CrossMark

ORIGINAL ARTICLE

Ultrasound texture-based CAD system for detecting neuromuscular diseases

Tim König · Johannes Steffen · Marko Rak · Grit Neumann · Ludwig von Rohden · Klaus D. Tönnies

Abstract

Purpose Diagnosis of neuromuscular diseases in ultrasonography is a challenging task since experts are often unable to discriminate between healthy and pathological cases. A computer-aided diagnosis (CAD) system for skeletal muscle ultrasonography was developed and tested for myositis detection in ultrasound images of biceps brachii.

Methods Several types of features were extracted from rectangular and polygonal image regions-of-interest (ROIs), including first-order statistics, wavelet-based features, and Haralick's features. Features were chosen that are sensitive to the change in contrast and structure for pathological ultrasound images of neuromuscular diseases. The number of features was reduced by applying different sequential feature selection strategies followed by a supervised principal component analysis. For classification, two linear approaches were investigated: Fisher's classifier and the linear support vector machine (SVM) as well as the nonlinear k-nearest neighbor approach. The CAD system was benchmarked on datasets of 18 subjects, seven of which were healthy, while 11 were affected by myositis. Three expert radiologists provided pre-classification and testing interpretations.

Results Leave-one-out cross-validation on the training data revealed that the linear SVM was best suited for discriminating healthy and pathological muscle tissue, achieving 85/87 % accuracy, 90 % sensitivity, and 83/85 % specificity, depending on the radiologist.

Conclusion A muscle ultrasonography CAD system was developed, allowing a classification of an ultrasound image by one-click positioning of rectangular ROIs with minimal user effort. The applicability of the system was demonstrated with the challenging example of myositis detection, showing highly accurate results that were robust to imprecise user input.

Introduction

Neuromuscular diseases, e.g., progressive spinal muscular atrophy (Werdnig–Hoffmann disease), Duchenne muscular dystrophy, or myositis are common neurological disorders. Incidence rates are 10/100,000 for progressive spinal muscular atrophy, 29/100,000 for Duchenne muscular dystrophy, and 1/100,000 for myositis.

Ultrasonography is a convenient technique to visualize healthy and pathological skeletal muscle tissue, providing high resolution scans with noninvasive real-time image acquisition. Neuromuscular disorders, like myositis, often cause structural muscle changes that can be observed in ultrasound images. It is assumed that the replacement of muscle tissue by fat and an increase or a thickening of fibrous connected tissue is the main reason for an increased muscle echo intensity in pathological tissue, since the number of reflections increases within the muscle resulting in partially higher intensities in the ultrasound image [29]. Hence, only few reflections occur in healthy muscle tissue during image acquisition, resulting in a low echo intensity and thus in rather dark images with more evenly distributed high-intensity peaks. As

T. König (✉) · J. Steffen · M. Rak · K. D. Tönnies
Department of Simulation and Graphics, Otto von Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany
e-mail: tim.koenig@isg.cs.ovgu.de

G. Neumann · L. von Rohden
Department of Radiology and Nuclear Medicine, Otto von Guericke University, Leipziger Straße 44, 39120 Magdeburg, Germany

a consequence, texture of healthy muscle tissue seems to be more oriented and structured due to higher contrast changes, while pathological muscle tissue appears to be diffuse and unstructured due to its smaller contrast changes.

We treat the issue of detection of neuromuscular diseases in ultrasound images at the challenging example of myositis detection. Myositis, or inflammatory myopathy, is a rare disease in which the immune system chronically inflames the body's muscle tissue. Persistent inflammation progressively weakens or destroys the muscle tissue and is commonly accompanied by pain. Over time, this process may also lead to a loss of muscle mass. Myositis most often appears in childhood, at the age of 5–15, and in adulthood, at the age of 40–60, although it might occur at any age. In contrast to other neuromuscular disorders like progressive spinal muscular atrophy or Duchenne muscular dystrophy, a texture analysis of myositis is more difficult, since the disease may only occur focally and brightness changes in pathological cases are often small [28,30]. For this reason, we developed a CAD system for detecting myositis in ultrasound images, which can be easily adapted to detect similar neuromuscular disorders.

Under the assumption that pathological muscle tissue appears to be diffuse and unstructured, a texture-based analysis seems adequate for detecting myositis in ultrasound images. A computer-aided approach serves several purposes: It provides an independent opinion, which can be taken into consideration, and it saves time for radiologists. Moreover, it helps young professionals with less experience to identify myositis in ultrasound images. A well-trained process may also be capable to support an early diagnosis.

Diagnosis from visual inspection can be difficult for the radiologist in early stages of neuromuscular diseases or if myopathies cause only little structural changes in muscle groups. Figure 1 shows ultrasound images of pathological muscle tissue with different degree of severity of myositis. Discrimination between healthy and strongly affected muscle tissue in Fig. 1a, d seems to be rather intuitive, compared to the more difficult discrimination between healthy and little affected muscle tissue in Fig. 1a, b, or little and medium affected muscle tissue in Fig. 1b, c. The latter distinctions are challenging tasks even for automated classification.
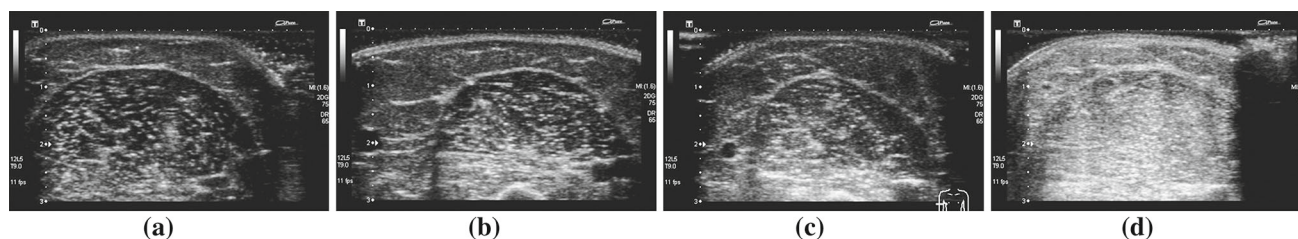
Related work

Texture features are well established in many medical applications incorporating ultrasound imaging. Most problems require a combination of statistical and spectral features, since they represent the nature of deterministic tissue-related variation and non-deterministic influences from image generation ,respectively, [16,25,31,35,38].

