Classification of Medical Ultrasound Transducer Using Neural Network

Laila M Aboughazala, Kamel. K. Mohammed*

Center for Virus Research and Studies, Al-Azhar University, Cairo, Egypt

***E-mail: vrsc@azhar.edu.eg**

Abstract

Diagnostic medical ultrasound makes a significant contribution to patient care and is increasingly used in a variety of clinical settings by many different professionals with varying technical backgrounds (e.g. hepatocellular carcinoma). Therefore, the importance of ultrasound quality control is not only necessary for patient and operator safety but is also essential for maintaining the performance of the equipment to the highest-level achievable and it is required by various regulatory and accrediting agencies. Ultrasound image degradation originates primarily from transducer defects and potentially undermines reliable image interpretation. The ultrasound probe in-air reverberation pattern is used in routine quality assurance. We produce a quantitative quality control based on in-air reverberation images. They are easily generated for any probe independent of the level of expertise of the operator. The results are available to the sonographer prior to clinical use and transducer status can be remotely monitored with trend analysis over time. The method presents a scheme for the classification of normal functioning and defect transducers Region of Interest selected "ROIs" of 65 probes based on texture analysis that automatically detects in-air reverberation regions and recognizes them as normal functioning and defect transducers. However, feature selection is done by Twin Support Vector Machine. The accuracy of these features in distinguishing normal functioning and defect transducers has been evaluated by artificial neural network, and linear support vector machine algorithms classifiers. From the analysis of results, it was found that artificial neural network classifier gave an overall classification accuracy of 100% with 100% sensitivity. The results show that it is feasible to identify defect transducers based on texture features extracted from in-air reverberation ultrasound images. This method is shown to be useful for increased accuracy and increased speed for classification of functioning and defect transducers for improving the quality assurance of ultrasound.

Keywords: Quality assurance, in-air reverberation ultrasound images, Twin Support Vector Machine (TWSVM), artificial neural network (ANN), support vector machine (SVM).

1. Introduction

Ultrasound is one of the most widely used non-invasive imaging techniques in medical diagnosis. For example, in the year from July 2016 to July 2017, over 9.2 million ultrasound scans were carried out on National Health Service patients in England. This is almost twice the number of computed tomography scans and almost three times the number of magnetic resonance imaging scans carried out in the same period [1]. Diagnostic ultrasound (US) images can be obtained from a quality device, in optimal working conditions, combined with the capable actions of the

operator in tweaking the equipment's characteristics. Therefore, performance evaluation or quality assurance (QA) of ultrasound (US) equipment is necessary, as for any other medical imaging equipment, for ensuring the safety of the patient and operator and for complying with the requirements by various regulatory and accrediting agencies (*e.g.,* American Institute of Ultrasound in Medicine AIUM 2008; Goodsitt et al. 1998; Kollmann et al. 2012; Russell et al. 2010) [2] then the lack of an effective QA program introduces an unnecessary risk of diagnostic errors. Several previous studies have shown high fault rates in US probes that are in clinical use. Hangiandreou et al. [3] reviewed the results of a four-year quality control program in a single large radiology department, including more than 300 probes. They found that 165 probes failed (defined as a problem that required immediate repair or replacement) over the period, with an average annual failure rate of 13.9%. Probe failure represented 88% of total failures, the remainder being scanner component failures and the probe failure detected using image uniformity that it was assessed by looking for artefacts in images of a tissue mimicking test object (TMTO) and the in-air reverberation pattern. Martensson et al. [4] tested 676 probes using an electronic probe tester and found faults in 269 probes (40%). Martensson et al. [5] tested 299 probes in 12 hospitals a year after their earlier study and found that 81 were defective (27.1%). These probes had either passed the test a year previously or had been purchased to replace defective probes. The distribution of fault types was similar to their previous study, with delamination (50%) and cable faults (35%) being the most common. They found the highest proportion of defective probes in radiology departments (36%) and that a factor in this was the way that probes were handled in different clinics, a higher failure rate being associated with probes being disconnected and stored away from the scanner when not in use, rather than being stored on the scanner. Sipila et al. [6] tested 135 probes using an electronic probe tester and with a TMTO to assess image uniformity and made a physical inspection of each probe, finding a total of 52 faulty probes (39%). A total of 21 faults (40% of total faulty probes) were detected by the electronic probe tester, 20 (38%) using the TMTO and 34 (65%) by physical inspection. Three faults (6% of total faulty probes) were demonstrated only by the electronic probe tester, eight (15%) only by the TMTO and 21 (40%) only by physical inspection. Probe faults are therefore common and important to detect. Electronic probe testers such as FirstCall (Unisyn, Golden, CO, USA) provide comprehensive results that both detect faults and indicate their likely origin (Martensson et al. 2009, 2010; Sipila et al. 2011), but it has its price, as a suitable adapter is required for each type of transducer. Since an average hospital is equipped with numerous ultrasound systems, each equipped with multiple transducers, systematic quantitative quality control is often neglected due to the limited resources available for such an extensive task. Quality of clinical diagnosis is at risk once transducer failure is unnoticed or visual assessment underestimates the severity. Without structural quantitative measurements, any form of trend analysis of transducer degradation is impossible. We propose an automated analysis approach of quantitative quality control based on the reverberation image with the transducer operated in-air to overcome these difficulties. This image consists of a series of consecutive lines which are generated by internal reflections of ultrasound due to the large acoustic impedance mismatch between the front face of the transducer and air. This image can be generated independent of the level of expertise of the operator for any (curvi-) linear transducer by simply activating the ultrasound system. Visual inspection of the in-air reverberation image is a standard qualitative routine assessment of scanner

