Turkish Journal of Science & Technology
16(1), 65-84, 2021

Performance Analysis of Various Classification Algorithms for Computer-Aided Breast Cancer Diagnosis System Using Thermal Medical Images

Muhammet BAYKARA

Software Engineering, Technology Faculty, Firat University, Elazig, Turkey
mbaykara@firat.edu.tr

(Geliş/Received: 03/01/2021; Kabul/Accepted: 31/01/2021)

Abstract: One of the most widespread cancer types is breast cancer all over the world. It affects both women and men. Detection of cancer in early-stage is very critical in terms of treatment success. Many studies have been done in image processing, for the detection of cancer using computer-aided diagnosis systems. In this study, the performance of various classification algorithms in cancer detection was analyzed on a thermal image dataset. For this purpose, a graphical user interface based system was developed using MATLAB. The developed system uses five different algorithms; Decision Tree, Support Vector Machine (SVM), Logistic Regression Analysis, K Nearest Neighborhood (kNN), Linear Discriminant Analysis (LDA). According to the obtained results, kNN and SVM provide the best performance. The developed system can be used as an assistant system to produce an objective result for the expert in breast cancer diagnosis with the 98.8% success rate. This study contributes to the literature by using different methods compared to studies using the same dataset and increasing the performance of existing methods with hybrid approaches.

Key words: Image processing, breast cancer detection, thermal image analysis, image classification, feature extraction.

Termal Tıbbi Görüntüler Kullanılarak Bilgisayar Destekli Meme Kanseri Tanı Sistemi İçin Çeşitli Sınıflandırma Algoritmalarının Performans Analizi

Öz: Tüm dünyada en yaygın kanser türlerinden biri meme kanseridir. Hem kadınları hem de erkekleri etkiler. Kanserin erken dönemde tespiti tedavi başarısı açısından çok önemlidir. Görüntü işlemede kanserin bilgisayar destekli teşhis sistemleri kullanılarak saptanması için birçok çalışma yapılmıştır. Bu çalışmada, kanser tespitinde çeşitli sınıflandırma algoritmalarının performansı termal bir görüntü veri setinde analiz edilmiştir. Bu amaçla, MATLAB kullanılarak grafik kullanıcı ara yüzü tabanlı bir sistem geliştirilmiştir. Geliştirilen sistem beş farklı algoritma kullanır; Karar Ağacı, Destek Vektör Makinesi (DVM), Lojistik Regresyon Analizi, K En Yakın Komşu (kNN), Lineer Diskriminant Analizi. Elde edilen sonuçlara göre kNN ve DVM en iyi performansı sağlar. Geliştirilen sistem, %98,8 başarı oranı ile meme kanseri teşhisinde uzmana objektif bir sonuç üretmek için yardımcı sistem olarak kullanılabilir. Bu çalışma literatüre, aynı verisetini kullanan çalışmalara göre farklı yöntemler kullanması ve mevcut yöntemlerin hibrit yaklaşımlarla başarımlarının arttırılması yönüyle katkı sağlamaktadır.

Anahtar kelimeler: Görüntü işleme, meme kanseri tespiti, termal görüntü analizi, görüntü sınıflandırma, özellik çıkarma.

1. Introduction

Today, breast cancer is still one of the most widespread cancer types [1], especially for women. Early detection of breast cancer is very vital in diagnosis and treatment. Studies have shown that new cancer cases and death rates due to cancer have been decreasing. This decrease is thought to be due to improvements in treatment, increased awareness, and early detection by screening [2]. To provide early detection many different biomedical imaging methods are used in cancer diagnosis. These methods include ultrasonography [1], mammography [2-4], thermography, tomography, and etc. techniques.

In this study, a novel software tool was developed to detect breast cancer using thermal images. This paper also presents a performance evaluation of classification algorithms for a computer-aided diagnosis system on thermal mammography images. For this purpose Decision Tree, SVM, Logistic Regression Analysis, kNN, Linear Discriminant Analysis algorithms are compared according to their accuracy performance. The aim of this paper is to develop a computer-aided diagnostic system that can detect cancer in the thermal breast image. In the proposed system, Machine Learning (ML) algorithms are used in the detection of cancer images and high performance is obtained. The developed system can provide high accuracy, autonomous control, speed, and objective evaluation for experts.

* Corresponding author: mbaykara@firat.edu.tr. ORCID: 0000-0001-5223-1343[1]

Breast cancer is the most frequently encountered type of cancer in women today. Therefore, it is the leading cause of death due to cancer. Early diagnosis of it increases the effectiveness of treatment and increases the probability of complete healing. In the detection and diagnosis of it; mammography, ultrasonography, and magnetic resonance imaging are used [1-3, 5]. Cancer symptoms can be detected by mammogram images obtained using low-level X-ray radiation [6, 7]. For this reason, it is aimed to make use of methods such as image processing techniques, classification methods, and texture analysis and so on to diagnose this cancer early and go to the treatment part without wasting time. Many studies based on computer evaluation have been performed to help the expert physician in the interpretation of radiological images. This computer-assisted detection and diagnosis system (CAD) provides an increased success rate in the identification of diseases, such as breast cancer, for a specialist physician [8].

This paper is organized as follows. Section 2 provides a literature review on breast cancer detection with various techniques. Section 3 describes the proposed methodology and application. Evaluation and benchmarking of the results and relevant studies are given in Section 4. The conclusions are given in the last section.

2. Related Work

Recently, several studies have been done to detect breast cancer. These studies used different imaging techniques [1-4]. Huang et al. developed a web-based application [1] for the early detection and diagnosis of breast cancer. The main goal here is to interpret ultrasound images of the breast for enhancing the skills of the physicians on interpretation. Chougrad et al. [2] analyze the importance of transfer learning instead of random initialization for previously proposed CNN models and explore the effect of a well-adjusted number of layers on the results. They achieved 97.35% accuracy on the Digital Database for Screening Mammography (DDSM) database, 95.50% accuracy on the INbreast database, and 96.67% accuracy on the Breast Cancer Digital Repository (BCDR) database. Dhahbi et al. [3] investigated various gray-level texture analysis algorithms for the characterization of mammograms. This study emphasized that gray level texture features should be used together in order to reduce the false-positive level in cancer diagnosis systems.

Rezk et al. [4] developed a new data sampling method used to create samples on data distribution of binary patterns. Their method was applied in the classification of breast cancer images. The results show that their proposed method is efficient. Aswathy et al. [5] presented the current status and future possibilities for the detection of breast cancer. They specified that a new algorithm called visiopharm is used recently for a more objective diagnosis. Khalilabad et al. [6] focused on the detection of breast cancer type in their study. For this purpose, they developed a system to classify raw data related to cancer from microarray images and they achieved an accuracy of 95.23% in diagnosis.

It is difficult to evaluate mammographic X-ray images at first. An image defect occurs in images that are transferred from analog media to digital media. Various image filtering methods are used to enhance these images or reduce noise in these images [9-14]. Some of these filtering types are “imnoise”, “average”, “unsharp”, “gaussian”, and median filtering [15-17]. Filtering is basically obtained from the result of evaluating new values generated by changing pixel values. According to the new pixel values created in the images, the image can be made healthier by providing blurring, sharpening, increasing the brightness, understanding the color levels, and so on [18-21].

These images, which are brought to a healthy state, are then kept for processing by many classification methods. There are many types of classifications in the literature. One of these classification types is Artificial Neural Networks (ANN), which only work with digital information, have features such as information storage, learning using examples and generating information about unknown examples, classification, and shape completion. ANN can be applied to many fields from financial matters to engineering science, from production applications to fault detection and analysis in our daily life, and it has an important use in medical science. The predictive feature that ANNs are used extensively is used to estimate output from input values [23-27]. ANN uses the information provided to the network to estimate the output value corresponding to this information. ANN is used in the classification stage to remove tumors from mammographic images. Nowadays, many ANN models (Perceptron, Adaline, MLP (multilayer perceptron), LVQ (Learning vector quantization), Hopfield, Recurrent, SOM (Self-organizing map), ART (Adaptive resonance theory), and PCA (Principal component analysis)) have been developed for use with specific purposes and in various fields [28-32].

In addition, the Random Forest (RF) decision tree used in breast cancer detection distinguishes all nodes from each other by choosing the best of the randomly acquired attributes in each node, instead of dividing nodes selected from the best attributes in the data set. Each dataset is generated with displacement from the original dataset. Trees are developed using random property selection and there is no pruning. This is the reason why the RF is faster and more accurate than the other algorithms [3, 6, 12].

The most common methods used to classify medical images and diagnose the disease are kNN, SVM, and MLP [3, 4, 9, 15, 17]. Among other preferred methods, the C4.5 algorithm is a sort of decision tree-based classifier that takes the source from Quinlan's ID3 algorithm. The classifier constructs decision trees from a set of tagged training data using the concept of information gain. C4.5 is an algorithm generally used in medical data analysis. KNN is a classifier often used in pattern recognition. It uses similarity information to find neighbors [16]. In order to be able to say that a patient in breast cancer diagnosis belongs to a cancer patient, most of the k nearest neighbors must have cancerous specimens. SVM is not only for an early diagnosis of breast cancer but is also successful in many other classification problems [19, 23, 33-35].

