MAMMOGRAM CLASSIFICATION TECHNIQUE BY USING NEURO FUZZY SVM FOR TUMOR EXTRACTION

1M PUNITHA, 2K PERUMAL

1Research Scholar, Madurai Kamaraj University, TamilNadu, India
2Professor, Madurai Kamaraj University, TamilNadu, India

ABSTRACT

Image classification is helping radiologists to improve the accuracy of tumor detection in mammogram images for better diagnostics. The main aim of this proposed work is to build an efficient Neuro-Fuzzy support vector machine Techniques to detect and extract the tumor in the mammogram images and get an efficient result. The proffered classifiers are to achieve a very fast, simple, and efficient breast cancer diagnosis. The edge-based image segmentation and Neuro-fuzzy support vector machine are to find the Abnormality of classification such as cysts, calcification, fibro adenomas, and scar tissue. The tumor pixel values are calculated easily in a short time. Based on this experimental result, the overall performance of the proposed method is improved significantly. Furthermore, it can be inferred and ensures that the best classification accuracy of 99.85% ratio. And it has been compared in various Existing methods. There is no other research has been done for this type of research.

Keywords: Neuro-Fuzzy support vector machine, Skeletonization, and Edge base Segmentation algorithm

1. INTRODUCTION

Breast cancer disease is accounted for as the second most fatal malignant growth among ladies on the world and the open mindfulness about it has been expanding during the most recent couple of decades [1]. In many countries, including Malaysia, breast cancer has become a major health problem. Early detection of breast cancer is only possible when regular screening examinations are conducted. The best way to detect breast cancer In its treatable and early-stage is through mammography. A mammogram is a special type of X-ray photograph that uses high-resolution film, high-contrast, and low-dose X-ray for imaging the breasts. It helps to detect and diagnose breast cancer effectively. However, not all breast cancers can be detected by mammograms [2]. Neural networks and fuzzy logics had been researched to the problems in digital image processing. A Neuro-Fuzzy System is an adaptable framework prepared by heuristic taking in methods got from neural systems that can be seen as a 3-layer neural system with fuzzy loads and unique activation functions are always interpretable as a fuzzy system uses constraint learning procedures is a function approximation (classifier, controller). In this paper, Adaptive Neuro-fuzzy Inference System (ANFIS) is introduced, which is a fuzzy deduction framework executed in the system of a versatile system. This ANFIS preparing calculation is proposed by Jang. By utilizing a hybrid learning system, the proposed ANFIS can develop an info yield mapping which depends on both human information (as fuzzy rules) and learning. The proposed work is carried in two stages. In the first stage, noisy image (i.e. received or captured by camera) is denoised by the median filter. The denoised picture is additionally upgraded by ANFIS. The structure of the proposed work, to make the process robust against noise, is a combination of nonlinear switching median filter and neuro-fuzzy systems. SVM is a learning machine utilized as a device for information order, function approximation, etc, due to its Generalization capacity and has discovered Achievement in Many applications [3]. The component of SVM is that it limits and upper bound of speculation mistake by growing the edge between confining hyper plane and dataset. SVM has an additional favorable position of programmed model choice as in both the ideal number and areas of the essential capacities are naturally acquired during preparing. The performance of SVM classification in the Mammogram image [4][5].
2. LITERATURE SURVEY

