International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V70I3P219

Volume 70 Issue 3, 170-178, March 2022

Original Article

Detection and Classification of Breast Cancer Using Machine Learning Techniques for Ultrasound Images

Akila Victor[1], Bhuvanjeet Singh Gandhi[2], Muhammad Rukunuddin Ghalib[3], Ramani Selvanambi[4]

[1]*Associate Professor Grade 1, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.*

[2] *Student, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.*

[3] *Senior Lecturer, School of Sciences, Engineering and Computing, De Montfort University, Dubai, UAE.*

[4]*Assistant Professor Senior Grade 1, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.*

[1]a.akilavictor@gmail.com, [2]bhuvanjeetsinghgandhi@gmail.com, [3]ghalib.it@gmail.com

Abstract — *One of the most common diseases in the world is cancer. There are many different forms of cancer, and one of the most frequent is breast cancer. Breast cancer can affect anybody, but it most usually affects women. Breast cancer can be cured quickly with early identification and a better knowledge of the illness. A computer-aided diagnostic (CAD) system allows us to uncover several ways to identify and diagnose cancer problems. The primary motivation is for accurate detection to detect cancer as soon as possible. Pre-processing, segmentation, feature extraction, and classification are the four critical steps of detection and identification. Pre-processing techniques employed in this study include median filtering and histogram equalization. For segmentation, a hybrid technique is utilized, and for feature extraction, fundamental methods are applied. For classification, the Support Vector Machine (SVM) is proposed and employed. SVM's accuracy is then compared to that of other machine learning approaches such as boosted tree (BT), random forest (RF), Naive Bayes (NB), and convolutional neural network (CNN). The results obtained are tabulated, and an accuracy of 93.4% is obtained from the SVM classifier.*

Keywords — *Ultrasound images, Histogram equalization, Support vector machine, Convolutional neural network, Accuracy.*

I. INTRODUCTION

Ultrasound imaging is a challenging technique. It uses high-frequency sound waves. It is one of the effective methods to diagnose breast cancer. The widespread disease in today's world is cancer. Breast cancer, lung cancer, skin cancer, liver cancer etc., are some of the popular forms in which cancer exists. Each has its own impacts. Cancer itself is a life-threatening disease, and early-stage detection leads to an easier cure rather than finding it in an advanced stage.

Cancer is divided into two types: primary cancer and secondary cancer. Primary cancer is cancer that develops for the first time in a specific region of the body. Secondary cancer invades from primary cancer, is quite challenging to cure when compared to primary cancer. The proposed research focuses on the identification and categorization of ultrasound pictures associated with breast cancer in order to aid in early detection and treatment. It will be extremely beneficial in today's environment if the procedure can be made quick and accurate.

There are various types of breast cancer as well. Broadly, there are four important stages in breast cancer, out of which the first stage has a better curing rate. However, the patients in other stages require vigorous medications, but they are durable too. So, it's always better with earlier detection. Although detection plays a major role, similarly, prevention also plays a vital role. It's always advised for a female of more than forty plus age to undergo a series of examinations once every year.

This paper proposes to automate the system for detection and the classification of breast cancer. The initial phase can help to reduce the noise in the input ultrasound image by means of the filtering technique. The segmentation phase aids in the identification of the region of interest (ROI). After that, the traits, i.e., the features, are selected and fed into a machine learning classifier to determine whether they are benign or cancerous. The study article starts with a discussion of related studies in the subject of breast cancer detection, then moves on to the overall architecture and technique proposed. It also discusses the findings and conclusions, as well as a comparison of various machine learning classifiers for the task of breast cancer detection and classification, such as the Boosted Tree (BT), Random Forest (RT), Naive Bayes (NB), and Convolutional Neural Network (CNN), as well as the scope of future research in this field.

II. RELATED WORKS

Breast cancers are classified as ductal carcinoma, invasive ductal carcinoma, lobular carcinoma, inflammatory carcinoma etc. It's always that the death rate keeps increasing every year [1]. [2] Speaks on the risk of cancer in the life of a female. [3] The CAD system helps the radiologist to find out even the missed malignant cases. It uses model-based visions in the CAD environment.

In the paper [4], the authors use a binary decision tree classification algorithm for the classification of lesions. Before running the mammograms through a computer algorithm that detects spiculated lesions, they are digitized and categorized. The essential characteristic for the detection of spiculated lesions is the examination of local-oriented edges (ALOE), a textural assessment based on an analysis of the histogram of edge orientations in local windows. The process involves collecting picture measurements (known as features) from each pixel's neighbours, resulting in each pixel growing into a vector of features. These feature vectors are used to train a binary decision tree (BDT) classifier, which assigns a likelihood of finding an anomaly to each pixel of a new mammogram. So, the CAD, as well as its stages, play serious roles. The study found that computer analysis of mammograms can improve radiologist screening effectiveness by a large and statistically significant amount.