Digitization and processing of pathologic data allow obtaining faster and more accurate results with computerized image analysis. With these techniques, diagnostic breast pathology assisted software has been continuously developed for years. In the literature, many techniques such as artificial intelligence, SVM, deep learning, ML, K-Nearest Neighbor algorithm (kNN), Naive Bayes, Linear Discriminant Analysis, and Logistic Regression have been used to achieve more accurate results [24, 33]. As can be seen from the literature studies, there are many systems used to provide datasets such as UCI [36-40]. In some studies, it has been focused on feature extraction methods and classification for malignant masses on mammograms image. For this purpose, methods such as the GLCM (gray level co-occurrence matrix), contourlet, discrete wavelet, ridgelet, and curvelet transform have been used additionally to help classification [8, 10, 11, 14].

The classification has an important place in ML and data mining. A decision tree is also one of the most popular learning models in data mining. Actually, the effectiveness of each algorithm used in classification depends on various configurations such as input property types and model parameters. Methods used to overcome model performance limitation, using a classification-based learning algorithm such as SVM to reduce diagnostic variance and improve diagnostic accuracy have been extensively used in breast cancer diagnosis [40-43].

Thermography is an imaging technique thet can easily detect cancerous masses more quickly than conventional mammography. Progress in the IR (infrared) cameras used to obtain thermal images of the breasts and in the calculation tools used to accurately model the heat transfer within the breast has significantly increased the accuracy of the thermography [36]. Current studies have also explored the progress of using thermal and benefited from this progress. Ultrasonography techniques such as ultrasonography elastography, contrast-enhanced ultrasound, 3-D ultrasound, automated breast ultrasound and chest ultrasound have also been used in the studies. In proposed knowledge-based systems, many clinical decision support systems have been developed to assist health practitioners. In addition, ultrasound-guided breast biopsy and other imaging modalities, especially MRI (Magnetic resonance imaging) and ultrasound fusion have been used to detect chest diseases [24, 25, 26, 27, 34].

Although the most frequently used method for detecting breast cancer is the investigation of mammography, thermography performs better in the analysis of dense tissues. However, in the last two decades, many computer-assisted diagnostic systems have been proposed for the early detection of cancer. Mammogram-based classification, which consists of many stages such as feature extraction, classification, and segmentation, is an important and effective way to diagnose breast cancer in computer-aided diagnosis. Ultrasound used in CAD systems is one of the most commonly used methods to detect and diagnose breast tumors due to its inoffensiveness and low cost [37-42].

These studies suggest that ML approaches to diagnosing breast cancer are a powerful alternative to clinical methods and new studies are needed for better results [44-48]. In this context, Logistic Regression Analysis and LDA methods, which are not used in the literature, have also been used in this study. In the developed study, the most successful result with a 98.8% success rate was found with kNN.

3. Proposed Methodology and Application

Mammography is the most important imaging method that can be used for the early detection of cancer. With the use of images obtained by mammography, even small changes that cannot be detected by palpation can be detected early. Therefore, mammography images should be analyzed in detail for the diagnosis of the disease. In the detection phase, breast cancer is often misdiagnosed by radiologists. Because mammograms are two-dimensional projections of a three-dimensional object. So superimposed textures may also produce some symptoms, as well as hide symptoms that do not exist. Hence, radiologists cannot detect real patients, and every year a certain number of patients suffer from it. Many biomedical image processing applications have been developed in order to simplify the work of the radiologists and to short the examination period.

In this study, a computer-aided diagnosis system was developed using thermal mammography images. Database for Mastology Research (DMR) database was used in the developed system. There are 3340 images of 287 people in the dataset images (Fig. 1), each with .txt extension at 640x480 pixels, taken at different numbers and angles. In the first stages of the study, various operations were performed on these images. These operations are preparation, pre-processing and segmentation, feature extraction, and classification. The aim of this section is to provide detailed information about these steps.

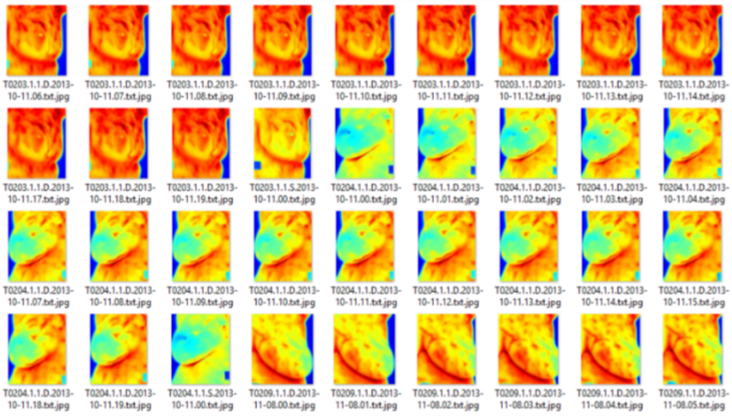

Figure 1. Image database.

During the preparation phase (step 1), filtering processes were performed for 3340 image files with the .txt extension in the dataset. First, the data in the form of a text file with a float value of 640x480 size has been imported into the system. This imported data is then transformed into a picture. However, when we convert the txt file to Image, it will be a grayscale image because it is a one-dimensional array. Since the operations will not perform on a grayscale image, the txt file will be read first and the highest value will be converted to red and the lowest value will be converted to blue to build an image file. Thus, thermal images will be obtained. After that, these thermal images cut out from the neck region and from the lower part in order to focus on the chest area. The obtained image is saved to use after the last phase of step 2, called "Horizontal projection analysis"

At the pre-processing and segmentation phase (step 2), it is aimed to obtain the image quality necessary for the system to be more accurate by eliminating the difficulties in diagnosis (noise, low contrast, etc.) on the digitized image. In addition, features not used on the image are removed. Cleaning the background noise will be useful for improving the image on the mammograms while protecting the details of the suspicious areas that can be detected as a tumor. For pre-processing, the image that was first converted to a thermal image was considered, and filtration was carried out. With filtering, each pixel value is recalculated as though there is a filter on the image. Filtering has done the processes such as sharpening the image, extracting certain details, smoothing the image, edge sharpening or edge detection. The obtained original image was converted to a gray tone to clearly select the chest area required for the detection step, and then noise removal was performed by centering the values of the pixels with this median filter method on the gray tone image. After the noise removal process, a clean image is obtained. It is then necessary to fill in the gaps in the image. For this, before the edge detection, the image is converted to a gray level. Later, the gray level image has been filled with white areas under the 50-pixel threshold. With the threshold process, objects in the image are separated from the image background. For thresholding, the image histogram that shows the gray level distributions in the image is utilized. After this method, edge detection is performed in the gray level image and the protrusions on the image are detected and the chest area is clearly determined by Sobel Edge Detection. Then, a graph was created using the Horizontal Projection Analysis method and the pixel values at the peak points were determined, and then the chest region was clearly extracted by "imcrop" method by matching on the original thermal image that was mentioned in the last step of Step 1. After this process, the resulting image is divided into two parts using the "imcrop" method to be used in the feature extraction phase. A general working diagram is given in Fig. 2.

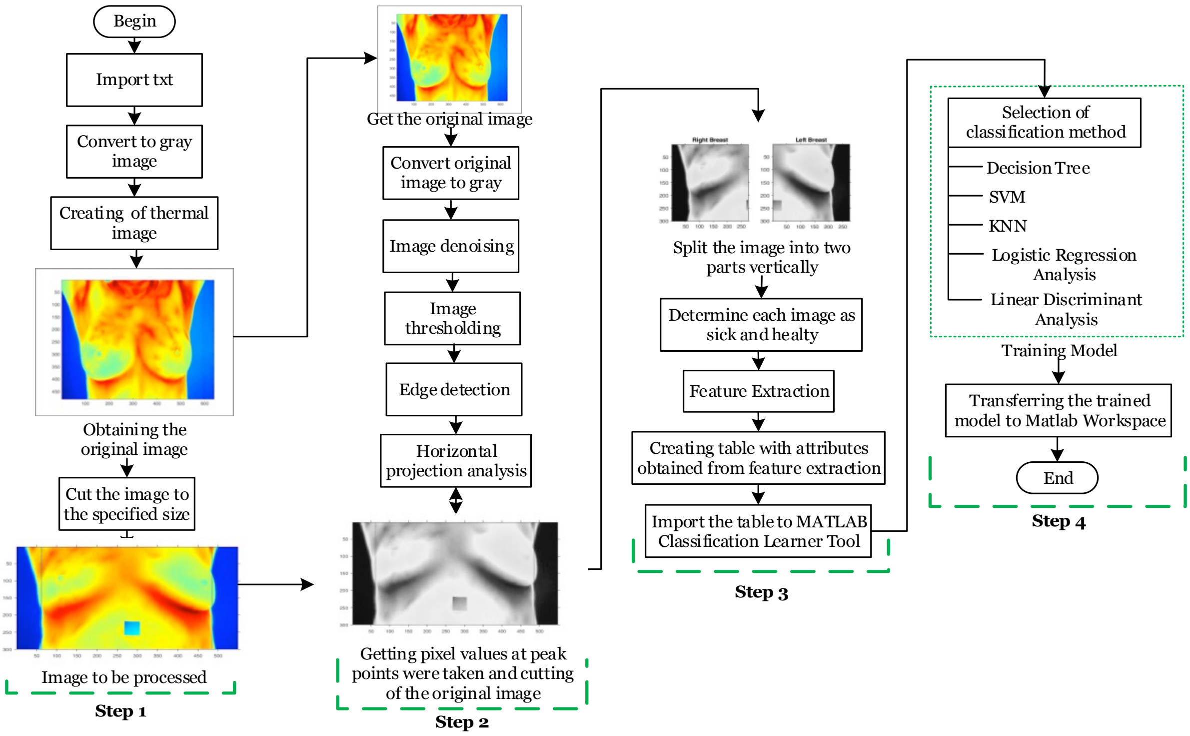

Figure 2. General working diagram of the developed system.

