SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I11P138

Volume 11 Issue 11, 409-423, November 2024

Original Article

Optimized DL Approach for Microcalcification Analysis in Digital Breast Tomosynthesis

K. Jasna[1*], S. Albert Jerome[2]

[1]*Department of Electronics& Instrumentation Engineering, Noorul Islam Centre for Higher Education, Tamilnadu, India.*
[2]*Department of Biomedical Engineering, Noorul Islam Centre for Higher Education, Tamilnadu, India.*

[1]*Corresponding Author : jasnak636@gmail.com*

Received: 21 September 2024 Revised: 22 October 2024 Accepted: 20 November 2024 Published: 30 November 2024

Abstract *- Breast cancer is the primary cause of cancer-related deaths in women globally, and increased survival rates and effective treatment depend on early identification. One of the earliest indications of breast cancer is frequently Microcalcifications (MC), making their accurate classification vital in breast cancer diagnosis. In order to classify MCs using data from Digital Breast Tomosynthesis (DBT), this study suggests an effective Deep Learning (DL) algorithm. Stacked Long Short-Term Memory (LSTM) with Bayesian Optimization was developed and evaluated on the DBT slice. Various features associated with MCs are extracted from the DBT slices with the help of image processing techniques. Extensive Exploratory Data Analysis (EDA) was performed on the DBT dataset to extract key features and identify patterns related to MC classification. Preprocessing techniques were applied to enhance image quality and remove noise. The proposed LSTM model achieved an accuracy of 97.27% and demonstrated superior classification performance, achieving better accuracy than traditional methods. The approach improves breast cancer detection speed and accuracy by automating the categorization process, which presents a possible path for early diagnosis and better patient outcomes. These results advance DL methods in medicine, especially in terms of better radiographic imaging analysis for breast cancer screening.*

Keywords *- Microcalcification, Stacked long short-term model, Digital breast tomosynthesis, Explanatory data analysis, Ductal carcinoma.*

1. Introduction

Breast cancer is a condition where abnormal breast cells grow uncontrollably, forming tumors [1]. It starts from the milk ducts or lobules and may spread to nearby tissues, becoming potentially life-threatening. Men can also develop it; however, it typically affects women. Risk factors include age, obesity, alcohol use, family history, and genetic mutations like BRCA1 and BRCA2. No known risk factors exist for many breast cancer patients. Early symptoms may include lumps or nipple discharge, though many cases are asymptomatic early on. Surgery, radiation, chemotherapy, or targeted medicines are available as forms of treatment; early detection improves results [2].

Breast calcifications are calcium deposits in breast tissue commonly detected by mammograms. Though usually harmless, they can occasionally indicate early cancer signs. There are two types of macrocalcifications: large and generally non-cancerous and microcalcifications, which are smaller and might indicate cancer or precancerous changes. Usually exhibiting no symptoms, calcifications are not perceptible when examined [3]. Causes include aging, past breast injuries, infections, or treatments. Routine mammograms are essential for detecting suspicious calcifications. If needed, further tests like diagnostic mammograms or biopsies help determine cancer risk, especially since microcalcifications may indicate early-stage cancers like Ductal Carcinoma in Situ (DCIS). Advanced imaging technology called DBT improves the identification and evaluation of microcalcifications in breast tissue. DBT enhances the detection of microcalcifications, which can indicate early breast cancer, especially Ductal Carcinoma in Situ (DCIS) [4]. By capturing multiple low-dose X-rays from different angles, DBT creates detailed 3D images of breast tissue, improving visualization compared to traditional 2D mammography. An overview of DBT is shown in Figure 1. A key advantage of DBT is its ability to reduce overlapping tissue issues common in standard mammograms, allowing radiologists to assess better the size, shape, and distribution of microcalcifications, particularly in dense breast tissue.

The analysis of microcalcifications in DBT using DL represents a significant advancement in breast cancer detection. Conventional imaging techniques frequently lack the degree of accuracy and sensitivity required for accurate verification, especially in dense breast tissue. DL algorithms,

utilizing neural networks, analyze extensive mammographic datasets to identify patterns associated with microcalcifications [5].

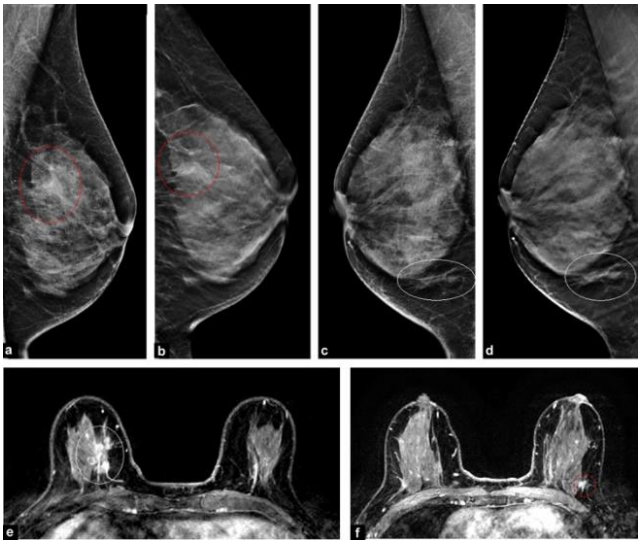


Fig. 1 Overview of DBT

By evaluating features like size, shape, and distribution, these algorithms enhance diagnostic accuracy and reduce interpretation time, leading to improved workflow and better patient outcomes.

LSTM networks offer substantial advantages in analyzing microcalcifications in DBT due to their design for sequential data. When combined with Convolutional Neural Networks (CNNs), LSTMs enhance feature extraction, leading to more accurate and reliable diagnostic outcomes in breast cancer detection. The major objectives of the proposed methodology include:

- To generate a new dataset by utilizing the collected DBT slices, ensuring a comprehensive foundation for analysis.
- To propose various image processing techniques and feature extraction methods to derive meaningful attributes from the DBT slices.
- To introduce and evaluate multiple feature optimization methods that improve the efficiency and accuracy of the extracted features.
- To propose a Stacked LSTM model as an effective approach for microcalcification classification.
- To assess how effectively proposed classification models perform in relation to pre-established metrics.

The research is organized into subsequent sections: Section 2 offers an overview of earlier studies, highlighting topics that require additional investigation. A thorough explanation of the approach is provided in Section 3. The detailed results of the suggested methodology are shown in Section 4. The analysis and conclusion of the study are presented in Section 5.

2. Literature Review

Francesco Prinzi et al. (2024) [6] discussed the exertion in detecting breast microcalcifications, which can become invasive cancers and show up in 80% of mammograms. In the ability to distinguish benign from malignant microcalcifications in healthy tissue, the study proposed a radiomic signal.

136 normal and 242 cancerous microcalcifications were identified from 380 healthy samples using radiomic characteristics. Support Vector Machine (SVM), XGBoost, and Random Forest (RF) were among the Machine Learning (ML) models utilized for detection. For classifying healthy, benign, and malignant patients, XGBoost obtained AUC-ROC values of 0.830, 0.856, and 0.876.

Marwa Ben Ammar et al. (2024) [7] conducted a study on early breast cancer detection through mammography, exploring the potential of DL techniques to improve diagnostic accuracy. The study examined architectures like Recurrent Neural Networks (RNN), CNN, and Generative Adversarial Networks (GAN) for detection and classification tasks. Pre-trained models like ResNet 50 and AlexNet were evaluated for image analysis, with CNNs showing the highest accuracy in detecting breast abnormalities.

Laxman Singh and Altaf Alam (2024) [8] proposed a hybrid methodology for detecting doubtful mass regions in digital mammograms by integrating a pixel-based low-level preprocessing technique with the Faster R-CNN model. The method addressed the limitations of Faster R-CNN by mitigating occlusion issues caused by normal breast tissue, pectoral muscles, and noise. The method's sensitivity, accuracy, specificity, and AUC were assessed. Results showed that the model achieved an accuracy of 94.2%, a specificity of 93.5%, and an AUC score of 0.983, outperforming other state-of-the-art algorithms such as SSD and R-FCN.

Paul Terrassin et al. [9] developed an annotation-free framework for detecting and classifying Microcalcifications (MCs) in mammograms using a deep CNN inspired by the UNet3+ architecture. This framework integrated multi-scale features to enhance context and detail, addressing challenges with MCs' small size. It applied a two-channel segmentation and three-class classification method to distinguish soft, benign, and malignant tissues. The study used synthetic MC data for self-supervised training and achieved an AUC of 0.93 for malignant MC classification. Validated with INBreast and BMCD datasets, the model significantly improved over ResNet-22 and ConvNeXt architectures.

