

## A Review on Feature Selection Techniques in Digital Mammograms

<sup>1</sup>L Kanya kumara, <sup>2</sup>B N Jagadesh

<sup>1</sup>Research Scholar, Department of Computer Science & Engineering, K L E F (Deemed to be University), Vaddeswaram, Guntur, A.P, India. kanyabtech@yahoo.com

<sup>2</sup>Professor, Department of Computer Science & Engineering, Srinivasa Institute of Engineering and Technology, Amalapuram, Andhra Pradesh, India. nagajagadesh@gmail.com

**Article History:** Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2020;

**Abstract:** The most of the women in the world are suffering from a deadly disease called Breast Cancer (BC). Breast cancer is analyzed by using imaging modalities such as mammograms, magnetic resonance imaging, ultrasound, and thermograms. Among all, mammograms are the low dosage, less cost, more effective, and accurate method to detect BC in early stages. There are many Computer-Aided Detection (CAD) systems for the automatic detection of masses in mammograms. These techniques are helping radiologists and physicians in diagnosing disease. The objective of this paper is to overview different CAD systems in which mainly we focused on feature selection, as feature selection techniques are used to reduce the complexity of the classifiers and also increase the accuracy. We conclude that suitable optimization techniques should be chosen to increase the accuracy of the classifier so that we can increase the survival rate of the patient.

**Keywords:** Computer Aided Detection, Breast cancer, feature selection, Mammograms.

### 1. Introduction

The women in the world are suffering from Breast Cancer (BC). According to American Institute for Cancer Research, there were 2 million new cases in 2018 (<https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics>) and suggested to go for screening once a year after 40 years of age (Lewis C, 1999). The reason behind this cancer is still a research topic (Mohanty, F, Rup, S, & Dash, B. , 2019). So, an efficient and cost-effective modality should be used for early detection of BC (Sri Hari Nallamala et.al, 2019). Mammography is the best screening modality used to detect tumors in breasts (M. Mohsin Jadoon, Qianni Zhang, Ihsan Ul Haq, Sharjeel Butt & Adeel Jadoon, 2017). The masses which are seen in the mammograms are classified as benign or malignant based on the shape and size. The shape of benign tumor is round or oval whereas partially round or irregular in shape for malignant tumors (U. Raghavendra, U. Rajendra Acharya, H. Fujita, A. Gudigar, J.H. Tan, & S. Chokkadi, 2016).

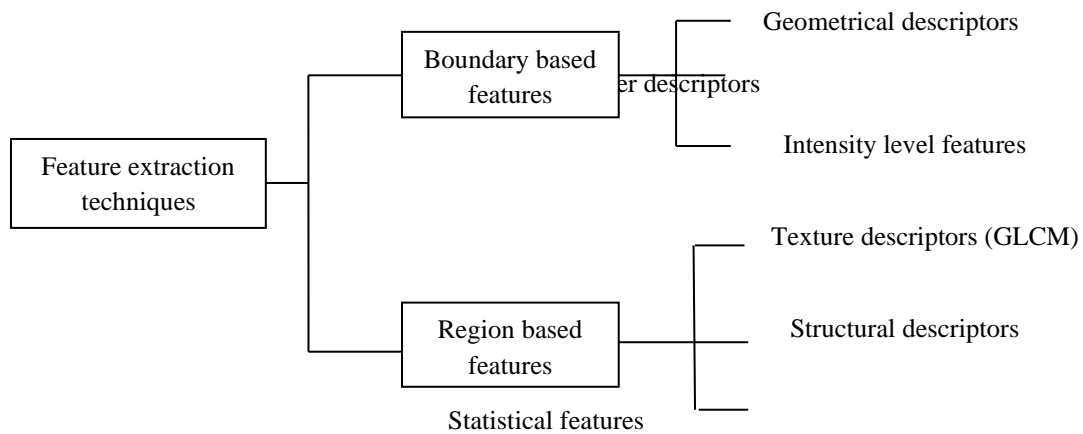
A Computer-Aided Detection (CAD) (Kayode AA, Akande NO, Adegun AA & Adebisi MO, 2019) is used to help radiologists and physicians to detect the BC in early stages. This CAD framework uses the features retrieved from the mammogram images to classify them into normal or benign or malignant. There are numerous techniques for mammogram classification, but this is still under research as an intrinsic representation of mammogram and classification (M. Mohsin Jadoon et.al, 2017).

The general method for CAD mainly constitutes feature extraction, feature selection, and classification. Some works additionally use pre processing stage before feature extraction for increasing the quality of an image (U. Raghavendra et.al, 2016). For pre processing and feature extraction image processing techniques, for feature selection meta-heuristic techniques, and for classification, machine learning is used. The extracted features are given as input to feature selection to select the optimal features. The CAD system without feature selection will affect the accuracy of the classifier. So, feature selection should be applied before the classification (Shankar Thawkar, 2020).

