Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 10, No. 3, October : 2023

INTELLIGENT BREAST CANCER DIAGNOSIS: A MULTILAYERED ARTIFICIAL INTELLIGENCE APPROACH FOR ENHANCED ACCURACY

Nagavarapu Sowmya[1], Kanigiri Kumari [2], Chilukuri Anusha[3], G Shivakumari[4], Jogu Ramesh[5]
[1,2,3,4,5] Assistant professor, Department of Electronics and Communication Engineering, St.Martin's Engineering College, Dhulapally, Secunderabad, Telangana, India.
nsowmyaece@smec.ac.in[1], kkumariece@smec.ac.in[2], chanushaece@smec.ac.in [3], gshivakumari@smec.ac.in [4], jrameshece@smec.ac.in [5]

Abstract
Breast cancer is a significant health concern for women, ranking as the second leading cause of death among all cancers affecting females. Early detection is crucial in reducing mortality rates. Mammography is a widely used radiological screening procedure for breast cancer when there are no apparent symptoms. Some machine learning techniques, such as Support Vector Machine (SVM), have been employed in state-of-the-art approaches for breast cancer detection and classification. However, SVM has shown limitations in achieving accurate detection and classification results. To enhance the precision of breast cancer detection and characterization, this article presents a novel approach. It utilizes the Nonsubsampled Contourlet Transform (NSCT) in combination with Adiabatic Probabilistic Fuzzy C-means (APFCM) clustering. Additionally, a Spatially Dependency Matrix (SDM) is employed to extract features from segmented breast cancer regions. Subsequently, a Back-Propagated Artificial Neural Network (BP-ANN), a concept from artificial intelligence, is employed for disease type classification. The proposed method outperforms conventional approaches across various evaluation metrics, showcasing its superior qualitative performance in breast cancer detection and classification.

Keywords: Mammography, breast cancer, image segmentation, NSCT, APFCM, SDM, SVM and BP-ANN.

1. Introduction

Breast cancer is one of the leading diseases that reflect an uncontrolled growth of abnormal cells in the breast. Due to the breast anomalies properties and the nature of the human visual perception, it is natural that, sometimes the abnormalities are missed or miss classified. As a result, unnecessary biopsies are taken. To mitigate this problem, a computer aided diagnosis (CAD) system [1-2] has emerged. The proposed CAD system is implemented as an integrated system using image processing techniques and machine learning algorithms. CAD aims at the detection and localization of abnormalities at an early phase, which avoids the further spread of the abnormality. Breast cancer [3-4] is one of the leading diseases that reflect an uncontrolled growth of abnormal cells in the breast. Due to the breast anomalies properties and the nature of the human visual perception, it is natural that, sometimes the abnormalities are missed or miss classified. As a result, unnecessary biopsies are taken. In breast, normal cells [5-6] grow and divide at a particular time but in case of cancerous cells, the cell growth is continuous and uncontrolled. Many researchers have addressed the issue of breast cancer detection from the past few years' later classification approaches also discussed and presented by several authors. A new breast cancer diagnostic system by employing PSO-SVM framework is presented in [7], where the PSO aimed at mitigating the simplification ability of the SVM classification by concurrently tackle the essential kernel constraint set and recognizes the majority discriminative characteristic feature separation. In scheming categorization accurateness, the object utility quantity of SVs and quantity of characteristic features are concurrently followed as deliberation. Principally, during a sequence of observed experiments on standard database, PSO-SVM organization not merely exploits the simplification presentation but too chooses the majority revealing characteristic features. Author in [8] reviewed supervised deep learning (SDL) area of research, conceptual groups and analyzes various techniques. They have proposed two unconventional combinations; the primary is founded on SDL representation mechanism utilized for feature drawing out, while the subsequent utilizes the imperative drawing out method. The analysis is followed by a comparative evaluation of the algorithms are relative performance as measured by several metrics. They have concluded by highlighting the potential research directions, such as the need for rule extraction methods. The recognition scheme for classification of tumor lesions appearing in mammographic X-ray images are addressed in [9]. Genetic algorithm (GA) utilized FSS is defiant from clamor up to a definite stage and categorization rate is enhanced for GA used FSS method. FCM has separation the huge amount contour group clusters such that the level of alliance is burly for the features inside the similar groups and weedy for the features in dissimilar groups. GA explored the important contour features by concerning the magnificence of usual dispute. Utilizing three operatives like imitation, crossover and mutation, GA is able to choose important feature division. However, due to lack of accurate detection, efficient extraction of features and classification accuracy, conventional breast cancer diagnosis systems failed to produce an acceptable outcome.