In the feature extraction section (step 3), feature extraction is performed on images divided into two parts. In the first phase of this section; Firstly, the images were compared with excel manually, looking at the ID numbers before dividing the images into two parts. According to the ID number of the patient whose breast was biopsied, the breast was determined to be sick. In other words, the breast that was biopsied determined as sick, and the others were determined as healthy by examining the attributes on excel. Thus, by dividing a patient's image into two, the total number of data, which has doubled in number, has become 3895. Here, in comparison with the excel table, 3098 healthy breasts and 797 diseased breasts were identified. After these operations, various textural attributes of each breast were obtained such as entropy, variance, standard deviation, contrast, homogeneity, kurtosis, skewness, correlation, energy. The obtained values were recorded in a table. The table was imported into the MATLAB Classification Learner Tool. The values calculated in the feature extraction stage are briefly explained in this section.

Entropy is the measure of image complexity. Complex tissues have higher entropy. Entropy is a statistical computation of randomness. It can be used to specify the texture of an image. Entropy is defined as -sum(p.*log2(p)), where p defines the normalized histogram counts returned from the function of imhist.

Variance is the sum of the squares of the deviations of the data from the arithmetic mean. In other words, the variance is defined as the square of the standard deviation. The formula bellowed can be given as a definition of variance for a feature vector called A made up of N scalar observations.

$$V = \frac{1}{N-1}\sum_{i=1}^{N} |A_i - \mu|^2 \quad (1)$$

In this equation, μ is the mean of A.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i. \quad (2)$$

Kurtosis is a measure of the probability distribution of a real-valued random variable and an indication of the amount of change in the class. Kurtosis includes different ways of quantification for a theoretical distribution and appropriate estimation methods from a population sample.

$$k = \frac{E(x-\mu)^4}{\sigma^4} \quad (3)$$

Skewness is defined as the measure of the asymmetry of the data around the sample mean. When the skewness is negative, the spread of the data is greater towards the left of the center relative to the right. When the skewness is positive, spread to the right is greater. The skewness for a normal distribution or any perfectly symmetric distribution is zero. We can define the disruption of distribution as follows:

$$s = \frac{E(x-\mu)^3}{\sigma^3} \quad (4)$$

Mean returns the average values of items in an array along their different dimensions. In the form, it is shown that for a random variable vector of N scalar observations, the average is found.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i. \quad (5)$$

Standard deviation can be defined as the square root of the average of the differences of each number on a group of numbers with the mean. To calculate the standard deviation;

• The arithmetic mean of the numbers is calculated.
• For each number, the difference from the arithmetic means is calculated.
• The square of each difference is calculated.
• The squares of the differences are summed.
• The sum obtained is divided by the previous number of the total count of the elements in the series
• The square root of the found number is taken.

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} |A_i - \mu|^2} \quad \text{where } \mu \text{ is the mean of A} \quad (6)$$

RMS (root mean square) is a statistical value used to measure the magnitude of varying amounts. It is useful in waves where the change is positive and negative. A changing function can be calculated for the continuous value series. The square root comes from taking the square root of the average of the mean squares. The average square root level of an x vector:

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |x_n|^2} \quad (7)$$

It is expressed as the sum performed over the specified dimension.

Contrast is a calculation of image variation and image contrast.

$$\sum_{i,j} |i-j|^2 \, p(i,j) \quad (8)$$

Correlation: Correlation is the calculation of image linearity. Linear structures in the Φ direction lead to large correlation values in this direction.

$$\frac{\sum_i \sum_j (i-\mu_i)(j-\mu_j)N_d[i,j]}{\sigma_i \sigma_j} \tag{9}$$

Energy is the measure of image homogeneity. For the more homogeneous image, energy has a larger value.

$$\sum_{i,j} p(i,j)^2 \tag{10}$$

Homogeneity is a measure of similarity in different regions of the image. As the object size gets smaller, the homogeneity in the object will increase accordingly. The heterogeneous structure of different spectral properties and classes in these objects during the selection of large objects influences the classification accuracy.

$$\sum_i \sum_j \frac{N_d[i,j]}{1+|i-j|} \tag{11}$$

In the classification section (step 4) ML methods are used. ML is a kind of method that aims to give computers the ability to learn. ML provides learning skills such as prediction, diagnosis, detection and recognition using statistics, optimization, programming and many more disciplines. ML is basically divided into two groups such as supervised and unsupervised learning. Firstly data is given to the system in supervised learning, then data that the system has never seen is expected to be known over the previous data, similar to the artificial neural network. In unsupervised learning, data is clustered with the help of the distribution of the data features, similar to k-means clustering. Artificial neural networks, Bayesian decision theory, clustering, statistical discriminant analysis, multi-layered networks, and hidden Markov models have widely used ML techniques, and each technique has its advantages and disadvantages in terms of performance.

As mentioned in the previous section, a total of 11 different features were extracted from each image. There are 3895 images in the image database. The size of the dataset is totally 3895x12 because one extra column is for the result (Fig. 3). The cross-validation method has been preferred to train and test the system.

With cross-validation, the data set is divided into a number of parts, one of which is used for testing, and the other parts are used for training. In this study, k value was chosen as 5 for cross validation. In the next stage, another part is used for testing while the other parts are used for training. As can be seen in Figure 4 this process continues for all parts.

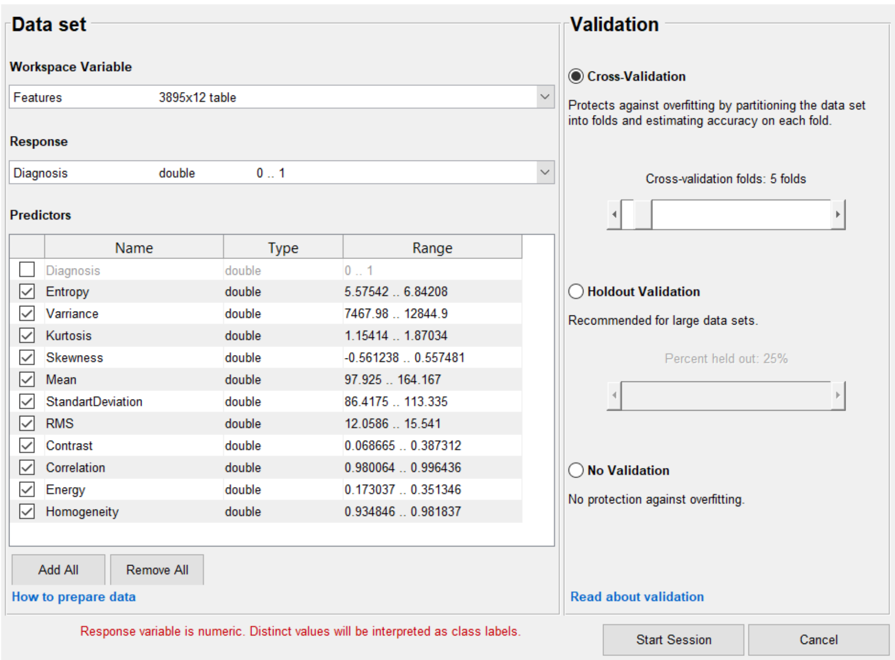

Figure 3. Feature extraction from the dataset.