Jinhua Wang et al. (2016) [10] created a model based on DL to enhance the accuracy of microcalcification detection for early breast cancer diagnosis. Using a semi-automated segmentation method, the model characterized microcalcifications and achieved 87.3% accuracy, outperforming the SVM with 85.8%. When combined with breast mass data, accuracy improved to 89.7%. The study demonstrated that DL models surpassed traditional methods in diagnostic accuracy and showed potential clinical value in detecting microcalcifications and breast masses for early breast cancer diagnosis.

Redona Brahimetaj et al. (2022) [11] evaluated the relationship between microcalcifications' shape and texture with malignancy. Biopsy samples from 94 female patients were scanned using micro-CT, and 3504 microcalcifications were analyzed. Radiomic features were extracted, and CNN was used to diagnose them. Individual microcalcifications were classified with 77.32% accuracy, while sample-level accuracy reached 84.04%. The study found texture features from 3D images to be more predictive of malignancy than shape features, marking the largest study evaluating Computer-Aided Detection (CAD) systems using 3D microcalcification data for breast cancer diagnosis.

Koushlendra Kumar Singh et al. (2022) [12] proposed a DL model for detecting and classifying breast microcalcifications into benign and malignant. Using CNNs with four optimizers and the InceptionResNetV2 model for feature extraction, the model was trained on CBIS–DDSM mammogram images. The model achieved 94% validation accuracy, 94% overall accuracy, and a 96% AUC. The ADADelta optimizer yielded the best results, outperforming traditional and DL methods, demonstrating the effectiveness of automating microcalcification detection without manual feature extraction.

Chrysostomos Marasinou et al. (2023) [13] created a multiscale method to lower diagnostic uncertainty in Full-Field Digital Mammography (FFDM) by segmenting Microcalcification (MC) data. The method used blob detection and Hessian analysis to delineate candidate objects, followed by a regression convolutional network trained to detect MCs. Trained on 435 FFDMs, the model outperformed state-of-the-art methods in intersection over union (0.670 vs. 0.524) and true positive rate (0.744 vs. 0.581). A case study of 248 amorphous MC cases classified cases as benign or malignant with an AUC of 0.763, improving MC segmentation and reducing unnecessary biopsies.

Ehtsham Rasool et al. (2023) [14] focused on reducing breast cancer mortality through early detection, highlighting that it affects one in eight women by age 80 in the U.S. The study utilized transfer learning with a VGG-19 neural network to detect mammogram MCs. Their automated algorithm pre-processed images to cluster MCs before classification, achieving 96% sensitivity on the CBIS-DDSM and DDSM databases, outperforming other ML models. The study demonstrated the efficiency of transfer learning in reducing false positives and improving breast cancer detection accuracy.

Qing Lina et al. (2023) [15] demonstrated the importance of identifying microcalcifications in mammograms for early breast cancer diagnosis. Their automated deep-learning pipeline detected and classified microcalcifications from a dataset of 4,810 mammograms with 6,663 lesions. Using a Faster RCNN model, the system achieved 0.8124 classification accuracy for benign vs. malignant cases in training and 0.7237 in testing, with a sensitivity of 0.8891 and 0.7778, respectively. The AI system excelled in visual interpretation and region extraction, achieving an AUC of 0.8042, though its performance in classifying pathological types and ER statuses highlighted room for improvement.

Zahra Maghsoodzadeh Sarvestani et al. (2023) [16] aimed to improve breast cancer detection by enhancing mammography images of masses and microcalcifications. Their study used fuzzy systems and Gabor filtering to improve microcalcification image quality and clustered them using a decision tree algorithm. The benign and malignant areas of interest (ROIs) were categorized by an Artificial Neural Network (ANN). Using the DDSM dataset, the method achieved 93% accuracy and over 95% sensitivity. The study showed that their fuzzy design improved image detection speed and quality, proving effective for automatic breast cancer diagnosis using mammograms.

Ana M. Mota et al. (2022) [17] investigated Microcalcification Clusters (MCs) as key biomarkers in breast cancer diagnosis. They implemented CNN to classify entire DBT images, focusing on noise reduction and contrast enhancement without prior region selection. Deep CNNs were trained in a new manner and tested on DBT data, with GoogLeNet achieving an AUC of 94.19% and their proposed model an AUC of 91.17%. The study filled a gap in automatic MC classification without preselecting regions.

Despite progress in DL for identifying breast lesions in Digital Breast Tomosynthesis (DBT), there is a scarcity of research focused on automatically classifying Microcalcification Clusters (MCs) throughout entire images without prior region selection. Most existing studies concentrate on larger lesions and rely on private databases, which restricts the applicability of their findings. While various preprocessing techniques have been investigated, their effects on MC detection within complete DBT images have not been comprehensively studied. There is a need for further research to create fully automated models that can accurately identify and classify MCs, ultimately enhancing early breast cancer detection in clinical environments.

3. Materials and Methods

The suggested approach comprises the data generation phase and classification model. In the data generation process, the dataset required for the proposed work has been generated from the DBT slices with the help of various image processing and feature extraction techniques. A microcalcification classification model is proposed using stacked LSTM. The thorough advanced computational framework for microcalcification analysis in DBT using the DL model is described in Figure 2.

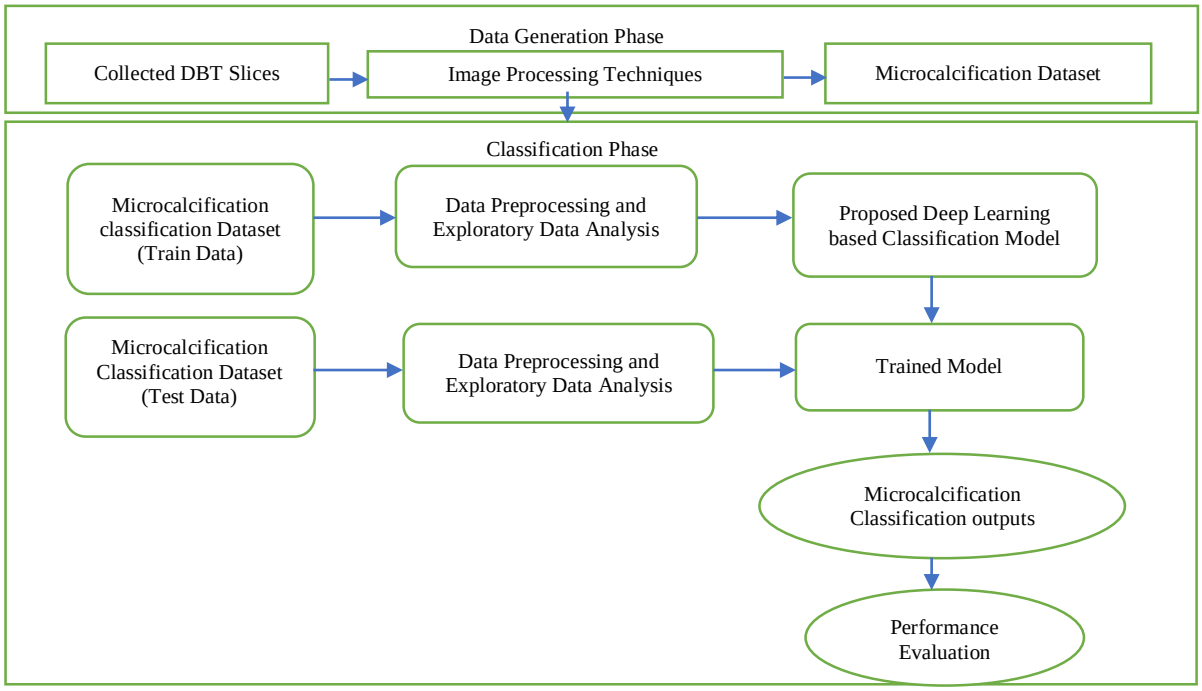


Fig. 2 Computational framework for microcalcification analysis in DBT using DL model

3.1. Data Generation Process

DBT is a 3D imaging technique that enhances breast cancer screening by generating multiple slices of breast tissue, reducing overlapping structures seen in conventional mammography. This method is particularly effective for detecting microcalcifications and tumors. The study focuses on the processes, challenges, and considerations in collecting DBT slices for automated detection systems. DBT captures low-dose X-ray images from various angles, reconstructing them into thin slices (1mm to 3mm) for detailed examination. Proper selection, preprocessing, and annotation of this volumetric data are crucial for training ML algorithms.

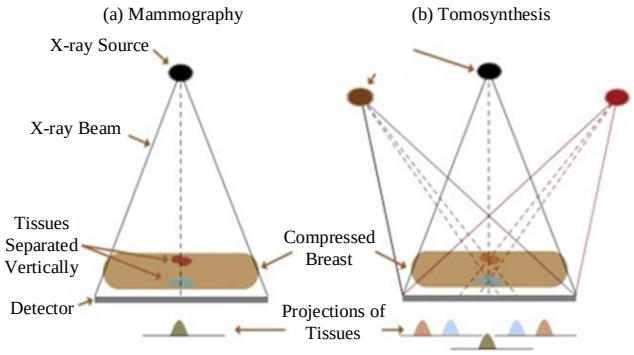


Fig. 3 DBT with mammography for the identification of breast cancer

DBT is an imaging method that uses a geometric premise very similar to that used in stratigraphic technique [18] to allow a small number of low-dose two-dimensional projections made with various X-ray tube orientations to create a volumetric reconstruction of the complete breast. Figure 3 indicates that in DBT, an arc is formed by the X-ray tube and a sequence of images is obtained, with each image being dosed at a proportion lower than that of a typical mammography.