This review will surely guide the researchers in choosing the most appropriate feature selection technique(s) required to detect tumors. It will also help the accurate classification of breast abnormalities thereby increasing the survival rate of the patient. This paper consists of different feature extraction techniques used by different researchers for detecting tumors which is discussed in section 2 and feature selection techniques are explained in section 3.

### 2. Feature Extraction techniques

Breast masses are identified as benign or malignant tumors based on the features extracted from mammogram images. Feature extraction is the process of extracting features within an image that are detected for further processing. There are many feature extraction techniques available in image processing such as boundary-based, region-based, and shape-based features. Boundary-based descriptors are geometrical, shape numbers, Fourier descriptors (Mohan Allam & M. Nandhini 2018) and statistical moments. Region-based features (Syed Jamal Safdar Gardezi, Ahmed Elazab & Tianfu Wang, 2019) are histogram-based texture descriptors, structural and statistical features. Different authors used different feature extraction techniques for mammogram classification which are represented in figure 1.



**Figure 1.** Feature extraction techniques used in mammogram classification

Digital mammogram classification was done based on GLCM features which were extracted from MIAS and DDSM datasets and these were optimized by Forest Optimization Algorithm (Mohanty, F, Rup, S, & Dash, B. , 2019 ). They achieved maximum classification accuracy. In (Fernando Soares S ervulo de Oliveira et.al, 2015), the authors have used texture features called Gray Level Co-occurrence Matrix (GLCM) was used to classify mammograms into normal or abnormal. GLCM (S. Jayaprada, JayaLakshmi & L. Kanya Kumari 2021) features include energy, contrast, correlation, homogeneity, entropy, etc. To classify breast regions into a mass or non-mass the authors (Radhika Mani,Jagadesh, Satyanarayana & Potukuchi D.M, 2021) used surface characteristics like size, shape, density, arrangement, and proportion of elementary parts. They have also used taxonomic indexes like taxonomic diversity index and taxonomic distinctness.

In (Li, H., Meng, X & Wang, T. et al. 2017), the authors designed a model for finding breast masses based on the root mean square roughness is only the feature considered to describe the irregular degree of one dimensional signature. It is not possible to select one technique as the best feature extraction for finding breast tissue (Daniel O.Tambasco Bruno et.al, 2016). Local Binary Patterns with curvelet transformation features are extracted from mammograms to describe the breast tissues.

An efficient method for mass segmentation and classification is achieved by combining shape, texture, and intensity features(Dong M, Lu X, Ma Y, Guo Y, Ma Y & Wang K,2015). The features extracted are mean, standard deviation, smoothness, skewness, uniformity, entropy, kurtosis, pixel value fluctuation, and conspicuity and achieved good accuracy for random forest (BN Jagadesh1& L Kanya Kumari, 2021) compared to state-of-art methods. A CAD was designed to find abnormal breasts using weighted type Fourier transform to achieve the unified time-frequency spectrum. Good classification accuracy was achieved for SVM(Nallamala, S.H., Mishra, & P., Koneru, S.V. ,2019) then compared with state of art methods in (Yu-Dong Zhang, Shui-Hua Wang, Ge Liu & Jiquan Yang ,2016). Breast image classification was done by extracting discrete cosine transform features and used KNN (L Kanya Kumari & Jagadesh B.N ,2020) and Na ive Bayes classifiers for classification (E. J. Kendall & M. T. Flynn ,2014). Automatic detection of BC was designed based on Gabor wavelet filter and locality sensitive discriminant analysis and achieved good classification accuracy for KNN (U. Raghavendra, U. Rajendra Acharya, H. Fujita & A. Gudigar ,2016). From the literature, we have observed that GLCM features provide better results to extract from mammogram images.

## 2. Feature Selection Techniques

Feature selection is an important task for analyzing the data to predict or classify the label for an image. Optimization algorithms use specific parameters with common parameters for evaluation. These parameters are playing an important role in selecting features and also in the performance of the classifiers. To design and develop an efficient classifier, feature selection techniques must be used to reduce both time and space complexities for the mammogram classification.

The main goal of this step is to remove unnecessary or irrelevant features from extracted feature vector (Shankar Thawkar & Ranjana Ingolikar ,2020). There are several optimization techniques to select the optimal features. They are Evolutionary Algorithms (EA) and Swarm Intelligence (SI). Forest Optimization Algorithm (FOA), Evolutionary Programming (EP), and Genetic Algorithm (GA) are the examples of EI whereas Particle

Swarm Optimization(PSO), Ant Colony Optimization (ACO) are the examples of SI (<https://sites.google.com/site/tlborao/>). All these techniques are used in mammogram classification.