Diagnosis	Entropy	Variance	Kurtosis	Skewness	Mean	StandartD	RMS	Contrast	Correlatio	Energy	Homogeneity
1	6.382179	9864.648	1.606388	-0.42113	152.6657	99.32071	14.32094	0.143684	0.991052	0.221453	0.969636816
1	6.294064	10354.97	1.508547	-0.3921	150.8107	101.7591	14.2102	0.104337	0.993821	0.22368	0.973263389
1	6.368711	9875.018	1.601532	-0.41447	152.1913	99.3729	14.29881	0.139078	0.991331	0.218729	0.970520497
1	6.273599	10343.87	1.508562	-0.3888	150.6141	101.7046	14.20311	0.103589	0.993855	0.222208	0.973894538
1	6.374852	9907.613	1.586333	-0.40043	151.4516	99.53676	14.2854	0.136741	0.991516	0.217708	0.970858363
1	6.279064	10413.14	1.492118	-0.38101	150.2886	102.0445	14.20555	0.103589	0.9939	0.223405	0.973470132
1	6.361803	9962.298	1.575009	-0.39503	151.0938	99.81108	14.26819	0.13698	0.991539	0.217335	0.970815717
1	6.27979	10459.7	1.482115	-0.37376	149.8565	102.2725	14.18872	0.102717	0.993978	0.223671	0.973913576
1	6.351475	10010.05	1.565464	-0.39162	150.9	100.05	14.25539	0.136549	0.991609	0.217478	0.970576034
1	6.277286	10462.69	1.479804	-0.36981	149.6448	102.287	14.18216	0.103839	0.993923	0.223418	0.973634181
1	5.868081	12422.43	1.218631	-0.19014	136.5699	111.4557	13.14352	0.090005	0.995528	0.292125	0.976404999
1	5.851778	12057.31	1.24339	-0.08595	130.4701	109.8055	12.97495	0.083451	0.995685	0.270957	0.975639049
1	6.110144	11314.2	1.368582	-0.28159	140.7132	106.368	13.41375	0.120606	0.993391	0.245043	0.963403044
1	6.228667	10802.83	1.41381	-0.23446	140.3816	103.9364	13.70881	0.123852	0.992893	0.230497	0.960153091
1	6.056771	11356.39	1.355762	-0.25235	139.1217	106.5661	13.35642	0.116476	0.99362	0.245642	0.964977037
1	6.235295	10893.77	1.393537	-0.21308	139.4956	104.373	13.68237	0.115629	0.993415	0.232331	0.962122779
1	6.079482	11353.37	1.341554	-0.21388	137.2334	106.5519	13.32566	0.118209	0.993528	0.24369	0.965118772
1	6.204321	11005.77	1.366239	-0.16311	136.9302	104.9082	13.6167	0.121007	0.993179	0.233309	0.962171623
1	6.075595	11344.29	1.341034	-0.20553	136.8762	106.5093	13.32264	0.11754	0.993552	0.242865	0.96575132
1	6.219178	11039.37	1.35722	-0.14943	136.2226	105.0682	13.605	0.121663	0.993169	0.233526	0.962218907
1	6.09722	11334.78	1.331888	-0.17159	135.2493	106.4647	13.31197	0.118848	0.993482	0.241913	0.964717418
1	6.217624	11149.85	1.334054	-0.11128	134.3651	105.5926	13.56511	0.118137	0.993444	0.234807	0.962471386
1	6.078416	11339.16	1.330881	-0.16457	134.8649	106.4853	13.29332	0.11702	0.993575	0.241284	0.96555677
1	6.207025	11166.33	1.333252	-0.10741	134.0603	105.6706	13.54251	0.1195	0.993374	0.235983	0.963168223
1	6.094008	11331.42	1.326978	-0.14622	134.0085	106.4489	13.28987	0.119184	0.993457	0.24006	0.964445255
1	6.201555	11214.58	1.324655	-0.09479	133.4761	105.8987	13.53539	0.118556	0.993456	0.236914	0.962901423
1	6.085801	11341.99	1.327341	-0.14967	134.1805	106.4985	13.29307	0.116594	0.993606	0.241707	0.965302876
1	6.208531	11213.94	1.3243	-0.09649	133.5159	105.8956	13.53364	0.119507	0.993402	0.236406	0.962853717

Figure 4. Feature values of the images from GLCM method.

3.1 Decision Tree

Decision trees are a commonly used data mining approach to classification and estimation. Although other methodologies, such as neural networks, can be used for classification, decision trees provide an advantage for decision-makers in terms of ease of interpretation and intelligibility. The decision tree technique is a two-step process using learning and classification. In the developed project, training data which is known in the learning step is analyzed by the classification algorithm in order to construct the model. The learned model is shown as a classification rule or decision tree. In the classification step, the test data is used to determine the accuracy of the classification rules or decision tree. If accuracy is acceptable, rules are used to classify new data. It should be determined which areas in the training data will be used in which order to build the tree. The most commonly used measurement for this purpose is entropy. The result obtained using the area where the entropy measure higher, is uncertain and unstable. For this reason, at the root of the decision tree, the fields with the least Entropy measure are used [3], [6].

3.2 Support Vector Machine

SVM is a ML algorithm based on convex optimization working by the principle of structural risk minimization. The SVM algorithm is a distribution independent learning algorithm because it does not need any joint distribution function knowledge related to the data. In this algorithm, each data item is drawn as a point in n-dimensional space (where n is the number of properties we have), along with the value of each feature whose value is a specific coordinate value. The classification is then performed by finding the hyperplane, which distinguishes the two classes very well. The SVM developed for binary classification essentially separates the data with a plane that can be expressed by a linear equation such as wTx+b=0. Here, w defines the d-dimensional coefficient vector, x defines the data, and b defines an offset value. Linear SVM provides finding the separation plane with optimizing the objective function using quadratic optimization [17, 19].

SVM is a structured learning procedure in statistical learning theory. Instead of minimizing the squared error of the data sets, SVM minimizes the limit on the generalization error. Therefore, it gives successful results also for the data apart from the training set. There are three main stages in SVM classification: (1) defining the training cells as feature vectors, (2) mapping feature vectors to feature space using kernel functions, and (3) creating n-dimensional hyperplanes that best separate classes. SVMs have been proposed for solving classification and curve-fitting problems based on statistical learning theory and minimizing the structural risk. This learning method is considered as supervised learning method. In supervised learning, it is known that the data belong to which classes. The SVM separates the data provided as input into two classes [23].

3.3 Logistic Regression Analysis

It is a regression method in which the expected values of the response variable are obtained as probabilities according to the explanatory variables. Simple and multiple regression analyses are used to analyze the mathematical relationship between the dependent variable and the explanatory variable or variables. Logistic regression analysis is a regression method that helps to perform classification and assignment. There is no assumption of normal distribution and continuity. The effects of the explanatory variables on the dependent variable are obtained as probabilities and the risk factors are determined as probabilities. Discriminant analysis is a method of classifying data and assigning it to specific classes according to certain possibilities. It is possible to determine the effects of the variables in the data set to the classification by the logistic regression [45, 46].

3.4 K Nearest Neighborhood (kNN)

KNN algorithm is in the supervised learning category between ML algorithms. KNN algorithm can be used for classification. In this algorithm, the output is a class membership. An element is classified by a majority vote of its neighbors, with the element being assigned to the class most common among its k nearest neighbors. It can also be used for the prediction of continuous values. This value is the average of the values of its k nearest neighbors.

The kNN algorithm can be summarized as [17]:

1. k value is specifies as a positive integer value

2. The k entries are selected among our data which is closest to the new sample (Euclidean distance)

3. The most common classification of these entries is found. The distances are sorted and k objects are selected at the nearest distance.

4. This is the classification that is given to the new sample

3.5 Linear Discriminant Analysis

LDA is a classification method developed by R. A. Fischer in 1936. Although it is a simple method, it provides good results in complex problems. Discriminant analysis is a common classification algorithm, and it is fast, accurate and easy to interpret. It is also good for large data sets. The discriminant analysis assumes that different classes form data based on different Gaussian distributions. To train a classifier, it estimates the parameters of a Gaussian distribution for each class. There are two discriminant types such as linear and quadratic discriminant. The LDA defines the following score function [47, 48].

$$Z = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_d x_d \tag{12}$$

$$S(\beta) = \frac{\beta^T \mu_1 - \beta^T \mu_2}{\beta^T C \beta} \quad \text{(Score function)} \tag{13}$$

$$S(\beta) = \frac{\overline{Z_1} - \overline{Z_2}}{The\ variance\ of\ Z} \tag{14}$$

Score function, predicts the linear coefficients that maximize the problem score. Formulation: where β is linear model coefficients; C1, C2 are covariance matrices; μ_1, μ_2 are average vectors

$$\beta = C^{-1}(\mu_1 - \mu_2) \quad \text{Model coefficients} \tag{15}$$

$$C = \frac{1}{n_1 + n_2}(n_1 C_1 + n_2 C_2) \text{ Covariance matrices} \tag{16}$$

The way to determine the best discrimination is to calculate the Mahalanobis distance between the two groups. The fact that Mahalanobis distance is less than three indicates that the probability of misclassification is very small.

$$\Delta^2 = \beta^T(\mu_1 - \mu_2), \tag{17}$$

$$\Delta: Mahalanobis\ difference\ between\ two\ groups$$

Finally, if the following condition is satisfied, a new incoming property is classified.

$$\beta^T\left(x - \left(\frac{\mu_1+\mu_2}{2}\right)\right) > log\frac{p(c_1)}{p(c_2)}, \tag{18}$$

Here, β is a coefficient vector, x is a data vector, μ_1, μ_2 are average vectors, p is class probabilities.

3.6 Using Gabor Filter for Feature Extraction

The Gabor filter method, developed by Dennis Gabor, describes the signals of frequency and time (or space) with minimum uncertainty. Gabor filters are band-pass filters that can select frequency and direction. Gabor filter is the product of Gaussian function and sinusoidal plane wave. The Gabor function is calculated by multiplying the Gaussian function with a complex exponential function. Gabor filters have different uses for an application in computer vision and image processing. Examples include tissue recognition and classification, type identification, texture separation, edge detection, image compression, motion estimation, object identification and shape recognition in the tissue. Visual cortex using the direction-selective feature of the Gabor filter performed increased the use of these methods in computer vision and digital image processing. The Gabor filter also allows access to local frequency information from an image. In contrast to the Fourier analysis, which determines the representation of a global frequency area of the whole image, the Gabor filter gives a result in the spatial field, calculating the power of certain frequency belts and the predictions in each position on the image. The number of Gabor filters varies for different applications. 5-scale and 10-oriented 39x39 filters are used for filtering. These filters, images to be feature extracted, and the number of steps are sent to the Gabor function for the feature extraction operation. The length of the property vector is calculated as (m*n*u*v) / (d1*d2). "m" and "n" represents the x and y dimension of the image. "u" represents the number of scales and v represents the number of orientations in the filters. "d1" and "d2" represent the sub-sampling factor along the row and column.