Most of the research work is focused on using optimization techniques to develop a Classification [10] and Diagnosis [11] of breast cancer from Mammographic images. The detection & classification of irregularities in Mammographic images are considered for investigation in this paper. Poor noise-to-signal ratio is a drawback in Mammographic images. The anatomically distinct structures are often seen with a very low contrast. Reliable standard image processing technique [12] is needed for its computation. Modification in image content is done in a highly controlled and reliable way without any compromise in clinical decision-making, but the presence of artifacts leads to 10 – 25% of tumors being missed by radiologists. Basic noise removal filters [13] cannot be applied on Mammographic images as they are not able to remove the artifacts effectively. If we use those fundamental filters then, image get corrupted and enhancement operation will not work. Image denoising is one of the significant topics in image enhancement [14] that deals with noise contained imagery, which are need to be preprocessed using various approaches. Nowadays, medical imaging field and its equipment's are improved noticeably. The existing mammographic image segmentation approaches failed to perform well in terms of sensitivity, false positive rate, accuracy, specificity and improved classification when processing the images generated from the advanced image generation sources. In order to overcome this issue, various classification and diagnosis of breast cancer approaches [15-18] are presented to process the images generated from mammographic images.

To achieve this research goal, specific objectives are set. The research objectives of this paper comprise of the following components.

- Implementation of NSCT to mitigate the artifacts and any noise components present in the mammographic X-ray images during image acquisition.
- Next, APFCM is employed for detection of breast cancer effectively with exact region of interest (ROI) extraction and feature extraction is done by utilizing SDM approach.
- Finally, BP-ANN is applied to classify whether the patient is normal or abnormal then form abnormal the type of cancer is classified as benign or malignant.

The rest of the paper is planned as follows, Section 2 deals about the detailed architecture of proposed methodology with its block wise operation. Section 3 deals with results and discussion with comparison to the state of art methods using quantitative metrics. Section 4 deals with the conclusion and future enhancements of proposed methodology followed by bibliography.

2. PROPOSED METHOD

The proposed method for breast cancer detection and classification shown in Figure 2, initially query image applied from image acquisition unit, and then it is applied to preprocessing stage. Here, by using NSCT to remove the artifacts and noises and performs the image enhancement. Then APFCM clustering applied for breast cancer detection and effective ROI extraction. Then by using the SDM feature matrix to achieve the features and create the database using features. Then by applying the BP-ANN classification methodology to detect the normal and abnormal stage of cancer, at the same time type of classification also recognized.