In another paper [5], as part of an integrated framework, a new supervised cell-image segmentation technique and a new touching-cell splitting strategy are proposed. The authors isolate the cell areas from the rest of the image in the segmentation stage by categorizing the image pixels as cell or extra-cellular. Using the local Fourier transform, a new colour space called the most discriminant colour space (MDC) is then employed to extract the colour texture at each pixel (LFT). For the splitting portion, they first assess if a connected component of the segmentation map is a touching-cell clump or a single non-touching cell. To distinguish the components, the distance between the most likely radial-symmetry centre and the geometrical centre of the connected component is employed by them. The entire proposed system delivers extremely desired cell-nuclei segmentation and touching-cell splitting outcomes for a variety of applications.

The paper [7] speaks about the classifications made with the help of convolution neural networks and how convolution neural network works effectively in classification methodology.

The automatic detection of breast cancer [6] is explained by means of the Nottingham method. The entire analysis procedure is explained in [8], the steps involved in the analysis and other details are explained clearly. The paper [9] focuses on the histopathological image dataset. It includes histogram equalization, explains various morphological processing and helps to classify the images as benign or malignant.

The paper [11] explains the GLCM feature extraction methodology. This paper explains the computer-aided diagnostic system, and it uses the feature selection algorithm. It states that not only histopathological images but thermal images can also be effective and can help in accurate detection. They include the statistical features for the proper feature extraction and fuzzy classifications. In paper [12], various classification algorithms are explained as fuzzy rule-based classifications. The features are calculated by comparing the left and right breast areas, which are then used to quantify the bilateral differences discovered. After this asymmetry study, the features are fed into a fuzzy classification system. This classifier extracts fuzzy if-then rules based on a training set of known cases. The experiments conducted on almost 150 cases show that the proposed approach is capable of accurately identifying around 80% of cases, a performance comparable to that of other imaging modalities such as mammography.

The paper [14] and [15] focuses on the pixel relationship analysis in detail and improves the efficiency. The proposed three-dimensional (3-D) ultrasound (US) can capture the geometry of a breast tumour while also overcoming the limitations of standard two-dimensional ultrasound. The performance of the pixel relation analysis algorithm with 3-D breast US images compared to 2-D versions of the images is evaluated in this study. The 3-D US imaging is done with a Voluson 530 scanner. The rectangular subpictures of the volume-of-interest (VOI) were manually picked, and the selected VOIs were delineated to include the whole region of the tumour margin.

In another paper [21], Monte Carlo algorithms are utilized to construct simulated clusters of microcalcifications that are placed on normal mammographic backgrounds. This allows for a quantitative evaluation of the computer method's detection accuracy and its link to the microcalcification's physical attributes. The proposed technique achieves a true-positive cluster identification rate of roughly 80% at a false-positive detection rate of one cluster per image. When combined with signal-extraction methods, the results show that a matched filter/contrast-reversal filter or a matched filter/median filter pair may obtain an almost 80% true positive detection rate for microcalcification clusters of moderate subtlety with just one false positive cluster per picture. Monte Carlo simulations of microcalcifications are effective for producing known signals on test pictures and hence for evaluating the performance of image processing systems in detail.

To find breast mass tumour candidates, a marker-controlled watershed segmentation approach is proposed by Samual H. Lewis and Aijuan Dong [30]. The method involved smoothing the images using morphological processes such as opening-by-construction and closing-by-construction, then identifying foreground and background markers, and then utilizing a watershed segmentation technique to isolate a tumour site from its surrounding tissue. Because watershed segmentation is based on pixel density fluctuation, which is evident in all bulk tumours, the proposed approach proved quite successful in recognizing tumours in almost all circumstances. In a study employing data from the Mammographic Image Analysis Society, the total detection percentage for bulk tumours was 90%.

In another paper [40], five feature extraction methods are compared, which are employed for the detection of breast cancer. The Gray-map and the set of fractions under the minimum (SFUM) methods obviously outperform the others, according to the data. Combining a basic feature extraction approach like the Gray-map with Principal Components Analysis appears to be a suitable strategy for this goal. It's also worth noting that, despite being a simpler unsupervised approach, the average fraction under the minimum (AFUM) algorithm performs admirably in this task. The recommended SFUM technique improves the behaviour of the original AFUM.

Comparison of mammogram Vs Thermography: Usually, the mammographic images are done by X-rays, whereas the thermal images are done by infrared rays. It speaks on structural imaging and notifies the suspicious part. It generally focuses on the abnormal part of the body. Thermal imaging mainly focuses on the temperature and how the heat is emitted. It focuses on the changes in the tissues of the breast.

III. GENERAL ARCHITECTURE

The general architecture in Fig. 1. speaks about the major steps involved in the detection and classification of breast cancer. The four major steps involved are Pre-Processing of the input image, Segmentation, Feature Extraction and Image Classification.

A. Pre-Processing of the input image

Pre-Processing is described as a technique for denoising and removing image distortions. The noise might be caused by air or a variety of other factors.

The actions performed on the pictures at the most fundamental level of abstraction, where both the input and output are intensity images, are referred to as pre-processing. It's used to enhance picture data by eliminating unwanted distortions and boosting specific visual features that are important for subsequent processing. The picture histogram is often represented as a matrix of image function values, and these images resemble the sensor's original data (brightness). To achieve this, a variety of noise reduction methods can be used.

