybern Biomed Eng* 2021; 38(2): 275-296. DOI: 10.1016/j.bbe.2018.01.001.
- [20] Shia WC, Chen DR. Classification of malignant tumors in breast ultrasound using a pretrained deep residual network model and support vector machine. *Comput Med Imaging Graph* 2021; 87: 101829. DOI: 10.1016/j.compmedimag.2020.101829.
- [21] Rodriguez-Cristerna A, Gomez-Flores W, de Albuquerque Pereira WC. A computer-aided diagnosis system for breast ultrasound based on weighted BI-RADS classes. *Comput Methods Programs Biomed* 2018; 153: 33-40. DOI: 10.1016/j.cmpb.2017.10.004.
- [22] Wu T, Sultan LR, Tian J, Cary TW, Sehgal CM. Machine learning for diagnostic ultrasound of triple-negative breast cancer. *Breast Cancer Res Treat* 2019; 173(2): 365-373. DOI: 10.1007/s10549-018-4984-7.
- [23] Xiong H, Sultan LR, Cary TW, Schultz SM, Bouzghar G, Sehgal CM. The diagnostic performance of leak-plugging automated segmentation versus manual tracing of breast lesions on ultrasound images. *Ultrasound* 2017; 25(2): 98-106. DOI: 10.1177/1742271X17690425.
- [24] Kolchev AA, Pasyukov DV, Egoshin IA, Kliouchkin IV, Pasyukova OO. Cystic (including atypical) and solid breast lesion classification using the different features of quantitative ultrasound parametric images. *Int J CARS* 2022; 17: 219-228. DOI: 10.1007/s11548-021-02522-x.
- [25] Yang Y, Li W, Kang Y, Guo Y, Yang K, Li Q, Liu Y, Yang C, Chen R, Chen H, Li X, Cheng L. A novel lung radiomics feature for characterizing resting heart rate and COPD stage evolution based on radiomics feature combi-

- nation strategy. *Math Biosci Eng* 2022; 19(4): 4145-4165. DOI: 10.3934/mbe.2022191.
- [26] Chowdhury A, Razzaque RR, Muhtadi S, Shafiullah A, Ul Islam Abir E, Garra BS, Kaisar Alam S. Ultrasound classification of breast masses using a comprehensive Nakagami imaging and machine learning framework. *Ultrasonics* 2022; 124: 106744. DOI: 10.1016/j.ultras.2022.106744.
- [27] Liu L, Li K, Qin W, Wen T, Li L, Wu J, Gu J. Automated breast tumor detection and segmentation with a novel computational framework of whole ultrasound images. *Med Biol Eng Comput* 2018; 56(2): 183-199. DOI: 10.1007/s11517-017-1770-3.
- [28] Wei M, Du Y, Wu X, Zhu J. Automatic classification of benign and malignant breast tumors in ultrasound image with texture and morphological features. 2019 IEEE 13th Int Conf on Anti-counterfeiting, Security, and Identification (ASID) 2019; 126-130. DOI: 10.1109/ICASID.2019.8925194.
- [29] Hsieh YH, Hsu FR, Dai ST, Huang HY, Chen DR, Shia WC. Incorporating the breast imaging reporting and data system lexicon with a fully convolutional network for malignancy detection on breast ultrasound. *Diagnostics* 2021; 12(1): 66. DOI: 10.3390/diagnostics12010066.
- [30] Zhou Z, Wu S, Chang KJ, Chen WR, Chen YS, Kuo WH, Lin CC, Tsui PH. Classification of benign and malignant breast tumors in ultrasound images with posterior acoustic shadowing using half-contour features. *J Med Biol Eng* 2015; 35(2): 178-187. DOI: 10.1007/s40846-015-0031-x.
- [31] Agafonova YuD, Gaidel AV, Zelter PM, Kaphishnikov AV. Efficiency of machine learning algorithms and convolutional neural network for detection of pathological changes in MR images of the brain. *Computer Optics* 2020; 44(2): 266-273. DOI: 10.18287/2412-6179-CO-671.
- [32] Vinokurov VO, Matveeva IA, Khristoforova YA, Myakinin OO, Bratchenko IA, Bratchenko LA, Moryatov AA, Kozlov SG, Machikhin AS, Abdulhalim I, Zakharov VP. Neural network classifier of hyperspectral images of skin pathologies. *Computer Optics* 2021; 45(6): 879-886. DOI: 10.18287/2412-6179-CO-832.
- [33] Zhu C. *Machine reading comprehension: Algorithms and practice*. Oxford: Elsevier; 2021. ISBN: 978-0-323-90118-5.
- [34] Kriti, Virmani J, Agarwal R. Deep feature extraction and classification of breast ultrasound images. *Multimed Tools Appl* 2020; 79(38): 27257-27292. DOI: 10.1007/s11042-020-09337-z.
- [35] Bochkovskiy A, Wang CY, Liao HY. YOLOv4: Optimal speed and accuracy of object detection. *arXiv Preprint*. 2020. Source: <<https://arxiv.org/abs/2004.10934>>.
- [36] Zeimarani B, Costa MGF, Nurani NZ, Bianco SR, De Albuquerque Pereira WC, Filho CFFC. Breast lesion classification in ultrasound images using deep convolutional neural network. *IEEE Access* 2020; 8: 133349-133359. DOI: 10.1109/ACCESS.2020.3010863.
- [37] Kalafi EY, Jodeiri A, Setarehdan SK, Lin NW, Rahmat K, Taib NA, Ganggayah MD, Dhillon SK. Classification of breast cancer lesions in ultrasound images by using attention layer and loss ensemble in deep convolutional neural networks. *Diagnostics* 2021; 11(10): 1859. DOI: 10.3390/diagnostics11101859.
- [38] Cao Z, Yang G, Li X, Chen Q, Wu J. Multitask classification method based on label correction for breast tumor ultrasound images. *Neural Process Lett* 2021; 53: 1453-1468. DOI: 10.1007/s11063-021-10455-4.
- [39] Cao Z, Duan L, Yang G, Yue T, Chen Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Med Imaging* 2019; 19(1): 51. DOI: 10.1186/s12880-019-0349-x.

