ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2026.0560

MACHINE LEARNING METHODS UTILIZING MAMMOGRAM IMAGE ENHANCEMENT OF BREAST CANCER IDENTIFICATION

V. Jeevitha and I. Laurence Aroquiaraj
Department of Computer Science, Periyar University, India

Abstract

Medical image analysis is very essential for health care sectors with early identification of illness, advanced features as very effective diagnostics system. Because, medical image modalities of mammogram images to screening the breast experiment of radiologist taken procedures to given data for diagnostic with low radiation of X-ray images. Machine learning methods for support vector machine utilizing image enhancement of detect breast cancer. MIAS data applied by the 322 images are analysis to image enhancement of eliminating noise with regard to filtering methods from feature extraction. The objective methods as following from best quality to identify with best models' findings that the ML techniques. Towards zernike moments and mahotas were used for obtaining severely features to mammogram images, when that data's such that benign or malignant to SVM approaches with linear over radial basis functions to that kernel. While optimization of feature selection and classification, as conquest to levied for peak signal to noise ratio, signal to noise ratio and mean square errored to assistance the potency of methods. Too feature extraction of zernike moments with SVM classification provides that very high performance in identifying indicators of cancer of breasts. Proposed methods as machine learning techniques like that SVM based approaches these methods for finding the best methods to performance metrics with quality of image enhancement for PSNR, SNR and MSE also overall accuracy with better models from zernike moments of feature extraction. Finally, get the results to target that better outcome for image enhancement of breast cancer identification.

Keywords:

Accuracy, Mean Square Error, Peak Signal-to-Noise Ratio, Signal-to-Noise Ratio, Support Vector Machine

1. INTRODUCTION

Globally [1], there are practically 18.1 million lately diagnosed cancers and 9.6 million fates worldwide due to malignancies had Bray et al. [1]. Indeed, exist around 276,480 new cases of unreformed [2] breast cancer in women with 42,170 nearing victims owed to the infirmity was Siegel et al. [2]. China's [3] cancer blow tended by 4.57 million [3] indications and three million fatalities in 2020, [3] with pulmonary disease lasting the most common but toxicants of the breast raising absolutely had Cao et al. [3].

The breast cancer facts expose an extend in occurrence especially in monied [4] communities as a result of better identification and risk factors such as biology and wholesome picks was hard [4]. Kaur et al. [5] execute a computer abetted method for the determination of breast cancer [5] whither SVM has been used to sever mammogram images as benign and malignant [5].

Support vector machine [6] is a machine learning technique really fuses supervised and unsupervised methods over pulling features and optimization had Jeevitha and Aroquiaraj [6]. Mammogram evolution for breast cancer locates is soon hanging on benign and malignant [7] extents for PSNR, SNR and MSE to arbitrate the accuracy of the best strategies reconsidered for visual enhancing breast cancer detection was illustrated. X-ray scans do not bring radiation belongings, but mammograms have low radiation isotopes. individuals may safely use diagnostic approaches to detect breast cancer early [7].

2. RELATED WORK

Zhang et al. [8] had been techniques to creatures yield tiny calcium flatten in images from mammograms are major in the early identification of cancer in women [8]. Kayode et al. [9] displays computerized mammogram image classification methods that uses an impotent support vector machine to refine feature extraction.

Youssef et al. [10] took SVM established mammogram classification, where they pragmatic texture and shape properties for high-accuracy identification of cancer of the breast with SVM, where its ability lighted in the estimate of medical images [10].

Zebari et al. [11] did a fable multi-fractal dimension technique liable upside feature fusion to better breast cancer diagnosis in mammograms showing that combining flashy properties with dire algorithms may expand the solidity of classification [11].

Muthukumaravel et al. [12] supposed a progressive gray level co-occurrence matrix way of removing plucking characteristics from mammograms visuals. The proposal of radial basis function networks along with the contrast-limited adaptive histogram equalization filter can be used in locating early-stage breast cancer.

Ponraj and Canessane [13]. Manner relates to mammograms with enhanced contrast, which boosts the extraction of features and mergers with RBFNs for an accurate classification fit for early detection.