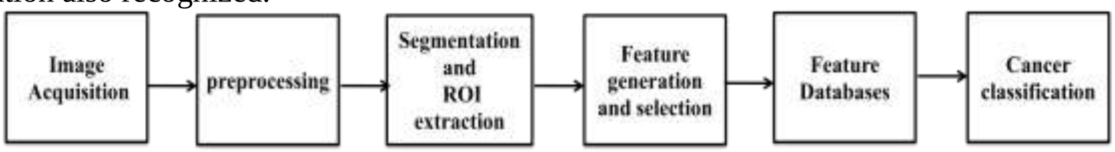

Figure 2. Proposed methodology

3.1Preprocessing

The query image is acquired from the image acquisition step, which includes background information and noise. Pre-processing is required and necessary to remove the above-mentioned unwanted portions. The pre-processing stage is mainly used for eliminating the irrelevant information such as unwanted background part, which includes noises, labels, tape and artifacts and the pectoral muscle (present in the left or right upper corner) from the breast image. The different types of noise occurred in the mammogram images are salt and pepper, Gaussian, and speckle and Poisson noise. When noise occurs in an image, the pixels in the image show different intensity values instead of true pixel values. So, by choosing the perfect method in the first stage of preprocessing, this noise removal operation will perform effectively. Figure 3 shows NSCT transform which is an extension of wavelet transform and useful to perform multidirectional and multidimensional analysis. Reduction of the noise mostly and avoiding the introduction visual artifacts by the analysis of pixels at various scales, NSCT denoising efforts to eradicate the noise presented in the pixel, as it conserves the image uniqueness, despite of its pixel satisfied. The excellence of the mammographic representation has a straight manipulate on analytic and management procedure. Proficient denoising method is complete to evade information sleaze. NSCT divides the image into multiple bands using high pass and low pass filters on rows and columns parallel to the non-sub sampled laplasian pyramid. In this research, three levels of pyramids are applied to remove the noise effectively and parallel enhance the image without loosing the original properties of image, because of the multidirectional and multidimensional nature of different cancer signs. If the input medical data is of low contrast, then this scheme provides an enhanced image. This proposed algorithm avoids the generation of annoying artifacts.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 10, No. 3, October : 2023

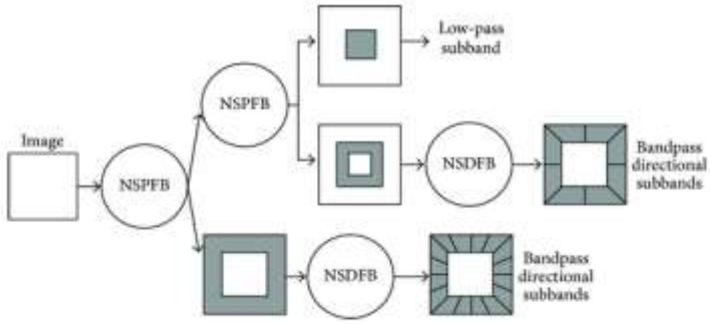

Figure 3. NSCT transform.

3.2 Segmentation and ROI extraction

In the second stage, automatic detection of suspicious patterns and classify the abnormalities using the APFCM. Image segmentation is the process of dividing the image into multiple clusters based on the region of interest presented to detect the mammographic tumor. Image segmentation is widely used in remote sensing, medical imaging, etc. Regions of interest are portions of breast images, which are used by radiologists to detect abnormalities like micro classifications (benign and malignant) and masses (benign and malignant). The APFCM is used in the proposed procedure for segmentation to a certain extent than FCM clustering approach because of its speed of operation while maintaining the highest accuracy. The APFCM procedure combines the properties of jointly possibility and FCM approaches as shown in figure 4. Here the membership functions are generated in a probability-based manner to get better detection. Among those detected tumors, the highest accurate tumor is considered as ROI. The automatic extraction of ROI is difficult. So, ROIs are obtained through possibility cropping, which are based on location of abnormality of original database mammogram images(1024×1024). Here the membership functions are generated in a probability-based manner to get better detection. Among those detected tumors, the highest accurate tumor is considered as ROI. The automatic extraction of ROI is difficult. So, ROIs are obtained through possibility cropping, which are based on location of abnormality of original database mammogram images(1024×1024).