In recent days, the challenging task in the research area is feature selection and classification of breast cancer. In this, Biogeography based optimization was used to select the features from the DDSM dataset, and ANN and Adaptive Neuro-Fuzzy Inference System (Sri Hari Nallamala, et.al, 2019) were used as the fitness function. The sensitivity, specificity, and Area Under the Curve(AUC) are 99.10%, 98.72%, and 0.99 respectively for Biogeography Based Adaptive Neuro-Fuzzy Inference System (BBO-ANFIS).

To classify the clusters of microcalcifications in the DDSM dataset, the authors (Khehra, B.S., Pharwaha, A.P.S, 2017) have used texture Fourier domain, shape and wavelet domain-based features were extracted. Totally 50 features were extracted. Optimal features were selected by using Genetic Algorithms (GA), Particle Swarm Optimization (PSO),and Biogeography Based Optimization (BBO). GA-based SVM, PSO based SVM, and BBO based SVM were used as classifiers. The accuracy was measured using random trails and cross-validation. They have concluded that PSO and BBO based feature selection techniques were better than GA-based feature selection.

The authors (Salabat Khana, Muhammad Hussainb, Hatim Aboalsamhb, Hassan Mathkourb, George Bebisc ,2016)optimized the Gabor features which were extracted from the DDSM dataset using PSO. The fitness function used was SVM with Gaussian kernel. The authors compared PSO with GA and concluded that PSO was suitable with continuous non-linear optimization problems than GA. PSO concentrates on local and global search but GA concentrates on Global search. Implementation of PSO is simple than GA.

In(Sh, Shahraki H, Rowhanimanesh AR, Eslami S,2016), the authors have done a feature selection step on 3 datasets: Wisconsin Breast Cancer (WBC) (Sri Hari Nallamala, et.al,2019), Wisconsin Diagnosis breast Cancer (WDBC), and Wisconsin Prognosis Breast Cancer (WPBC) for BC diagnosis. The classifiers used to measure the accuracy were Polynomial Classifier (PC), ANN, and GA based classifiers. The evaluation metrics used were sensitivity, specificity, and accuracy. The authors concluded that PC works better than ANN and GA in diagnosing WBC dataset where as ANN works better for diagnosing WDBC and WPBC datasets than the other two classifiers. The authors have also concluded that feature selection will improve the classifier accuracy results.

The authors (. Rouhi, M. Jafari, S. Kasaei& P. Keshavarzian, 2015) have extracted texture, intensity and shape features from MIAS and DDSM datasets. To select optimal features GA was used. The selected features were given as input to the classifier. As the Multi Layer Perceptron (MLP) classifier is good in pattern recognition, it was used to evaluate the performance of the proposed method(Chauhan, S., Goel, V., & Dhingra, S,2012 ; Cheng, H. D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P., & Du, H. N ,2006; Kuo, S. J., Hsiao, Y. H., Huang, Y. L., & Chen, D. R. ,2008; Pawar, P. S., & Patil, D. R ,2013; Tahmasbi, A., Saki, F., & Shokouhi, S. B ,2011). The evaluation metric used was 10-fold cross-validation. The authors concluded that their proposed method helps the radiologist and has the advantage of an increase in the survival rate of the patient with early detection of breast cancer.

To classify the clusters of microcalcifications in the DDSM dataset, the authors (Khehra, B.S., Pharwaha, A.P.S, 2017) have used texture Fourier domain, shape and wavelet domain-based features were extracted. Totally 50 features were extracted. Optimal features were selected by using Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Biogeography Based Optimization (BBO). GA based SVM, PSO based SVM, and BBO based SVM were used as classifiers. The accuracy was measured using random trails and cross-validation. They have concluded that PSO and BBO based feature selection techniques were better than GA-based feature selection.

A multi-scale GLCM (Al Mutaz M. Abdalla, Safaai Dress & Nazar Zaki, 2011) and Second-Order Statistics of Wavelet Co-efficient (SOSWC) are calculated from MIAS and DDSM mammogram images to detect breast abnormalities (Zyout I, Czajkowska J& Grzegorzec M. 2015). To optimize the features PSO was used and classifier used was SVM, and concluded that PSO based model gives better results.