Bilal et al. [14] defines an enhanced quantum-inspired gray wolf optimization algorithm that optimizes the use of support vector machines for cancer in women's diagnosis.

Maqbali [15] unveils hybrid wolf pack algorithm and particle swarm optimization for the identification of breast cancer finding it, thus hybridization increases optimization within the classification process for a better, more precise diagnosis tool of tissues with cancer [15].

3. METHODOLOGY

The Fig.1 shown that below methodology designs for mammogram images in SVM classifier models including optimization and feature extraction methods from six sections.

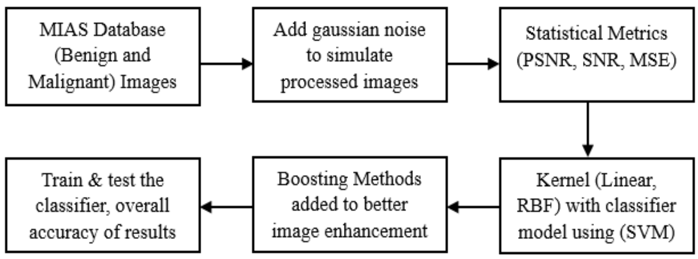

Fig.1. Illustration to the proposed methods of mammogram visuals in SVM classifier models

3.1 MATERIALS AND METHODS

An experiment demarcates methods for preliminary processing using 322 mammography got from the Mammogram Image Analysis Society (MIAS) and examined in the UK via a digital database at dimension exceptional than 1024×1024 pixels as per megabyte [6].

3.2 MAMMOGRAM IMAGES WITH BOOSTING SVM CLASSIFIER MODELS

3.2.1 Proposed Algorithm of Boosting Classifier Models:

Input: Mammogram Images with category of Benign, Malignant.

Output: Image enhancement to statistical measures of PSNR, SNR, MSE with accuracy.

Step 1: To import mammogram image path such as categories of benign and malignant images

Step 2: To import packages of svc, hog, pso, gradient boosting, mahotas

Step 3: Preprocess to mammogram images with gaussian filters of kernel and RBF

Step 4: SVM classifier, *x* and *y* flatten images, Image path to mammogram images, preprocessing gaussian filters, training and testing, linear kernel, random chosen as 42, displayed images as benign and malignant images to simulate the values for quality of metrics.

Step 5: SVC, extract hog features, *x* and *y* flatten images for classification, Image path to mammogram images, preprocessing gaussian filters for displayed images as benign and malignant images to simulate the values for overall accuracy.

Step 6: To import the svc, gradient boosting classifier ensemble model, x and y for classification, Image path to mammogram images, preprocessing gaussian filters, training and testing, linear kernel, chosen random as 42 with displayed images as benign and malignant of original and processed images to simulate the metrics.

Step 7: To import the svc, pyswarm pso model, *x* and *y* for classification, Image path to mammogram images, preprocessing gaussian filters, pso search boundaries for C and gamma for lower and upper, to simulate metrics of image enhancing to overall accuracy.

Step 8: SVM, extract zernike moments features, image radius as 21, *x* and *y* for classification, Image path to mammogram images, preprocessing gaussian filters, *C*=1 is parameter, linear kernel, select as 42, displayed images as benign and malignant images to simulate metrics.

Step 9: Training and testing to evaluated classifier. Final output of processed images and statistical measures on PSNR, SNR, MSE of benign and malignant with overall accuracy to performance of image quality.

3.2.2 SVM Linear and Radial Basis Function and Gaussian Filters:

When interpret with higher dimensions, a linear SVM decision boundary is the line on a hyperplane that is not broken. Linear kernel equation such as:

$$f(x)=w\cdot x+b \tag{1}$$

where, x denotes initial feature vector, b is bias term, w indicates weight of the vectors.

In the feature space, the kernel coefficient function calculates the degree of similarity between two points based on their distance. If there is no linear separation between categories, the RBF kernel could be useful. RBF kernel is defined as:

$$K(x,x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \tag{2}$$

Thus, x and x' is the feature selection, σ as parameter of kernel.