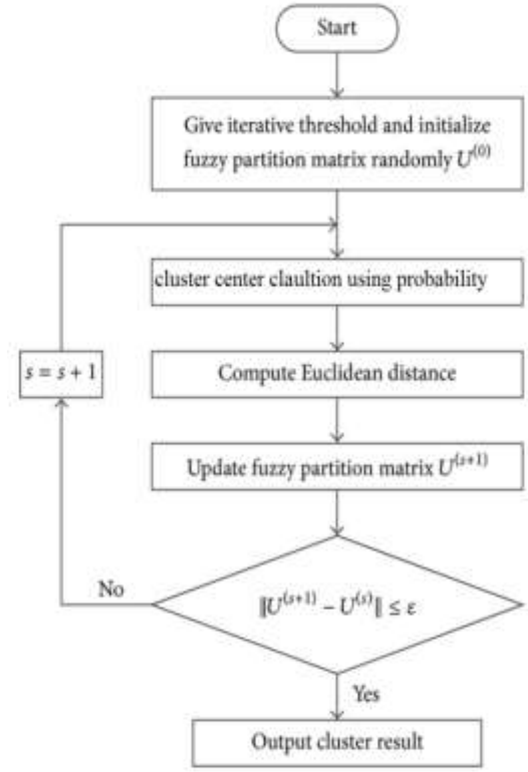

Figure 4: Adiabatic probabilistic fuzzy C-means

3.3 Feature generation and selection

Feature extraction is nothing but the process of mapping the huge information data space into a reduced feature space. The principle behind the feature extraction is that all the calculations are performed in an easiest way by using the reduced feature space. Normally, feature extraction procedure consists of three major steps. The steps are the construction of possible features using the various transformation techniques, the selection of the optimal features from the possible features which leads to better performance of the system and finally, it matches the features using various classifiers for recognizing the objects. Features or descriptors are called as the original measurement variable function, which are mainly used for the classification and pattern recognition task. The various characteristics or properties of the features are that they are robust, reliable, and independent. Feature extraction is nothing but extracting the image properties or information from the images. The objective of feature extraction is to increase the performance of the classification and prediction problem. The most important regional descriptor is texture which is extracted using SDM. Image consistency is a significant exterior feature-based attributes utilized to recognize tumors. Texture is derived as a configuration which is composed of many ordered similar patterns or structures. Based on the number of pixels used to describe the texture feature, statistical texture features are further classified as the first order statistical texture features, the second order statistical texture features and the higher order statistical texture features. In first order statistical texture features, statistical measures are calculated by using an individual pixel of an image. In case of second order statistical texture features, statistical measures are calculated by consider the relationship between neighboring pixels. Intensity histogram and intensity features are the first order statistical texture features. Hence, the shearlet moments of 1st, 2nd, 3rdand 4th order are used as features. In this research work, the texture regional descriptors are extracted by using the statistical approach. In this work, the shearlet moments of 1st, 2nd, 3rd and 4th order features are extracted from the decomposed

ROI image and the definitions of these features. Once, the wavelet, curvelet and shearlet coefficients are generated, the mean, variance, skewness and kurtosis features are extracted from the shearlet coefficients of ROI image, curvelet coefficients of ROI image and wavelet coefficients of ROI image. Similarly, the above said features are extracted for all the training images of normal or abnormal as benign and malignant types. The output of the feature extraction phase has two different feature databases.

