Tumor detection in mammography using a combination of texture and local features

Abstract

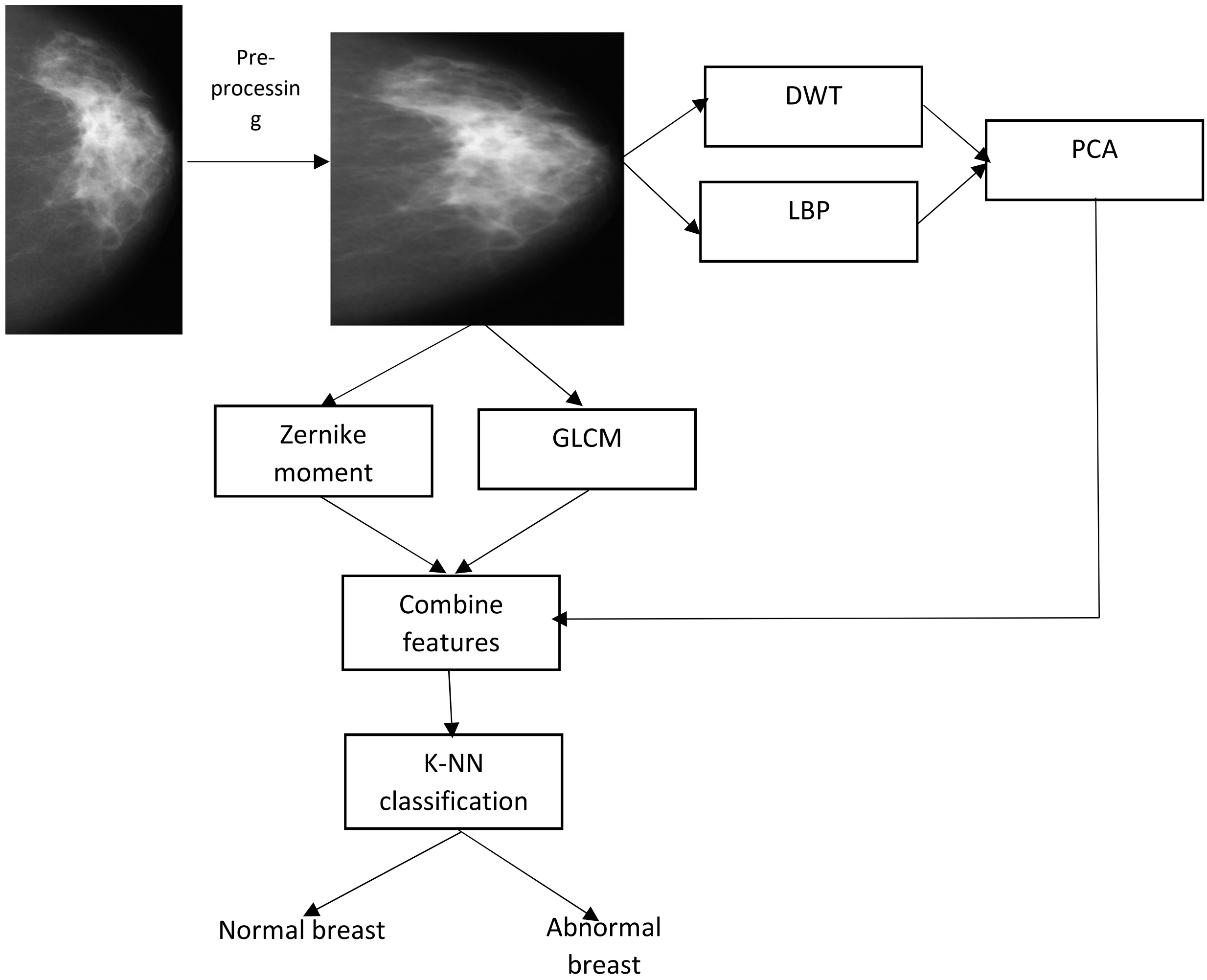
In this research, DWT combination and local binary pattern (LBP) as well as Zernike moment combination and gray level co-occurrence matrix (GLCM) are used for feature extraction.

Also, this method works based on local features and local binary features. The most important feature of the features of local binary patterns is their changes in contrast to the brightness and simplicity of their calculations. Compared to the Gabor filter, the features of local binary patterns are much faster. DWT decomposes images into different frequency components at multiple resolutions, enabling detailed analysis in medical imaging. It provides spatial and frequency information, making it easier to detect localized anomalies. Unlike Fourier Transform, DWT separates noise from important features, ensuring high precision in medical imaging. It extracts essential features like edges, textures, and shapes for accurate analysis and classification. The advantage of using GLCM simple histogram matrix is that, compared to the simple histogram, where the spatial information of the pixels is lost and only the frequency of the gray values of the pixels is calculated, in this matrix the spatial position of pixels are also considered . So that the wider the distribution of gray values is, the more variance will be seen in the matrix.