Fathima Abubacker et.al [6] worked with Genetic Association Rule Miner (GARM) with the learning ability of neural systems and highlights are removed. Punitha et.al [7] using the rough fuzzy feature extraction in mammogram images by using FRIS and FRNN Algorithms for segmentation and feature extraction. Ahmad Taher Azar et.al [8] introduced the combination of Takagi-Sugeno-Kang (TSK) fuzzy inference system Adaptive neuro-fuzzy inference system (ANFIS) Subtractive clustering Scaled conjugate gradient Linguistic hedge (LH) Feature selection (FS) for segmentation and feature extractions are used in different classification methods. Indrajeet Kumar et.al [9] focuses on Texture based feature extraction with SVM Hierarchical classifiers which has a highly accurate range in the images. Ahmad Taher Azar et.al [10] proposed different types of SVM Classifiers to extract the features in mammogram images and 97% accuracy radio has been active. Cheng-Min Chao et.al [11] used data mining technology to establish a classification of breast cancer survival patterns and the SVM classifier established a high accuracy rate. The mining rules-based feature extractions are mainly forced in this work. Punitha et.al [13] say histogram orientation gradient with Hough transform for feature extraction for tumor detection in mammogram images and segmentation they using double thresholding methods. Samuel Rud et.al [14] mainly focused on detecting the low-frequency vibrations by using the SVM classification technique and concentrated on the normal and abnormal conditions of the cardiac conditions. Yu Zhang et.al[15] introduce the new machine learning algorithms for classifying the masses in the mammogram images and they compared various classification techniques for getting better classification accuracy and they conclude the decision tree is the best classification results. M. R. Senapati et.al [16] is proposed a firefly algorithm with a local linear wavelet neural network for identifying cancer and compared in various classification techniques. Mouna Zouari Mehdi et.al[17] mainly concentrate the microcalcifications are tiny deposits of calcium located in breast tissue and also finding the surrounding tissue by using CAD with stretching followed by hybrid enhancement approaches. R.R. Janghel et al[18] here they used a soft computing technique with six types of models of a neural network to find cancer in the mammogram images and they have taken the datasets from the UCI repository and all the models solve the problem to a reasonable extent. A. H. El-Baz et.al [19] here they followed hybrid intelligent systems based on rough set and ensemble classifier for breast cancer diagnosis and f-nearest neighbor algorithm are helped to find the cancer cells in the breast images. P. Foggia et.al [20] this paper mainly focused on a graph theoretical clusters analysis to found the tumors in the mammographic images and they prove the cluster classification with the brightness of the given input images; cluster detection is helped to found cancer in the mammogram images. Yong Fan et.al[21] this paper presents machine learning and deformation based morphometry and also support vector machine recursive feature elimination technique is used to computed the features from the extracted regions and got the low accuracy of 91.8 ratios. JeongHee et.al [22] introduced the novel algorithms for classifying the medical images, especially x-ray images and they followed color descriptor and histogram descriptor for extracted the global and local parts of the images and the support vector machine supports the membership scores for each image. Arbab Masood Ahmad et.al [23] they developed a system for classifying the mass in a mammogram as either benign or malignant and used genetic programming evolved artificial neural network for classification process and also found the gray level co-occurrence matrix was used in region boundary and got the testing accuracy of 100% ratio on their research. Hans Bornefalk et.al [24] introduced quadrature filters and using level set methods for segmented and filters are helped to getting the accuracy and sensitivity and got the 90% sensitivity.
3. PROPOSED METHODOLOGY

The proposed work comprises of developing new Feature Extraction methods. This is used for extracting the features in the mammography images. Here we used neuro-fuzzy with SVM Classifiers for extracting the tumor portion in the mammography images. The primary part of this work is using the skeletonization for preprocessing. This is used to reduce the noise in the input images then edge-based segmentation can be used for tumor segmentation part after that neuro fuzzification can be done and finally, the neuro-fuzzy with SVM classifiers are used to classify the tumor from the features extracted from the mammogram images. The progression of our proposed work is clarified by methods for Figure 2.

3.1 Pre-processing

The Skeletonization can be done to remove each connected component in a binary image to a single-pixel wide skeleton and it acts as an erosion process. This includes contracting the picture until the region of intrigue is 1 pixel wide. It can allow accurate image processing on other operations. The result of this skeletonization is very thinning and accurate and it removes the wanted and unwanted pixels, while dissimilar in removing time.