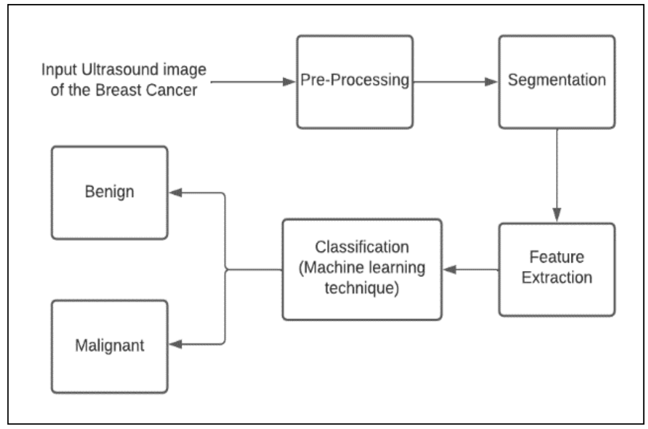

Fig. 1 General architecture of the breast cancer detection and classification system

The nature of apriori information is significant if the goal of pre-processing is to rectify certain image degradation:

- The first group of approaches assumes general degradation properties without knowing the nature of the deterioration. Pre-processing generally is one of the most important steps which could help in removing the noise and also is used to restore any images if necessary.
- A second group requires information on the image capturing device's properties as well as the parameters under which the information was captured. Noise's nature (typically its spectrum properties) is occasionally understood.
- A third technique relies on prior knowledge of the items being searched for in the image, which can greatly simplify the pre-processing. If object knowledge is not accessible ahead of time, it might be approximated during the procedure. It is feasible to use the following strategy:

Initially, the picture is coarsely processed to decrease the amount of data and identify what is needed to build a hypothesis. This idea is then confirmed in a higher-resolution picture. An iterative procedure like this can be performed until the presence of knowledge is confirmed or disapproved. Since segmentation also produces semantic knowledge about objects, this feedback may extend beyond pre-processing. As a result, feedback might begin after the object segmentation.

B. Segmentation

It consists of dividing the entire image into multiple parts. The segmented image attributes can be collected and checked for the required region of interest (ROI).

Segmentation seeks to make a picture simpler to interpret and analyze by decreasing and/or changing its representation. It's often used in images to locate items and borders. Labels are assigned to each pixel in an image such that the pixels having the same labels implies they have comparable features.

Segmentation produces a group of segments or contours that covers the entire image. Each pixel in a region has similar characteristics or attributes such as texture, intensity etc. Adjacent regions have drastically different colours when it comes to considering the same feature. The marching cubes method can be used to generate three-dimensional (3-D) reconstructions from the edges obtained by applying segmentation to a stack of images. This is common in the case of medical imaging.

C. Feature Extraction

Nowadays, data sets have become massive. They have a high number of variables to process, which necessitates a lot of computer power. Feature extraction is used to solve this problem. It is basically a step in the dimensionality reduction process that splits the data into smaller groups. It helps in choosing the best feature from big data sets by combining variables into features, thereby reducing the amount of data required. These features are simple to use while still characterizing the underlying data set precisely.

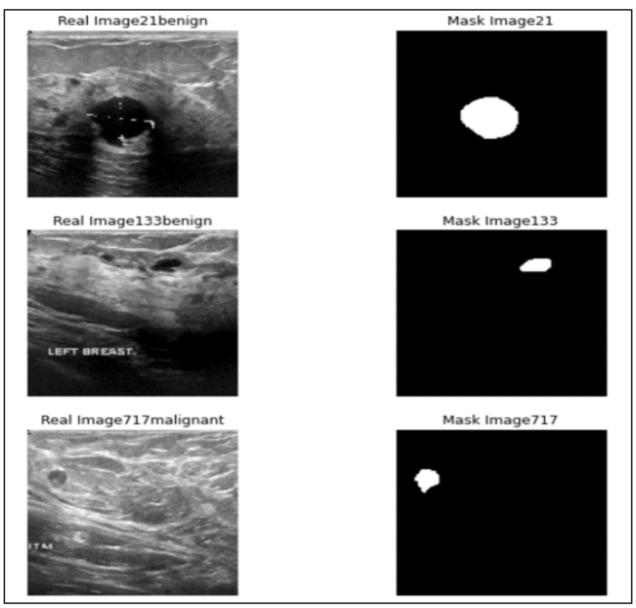

Fig. 2 Image after masking – real image and its corresponding masked image

D. Classification

In classification, a class label is predicted for a given example of input data. Various features are gathered, such as texture, colour etc., as input data for classification, based on which image is accurately classified as benign or malignant. The Support Vector Machine is proposed here to perform the classification task. Its purpose is to find a hyperplane that separates data points in an N-dimensional space, where N is the number of attributes or qualities.

Many hyperplanes are used in SVM to separate and divide the two types of data points. The primary goal is to locate the plane with the greatest margin, which is the distance between the data points from both classes. As the margin increases, the ease to recognize the subsequent data points also increases. Thus, the objective is to maximize this margin distance.

Few of the systems don't use image pre-processing and image segmentation components, and rather they use only the feature extraction techniques as inputs. The computer-aided diagnostic (CAD) systems are very simple and fast in nature, whereas the feature extraction, when directly applied, does not give an accurate or the best performance as expected.