Zernike moments are effective for feature extraction in mammogram images because they provide rotation-invariant features, ensuring consistency regardless of image orientation. They are robust to noise, capturing significant image details while minimizing the impact of noise. Zernike moments excel in representing the shape of objects, which is crucial for identifying tumors and other anomalies. Their orthogonality ensures a compact and non-redundant feature set, and their ability to capture both global and local features enhances the performance of machine learning algorithms in computer-aided diagnosis systems. These properties make Zernike moments highly suitable for accurate and detailed analysis of mammograms.

In the next step, these two feature vectors (the feature vector obtained by applying the LBP and the DWT feature vector) are combined into one vector and finally a feature vector is obtained. This feature vector is fed to the Principal Component Analysis (PCA) algorithm to select the best feature. Finally, this feature vector is combined with GLCM and Zernike moment feature vectors and the final feature vector is obtained. Then the extracted features are entered into K-NN classification. The tests have been evaluated with sensitivity, specificity and accuracy parameters. The results of the tests on MIAS and DDSM database have 84.47% and 82.50% recognition accuracy, respectively.

1- Proposed method



1-1- Pre-processing breast images

- Convert image2grayscale
- Resizing to 250×250

1-2- Feature extraction using combination of local and texture features

- Discrete wavelete transform (DWT)

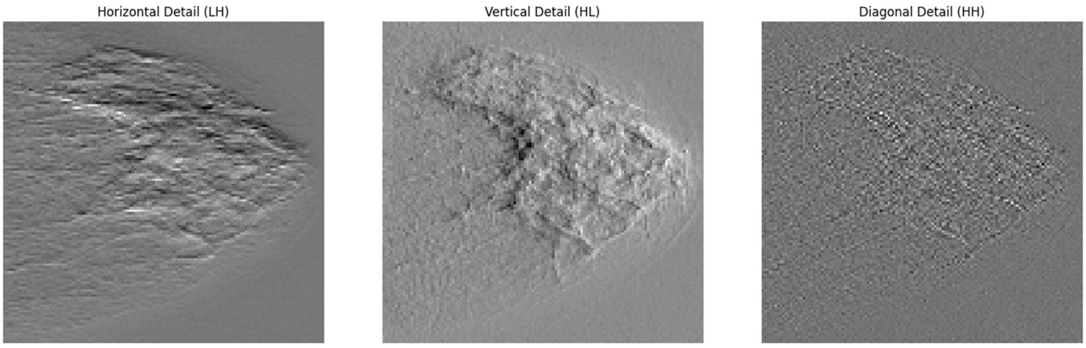


Figure 1. Output of DWT

- Local binary pattern (LBP)

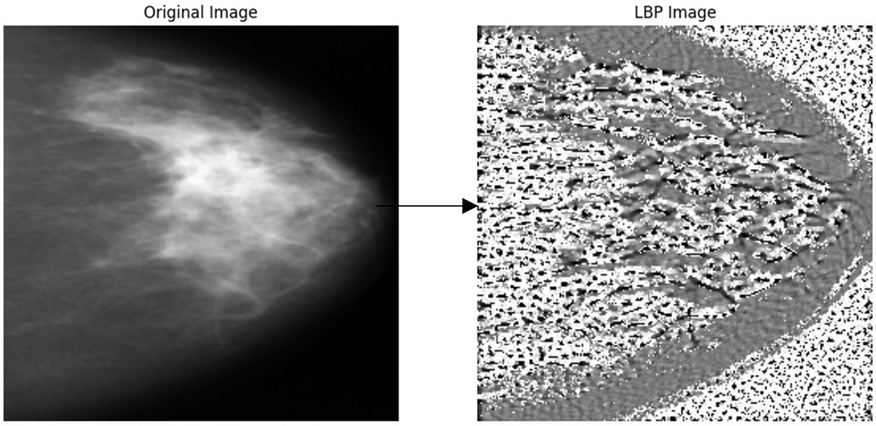


Figure 2. Output of LBP

- Zernike moment
- Gray level co-occurrence matrix (GLCM)