These problems include diagnosis and classification of, for example, diffuse, chronic, alcoholic, fatty liver diseases, and liver cancer in ultrasound images [4,15,19, 20,24,26]. Such methods often require texture measures that rely on spatial gray level dependence, gray level co-occurrence, fractal dimension, Fourier transform, Law's texture energy, discrete wavelet transform, or wavelet packet decomposition features combined with, for example, probabilistic neural networks or support vector machines as classifiers. More recent methods also include an automatic ROI detection or segmentation of the liver [22, 32].

Another important research branch of texture analysis in ultrasound imaging is the detection of breast cancer. Earlier work uses first-order statistics or the Fourier transform to extract texture features [12,13]. Autocorrelation [6], morphological [5], auto-covariance [16], or Haralick's [21] features of the breast tissue have also been used. More recently, approaches combine morphological features, generated by active contour or level set segmentation, with texture features derived from autocorrelation functions, first-order statistics, or spatial gray level dependences [34,39] as well as shearlet transform [40], which have proven to produce good and stable classification rates.

There are several other research areas of texture analysis in ultrasound imaging, e.g., the diagnosis of prostate cancer in transrectal images. Here, methods use, e.g., combinations of features derived from gray level co-occurrence matrices, Gabor transform, or discrete wavelet transform [3,25,33]. To detect carotid atherosclerotic plaque in ultrasound images, different approaches exist, which for instance use features based on first- and second-order statistics, fractal dimensions, discrete wavelet transform, wavelet package decomposition,

Fig. 1 Exemplary stages of myositis disease on ultrasound images of biceps brachii showing healthy (**a**) and pathological muscle tissue with different degree of severity: **b** little, **c** medium, and **d** strongly affected muscle tissue

gray level co-occurrence matrix, Law's filter masks, and the Fourier transform [1,2,8,35,37]. Some geometrical, Gabor, and statistical features as well as combinations of these have also been used for Parkinson's disease diagnostics in transcranial sonography [7,17,31].

Regarding skeletal muscle ultrasound, Pohle et al. [28] proposed an approach for examining a number of different neuromuscular diseases. Although they obtained good classification rates, they did not use a standardized examination procedure. Scan and ROI estimation were performed subjectively by one person for already preselected images. Moreover, it was necessary to record digital ultrasound images on analog S-VHS cassettes during that time. A re-digitization, which is often a source of error, was required for a computer-aided evaluation. However, it was shown that a texture-based analysis of ultrasound images may work well for detecting progressive spinal muscular atrophy and Duchenne muscular dystrophy.

Pillen and van Alfen [27] also investigated neuromuscular diseases in ultrasound images. They proposed a quantitative gray scale analysis using first-order histogram-based statistical features, e.g., the mean intensity values of the ROIs. This approach performs well for diseases that cause a regular, increased echo intensity like Duchenne muscular dystrophy. However, methods using only first-order statistics most likely fail for irregular or focally located diseases, as in some cases of myositis.

König et al. [18] proposed a classification approach for detecting myositis with an overall accuracy of $\geq 85\%$. They used a combination of two statistical and two wavelet features, measuring both, brightness and structural changes of pathological tissue in ultrasound images. Moreover, they demonstrated that too small ROI sizes ($<1.2\,\text{cm}^2$) are unsuitable for such tasks. But their dataset was sparse and the approach was too sensitive with respect to tissue surrounding the muscle tissue of interest. Furthermore, texture features were only determined by visual inspection of the feature space. To sum up, this approach is not really applicable for clinical practice.

Materials and methods

We performed texture analysis on a dataset of biceps brachii (BB) of 7 healthy subjects and 11 subjects affected by myositis. Subjects aged from 40 to 80 years. Images were captured with a Toshiba Aplio XG system using a 9 MHz linear-array transducer with a scan width of about 50 mm, viewing a field depth of 30 mm. We used a standardized acquisition protocol (75 gain, 65 dynamic range) and smooth transducer movement in distal direction and with respect to orientations in transverse plane. All scans were captured in the transverse plane that is perpendicular to the long axis of the muscle, resulting in a more speckled appearance because of reflections caused by the perimysial connective tissue [27]. Two to four frames were captured per subject at a pixel resolution of 716×537 px from the ultrasound video sequences of the skeletal muscle. To ensure that frames of the same subject are not highly correlated, we captured them at different distal positions on the muscle. Another reason is that myositis may only appear focally and may be affected to varying degrees within the muscle. Figure 2 exemplary shows the different appearance of muscle tissue between captured frames of the same subject. Frames were saved in the Tagged Image File Format (TIFF), resulting in a database of 60 ultrasound images.

We decided to use a twofold experimental design, which means that our pre-studies in the following subsections (Sects. 3.2 to 3.5) were based on pre-classified images of an expert radiologist E0 to optimize the classification process. Additionally, the expert E0 evaluated 25 of the 60 images as visually hard classifiable, whereby 18 of 25 were healthy and seven of 25 were pathological cases. Pathological cases were additionally verified by the existing results of diagnostic tests. Later experiments evaluate the performance of our eventually designed CAD system and were based on the decisions of two further expert radiologists E1 and E2 (see Sect. 3.6).

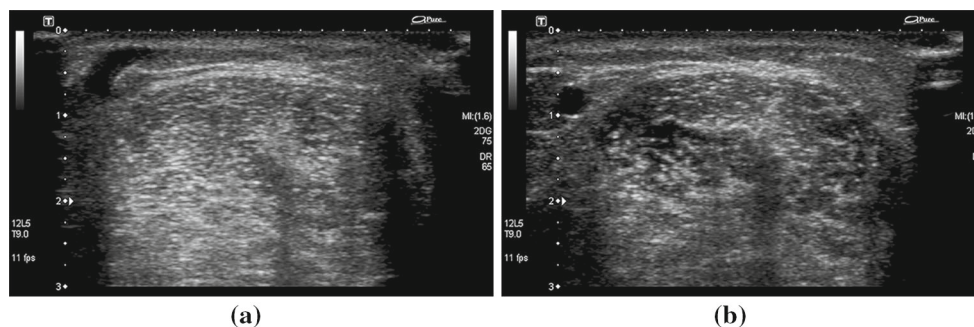All methods were implemented in MATLAB®, partially using the freely available PRTools [11] toolbox.