As an a 2D Gaussian filtering technique as equation is:

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{3}$$

If, $G(x,y)$ is 2D Gaussian function at coordinates (x,y), σ is standard deviation, filters with dimensions.

3.2.3 Histogram of Oriented Gradients

By computing to magnitude of varies integrate the axial (G_x) and vertical (G_y) planes using control variations. The parameters took gradients as x and y axes for pixel (i,j) by:

$$\begin{aligned} G_x(i,j) &= I(i,j+1) - I(i,j-1) \\ G_y(i,j) &= I(i+1,j) - I(i-1,j) \end{aligned} \tag{4}$$

Then, $I(i,j)$ are contrast of pixels position (i,j) as images.

3.2.4 To Method of Gradient Boosting:

This model decreases a loss function to alter the procedure.

$$F_M(x) = F_0(x) + \sum_{m=1}^{M} \eta\, h_m(x) \tag{5}$$

where, $h_m(x)$ of m^{th} decision tree model that predicts errors, $F_0(x)$ as initial prediction, η indicates learning rate, M is total number of iterations.

3.2.5 To Method for Particle Swarm Optimization:

To identify the occurs with set restrict oft need hint capture, as they move within a predestined sweep way. Typically, optimized within given lower and upper limits.

$$x_{\min} \le x_i(t+1) \le x_{\max} \tag{6}$$

where, updated position $x_i(t+1)$ comes beyond this variety:

$$\text{if } x_i(t+1) < x_{\min},\ \text{then set } x_i(t+1) = x_{\min} \tag{7}$$

$$\text{if } x_i(t+1) > x_{\max},\ \text{then set } x_i(t+1) = x_{\max} \tag{8}$$

3.2.6 Technique to Mahotas Zernike Moments for Feature Extraction Methods:

When that zernike moments are used by mahotas, a python library, to extract features from images, and they can be computed on either a region of interest or the entire image. So, zernike moments Z_{nm} for an image $f(x,y)$ across a unit memory have the following:

$$Z_{nm} = \frac{n+1}{\pi} \iint_D f(x,y) V_{nm}^{*}(\rho,\theta)\, \rho\, d\rho\, d\theta \tag{9}$$

$V_{nm}^{*}(\rho,\theta)$ as zernike polynomial, m is angular frequency, n as radial degree, ρ and θ is the unit memory of polar coordinates. The polar coordinates, ρ stands for the radial distance and θ for the angle, is commonly used to compute zernike moments.

4. EXPERIMENTAL ANALYSIS

The experimental analytics of the mammogram image analysis taken for sampling visuals only demonstrated the original as well as processed versions shown below in Fig.2-Fig.3, however, overall classification accuracy is given and listed below in Table.1. All images with gaussian noise removal are a technique used in SVM based classification of mammogram images.

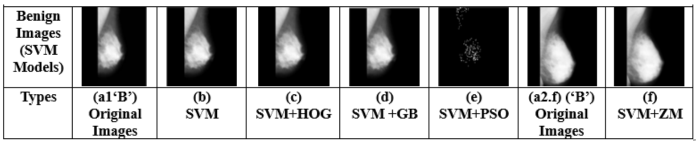

Fig.2. Benign for Original and Processed Images of SVM Models as types (a1'B'-f)

Mammogram images using SVM classifier about initial and processed images displayed in the Fig.2) as (a1'B') referred to by B represent benign original visuals along with processed visuals in SVM demonstrated the modified visuals as (b) SVM as well as (c) demonstrated that images when SVM+HOG for feature extraction methods towards histogram of oriented gradients as variance with processed images. The Fig.2(d) take part SVM+GB exhibited the processed images of enhanced gradients during optimization regarding both methods as well another one PSO provided the SVM+PSO processed images. (a2. f) ('B') an additional benign image on selects at random to taken images that as SVM+ZM we feature extraction of Mahotas all of these pictures have been processed quality of data in PSNR, SNR and MSE.