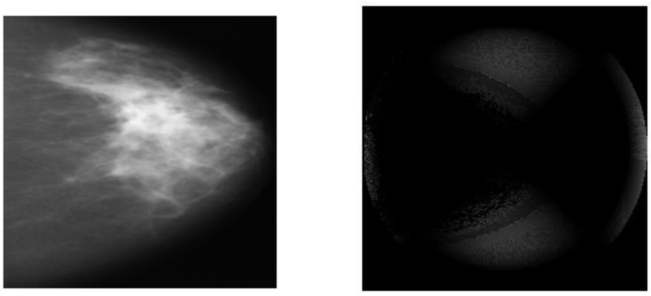


Figure 3. Output of Zernike moment

1-3- Feature selection using principal component analysis (PCA)
1-4- Feature classification using K-nearest neighbour (K-NN)

2- Experimental Result

- Dataset: MIAS (322 images): https://www.kaggle.com/datasets/kmader/mias-mammography
- DDSM : (1000 images) https://www.kaggle.com/datasets/skooch/ddsm-mammography

Table 1. Evaluation results of the proposed method for the MIAS database

Abnormal =115 images		**Normal images=207**	
True Negative (TN)	False Positive (FP)	True positive (TP)	False Negative (FN)
97	18	175	32

Table 2. Evaluation results of the proposed method for the DDSM database

Abnormal =400 images		**Normal images=200**	
True Negative (TN)	False Positive (FP)	True positive (TP)	False Negative (FN)
330	70	165	35

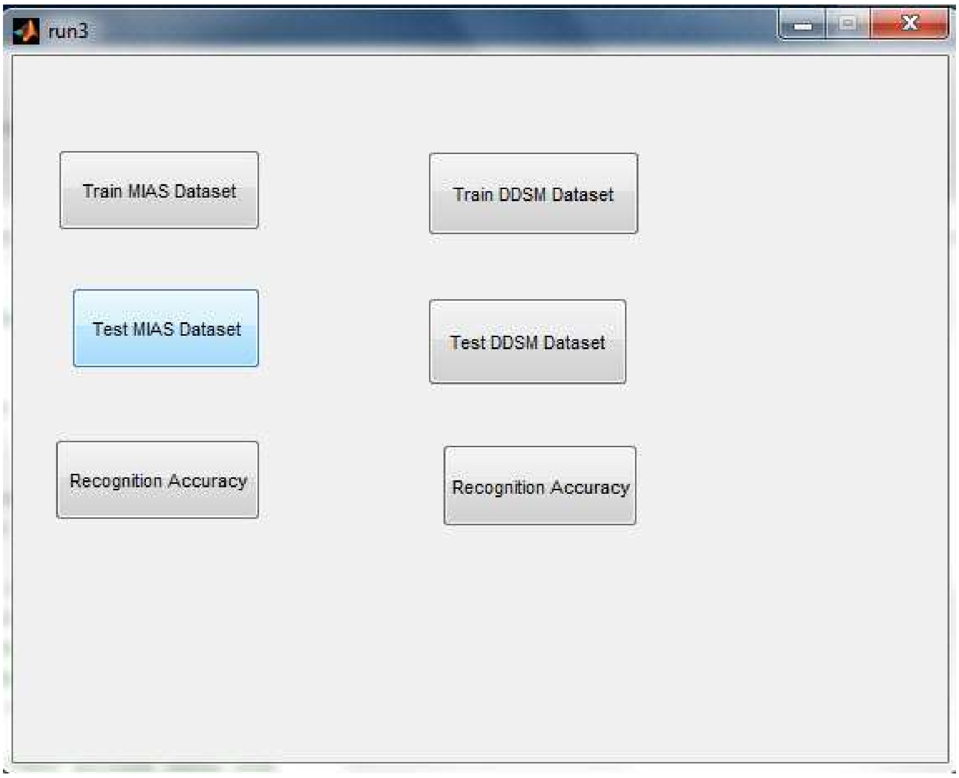
Table 3. Comparison of accuracy, precision and recall metrics of DDSM and MIAS database

	Sensitivity	specificity	accuracy
DDSM dataset	82.50%	82.50%	82.50%
MIAS dataset	84.54%	84.34%	84.47%

3- How to run:

- Open run.m file and click run button
- First click train MIAS dataset button
- Then click test MIAS dataset
- First click train DDSM dataset button
- Then click test DDSM dataset

You can see the accuracy of test data on MIAS and DDSM datasets.


run3
Train MIAS Dataset
Train DDSM Dataset
Test MIAS Dataset
Test DDSM Dataset
Recognition Accuracy
Recognition Accuracy