This section describes the general stages involved in cancer detection and categorization. The proposed approach is explored in further depth in the next section.

IV. METHODOLOGY

In this section, the proposed methodology is described. Also, all the steps described in the above section are dealt in here, and how these steps are performed in detecting and classifying cancer in ultrasound images is explained. Median filtering and histogram equalization are used for pre-processing. For segmentation, a hybrid technique is utilized, and for feature extraction, fundamental methods are applied. For classification, the Support Vector Machine (SVM) is suggested and used. SVM's accuracy is then compared to that of other machine learning approaches, including boosted tree (BT), random forest (RF), naive Bayes (NB), and convolutional neural networks (CNN). Fig. 3. depicts the Methodology workflow.

The dataset used consists of 1578 ultrasound images of the breast, out of which 891 images consists of benign cancer, 266 are normal, and 421 consists of malignant cancer. The dataset has been taken from Kaggle. Since the number of images in the dataset is less, data augmentation is used - to produce more ultrasound images, to solve the data imbalance issue and to do regularization as it helps in preventing overfitting to some extent.

A. Pre-Processing

The image enhancement helps to get the fine details of an image by means of various noise reduction algorithms. It improves the quality of the image with low contrast. In general, filters are used to enhance image quality. Filters are usually applied with a mask or kernel. A median filter and histogram equalization are used for pre-processing.

The input picture is filtered with the median filter. The median filter [15,16,17,18,19] is one of the widely used filters which is applied for the treatment of salt and pepper noise. It smoothens the image. The median is computed by arranging all of the pixels in ascending order and then substituting the element's middle value with it. This technique is preceded till the end. And it is repeated again and again until the end is reached. We measure it with the metrics which are discussed in Table I. It measures the various metrics of the pre-processing function and thereby gives the various values and measurements to it.

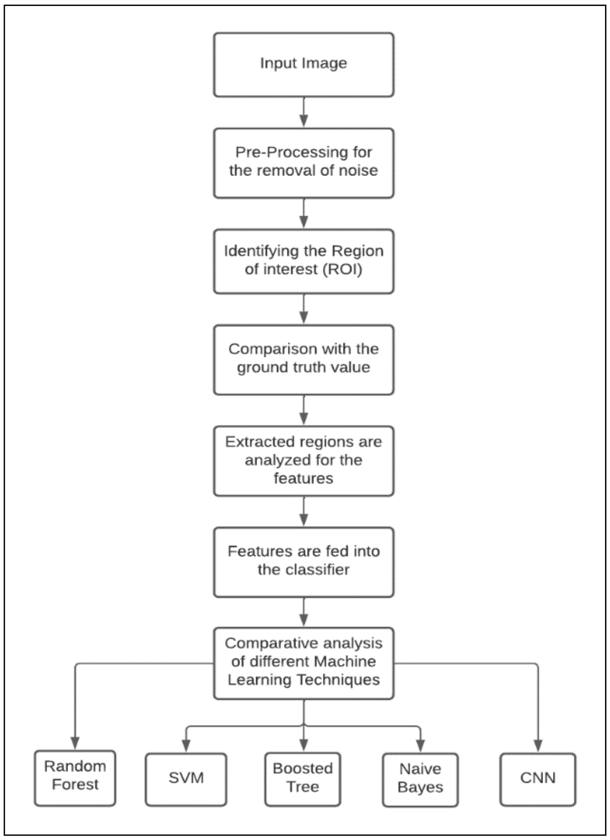

Fig. 3 Methodology workflow

The output of the median filter is fed for the histogram equalization methodology [20, 21] to get the uniform histogram, i.e., the median filtered image is equalized by the histogram equalization. The process helps to find if the image is free from noise. They are arranged accordingly to the number of pixels allotted. Then the corresponding running sum is obtained. Then the total number of values are divided by the levels of pixel intensities and is multiplied by its highest value to equalize the image and to obtain the best result for the images fed into.

Various metrics are used to calculate and assess the results obtained. The results obtained are further evaluated by comparing it with various other algorithms and are proved to be quite efficient comparatively. The evaluation metrics of pre-processing are [22, 23, 24, 25] used, and the evaluation parameters include the PSNR, MSE and the SNR as their prime components. The pre-processing gives the best of the various results, which could act as a preliminary step for the other steps. The better the result pre-processing step brings, the better the final classification result we shall expect. The major steps will obtain noise-free images. The pre-processing factors are clearly expressed in Table 1. which explains the various metrics of the pre-processing. The other important function of pre-processing is to restore images if lost, so even if the restoration process also happens with the effective replacement of the images, the pre-processing plays a crucial role.