performance regarding uniformity and sensitivity as recommended by the Institute of Physics and Engineering in Medicine [7]. In this paper, the senior contribution focused on the automated classification of ultrasound transducer images as the normal or defect transducer. The proposed method subdivided into four sequential phases which are preprocessing phase, feature extraction, feature selection, and classification methods independent of ultrasound vendor. Proof of principle of the technique is demonstrated based on data obtained from a series of transducers. The presented approach resolves the current limitations of conventional periodic quality control and enables hospital wide status monitoring of ultrasound systems.

2. MATERIALS AND METHODS

2.1. Image acquision

The in-air reverberation images can be obtained with the transducer in the holder on most systems. In the air test, the transducer is put in function in the air with a clean and dry face in its routine clinical pre-set with the frequency in its lowest value and the gain in the maximum level (both overall gain and time gain compensation) [8]. The image obtained is one of a series of horizontal bands of reverberation due to the transducer–air impedance mismatch. These series of reverberations are used for a first line subjective evaluation assessing eventual faulty crystal elements (vertical line of reduced echoes is seen), uniformity of image, transducer sensitivity and image noise. The transducer surface should be clean of any ultrasonic gel remnants prior to imaging as shown in figure 1. The imaging protocol can be extended to apply a standardized form of manual moving and rotating the transducer to evoke possible intermittent failures due to cable breaks or crystal contact issues before image acquisition. The resulting dataset consists of 24 no fault ultrasound transducers and

41 fault ultrasound transducers by using 65 in-air reverberation ultrasound images that obtained from [9] as shown in Figure 2.
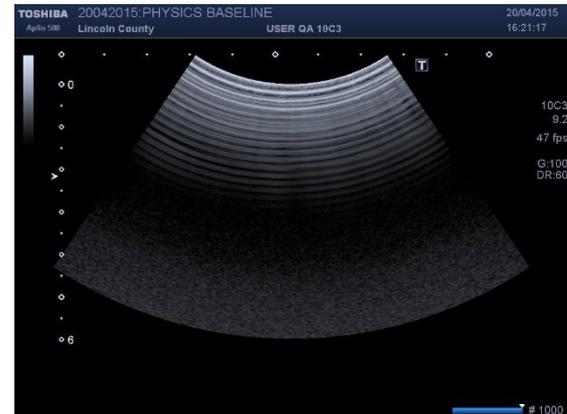
Figure 1. A normal an in-air reverberation image.
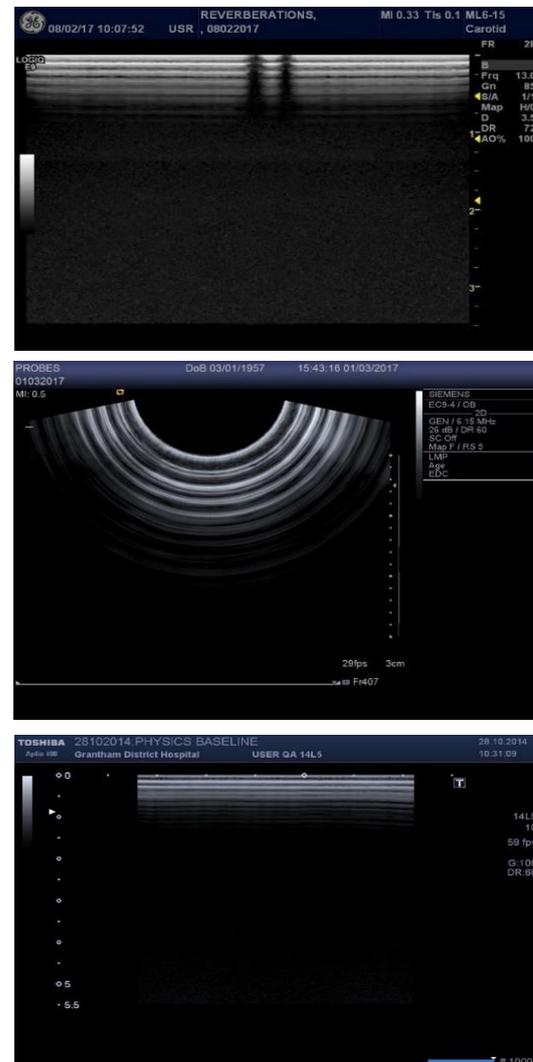
Figure 2. An in-air reverberation ultrasound images from different vendors.