Authors' information

Alexey Anatolevich Kolchev (b. 1965) Dr, PhD, associate professor in Kazan (Volga region) Federal University, Ministry of Education and Science of Russian Federation. Research interests: mathematical modeling, digital signal and image processing, deep learning, machine learning. E-mail: kolchevaa@mail.ru. ORCID: 0000-0002-1692-2558.

Dmitry Valerevich Pasyukov (b. 1975) MD, PhD, associate professor, head of Radiology and Oncology department in Mari State University, Ministry of Education and Science of Russian Federation, senior researcher in Kazan (Volga region) Federal University, Ministry of Education and Science of Russian Federation, assistant professor of Ultrasound Diagnostics department in Kazan State Medical Academy - Branch Campus of the Federal State Budgetary Educational Institution of Further Professional Education «Russian Medical Academy of Continuous Professional Education», Ministry of Healthcare of the Russian Federation. Research interests: radiology in oncology, radiological image analysis. E-mail: passynkov@mail.ru. ORCID: 0000-0003-1888-2307.

Ivan Aleksandrovich Egoshin (b. 1991) junior researcher at Mari State University, Ministry of Education and Science of Russian Federation. Also he is a junior researcher at Kazan (Volga region) Federal University, Ministry of Education and Science of Russian Federation. Research interests: digital signal and image processing, radiomics, artificial intelligence, machine learning, programming. E-mail: jungl91@mail.ru. ORCID: 0000-0003-0717-0734.

Ivan Vladimirovich Klioushkin (b. 1948) MD, DSc, Professor of General Surgery department in Kazan Medical University, Ministry of Health of Russian Federation. Research interests: radiology, medical image analysis. E-mail: hirurgivan@rambler.ru. ORCID: 0000-0001-5052-2921.

Olga Olegovna Pasyukova (b. 1977) MD, PhD, associate professor of Fundamental Medicine department in Mari State University, Ministry of Education and Science of Russian Federation. Research interests: fundamental and clinical medicine, ultrasound diagnostics, mathematical modeling. E-mail: olgaved@inbox.ru. ORCID: 0000-0001-9117-8151.

Code of State Categories Scientific and Technical Information (in Russian – GRNTI): 28.23.15
Received October 23, 2022. The final version – November 30, 2022.