In the previous part, breast images were divided into two groups as right and left. The sizes of these images were not very important for feature extraction. However, when sub-sampling with the Gabor filter, the number of properties (or feature vector size) extracted from each image is variable due to the size of these images. For this reason, while x features were extracted from an image, x + y (y! = 0) or x-y (y! = 0) features were extracted from another image. When these properties were extracted and collected in a data set, the resultant part of the data set in a given column or row was different in the feature vector.

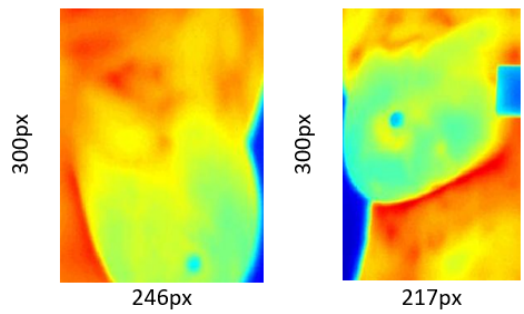

Figure 5. Preparing images for Gabor filter.

3000 features were taken from constant Gabor variables determined from Figure 5, Image 1 above, while 2700 different properties were extracted from the 2nd image using the same variables. The dataset formed by feature vectors has an irregular structure as all images may vary in size. To solve this problem, the average height and width values of all images were taken. These values were observed to be 300x220. The image size is resized to 300x220 before the feature is extracted from an image. Thus, data should be recovered from the irregular form as given in Table 1. So it is made regular as in Table 2.

Table 1. Non-Uniform dataset

X1	X2	X3	X4	…	y
X1	X2	…	y		
X1	X2	X3	…	y	

Table 2. Uniform dataset

X1	X2	X3	X4	…	y
X1	X2	X3	X4	…	y
X1	X2	X3	X4	…	y

Gabor filter is created after all the images are fixed to the same size. This filter takes 4 variables. These variables are (m, n, u, v).

u: Number of scales (generally set to 5)

v: Number of directions (generally set to 8)

m: Number of lines in the 2-D Gabor filter (generally set to 39 and preferred to be an odd number)

n: Number of columns in a 2-D Gabor filter (generally set to 39 and preferred to be an odd number)

In this section, the parameters m, n, u, v are set as 5, 10, 39, 39 respectively. Figure 6 shows Gabor Filters with five scale and ten orientations.

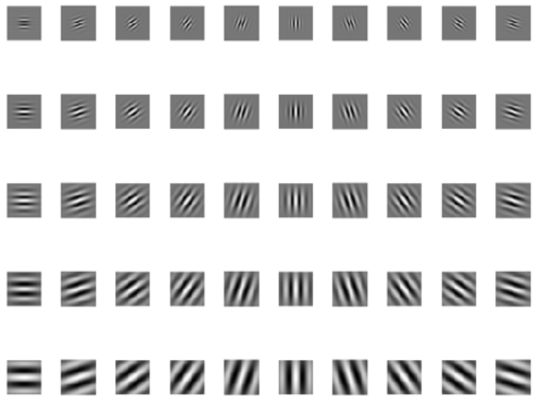

Figure 6. Gabor filters

These filters are sent to the Gabor feature extraction function. This function takes 4 parameters as input and returns the feature vector. These parameters which taken as input are named as I, Gabor Filter, d1, and d2 respectively. Here "I" represents the image, "gaborFilter" represents the filters created in the previous function, “d1” and “d2” represent the sub-sampling factor to be applied across the rows and columns of the image. In addition, this function accepts a grayscale image as input. Filters are applied to the image in the matrix in which the filters are kept.

17	24	1	8	15
23	5	7	14	16
4	6	13	20	22
10	12	19	21	3
11	18	25	2	9

+

-1	0	1

=

24	-16	-16	-14	-8
5	-16	9	9	-14
6	9	14	9	-20
12	9	9	-16	-21
18	14	-16	-16	-2

Figure 7. Gabor filter example.

As it can be seen in Figure 7, the last two cells of the filter and the values in the first two pixels of the image are multiplied first and then collected and written to the first pixel of the filtered image. So (17 * 0) + (24 * 1) = 24 is in the form. Then the filter is shifted one step to the right and the same operation is repeated. It continues as like (17 * (-1)) + (24 * 0) + (1 * 1) = -16. In this section, all the filters created in the mammography image are applied in order (Figure 8).

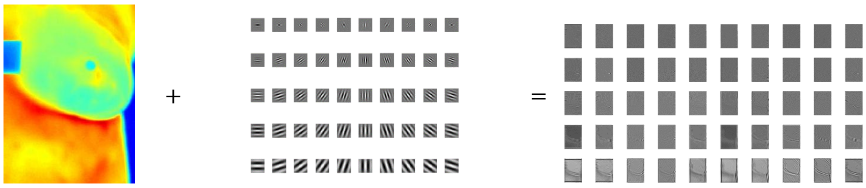

Figure 8. Gabor filter application to the mammography image.

The following operations are applied to the filtered images. *(Step 1-5)*

1. *The absolute value of the image is taken. This eliminates negative values in the pixel values.*
2. *Sub-samples of all rows are taken according to the d1 variable specified in the image.*
3. *Transpose of sub-samples are made, and sub-samples are taken according to d2 variable.*
4. *The obtained two-dimensional sub-sample space is made one-dimensional and added to the feature vector.*
5. *Move to the next filtered image and the same steps are performed.*

The above application of operations is as follows. Consider the image in row 5 and column 1 of the filtered images.

3.33128395992207 + 3.98315128827445i	3.54133485101283 + 4.23872581298265i	3.75074056426018 + 4.49636486896177i	…
5.10909916923280 + 3.24485568932152’i	5.75652298999523 5.43282788617142 +	3.45382524842210i + 3.66462216297547i	…
6.45916945830799 + 1.89067945018052i	6.86953182792018 + 2.01382989990496i	7.28034023028880 + 2.13835573922378i	…
7.17776964942566 + 0.137863599153909i	7.63463771635483 + 0.149405016347111i	8.09221333491624 + 0.161700819108754i	…
…	…	…	…

Figure 9. The pixel values of the filtered image.

As shown in Figure 9, the pixel values of the filtered image consist of complex numbers. The absolute value of the image is taken to eliminate this complexity. The absolute value for complex values is defined as

$$|a+bi| = \sqrt{a^2 + b^2} \tag{19}$$

5.19258577270674	5.52339108199333	5.85536948579851	…
6.05243610173437	6.43774243717468	6.82400267670776	…
6.73019605022571	7.15862964543911	7.58787975235724	…
7.17909350212043	7.63609945711143	8.09382874866381	…
…	…	…	…

Figure 10. Absolute value for the image.

The sub-samples are taken at the d1 rate of each column of the absolute value as given in Figure 10. “d1” and “d2” were taken as 8 in this section. The sub-samples will be taken as 1., 9., 17., 25...., and all values will be taken by increasing in d1 ratio. In this way, only the rows with multiples of 8 will be kept. As a result of this process, a picture with a size of 300x220 is reduced to a size of 38x300. Transposition of the 38x300-sized image created in “Step 2” is taken. The result is 300x38-sized. The same steps as in step 2 are performed again. But the ratio here is “d2” and may differ from “d1”. D1 is the same for this project. As a result of the process, a 28x38 image was created. This image matrix will then be called the future. The values in the 2-dimensional (28x38) matrix created in step 3 are converted to a one-dimensional matrix. As a result of this process, a 1064x1 matrix is formed. This matrix is the feature vector obtained from the filtered image. The next filtered image is selected and all steps are repeated. The feature vector of that filtered image obtained in “Step 4” is added to the feature vector of the actual image. These operations continue and the feature vector of the image (actual image) to which the filters are applied is obtained. In total (1064+1064+…) x1 = 53200 features are obtained. The above steps continue with the next image. After all the images are completed, a feature vector consists of 3895 x 53200. After this step, the column 53201 is added to decide whether the breast is diseased or not. (0 for patient breasts, 1 for healthy breasts). This data set is divided into a 5-fold cross-validation method and training and test data are separated by the Fine kNN algorithm. As a result of the training, 98.7% (Figure 11) success rate was achieved.

Figure 11. The accuracy result of Gabor-Fine kNN.

In order to increase the success rate, a second training was carried out on the same data set. In this training, 69% of the data using PCA were kept in those with a variance of 69%. In the above study, the number of the nearest neighbors in the kNN algorithm was 1, while this rate was 3 in the second training. At the end of the second training, the success rate increased from 98.7% to 98.8% (Figure12). The details of the second training are given below.