Fig. 2 Images of the same subject affected by myositis were captured at different positions in distal direction. The muscle tissue of the images (**a**) and (**b**) is visually well distinguishable and not highly correlated
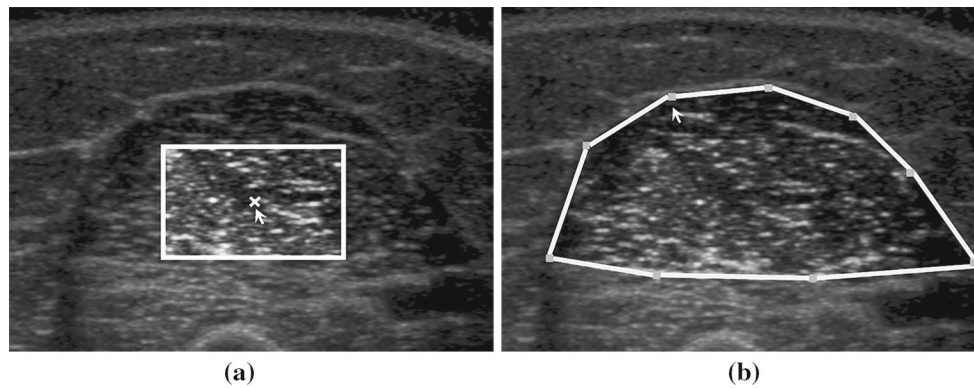
Fig. 3 Two different ROI shapes used for feature extraction in our CAD system. An rectangular ROI (**a**) is used to minimize user effort. A closed polygonal ROI (**b**) can be used on demand as it is more accurate with respect to delineation of muscle tissue. However, it requires slightly higher user effort (8–12 mouse clicks)

ROI generation

To ensure that texture analysis is performed on areas of muscle tissue, a segmentation of the muscle would be beneficial. However, automatic segmentation relies heavily on prior knowledge about the muscle and joint. Therefore, we let the radiologist E0 define a ROI around the tissue of interest. We implemented two kinds of ROI placements, shown in Fig. 3, requiring low and medium user effort ,respectively.

The first step of our CAD system is based upon an one-click approach where the radiologist selects the center of the relevant muscle area. From this selection, a rectangular ROI is generated, which covers the majority of the clinically relevant muscle tissue. This ROI size is based on the average muscle size of biceps brachii and represents only an initial guess. If the suggested ROI is over- or undersized, the radiologist may adjust sizes by using the mouse's scroll wheel. Figure 3a outlines the rectangular ROI approach.

Additionally, we investigated a more precise polygonal ROI definition which is shown in Fig. 3b. The amount of user interactions for polygonal ROIs depends on the desired precision but typically requires 8–12 mouse clicks. Such a quasi-segmentation of the muscle is more reliable, since no other types of (surrounding) tissue are included in the texture analysis. The following pre-studies were based on the rectangular and polygonal ROI definitions of expert radiologist E0.

Feature extraction

We investigated several first-order statistical features derived from the intensity histogram, Haralick's features, derived from the gray level co-occurrence matrix (GLCM) [14], and features obtained from two-dimensional discrete wavelet transform (2D-DWT) [23] due to the following reasons:

– First-order statistics, of which some have already been successfully applied on muscle ultrasound images [27], should describe the intensity changes of pathological muscle tissue appropriately.
– Haralick's features have been used on ultrasound images of similar neuromuscular diseases [28] and seem promising for describing structural changes of pathological muscle tissue.
– Wavelet features are able to describe both intensity and structural information similar to other spectral features that have been successfully applied in other ultrasound image classification tasks (see Sect. 2).

For first-order statistics, we used five features: the mean intensity value, standard deviation, skewness, kurtosis, and entropy of the ROI's histogram. A bias-corrected formula is used to compute standard deviation, skewness, and kurtosis to cope with differently sized ROIs.

Haralick's texture features are based on the GLCM from which several statistical measures can be derived. GLCM is usually calculated from intensity quantized images to obtain a statistically reliable estimate of the joint intensity probability distribution. We computed the GLCM from ROIs quantized to 8, 16, and 32 intensity levels resulting in GLCMs with sizes of 8×8, 16×16, and 32×32 for each direction α and distance d. The geometric arrangement of considered pixel pairs (p, q) is given by a displacement vector $\delta = (d, \alpha)$ in polar coordinates, where $d \in \{1, 2, 3, 4, 5\}$ defines the distance from the pixel of interest p to any adjacent pixel q. $\alpha = \{0°, 45°, 90°, 135°\}$ defines the horizontal, vertical, and both diagonal directions of δ. After normalization of the GLCM to relative values, we derived contrast, correlation, energy, and homogeneity as statistical features. A detailed description of these measures can be found in Haralick et al. [14]. The combination of intensity quantization and displacement vectors δ results in a 240-dimensional feature space.
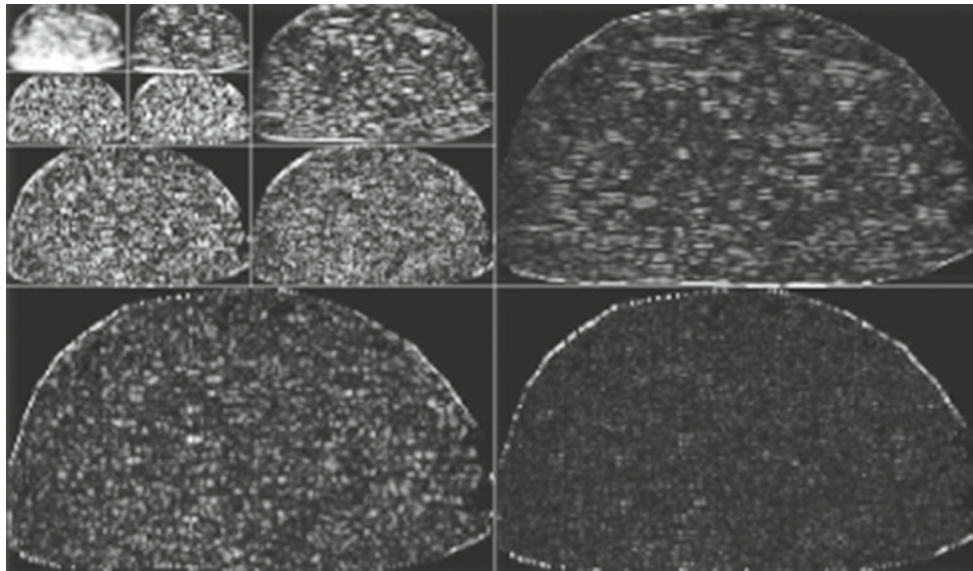
Fig. 4 Results of a three-level 2D-DWT. The LL_3 approximation coefficients of level 3 are located in the *upper left corner*. Horizontal detail coefficients (HL_x) of each level x are shown in descending order from this point toward the *upper right corner*. The same applies for the vertical detail coefficients (LH_x), which are positioned toward the *lower left corner* and for the diagonal detail coefficients (HH_x), which are arranged toward the *lower right corner*