Table 1. Metrics for pre-processing

Input image	PSNR	SNR	MSE
1	41.38	37.87	36.91
2	44	39.96	16.83
3	33.02	28.36	55.56
4	29.05	22.67	85.63
5	41.43	38.67	73.86

B. Segmentation

This defines splitting the image into multiple regions. The filtered image is now given to the segmentation phase as input [26, 27, 28, 9, 30]. After applying the filters, the region of interest is found by means of the segmentation methods like the Marker Controlled Watershed Algorithm and the J Segmentation (JSEG) algorithm. Colour quantization and spatial segmentation are now segregated from the output [34, 35]. The foremost step helps us to differentiate the various regions in the image. The next step, using the class maps, helps us to differentiate the regions and the border images [36, 37]. The region-growing technique is used to expand the area in which the appropriate content may be found. The colour space quantization act as an input which is step 1, and then step 2 includes spatial segmentation, which includes J image and region growing [41,42, 43]. Finally, the similar coloured regions merge together. The marker control algorithm takes the filtered image as an input, reads the image and performs its corresponding gradient magnitude so that the borders appear sharp, comparatively to the region. [31, 32, 33].

The J calculations are done as follows:

$$k_{i\,=} \quad \frac{1}{L_i}\sum_{n=y_i} Y$$

$$R_t \;=\; \sum_{n=y} \mathrm{mod}\,(Y-K)^2$$

Where R stands for the total variance in an image and K is mean to be calculated for the region.

$$R_w = \sum_{i=1}^{c} R_i$$

where,

$$R_i = \sum_{i=1}^{a} \sum_{y=Y} \mathrm{mod}\,(Y - K_i)^2$$

$$d = \frac{R_b}{R_i}$$

Now the watershed transformation is applied to the image with gradients. The foreground regions are marked by means of the morphological functions, namely opening and closing, by the reconstruction operations. Now similar to the foreground, the background regions are obtained. Gradients are modified, and the results are obtained. The segmentation is compared with the ground truth dataset and is evaluated in Table 2. by using the Jaccard coefficient and the Dice coefficient. These act as metrics for the segmentation algorithm. The Pre-Processed image is applied with various hybrid algorithms where the Dice coefficient and the Jaccard index are calculated.

Table 2. Jaccard coefficient

Algorithm	**Accuracy**	**Jaccard Index**	**Dice Coefficient**
Proposed (SVM)	93.4	93.76	93.56

C. Feature extraction and classification

It consists of extracting various components in an image, for example, colour, texture, size, shape, mean, variance etc. So, we could include entropy, standard deviation, dissimilarity, homogeneity, number of pixels [38, 39, 40], autocorrelation, contrast, variance, mean, shape, and margin.

Equation 1 and 2 shows a sample of how the calculations for the features are made, and using these, the values in the tables are calculated accordingly.

$$\text{Homogeneity} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \{F(x, y)\}^2$$

Eq. 1. Homogeneity equation

$$\text{Entropy} = -\sum_{x=1}^{N-1} \sum_{y=1}^{N-1} f(x, y) \times \log(f(x, y)$$

Eq. 2. Entropy equation

The extracted features are explained and are tabulated in Table 3. The collected features are then fed into the classifier. The process of classification helps in classifying the images into benign or malignant images. The metrics of sensitivity and selectivity are tabulated with the help of equations 3, 4 and 5. They calculate the accuracy and other following factors based on the tabulated values from the feature extraction and are fed into the classifier. Here we use an SVM classifier that includes ten-fold cross-validation, which improves to be 93.4%.

Table 3. Feature extraction values

Features	**Image 1**	**Image2**
Entropy	0.00245	0.00231
Standard Deviation	0.4145	0.39876
Dissimilarity	0.02341	0.02981
Homogeneity	0.99871	0.99876
Autocorrelation	67.876	64.9876
Contrast	0.123	0.1
Variation	0.1321	0.234
Mean	0.6789	0.5678

The values of Sensitivity and Specificity for the SVM Classifier are calculated and are tabulated in Table 4. Fig. 4. explains the Receiver Operating characteristic curve (ROC) for the retrieved characteristics that are given into the Support Vector Machine classifier. From the ROC curve, it is observed that the improvement could be made in the proposed methodology by increasing the number of dataset elements. The accuracy rate still is effective.

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

Eq. 3

$$\text{Specificity} = \frac{TN}{TN+TP}$$

Eq. 4

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

Eq. 5

Where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.