2.2. Neural Network

A back propagation neural network (BPNN) is used for classification of normal or defect transducer. Back-propagation is actually the nearly common method for accomplishing the supervised learning task. In supervised learning, we attempt to adjust an artificial neural network in which its output reaches close to specific target output for a training set. The aim is to adjust the parameters of the network (i.e. weights and biases) where it implements well for patterns from outside the training set. Hidden layer size play an important role in achieving the better desired performance. Smaller number of hidden neurons makes the hidden neurons unstable, whereas larger number of hidden neurons makes the output neurons unstable. An upper bound on the number of hidden neurons to avoid over fitting is considered as:

$$N_h = \frac{N_s}{\alpha(N_i + N_o)}$$

Where N_i is number of input neurons, N_o is number of output neurons, N_s is number of samples in training and α is between 2 to 10. Lower bound is considered as the mean value of the input and output neurons. The designed neural network classifier was used consists of twenty hidden layer feedforward back propagation network and one output neuron. Twenty hidden layer feedforward network with sigmoid transfer function and one output neuron with linear transfer function.

2.3. Support Vector Machine

It is a supervised learning classifier commonly used for binary classification. It tries to linearly discriminate the two class features by fitting a hyperplane between them. The training phase of SVM aims to find the hyper plane with the largest margin that helps to separate the high dimensional feature space with less classification error.

The weight and bias vectors are derived by using a cost minimization function. For good generalization performance with SVM classifier the correct choice of kernel function is important. In this paper linear kernel function is found to be most appropriate. Consider training N points, where each input x_i has "A" attributes and is in one of two classes. $y_i = +1$ or -1 (i.e. training data is of the form: (x_i, y_i), i = 1,2,...N , $y_i \in \{+1, -1\}$. The main purpose of using Support Vector Machine (SVM) is to orientate hyperplane in such a way as to be as far as possible from the closest members of both classes. Training data can be described by:

(w. x_i+b) $\geq$ 1 for $y_i = +1$

(w. x_i+b) $\leq$ 1 for $y_i = -1$
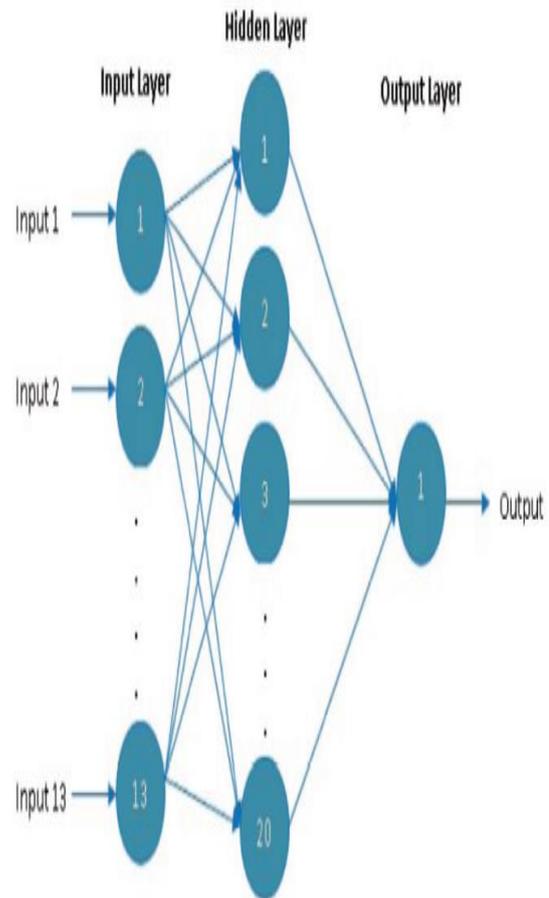
Figure 3. Architecture of the proposed back propagation neural network.
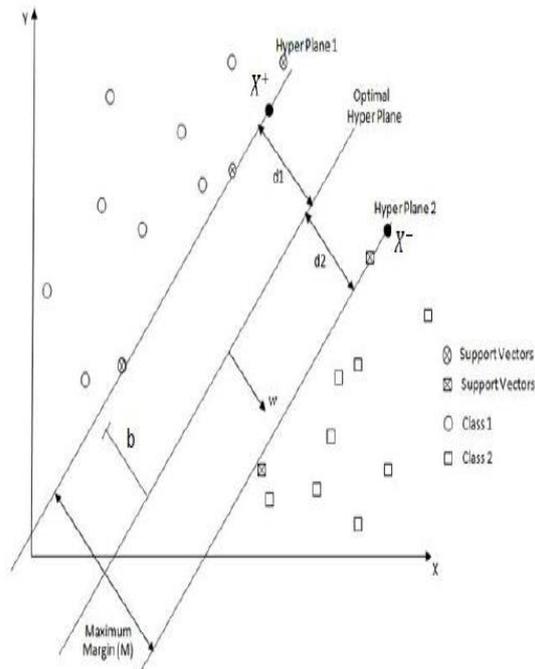
Figure 4: Linear separable binary classification.

3. The Proposed method system for ultrasound transducer classification

The proposed model for classification of normal or defect transducer database is summarized as in Fig.5.

3.1.Preprocessing Phase

The software will require some parameters immediately after the ultrasound instrument has been installed and accepted. The parameters will be related to the instrument vendor (e.g. TOSHIBA, PHILIPS, SIEMENS and GE) to represent the ultrasound data of all transducers in a data field that includes only in-air reverberation image as shown in figure 6. The parameters are x and y coordinates of in-air reverberation image that are top-left corner of a rectangular in-air reverberation image and the size of a rectangular data field image. The parameters are only defined one time immediately after the ultrasound instrument has been installed and accepted.

Figure 5. Proposed model for classification of in-air reverberation ultrasound images database