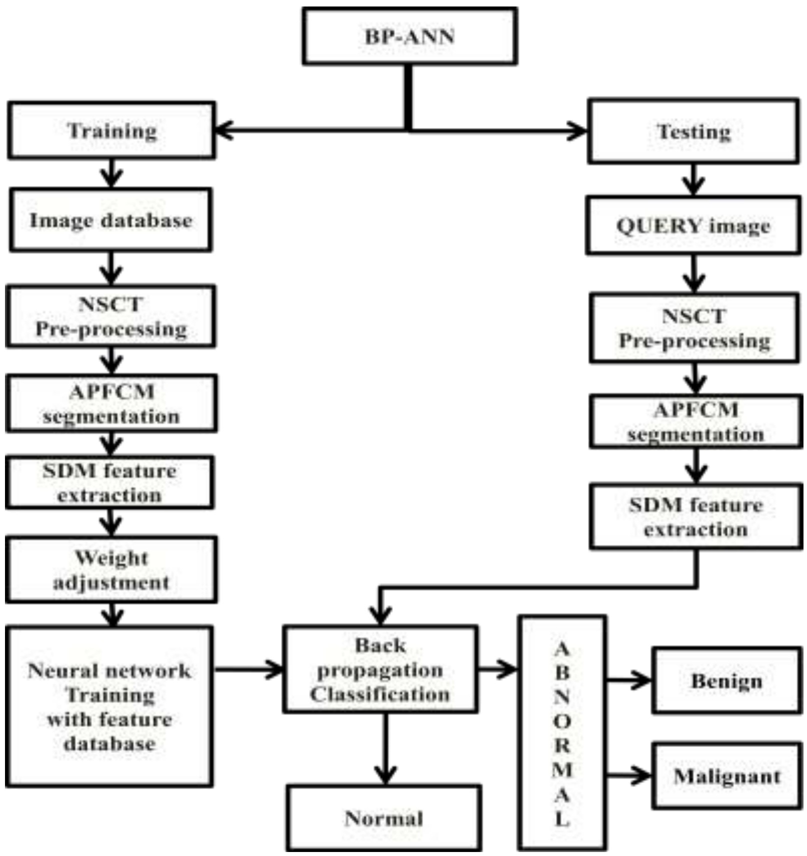

Figure 5: Training and testing process of proposed hybrid BP-ANN model

3.4 Image classification
Figure 5 disclose the complete architecture of proposed hybrid BP-ANN model for breast cancer detection and classification. Neural networks have been effectively applied across a range of problem domains like finance, medicine, engineering, geology, physics and biology. From a statistical viewpoint, neural networks are interesting because of their potential use in prediction and classification problems. BP-ANN is a method developed using emulation of birth neural scheme. The neurons are connected in the predefined architecture for effectively performing the classification operation. Depending on the SDM features, the weights of the neurons are created. Then, the relationships between weights are identified using its characteristic features. The quantity of weights decides the levels of layers for the proposed network. Figure 6 represents the architecture of artificial neural networks. BP-ANN basically consists of two stages for classification such as training and testing. The process of training will be performed based on the layer based architecture. The input layer is used to perform the mapping operation on the input dataset; the features of this dataset are categorized into weight distributions.

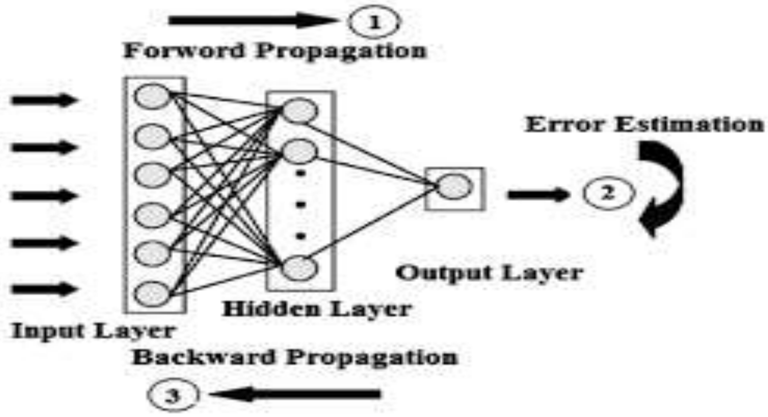

Figure 6. Layered architecture of BP-ANN model

Then the classification operation was implemented in the two levels of hidden layer. The two levels of hidden layer hold individually normality and abnormalities of the breast cancer characteristic information. Based on the segmentation criteria, it is categorized as normal and abnormal classification stage. These two levels are mapped as labels in output layer. Again the hidden layer also contains the abnormal cancer types separately; it is also holds the benign and malignant cancer weights in the second stage of hidden layer. Similarly, these benign and malignant weights are also mapped as label into output layer. When the test image is applied, its SDM features are applied for testing purpose in the classification stage. Based on the maximum feature matching criteria utilizing Euclidean distance manner it will function. If the feature match occurred with hidden layer 1labels, then it is classified as normal mammographic image. If the feature match occurred with hidden layer 2 labels with maximum weight distribution, then it is classified as benign effected cancer image. If the feature match occurred with hidden layer 2 labels with minimum weight distribution, then it is classified as malignant affected cancer image.

