Transactions on Machine Intelligence 2(4) (2019) 213-220

ISSN Online: 2821-1693

Transactions on Machine Intelligence

Journal Homepage: https://tmachineintelligence.ir/index.php/TMI

# Breast Cancer Diagnosis Using Scattering Wavelet Transform and Hierarchical Multilayer Perceptron Neural Network

M. Nouri [1], N. Nematbakhsh [2,*], S. Farrokhi [3]

[1] Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Isfahan, Iran
[2] Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Isfahan, Iran
[3] Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Isfahan, Iran

**ARTICLE INFO**

Article History:
Received 10 July 2019
Received in revised form 14 October 2019
Accepted 5 December 2019
Available online 11 December 2019

Keywords:
Breast Cancer, Scattering Wavelet Transform, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Hierarchical Classification

**ABSTRACT**

Breast cancer has been one of the leading causes of mortality among women in the past decade. Although this type of cancer cannot be prevented due to the unknown nature of its primary causes, early diagnosis can significantly improve a patient's chances of full recovery. Mammography is a well-established tool that aids in the early detection of this disease. Various studies have been conducted to develop breast cancer detection methods; however, these efforts have often failed to achieve sufficient accuracy due to the lack of an effective feature extraction method capable of capturing essential texture characteristics and the absence of a robust classifier. In this study, scattering wavelet transform is employed to extract texture-based features from medical images. The use of multiple features increases the dimensionality of input data for the classifier, necessitating an effective dimensionality reduction approach. To address this, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) have been applied. Finally, a hierarchical multilayer perceptron (MLP) neural network is utilized as the classifier for cancer detection. To evaluate the proposed method, the Mini-MIAS dataset has been used, achieving an accuracy of 97.57%.

## 1. INTRODUCTION

Breast cancer is the oldest known type of cancer in humans, with its earliest symptoms observed in Egypt around 1600 BCE. Despite extensive efforts in the diagnosis and treatment of breast cancer, it remains a leading cause of mortality. Various methods, including image processing, have been developed to facilitate early detection of this disease [1].

Timely diagnosis and appropriate treatment significantly enhance patient recovery and survival rates. While conventional cancer detection relies on invasive methods such as surgery, radiology, and chemotherapy, studies indicate that modern computational technologies particularly image processing mechanisms have proven effective in cancer diagnosis and classification [2]. Given the diverse nature of medical images, specific preprocessing techniques are required before diagnosis. Additionally, various imaging devices capture different anatomical

* Corresponding Author: nemat@eng.ui.ac.ir
Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Isfahan, Iran



http://dx.doi.org/10.47176/TMI.2019.213

regions, necessitating distinct diagnostic processing techniques [3]. This study reviews existing methods and previous research, analyzing their advantages and limitations.

## 2. LITERATURE REVIEW

Several approaches for breast cancer detection have been introduced in recent years. Below, we discuss some of these methods.

In 2016, Abolghasem et al. proposed a simple method for breast cancer detection using digital mammography images. Their approach consists of three primary steps: segmentation of the breast region, removal of the pectoral muscle, and classification of the breast tissue into cancerous and non-cancerous categories. The pectoral muscle segmentation was performed using the Otsu thresholding technique, followed by pectoral muscle removal through arc edge detection and an approximate straight-line technique. Next, gray-level co-occurrence matrices (GLCM) were employed for feature extraction, and a support vector machine (SVM) classifier was trained to distinguish between cancerous and non-cancerous tissues. The proposed method was evaluated using the Mini-MIAS dataset [4].

In 2014, Elnaz Alafati and colleagues developed a breast cancer detection system that utilizes Principal Component Analysis (PCA) for feature dimension reduction, a genetic algorithm for feature selection, and an SVM classifier for classification. However, a drawback of this approach is the genetic algorithm's inherent tendency to perform a global search while exhibiting weak local search capabilities. The proposed method was tested using the Wisconsin Breast Cancer Dataset (WBCD) [5].

Ranjit Biswas et al. employed various preprocessing techniques, including artifact removal for region-of-interest extraction, two-dimensional median filtering for noise reduction, and Contrast-Limited Adaptive Histogram Equalization (CLAHE) for image enhancement. For feature extraction, they utilized GLCM, and classification was performed using k-nearest neighbors (KNN), SVM, and artificial neural networks (ANN). Their proposed approach was evaluated on the MIAS dataset [6].

Seryasat and colleagues have proposed several approaches in four distinct studies related to breast cancer detection. In the first study, a novel ensemble learning framework for classifying breast masses in mammograms was introduced. This method integrates multiple machine learning algorithms to analyze texture and intensity features, thereby improving diagnostic accuracy. The results indicate that this approach achieves higher accuracy compared to traditional methods [7].