.
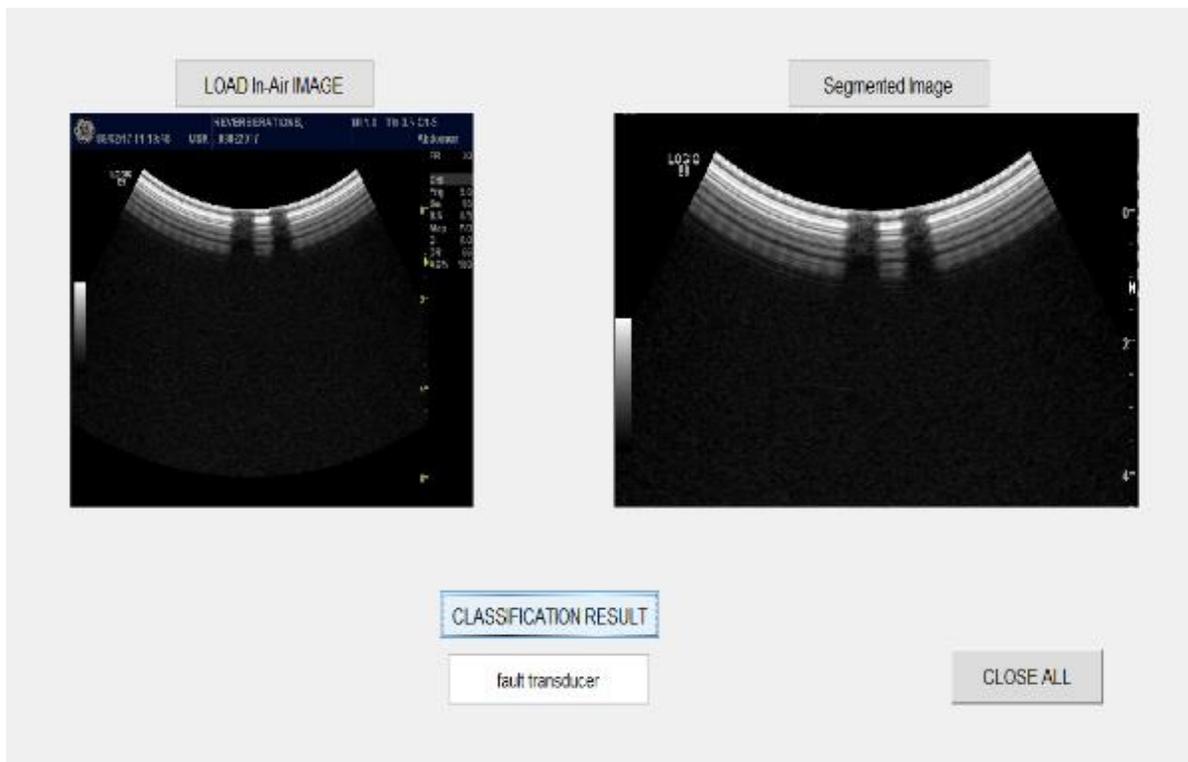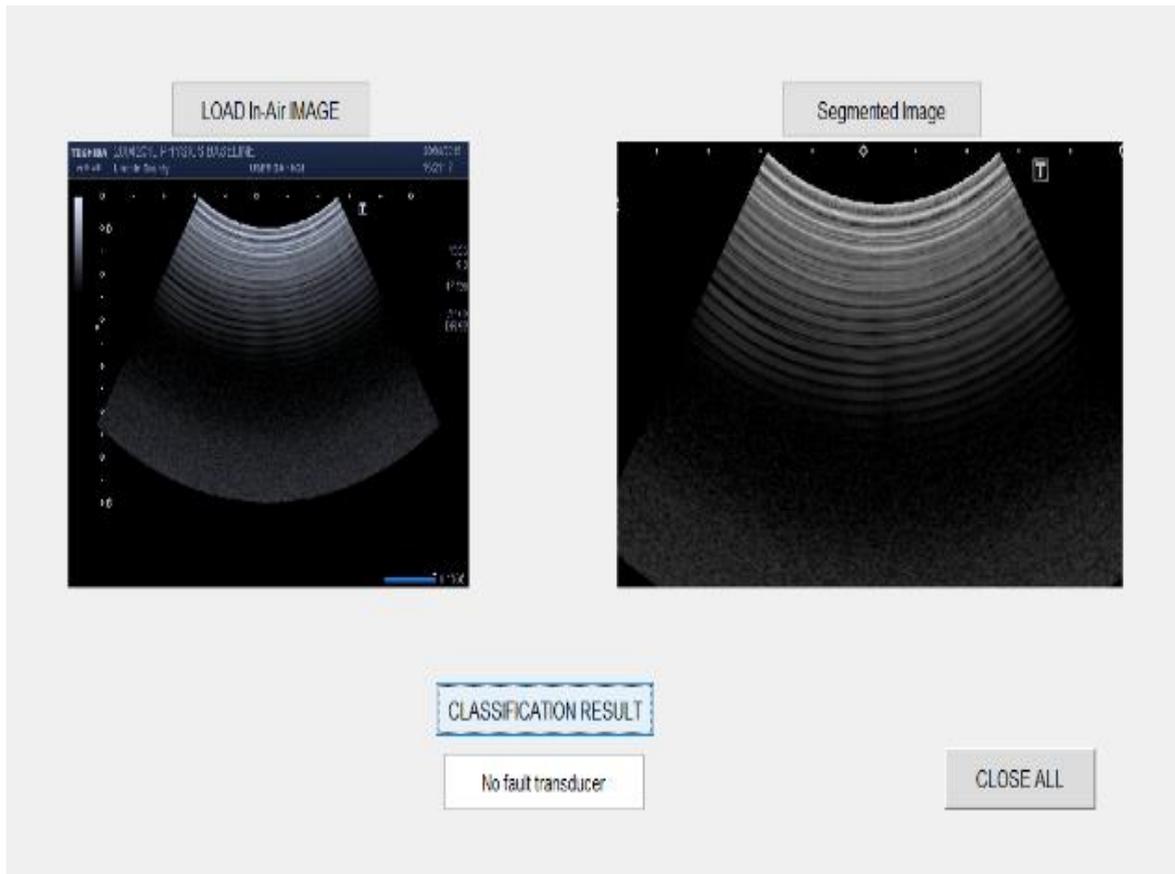
Figure 6. Proposed model for classification of in-air reverberation ultrasound images database.