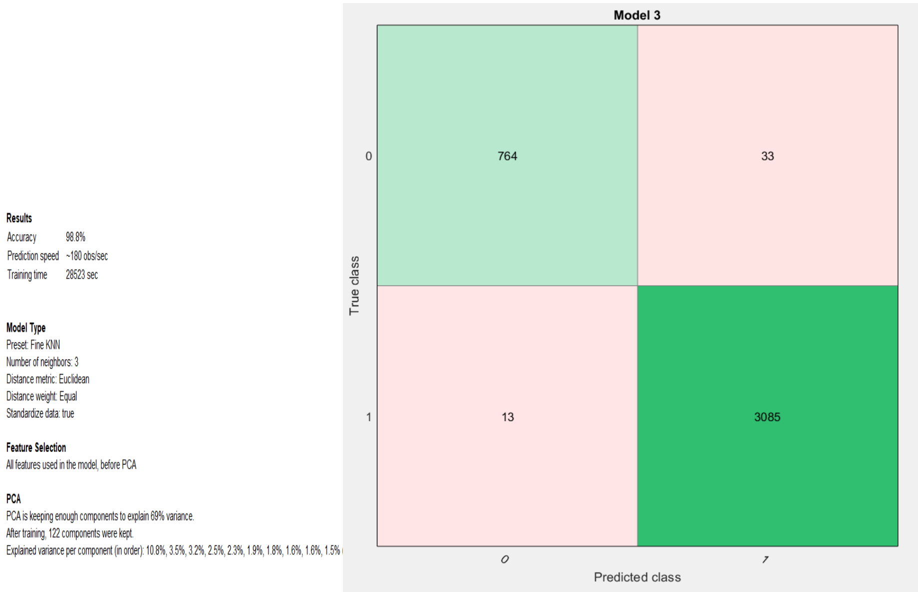

Figure 12. The accuracy result of Gabor-Fine kNN with PCA.

Figure 13 shows the implementation of the 2D-Gabor filter and kNN. Figure 14 shows the MATLAB GUI for the developed software.

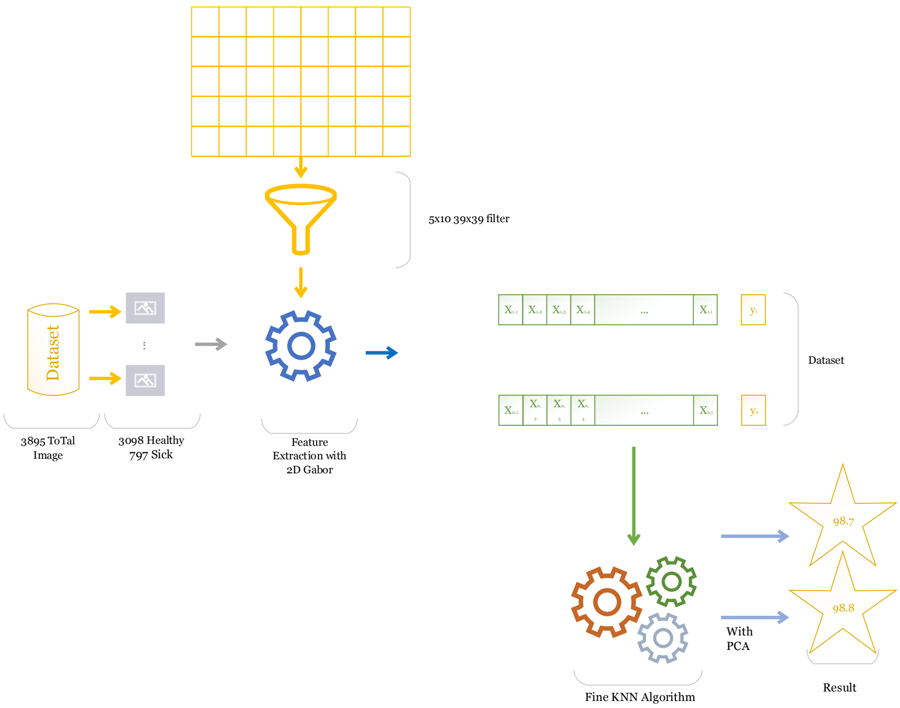

Figure 13. Implementation of 2D-Gabor filter and kNN.

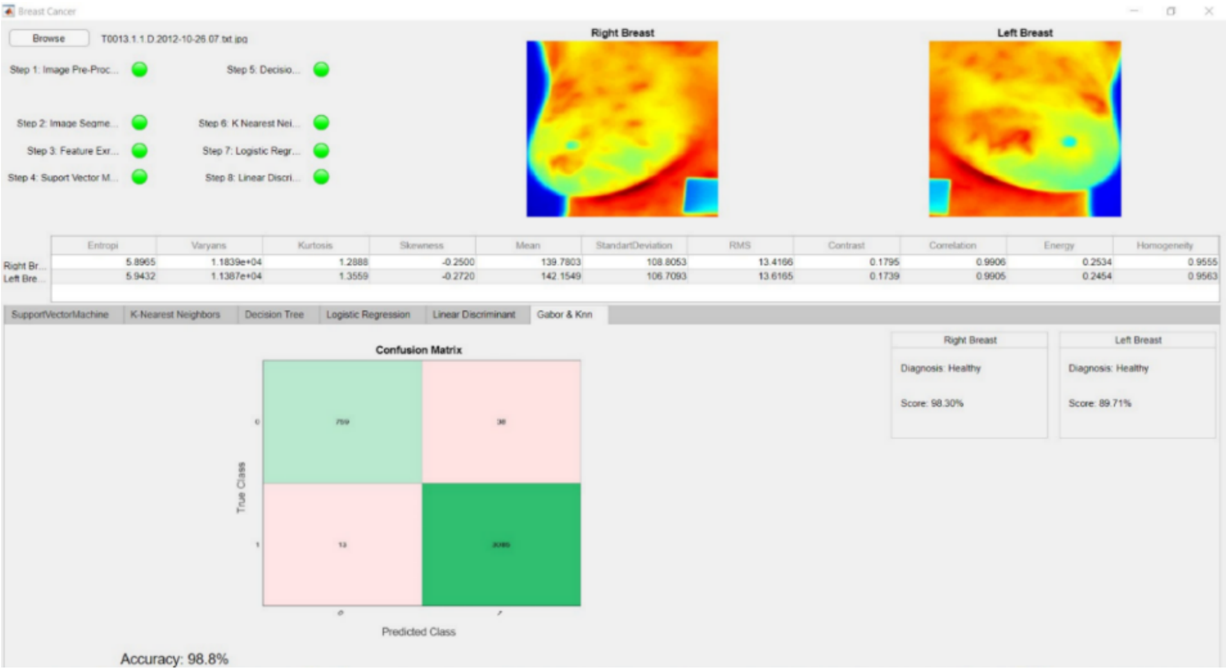

Figure 14. MATLAB GUI for the developed software

4. Evaluation and Benchmarking

In the classification of medical images and disease diagnosis; kNN, SVM and MLP are widely used. Of the preferred methods, the C4.5 algorithm is a sort of decision-tree-based classifier originating from the Quinland ID3 algorithm. The classifier constructs decision trees from a set of labeled training data using the concept of knowledge gain. C4.5 is an algorithm commonly used in medical data analysis. kNN is a classifier based on majority voting, which is commonly used in pattern recognition. KNN, which is classified according to the class of neighbors, uses similarities in finding neighbors. In order to be able to tell that a patient is a cancer patient in the diagnosis of breast cancer, most of the nearest neighbors should be cancerous. When the studies were examined, it was observed that the SVM method was not only useful for the result in early diagnosis of breast cancer but also many other classification problems. In addition, studies in the literature have shown that ML approaches in breast cancer diagnosis provide a strong alternative to clinical methods and new studies are needed for better results. Table 3 gives a detailed comparison between the studies have done before and the current study present in this paper. In this study, methods such as Logistic Regression Analysis and LDA which were not used before in early diagnosis of breast cancer were used.

In this study, both GLCM and Gabor were investigated and performed in the feature extraction section. Similarly, many methods have been tried in the classification phase. When GLCM is used in feature extraction, performance is lower (96.1%). So Gabor filter is more effective than the GLCM. The most successful result with the 98.8% performance rate was obtained with kNN in the classification. Table 4 gives a comparison of the performance analysis of the relevant studies. As can be seen from Table 3 and Table 4, the current study has more methods than the others. And the performance is higher than the related studies.