3.2 Edge Based Segmentation

After skeletonization, the image can be segmented and it is mainly concentrated to detect the tumor part and discontinuities in brightness. So the tumor can easily be segmented from the input mammogram image. The below steps can be followed by finding the tumor segmentation in mammogram images.

**Pseudocode for Edge Based Segmentation**

Step 1: Read the input image
Step2: Apply horizontal mask Ga and vertical mask Gb to the input image (for Smoothing)
Step3: Apply different edge detection algorithms and find the gradient
Step4: Construct separate image for Ga and Gb
Step5: Results are consolidated to locate the magnitude of the gradient |G| = √Ga² + Gb²
Step6: The total magnitude is the output slope magnitude image
Step7: For some slope magnitude images, the pixel's esteems are excessively little or excessively high.

To improve the visibility of those images, scaling has to be done for small values, it has to be scaled up by an appropriate factor. For large values, it has to be scaled down by the appropriate factor.

The tumor range depends upon the image that we are given to the input image.

3.3 Neuro-Fuzzy

It is a hybrid combination of Fuzzy systems as well as a neural network so it's widely called it us Neuro-fuzzy systems and it's having the universal approximations of if-then rules. Neuro-fuzzy systems are usually having three-layer neural network one input layer one output layer and one hidden layer or middle layer represents fuzzy rules. and it used the learning algorithms derived from...
neural network theory. Fuzzy sets encoded the connection weights.

3.4 Support Vector Machine

Support vector machine mainly used in classification problems and it's used for both classification and regression challenges. Most of the problem in SVM is based machine learning algorithm. In image processing SVM is fundamentally a binary classification algorithm that should be followed and it maximizes a margin around the separating hyperplane.

3.5 Feature Extraction by using neuro Fuzzy based SVM classification

The crux in choosing this Neuro-Fuzzy Support Vector Machine (NFSVM) is simplicity and smooth implementation in comparison with other feature extraction techniques. In the context of our proposed work, NFSVM is incorporated in reducing the computational time and selecting the optimal neighborhood pixel from the original image. Tumors are distinctive properties of input patterns that help in identifying the tumor easily and Features are played a vital role in image processing. The main purpose of this proposed work is to find the imperfect or tumor instances of an object in the image. Here predominantly features are extracted from the segmented mammogram image and mainly focused on the four features are extracted such as CALC, CIRC, SPIC, and MISC. The resultant feature extraction image which is very important for deciding either the mammogram image should have the tumor or not. The features are introduced in table 1.

<table>
<thead>
<tr>
<th>Features</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALC</td>
<td>Calcification</td>
</tr>
<tr>
<td>CIRC</td>
<td>Well-defined/ Circumscribed mass</td>
</tr>
<tr>
<td>SPIC</td>
<td>Speculated mass</td>
</tr>
<tr>
<td>MISC</td>
<td>Ill-defined mass</td>
</tr>
</tbody>
</table>

Table 1 Each feature is having the unic methods in CALC features like small dots of calcium segmented in the breast and it has no pain and no feel so it's in a benign stage. CIRC is also benign stage and it is clearly defined along with at least 80% of its surface and remaining 20% and, at most is masked by the adjacent gland and it is in oval and round shape masses, an irregular shape suggests a great likelihood of malignancy. Then SPIC is a mass centrally dense lesion seen on mammogram images with sharp lines radiating from its margin. The MISC is an ill-defined mass circumscribed oval and round masses are usually benign and irregular shape suggests a greater like the hood of malignancy. In this study, we use NFSVM method which is used to resolve the classification problem and it is the very accurate and fastest method to extract the tumor in the mammogram images. According to the features the decision has been taken and the machine learning support vector machine hyperplane extract the features in short computational time.