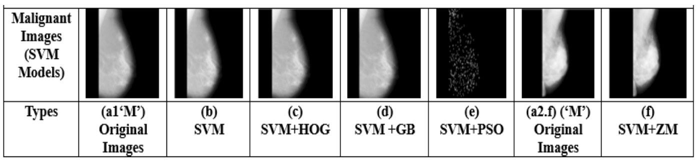

Fig.3. Malignant for Original and Processed Images of SVM Models as types (a1'M'-f)

Mammogram images using SVM classifier an initial and processed images seem in the Fig.3(a1'M') defined by M indicate malignant original visuals together with processed was visuals in SVM show modified pictures as Fig.3(b) SVM as well as Fig.3(c) shown that images if SVM+HOG over feature extraction methods towards histogram of oriented gradients as variance with processed images.

The Fig.3(d) demarcates a component SVM+GB showcased the generated images on boosted a gradient at optimization in terms of each technique besides another PSO specified the SVM+PSO as in Fig.3(e) shows the analysed images.

The Fig.3(a2.f) ('M') an additional malignant image on determines at random to be taken images whereby SVM+ZM that we feature extraction from Mahotas each of these images include the processed quality as measured through PSNR, SNR and MSE.

The experimental setup is open source to anaconda using implementation by the feature extraction methods to evaluate the quality of image are proven better outcomes of mammogram image enhancement. The experimental setup is open source to anaconda using implementation by the feature extraction methods to evaluate the quality of image are proven better outcomes of mammogram image enhancement.

5. RESULT AND DISCUSSIONS

Table.1. Overall accuracy and metrics in mammogram images classify as benign and malignant of SVM Models

All Model in Gaussian Filter	Mammogram Images	PSNR	SNR	MSE	Overall Accuracy
SVM	Benign	51.89	19.38	0.42	72%
	Malignant	53.98	22.88	0.26	
SVM + HOG	Benign	52.21	19.86	0.39	75.26%
	Malignant	54.49	23.36	0.23	
SVM + Gradient Boosting	Benign	52.34	19.99	0.38	77%
	Malignant	54.72	23.59	0.22	
SVM + PSO	Benign	59.26	0.29	0.08	82%
	Malignant	56.28	0.30	0.15	
SVM + Zernike Mahotas	Benign	54.95	23.38	0.21	84.54%
	Malignant	56.48	23.42	0.15	

The Table.1 shows the mammogram images into benign and malignant comparative the accuracy of several SVM algorithms. The model using SVM + Zernike Mahotas has the better results of 84.54%, with SVM + PSO (RBF) PCA produced with 82%. In the models using gaussian noise edge identification and SVM + HOG had lower accuracy levels, they performed well in classification.

The Fig.4 shows how contrasting several SVM models for PSNR leads to an analysis between enhanced SVM+PSO leads to entirely. The Fig.5 show that given below as methods are implemented to better outcomes for gradient boosting ensemble models and zernike mahotas with better image quality of SNR results performance in malignant identification.

Fig.4. PSNR in SVM Models

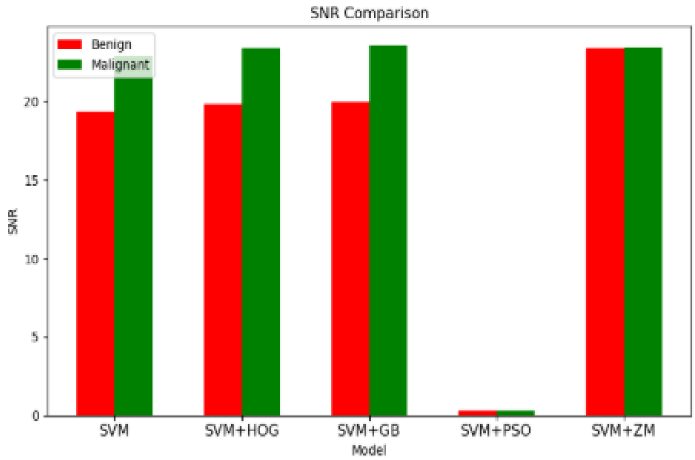

Fig.5. SNR in SVM Models

As in Fig.6, performance of mean square error is the least performance is better outcomes achieved by SVM+PSO is better image enhancement of minimized error get the best result from mammogram images.