In 2D discrete wavelet transform (2D-DWT), we decomposed the images at different levels with preselected wavelet functions using low-pass (L) and high-pass (H) filters. The 2D-DWT outputs a decomposition vector of approximation coefficients (LL) of the decomposition at the highest level as well as horizontal (LH_x), vertical (HL_x), and diagonal (HH_x) detail coefficients of the decomposition at each level x. Decomposition and reconstruction low-pass and high-pass filters can be derived from different wavelet function families, each having different properties such as filter lengths, symmetry, and bi-/orthogonality. We investigated ten wavelets including Haar/Daubechies (Db: 4, 8), coiflets (Coif: 1, 4), and reverse biorthogonal (RBio: 1.3, 3.1, 3.3, 3.5, 4.4) wavelet families. A detailed description of wavelet families and related properties can be found in the work of Daubechies [9]. We used a three-level 2D-DWT (shown in Fig. 4 for an exemplary ROI) for which we compute the energy that corresponds to approximation coefficients (LL), horizontal (LH_x), vertical (HL_x), and diagonal detail coefficients (HH_x). More precisely, we summed up the squared coefficients for each of the ten sub-images for each wavelet function yielding 100 features. Additionally, we summed up horizontal, vertical, and diagonal energies over all levels for each wavelet function, ultimately resulting in a 130-dimensional feature space.

Feature selection

Overall, we constructed a 375-dimensional feature space, which cannot directly be used for classification because feature vectors will most likely contain unnecessary, correlated, or even misleading information. Therefore, reduction is necessary. Here, the main purpose is the removal of features not contributing to the classification (feature selection). We investigated two feature selection methods: sequential forward feature selection (SFFS) and sequential backward feature selection (SBFS) to determine the features among first-order statistics, Haralick's features, and wavelet-based features that perform best with respect to class discrimination in feature space.

We used two different objective functions for each strategy in order to detect whether the greedy selectors got trapped in a local optimum. We applied sum of squared differences between any two samples of the two classes as our first objective. This leads to selectors for compact features. As a second objective, we applied classification with linear SVM, providing selectors that target for linear separation. The former selectors can be executed much faster than the latter because they evaluate intrinsic properties instead of interactions with a special classifier. Thus, their results exhibit more generality. The latter, however, can achieve higher accuracy, although their solution often lacks generality.

Since the range of feature values varies greatly, the feature selection will not work properly without prior normalization (also known as pre-whitening). Hence, we normalized the range of features by shifting the values to zero mean and rescaling them to unit class variances.

Dependent on the selection strategy and objective function, different wavelet-based and first-order statistical features were ranked with highest score in almost all investi-
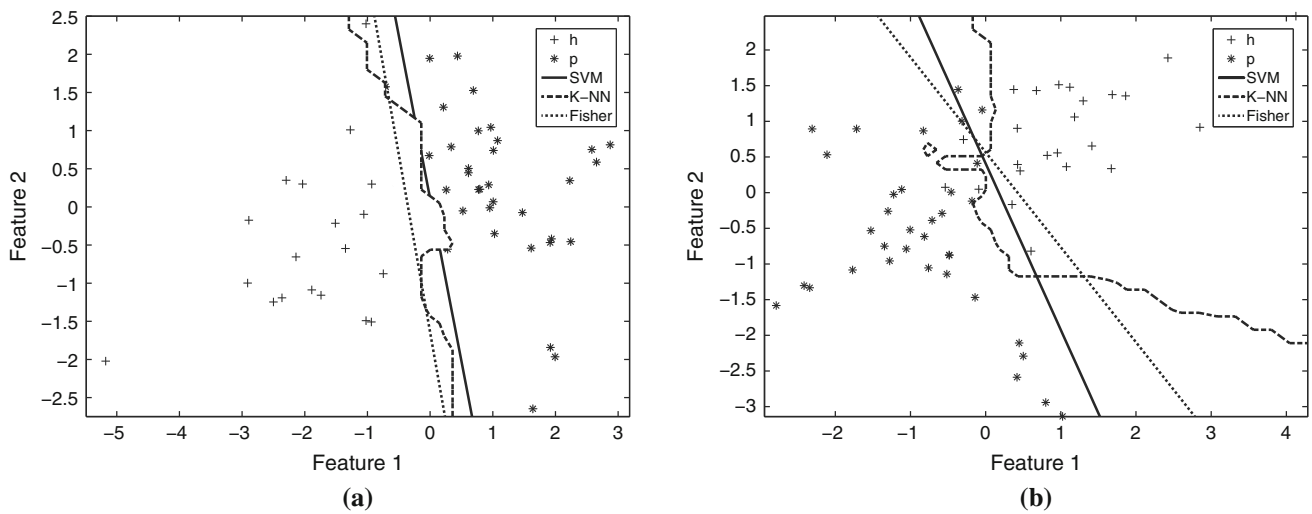
Fig. 5 Scatter plots of a two-dimensional projection of our feature space. **a** Feature combination for polygonal and **b** for rectangular ROIs. The stars represent samples of pathological (p) and the plus samples of healthy (h) muscle tissue. Obtained decision boundaries of k-NN, Fisher's classifier, and the linear SVM are depicted after a training on the whole training dataset

gated cases. However, a stable feature combination could not be found. Contrarily, Haralick's texture features seem to be less suitable to distinguish between healthy and pathological muscle tissue.

In summary, we selected a 135-dimensional feature subspace of wavelet-based features and first-order statistics, excluding 240 Haralick's features from further analysis because they performed inferior in feature selection compared to the former. These results were independent of the ROI definition approach.