3.2.Feature Extraction Phase

After image preprocessing, the statistical features extraction are 46 features that divided into 6 features for intensity histogram [10], 22 features for gray-level co-occurrence matrix (GLCM) [11], 11 features for gray-level run length matrix (GLRLM) [12], and 7 features for invariant moments [13]. The intensity histogram features consist of mean, standard deviation, average energy, entropy, skewness, contrast and kurtosis. GLCM features consist of autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity energy, entropy, homogeneity (1), homogeneity (2), maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation (1), information measure of correlation (2), and inverse difference normalized.

The GLRLM features consist of short run emphasis (SRE), long run emphasis (LRE), run length nonuniformity (RLN), gray level nonuniformity (GLN), run percentage (RP), low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), short run low gray-level emphasis (SRLGE), short run high graylevel emphasis (SRHGE), long run low gray-level emphasis (LRLGE), and long run high gray-level emphasis (LRHGE). The Invariant moments are invariant under translation, changes in scale, and rotation.

3.3.Feature Selection Phase

All features extraction are transferred to a feature selection algorithm that based on Twin Support Vector Machine (TWSVM) [14] to select the optimal features for functioning and defect transducers images. The TWSVM algorithm is based on study two nonparallel hyperplanes such that each hyperplane is close to one of the two classes and leave from the other together which corresponds to two weights for each feature resulting some difficulty for feature selection. Thus, a feature selection matrix is introduced, aiming at realizing feature selection. This approach leads to solve a multi-objective mixed integer programming problem by an alternate iterative greedy algorithm

3.4.Classification phase

The classification step is the final step in our model where the extracted features from feature selection phase are the input to the classifier, while the output is in-air reverberation ultrasound transducer image class. This classifier uses only thirteen features selected based on (TWSVM) as shown in figure 3. All feature selection is applied to two classifiers such as artificial neural network (ANN) and linear Support Vector Machine (SVM) to differentiate functioning and defect transducers. In the proposed model, the ANN [15] is used for back-propagation algorithm and SVM [16] is used for binary classification. The performance of the ANN classifier is determined by accuracy level and confusion matrix while the performance of the SVM classifier is determined by accuracy level.

3.5.Performance Measures

Since we are studying a detection problem, we estimate the performance of classifiers by considering the confusion matrix for a dichotomous problem where the testing results can be divided into four categories. Usually, an image region can be called normal transducer (negative) or defect transducer (positive), and a decision for a detection result can be either correct (true) or incorrect (false).

True positive (TP): The classifier yields positive result for an in-air reverberation image of ultrasound transducer and an in-air reverberation image of ultrasound transducer has actually fault transducer.

False positive (FP): The classifier yields positive result for an in-air reverberation

image of ultrasound transducer, but an in-air reverberation image of ultrasound transducer does not actually defect transducer.

True negative (TN): The classifier yields negative result for an in-air reverberation image of ultrasound transducer and an in-air reverberation image of ultrasound transducer has actually normal transducer.

False negative (FN): The classifier yields negative result for an in-air reverberation image of ultrasound transducer, but an in-air reverberation image of ultrasound transducer has actually defect transducer.

Sensitivity is the conditional probability of detecting defect transducer head while there are really defect transducer.

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity is the conditional probability of detecting normal transducer head while the true state of normal transducer.

$$Specificity = \frac{TN}{TN + FP}$$

Accuracy: relates to the classifier's ability in classifying the whole set. Higher the accuracy, the better the classifier is performing. The performance of the designed neural network classifier is measured in terms of accuracy.

$$Accuracy = \frac{(TP + TN)}{(TP + FN) + (TN + FP)}$$

4. Results and Discussion

In this paper, the proposed method is developed by MATLAB 2017b environment in a standalone personal computer using Intel i7 3770 processor @3.40 GHz with 6 GB RAM and 64-bit Windows 10 operating system. A total of 65 probes were tested: 32 convex abdominal arrays; 31 linear arrays; 2 transvaginal probes. Probes came from four manufacturers: 26 GE (GE Healthcare, Hatfield, UK); 3 Philips (Philips, Amsterdam, The Netherlands); 10