![Figure 3: Proposed Method following diagram](image)

The proposed NFSVM Algorithm

Step 1: Load image
Step 2: Classify Features based on class labels
Step 3: Estimate tumor Support Value
While (instances! =null) Do
Stage 4: Support Value=Similarity between each occasion in the property Find Total Error Value
Step 5: If any instance < 0
Estimate Decision value = Support Value/Total Error
Rehash for all focuses until it will discharge
End If

The proposed algorithm has been classifying the features in the input images and estimates the tumor image and non-tumor images. If it's a tumor image, then classify the features whether it is in which type of tumor image until all the conditions should be satisfied and it gives the correct Feature Extraction from the mammogram image. Totally 60 images are used in this paper and 30 images are taken as
training sets remaining 30 images are taken for the testing set. Then we have to predict the image which are the images affected by a tumor or not. If the image found the tumor whether it is benign or malignant. According to the feature, extraction predicts value the tumor has been diagnosed according to the feature classifications. finally, the prediction has been done. The prediction has been done by the below-mentioned formula.

\[
\text{Predictive value} = \text{Original Value} - \text{Predictive Value}
\]

Our calculation of the probability of whether a tumor is malignant or benign is based on 4 features, like CALC, CIRC, SPIC, and MISC. With appropriate values substituted by the formula and we found the probability value is malignant or benign. Let consider the probability is more than 50% and Consider the Feature a shape matching the tumor range are calling it malignant (cancerous), and if the probability is less than 50%, it is benign (non-cancerous).

4. RESULTS AND DISCUSSION

In this section, we present the experimental results of applying different feature extraction for image classification and it has been implemented R Tool and all tested images are got from Madurai Rajaji Government Hospital.

Figure 4 shows the Feature Extraction images which are implemented in this paper

![Feature Extraction Images](image)

Figure 4: different Feature extractions from the input mammogram image

Table 2 Results of improved noise removal approach with segmentation

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>INPUT IMAGE</th>
<th>DENOISED IMAGE</th>
<th>SEGMENTED IMAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMAGE 1</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td>IMAGE 2</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td>IMAGE 3</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td>IMAGE 4</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td>IMAGE 5</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
</tbody>
</table>

Image (a) describes the calcification of the breast tissue small dots are segmented in the circle forms and it is considered as benign. Image (b) describes oval and round masses are usually benign are a malignancy. Image (c) shows the dense mass on the mammogram extracted image however in rare cases, certain begin lesion may appear in the form of speculated mass. Image (d) shows the unstructured mass appears in the mammogram image and it's in an oval as well as round masses are usually an irregular shape and this is also found as malignancy. Each image should have different features and it may depend upon the tumor segmentation.
In Table 2, the input image has been denoised and segmented, and this segmented portion has been evaluated, and the features are extracted from this image.

Table 3: Pixel Calculation

<table>
<thead>
<tr>
<th>Original Image</th>
<th>After Denoised pixel</th>
<th>Segmented pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1</td>
<td>0.38461</td>
<td>0.36540</td>
</tr>
<tr>
<td>I2</td>
<td>0.64587</td>
<td>0.60451</td>
</tr>
<tr>
<td>I3</td>
<td>0.86540</td>
<td>0.75248</td>
</tr>
<tr>
<td>I4</td>
<td>0.72453</td>
<td>0.70214</td>
</tr>
<tr>
<td>I5</td>
<td>0.45985</td>
<td>0.37564</td>
</tr>
</tbody>
</table>

Table 3 shows the denoised pixel values after Skeletonization for each image should have a different denoised pixel range. After that, the denoised image has segmented the tumor portion without noise for all input mammogram images. The segmented image has been trained and tested for feature extraction. After this denoised segmented image should have them cleared and accurate feature extraction for tumor part of the given input mammogram images that extracted feature pixel range has been calculated for accuracy purpose and it should evaluate in some quality measures to get the better accuracy result.

The above Pie chart shows the denoised segmented pixel range values and mentioned each image pixel values in different colors.