The authors (Mei-Ling Huang ,Yung-Hsiang Hung , Wen-Ming Lee, R. K. Li & Tzu-Hao Wang ,2012) have done an experiment for classification of the mammographic Mass dataset from UCI repository using PSO with ANN (PSO-ANN) and Adaptive Neuro-Fuzzy Inference System (PSO-ANFIS), Case-Based Reasoning with Logistic Regression model (CBR-LR), and decision tree(Sri Hari Nallamala, et al., 2018) model (CBR-DT) were used. The experimental results showed that PSO-ANN and PSO-ANFIS performed well compared to the CBR-LR and CBR-DT.

The authors (Eltoukhy MM, Faye I, Samir BB, 2012) proposed a hybrid approach to diagnose breast cancer by using wavelet (Min Dong, Zhe Wang, Chenghui Dong, Xiaomin Mu & Yide Ma, 2017) and curvelet transforms. The optimal features were selected using t-test (Aijuan Dong & Baoying Wang, 2009) approach and classified based on SVM. Performance was evaluated using 5-fold cross validation and obtained 96.56% to classify mammograms into normal or abnormal whereas 95.58% obtained for benign or malignant.

To classify the mammograms the authors (Beura S, Majhi B & Dash R. ,2015) have used GLCM and 2D Discrete Wavelet Transformation techniques were used to extract the features from MIAS and DDSM datasets. For the selection of the most relevant features, filter-based methods called t-test and F-test were used. Back Propagation Neural Networks (BPNN) classifier was used for classifying the images.

A CAD model was designed called Fast Finite Shearlet Transform (FFST) for feature extraction from MIAS and DDSM datasets (Gedik N ,2016). Ranks were given for extracted features using t-test filter. The SVM classifier was used for classification and 5-fold cross validation for performance measure. They got 98.29%, 98.08% for MIAS and DDSM respectively.

To classify digital mammograms the authors (Figlu Mohanty, Suvendu Rup, Bodhisattva Dash, Banshidhar Majhi, M. N. S. Swamy, 2020) have combined 2D block discrete wavelet and GLCM techniques for feature extraction from MIAS and DDSM datasets. To reduce the dimensions of the feature vector, PCA was used and Forest Optimization Algorithm (FOA) was used to select the optimal features. Their methodology gained good accuracy for SVM, KNN, and C4.5 classifiers to classify the mammograms into benign or malignant and normal or abnormal.

A CAD system (Wang S, Rao RV, Chen P, Zhang Y, Liu A, Wei L ,2017) was designed using Weighted-type Fractional Fourier Transformation (WFRFT) to extract features from the mini-MIAS dataset. PCA was used to reduce the features. Jaya algorithm-based Feed Forward Neural Network (FFNN) was used to train the classifiers and compared the results with state-of-art classifiers. The authors concluded that the proposed algorithm was superior to the state of art methods.

A CAD system was constructed to classify mammograms based on contourlet features and optimal features were selected by using FOA from MIAS and DDSM benchmark datasets. The classifiers used were SVM, Naïve Bayes (NB), C4.5 and KNN. The classification accuracies obtained for SVM, KNN and C4.5 were 100% except NB to classify into normal or abnormal. Similarly, the authors achieved a maximum 98.74% for C4.5 to classify into normal or abnormal.

The authors (Shankar Thawkar, 2020) have chosen the optimal features by using Teaching Learning Based Optimization (TLBO) technique from the Wisconsin Diagnostic Breast Cancer dataset (WDBC). The performance is evaluated by using Discriminant Analysis, Naïve Bayes, Decision trees, Support Vector Machines (SVM), and KNN. SVM has given better results than other classifiers.

A new model (M. N. Sudha & S. Selvarajan, 2016) was designed to classify mammograms. They extracted texture, intensity histogram, radial distance, and shape features. The optimal features were selected by using Enhanced Cuckoo Search (ECS). The performance was evaluated using k-fold cross validation for minimum distance classifier, KNN classifier, and achieved 98.75% and 99.13% accuracy respectively.

The authors (<https://sites.google.com/site/tlborao/>) used shape, texture and intensity based features to extract the features from the DDSM dataset. Total of 25 features were extracted. From these 25 features, 11 features were selected by applying the genetic ensemble method. The parameters used were: number of iterations= 50, population sizes considered were 10, 20 and 30, crossover probability = 0.9 and mutation probability = 0.1. The experiment was done by using Adaboost, random forest, and decision tree to classify the masses in digital mammograms. The authors concluded that the Adaboost classifier has given better accuracy if optimal features were selected and RF is better if all the features were considered. The Misclassification rate of Adaboost, RF and decision tree was 3.85, 4.92 and 14.6 respectively.

GLCM with Genetical Swarm Optimization (GSO) was designed to classify the mini-MIAS mammogram images into normal or abnormal. The classifier used was SVM to measure the performance of the proposed method. They have compared the results with GA-SVM and PSO-SVM and concluded that GSO-SVM has given good performance than compared GA and PSO (Jona J & Nagaveni N ,2012).