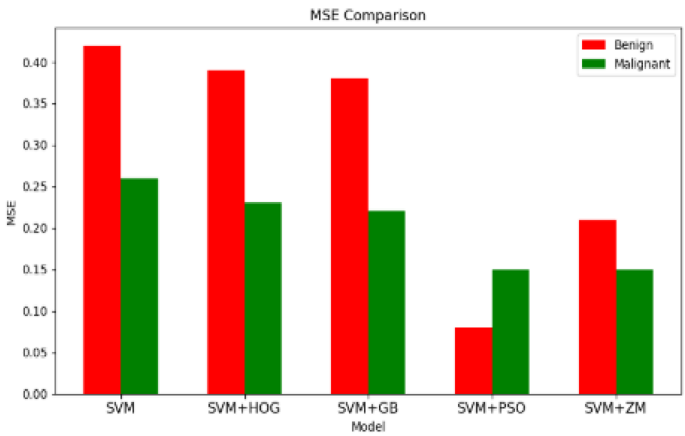

Fig.6. MSE in SVM Models

Performance evaluation of support vector machine is a machine learning algorithm of linear and non-linear classification with processing by mammogram images classified as benign and malignant visuals by gaussian filters. And then, five SVM models performed with PSNR, SNR and MSE metrics are image enhanced and processed with the removal of noise to minimize the error.

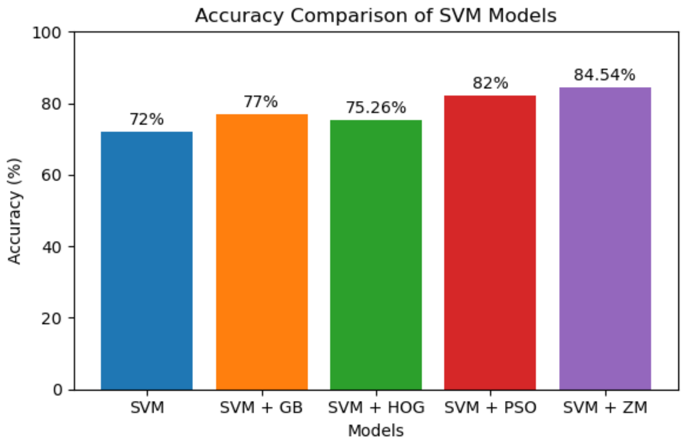

Fig.7. Illustration with the overall accuracy of SVM models in mammogram images

Finally, SVM classification of accuracy as shown in Fig.7, shows the overall performance of accuracy. Comparison of models as five categories with SVM, SVM+GB, SVM+HOG, SVM+PSO and SVM+ZM. Producing one is SVM+Zernike Mahotas (ZM) for feature extraction with better outcomes than other comparison algorithms evaluated.

6. CONCLUSION

Medical image analysis is a key term in wellness consciousness for prevention and is better than cure for breast cancer identification with image enhancement besides with classification methods as the benchmark algorithm for both linear and non-linear classification to support vector machine model. Gaussian filters that use all optimization and model based SVM removal of features aid as image processing techniques. Metrics such as PSNR, SNR and MSE show that peak signal to noise ratios took 40–50 dB produce better outcomes. The Fig.4 seen that PSNR metrics have ranges and that the higher quality of SVM models as well as SNR also result in improved results.

To normalization method processing was completed with the perversion ratio of the gaussian image filters was improved as had been the general accuracy of the kernel-based SVM classifier of linear and radial basis function. So, PSO has too little SNR prior to MSE for all models reducing error values.