Siemens (Siemens Healthcare, Erlangen, Germany); and 26 Toshiba (Toshiba Medical Systems, Tokyo, Japan). The proposed method of transducer status detection was tested on 65 in-air reverberation images having no fault and fault ultrasound transducer and they were getting from [9]. The presented method consists of two main stages ultrasound transducer an in-air reverberation image separation and ultrasound transducer classification and performance of the classifier is evaluated in terms of Sensitivity, Specificity, Accuracy. In first stage, the software will require some parameters immediately after the ultrasound instrument has been installed and accepted. The parameters will be related to the instrument vendor (e.g. TOSHIBA, PHILIPS, SIEMENS and GE) to represent the ultrasound data of all transducers in data field that includes only in-air reverberation image as shown in figure 5. The parameters are x and y coordinates of in-air reverberation image that are top-left corner of a rectangular in-air reverberation image and the size of a rectangular data field image. The parameters are only defined one time immediately after the ultrasound instrument has been installed and accepted. In the second stage, the separated ultrasound transducer an in-air reverberation images are further subjected to 46 features. Four different texture sets (a total of 46 features) were extracted from the automatically an in-air reverberation image. Additionally, the conventional classifiers (i.e., SVM and BPNN) depend on the selection feature algorithms to select the most prominent features. The main purpose of using feature selection technique is to reduce the dimension of features to make the classifiers strong enough for distinguishing between the different functioning and defect transducers. Thirteen features were selected from twin support vector machine are contrast, Sum average, Sum entropy, Inverse difference normalized, Run Length

Nonuniformity, Invariant moment1, Invariant moment7, Mean of intensity histogram, Variance of intensity histogram, Entropy of intensity histogram, Skewness of intensity histogram, Kurtosis of intensity histogram and standard deviation of intensity histogram. Finally, classification results: the classification of fault / no fault transducer were done using SVM and BPNN techniques. The performance of the designed neural network classifier is measured in terms of accuracy. This term refers to the ability of the model to correctly predict the class of new unseen data. Classification accuracy is calculated by determining the percentage of cases in which the test sets are correctly classified. The performance of the neural network was calculated by analysis of confusion matrix. Results of the training data show that selected features with twin support vector machine yield an accuracy of 100%, sensitivity of 100%, specificity of 100%, false positive rate computed of 0%,

false negative rate computed of 0%, and misclassification rate of 0%. Size of the input dataset loaded in the network was of 65 samples. For example, 24 no fault ultrasound transducers are correctly classified as no fault ultrasound transducer. This corresponds to 36.9% of all 65 probes. Similarly, 41 fault ultrasound transducers are correctly classified as fault ultrasound transducer. This corresponds to 63.1% of all probes. Overall, 100% of the predictions are correct and 0% are wrong classifications as shown in figure 7. ROC graph shows the plotting of true positive rate against false positive rate (1-specificity). ROC graph of this network shows a perfect classification between the two categories, as the curves lie in the region of upper-left as shown in Figure. Results of classification performance of linear SVM is 65.63%. From the results, they are clear that the proposed method achieves highest sensitivity and smallest false positives.
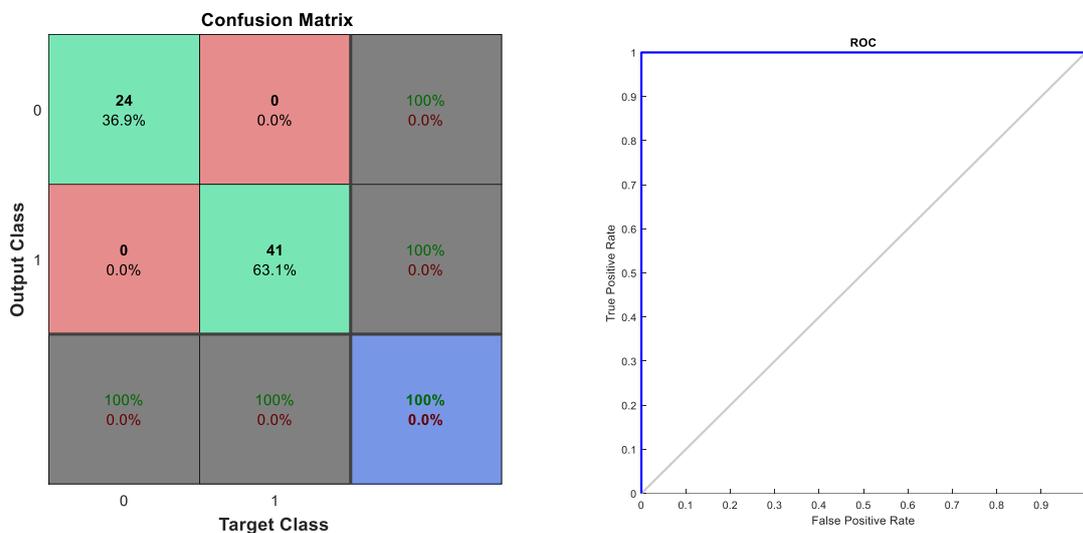
Figure 7. Confusion matrix and ROC plot of mixed features training data of ANN.