In some systems, a conventional 2D Full-Field Digital Mammography (FFDM) image is also captured alongside the 3D slices. These can be combined for more comprehensive data analysis and DL models. Table 1 includes various features associated with DBT slices for microcalcification classification, such as “area”, “gray level”, “gradient strength”, “noise fluctuation”, “contrast”, and “shape descriptor”.

Several preprocessing techniques are employed to prepare DBT slices for analysis. In grayscale conversion, DBT images are converted from color to grayscale by reducing the color depth to shades of gray. This process uses weighted averages of the RGB channels, emphasizing the green channel due to human visual sensitivity.

Table 1. Feature description

Features	Description
Gray Level	The average gray level of the object.
Contrast	The object's average grayscale value is less than the average of the two-pixel-wide border surrounding it.
Gradient Strength	The object's perimeter pixels have a gradient strength.
Shape Descriptor	A low order moment based on shape descriptor.
Area	Area of object in pixels
Noise Fluctuation	Root means square noise fluctuation in the object.

By normalization, pixel values are standardized to a range from 0 to 255, ensuring uniform intensity across all images and transforming the image so that pixel values range from 0 to 255.

$$Gray\ Scale\ Conversion = 0.299 \cdot Red + 0.587 \cdot Green + 0.114 \cdot Blue \quad (1)$$

Gaussian blurring is used in image processing to reduce noise and detail by applying a Gaussian function to the image. The level of blurring is controlled by the kernel's standard deviation (sigma). Gaussian function in one dimension is represented as

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (2)$$

Gaussian function in two dimensions is expressed as

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

The Histogram Equalization method enhances global contrast by redistributing pixel intensities, making Microcalcifications (MCs) stand out more against the background. The mathematical expression of histogram equalization is expressed as

$$s_k = T(r_k) = \sum_{k=0}^{k} \frac{n_k}{n} = \sum_{k=0}^{k} p_k(r_k) \quad (4)$$

Otsu's Thresholding is an effective image segmentation technique [20] that establishes the ideal threshold automatically for converting a grayscale image to a binary image. It aims to separate the foreground from the background by maximizing the variance between these regions. The process involves computing the image's histogram, which illustrates the distribution of pixel intensities. Otsu's method iterates through all possible threshold values (0-255) to determine the maximum between-class variance threshold. The image is converted to the binary format using the ideal threshold, classifying pixels above the threshold as foreground (e.g., microcalcifications) and those below as background. The thresholding process is formulated as

$$Binary\ Image\ (x,y) = \begin{cases} 255, & if\ G(x,y) \geq Threshold \\ 0, & otherwise \end{cases} \quad (5)$$

The Sobel operator is a popular edge detection method in image processing [21], intended to draw attention to regions of the image with sudden fluctuations in intensity, usually found around edges. The Sobel operator, a discrete differentiation operator, roughly represents the gradient of the intensity function of the image. By detecting edges, it helps identify boundaries and structures within the image. The two components are combined to compute the overall gradient magnitude and direction. Sobel Kernels, the horizontal gradient measures changes along the rows (left to right) and the vertical gradient measures changes along the columns (top to bottom).

Horizontal gradient (Gx): $$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (6)$$

Vertical gradient (Gy): $$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (7)$$

After applying the Sobel kernels to the image, the gradient magnitude at each pixel is calculated as

$$G = \sqrt{G_x^2 + G_y^2} \quad (8)$$

The gradient direction, or orientation of the edges, is calculated as

$$\theta = \tan^{-1} \frac{G_y}{G_x} \quad (9)$$

The gradient direction θ shows the edge's direction, while the gradient magnitude G stands for the edge's strength. The Sobel operator is particularly effective for detecting edges with smooth intensity variations, making it useful in applications like image segmentation and feature extraction. The Sobel operator uses convolution, which helps suppress noise while detecting edges. It is less noise-sensitive than other methods but still effective at highlighting edges. The 3x3 convolution makes it computationally efficient but is limited in detecting very fine details. Based on the Prewitt operator, the Sobel operator uses weighted calculations to count differences across four neighbourhoods. The Sobel operator performs effectively when processing images with much noise and a gray gradient.

$$S_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \quad (10)$$

$$S_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4) \quad (11)$$

$$M = \sqrt{(S_x)^2 + (S_y)^2} \quad (12)$$

The direction of this point is expressed as,

$$\alpha(x, y) = \tan^{-1}(\frac{S_x}{S_y}) \quad (13)$$

The Euclidean norm is a key mathematical means in image processing, especially for calculating gradient magnitudes in edge detection and feature extraction tasks. It helps combine horizontal (G_x) and vertical (G_Y) gradients into a single value, representing the overall intensity change at a pixel. This is essential for identifying transitions like edges or boundaries in medical images, such as DBT slices, where detecting features like microcalcifications is critical. For a given pixel, the Euclidean norm combines these gradients using the equation,

$$G = \sqrt{G_x^2 + G_y^2} \quad (14)$$

To evaluate the texture and noise levels within an object, Root Mean Square (RMS) noise fluctuation quantifies the variability of pixel intensities, giving an idea of the texture and possible noise levels in the image. Using a mask to isolate the object, it computes the pixel values' standard deviation, representing how dispersed the pixel intensities are from the mean.

A high standard deviation indicates more noise or variation, while a low value suggests a smoother region. Subtraction of the object's brightest and darkest pixels is required to compute contrast, revealing its distinctness against the background. Shape descriptors like moments are employed to extract more structural features.

Moments, such as the centroid (from first-order moments) and the variance along the x- and y-axes (from second-order moments), provide detailed information about the shape and orientation of the object. A new dataset is generated from these extracted features, allowing a medical expert to offer insight for further processing and analysis. This dataset forms the foundation for classification and diagnostic tasks. The generated dataset is given in Figure 4.

3.2. Microcalcification Classification Model for DBT

The suggested breast cancer detection model uses LSTM as its base model. The microcalcification classification model using Stacked LSTM involves a multi-layered LSTM network architecture designed for sequential analysis of DBT slices.

	Area	Grey Level	Gradient Strength	Noise Fluctuation	Contrast	Shape Descriptor	Microcalcification
0	0.230020	5.072578	-0.276061	0.832444	-0.377866	0.480322	'-1'
1	0.155491	-0.169390	0.670652	-0.859553	-0.377866	-0.945723	'-1'
2	-0.784415	-0.443654	5.674705	-0.859553	-0.377866	-0.945723	'-1'
3	0.546088	0.131415	-0.456387	-0.859553	-0.377866	-0.945723	'-1'
4	-0.102987	-0.394994	-0.140816	0.979703	-0.377866	1.013566	'-1'

Fig. 4 Sample dataset

The input consists of pre-processed DBT images, which are converted to grayscale and enhanced using image processing techniques. Features like area, contrast, gradient, and texture are extracted from each slice. These features are fed sequentially into the Stacked LSTM layers, which capture temporal and spatial dependencies across multiple DBT slices. Figure 5 illustrates the complete block diagram of the proposed stacked LSTM based classification model.

3.2.1. Data Preprocessing and Exploratory Data Analysis

Data preprocessing, Exploratory Data Analysis (EDA), and microcalcification classification using stacked LSTM networks are essential to improve the model's performance [22]. The preprocessing phase involves several steps to ensure the dataset's quality and relevance. Raw DBT images are initially converted to grayscale to simplify the data while retaining essential intensity information.

Techniques like Gaussian blurring and histogram equalization reduce noise and enhance contrast, making microcalcifications more distinguishable. Otsu's thresholding is then applied to create binary images, effectively separating the areas of interest from the background.

Following preprocessing, EDA aims to analyze the dataset's properties, such as the distribution of microcalcifications and the balance between classes (microcalcifications present or absent). Visualization techniques such as scatter plots, box plots, and histograms identify patterns, anomalies, and correlations, guiding further model development and feature selection strategies. This comprehensive approach ensures that the data fed into the stacked LSTM model is well-prepared, ultimately improving classification accuracy.