Feature reduction

Since our feature space is still high dimensional, we subsequently applied feature reduction to reduce it to a lower-dimensional representation yielding a better generalization. Such data transformation may be linear or nonlinear depending on the class distinction in feature space. From Fig. 5a, b, we can deduce that linear transformation is sufficient. Hence, we decided to apply a supervised principal component analysis (SPCA) on the averaged within-class scatter matrix S_w for our 135-dimensional feature space:

$$S_w = \sum_{i=1}^{c} p_i C_i \qquad (1)$$

with C_i being the covariance matrix of class i, c the number of classes, and p_i the probability of a sample to belong to class i. In the next step, we solve the eigenproblem

$$S_w \mathbf{a} = \lambda \mathbf{a} \Leftrightarrow (S_w - \lambda I)\mathbf{a} = 0 \qquad (2)$$

Table 1 Comparison of the size of the feature subspace after a linear transformation to their preserved variance for both, rectangular and polygonal ROIs

Preserved variance in %	99	95	90	85	80	70	50
Dimensions							
Rect. ROI	22	9	6	4	3	2	1
Poly. ROI	22	9	5	3	2	2	1

and permute results such that eigenvalues are in descending order.

The aim of this reduction step was to find a trade-off between the subspace dimension and the preserved variance in the data. Table 1 details this trade-off for both, feature space of rectangular and polygonal ROIs. It can be seen that 99 % of the data's variance could be preserved using only the 22 most significant eigenvalue directions. This means the feature space could be already reduced by a factor of about six. At 80–90 % of the preserved variance, reduction was saturated. Therefore, we used five most significant eigenvalue directions as features for the next step.

Classification

For classification, we tested three approaches: a nonlinear k-nearest neighbor (k-NN), Fisher's linear discriminant, and a linear support vector machine (SVM). k-NN selects the class membership of a sample according to the majority of class memberships of its k nearest neighbors in the training data. As distance metric, we used the Euclidean distance. We optimized parameter k using the leave-one-out error of the dataset for both the rectangular and polygonal ROIs. Fisher's

Table 2 Leave-one-out error rates for trained classifiers

Classifier	Poly. ROI	Rect. ROI
SVM	0.050	0.083
k-NN	0.050	0.083
Fisher	0.100	0.117
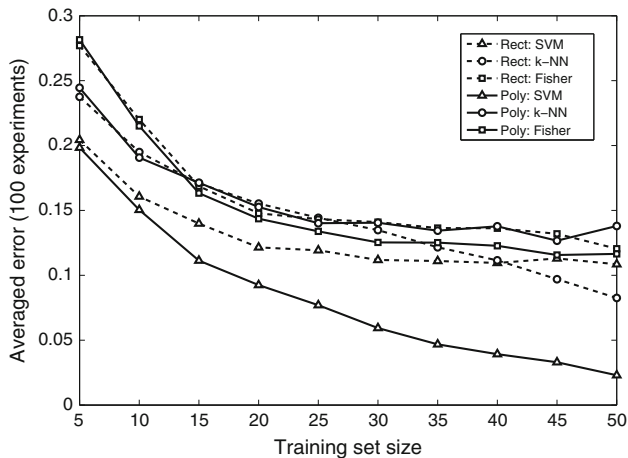
Fig. 6 Averaged classification error for x-fold cross-validation on rectangular (*dashed lines*) and polygonal ROIs (*solid lines*) compared for 100 stratified repetitions each

least square linear classifier attempts to find a linear discriminant function between the two classes along which the ratio of between- and within-class scatter is maximized in the least squares sense [10,36]. The linear SVM seeks a decision boundary that maximizes the distance to the nearest class sample to both of its sides. Therefore, a quadratic programming problem must be solved.

Decision boundaries, resulting from training k-NN, Fisher's classifier, and linear SVM on a selected feature subset of our training dataset, are shown in Fig. 5. It can be seen that the linear SVM adequately separates healthy and pathological muscle tissue. The decision boundary of k-NN also seems to provide good separation in contrast to that of the Fisher's classifier. At first glance, separation of polygonal ROIs (see Fig. 5a) seems to be slightly better than that of the rectangular ROIs (see Fig. 5b). The leave-one-out classification results are shown in Table 2. Feature reduction and classifier training were performed on the $N - 1$ images of the training dataset in each round of validation. Figure 6 indicates the general trend for cross-validation with smaller training set sizes. Here, the linear SVM outperforms Fisher's classifier yielding almost similar results when compared to k-NN. Hence, we chose a linear SVM, which provides a more generalized and less overfitted decision.

Performance measurements

To obtain a fair judgement of the performance of our CAD system, our performance measurements were undertaken

independently by two expert radiologists (E1, E2) classifying our data by the single rectangular and polygonal ROI approach as well as by our CAD system. The resulting dataset was denoted as validation dataset. Note that all pre-study experiments regarding the classification process, i.e., feature selection, reduction, and classifier training (see Sects. 3.3, 3.4, or 3.5), were based on the training dataset and thus on the ROI definition of another field expert (E0). Furthermore, to reduce any mutual dependence, we performed leave-one-out cross-validation. Hence, we excluded the current ROI of the training dataset, our ground truth, from feature reduction and classifier training in a leave-one-out manner. Therefore, our SPCA and linear SVM training were performed on $N - 1$ ROI definitions of our training data defined by expert E0, where N is the number of total ROIs. Performance was measured for each ROI of our validation dataset defined by experts E1 and E2, finally producing averaged values per expert after all rounds of validation. An ROI-wise leave-one-out cross-validation should be appropriate since texture features of images of the same subject should not be highly correlated due to our proposed standardized image acquisition procedure (see, e.g., Fig. 2).

We used well-established performance measures namely, sensitivity and specificity for the evaluation. Sensitivity, also called true positive rate, measures the proportion of pathological cases which are correctly classified as such. Specificity, also called true negative rate, measures the proportion of healthy cases which are correctly classified as such by our CAD system. Additionally, we calculated the accuracy of our CAD system as overall performance measure since the number of class instances of healthy almost equals that of pathological ROIs.