4. RESULTS AND DISCUSSIONS

4.1. Dataset

Total 1000 mammographic images are adopted for this experiment analysis where 400 of malignant, 400 of benign and 200 of normal mammographic X-ray images with the consideration of patient mean age around differerent ages and ranging from 18 to 81. The breast grazes assortment from 2mm to 20mm in mass and several patients contain several grazes whereas some other patients might have merely one. Figure 7a shows the original input image datasets, Figure 7b shows the preprocessed noise free enhanced images using NSCT transform. Figure 7c shows the detection of breast cancer using APFCM clustering method and classified using BP-ANN methodology respectively.

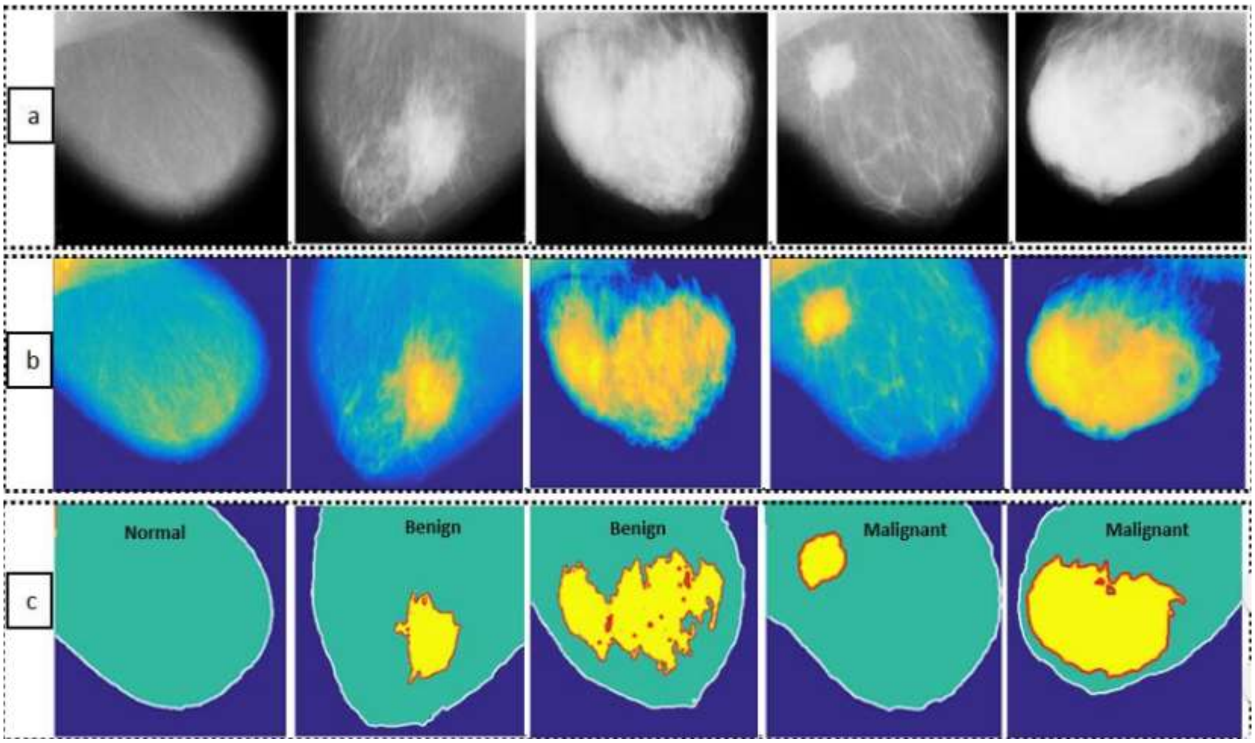

Fig.7. detection and classification procedure of proposed method
(a) Original images, (b) Preprocessed images, (c) segmented and classified image