Table 4. Classification Values

Classifier	**Sensitivity**	**Specificity**
SVM	82.64%	91.09%

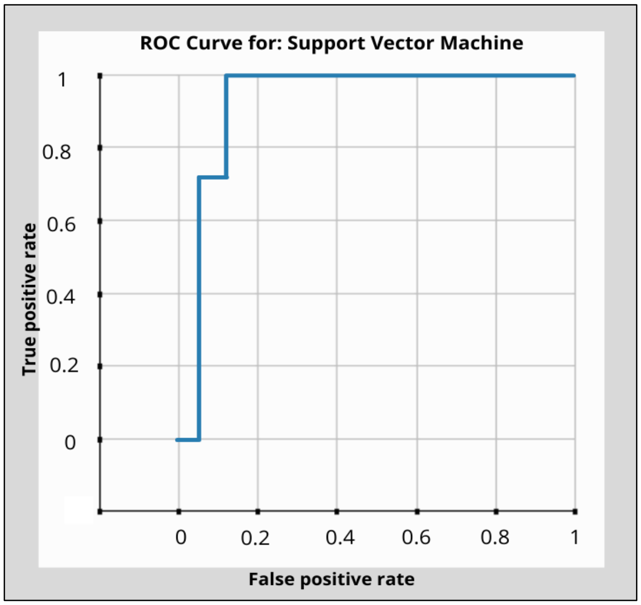

Fig. 4 ROC curve for SVM

The summary of the CNN model is shown in Fig. 5. Table 5. represents the accuracy metric for each classifier. Fig. 6. makes an overall comparison of the classifiers by feeding appropriate features to the corresponding classifiers. The SVM shows a better accuracy of detection of cancer in the testing dataset of ultrasound images when compared to the other classifiers.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 126, 126, 32)      320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 63, 63, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 61, 61, 32)        9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 30, 30, 32)        0
_________________________________________________________________
dropout (Dropout)            (None, 30, 30, 32)        0
_________________________________________________________________
flatten (Flatten)            (None, 28800)             0
_________________________________________________________________
dense (Dense)                (None, 128)               3686528
_________________________________________________________________
dense_1 (Dense)              (None, 3)                 387
=================================================================
Total params: 3,696,483
Trainable params: 3,696,483
Non-trainable params: 0
_________________________________________________________________
```

Fig. 5 Summary of CNN Model

Table 5. Accuracy metric for different classifiers

Classifier	Accuracy (On testing dataset)
Random Forest	73.1%
SVM	93.4%
Boosted Tree	82.4%
Naive Bayes	86.2%
CNN	76.8%

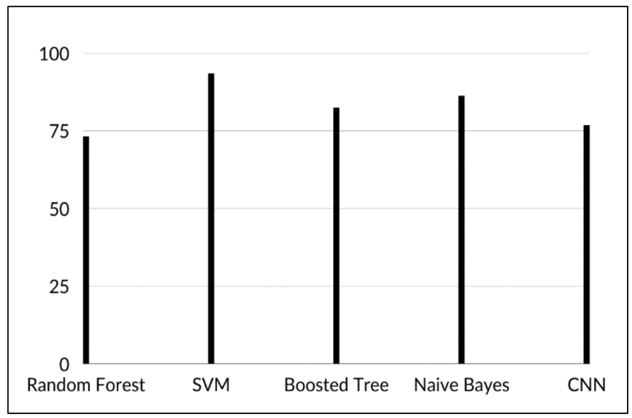

Fig. 6 Accuracy (Y) vs Classifier (X) Bar chart

D. Results and Discussion

The algorithm in the proposed methodology gives an improved accuracy. The algorithm thus helps in the earlier detection, which could be a life saving one. Noise removal using the filtering methods helps in removing the noise present in the image. It also helps in the segmentation methodology by identifying the proper region of interest and collecting the features from the segmented image. These features are then fed into the classifier. Based on the features identified, the classifier classifies the images as cancerous or non-cancerous. The median filtered image is equalized by the histogram equalization. The process helps us to find if the image is an error-free image from noise. The pre-processing factors and various other metrics of the pre-processing are also discussed in the tables above.

V. CONCLUSION AND FUTURE WORK

To conclude, the system receives a digital ultrasound image as an input. Support vector machine is used to classify the images into benign and malignant, which includes ten-fold cross-validation and gives an accuracy of 93.4%. It gets the highest accuracy when compared to the other popular classifiers, as discussed. It clearly shows that it is not always that neural networks will yield to best results. Also, considering the complexity and computation constraints of the neural networks, using SVM is a better choice. Thus, the work presented in this paper is quite satisfactory. Still, there are places we could focus on better improvement.

The future work will focus on a large dataset of ultrasound images and can be fully automated, which could lead to an increase in the accuracy, and the present work shows comparatively a better result with existing segmentations that were carried out, considering it to be partially automated. Future work could focus on the improved segmentation algorithm so that the classifier could yield a better result. The semi-automated system here can be converted to a fully automated system to reduce the time limit and to improve its efficiency as well.

Thus, this paper presents a comparative analysis and will serve as a base paper for future studies in the field of detection and classification of breast cancer.

REFERENCES

[1] J. H. Tanne, Everything you need to know about breast cancer...but were afraid to ask, New York [GNYC], 26 (1993) 52–62.

[2] R. A. Smith, Epidemiology of breast cancer, RSNA Categorical Course Phys., (1994) 21–33.

[3] E. Zelnio, ATR legacy, presented at MSTAR Program Initiation Meeting, Wright Laboratory, Wright-Patterson AFB OH, Sept. (1995).

[4] W. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, Computer-aided mammographic screening for spiculated lesions, Radiol., (1991) 331–337.

[5] Hui Kong, Metin Gurcan, and Kamel Belkacem-Boussaid, Partitioning Histopathological Images: An Integrated Framework for Supervised Color-Texture Segmentation and Cell Splitting, IEEE Transactions On Medical Imaging, 30(9) (2011) 1661-77

[6] Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L., Breast cancer histopathological image classification using convolutional neural networks. In 2016 joint international conference on neural networks (IJCNN), (2016) 2560-2567.IEEE.