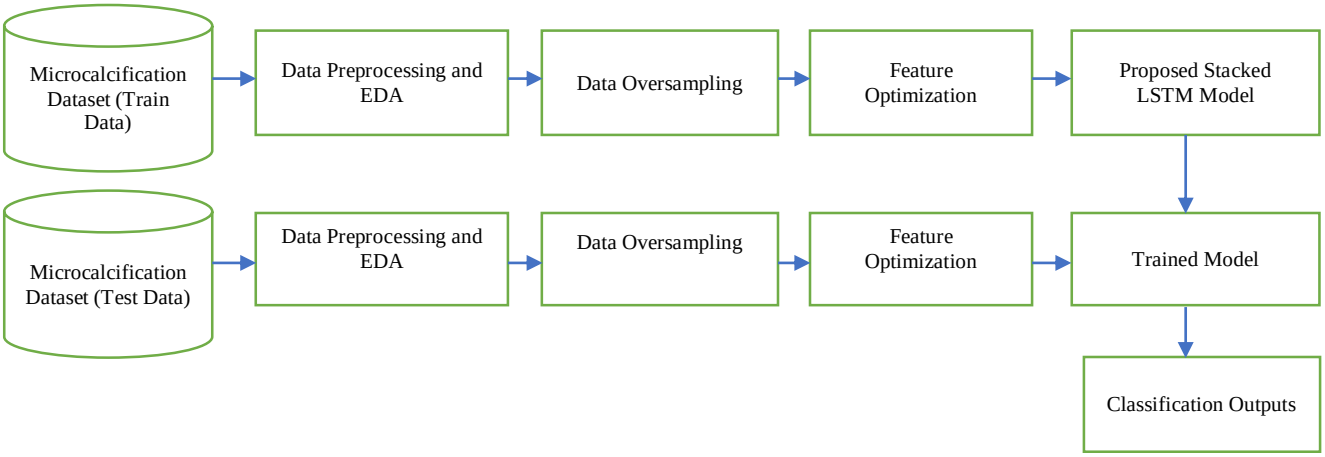
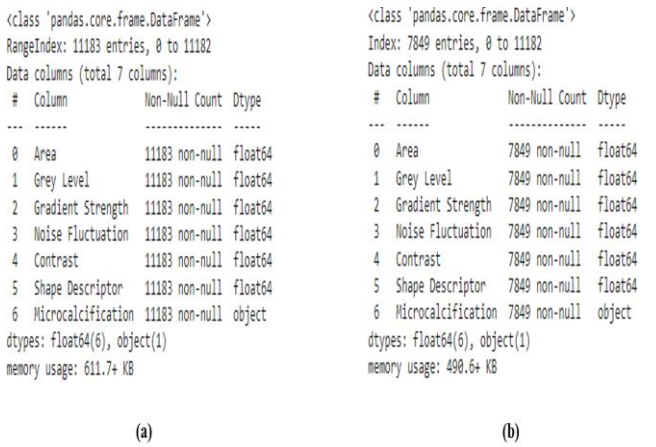


Fig. 5 Comprehensive block diagram of proposed stacked LSTM based classification model

After checking for null and duplicate values, the dataset is additionally refined to extract valuable insights. Missing data is addressed using techniques like mean, median, or advanced imputation methods like k-Nearest Neighbors (KNN). Excessive null data is removed if necessary. Duplicates are eliminated to prevent biased results, particularly in ML models. Statistical summaries such as mean, median, standard deviation, and range help identify patterns, outliers, and distributions. Multiple variable correlation coefficients are displayed in a correlation matrix ranging from -1 to +1. It helps identify relationships, feature selection, and potential multicollinearity in datasets. Visualizing these matrices with heatmaps reveals patterns, making them essential for data analysis and guiding ML decision-making, as shown in Figure 7.



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11183 entries, 0 to 11182
Data columns (total 7 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Area                11183 non-null  float64
 1   Grey Level          11183 non-null  float64
 2   Gradient Strength   11183 non-null  float64
 3   Noise Fluctuation   11183 non-null  float64
 4   Contrast            11183 non-null  float64
 5   Shape Descriptor    11183 non-null  float64
 6   Microcalcification  11183 non-null  object
dtypes: float64(6), object(1)
memory usage: 611.7+ KB
```

(a)