The patterns from segmented mammograms were classified using SVM and General Regression Neural Networks (GRNN) and obtained AUC as 0.98 and 0.9780 respectively (Fu J, Lee S, Wong S, Yeh J, Wang A & Wu H, 2005). To achieve this, the authors have used Sequential Forward Search (SFS) to select the features. The authors (Dheeba J, Singh NA & Selvi ST, 2014) evaluated a model for diagnosing breast cancer using Particle Swarm Optimized Wavelet Neural Networks (PSOWNN). This algorithm extracted the features using laws texture energy measures from mammograms collected from screening centres. The performance was measured using AUC and also calculated sensitivity and specificity. Below table 3.1 gives a summary of different papers and limitations/ drawbacks they may possess that may have arisen in their methodology.

The authors (B. Bektaş, İ. E. Emre, E. Kartal & S. Gulsecen, 2018) classified the mini-MIAS mammogram database into benign or malignant by applying Gaussian, median, Wiener filters and increased the contrast of images using the CALHE technique. The features are extracted using GLCM and Linear Binary Pattern (LBP). The best features are selected using correlation. The classifiers applied are NB, CART and RF. They concluded that CLAHE+GLCM+CORR+NB give better results.

**Table 1. Mammogram classification Techniques**

Reference no	Feature selection	Feature extraction	Classifier	Dataset	Performance measure	Conclusion	Limitation/ future scope
[8]	Biogeography based optimization (BBO)	Shape, texture and intensity	ANN and Adaptive Neuro Fuzzy Inference System (ANFIS)	DDSM	Accuracy, AUC, correlation coefficient, Mean Square Error (MSE) and Root Mean Square Error (RMSE)	BBO-ANFIS has given the better classification than BBO-ANN	To be improved the computational time for feature selection and classification
[26]		Texture, Fourier domain, shape and wavelet domain based features	SVM	DDSM	Random trails and k-fold cross validation	BBO-SVM was performed well than GA-SVM	To be applied on other classifiers. TLBO and Firefly algorithm to be used for feature selection
[27]	GA	Gabor features	SVM	DDSM	Sensitivity, specificity and accuracy	Not performed well than compared to PSO	Cuckoo optimization may be applied as future work
[26]		Texture, Fourier domain, shape and wavelet domain based features	SVM	DDSM	Random trails and k-fold cross validation	PSO-SVM and BBO-SVM were performed well than GA-SVM	To be applied on other classifiers. TLBO and Firefly algorithm to be used for feature selection
[28]		--	PS, ANN and GA based classifier	WBC WDBC WPBC	Sensitivity, specificity and accuracy	PS is good in diagnosing WBC. ANN is good in diagnosing	--

			r			WDBC and WPBC	
[29]		Texture, shape and intensity features	MLP	MIAS DDSM	10-fold cross validation	Proposed method increases the survival rate of patient	Other neural networks and other EA has to be applied
[26]	PSO	Texture, Fourier domain, shape and wavelet domain based features	SVM	DDSM	Random trails and k-fold cross validation	PSO-SVM is performed well than GA-SVM	To be applied on other classifiers. TLBO and Firefly algorithm to be used for feature selection
[27]		Gabor features	SVM	DDSM	Sensitivity, specificity and accuracy	PSO based classification is better than GA based classification	Cuckoo optimization may be applied as future work
[35]		Multi scale GLCM and SOSWC	SVM	MIAS DDSM	AUC	PSO based model is better to reduce FPR problem	--
[37]		Mammographic Mass data set from UCI	PSO-ANN PSO-ANFIS	PSO-ANN PSO-ANFIS	Area under ROC curve (AROC)	PSO-ANN and PSO-ANFIS better results than CBR-LR CBR-DT	May not be applicable for other datasets
[39]		Wavelet and curvelet transforms	SVM	MIAS	5-fold cross validation	--	--
[41]	t-test	GLCM and 2D Discrete Wavelet Transformation	BPN	MIAS DDSM	AUC	Good accuracy for classification	--
[40]		FFST	SVM	MIAS DDSM	5-fold cross validation	Good accuracy for their proposed methodology	--
[44]	PCA	2D Block discrete wavelet transform and GLCM	SVM, KNN and C4.5	MIAS DDSM	Confusion matrix, area under curve, and Matthews correlation coefficient	Good accuracy	Good for smaller datasets and need to apply for large datasets
[45]		WFRFT	Jaya-FFNN	Mini-MIAS	10-fold stratified cross validation	Superior to state-of-art methods	Wavelet entropy and other features may