4.2. Evaluation criteria

For valuation of classification outcomes, we utilized three qualitative metrics such as specificity, accuracy and sensitivity. The accuracy can be defined as out of certain random test cases, how many outcomes give the perfect classification output. The sensitivity is defined as individual classification accuracy, how much the method is sensitive towards the malignant and benign cancers. And specificity is defined as the how much accurately the location of tumor is recognized.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

where TP conveys the amount of test cases properly recognized as malignant, FP conveys the amount of test cases improperly recognized as malignant, TN conveys the amount of test cases properly recognized as benign and FN is conveys the amount of test cases improperly recognized as benign. In the training procedure, network limits were attuned by the preparation slaughter and after that the justification dataset would be utilized to check the matching amount of the attuned system. The matching curvatures of system depend on network testing slaughter and training loss slaughter. In order to additionally calculate the planned technique, we contrasted it with pair of NN-contained methods utilized in [14], [15]. For the categorization, we adopted SVM [17], multi-kernel SVM [18] and K-nearest neighbor (KNN) [16] classifiers from the literature for comparison with the proposed hybrid BP-ANN classifier model. Table 1demonstrates that quality evaluation criteria of existing and proposed classifiers, where proposed hybrid BP-ANN classifier outperforms the conventional SVM, MK-SVM and KNN classifiers to distinguish the benign and malignant from the mammographic X-ray images. Figure 8 represents the

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 10, No. 3, October : 2023

graphical representation of different comparison schemes in contrast with proposed method using the quantitative parameters such as specificity, accuracy and sensitivity.

Table 1. Performance of quality metrics using existing and proposed hybrid BP-ANN model

Method	Accuracy (in %)	Specificity (in %)	Sensitivity (in %)	Recall (in %)
KNN [16]	76.09	75.51	77.40	78.40
SVM [17]	80.01	79.18	80.72	81.72
ANN [1]	80.42	80.18	81.81	83.81
CNN[13]	90.11	89.28	90.36	94.36
Proposed	95.91	95.81	96.34	97.34

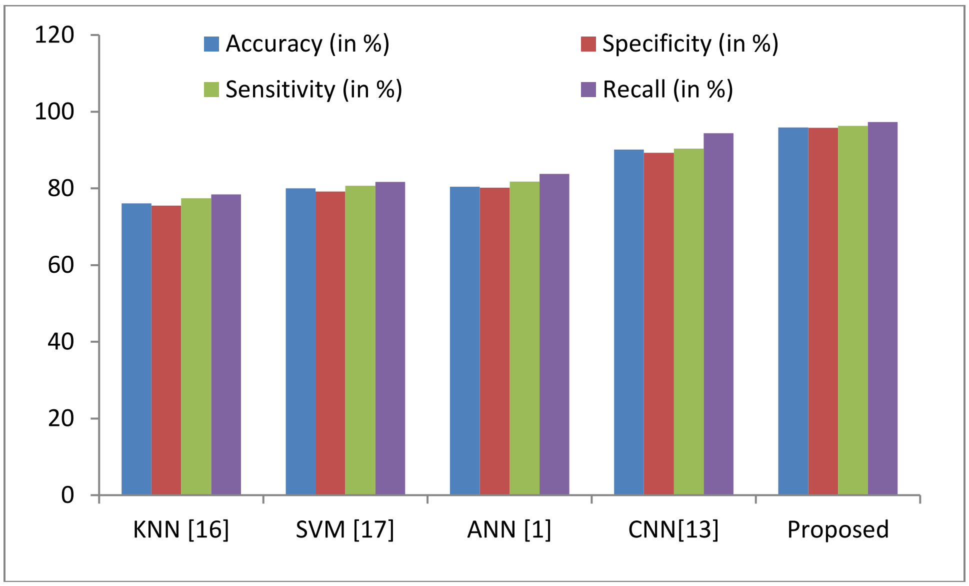

Figure 8. Performance evaluation of existing and proposed hybrid BP-ANN classifier model

5. CONCLUSION

This article presented a computational methodology for detection & classification of breast cancer from mammographic X-ray images using NSCT-based APFCM with SDM and BP-ANN, where NSCT is utilized for preprocessing, which eliminates any unwanted noise elements or artifacts innovated while image acquisition. Then APFCM clustering is employed for ROI extraction and detection of cancerous cells. In addition, SDM is used to extract the features from the segmented mammographic X-ray image to form a feature vector.Finally; BP-ANN is employed to classify the type of cancer as normal, benign or malignant using trained network model. In future, this work can be extended by implementing a greater number of network layers into the BP-ANN and can also be applied for other type of cancers.