```
<class 'pandas.core.frame.DataFrame'>
Index: 7849 entries, 0 to 11182
Data columns (total 7 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Area                7849 non-null   float64
 1   Grey Level          7849 non-null   float64
 2   Gradient Strength   7849 non-null   float64
 3   Noise Fluctuation   7849 non-null   float64
 4   Contrast            7849 non-null   float64
 5   Shape Descriptor    7849 non-null   float64
 6   Microcalcification  7849 non-null   object
dtypes: float64(6), object(1)
memory usage: 490.6+ KB
```

(b)

Fig. 6 Data analysis (a) after null value checking, and (b) after duplicate value checking.

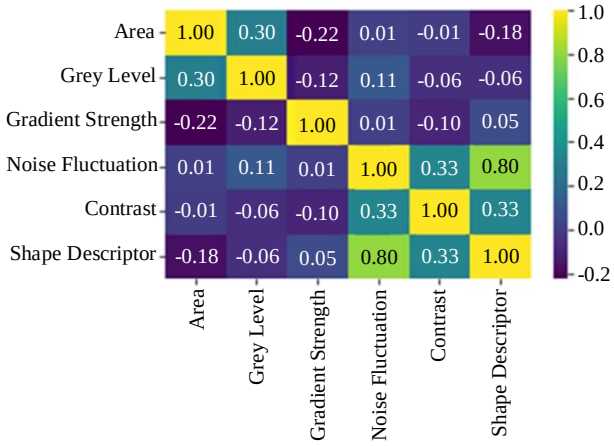


Fig. 7 Correlation matrix

Feature distribution in Figure 8 refers to how individual features (or variables) are spread across their possible values in a dataset [23]. Understanding the distribution of features is crucial in data analysis, as it provides insights into the nature of the data, including skewness, central tendencies (mean, median, mode), and the presence of outliers. Visualizing feature distribution using histograms, box plots, or kernel density plots helps to identify patterns and anomalies.

A pair plot is an efficient visualization tool that displays relationships between pairs of features in a dataset through a grid of scatter plots and histograms. A relationship between two variables is displayed in each grid scatter plot, helping identify correlations, trends, or potential outliers. Diagonal elements typically show histograms or density plots for each feature, as shown in Figure 9, offering insight into the distribution of individual variables. Figure 10 shows the contrast, noise fluctuation, gradient strength, and shape descriptor box plot.

3.3.2. Oversampling and Data Splitting

The data imbalance problem in the dataset is addressed using the random oversampling method. An approach is applied when there are substantially fewer instances of one class (minority class) than the other (majority class). Figure 11 shows oversampling approaches for dealing with imbalanced

classification in EDA. Because of this imbalance, the model may become biased in favour of the majority class, impairing model performance. Random oversampling tackles this by duplicating examples from the minority class at random, effectively balancing the class distribution. This approach increases the representation of the minority class without altering the existing data. After balancing the data, an 80:20 split of the dataset comprises training and test sets. Ensuring both sets have representative class distributions for model evaluation.

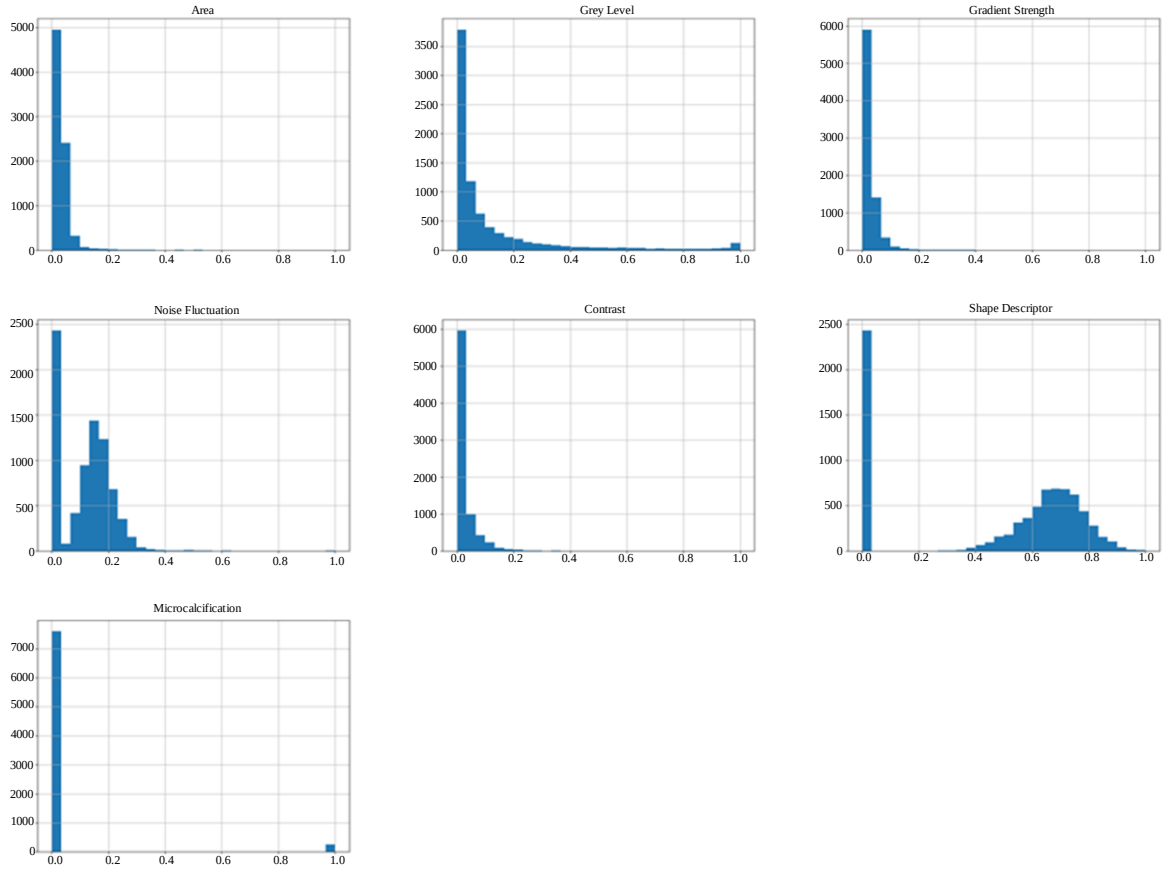


Fig. 8 Feature distribution

3.3.3. Feature Optimization Using Bayesian Optimization (BO) Approach

Bayesian Optimization (BO) was employed to optimize the hyperparameters of the Stacked LSTM model, including the number of layers, dropout rates, and learning rates. Unlike grid or random search methods, BO intelligently balances exploration and exploitation by constructing a probabilistic objective function model. Using a Gaussian process, BO identifies the most promising hyperparameter combinations with fewer iterations, reducing computational overhead. This approach significantly improved the model's performance by fine-tuning critical parameters, achieving high accuracy and efficient convergence. BO's application is useful in refining deep learning models for complex tasks like microcalcification classification. Because Bayesian Optimization (BO) can balance exploration and exploitation, it is a popular approach to transportation optimization problems [29].

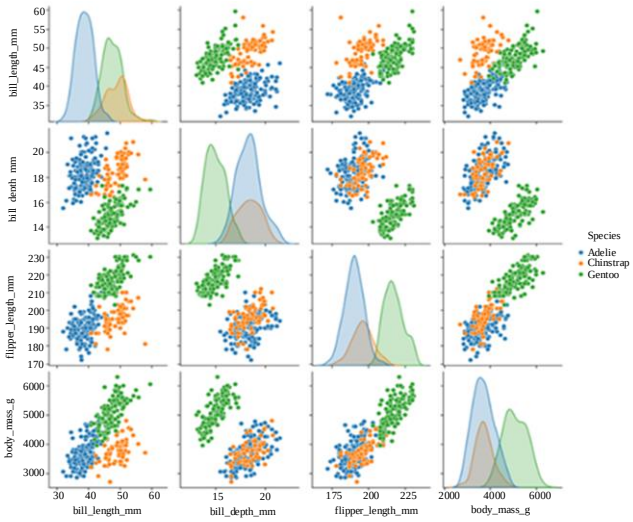


Fig. 9 Pair plot of features

Rather than directly assessing the function at each point, BO constructs a probabilistic model, typically a Gaussian process, to estimate the function's behavior based on previously evaluated points. This model allows BO to explore and exploit the search space intelligently. An acquisition function selects the next sampling point by balancing exploring uncertain areas with exploiting regions likely to yield improvements and iterative optimization that samples new points to refine the model's understanding of the function.

3.3.4. Proposed Microcalcification Classification Model Using Stacked LSTM

RNNs with stacked LSTM architecture are sophisticated and can efficiently identify temporal connections in sequential input [30]. Because of vanishing gradient issues, stacked LSTMs are not like standard RNNs; instead, LSTM cells are stacked on top of one another in several layers, which enables deeper learning and greater feature abstraction. Each layer of the LSTM keeps its capacity to learn from earlier time steps while transferring data to the subsequent layer. Thus enhancing the model's capacity to understand complex patterns over time.

The input data is processed sequentially, and the memory cells in LSTMs enable the model to maintain long-term information while forgetting irrelevant details. This architectural design is especially successful. Understanding context and sequential dependencies is essential for speech recognition, natural language processing, and time series forecasting. The model's prediction performance on various tasks is enhanced by stacking numerous LSTM layers, which provide a richer representation of the input data.

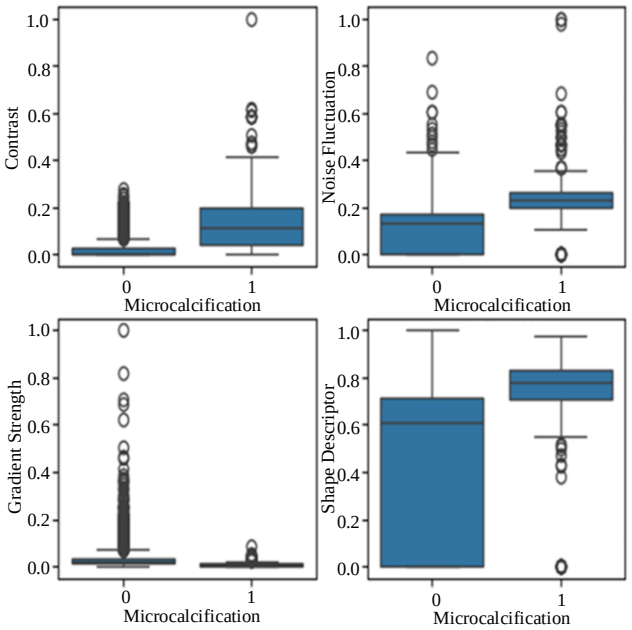


Fig. 10 Box plot of selected features

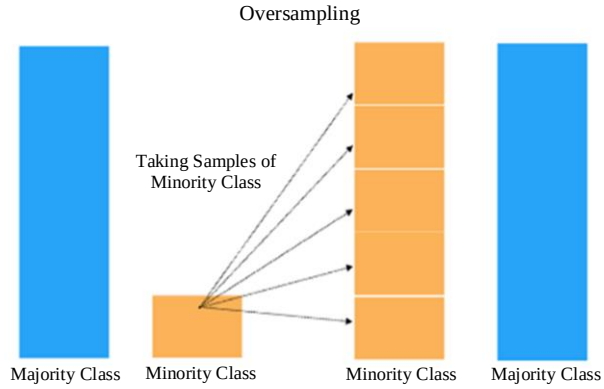


Fig. 11 Oversampling techniques for addressing unequal classification in EDA

Figure 12 shows that An LSTM unit usually consists of a cell, an input gate, an output gate, and a forget gate. Basic structure of LSTM network. A cell, an input gate, an output gate, and a forget gate are the usual components of an LSTM unit. The gates control how data enters and exit the cell, and the cell stores data for variable lengths of time. Using the previous state and current input as a guide, the forget gate assigns values between 0 and 1, where 1 denotes keeping the information, and 0 denotes discarding it. This makes it possible for the forget gate to select which data from the prior state to ignore. The input gate chooses which new data to store in the cell using the same method. Considering both the current and past states, it is possible for the forget gate to determine which data from the prior state to ignore. By maintaining essential long-term dependencies, this selective output improves prediction accuracy for both present and future time steps of the LSTM.

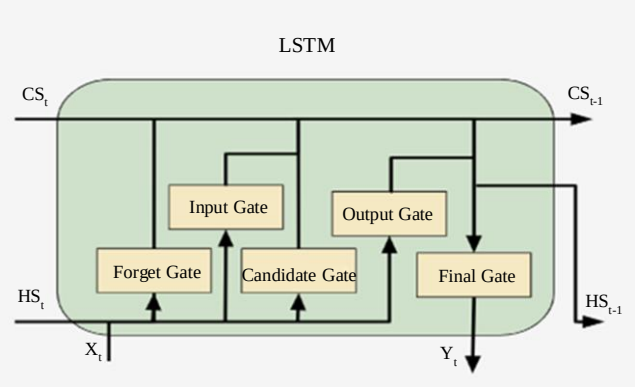


Fig. 12 Basic structure of LSTM network

In LSTM, the internal cell state is forwarded along with the hidden state. Where forget gate,

$$f_t = \sigma(X_t * U_f + H_{(t-1)} * W_f) \quad (15)$$