Table 3. A comparison of the relevant studies on breast cancer detection

Author(s)	Year	Specification	Used Database	Method	Application Environment
Chougrad H. et al. [2]	2018	Computer-aided Diagnosis (CAD) to help classify mammography mass lesions	DDSM, BCDR, INbreast, The Merged Dataset (MD)	Convolutional Neural Networks (CNN)	Breast Cancer
Dhahbia S. et al. [3]	2018	Reducing false positives in CAD systems for cancer diagnosis	DDSM	Random Forest, SVM, Decision Tree	Breast Cancer
Rezk E. et al. [4]	2017	Data sampling and classification for cancer	MITOS 2012 dataset3	NB, SVM, Pattern net (PN), Cascade forward net (CFN), Feed forward net (FFN)	Breast Cancer
Aswathy M. A. et al. [5]	2016	Cancer detection on digital images	Special Collected Data	General Classifier Neural Network (GCNN)	Breast Cancer
Khalilabad N. D. et al. [6]	2016	Fully automatic classification of microarray images	Stanford Microarray Database (Stanford, 2015)	Decision Tree	Breast Cancer
Saha M. et al. [7]	2017	Efficient deep learning model for mitosis detection	MITOS-ATYPIA-14, ICPR-2012, AMIDA-13	F-score	Breast Cancer
Wang P. et al. [9]	2016	Automatic quantitative image analysis technique of BCH images	Dataset not specified	Wavelet decomposition and multi-scale region-growing (WDMR), Double-strategy splitting model (DSSM), Curvature Scale Space (CSS), SVM, Chain-like agent genetic algorithm (CAGA).	Breast Cancer
Berbar M. A. et al. [11]	2018	Hybrid methods for feature extraction in classification	DDSM, MIAS	ST-GLCM, GLCM, Wavelet-CT1 and Contourlet (CT2)	Breast Cancer
Carvalho E. D. et al. [12]	2018	Method of differentiation of benign and malignant masses in digital mammograms using texture analysis based on phylogenetic diversity	DDSM, MIAS	Random Forest, Neural Network, MLP, SMO	Breast Cancer
Sun W. et al. [15]	2017	A graph based semi-supervised learning scheme using deep CNN for cancer detection	L denote labeled dataset and U denote unlabeled dataset	ANN, SVM and CNN algorithms	Breast Cancer
Majid A. et al. [17]	2014	Prediction of breast and colon cancers from imbalanced data	Cancer/non-cancer (C/NC), breast/nonbreast cancer (B/NBC) and colon/non-colon cancer (CC/NCC)	K-Nearest Neighbor and SVM	Breast and Colon Cancers
Sampaio W. B. et al. [19]	2015	Detection of masses in mammograms with adaption to breast density	DDSM	Genetic Algorithm, Phylogenetic Trees, LBP and SVM	Breast Cancer
Cheng H.D. et al. [23]	2010	Automated detection and classification using ultrasound images	Dataset not specified	Linear classifiers, Artificial neural networks, Bayesian neural networks, Decision tree, SVM, Template matching, Human classifier	Breast Cancer
Nilashi M. et al. [26]	2017	A knowledge-based system	Wisconsin Diagnostic Breast Cancer (WDBC) and Mammographic mass datasets	Fuzzy logic method	Breast Cancer
Rastghalam R. et al. [28]	2016	Cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images	Dataset not specified	Hidden Markov Model (HMM)	Breast Cancer
Mohammed M. A. et al. [32]	2018	Neural network and multi-fractal dimension features	Dataset of Breast Cancer Department of the Oncology Specialist Hospital in Baghdad, Iraq	Neural network and multi-fractal dimension features	Breast Cancer
Lee M. Y. et al. [34]	2010	For breast cancer diagnosis using standardized thermograph images	Thermograph images of unspecified sources	Entropy and decision tree induction	Breast Cancer
Francis S. V. et al. [35]	2014	Detection of cancer from rotational thermography images	Dataset not specified	SVM	Breast Cancer
EtehadTavakol M. et al. [36]	2013	Cancer detection thermal images using bispectral invariant features	Dataset not specified	Adaboost classifier	Breast Cancer
Abdel-Nasser M. et al. [37]	2016	Automatic nipple detection in breast thermograms	Proeng database	Multi-layer perceptron (MLP)	Breast Cancer
Elyasi I. et al. [41]	2016	Speckle reduction in breast cancer ultrasound images	Dataset not specified	Homogeneity Modified Bayes Shrink (HMBS)	Breast Cancer
Singh B. K. et al. [43]	2017	Risk stratification of 2D ultrasound-based breast lesions		Back-propagation artificial neural network (BPANN) and SVM	Breast Cancer

			The database of 178, B-mode breast ultrasound images		
Baykara M.* [current study]	2021	Computer assist system design in breast cancer diagnosis	DMR (Database For Mastology Research)	Decision Tree, SVM, Logistic Regression Analysis, kNN, Linear Discriminant Analysis	Breast Cancer

Table 3. A comparison of the performance analysis of the relevant studies

Relevant Studies	Classification Method				
	SVM	kNN	Decision Tree	Logistic Regression	Linear Discriminant
Baykara M. [Current Study]	95,6	98,8	91,2	80,3	80,5
Dhahbia S. et al. [3]	80,01	X	79,12	X	X
Rezk E. et al. [4]	78	X	X	X	X
Khalilabad N. D. et al. [6]	X	X	95.23	X	X
Wang P. et al. [9]	96.19	X	X	X	X
Sun W. et al. [15]	85.52	X	X	X	X
Majid A. et al. [17]	95.18	93.47	X	X	X
Sampaio W. B. et al. [19]	92,99	X	X	X	X
Cheng H.D. et al. [23]	94.25	X	96	X	X
Lee M. Y. et al. [34]	X	X	90	X	X
Francis S. V. et al. [35]	83.3	X	X	X	X
Singh B. K. et al. [43]	94,4	X	X	X	X

5. Conclusion

Image processing methods provide the development of a computer-aided diagnostic system to assist the expert in cancer diagnosis. With the use of intelligent methods, the performance of these diagnostic systems is increasing in recent years. In particular, methods such as ANN, CNN, and SVMs are the most popular methods that increase diagnostic success.

In this study, new effective software is developed for detecting breast cancer. Additionally, this paper presents a performance evaluation of classification algorithms for a computer-aided diagnosis system on thermal mammography images. The implemented software can be used detection and diagnosis of breast cancer in an early stage. It provides an objective result with good accuracy for the expert in breast cancer diagnosis.

In order to reach the best classification results in this study, very different structures have been tried both as a feature extraction method and as a classification method. In the feature extraction phase, both the GLCM-derived properties and the Gabor filter were used. For classification, SVM, kNN, Decision Tree, Logistic Regression, Linear Discriminant methods were used given in Table 4. In the study, feature extraction was first done with GLCM and the classification was done by the methods mentioned above. Then the attributes extracted by the Gabor filter are used for feature extraction and the classification was done by the methods mentioned above. The success rates achieved by using the GLCM method are lower. This means that the values obtained with Gabor are enhancing the classification performance.

The developed system is a sufficiently high-performance expert system to classify thermal medical images. In future works, it is planned to develop high-performance computer-aided diagnostic systems by using some deep learning methods.

References

[1]. Huang, Q., Huang, X., Liu, L., Lin, Y., Long, X., & Li, X. (2018). A case-oriented web-based training system for breast cancer diagnosis. *Computer Methods and Programs in Biomedicine*, *156*, 73–83. http://doi.org/10.1016/j.cmpb.2017.12.028.

[2]. Chougrad, H., Zouaki, H., & Alheyane, O. (2018). Deep Convolutional Neural Networks for breast cancer screening. *Computer Methods and Programs in Biomedicine*, *157*, 19–30. http://doi.org/10.1016/j.cmpb.2018.01.011

[3]. Dhahbi, S., Barhoumi, W., Kurek, J., Swiderski, B., Kruk, M., & Zagrouba, E. (2018). False-positive reduction in computer-aided mass detection using mammographic texture analysis and classification. *Computer Methods and Programs in Biomedicine*, *160*, 75–83. http://doi.org/10.1016/j.cmpb.2018.03.026

[4]. Rezk, E., Awan, Z., Islam, F., Jaoua, A., Al Maadeed, S., Zhang, N., Rajpoot, N. (2017). Conceptual data sampling for breast cancer histology image classification. Computers in Biology and Medicine, 89(July), 59–67. http://doi.org/10.1016/j.compbiomed.2017.07.018

[5]. Aswathy, M. A., & Jagannath, M. (2016). Detection of breast cancer on digital histopathology images: Present status and future possibilities. Informatics in Medicine Unlocked, 8(October 2016), 74–79. http://doi.org/10.1016/j.imu.2016.11.001

[6]. Khalilabad, N. D., Hassanpour, H., & Abbaszadegan, M. R. (2016). Fully automatic classification of breast cancer microarray images. Journal of Electrical Systems and Information Technology, 3(2), 348–359. http://doi.org/10.1016/j.jesit.2016.06.001

[7]. Saha, M., Chakraborty, C., & Racoceanu, D. (2018). Efficient deep learning model for mitosis detection using breast histopathology images. Computerized Medical Imaging and Graphics, 64(June 2017), 29–40. http://doi.org/10.1016/j.compmedimag.2017.12.001

[8]. Robertson, S., Azizpour, H., Smith, K., & Hartman, J. (2018). Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. Translational Research, 194, 19–35. http://doi.org/10.1016/j.trsl.2017.10.010

[9]. Wang, P., Hu, X., Li, Y., Liu, Q., & Zhu, X. (2016). Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Processing, 122, 1–13. http://doi.org/10.1016/j.sigpro.2015.11.011

[10]. Peng, L., Chen, W., Zhou, W., Li, F., Yang, J., & Zhang, J. (2016). An immune-inspired semi-supervised algorithm for breast cancer diagnosis. Computer Methods and Programs in Biomedicine, 134(61472092), 259–265. http://doi.org/10.1016/j.cmpb.2016.07.020

[11]. Berbar, M. A. (2018). Hybrid methods for feature extraction for breast masses classification. Egyptian Informatics Journal, 19(1), 63–73. http://doi.org/10.1016/j.eij.2017.08.001

[12]. Carvalho, E. D., de Carvalho Filho, A. O., de Sousa, A. D., Silva, A. C., & Gattass, M. (2018). Method of differentiation of benign and malignant masses in digital mammograms using texture analysis based on phylogenetic diversity. Computers & Electrical Engineering, 67(November 2017), 210–222. http://doi.org/10.1016/j.compeleceng.2018.03.038

[13]. Cherif, W. (2018). Optimization of K-NN algorithm by clustering and reliability coefficients: application to breast-cancer diagnosis. Procedia Computer Science, 127, 293–299. http://doi.org/10.1016/j.procs.2018.01.125

[14]. Li, Y., Chen, H., Rohde, G. K., Yao, C., & Cheng, L. (2015). Texton analysis for mass classification in mammograms. Pattern Recognition Letters, 52, 87–93. http://doi.org/10.1016/j.patrec.2014.10.008