							extract and other classifiers such as radial Basis Function, SVM and fuzzy SVM has to be applied
[44]	FOA	2D Block discrete wavelet transform and GLCM	SVM, KNN and C4.5	MIAS DDSM	Confusion matrix, area under curve, and Matthews correlation coefficient	Good accuracy	Good for smaller datasets and need to apply for large datasets
[3]		Contourlet Features	SVM, Naïve Bayes (NB), C4.5 and KNN	MIAS DDSM	Accuracy	--	1.Can be applied on large databases 2.Deep Learning techniques may be applied 3.Multi resolution transformation techniques to get more robust features 4.may be applied to classify other types of cancers
[9]	TLBO	--		WDBC	Accuracy	SVM with BTLBO optimal features have given the better results	New hybrid algorithm to be proposed to increase the performance
[46]	Enhanced Cuckoo Search	Texture, intensity histogram, radial distance and shape features	Minimum distance Classifier and KNN	DDSM	k-fold cross validation	Performance is good in terms of accuracy with the few features.	--
[23]	Genetic ensemble	Shape, texture and intensity	Random forest Decision tree Ada boost	DDSM	Accuracy and misclassification rate	Adaboost is better if optimal features are considered RF is better if all features are selected.	To be focused to reduce the misclassification error
[47]	Genetic Swarm Optimization	GLCM	SVM	Mini-MIAS	AUC	Compared with GA and PSO	--
[48]	SRS	Spatial and spectral domain	SVM, GRNN	Nijmegen University Hospital (Netherlands)	AUC	SVM is better than GRNN	SRS can be applied to content based retrieval of mammograms

		features		)			
[49]	PSOW NN	Laws texture energy measure	WN N	Database from screening centers	AUC	Better results than traditional classifiers	--

The above literature provides information about breast cancer in the patient which can help to overcome certain limitations. From the above study, it was observed that most of the researchers were used meta-heuristic techniques to optimize the features to get better classification results. They have used BBO,GA,PSO, TLBO,FOA,PCA, and t-test on shape, texture and intensity-based features were extracted. In some literature, the authors have used to apply optimization techniques for feature selection to increase the classification accuracy. This paper gives information about the existing methods and also very much helpful to the researchers in the following aspects like: choosing of optimal feature selection technique for other modalities, combining different feature selection techniques to design a hybrid approach and selection of efficient feature selection technique based on the features extracted for better classification.

### Conclusion

Early detection of BC can be done by finding the cancer cells in the breast. To achieve this, the general steps followed in every CAD system are image preprocessing, feature extraction, feature selection and classification. Different researchers have used different techniques for early detection.. This paper reviewed several feature selection algorithms like GA, PSO, PCA, FOA, TLBO and some authors have applied hybrid approaches too for achieving good classification accuracy. So, keeping all these observations, we can conclude that the classification accuracy is depending on choosing a suitable feature selection technique which helps the radiologists and physicians to detect the tumors so that the survival rate of the patient can be increased. We too extracted features from MIAS dataset based on texture, shape and intensity. From the literature, it is observed that meta-heuristic techniques plays a vital role in feature selection to classify mammogram images as they can improve the accuracy of the classifiers.

### References

1. Aalaei Sh, Shahraki H, Rowhanimanesh AR, Eslami S .(2016). Feature selection using genetic algorithm for breast cancer: experiment on three different datasets. *Iran J Basic Med Sci*, 19, 476-482.
2. Aijuan Dong Baoying Wang. (2009).Feature Selection and Analysis on Mammogram Classification. *Communications, Computers and Signal Processing*.
3. Beura S, Majhi B, Dash R. (2015). Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. *Neuro computing* ,154,1–14.B. Bektaş, İ. E. Emre, E. Kartal and S. Gulsecen. (2018).Classification of Mammography Images by Machine Learning Techniques. In Proc. International Conference on Computer Science and Engineering, Sarajevo, Bosnia and Herzegovina, pp. 580-58.
4. BN Jagadesh1, L Kanya Kumari.(2021).A GLCM based Feature Extraction in Mammogram Images using Machine Learning Algorithms, *International Journal of Current Research and Review*, 13(5).
5. Chauhan, S., Goel, V., & Dhingra, S. (2012). Pattern recognition system using MLP neural networks. *Pattern Recognition*,4(9), 43–46.
6. Cheng, H. D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition*, 39(4), 646–668.
7. Daniel O.Tambasco Bruno,Marcelo Z.do NascimentoabRodrigo P.Ramos,Valério R.Batista1aLeandro A.NevesAlessandro S.Martins. (2016). LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues. *Expert Systems with Applications*, 55, 329-340.
8. Dheebea J, Singh NA, Selvi ST.Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *J Biomed Inform.* 49, 45–52.
9. Dong M, Lu X, Ma Y, Guo Y, Ma Y, Wang K (2015). An Efficient Approach for Automated Mass Segmentation and Classification in Mammograms. *J Digit Imaging*,28(5),613-25.
10. E. J. Kendall and M. T. Flynn.(2014) .Automated breast image classification using features from discrete cosine transform, *PLoS ONE* 3.