Whereas, Input Gate,

$$i_t = \sigma(X_t * U_i + H_{t-1} * W_i) \quad (16)$$

Updating cell state,

$$C_t = f_t * C_{t-1} + i_t * N_t \quad (17)$$

Where, $N_t = \tanh(X_t * U_c + H_{(t-1)} * W_C)$ (18)

Output gate,

$$O_t = \sigma(X_t * U_o + H_{(t-1)} * W_o) \quad (19)$$

Output relation is,

$$Output = Softmax(H_t) \quad (20)$$

Where input at the current timestamp t is represented as X_t, U_f *and* W_f *are* the weight matrix of the forget gate, and the activation function is denoted as σ, $H_{(t-1)}$ is hidden state of the previous timestamp, the weight matrix of input associated with the hidden state is denoted as W_i

The basic procedure of an LSTM network revolves around its ability to store and regulate information through a series of gates. Each LSTM cell maintains a cell state C_t, which acts as a memory and an output state h_t, representing the information passed to the next time step. These gates manage the information flow and decide what to retain, update, or discard. The model architecture of the proposed stacked LSTM.

In the proposed classification model, the LSTM layers are arranged sequentially, one on top of the other, with the prime objective of capturing complex patterns in the input data. The architecture starts with an LSTM layer consisting of 128 units and a tanh activation function. To address overfitting, a dropout layer with a dropout rate 0.1 comes after this LSTM layer.

Next, the second LSTM layer is introduced, consisting of 64 units, again using the tanh activation function, followed by another dropout layer with the same 0.1 dropout rate. The third LSTM layer contains 32 units, employing the tanh activation function. It is succeeded by a dropout layer with a 0.1 dropout rate. Finally, the model concludes with a dense layer using a sigmoid activation function.

It is perfect for binary classification applications because it offers an output probability ranging from 0 to 1. The basic algorithm for microcalcification classification using stacked LSTM with Bayesian optimization is given below.

Algorithm 1: Microcalcification detection using Stacked LSTM model

Input:
- DBT Slices
- Generated MC Classification dataset

Output: MC Classification Model

Begin:

1) Dataset Generation:
- Grayscale Conversion, $\text{Gray}(x,y) = 0.299.\text{R}(x,y) + 0.587.\text{G}(x,y) + 0.114.\text{B}(x,y)$
- Gaussian Blurring, $\text{G}(x,y) = \frac{1}{2\pi\sigma^2}\exp^{-\frac{x^2+y^2}{2\sigma^2}}$
- Histogram Equalization, $s_k = T(r_k) = \sum_{i=0}^{k}\frac{n_i}{n}$
- Binary Image Generation, $\text{Binary Image}(x,y) = \begin{cases} 255, \text{if } G(x,y) \geq T \\ 0, \text{Otherwise} \end{cases}$
- Contour Detection, $\text{Area of Object} = \sum_{p \epsilon \text{Contour}} 1$
- Average Gray Level of Object, $\text{Average Gray Level} = \frac{1}{|R|}\sum_{(x,y)\in R}\text{Gray}(x,y)$
- Gradient Strength using Sobel Operator:
- $S_x = (a_2 + c.a_3 + a_4) - (a_0 + c.a_7 + a_6)$
  $S_y = (a_0 + c.a_1 + a_2) - (a_6 + c.a_5 + a_4)$
- $M(x,y) = \sqrt{S_x^2 + S_y^2}$
- $\text{Gradient Strength} = \frac{1}{|R|}\sum_{(x,y)\in R} M(x,y)$
- RMS Noise Fluctuation, $\sqrt{\frac{1}{|R|}\sum_{(x,y)\in R}(\text{Gray}(x,y) - \mu_R)^2}$
- Contrast, $\text{Contrast} = \max_{(x,y)\in R}\text{Gray}(x,y) - \min_{(x,y)\in R}\text{Gray}(x,y)$
- Shape Descriptors using Moments
- Feature Aggregation, Features [area of object, gray level, gradient strength, noise fluctuation, contrast, shape descriptor]
- Save the feature vectors for all DBT Slices into a data frame.
- Save the data frame as a CSV file

2) Loading and Preprocessing of Generated Dataset:
- $D = \{(e_i, y_i)\}_{i=1}^{n}$ where $y_i \epsilon \{0,1\}$
- D = D. drop_duplicates ()

3) Label Encoding, Data Resampling and Data Splitting:
- Encode.fit_transform(D)
- D_class_1_samp = D_class_1. Sample (n = 7595, replace = True)
- D_test_samp = concat ([D_class_0, D_class_1_samp])
- $D_{train}, D_{test} = \text{train\_test\_split}(D, \text{test}_{size} = 0.2)$

4) Feature Optimization:

- $\max_{\theta \epsilon \Theta} f(\theta) = \text{Accuracy}$
- Bayesian optimization bounds, $\eta \epsilon [10^{-4}, 10^{-2}], \text{dropout}_{\text{rate}} \epsilon [0.05, 0.5]$
- Optimizer.step (θ)
- Hyperparameter search, $\theta_i = \arg\max f(\theta_i)$

5) Define Stacked LSTM Model:

M = Sequential ()
M. add (LSTM (128, input (), activation = 'tanh'))
M. add (Dropout (0.1))
M. add (LSTM (64, activation = 'tanh'))
M. add (Dropout (0.1))
M. add (LSTM (32, activation = 'tanh'))
M. add (Dropout (0.1))
M. add (Dense (1, activation = 'sigmoid'))

Model Compilation:

M. compile (Optimizer = 'adam' , loss='binary_crossentropy' , metrics='accuracy')

1) Model Training:

M. fit(train_data, val_data, Epochs, Batch Size)

2) Model Evaluation:

M. evaluate (test_data)

3) Save the Model:

End

Table 2. The proposed stacked LSTM model's hyperparameter specifications

Hyperparameters	Value
Batch Size	500
Number of Epochs	500
Activation Function	Sigmoid, tanh
Loss Function	Binary Cross Entropy
Optimizer	Adam

3.3.5. Hardware and Software Setup

An NVIDIA GeForce GTX 1080Ti GPU, an Intel Core i7 processor, 32GB of RAM, and the Python-based Kera's library combined with the TensorFlow framework comprised the extensive configuration used in the study. The combination of Google Collab's extensive computing capacity and Keras's interface ensures improved model construction with GPU acceleration. The proposed classification model has been trained and tested on the Google Collaboratory platform with Python. The hyperparameters utilized in the training, which are essential for maximizing the accuracy and effectiveness of the model, are shown in Table 2. These settings are made before the model's execution to control the training procedure, guarantee effective dataset management, and promote quicker convergence.

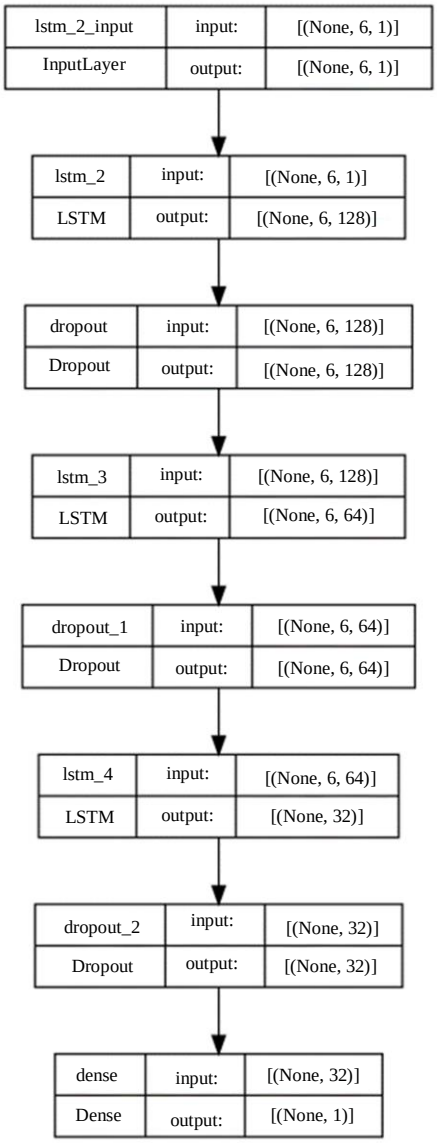


Fig. 13 Architecture of the proposed stacked LSTM model

4. Results and Discussion

Performance parameters for the model are described by measuring accuracy, precision, recall, and f1 score. Following is the mathematical expression for it,

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (21)$$

$$Precision = \frac{(TP)}{(TP+FP)} \quad (22)$$

$$Recall = \frac{(TP)}{(TP+FN)} \quad (23)$$

$$F1-Score = 2\times \left(\frac{(Precision \times Recall)}{(Precision+Recall)}\right) \quad (24)$$

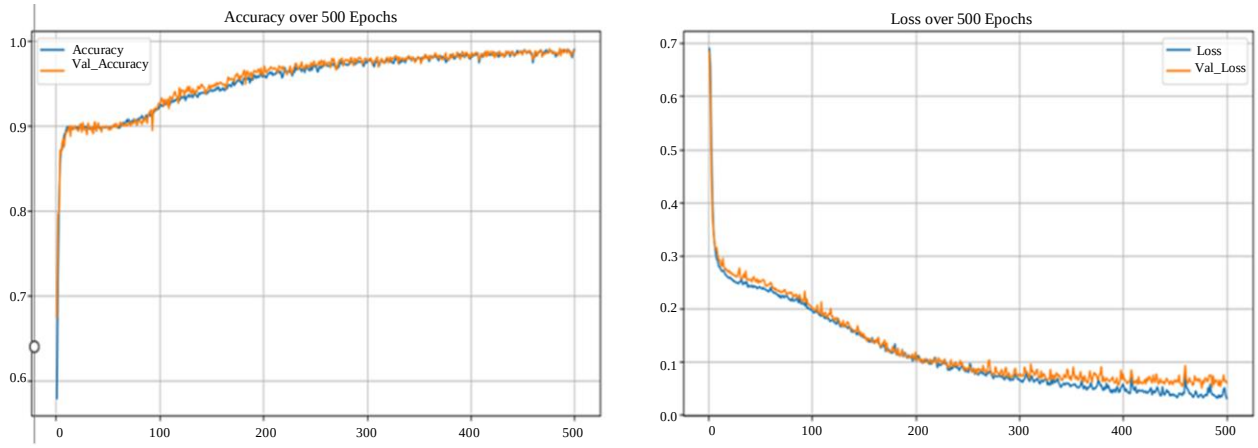


Fig. 14 (a) accuracy plot, and (b) loss plot.

An accuracy plot visualizes the model's performance over time, showing how well it predicts correctly as training progresses, as shown in Figure 14. The loss plot tracks the error or loss during training, with lower values indicating better performance. The model demonstrated a strong performance during training, achieving a final training accuracy of 98.92% and a validation accuracy of 98.75% by the 500th epoch.