[7] Dalle, J. R., Leow, W. K., Racoceanu, D., Tutac, A. E., & Putti, T. C., Automatic breast cancer grading of histopathological images. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2008) 3052-3055. IEEE.

[8] Veta, M., Pluim, J. P., Van Diest, P. J., & Viergever, M. A., Breast cancer histopathology image analysis: A review. IEEE Transactions on Biomedical Engineering, 61(5) (2014) 1400-1411.

[9] Singh, S., Gupta, P. R., & Sharma, M. K., Breast cancer detection and classification of histopathological images. International Journal of Engineering Science and Technology, 3(5) (2010) 4228.

[10] Nithya, R., & Santhi, B., Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast cancer. International journal of computer applications, 28(6) (2011) 21-25.

[11] Cheng, H. D., Shan, J., Ju, W., Guo, Y., & Zhang, L., Automated breast cancer detection and classification using ultrasound images: A survey. Pattern recognition, 43(1) 299-317.

[12] Schaefer, G., Závišek, M., & Nakashima, T., Thermography based breast cancer analysis using statistical features and fuzzy classification. Pattern Recognition, 42(6) (2009) 1133-1137.

[13] Wolberg, W. H., Street, W. N., & Mangasarian, O. L., Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine-needle aspirates. Cancer letters, 77(2-3) (1994) 163-171.

[14] Bassett, L. W., & Kimme-Smith, C., Breast sonography. AJR. American journal of roentgenology, 156(3) (1991) 449-455.

[15] Chen, W. M., Chang, R. F., Moon, W. K., & Chen, D. R., Breast cancer diagnosis using three-dimensional ultrasound and pixel relation analysis. Ultrasound in medicine & biology, 29(7) (2003) 1027-1035.

[16] Das, S., & Mohan, A., Medical image enhancement techniques by the bottom hat and median filtering. Int J Electron Commun Comput Eng, 5 (2014) 347-351.

[17] Schmitt, R. M., Meyer, C. R., Carson, P. L., & Samuels, B. I., Median and spatial low-pass filtering in ultrasonic computed tomography. Medical Physics, 11(6) (1984) 767-771.

[18] Arastehfar, S., Pouyan, A. A., & Jalalian, A., An enhanced median filter for removing noise from MR images. Journal of Artificial Intelligence & Data Mining, 1(1) (2013) 13-17.

[19] Cheng, H. D., Shan, J., Ju, W., Guo, Y., & Zhang, L., Automated breast cancer detection and classification using ultrasound images: A survey. Pattern recognition, 43(1) (2010) 299-317.

[20] Thangavel, K., & Karnan, M., Computer-aided diagnosis in digital mammograms: detection of microcalcifications by metaheuristic algorithms. VIP Journal, 5(7) (2005) 41-55.

[21] Chan, H. P., Doi, K., Galhotra, S., Vyborny, C. J., MacMahon, H., & Jokich, P. M., Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. Medical Physics, 14(4) (1987) 538-548.

[22] Sundaram, K. M., Sasikala, D., & Rani, P. A., A study on pre-processing a mammogram image using Adaptive Median Filter. International Journal of Innovative Research in Science, Engineering and Technology, 3(3) (2014) 10333-10337.

[23] Kumar, P., & Vijayakumar, B., Brain tumour Mr image segmentation and classification using PCA and RBF kernel-based support vector machine. Middle-East Journal of Scientific Research, 23(9) (2015) 106-2116.

[24] Anitha, V., & Murugavalli, S., Brain tumour classification using a two-tier classifier with adaptive segmentation technique. IET computer vision, 10(1) (2016) 9-17.

[25] Viji, K. A., & Jayakumari, J., Performance evaluation of standard image segmentation methods and clustering algorithms for segmentation of MRI brain tumour images. European Journal of Scientific Research, 79(2) (2012) 166-179.

[26] Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B., Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clinical imaging, 37(3) (2013) 420-426.

[27] Roy, S., Bhattacharyya, D., Bandyopadhyay, S. K., & Kim, T. H. ., Artefacts and skull stripping: an application towards the pre-processing for brain abnormalities detection from MRI. Int J Control Autom SERSC, 10(4) (2017) 147-160.

[28] Ancy, C. A., & Nair, L. S., An efficient CAD for detection of tumours in mammograms using SVM. In 2017 International Conference on Communication and Signal Processing (ICCSP), (2017) 1431-1435, IEEE.

[29] Sandabad, S., Benba, A., Tahri, Y. S., & Hammouch, A., Novel extraction and tumour detection method using histogram study and SVM classification. International Journal of Signal and Imaging Systems Engineering, 9(4-5) (2016) 202-208.

[30] Lewis, S. H., & Dong, A., Detection of breast tumour candidates using marker-controlled watershed segmentation and morphological analysis. In 2012 IEEE Southwest Symposium on Image Analysis and Interpretation , (2012) 1-4. IEEE.

[31] Filipczuk, P., Kowal, M., & Obuchowicz, A., Multi-label fast marching and seeded watershed segmentation methods for diagnosis of breast cancer cytology. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (2013) 7368-7371.IEEE.

[32] Hefnawy, A., An improved approach for breast cancer detection in mammograms based on watershed segmentation. International Journal of Computer Applications, 75(15) (2013).