Results and discussion

CAD system

Figure 7 shows the sequence of operations performing in our CAD system. The first step (1) of our system is based upon an one-click approach where the radiologist selects the center of the relevant muscle area. To overcome imprecise ROI definitions and to increase the robustness of our approach, we additionally evaluate eight new ROIs which are shifted versions (along x-, y- and both diagonal axes) of the placed single rectangular ROI. Figure 8 outlines the shifting of the ROI. The amount of shifting depends on the chosen size of the single rectangular ROI and typically ranged from 1.5 to 4.0 mm. Because of the curved shapes of biceps brachii in the ultrasound images, our shifted ROIs cover most of the relevant muscle tissue. For each of the ROIs, five first-order statistical and 130 wavelet-based features are extracted; the feature space is reduced by the trained SPCA and then clas-
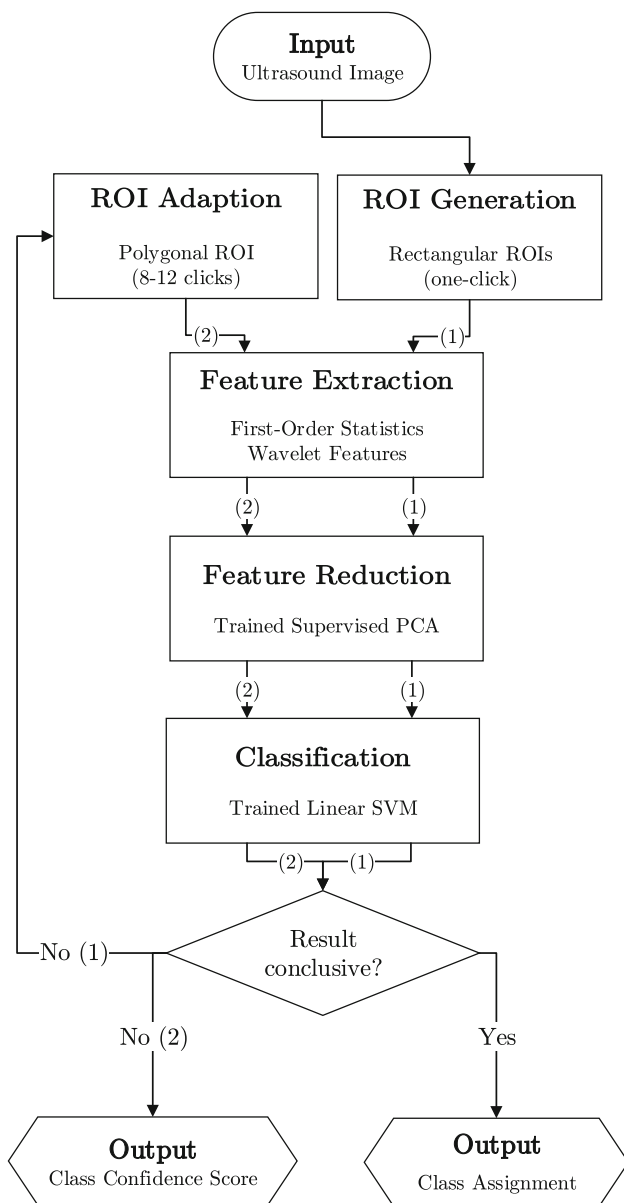
Fig. 7 Workflow of the proposed CAD system

Fig. 8 ROI shifting scheme used by our CAD system. The *arrows* illustrate the eight directions in which the rectangular ROI was shifted, producing eight new shifted versions of the ROI with unchanged size

final decision depends on the majority of the non-discarded ROIs. We used this majority voting scheme to overcome imprecise ROI definitions and to increase the robustness of our approach. If the resulting decision of the voting scheme is conclusive, our CAD system outputs the class assignment for the defined ROI. If not, the radiologist is prompted to define a more exact polygonal ROI (2) in a second step. The feature extraction, feature reduction, and classification methods used for the polygonal ROI are equal to those of our one-click voting scheme. In case that the CCS of the polygonal ROI is still inconclusive after classification, no class assignment is produced by our CAD system, but the resulting CCS is displayed as a guidance for the radiologist.

Experimental results

Results of our experiments are given in Table 3. We also included the results of our one-click voting scheme without the on-demand polygonal ROI definition (voting only) to show the impact of the voting scheme. Moreover, the number of inconclusive ROIs is indicated and all results are given for both radiologists, E1 and E2.

As expected, accuracy, sensitivity, and specificity was worst using single rectangular ROIs (72–81 %) and best using polygonal ROIs (92–97 %). However, the user effort of polygonal ROI definition is substantially higher than that of single rectangular ROIs, which renders its application impractial in clinical routine. It can be seen that the accuracy, sensitivity, and specificity of the voting scheme (82–89 %) were higher than that of a single rectangular ROI, which is especially true for the second radiologist, E2. This confirms the need for an improved stability of the user input. Moreover, our voting scheme produced a number of inconclusive images, which could not be reliably classified. These images

sified by the trained linear SVM. Classifying each of the nine ROIs separately allows us to stabilize the final decision.

To estimate whether the classification is reliable, we map signed distances to the linear SVM boundary using the sigmoidal function $(1 + \exp(-d))^{-1}$. This provides us with a class confidence score (CCS) allowing us to judge how likely it is that an unknown sample belongs to either of the classes. In our CAD system, we used the CCS to identify inconclusive ROIs for which a distinct class assignment was not feasible. If the CCS ranges from $\frac{1}{3}$ to $\frac{2}{3}$, the considered ROI is set as inconclusive.

Immediately after the classification of the nine ROIs, our CAD system evaluates the CCS for each of them, whereby the
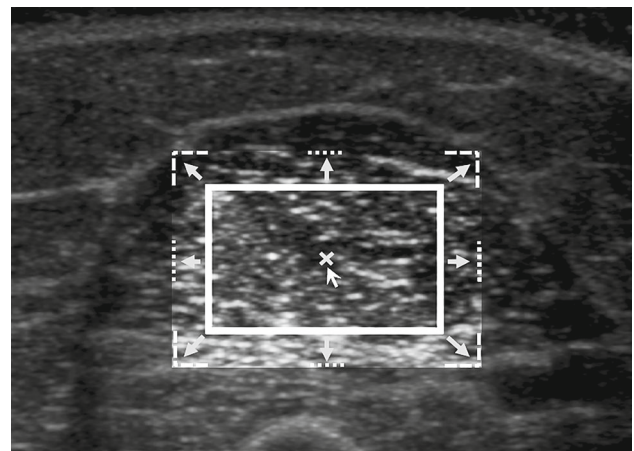
Table 3 Accuracy, sensitivity, specificity, and the number of inconclusive images (#Inconclusive) for single rectangular and polygonal ROIs, our voting scheme only, and our CAD system

Approach	#Inconclusive	Accuracy (%)	Sensitivity (%)	Specificity (%)
Rectangular ROI	–	81/77	80/72	81/80
Polygonal ROI	–	95/95	92/92	97/97
Voting scheme only	4/6	84/85	89/88	82/84
CAD system	0/0	85/87	90/90	83/85