The loss steadily decreased, with the final training loss reaching 0.0308 and the validation loss dropping to 0.0604. These results indicate that the proposed model generalizes well to unseen data, showcasing its effectiveness for accurate classification tasks.

The classification report in Table 3 indicates high performance, with precision, recall, and F1 scores of 0.99, 0.96, and 0.97 for class 0 and 0.96, 0.99, and 0.97 for class 1. The overall accuracy is 0.9727, with macro and weighted averages also at 0.97, reflecting balanced model performance.

Table. 3 Classification Report of Proposed Stacked LSTM Model

	F1 Score	Recall	Precision
0	0.97	0.96	0.99
1	0.97	0.99	0.96
Accuracy	0.9727		
Macro avg	0.97	0.97	0.97
Weighted avg	0.97	0.97	0.97

For assessing the effectiveness of a classification model, a confusion matrix is used. It offers a graphic depiction of true positive, true negative, false positive, and false negative rates, enabling a more thorough comprehension of how well the model differentiates across classes. Out of 1519 true negatives (class 0), the model correctly predicted 1454 but misclassified 65 instances. Similarly, for the 1519 true positives (class 1), the model correctly classified 1501, with only 18 misclassifications. This suggests high precision in classifying both groups with few errors.

The Receiver Operating Characteristics (ROC) curve in Figure 16 illustrates the model's performance across several thresholds. The ROC curve expresses the trade-off between sensitivity and specificity in terms of probability. An AUC number closer to 1 indicates better performance and shows the model's performance across all threshold levels. The threshold in a perfect model has a TP rate of "1" and an FP rate of "0."

The simulation results indicated that the proposed classification models attained impressive performance. The Stacked LSTM model with the Bayesian Optimization approach attained an accuracy of 97.27 %. The utilization of DL in DBT analysis ultimately leads to improved outcomes for patients undergoing breast cancer screening and diagnosis by improving diagnostic accuracy and potentially streamlining workflow efficiency in clinical settings.

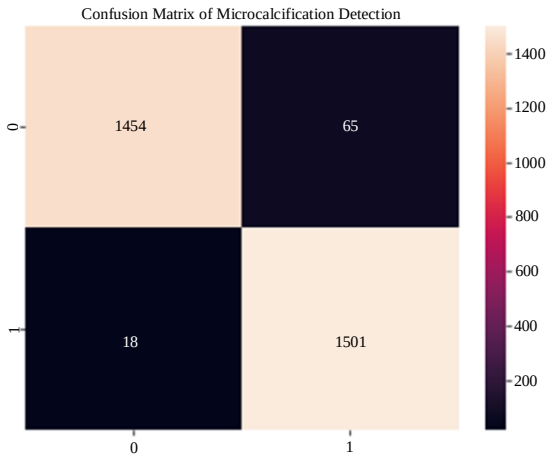


Fig. 15 Confusion matrix of the proposed stacked LSTM model

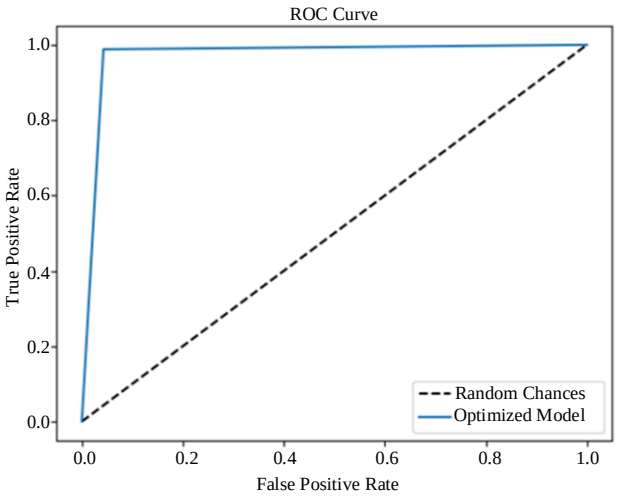


Fig. 16 ROC curve of proposed stacked LSTM model

Table 4 summarizes various methodologies and their respective accuracy rates in the classification of breast cancer, showcasing diverse ML and DL models used by different researchers. CNN achieved an accuracy of 92%, demonstrating the effectiveness of convolutional networks in image analysis. They obtained 74.7% accuracy using ResNet101, which, while lower than others, still highlights the utility of deep residual networks. Hybrid models like SVM, Random Forest, and XGBoost reached 90%, while ANN outperformed others with an accuracy of 93%. The combination of SVM, RF, and ANN, as demonstrated, reached an impressive 95%. The highest accuracy of 97.27% was achieved by the proposed stacked LSTM method, outperforming traditional CNNs and other DL models. These results in Figure 17 reflect the increasing success of DL methods in improving diagnostic accuracy in breast cancer detection.

Table 4. Performance comparison in terms of detection accuracy

Authors & Year	Methodology	Accuracy (%)
Oladasu et.al (2024) [24]	CNN	92
Francesco Prinz (2024) [6]	SVM, RF, and XGBoost.	90
A Mračko et al (2024) [25]	ResNet101	74.7
JR Teoh et al. (2024) [26]	DL models (AlexNet, ResNet-50, VGG16, GoogLeNet)	90
Filippo Pesapane et al. (2023) [27]	AlexNet, ResNet18, and ResNet34	87
Zahra Maghsoodzadeh et al. (2023) [16]	ANN	93
M.C. Shanker et al. (2023) [28]	SVM, RF, and (ANN)	95
Qing Lina et al. (2023) [15]	FPNNet and SPPNet	81
Proposed method (stacked LSTM with BO)		**97.27**

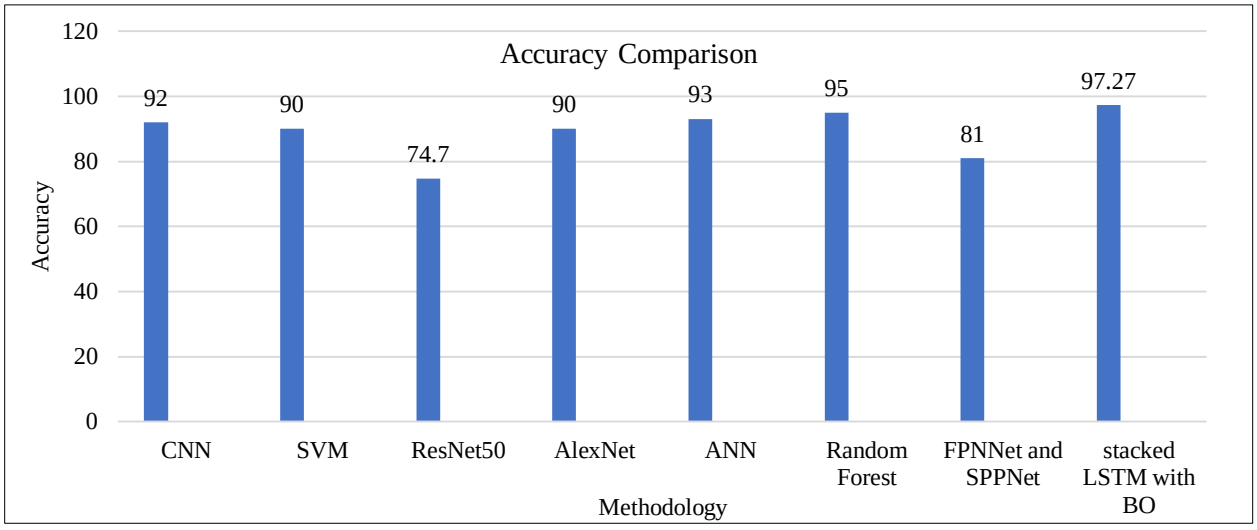


Fig. 17 Graphical visualization of performance comparison of an existing model in terms of detection accuracy

5. Conclusion

Worldwide, breast cancer is the leading cause of cancer-related deaths in women. Early diagnosis is essential for both successful therapy and higher survival rates. The advancement of radiographic breast imaging and screening has enabled more precise diagnoses, helping to identify breast diseases. In order to improve early breast cancer detection, an advanced DL framework is developed in this study for categorizing Microcalcifications (MCs) in Digital Breast Tomosynthesis (DBT) images. The research utilized a stacked LSTM model optimized with Bayesian methods, and its performance was compared against various other ML and DL models. The proposed method achieved a remarkable accuracy of 97.27%, surpassing other models. These findings demonstrate the efficacy of DL models, particularly the stacked LSTM, in automating and improving the precision of MC classification.

The study underscores the potential of the proposed methodology in clinical applications, offering a significant advancement in breast cancer detection and diagnosis. Extending the model to analyze other imaging modalities, such as MRI and ultrasound, could also open new avenues for comprehensive cancer detection. Investigating techniques to reduce computational requirements will also facilitate the deployment of the model in resource-limited clinical settings, broadening its impact on early cancer diagnosis.

Acknowledgments

I want to express my sincere gratitude to all those who contributed to completing this research paper. I extend my heartfelt thanks to my supervisor, family, colleagues and fellow researchers for their encouragement and understanding during the demanding phases of this work.

References

[1] Asita Elengoe et al., "A Short Review on Breast Cancer," *International Journal of Biotechnology and Biomedicine*, vol. 1, no. 1, pp. 1-11, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Kyril L. Cole et al., "Assessing Survival in Non-Small Cell Lung Cancer Brain Metastases after Stereotactic Radiosurgery: Before and after the Start of the Targetable Mutation Era," *Journal of Neuro-Oncology*, vol. 169, no. 3, pp. 671-681, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] J. Haskell, "*Raman Spectroscopy Techniques for the Detection and Management of Breast Cancer*," Doctoral Dissertation, University of Exeter, United Kingdom, 2024. [Google Scholar]