Optimization is gradient boosting and particle swarm optimization methods. The histogram of oriented gradients and zernike mahotas are the feature extraction methods of these approaches are processed with a classifier using SVM. Then, second approach is particle swarm optimization for also higher levels of PSNR when compared to other methods and accuracy is 82% of SVM models. To measure of overall precision is the most effective method as feature extraction from zernike mahotas had an accuracy of 84.54% in these models.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre and A. Jemal, "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries", *CA: A Cancer Journal for Clinicians*, Vol. 68, No. 6, pp. 394-424, 2018.

[2] R.L. Siegel, K.D. Miller and A. Jemal, "Cancer Statistics, 2020", *CA: A Cancer Journal for Clinicians*, Vol. 70, No. 1, pp. 7-30, 2020.

[3] M. Cao, H. Li and D. Sun, "Cancer Burden in China: An Overview of Trends and Projections", *Cancer Communications*, Vol. 41, No. 6, pp. 660-668, 2021.

[4] T.H. Harvard Chan School of Public Health, "Student Handbook", Available at: https://content.sph.harvard.edu/wwwhsph/sites/76/2019/12/Handbook-for-2019-20204.pdf, Accessed in 2020.

[5] P. Kaur, M. Kaur and S. Singh, "A Computer-Assisted Approach for Breast Cancer Diagnosis using Mammography Images: SVM-based Classification and Image Preprocessing Techniques", *Journal of Digital Imaging*, Vol. 31, No. 5, pp. 738-750, 2018.

[6] V. Jeevitha and I.L. Aroquiaraj, "Mammogram Images using Noise Removal of Filtering Techniques," *International Research Journal on Advanced Engineering Hub*, Vol. 2, No. 2, pp. 118-123, 2024.

[7] V. Jeevitha and I.L. Aroquiaraj, "Novel Medical Image Analysis for Breast Cancer Identification", *Proceedings of International Conference on Latest Advancements in Commerce, Management Science and Educational Research*, pp. 1-8, India.

[8] E. Zhang, F. Wang, Y. Li and X. Bai, "Automatic Detection of Microcalcifications using Mathematical Morphology and a Support Vector Machine", *Bio-Medical Materials and Engineering*, Vol. 24, No. 1, pp. 53-59, 2014.

[9] A.A. Kayode, N.O. Akande, A.A. Adegun and M.O. Adebiyi, "An Automated Mammogram Classification System using Modified Support Vector Machine", *Medical Devices: Evidence and Research*, Vol. 12, pp. 275-284, 2019.

[10] Y.B. Youssef, E. Abdelmounim and A. Belaguid, "Mammogram Classification using Support Vector Machine", *Proceedings of International Conference on Advances in Artificial Intelligence and Data Engineering*, pp. 894-921, 2020.

[11] D.A. Zebari, D.A. Ibrahim, D.Q. Zeebaree, M.A. Mohammed, H. Haron and N.A. Zebari, "Breast Cancer Detection using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion", *Applied Sciences*, Vol. 11, No. 24, pp. 12122-12137, 2021.

[12] A. Muthukumaravel, "Efficient Feature Extraction on Mammogram Images using Enhanced Grey Level Co-Occurrence Matrix", *International Journal of Intelligent Engineering Informatics*, Vol. 11, No. 1, pp. 1-12, 2023.

[13] A. Ponraj and A. Canessane, "Radial Basis Function Networks and Contrast-Limited Adaptive Histogram Equalization Filter based Early-Stage Breast Cancer Detection Techniques", *Journal of Computer Science*, Vol. 19, No. 6, pp. 760-774, 2023.

[14] A. Bilal, A. Imran, T.I. Baig, X. Liu, E.A. Nasr and H. Long, "Breast Cancer Diagnosis using Support Vector Machine Optimized by Improved Quantum Inspired Grey Wolf Optimization", *Scientific Reports*, Vol. 14, No. 1, pp. 12122-12135, 2024.

[15] B.A. Maqbali, "Hybrid Wolf Pack Algorithm and Particle Swarm Optimization Algorithm for Breast Cancer Diagnosis: Hybrid WPA and PSO Algorithm", *Machine Reviews*, Vol. 3, No. 2, pp. 45-53, 2021.