[15]. Sun, W., Tseng, T. L. B., Zhang, J., & Qian, W. (2017). Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Computerized Medical Imaging and Graphics, 57, 4–9. http://doi.org/10.1016/j.compmedimag.2016.07.004

[16]. Yassin, N. I. R., Omran, S., El Houby, E. M. F., & Allam, H. (2018). Machine learning techniques for breast cancer computer-aided diagnosis using different image modalities: A systematic review. Computer Methods and Programs in Biomedicine, 156, 25–45. http://doi.org/10.1016/j.cmpb.2017.12.012

[17]. Majid, A., Ali, S., Iqbal, M., & Kausar, N. (2014). Prediction of human breast and colon cancers from imbalanced data using nearest neighbor and support vector machines. Computer Methods and Programs in Biomedicine, 113(3), 792–808. http://doi.org/10.1016/j.cmpb.2014.01.001

[18]. Sayed, A. M., Zaghloul, E., & Nassef, T. M. (2016). Automatic Classification of Breast Tumors Using Features Extracted from Magnetic Resonance Images. Procedia Computer Science, 95, 392–398. http://doi.org/10.1016/j.procs.2016.09.350

[19]. De Sampaio, W. B., Silva, A. C., De Paiva, A. C., & Gattass, M. (2015). Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. Expert Systems with Applications, 42(22), 8911–8928. http://doi.org/10.1016/j.eswa.2015.07.046

[20]. Dora, L., Agrawal, S., Panda, R., & Abraham, A. (2017). Optimal breast cancer classification using Gauss–Newton representation based algorithm. Expert Systems with Applications, 85, 134–145. http://doi.org/10.1016/j.eswa.2017.05.035

[21]. Pak, F., Kanan, H. R., & Alikhassi, A. (2015). Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution. Computer Methods and Programs in Biomedicine, 122(2), 89–107. http://doi.org/10.1016/j.cmpb.2015.06.009

[22]. Dehghan Khalilabad, N., & Hassanpour, H. (2017). Employing image processing techniques for cancer detection using microarray images. Computers in Biology and Medicine, 81(December 2016), 139–147. http://doi.org/10.1016/j.compbiomed.2016.12.01

[23]. Cheng, H. D., Shan, J., Ju, W., Guo, Y., & Zhang, L. (2010). Automated breast cancer detection and classification using ultrasound images: A survey. Pattern Recognition, 43(1), 299–317. http://doi.org/10.1016/j.patcog.2009.05.012

[24]. Bhardwaj, A., & Tiwari, A. (2015). Breast cancer diagnosis using Genetically Optimized Neural Network model. Expert Systems with Applications, 42(10), 4611–4620. http://doi.org/10.1016/j.eswa.2015.01.065

[25]. Kandlikar, S. G., Perez-Raya, I., Raghupathi, P. A., Gonzalez-Hernandez, J. L., Dabydeen, D., Medeiros, L., & Phatak, P. (2017). Infrared imaging technology for breast cancer detection – Current status, protocols and new directions. International Journal of Heat and Mass Transfer, 108, 2303–2320. http://doi.org/10.1016/j.ijheatmasstransfer.2017.01.086

[26]. Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). A knowledge-based system for breast cancer classification using fuzzy logic method. Telematics and Informatics, 34(4), 133–144. http://doi.org/10.1016/j.tele.2017.01.007

[27]. Guo, R., Lu, G., Qin, B., & Fei, B. (2018). Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review. Ultrasound in Medicine and Biology, 44(1), 37–70. http://doi.org/10.1016/j.ultrasmedbio.2017.09.012

[28]. Rastghalam, R., & Pourghassem, H. (2016). Breast cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images. Pattern Recognition, 51, 176–186. http://doi.org/10.1016/j.patcog.2015.09.009

[29]. Karabadji, N. E. I., Seridi, H., Bousetouane, F., Dhifli, W., & Aridhi, S. (2017). An evolutionary scheme for decision tree construction. Knowledge-Based Systems, 119, 166–177. http://doi.org/10.1016/j.knosys.2016.12.011

[30]. Wang, H., Zheng, B., Yoon, S. W., & Ko, H. S. (2018). A support vector machine-based ensemble algorithm for breast cancer diagnosis. European Journal of Operational Research, 267(2), 687–699. http://doi.org/10.1016/j.ejor.2017.12.001

[31]. Kaymak, S., Helwan, A., & Uzun, D. (2017). Breast cancer image classification using artificial neural networks. Procedia Computer Science, 120, 126–131. http://doi.org/10.1016/j.procs.2017.11.219

[32]. Mohammed, M. A., Al-Khateeb, B., Rashid, A. N., Ibrahim, D. A., Abd Ghani, M. K., & Mostafa, S. A. (2018). Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. *Computers and Electrical Engineering*, *0*, 1–12. http://doi.org/10.1016/j.compeleceng.2018.01.033

[33]. Karabatak, M. (2015). A new classifier for breast cancer detection based on Naïve Bayesian. *Measurement: Journal of the International Measurement Confederation*, *72*, 32–36. http://doi.org/10.1016/j.measurement.2015.04.028

[34]. Lee, M. Y., & Yang, C. S. (2010). Entropy-based feature extraction and decision tree induction for breast cancer diagnosis with standardized thermograph images. *Computer Methods and Programs in Biomedicine*, *100*(3), 269–282. http://doi.org/10.1016/j.cmpb.2010.04.014

[35]. Francis, S. V., Sasikala, M., Bhavani Bharathi, G., & Jaipurkar, S. D. (2014). Breast cancer detection in rotational thermography images using texture features. *Infrared Physics and Technology*, *67*, 490–496. http://doi.org/10.1016/j.infrared.2014.08.019

[36]. Etehadtavakol, M., Chandran, V., Ng, E. Y. K., & Kafieh, R. (2013). Breast cancer detection from thermal images using bispectral invariant features. International Journal of Thermal Sciences, 69, 21–36. http://doi.org/10.1016/j.ijthermalsci.2013.03.001

[37]. Abdel-Nasser, M., Saleh, A., Moreno, A., & Puig, D. (2016). Automatic nipple detection in breast thermograms. Expert Systems with Applications, 64, 365–374. http://doi.org/10.1016/j.eswa.2016.08.026

[38]. Li, J. B., Wang, Y. H., & Tang, L. L. (2012). Mammogram-based discriminant fusion analysis for breast cancer diagnosis. Clinical Imaging, 36(6), 710–716. https://doi.org/10.1016/j.clinimag.2012.01.041

[39]. Li, J. B., Peng, Y., & Liu, D. (2013). Quasiconformal kernel common locality discriminant analysis with application to breast cancer diagnosis. Information Sciences, 223, 256–269. https://doi.org/10.1016/j.ins.2012.10.016

[40]. Shi, Y., Dai, D., Liu, C., & Yan, H. (2009). Sparse discriminant analysis for breast cancer biomarker identification and classification. Progress in Natural Science, 19(11), 1635–1641. https://doi.org/10.1016/j.pnsc.2009.04.013

[41]. Elyasi, I., Pourmina, M. A., & Moin, M. S. (2016). Speckle reduction in breast cancer ultrasound images by using homogeneity modified bayes shrink. Measurement: Journal of the International Measurement Confederation, 91, 55–65. https://doi.org/10.1016/j.measurement.2016.05.025

[42]. Amin, K. M., Shahin, A. I., & Guo, Y. (2016). A novel breast tumor classification algorithm using neutrosophic score features. Measurement: Journal of the International Measurement Confederation, 81, 210–220. https://doi.org/10.1016/j.measurement.2015.12.013

[43]. Singh, B. K., Verma, K., Thoke, A. S., & Suri, J. S. (2017). Risk stratification of 2D ultrasound-based breast lesions using hybrid feature selection in machine learning paradigm. Measurement: Journal of the International Measurement Confederation, 105, 146–157. https://doi.org/10.1016/j.measurement.2017.01.016

[44]. Fenster, A., Surry, K., Smith, W., & Downey, D. B. (2004). The use of three-dimensional ultrasound imaging in breast biopsy and prostate therapy. Measurement: Journal of the International Measurement Confederation, 36(3–4), 245–256. https://doi.org/10.1016/j.measurement.2004.09.013

[45]. Masamha, T., Mnkandla, E., Jaison, A. (2017, September). Logistic regression analysis of information communication technology projects' critical success factors: A focus on computer networking projects. InAFRICON, IEEE, pp. 963-967, 2017.

[46]. Hai, Y., Wong, S. Y., Yuen, K. Y., Cheng, V. C., Tsui, K. L. (2011, July). Logistic regression analysis for Predicting Methicillin-resistant Staphylococcus Aureus (MRSA) in-hospital mortality. In Intelligence and Security Informatics (ISI), 2011 IEEE International Conference, pp. 349-353, 2011.

[47]. Li, M., Yuan, B., 2D-LDA: A statistical linear discriminant analysis for image matrix. Pattern Recognition Letters, 26(5), 527-532, 2005.

[48]. Gyamfi, K. S., Brusey, J., Hunt, A., Gaura, E., Linear classifier design under heteroscedasticity in Linear Discriminant Analysis. Expert Systems with Applications, 79, 44-52, 2017.

[49]. Internet: Database for Mastology Research, http://visual.ic.uff.br/dmi/.