[4] M.T. Ramli Hamid et al., "Abbreviated Breast Magnetic Resonance Imaging (MRI) or Digital Breast Tomosynthesis for Breast Cancer Detection in Dense Breasts? A Retrospective Preliminary Study with Comparable Results," *Clinical Radiology*, vol. 79, no. 4, pp. e524-e531, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ashish Shiwlani et al., "BI-RADS Category Prediction from Mammography Images and Mammography Radiology Reports Using DL: A Systematic Review," *Jurnal Ilmiah Computer Science*, vol. 3, no. 1, pp. 30-49, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Francesco Prinzi et al., "Interpretable Radiomic Signature for Breast Microcalcification Detection and Classification," *Journal of Imaging Informatics in Medicine*, vol. 37, pp. 1038-1053, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Marwa Ben Ammar et al., "Harnessing Deep Learning for Early Breast Cancer Diagnosis: A Review of Datasets, Methods, Challenges, and Future Directions," *International Journal of Computing and Digital Systems*, vol. 16, no. 1, pp. 1643-1661, 2024. [Google Scholar]

[8] Laxman Singh, and Altaf Alam, "An Efficient Hybrid Methodology for an Early Detection of Breast Cancer in Digital Mammograms," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 1, pp. 337-360, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Paul Terrassin et al., "Annotation-Free Deep-Learning Framework for Microcalcifications Detection on Mammograms," *Medical Imaging 2024: Computer-Aided Diagnosis*, vol. 12927, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Jinhua Wang et al., "Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning," *Scientific Reports*, vol. 6, no. 1, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[11] Redona Brahimetaj et al., "Improved Automated Early Detection of Breast Cancer Based on High Resolution 3D Micro-CT Microcalcification Images," *BMC Cancer*, vol. 22, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Koushlendra Kumar Singh et al., "Deep Learning Capabilities for the Categorization of Microcalcification," *International Journal of Environmental Research and Public Health*, vol. 19, no. 4, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Chrysostomos Marasinou et al., "Improving the Quantitative Analysis of Breast Microcalcifications: A Multiscale Approach," *Journal of Digital Imaging*, vol. 36, no. 3, pp. 1016-1028, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ehtsham Rasool et al., "Breast Microcalcification Detection in Digital Mammograms Using Deep Transfer Learning Approaches," *Proceedings of the 2023 9th International Conference on Computing and Data Engineering*, pp. 58-65, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Qing Lin et al., "Artificial Intelligence-Based Diagnosis of Breast Cancer by Mammography Microcalcification," *Fundamental Research*, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Zahra Maghsoodzadeh Sarvestani et al., "A Novel Machine Learning Approach on Texture Analysis for Automatic Breast Microcalcification Diagnosis Classification of Mammogram Images," *Journal of Cancer Research and Clinical Oncology*, vol. 149, no. 9, pp. 6151-6170, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ana M. Mota et al., "Automatic Classification of Simulated Breast Tomosynthesis Whole Images for the Presence of Microcalcification Clusters Using Deep CNNs," *Journal of Imaging*, vol. 8, no. 9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Nicholas T. Smith, Jonathan W. Merritt, and Emrys R. Phillips, "High-Resolution 3D Geological Modelling of Heterogeneity in Poorly Exposed Glacial Deposits Using Sedimentary and Glaciotectonic Architectural Element Analysis: A Case Example from Sellafield in West Cumbria, UK," *Quarterly Journal of Engineering Geology and Hydrogeology*, vol. 56, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Barsha Abhisheka et al., "Recent Trend in Medical Imaging Modalities and their Applications in Disease Diagnosis: A Review," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 43035-43070, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Deepak S. Uplaonkar, Virupakshappa, and Nagabhushan Patil, "Modified Otsu Thresholding-Based Level Set and Local Directional Ternary Pattern Technique for Liver Tumor Segmentation," *International Journal of System Assurance Engineering and Management*, vol. 15, no. 1, pp. 73-83, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Taoufik Saidani et al., "Design and Implementation of a Real-Time Image Processing System Based on Sobel Edge Detection Using Model-Based Design Methods," *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 3, pp. 273-278, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mohan Raparthi et al., "Exploratory Data Analysis Techniques-A Comprehensive Review: Reviewing Various Exploratory Data Analysis Techniques and their Applications in Uncovering Insights from Raw Data," *Australian Journal of Machine Learning Research & Applications*, vol. 4, no. 1, pp. 215-225, 2024. [Google Scholar] [Publisher Link]

[23] Enrique R. Sebastian, Julio Esparza, and Liset M. de la Prida, "Quantifying the Distribution of Feature Values over Data Represented in Arbitrary Dimensional Spaces," *PLOS Computational Biology*, vol. 20, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Oladosu Oyebisi Oladimeji, Ian McLoughlin, and Saritha Unnikrishnan, "Volumetric Attention Mechanism-Based Deep Learning for Breast Cancer Diagnosis in Digital Breast Tomosynthesis," *Proceedings of Ninth International Congress on Information and Communication Technology*, pp. 231-241, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Adam Mračko, Ivan Cimrák, and Lucia Vanovcanová, "Enhancing Breast Microcalcification Classification: From Binary to Three-Class Classifier," *2024 35th Conference of Open Innovations Association (FRUCT)*, Tampere, Finland, pp. 473-481, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Jing Ru Teoh et al., "Enhancing Early Breast Cancer Diagnosis through Automated Microcalcification Classification Using an Optimized Ensemble Deep Learning Framework," *PeerJ Computer Science*, vol. 10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] Filippo Pesapane et al., "Deep Learning Performance for Classification and Classification of Microcalcifications on Mammography," *European Radiology Experimental*, vol. 7, no. 1, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] M.C. Shanker, and M. Vadivel, "Micro-Calcification Classification Analysis in Mammogram Images with Aid of Hybrid Technique Analysis," *Wireless Personal Communications*, vol. 128, no. 2, pp. 1287-1307, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Zi Wang et al., "Pre-Trained Gaussian Processes for Bayesian Optimization," *Journal of Machine Learning Research*, vol. 25, pp. 1-83, 2024. [Google Scholar] [Publisher Link]

[30] Arpita Maharatha et al., "Employing Sequence-to-Sequence Stacked LSTM Autoencoder Architecture to Forecast Indian Weather," *Procedia Computer Science*, vol. 235, pp. 2258-2268, 2024. [CrossRef] [Google Scholar] [Publisher Link]