The second study presents a new method for classifying and analyzing breast cancer tumors. By leveraging image processing and texture feature analysis, the method effectively distinguishes between benign and malignant masses. The findings suggest that this approach can enhance diagnostic precision and aid medical decision-making [8].

In the third paper, a novel computer-aided diagnosis (CAD) system for detecting breast masses in mammographic images is evaluated. The method utilizes image processing and machine learning techniques to extract precise features. The system's performance is compared to conventional techniques, demonstrating high sensitivity and accuracy in early breast cancer detection [9].

The fourth study proposes an efficient method for identifying breast masses in mammograms. This method employs advanced image processing algorithms to enhance accuracy and reduce false positive rates. The results indicate that this approach can serve as a valuable tool for improving the reliability of breast cancer diagnosis [10].

Rajani and colleagues introduced a novel method for tumor segmentation, which involves multiple stages:

- Enhancing mammographic image quality through techniques such as filtering and discrete wavelet transform (DWT).
- Isolating the region suspected of containing a tumor.
- Extracting features from the segmented region.
- Classifying the extracted features using a support vector machine (SVM).

This method was tested on 75 mammographic images from the MINI-MIAS database, achieving a sensitivity of 88.75% [11].

Although various studies have been conducted in this domain, they often lack sufficient accuracy. Given the critical importance of precision in breast cancer diagnosis, this research aims to enhance detection accuracy compared to previous studies. The proposed approach utilizes digital mammography images for breast cancer identification, employing texture-based features such as the scattering wavelet transform.

The use of multiple features increases the dimensionality of input data, making classification more challenging. Therefore, it is essential to reduce feature dimensions in a manner that preserves classification accuracy. To achieve this, after feature extraction, dimensionality reduction techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) will be applied. Finally, a hierarchical multilayer perceptron (MLP) neural network classifier will be used to improve accuracy compared to previous methods.

## 3. SCATTERING WAVELET TRANSFORM

A locally invariant image descriptor, similar to scale-invariant feature transforms, provides a more effective representation for image classification. This feature vector, along with a multi-scale descriptor based on the mean distances of wavelet coefficient domains, is computed accordingly [12]. The mean reduces feature variability while ensuring local translation invariance, although it also decreases information content.

The scattering operator enhances the lost high-frequency components and redistributes them into coefficients across multiple scales and orientations. These coefficients transform small deformations into longitudinal structures. They are computed using a convolutional network that arranges contractive wavelet transforms and modulus operators in sequence. Scattering operators generate a novel representation of texture features, enabling the differentiation of textures with similar spectral power [13]. Expressing these feature vectors in terms of wavelet coefficients aids in better comprehension and enhances feature completeness.

If $R_x x$ represents rotation in $x \in \mathbb{R}^2$ by an angle $\gamma$, the directional wavelets are obtained by rotating a wavelet $\psi$, in length $k$ and angle ($\gamma \in \Gamma$). By scaling them with $2^j$, we obtain:

$$\psi_{j,\gamma} = 2^{-2j}\psi\left(2^{-j}R_\gamma x\right) \tag{1}$$

The directional wavelet transform $f$ at position $x$ for scales $2^j \leq 2^J$ is represented by a vector with the following coefficients:

$$W_J f(x) = \begin{pmatrix} f * \psi_{j,\gamma}(x) \\ f * \phi_J(x) \end{pmatrix}_{j \leq J,\, \gamma \in \Gamma} \tag{2}$$

When $\phi(x) = 2^{-2J}\phi(2^{-J}x)$ is a low-pass filter that captures frequencies of $f$ higher than the scale $2^J$,

$$\int \phi(x)dx = 1 \tag{3}$$

The scattering operators recover part of the lost information through averaging, using simultaneous coefficients with similar deviation characteristics. The wavelet transform (2) shows that the high frequency in $|f \times \psi_{j1,\gamma1}| \times \phi_J$ is removed by twisting with $\phi_J$, and is enhanced by twisting with wavelets $|f \times \psi_{j1,\gamma1}| \times \psi_{j2,\gamma2}$ at scales $2^{j2} \leq 2^J$.

## 4. PROPOSED METHOD

The general outline of the proposed method is shown in Figure 1, which includes preprocessing, segmentation, feature extraction, and classification. This section provides a summary of each of these steps.

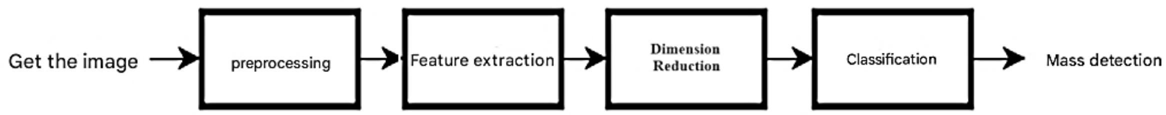
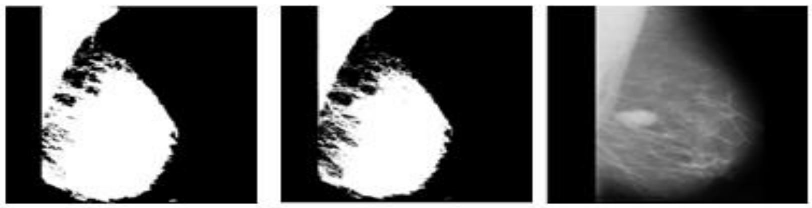



**Fig. 1.** Image processing techniques used for digital mammography images [14]

### 4.1. Preprocessing

In the first preprocessing step, the color image is converted to a grayscale image. Next, the breast region needs to be extracted, and any additional regions in the image should be removed as much as possible. To do this, thresholding is applied, and after thresholding, due to the varying textures in individuals and the differences in intensity levels within the breast region, some cavities may appear. To smooth and unify the region and create a connected component for the breast area, morphological dilation operations are used. A sample of the preprocessing operations is shown in the figure [15]. Finally, to enhance the tissue differentiation, histogram equalization is applied to expand the intensity levels of the image across the entire range, improving contrast [16]. Afterward, the images are resized to a uniform size of 150×200 pixels based on the different dimensions after selecting the breast region.



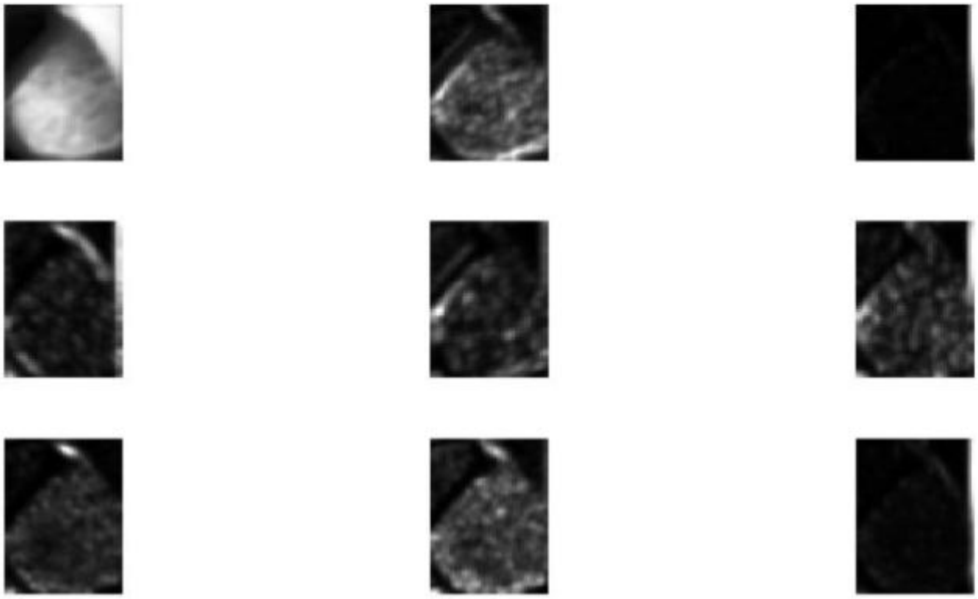
c b a

**Fig. 2**. Preprocessing, a) Input grayscale image, b) Thresholding applied, c) Morphological dilation.

### 4.2. Feature Extraction

In this section, the Discrete Wavelet Transform (DWT) is used for feature extraction. A sample of the DWT output images is shown in Figure 2. The parameters considered for the DWT, for scale and orientation, are 3 and 6, respectively, and the transform progresses up to two stages. The input image size for the wavelet extractor is initially 200×150 pixels. In the first stage, the image size is reduced to 50×35 pixels, and in the second stage, it is resized to 33×50 pixels.

(a)



(b)

**Fig. 3**. Discrete Wavelet Transform: a) Input image, b) Examples of DWT output images.

### 4.3. Dimensionality Reduction

The use of multiple features, considering the need for high-quality images, increases the input data volume for the classifier, making the classification process more challenging. Therefore, it is necessary to reduce the dimensions of the features in an appropriate manner. To achieve this, after feature extraction, dimensionality reduction techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) will be used to improve classification accuracy [18, 17].

### 4.4. Classification

After extracting the breast region and reducing dimensions, two independent Multi-Layer Perceptron (MLP) neural networks are used hierarchically to improve classification accuracy [19]. Initially, the reduced features are fed into the first neural network, which makes a decision about the input. The output of this neural network at this stage consists of two classes: either normal or diagnosed as cancer. This network does not comment on the type of cancer. If a cancerous sample is detected, it is passed to the second neural network, which determines whether the cancer is benign or malignant. For images that are not diagnosed as cancerous, the proposed method does not utilize the second neural network. Figure 3 shows the hierarchical process.

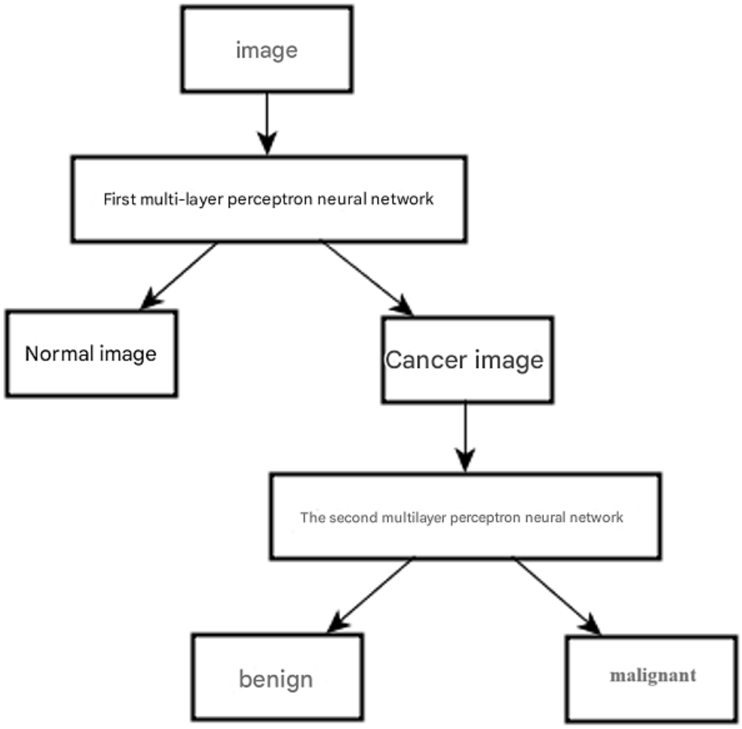



**Fig. 3**. Hierarchical process of the proposed method.

### 4.5. Evaluation of the Proposed Model and Comparison with Other Methods

To evaluate the proposed method, the MIAS dataset, which contains 322 images, is divided into three parts: 30% for training, 20% for validation, and 50% for testing. The dataset is divided into training and testing sets, and comparisons are made with the methods in [6] and [19]. The evaluation is based on the recognition rate, which is defined as follows:

$$Recognition\ Rate = \frac{(Number\ of\ Correctly\ Identified\ Samples)}{Total\ Number\ of\ Samples} \times 100 \quad (4)$$

The parameters chosen for the Discrete Wavelet Transform (DWT) for scale and direction are 3 and 6, respectively. The number of neurons in the first and second hidden layers of each Perceptron neural network is 15 and 10, respectively. The achieved accuracy for the image set is 98.13%, indicating correct output selection for the three classes: normal, benign, and malignant. Table 3 compares the accuracy of the proposed method with other methods. Previous methods failed to achieve high accuracy due to the lack of an effective feature extraction technique and the absence of a hierarchical classifier. Additionally, methods [6, 4] did not address the three-output class issue.

**Table 3**. Comparison of the proposed method with other methods.

| Method name | Accuracy |
|---|---|
| The method used in [ ۶ ] | ۹۵٪ |
| The method used in [ ۴ ] | ۹۶/۵۵٪ |
| Suggested method | ۹۸/۱۳٪ |

## 5. CONCLUSION

In this study, an accurate method for breast cancer detection using mammography images has been proposed. This method utilizes the Discrete Wavelet Transform (DWT) for feature extraction, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) algorithms for improving classifier accuracy, and a multilayer Perceptron neural network in a hierarchical manner for breast cancer detection. By testing this method on mammography images, an accuracy of 97.57% was achieved. Based on the obtained results and comparison with conventional methods, it can be concluded that the proposed method, due to the choice of an appropriate feature extractor and hierarchical classifier, performs exceptionally well in distinguishing between benign and malignant masses.

### Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

### Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

### Declaration of Interest

The authors declare that they have no competing interests.

### Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

## REFERENCES

[1] R. M., Ayres, F. J., & Desautels, J. E. L. (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute, 344(3), 312–348. https://doi.org/10.1016/j.jfranklin.2006.09.003

[2] Glotsos, G., Kalatzis, D., Theocharakis, P., Daskalakis, A., & Ninos, K. (2010). A multi-classifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images. Computer Methods and Programs in Biomedicine, 97(1), 53–61. https://doi.org/10.1016/j.cmpb.2009.07.003

[3] Yasmin, M., Sharif, M., & Mohsin, S. (2013). Survey paper on diagnosis of breast cancer using image processing techniques. Research Journal of Recent Sciences, 2(10), 88–98.

[4] Qayyum, A., & Basit, A. (2016). Automatic breast segmentation and cancer detection via SVM in mammograms. In International Conference on Emerging Technologies (ICET) (pp. 1–6). Islamabad, Pakistan. https://doi.org/10.1109/ICET.2016.7813261

[5] Olfati, E., Zarabadipour, H., & Shoorehdeli, M. A. (2014). Feature subset selection and parameters optimization for support vector machine in breast cancer diagnosis. In 2014 Iran Conference on Intelligent Systems (pp. 1–6). Bam, Iran. https://doi.org/10.1109/IranianCIS.2014.6802601

[6] Biswas, R., Nath, A., & Roy, S. (2016). Mammogram classification using gray-level co-occurrence matrix for diagnosis of breast cancer. In International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE) (pp. 161–166). Ghaziabad, India. https://doi.org/10.1109/ICMETE.2016.85

[7] Seryasat, O. R., & Haddadnia, J. (2018). Evaluation of a new ensemble learning framework for mass classification in mammograms. Clinical Breast Cancer, 18(3), e407–e420. https://doi.org/10.1016/j.clbc.2017.05.009

[8] Rahmani-Seryasat, O., Haddadnia, J., & Ghayoumi-Zadeh, H. (2015). A new method to classify breast cancer tumors and their fractionation. Ciência e Natura, 37(4), 51–57. https://doi.org/10.5902/2179460X19428

[9] Rahmani-Seryasat, O., Haddadnia, J., & Ghayoumi-Zadeh, H. (2016). Assessment of a novel computer-aided mass diagnosis system in mammograms. Iranian Journal of Breast Diseases, 9(3), 31–41.

[10] Haddadnia, J., Rahmani-Seryasat, O., Ghayoumi-Zadeh, H., & Rabiee, H. (2015). An efficient method for detection of masses in mammogram images. Cumhuriyet Üniversitesi Fen Edebiyat Fakültesi Fen Bilimleri Dergisi, 36(3), 2269–2277.

[11] Ireaneus, Y. A., & Rejani, S. (2009). Early detection of breast cancer using SVM classifier. Computer Science and Engineering, 1(3), 127–130.

[12] Bruna, J., & Mallat, S. (2013). Invariant scattering convolution networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1872–1886. https://doi.org/10.1109/TPAMI.2012.230

[13] Bruna, J., & Mallat, S. (2013). Classification with scattering operators. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1872–1886. https://doi.org/10.1109/TPAMI.2012.230

[14] Burshi, P. B., & Kulkarni, D. A. (2016). Digital mammography: A review on detection of breast cancer. 5(1), 386–390.

[15] Sivarajan, U., Jayapragasam, K., Aziz, A., Rahmat, K., & Bux, S. (2009). Dynamic contrast enhancement magnetic resonance imaging evaluation of breast lesions: A morphological and quantitative analysis. JHK Coll Radiol, 12, 43–52.

[16] Liu, Y., Cheng, H., Huang, J., Zhang, Y., Tang, X., Tian, J.-W., & Wang, Y. (2012). Computer-aided diagnosis system for breast cancer based on color Doppler flow imaging. Journal of Medical, 36(6), 3975–3982. https://doi.org/10.1007/s10916-012-9869-4

[17] Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

[18] Martinez, A. M., & Kak, A. C. (2001). PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2), 228–233. https://doi.org/10.1109/34.908974

[19] Isa, I., Saad, Z., Omar, S., Osman, M., Ahmad, K., & Sakim, H. M. (2010). Suitable MLP network activation functions for breast cancer and thyroid disease detection. In International Conference on Computational Intelligence, Modelling and Simulation (CIMSiM) (pp. 30–44). Tuban, Indonesia. https://doi.org/10.1109/CIMSiM.2010.93