ROIs were placed by two expert radiologists (E1/E2)
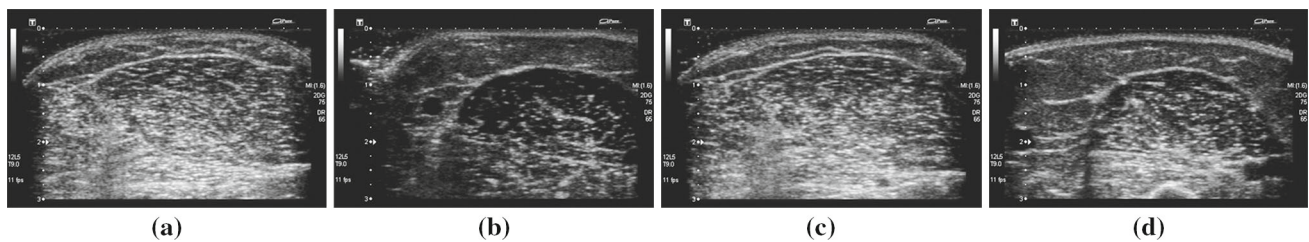
Fig. 9 Examples of incorrectly classified images of healthy (**a**) and pathological (**b**) muscle tissue in comparison with correctly classified images of healthy (**c**) and pathological (**d**) muscle tissue

were further processed with our CAD system, using polygonal ROIs.

Our system finally achieved 85/87 % accuracy, 90 % sensitivity, and 83/85 % specificity, depending on the expert radiologist. Moreover, the number of inconclusive images could be reduced to zero for both expert radiologists. Since the achieved values for both expert radiologists are almost equal, it can be concluded that the influence of user input variability was successfully reduced by our CAD system, although, mouse click positions of both expert radiologists were quite different. Moreover, nearly all of the falsely classified images are the same for both radiologists.

Referring to the problems discussed in Sect. 1 (see Fig. 1), we were still able to classify 17 of 25 (68 %) visually hard distinguishable muscle images correctly, e.g., the ultrasound images shown in Fig. 9c, d. However, the figure also shows exemplary images for which our classification approach fails. The images shown in Fig. 9a, b were falsely classified due to the fact these do not meet our previously defined assumptions: texture of healthy muscle tissue is more oriented and structured, whereas pathological muscle tissue appears to be diffuse and unstructured. In Fig. 9b, the scan of the pathological muscle tissue might be mistaken to appear healthy, because myositis was captured in an early stage or because there is no affected muscle tissue in this particular scan of the muscle due to the fact that myositis may often occur focally. Contrarily, the healthy scan from Fig. 9a seems to be very diffuse caused by the high amount of fat and connective tissue occurring in elderly people. To overcome this problem, we might include a patient-specific weight or fat content parameter into the classification process.

Conclusion

We developed a computer-aided diagnosis (CAD) system for detecting neuromuscular diseases in ultrasound images, providing radiologists an independent opinion and a tool that helps training young professionals. We investigated this at the challenging example of myositis detection in scans of the biceps brachii. In our CAD system, rectangular regions-of-interest (ROIs) are placed with a single click, triggering an automatic classifier that was learned on training samples. To deal with an potential misplacement of the ROI, e.g., when non-muscle tissue is included, we evaluated several spatially shifted ROIs and established a majority voting scheme to provide a stabilized decision. The decision of our CAD system is based on a class confidence score, classifying into healthy, pathological, and inconclusive cases. Regarding the latter, the radiologist is asked to specify a more detailed polygonal ROI using a small number (8–12) of clicks on which the actual decision is based. Thus, we combined minimum effort of the one-click solution, which works for most cases, with a highly accurate polygonal ROI, when needed.

For automatic classification, our CAD system uses a combination of feature extraction, feature reduction, and classification. Within this work, we investigated three kinds of features for the classification task: first-order statistics, wavelet-based features, and Haralick's features, all of which measure different important aspects of the texture of the muscle tissue, i.e., its contrast and structure, and therefore were used in several other classification tasks as well [21,24,33]. In a preliminary feature selection step, sequential forward and backward feature selection were used and it was shown that

Haralick's features contribute little, if at all, to the classification. Therefore, we proposed to use only first-order and wavelet-based features. Moreover, we showed that most of the information can be preserved with as little as five linearly combined feature space dimensions, using a supervised principal component analysis for feature reduction. The final five-dimensional feature space separates well between healthy and pathological classes, which was shown for three classifier: Fisher's linear classifier, a linear support vector machine, and a nonlinear k-nearest neighbor algorithm. The latter performed equally well in our comprehensive experiments, while the former performed worse. We proposed to use a linear support vector machine due to its beneficial in generalizability (see Sect. 3.5). Finally, our approach achieves high sensitivity, specificity, and accuracy for single rectangular ROIs (72–81%, radiologist-dependent) and even higher values for single polygonal ROIs (92–97%, radiologist-dependent). We also investigated the one-click voting scheme that our CAD system uses. Here, the user is asked to specify the polygonal ROI only if the decision for the one-click voting scheme is inconclusive. As expected, performance of our CAD system (83–90%, radiologist-dependent) lies midway between those of single rectangular and polygonal approaches, but requires substantially lower effort than the polygonal ROI definition. Further work will include the application of our CAD system to other muscles, other types of neuromuscular diseases, and other ultrasound devices. Moreover, an inclusion of patient-specific parameters, e.g., fat content and weight, into the classification process will be of interest.