Table 4: Resultant Image for the Proposed Method

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>INPUT IMAGE</th>
<th>SEGMENTED IMAGE</th>
<th>PROPOSED METHOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMAGE 1</td>
<td><img src="image1.png" alt="Image 1" /></td>
<td><img src="segmented_image1.png" alt="Segmented Image 1" /></td>
<td><img src="proposed_image1.png" alt="Proposed Image 1" /></td>
</tr>
<tr>
<td>IMAGE 2</td>
<td><img src="image2.png" alt="Image 2" /></td>
<td><img src="segmented_image2.png" alt="Segmented Image 2" /></td>
<td><img src="proposed_image2.png" alt="Proposed Image 2" /></td>
</tr>
<tr>
<td>IMAGE 3</td>
<td><img src="image3.png" alt="Image 3" /></td>
<td><img src="segmented_image3.png" alt="Segmented Image 3" /></td>
<td><img src="proposed_image3.png" alt="Proposed Image 3" /></td>
</tr>
<tr>
<td>IMAGE 4</td>
<td><img src="image4.png" alt="Image 4" /></td>
<td><img src="segmented_image4.png" alt="Segmented Image 4" /></td>
<td><img src="proposed_image4.png" alt="Proposed Image 4" /></td>
</tr>
</tbody>
</table>

Table 4, results are clearly shown the Features are extracted correctly and each feature shows their unique features. Image 1 shows the unstructured tumor extraction and it is considered as Malignant. Image 2 shows the CALC Features and describes the small rounded calcium segmentation and it may consider as benign. In image 3 shows the MISC feature extraction methods and it's having circular form according to the tumor prediction it is also considered as benign. In image 4 CIRC features are shown the oval shape and the prediction is considered as benign.
4.1 Performance evaluation metrics

The quantitative findings of the performance of a classification test data can be evaluated by the statistical measures such as accuracy, sensitivity, and specificity and all images are find whether tumor as normal (N), benign (B), and malignant (M). The classifications made for each class. In the Classification stage, the class for each image in the test group is identified, and the average accuracy measure (AC) of the classification which is the overall correctness of the model is calculated as the sum of true negatives and a true positive divided by the total number of classifications (N) and is given by Accuracy. Here we evaluated the accuracy, sensitivity, specificity, and precision for image quality.

TP: True Positive predicted values.
TN: True Negative predicted values
FP: False Positive predicted values
FN: False Negative predicted values

For the above formula has been followed to predict the tumor values in the mammogram images.

Accuracy: To estimate the exactness of a test, we should calculate the proportion of true positive, and true negative in all evaluated images, and it can be defined as:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

Sensitivity: It mainly concentrates on the percentage of positive labeled instances that were predicted as positive values.

\[
\text{Sensitivity} = \frac{TP}{TP + FN}
\]

Specificity: It is mainly concentrated on the percentage of negative values labeled instances that were predicted as negative.

\[
\text{Specificity} = \frac{TN}{TN + FP}
\]

Precision: precision is the percentage of positive prediction that is correct.

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

4.2 Confusion matrix

It helps us to get the accurate classification result of the input mammogram images and predict the accurate feature extraction.

Table 5: Representation of Confusion matrix

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted Positive</th>
<th>Predicted Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>X1</td>
<td>X2</td>
</tr>
<tr>
<td>N</td>
<td>Y1</td>
<td>Y2</td>
</tr>
</tbody>
</table>

Where the row represents the actual values and column represents the classified values. This matrix helps the extract of the features in the input mammogram images and predicts the correct positive and correct negative values. All the positive predicted images should have cancer images and negative predicted values should consider non-cancer images. According to the predicted value, we concentrate the main four feature extraction class which is identified after the prediction of positive and negative images.