5. Conclusion

An automated method of ultrasound quality control that enables ultr asound status observing of (curvi-) linear transducers offers a first sign of consistent image degradation and can be easily implemented throughout the hospital and

performed sonographer. The effort to regularly monitor the quality of ultrasound no longer relies on additional trained personnel and the availability of a dedicated ultrasound QC phantom since the clinical user acquires the QC image. It is reduced to a simple acquisition of an in-air image that requires only the selection

and pre-setting of a preconfigured QC. The whole QC procedure can be implemented in less than 1 minutes, with the results readily available to the user. In addition, the objective of this study is to confirm th e feasibility of applying texture features to quality control characteristics of ultrasound transducer from an in-air reverberation ultrasound images, and to determine the optimal features. It is anticipated that in the near future, most ultrasound instruments will incorporate facilities for performing such image analysis. Our study results show 100% of accuracy, 100% of sensitivity, and 100% of specificity, by using 13 texture features extracted from in-air reverberation ultrasound images for the evaluation of ultrasound transducer. We proved that our method covered the following characteristics: 1) It is reproducible and 2) It speeds up the evaluation of ultrasound transducer.

References

[1] Hoskins P, Martin K, Thrush A. Diagnostic ultrasound: physics and equipment, CRC Press, 2019.

[2] Rownd JJ, Madsen EL, Zagzebski JA, et al. Ultrasound Probe Acceptance Testing Using the In-Air Reverberation Pattern. J. Ultrasound in Med Biol. 2019; 23:1–8.

[3] Hangiandreou NJ, Stekel SF, Tradup DJ, et al. Four-year experience with a clinical ultrasound quality control program. Ultrasound in Med Biol. 2011; 37:1350–57.

[4] Martensson M, Olsson M, Segall B, Fraser AG, Winter R, Brodin L. High incidence of defective ultrasound probes in use in routine clinical practice. Eur J Echocardiogr 2009; 10:389–394.

[5] Martensson M, Olsson M, Brodin L. Ultrasound probe function: Annual testing is not sufficient. Eur J Echocardiogr 2010; 11:801–805.

[6] Sipila O, Mannila V, Vartiainen E. Quality assurance in diagnostic ultrasound. Eur J Radiol 2011;80: 519–25.

[7] Russell S. Quality assurance of ultrasound imaging systems, Institute of Physics and Engineering in Medicine Report 102. York, UK: IPEM, 2010.

[8] H. Grazhdani, E. David, O. Ventura Spagnolo, F. Buemi, A. Perri, N. Orsogna, S. Gigli, R. Chimenz. Quality assurance of ultrasound systems: current status and review of literature. Journal of Ultrasound, 2018.

[9] Dudley NJ, Woolley DJ. A multicenter survey of the condition of ultrasound probes. Ultrasound 2016a; 24:190–197.

[10] K.K. Mohammed, H.M. Afify, F. Fouda, A.E. Hassanien, S. Bhattacharyya, S. Vaclav, Classification of human sperm head in microscopic images using twin support vector machine and neural network. Int. Conf. Innov. Comput. Commun. (2020).

[11] M. Vasantha, V. Bharathi, and R. Dhamodharan, "Medical image feature, extraction, selection and classification," International Journal of Engineering Science and Technology, vol. 2, no. 6, pp. 2071–2076, 2010.

[12] S. Selvarajah and S. Kodituwakku, "Analysis and comparison of texture features for content-based image retrieval," International Journal of Latest Trends in Computing, vol. 2, no. 1, 2011.

[13] M. Hu, "Visual pattern recognition by moment invariants," IRE Transactions on InformationTheory, vol. 8, pp. 179–187, 1962.

[14] J. R. Khemchandai and S. Chandra, "Twin support vector machine classification for pattern classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 5, 2007, pp. 905-910.

[15] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," Journal of Microbiological Methods, vol. 43, no. 1, pp. 3–31, 2000.

[16] C. Cortes and V. N. Vapnik, "Support vector machine", Machine Learning, vol. 20, no. 3, 1995, pp. 273-297.