References

1. Asvestas P, Golemati S, Matsopoulos GK, Nikita KS, Nicolaides AN (2002) Fractal dimension estimation of carotid atherosclerotic plaques from B-mode ultrasound: a pilot study. Ultrasound Med Biol 28:1129–1136

2. Awad J, Krasinski A, Parraga G, Fenster A (2010) Texture analysis of carotid artery atherosclerosis from three-dimensional ultrasound images. Med Phys 37:1382–1391

3. Basset O, Sun Z, Mestas J, Gimenez G (1993) Texture analysis of ultrasonic images of the prostate by means of co-occurrence matrices. Ultrason Imaging 15:218–237

4. Basset O, Duboeuf F, Delhay B, Brusseau E, Cachard C, Tasu J (2008) Texture analysis of ultrasound liver images with contrast agent to characterize the fibrosis stage. In: Proceedings of IEEE ultrasonics symposium, pp 24–27

5. Chang RF, Wu WJ, Moon W, Chen DR (2005) Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. Breast Cancer Res Treat 89:179–185

6. Chen DR, Chang RF, Huang YL (1999) Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. Radiology 213:407–412

7. Chen L, Seidel G, Mertins A (2010) Multiple feature extraction for early Parkinson risk assessment based on transcranial sonography image. In: Proceedings of international conference on image processing, pp 2277–2280

8. Christodoulou C, Pattichis C, Pantziaris M, Nicolaides A (2003) Texture-based classification of atherosclerotic carotid plaques. IEEE Trans Med Imaging 22:902–912

9. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics

10. Duda R, Hart P, Stork D (2001) Pattern classification. Wiley, London

11. Duin R, Juszczak P, Paclik P, Pekalska E, de Ridder D, Tax D, Verzakov S (2013) A MATLAB toolbox for pattern recognition. Delft Univ. Techn., http://prtools.org/, pRTools 4.2.4

12. Finette S, Bleier A, Swindell W (1983) Breast tissue classification using diagnostic ultrasound and pattern recognition techniques: I. Methods of pattern recognition. Ultrason Imaging 5:55–70

13. Good MS, Rose JL, Goldberg BB (1982) Application of pattern recognition techniques to breast cancer detection: Ultrasonic analysis of 100 pathologically confirmed tissue areas. Ultrason Imaging 4:378–396

14. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3:610–621

15. Huang Y, Han X, Tian X, Zhao Z, Zhao J, Hao D (2010) Texture analysis of ultrasonic liver images based on spatial domain methods. In: Proceedings of international congress on image signal processing, vol 2, pp 562–565

16. Huang YL, Wang KL, Chen DR (2006) Diagnosis of breast tumors with ultrasonic texture analysis using support vector machines. Neural Comput Appl 15:164–169

17. Kier C, Seidel G, Brüggemann N, Hagenah J, Klein C, Aach T, Mertins A (2009) Transcranial sonography as early indicator for genetic Parkinson's disease. In: European conference of the international federation for medical and biological engineering, vol 22. Springer, Berlin, pp 456–459

18. König T, Rak M, Steffen J, Neumann G, von Rohden L, Tönnies KD (2013) Texture-based detection of myositis in ultrasonographies. Proceedings of Bildverarbeitung für die Medizin (BVM), Informatik aktuell, pp 81–86

19. Layer G, Zuna I, Lorenz A, Zerban H, Haberkorn U, Bannasch P, van Kaick G, Räth U (1990) Computerized ultrasound B-scan texture analysis of experimental fatty liver disease: influence of total lipid content and fat deposit distribution. Ultrason Imaging 12:171–188

20. Lee WL, Chen YC, Hsieh KS (2003) Ultrasonic liver tissues classification by fractal feature vector based on m-band wavelet transform. IEEE Trans Med Imaging 22:382–392

21. Liao YY, Tsui PH, Li CH, Chang KJ, Kuo WH, Chang CC, Yeh CK (2011) Classification of scattering media within benign and malignant breast tumors based on ultrasound texture-feature-based and Nakagami-parameter images. Med Phys 38:2198–2207

22. Lu NH, Kuo CM, Ding HJ (2013) Automatic ROI segmentation in B-mode ultrasound image for liver fibrosis classification. In: Proceedings of international symposium on biometrics and security technologies (ISBAST), pp 10–13

23. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693

24. Minhas FU, Sabih D, Hussain M (2012) Automated classification of liver disorders using ultrasound images. J Med Sys 36:3163–3172

25. Mohamed S, Salama M (2008) Prostate cancer spectral multi-feature analysis using TRUS images. IEEE Trans Med Imaging 27:548–556

26. Ogawa K, Fukushima M, Kubota K, Hisa N (1998) Computer-aided diagnostic system for diffuse liver diseases with ultrasonography by neural networks. IEEE Trans Nucl Sci 45:3069–3074

27. Pillen S, van Alfen N (2010) Skeletal muscle ultrasound. Eur J Transl Myol 1:145–155

28. Pohle R, Fischer D, von Rohden L (2000) Computer-supported tissue characterization in ultrasound images of neuromuscular diseases. Ultrasound Med 21:245–252

29. Reimers K, Reimers CD, Wagner S, Paetzke I, Pongratz DE (1993) Skeletal muscle sonography: a correlative study of echogenicity and morphology. J Med Ultrasound 12:73–77

30. von Rohden L, Wien F, Ptzsch S (2007) Myosonographie neuro-muskulärer Erkrankungen unter besonderer Berücksichtigung des Kindes- und Jugendalters. Klin Neurophysiol 38:141–154

31. Sakalauskas A, Lukoševičius A, Laučkaite K (2011) Texture analysis of transcranial sonographic images for Parkinson disease diagnostics. Ultragarsas 66:32–36

32. Sakr AA, Fares ME, Ramadan M (2014) Automated focal liver lesion staging classification based on Haralick texture features and multi-SVM. Int J Comput Appl 91(8):17–25

33. Sheppard M, Shih L (2007) P6C-2 image texture clustering for prostate ultrasound diagnosis. In: Proceedings of IEEE ultrasonic symposium, pp 2473–2476

34. Su Y, Wang Y, Jiao J, Guo Y (2011) Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features. Open Med Inf J 5:26–37

35. Tsiaparas NN, Golemati S, Andreadis I, Stoitsis JS, Valavanis I, Nikita KS (2011) Comparison of multiresolution features for texture classification of carotid atherosclerosis from B-mode ultrasound. IEEE Trans Inf Technol Biomed 15:130–137

36. Webb AR (2002) Statistical pattern recognition. Wiley, New York

37. Wilhjelm J, Gronholdt ML, Wiebe B, Jespersen S, Hansen L, Sillesen H (1998) Quantitative analysis of ultrasound B-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination. IEEE Trans Med Imaging 17:910–922

38. Wu CM, Chen YC, Hsieh KS (1992) Texture features for classification of ultrasonic liver images. IEEE Trans Med Imaging 11:141–152

39. Wu WJ, Lin SW, Moon WK (2012) Combining support vector machine with genetic algorithm to classify ultrasound breast tumor images. Comput Med Imaging Graph 36(8):627–633

40. Zhou S, Shi J, Zhu J, Cai Y, Wang R (2013) Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image. Biomed Signal Process Control 8(6):688–696