24. Eltoukhy MM, Faye I, Samir BB. (2012). A statistical based feature extraction method for breast cancer
25. diagnosis in digital mammogram using multi resolution representation. *Comput Biol Med*, 42(1),123–128.
26. Fernando Soares Sérvulo de Oliveira, Antonio Oseas de Carvalho Filho, Aristófanés Corrêa Silva, Anselmo
27. Cardoso de Paiva , Marcelo Gattass. (2015). Classification of breast regions as mass and non-mass based
28. on digital mammograms using taxonomic indexes and SVM. *Computer in Biology and Medicine*, 1, 42-53.
29. Figlu Mohanty, Suwendu Rup, Bodhisattva Dash, Banshidhar Majhi, M. N. S. Swamy. (2020). Digital
30. mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection
31. approach. *Neural Comput & Applic* .32, 7029–7043.
32. Fu J, Lee S, Wong S, Yeh J, Wang A, Wu H. (2005). Image segmentation feature selection and pattern
33. classification for mammographic microcalcifications. *Comput Med Imaging Graph*, Vol. 29(6),419–429.
34. Gedik N, A new feature extraction method based on multiresolution representations of mammograms. *Appl Soft*
35. *Computing*, 44, 128–133.
36. <https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics>. [ Accessed on 13/1/2020].
37. <https://sites.google.com/site/tlborao/>
38. Jona J, Nagaveni N. A hybrid swarm optimization approach for feature set reduction in digital mammograms.
39. *WSEAS Trans Inf Sci Appl*, 9,340–349.
40. Kayode AA, Akande NO, Adegun AA, Adebisi MO. (2019). An automated mammogram classification system
41. using modified support vector machine. *Med Devices* 12, 275-284.
42. Khehra, B.S., Pharwaha, A.P.S.(2017). Comparison of Genetic Algorithm, Particle Swarm Optimization and
43. Biogeography-based Optimization for Feature Selection to Classify Clusters of Micro calcifications. *J. Inst.*
44. *Eng. India Ser. B* ,98,189–202.
45. Kuo, S. J., Hsiao, Y. H., Huang, Y. L., & Chen, D. R. (2008). Classification of benign and 879 malignant breast
46. tumors using neural networks and three-dimensional power Doppler ultrasound. *Ultrasound in Obstetrics &*
47. *Gynecology*, 32(1), 97–102.
48. Lewis C. (1999) .Fda sets higher standards for mammography. *FDA Consum*,33(1),13–17.
49. Li, H., Meng, X., Wang, T. et al. (2017). Breast masses in mammography classification with local contour
50. features. *BioMed Eng OnLine* .
51. L Kanya Kumari , Jagadesh B.N. (2020) .A Novel Approach for Detection of Tumors in Mammographic Images
52. Using Fourier Descriptors and KNN. *Lecture Notes in Electrical Engineering*, 601.
53. Mei-Ling Huang & Yung-Hsiang Hung & Wen-Ming Lee & R. K. Li & Tzu-Hao Wang.(2012). Usage of Case-
54. Based Reasoning, Neural Network and Adaptive Neuro-Fuzzy Inference System Classification Techniques
55. in Breast Cancer Dataset Classification Diagnosis. *J Med Syst* ,36,407–414.
56. Mohanty, F., Rup, S., Dash, B. et al.(2019). Mammogram classification using contourlet features with forest
57. optimization-based feature selection approach. *Multimed Tools Appl* 78, 12805–12834.
58. Mohan Allam , M. Nandhini.(2018). Optimal feature selection using binary teaching learning based
59. Optimization Algorithm. *Journal of King Saud University - Computer and Information Sciences*.
60. M. Mohsin Jadoon, Qianni Zhang, Ihsan Ul Haq, Sharjeel Butt, Adeel Jadoon.(2017) .Three-Class Mammogram
61. Classification Based on Descriptive CNN Features. *BioMed Research International*, 11 pages.
62. Min Dong, Zhe Wang, Chenghui Dong, Xiaomin Mu, Yide Ma. (2017). Classification of Region of Interest in
63. Mammograms Using Dual Contourlet Transform and Improved KNN, *Journal of Sensors*, 15 pages.
64. M. N. Sudha, S. Selvarajan, Feature Selection Based on Enhanced Cuckoo Search for Breast Cancer
65. Classification in Mammogram Image, *Circuits and Systems*, 7, 327-338.
66. Mutaz M. Abdalla, Safaai Dress, Nazar Zaki.(2011). Detection of Masses in Digital Mammogram Using
67. Second Order Statistics and Artificial Neural Network. *International Journal of Computer Science &*
68. *Information Technology*, 3(3).
69. Nallamala, S.H., Mishra, P., Koneru, S.V.(2019). Breast cancer detection using machine learning approaches.
70. *International Journal of Recent Technology and Engineering*,7(5), 478–481.
71. Pawar, P. S., & Patil, D. R.(2013). Breast cancer detection using neural network models. *In Proc. International*