Table 6: Confusion matrix for Training image in NFSVM feature Classification

<table>
<thead>
<tr>
<th>NFSVM Feature class</th>
<th>CALC</th>
<th>CIRC</th>
<th>SPIC</th>
<th>MISC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Benign</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Malignant</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Above table 6 has shown the result of this implementation results which has the feature class whether having the tumor of benign, normal, and malignant. The out of thirty mammogram training images should have the 18 as the normal in the class of CALC and CIRC. 7 images should have the benign and 5 images have malignant out of thirty images. All predicted features have been correctly identified with the help of a neuro-fuzzy SVM classifier. If we using this
Figure 6: Graphical representation NFSVM feature Classification

Figure 6 represents the feature class in a different classification. It shows a clear idea about the tumor class representation and accurate graphical results. It represents each feature class should have how much normal cases, how much benign cases, and how many malignant cases are in given input images we are taken from the datasets.

6.1 Database

In this Experiment, all mammogram images are taken from the Mini-MIAS database [11] containing left and right breast images for a total of 60 is used in this study. All images containing hypothesized masses of the mentioned features and a selection of normal types are considered. Images of 1024 X 1024 pixels and at 8-bit greyscale level.

6.2 Tools

This paper has implemented in the R Tool. This tool is very user friendly for image processing with a high-quality tool. By using this tool we got an accurate result.

Table 7: Comparison of various Existing Feature Extraction Techniques

<table>
<thead>
<tr>
<th>METHODS/QM</th>
<th>SVM</th>
<th>L9SV</th>
<th>SSVM</th>
<th>NFSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURACY</td>
<td>97.1429</td>
<td>95.2456</td>
<td>96.5174</td>
<td>99.8451</td>
</tr>
</tbody>
</table>

Table 7 shows the performance measures in different methods already used in different papers but our proposed methods of Neuro-Fuzzy SVM should have the Highest accuracy ratio of 99.84% and no other researchers are used this type of Classification methods. Column two shows the Support Vector Machine (SVM) only and then maximum accuracy is 98.24. Column three shows the Lagration support vector Machine (LSVM) the maximum accuracy of this is 96.52% column four shows the Smooth Support Vector Machine (SSVM) is the maximum accuracy ratio 96.56% and the last column is the proposed Neuro-fuzzy with support vector machine (NFSVM) is got the highest accuracy ratio of 99.85%.

Figure 7 shows a comparison of the Graphical representation of image quality measures. It clearly shows the high accuracy in the proposed method.
Figure 8: Graphical representation of accuracy in various comparison methods.

Figure 8 shows the clear high accuracy ratio to compare in various existing methods and also the proposed method of NFSVM shows a high accuracy ratio of 99.85%.

Figure 9: Graphical representation of the sensitivity in various comparison methods.

Figure 9 shows the clear sensitivity ratio report for existing and proposed methods clearly.

Figure 10: Graphical representation of Specificity in various comparison methods.

Figure 10 and 11 shows the Clear High specificity and high precision ratio for proposed methods. So this the NFSVM is having High-Quality measures for in this research.

Figure 11: Graphical representation of Precision in various Comparison methods.
5. CONCLUSION AND FUTURE WORK

The ultimate goal of this paper is to show an efficient classification technique to detect the presence of the tumor and extracting the features easily like CALC, CIRC, SPIC and MISC. According to the classification result, we easily find the tumor prediction like normal or abnormal and we have discussed and compared some image quality measures with other existing feature extraction methods. This research has shown that our proposed methods for feature extraction are very effective for the classification of normal and abnormal mammogram images. The evaluation of the methods was carried out in real-time images. The best-obtained accuracy is 98.84%. In future work, feature selection techniques like the one in will be applied to reduce the dimension of a feature Vectors and the time needed for classification.

REFERENCES

[17] Mouna Zouari Mehdi1 ꞏ Norhene Gargouri Ben Ayed2 ꞏ Alima Damak Masmoudi3 ꞏ Dorra Sellami4 ꞏ Riadh Abid5,” An efficient microcalcifications detection based on dual Spatial /spectral processing” Published online:


