SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I5P114

Volume 11 Issue 5, 130-138, May 2024

Original Article

A Deep Learning Approach for Efficient Breast Cancer Diagnosis Using Hybrid CNN-BILSTM with Soft Attention Mechanism

M. Sarathkumar[1*], K. S. Dhanalakshmi[2], D. Madhivadhani[3], Revathi V[4]

[1]*University Science Instrumentation Centre (USIC), Madurai Kamaraj University, Tamil Nadu, India.*
[2]*Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Tamilnadu, India.*
[3]*Department of Electronics and Communication Engineering, Kings Engineering College, Tamilnadu, India.*
[4]*Department of Applied Sciences, New Horizon College of Engineering, Karnataka, India.*

[*]*Corresponding Author : msksarathkumar01@gmail.com*

Received: 17 March 2024 Revised: 21 April 2024 Accepted: 13 May 2024 Published: 31 May 2024

Abstract *- In recent years, the rising prevalence of breast cancer among women has underscored the critical importance of timely and accurate disease prediction. The ability to predict breast cancer efficiently empowers healthcare professionals with a robust decision-making system, enabling optimal treatment strategies. This study presents a sophisticated diagnostic framework, integrating various techniques from preprocessing to classification to enhance the precision and effectiveness of breast cancer detection. The proposed methodology begins with preprocessing using a Butterworth filter to enhance the quality of input data. The preprocessed output is segmented utilizing Cascaded Fuzzy C-Means segmentation. Subsequently, extracting features through the (GLCM), capturing intricate patterns and texture information crucial for discriminating between benign and malignant tissues. The classification step employs a Hybrid CNN- BiLSTM architecture augmented by a Soft Attention Mechanism. This hybrid model is designed to effectively leverage spatial hierarchies and sequential dependencies in medical images, allowing for a more comprehensive analysis of complex forms related to breast cancer. Soft Attention Mechanism enhances the interpretability of the model by assigning varying weights to different regions of the input data, emphasizing salient features crucial for exact analysis. The attention mechanism donates to the model's ability to focus on relevant information, improving both sensitivity and specificity in breast cancer classification. The proposed methodology is validated through extensive experiments, demonstrating its superior performance compared to traditional approaches. The results highlight the possibility of this deep learning-based strategy valuable enabling healthcare professionals, facilitating precise and timely decision-making in breast cancer treatment optimization.*

Keywords *- Breast cancer, Butterworth filter, Cascaded Fuzzy C-Means, Gray-Level Co-occurrence Matrix (GLCM), Hybrid CNN-BiLSTM.*

1. Introduction

Women worldwide lose their lives to cancer, primarily due to breast cancer. Research indicates that prompt detection and management of this specific cancer can significantly lower the death rates from these crippling disease illnesses [1]. Breast cancer is more common in women with thick breasts, and younger women have denser breasts than older women, indicating a relationship between density and age [2].

Abnormal cell proliferation that permeates the surrounding tissues of the body is associated with malignant tumours. Tumors are categorized as either benign or malignant, and Non-cancerous cells that grow locally but do not spread across the body make up benign tumors. On the other hand, cancerous cells that can proliferate uncontrollably, spread throughout the body, and invade other tissues make up a malignant tumor [3, 4].

Prior to applying any image processing technology, preprocessing approaches are essential to improve outcomes. A variety of preprocessing techniques has addressed the problem of inaccurate diagnosis in mammography images. Many approaches are included in these algorithms, such as mean [5], median [6], Gaussian [7], and wiener [8]. These experiments demonstrated the effects of several preprocessing methods on classification performance to get around issues with mammography images, such as excessive noise and poor contrast [9, 10].

The segmentation approach divides an image into multiple portions or areas, removes or modifies its representation to something more appropriate, and simplifies the evaluation process. Despite being one of the easiest methods to calculate threshold pictures, the Otsu approach makes segmentation errors when the histogram stops being bimodal [11]. The learning technique requires that the number of cluster centers is prioritized, although K-means performs best when the data set is distinct. The best outcome for the segmentation procedure is obtained using the Fuzzy C means approach with regard to the overlap between the dataset and the data point [12, 13]. A crucial step in the diagnosis of breast cancer is feature extraction. It involves selecting and altering relevant data from medical images or other data sources to represent important tissue properties. Breast cancer classification models can detect cancer more accurately if effective feature extraction techniques are used [14].

Modern methods have been employed recently by medical imaging experts to address issues with breast cancer analysis. Numerous studies have been conducted about automated breast cancer exposure. Breast cancer accounts for around one out of every four recently diagnosed instances of cancer. According to the World Health Organization (WHO), it is the most prevalent illness diagnosed in women [15]. It describes in great detail the numerous classification methods used to group BC in this study. Deep Learning (DL)-based [16] Artificial Neural Network (ANN) techniques, traditional Machine Learning (ML) [17]-based Support Vector Machines (SVM) techniques, K-Nearest Neighbor (KNN) and Decision Tree (DT) algorithms are our main areas of interest [18-20]. However, various shortcomings arise, such as data bias, lack of transparency, false positives and negatives, and evolution of cancer characteristics. As a result, the proposed work establishes a hybrid approach for the effective classification of breast cancer. The main contributions of research involve,

- To eliminate noise in the image butter worth filter is used.
- To enrich the accuracy, Cascaded Fuzzy C-means segmentation is done.
- The feature extraction is used to reduce redundant data from the dataset using GLCM.
- The network architecture, as well as the volume and caliber of data, allow CNN-BiLSTM to diagnose breast cancer.

The following list illustrates the structure of the paper: The introduction is described in Section I. Section II provides a proposed part of this work. Part III gives the result and discussion of this work. The conclusion is provided in Section IV.

2. Proposed System Modelling

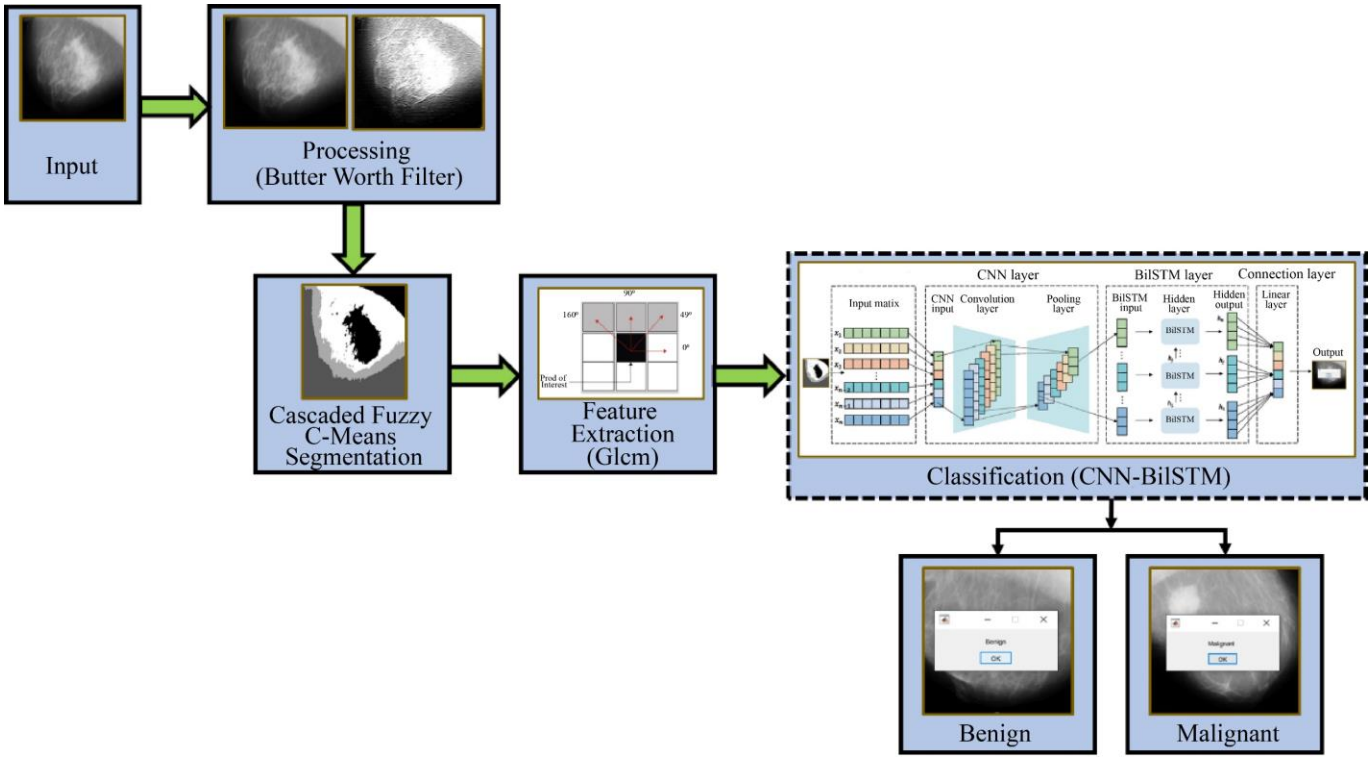

Fig. 1 Architecture of the proposed diagram

2.1. Preprocessing (Butterworth Filter)

Elimination of noise in the input image is the initial stage. The Butterworth Low Pass Filter (BLPF) is a tool used in image processing to smooth images in the frequency domain. It preserves low-frequency components and eliminates high-frequency noise from digital images. The definition of the BLPF of the order transfer function is:

$$H(u,v) = \frac{1}{1+\left(\frac{D(u,v)}{D_0}\right)^{2n}} \quad (1)$$

Where D the cut-off frequency, n is the filter's order, and $D(u,v)$ is the distance in Euclidean terms between any point (u,v) frequency plane origin. BLPF passes through all frequencies below the value without any attenuation, while all frequencies above the value are blocked. Cut-off frequency refers to this as the point at which $H(u,v) = 1$ and $H(u,v) = 0$ transition. To lessen ringing effects, however, it introduces a gradual transition from $1\ to\ 0$ rather than a fast cut-off. This step aims to optimize subsequent segmentation and feature extraction processes.

2.2. Cascaded Fuzzy C-Mean Segmentation

The k-means clustering method serves as a foundation for the fuzzy c-means algorithm. The k-means method introduces a fuzzy concept as a result of the data's imprecision. Bezdek and Dunn created the fuzzy c-means method in 1981. Centroids define the classes in this algorithm. A degree of membership between 0 and 1 can be assigned to each observation. Here, fuzzy c-means is based on iteratively reducing the function cost,J_{fcm} Provided by the calculation.

$$J_{fcm} = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^q d^2(x_k v_i) \quad (2)$$

Where the total sum of the dataset's data is$n$, c is the cluster number that falls between 2 and$n$, u_{ik} is the level of affiliation for the i^{th} group, q the weighting function, and $X = \{x_1, x_2, \cdots, x_n\} \subseteq R^P$ is the vector population in the dataset with p-dimension. The Euclidean distance is supplied by$d^2(x_k v_i)$, while the cluster center is given byv_i.

Here is the algorithm:

Step 1: Set the initial values forc, q andε.

Step 2: Determine the values of$\{u_{ik}\} = U$.

Step 3: Set up the count for the loop initially as$\begin{Bmatrix} b_{(b)} = 0 \\ v_i \end{Bmatrix}$.

Step 4: Calculate the cluster centers using$U^{(b)}$.

Step 5: Estimate the membership function$U^{(b+1)}$.

For k=1 to n, estimate $I_k = \{i | 1 \le i \le c, d_{ik} = ||x_k - v_i|| = 0\}$, or I. Compute new membership values for k^{th} column for the equation's given conditions shown below (3):

If $I_k = \xi$ then,

$$u_{ik}^{(b+1)} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{(q-1)}}} \quad (3)$$

Else $u_{ik}^{(b+1)} = 0$ for all $i \notin I$ and $\sum_{i \in I_k} {u_{ik}}^{(b+1)} = 1$;

Step 6: If $u_{ik}^{(b+1)} = 0$ for all$i \notin I$, then end; else fix $b = b + 1$ and go to step 4. This segmentation technique extracts meaningful information from images and enhances the efficiency of upcoming image analysis operations.

2.3. Feature Extraction (GLCM)

Features are the bits of data needed to solve particular problems and convey important aspects of an image. The choice of input features significantly impacts the requirements for the training set and classification accuracy. The technique of extracting an image's visual content, feature extraction, is the process of reducing the amount of resources required. GLCM is created in this investigation to retrieve statistical texture data. Texture features are important low-level characteristics utilized to quantify an image's perceived texture and define its contents.

2.3.1. GLCM

It is a recognized method utilizing numerical distributions of intensity value combinations at various locations in relation to one another in an image to extract second-order statistical texture features. A picture's intensity points determine which category of statistics is considered first, second, or higher order. Although they are theoretically feasible, computational complexity prevents the application of higher-order statistics. Texture features hold details on how surfaces are arranged structurally and how they interact with the environment. In total, the following texture-based attributes are produced: energy, correlation, entropy, Inverse Difference Moment (IDM), homogeneity, sum variance, autocorrelation, contrast, maximum probability, dissimilarity, and IDM normalized. These are some examples of how they are expressed:

2.3.2. Energy

"Uniformity" or "angular second moment" are alternate terms for vitality. The sum of the square components of the GLCM matrix is given. This is how regions that are homogeneous to non-homogeneous are converted. It is high when repeated image pixels occur frequently.

$$Energy = \sum_{i,j=0}^{N-1}\left(P_{ij}\right)^2 \quad (4)$$

2.3.3. Entropy

It establishes how unpredictable the image is. A homogeneous image will, therefore, have a lower entropy value.

$$Entropy = \sum_{i,j=0}^{N-1} -ln\left(P_{ij}\right)P_{ij} \quad (5)$$

2.3.4. Contrast

It gauges the strength of contrast that connects a pixel to its neighbor across the entire picture.

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}\,(i-j)^2 \quad (6)$$

2.3.5. Correlation

It is a measurement of the linear gray tone dependence of an image. It indicates how a pixel is connected with its neighbor.

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(i-\mu)}{\sigma^2} \quad (7)$$

2.3.6. Variance

The sum of squares variance measures the spreading of the image's gray level sum distribution. The standard deviation, a first-order statistical quantity, and this heterogeneity have a strong correlation. The variance of the image increases when the gray level value deviates from its mean value. Variance characteristics of an image are as follows:

$$Variance = \sum_{i=1}^{N_g} (i-\mu)^2 p(i,j) \quad (8)$$

Where, $P_{ij} = (i,j)$ The normalized GLCM matrix's element, μ GLCM matrix mean determined by applying the formula.

$$\mu = \sum_{i,j=0}^{N-1} iP_{ij} \quad (9)$$

$\sigma =$ The variance of each pixel's intensity determined using

$$\sigma^2 = \sum_{i,j=0}^{N-1} iP_{ij} \, (i-\mu)^2 \quad (10)$$

N=the image's amount of gray levels.

2.4. CNN-BiLSTM

Convolution kernels in CNN, the well-known deep neural network, only scan the logging curves' depth direction. A One-Dimensional (1-D) CNN is used in the proposed work. Three layers make up the CNN: convolutional, pooling, and fully connected. The CNN uses pooling and convolution techniques to extract implicit characteristics from the input data. The collected features are then combined and added to a fully connected layer. Lastly, the output of a neuron can be made non-linear by applying an activation function. The convolution layer is an essential part of the CNN. Each convolutional layer possesses an abundance of convolutional kernels, which are convolved with input data in order to extract latent features and generate feature maps. The feature maps are passed through a non-linear activation function to produce the convolutional layer's output. Following is an expression for the convolutional layer:

$$C_i = f(h_i) = max(0, h_i) \quad (11)$$

Where h_i is the feature map element that results from convolutional processes.

Reducing feature map dimensions and avoiding overfitting are the goals of the pooling operation. One of the most used pooling techniques is max pooling. This is accomplished by using Equations (12) and (13) to determine the determined value of a designated area in feature maps.

$$\gamma(C_i, C_i - 1) = max(C_i, C_i - 1) \quad (12)$$

$$P_i = \gamma(C_i, C_i - 1) + \beta_i \quad (13)$$

Where the function of max pooling subsampling is denoted by $\gamma(\cdot)$. The prejudice is denoted by β_i. P_i Stands for the maxpooling layer's output. Ultimately, the fully connected layer receives the final output vector is computed using the feature maps generated by the convolutional and pooling algorithms, as shown below:

$$y_i = f(t_i P_i + \delta_i) \quad (14)$$

Where t_i a weight matrix, y_i is the final vector output, and δ_i is the bias.

2.5. Bidirectional LSTM Network

Only prior is utilized by conventional LSTMs. The BiLSTM is provided as a way to grasp two-direction contextual dependencies better and obtain long-range information. The contextual data from both could be extracted by bidirectional architecture simultaneously with backward and forward hidden layers in both directions. Figure 2 displays the BiLSTM architecture.

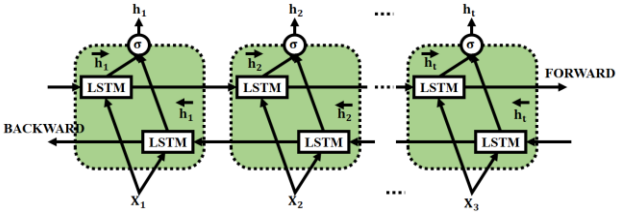

Fig. 2 Architecture of BiLSTM

Both $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ in Figure are the results of the hidden layers that go forward and backward, in turn. For the forward layer, the hidden sequences and their outputs are computed iteratively from $step\ t$ to $step\ 1$ using an orderly series of inputs; for the backward layer, the iteration is from step t to step 1. The conventional LSTM computes the results of both the forward and backward layers, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$. With each element calculated in accordance with Equation (15), the BiLSTM layer produces an output vector Y.

$$y_t = \left(\overrightarrow{h_t}, \overleftarrow{h_t}\right) \quad (15)$$

Where the two $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are coupled using a σ function sequence. A summing function, multiplication function, concatenating function, or average function can all be represented by the σ function. When a BiLSTM layer is used to conduct well log prediction, the final output can be represented as a vector $Y = [y1, y2 \ldots yt]$, where the previous part, yt predicted well logging value for subsequent depth.

2.6. Attention Mechanism

Neural networks with an attention mechanism have proven effective in a variety of tasks recently. When a BiLSTM network has an attention mechanism, the attention approach makes use of the most recent cell state of BiLSTM or uses the BiLSTM's implicit state to align with the input's cell state at the current step. Subsequently, the correlation coefficient is calculated between the output state and these potential intermediate stages. To improve the precision and effectiveness of prediction, pertinent information can be emphasized during the learning process while unnecessary information might be hidden. The following Equations (16) - (18) produce the attentive BiLSTM network's output A, which is formed by the attention layer:

$$M = \tan h(Y) \tag{16}$$

$$a = softmax(w_a^T M) \tag{17}$$

$$Y = Y\alpha^T \tag{18}$$

Where Y is a matrix that encodes features that the BiLSTM model has detected, an example of this would be the matrix$Y = [y1, y2 \ldots yt]$. The vector α denotes the attention weights assigned to certain features$Y$. $W_a$is the weight matrix of layer of attention. T Stands for transposition operation.

3. Result and Discussion

Based on the validation results of the methodology, mammograms are generated using Matlab. The data set consists of five hundred records, of which eighty percent are used for testing, and twenty percent are for training. Since the remaining 400 records are a part of the training process, the first 100 records make up the testing phase. Images are classified based on the number of nodules present; fewer nodules in an image present benign conditions, whereas more nodules in an image indicate malignant conditions.

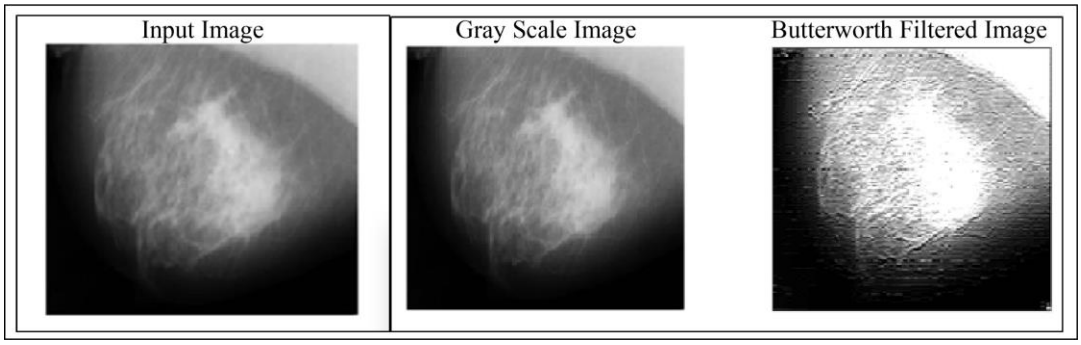

Fig. 3 Filtered image

Despite noise contamination in the input image, Figure 3 shows how the Butterworth filter successfully preserves edge information during noise reduction. Due to the intricacies of the input image and its preservation of little features, the Butterworth filter yields higher visual quality than other filters. The impacted area is then suitably partitioned to improve breast cancer detection. The segmentation process is managed by Cascaded Fuzzy C-mean, which measures the correspondence between object pairs.

Figure 4 illustrates how the Fuzzy C-mean distinguishes between sparse and dense areas, revealing overall distribution patterns and connections among various visual elements.

Table 1. Extracted features by GLCM

Features	Value
Correlation	0.020000
Contrast	0.000000
Energy	1.000000
Entropy	-0.000000
Variance	3.577532

Fig. 4 Segmented output

The segmented image's energy, correlation, variance, entropy, and contrast properties have all been extracted using the GLCM technique. The extracted output features are listed in Table 1.

Finally, Figure 5 demonstrates the output categorization result using CNN-BiLSTM. The classifier demonstrates that the input image classified reveals a malignant stage of cancer.

Similarly, the proposed system is tested on other 4 images detecting cancer in mammography images of breast cancer, and the corresponding outcomes accomplished are illustrated in Figure 6.

The extracted features of these four images, 1, 2, 3, and 4 utilizing GLCM, are listed in Table 2, which supports in accurate classification of breast cancer images.

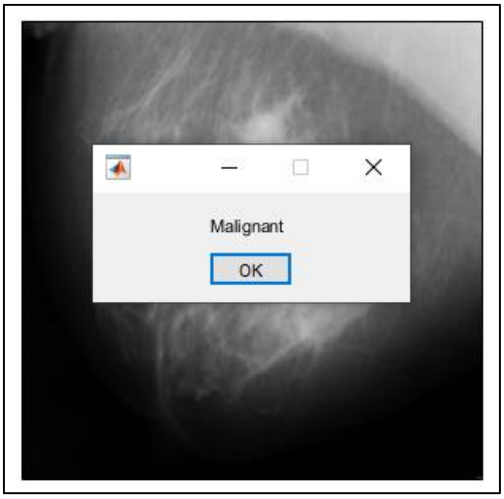

Fig. 5 Classification output

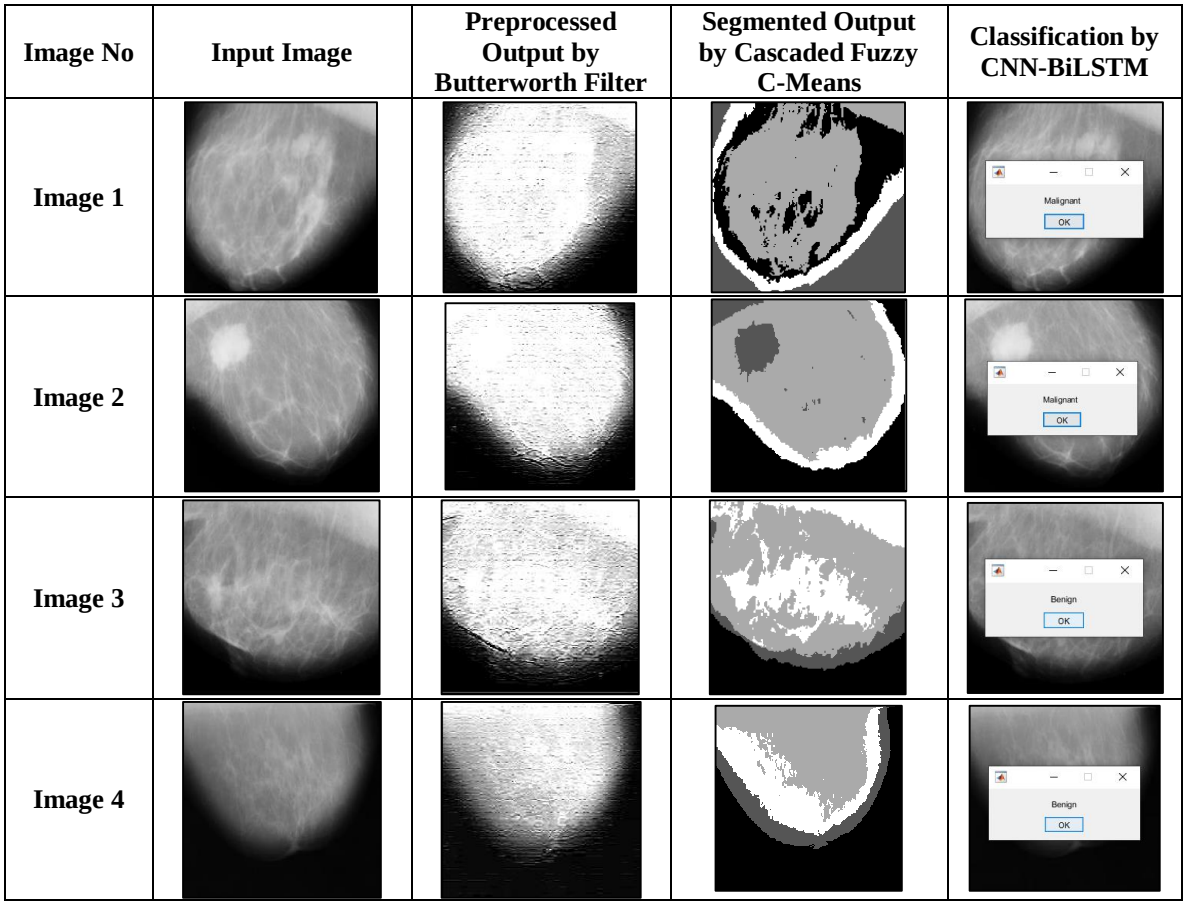

Fig. 6 Classification of outputs

Table 2. Features extracted with GLCM

Features	Image 1	Image 2	Image 3	Image 4
Correlation	0.020000	0.020000	0.020000	0.020000
Contrast	0.000000	0.000000	0.000000	0.000000
Energy	1.000000	1.000000	1.000000	1.000000
Entropy	-0.000000	-0.000000	-0.000000	-0.000000
Variance	3.746001	3.144654	5.285749	3.432792

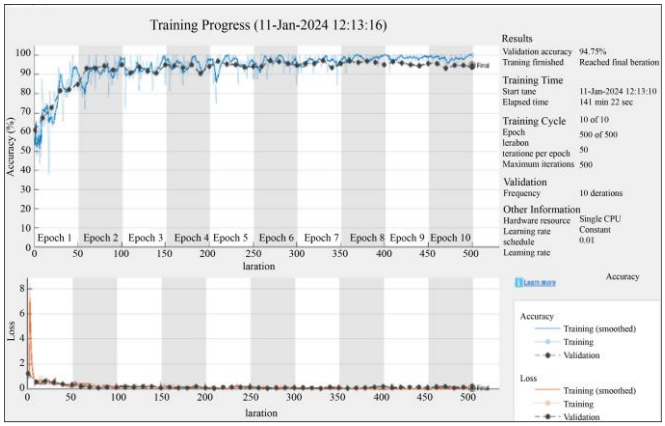

Fig. 7 Training accuracy

The accuracy and loss accomplished by the CNN-BiLSTM classifier is illustrated in Figure 7. It is noticed that the proposed system results with 94.75% accuracy and minimized loss measure in contrast to the training value. Table 3 demonstrates the performance metrics of the proposed CNN-BiLSTM classifier.

Table 3. Classification by CNN-BiLSTM

Performance Metrics	Percentage (%)
Accuracy	94.75
Precision	94.22
Recall	94.66
Sensitivity	93.5

Table 4. Comparison of classification techniques

Algorithms	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	Recall (%)
KNN [21]	74	76	84	76	88
CNN [22]	93	92.73	93.85	92.29	93.85
CNN-BiLSTM	94.75	94.22	94.01	93.5	94.66

The comparative analysis of various classifiers, including KNN [21], CNN [22] and the proposed CNN-BiLSTM approach, is listed in Table 4. On the observation the proposed CNN-BiLSTM ranks with superior accuracy of 94.75%, Precision measure of 94.22 and sensitivity value of 94.01%. In addition, the proposed classifier results with the specificity of 93.5% and 94.66% recall measure, respectively.

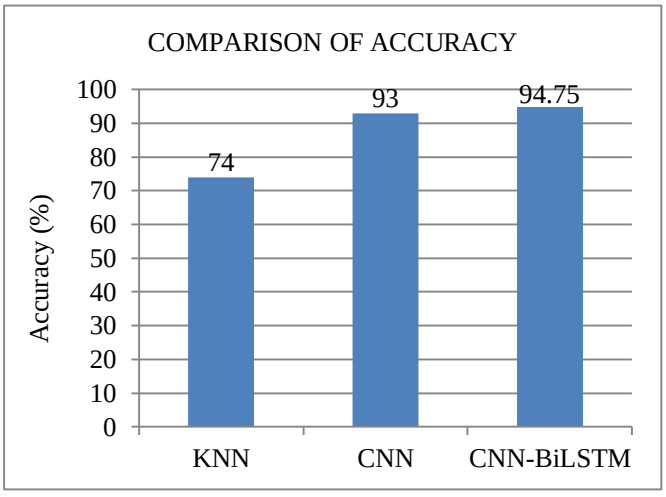

Fig. 8 Comparison of accuracy

The accuracy analysis of approaches such as KNN [21], CNN [22] and the proposed CNN-BiLSTM is represented in Figure 8. To be noticed that KNN shows an accuracy value of 74%, with improved CNN having an accuracy ranging from 93% and the proposed CNN-BiLSTM demonstrating a superior accuracy measure value of 94.75%, offering enhanced classification.

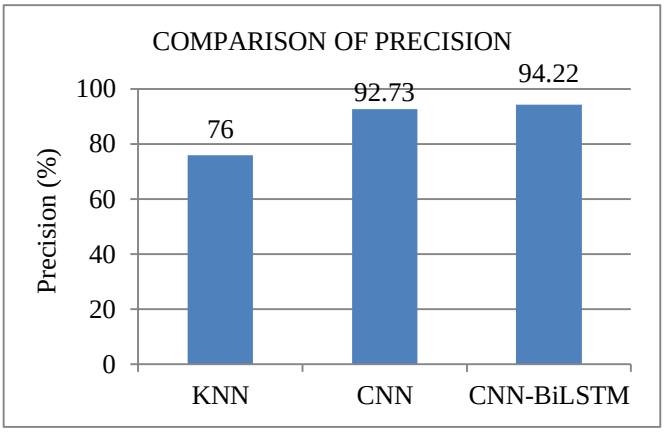

Fig. 9 Analysis of precision

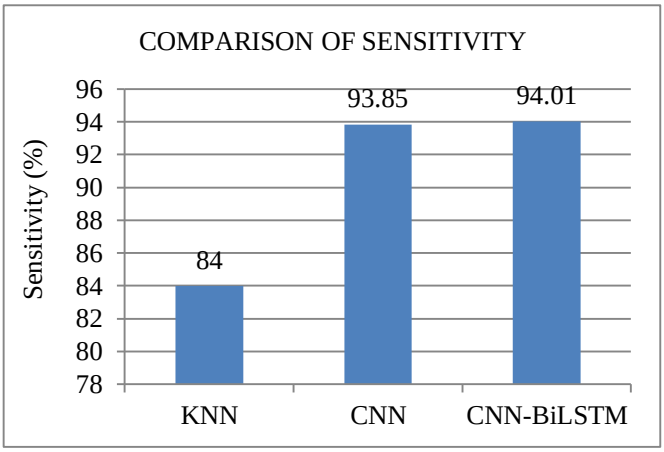

Fig. 10 Comparative study of sensitivity

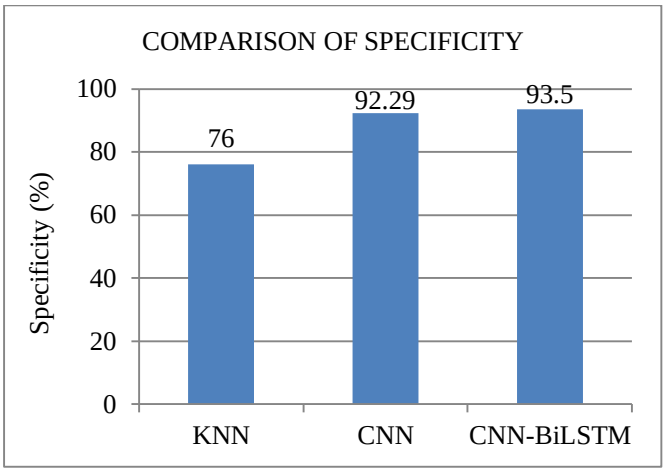

Fig. 11 Comparative assessment of specificity

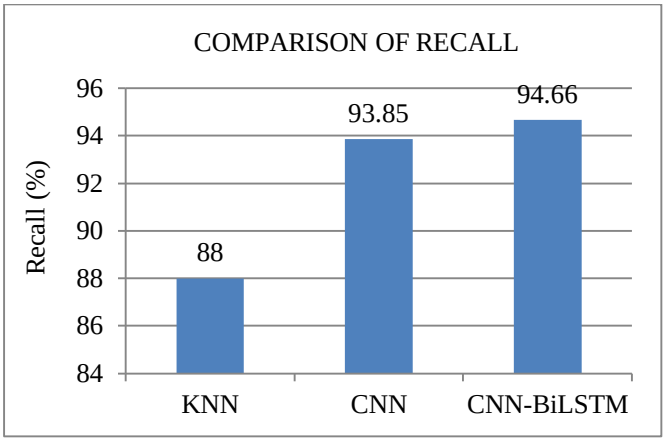

Fig. 12 Assessment of recall

The analysis of precision measures for 3 classification techniques is illustrated in Figure 9. On comparing the KNN [21] and CNN [22] classifier, an efficiency value of 76% and 92.73% is obtained. While the proposed CNN-BiLSTM shows superior precision performance with a value of 94.22%, respectively.

Figure 10 shows the gaining of the proposed CNN-BiLSTM, attaining 94.01%, while comparing the other existing methods, such as KNN [21] and CNN [22], it exhibits a percentage of 84, 93.85 which is low in contrast to the proposed CNN-BiLSTM technique.

The assessment of specificity for approaches like KNN [21], CNN [22] and the proposed CNN-BiLSTM is represented in Figure 11. It is observed that KNN reveals a specificity value of 76%, CNN shows 92.29%, and the proposed CNN-BiLSTM demonstrates enhanced specificity of 93.5%, accordingly.

Figure 12 demonstrates the recall measure values for different methods, indicating a recall of 88% and 93.85% for KNN [21] and CNN [22], respectively. Notably, the proposed CNN-BiLSTM outperforms both, achieving a higher recall measure of 94.66%. This suggests that the CNN-BiLSTM model is more effective in capturing instances of true positives compared to the CNN and KNN models.

4. Conclusion

In conclusion, an advanced methodology for breast cancer prediction combining various ML models detecting cancer in breast stages is established. The significance of this study lies in its potential impact on improving early breast cancer detection, which is a crucial time for intervention and treatment. The utilization of Butterworth filters for denoising, Cascaded FCM for effective image segmentation, and GLCM for feature extraction demonstrates a comprehensive and advanced approach to preprocessing the input images. The integration of the distinctive CNN-BiLSTM method further refines the categorization of cancer stages.

The comparison with traditional approaches such as KNN CNN provides a valuable benchmark for evaluating the proposed methodology's efficacy. The obtained outcomes from the comparative analysis indicate that the proposed approach outperforms the others in terms of accuracy at 94.75%, specificity at 93.5%, and sensitivity revealing 94.01%. This proposes that the combination of image processing techniques and the CNN-BiLSTM model offers a superior solution for breast cancer prediction. The robustness and generalizability of the proposed method should be improved by further research and validation using bigger datasets and a wider range of people.

References

[1] Yi Wang et al., "Deeply-Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 866-876, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[2] Usman Naseem et al., "An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers," *IEEE Access*, vol. 10, pp. 78242-78252, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Abeer Saber et al., "A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique," *IEEE Access*, vol. 9, pp. 71194-71209, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Zexian Huang, and Daqi Chen, "A Breast Cancer Diagnosis Method Based on Vim Feature Selection and Hierarchical Clustering Random Forest Algorithm," *IEEE Access*, vol. 10, pp. 3284-3293, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Neha N. Ganvir, and D.M. Yadav, "Filtering Method for Pre-Processing Mammogram Images for Breast Cancer Detection," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 4222-4229, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] R. Ramani et al., "The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images," *International Journal of Image, Graphics and Signal Processing*, vol. 5, no. 5, pp. 47-54, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[7] Nijad Al-Najdawi, Mariam Biltawi, and Sara Tedmori, "Mammogram Image Visual Enhancement, Mass Segmentation and Classification," *Applied Soft Computing*, vol. 35, pp. 175-185, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[8] Milos Radovic et al., "Parameter Optimization of a Computer-Aided Diagnosis System for Detection of Masses on Digitized Mammograms," *Technology and Health Care*, vol. 23, no. 6, pp. 757-774, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[9] Hanife Avcı, and Jale Karakaya, "A Novel Medical Image Enhancement Algorithm for Breast Cancer Detection on Mammography Images Using Machine Learning," *Diagnostics*, vol. 13, no. 3, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] P. Esther Jebarani et al., "A Novel Hybrid K-Means and GMM Machine Learning Model for Breast Cancer Detection," *IEEE Access*, vol. 9, pp. 146153-146162, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Rajeshwari S. Patil, and Nagashettappa Biradar, "Improved Region Growing Segmentation for Breast Cancer Detection: Progression of Optimized Fuzzy Classifier," *International Journal of Intelligent Computing and Cybernetics*, vol. 13, no. 2, pp. 181-205, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Naglaa F. Soliman et al., "An Efficient Breast Cancer Detection Framework for Medical Diagnosis Applications," *Computers, Materials & Continua*, vol. 70, no. 1, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Milosevic, Marina et al., "Early Diagnosis and Detection of Breast Cancer," *Technology and Health Care*, vol. 26, no. 4, pp. 729-759, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Selina Sharmin et al., "A Hybrid Dependable Deep Feature Extraction and Ensemble-Based Machine Learning Approach for Breast Cancer Detection," *IEEE Access*, vol. 11, pp. 87694-87708, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Amita Das et al., "Deep Learning Based Liver Cancer Detection Using Watershed Transform and Gaussian Mixture Model Techniques," *Cognitive Systems Research*, vol. 54, pp. 165-175, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Kumari Nibha Priyadarshani, and Sangeeta Singh, "Ultra-Sensitive Breast Cancer Cell Lines Detection Using Dual Nanocavities Engraved Junctionless FET," *IEEE Transactions on NanoBioscience*, vol. 22, no. 4, pp. 889-896, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Irum Hirra et al., "Breast Cancer Classification from Histopathological Images Using Patch-Based Deep Learning Modeling," *IEEE Access*, vol. 9, pp. 24273-24287, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jiande Wu, and Chindo Hicks, "Breast Cancer Type Classification Using Machine Learning," *Journal of Personalized Medicine*, vol. 11, no. 2, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Kapil Juneja, and Chhavi Rana, "An Improved Weighted Decision Tree Approach for Breast Cancer Prediction," *International Journal of Information Technology*, vol. 12, no. 3, pp.797-804, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Muhammad Danish Ali et al., "Breast Cancer Classification through Meta-Learning Ensemble Technique Using Convolution Neural Networks," *Diagnostics*, vol. 13, no. 13, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Mona Alfifi et al., "Enhanced Artificial Intelligence System for Diagnosing and Predicting Breast Cancer Using Deep Learning," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mehedi Masud, Amr E. Eldin Rashed, and M. Shamim Hossain, "Convolutional Neural Network-Based Models for Diagnosis of Breast Cancer," *Neural Computing and Applications*, vol. 34, no. 14, pp. 11383-11394, 2022. [CrossRef] [Google Scholar] [Publisher Link]