[33] Al-Tarawneh, M. S., Lung cancer detection using image processing techniques. Leonardo Electronic Journal of Practices and Technologies, 11(21) (2012) 147-58.

[34] Jaffery, Z. A., & Singh, L., Performance analysis of image segmentation methods for the detection of masses in mammograms. International Journal of Computer Applications, 82(2) (2013).

[35] Victor, A., Roselin, J., & Kavitha, V., Preferential image segmentation using j segmentation based on colour, shape and texture. International Journal of Engineering and Technology, 2(2) (2010) 131-135.

[36] Zhen Yu Chan., JSEG - Unsupervised Segmentation of Color-Texture Regions in Images (www.mathworks.com/matlabcentral/fileexchange/64123-jseg-unsupervised-segmentation-of-color-texture-regions-in-images), MATLAB Central File Exchange. Retrieved February 27, (2020).

[37] Lugli, Luciano & Tronco, Mario & Porto, Vieira., JSEG Algorithm and Statistical ANN Image Segmentation Techniques for Natural Scenes. 10.5772/14622, (2011).

[38] Verma, B., & Zakos, J., A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. IEEE transactions on information technology in biomedicine, 5(1) (2001) 46-54.

[39] Wroblewska, A., Boninski, P., Przelaskowski, A., & Kazubek, M., Segmentation and feature extraction for reliable classification of microcalcifications in digital mammograms. Optoelectronics Review, (3) (2003) 227-236.

[40] Llobet, R., Paredes, R., & Pérez-Cortés, J. C., Comparison of feature extraction methods for breast cancer detection. In Iberian Conference on Pattern Recognition and Image Analysis, Springer, Berlin, Heidelberg, (2005) 495-502.

[41] Victor, A., & Ghalib, M. R., Detection of Skin Cancer Cells-A Review. Research Journal of Pharmacy and Technology, 10(11) (2017) 4093-4098.

[42] Victor, A., & Ghalib, M., Automatic detection and classification of skin cancer. International Journal of Intelligent Engineering and Systems, 10(3) (2017) 444-451.

[43] Victor, A., & Ghalib, M. R., A hybrid segmentation approach for detection and classification of skin cancer, (2017).

[44] Azadeh NH, Adel A, Afsaneh NH. Comparing the performance of various filters on skin cancer images. International Conference on Robot PRIDE 2013-2014- Medical and Rehabilitation Robotics and Instrumentation 2013-2014 Procedia. Computer Science, 42 (2014) 32-37.

[45] Abdul JJ, Sibi S, Aswin RB. Artificial neural network-based detection of skin cancer. Int J Adv Res Electr Electron Instr Eng, 1 (2012) 200-205.

[46] Mariam A, Mai SM, Amr S. Automatic detection of melanoma skin cancer using texture analysis. Int J Comp Appl, 42 (2012).

[47] Silveira M, Nascimento JC, Marques JS. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. IEEE J Signal Proc, 3 (2009) 35-45.

[48] Chiem A, Al-Jumaily A, Khushaba RN. A novel hybrid system for skin lesion detection. Proceedings of the 3rd International Conference on Intelligent Sensors, Sensor Networks, and Information Processing, (2007) 567-572.

[49] Emre Celebi M, Hassan AK, Bakhtiyar U. A methodological approach to the classification of dermoscopy images. Comp Med Imag Graph , 31 (2007) 362-373.

[50] J.C. Perez-Cortes, J. Arlandis, and R. Llobet. Fast and accurate handwritten character recognition using approximate nearest neighbours search on large databases. In Workshop on Statistical Pattern Recognition SPR-2000, volume 1876 of Lecture Notes in Artificial Intelligence, (2000) 767–776, Alicante (Spain).

[51] J. Suckling and J. Parker et al. The mammographic images analysis society digital mammogram database. In Excerpta Medica. International Congress Series, 1069 (1994) 375–378.

[52] G.M. te Brake and N. Karssemeijer. Automated detection of breast carcinomas that were not detected in a screening program. Radiology, 207 (1998) 465–471.

[53] M. Wallis and M. Walsh et al. A review of false-negative mammography in an asymptomatic population. Clin Radiol, 44 (1991) 13–15.

[54] Nidhi Mongoriya, Vinod Patel., Review The Breast Cancer Detection Technique Using Hybrid Machine Learning. SSRG International Journal of Computer Science and Engineering 8(6) (2021) 5-8.

[55] Brotobor Deliverance, Edeawe Isaac Osahogie, Brotobor Onoriode.,Awareness And Practice of Student Nurses on Breast Self Examination: A Risk Assessment Tool For Breast Cancer. SSRG International Journal of Nursing and Health Science 6(1) (2020) 66-69.

[56] Dharampal Singh., Review on Breast Cancer. SSRG International Journal of Humanities and Social Science 7(4) (2020) 31-37. 10.14445/23942703/IJHSS-V7I4P106

[57] B.Johnson, P.Keerthi Vasan, V.Thillaivendan., Performance of Hyperthermia for Breast Cancer. SSRG International Journal of Applied Physics 3(2) (2016) 1-5.

[58] Mariam Shadan, Nazoora Khan, Mohammad Amanullah Khan, Hena Ansari and Sufian Zaheer., Histological categorization of stromal desmoplasia in breast cancer and its diagnostic and prognostic utility. SSRG International Journal of Medical Science 4(6) (2017) 18-11.