72. *conference on communication systems and network technologies*, 568–572.
73. Radhika Mani M., Jagadesh B.N., Satyanarayana C., Potukuchi D.M. (2021). Analysis of Shape Signature in
74. First and Second Derivatives by Using Wavelet Transformation. *Lecture Notes in Electrical Engineering*,
75. 698.
76. R. Rouhi, M. Jafari, S. Kasaei, P. Keshavarzian. (2015). Benign and malignant breast tumors classification based
77. on region growing and CNN segmentation. *Expert Syst. Appl.*, 42 (3), 990–1002.
78. Salabat Khana, Muhammad Hussainb, Hatim Aboalsamhb, Hassan Mathkourb, George Bebis, Mohammed
79. Zakariahd (2016). Optimized Gabor features for mass classification in mammography, *Applied Soft*
80. *Computing*, 267-280.
81. Shankar Thawkar.(2020). Classification of Masses in Digital Mammograms Using biogeography based
82. optimization. *Journal of King Saud University - Computer and Information Sciences*, 32(10), 1140-1148.
83. Shankar Thawkar and Ranjana Ingolikar.(2020). Classification of Masses in Digital Mammograms Using the
84. Genetic Ensemble Method, *J. Intell. Syst.*, 29(1), 831-845.
85. S Jayaprada<sup>1</sup>, G. JayaLakshmi<sup>2</sup> and L. Kanya Kumari.(2021). Fast Hybrid Adaboost Binary Classifier for Brain
86. Tumor Classification, *IOP Conf. Ser. Mater. Sci. Eng.* 1074 012016.
87. Sri Hari Nallamala, et.al.(2019). Pedagogy and Reduction of K-nn Algorithm for Filtering Samples in the Breast
88. Cancer Treatment. *International Journal of Scientific and Technology Research*, 8(11), 2168– 2173.
89. Sri Hari Nallamala, et.al.(2019). Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference
90. Systems. *International Journal of Advanced Trends in Computer Science and Engineering*, 8(2), 259 – 264.
91. Sri Hari Nallamala, et al.(2018). Literature Survey on Data Mining Approach to Effectively Handle Cancer
92. Treatment. *International Journal of Engineering and Technology*, 7(2.7), 729 – 732.
93. Sri Hari Nallamala, et.al.(2019). Breast Cancer Detection using Machine Learning Way. *International Journal of*
94. *Recent Technology and Engineering*. 8(2S3), 1402 – 1405.
95. Syed Jamal Safdar Gardezi, Ahmed Elazab, and Tianfu Wang.(2019). Breast Cancer Detection and Diagnosis
96. Using Mammographic Data: Systematic Review. *Journal of medical Internet Research*, 21(7).
97. Tahmasbi, A., Saki, F., & Shokouhi, S. B.(2011). Classification of benign and malignant masses based on
98. Zernike moments. *Computers in Biology and Medicine*, 41(8), 726–735.
99. U. Raghavendra, U. Rajendra Acharya, H. Fujita, A. Gudigar, J.H. Tan, and S. Chokkadi. (2016). Application of
100. Gabor wavelet and Locality Sensitive Discriminant Analysis for automated identification of breast cancer
101. using digitized mammogram images. *Applied Soft Computing*, 46, 151–161.
102. Wang S, Rao RV, Chen P, Zhang Y, Liu A, Wei L. (2017). Abnormal breast detection in mammogram images
103. by feed-forward neural network trained by Jaya algorithm. *Fundam Inform.* 151, 191–211.
104. Yu-Dong Zhang, Shui-Hua Wang, Ge Liu, Jiquan Yang.(2016). Computer-aided diagnosis of abnormal breasts
105. in mammogram images by weighted-type fractional Fourier transform, *Advances in Mechanical*
106. *Engineering*.
107. Zyout I, Czajkowska J, Grzegorzec M. (2015). Multi-scale textural feature extraction and particle swarm
108. optimization based model selection for false positive reduction in mammography. *Comput Med Imaging*
109. *Graph*, 46, 95–107.