REFERENCES

[1] Kaymak, Sertan, Abdulkader Helwan, and DilberUzun. "Breast cancer image classification using artificial neural networks." Procedia computer science 120 (2017): 126-131.
[2] Singh, Anuj Kumar, and Bhupendra Gupta. "A novel approach for breast cancer detection and segmentation in a mammogram." Procedia Computer Science 54 (2015): 676-682.

[3] Danala, Gopichandh, et al. "Classification of breast masses using a computer-aided diagnosis scheme of contrast enhanced digital mammograms." BP-ANNals of biomedical engineering 46.9 (2018): 1419-1431.
[4] Jouni, Hassan, et al. "Neural Network architecture for breast cancer detection and classification." 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET). IEEE, 2016.
[5] Abdel-Ilah, Layla, and Hana Šahinbegović. "Using machine learning tool in classification of breast cancer." CMBEBIH 2017. Springer, Singapore, 2017. 3-8.
[6] Li, Xingyu, et al. "Discriminative pattern mining for breast cancer histopathology image classification via fully convolutional autoencoder." IEEE Access 7 (2019): 36433-36445.
[7] Wang, Zhiqiong, et al. "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features." IEEE Access 7 (2019): 105146-105158.
[8] Sebai, Meriem, Tianjiang Wang, and Saad Ali Al-Fadhli. "PartMitosis: A Partially Supervised Deep Learning Framework for Mitosis Detection in Breast Cancer Histopathology Images." IEEE Access 8 (2020): 45133-45147.
[9] Saha, Monjoy, and Chandan Chakraborty. "Her2net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation." IEEE Transactions on Image Processing 27.5 (2018): 2189-2200.
[10] Zhang, Xiaofei, et al. "Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks." IEEE transactions on nanobioscience 17.3 (2018): 237-242.
[11] Eltrass, Ahmed S., and Mohamed S. Salama. "Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitized mammograms." IET Image Processing 14.3 (2019): 495-505.
[12] Das, Asha, Madhu S. Nair, and S. David Peter. "Sparse representation over learned dictionaries on the riemBP-ANNian manifold for automated grading of nuclear pleomorphism in breast cancer." IEEE Transactions on Image Processing 28.3 (2018): 1248-1260.
[13] Khan, Hasan Nasir, et al. "Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network." IEEE Access 7 (2019): 165724-165733.
[14] K. Sohn et al., “Learning and selecting features jointly with point-wise gated BoltzmBP-ANN machines,” in International Conference on International Conference on Machine Learning, 2013, pp. II–217.
[15] Q. Zhang et al., “Deep learning based classification of breast tumors with shear-wave elastography,” Ultrasonics, vol. 72, pp. 150–157, 2016.
[16] S. Z. Erdogan and T. T. Bilgin, “A data mining approach for fall detection by using k-nearest neighbour algorithm on wireless sensor network data,” IET Communications, vol. 6, no. 18, pp. 3281–3287, 2013.
[17] E. Karthikeyan and S. Venkatakrishnan, “Beast Cancer Classification using SVM Classifier”, International Journal of Recent Technology and Engineering, vol. 8, no. 4, pp. 527-529, Nov. 2019.
[18] P. B. Chandra and S. K. Sarkar, “Detection and classification technique of breast cancer using multi kernel SVM classifier approach”, In Proc. of International Conference on Applied Signal Processing, Kolkata, India, IEEE, Jul. 2019.