INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Enhancing Breast Cancer Data Classification Performance in Medical Data by Employing GRU-LMA Deep Learning Classifier for Accurate Diagnosis

S Vibin Chandar[1]*, **V Krishnapriya**[2]

1 Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India
2 Associate Professor & Head, Department of Computer Science with Cognitive Systems, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India

OPEN ACCESS

Received: 19-09-2023
Accepted: 01-10-2023
Published: 20-12-2023

Citation: Chandar SV, Krishnapriya V (2023) Enhancing Breast Cancer Data Classification Performance in Medical Data by Employing GRU-LMA Deep Learning Classifier for Accurate Diagnosis. Indian Journal of Science and Technology 16(47): 4490-4503. https://doi.org/10.17485/IJST/v16i47.2387

* **Corresponding author**.

vibinchandars@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To suggest a new data classifier to enhance breast cancer data classification in mammography digital images (MDI) in order to reduce the error rate and training time and improve accuracy in breast cancer prediction. A deep learning approach is utilized to identify the significant deep spots in MDI for accurate diagnosis. **Methods:** The Gated Recurrent Unit (GRU)-based DL approach is employed to speed up the learning phase in the DRNN classifier, and the Levenberg-Marquardt algorithm (LMA) is used to identify and reduce errors during classification. For feature extraction, noise removal, & filtering, Adaptive Median Filtering (AMF) and Tetrolet transform (TTM) algorithms are utilized, and the features are separated into two categories (morphology & texture). The King Abdulaziz University Breast Cancer Mammogram Dataset (KAUBCMD) dataset is used for this research work. KAUBCMD has 1416 case reports, with 2 views for both the right and left breasts, and 5662 MDI with clinical results of 3 target classes (negative, benign, and malignant) and 2 subclasses (incomplete and suspected malignant). MATLAB tool is used to evaluate the DL-based GRU-LMA approach based on performance metrics. The detailed comparative analysis is done with prevailing classifier models such as linear SVM, KNN-HP, and K-Means GRI. **Findings:** The suggested DL-based GRU-LMA classifier model shows proven results with 95.19% accuracy, 96.03% TPR and 95.09% TNR, 95.07% sensitivity, 97.81% specificity, 96% precision, 97.03% recall, 17 minutes of training time, 3% error rate, 3% FPR, and 4% FNR. GRU-LMA outperforms the current classifier models L-SVM, KNN-HP, and K-Means GRI. **Novelty:** The proven results show breast cancer prediction and classification accuracy with minimal time and error rate, which helps the radiologist and other BC clinical experts make diagnoses easily. The GRU-LMA addresses the limitations of the prevailing classifier models, L-SVM, KNN-HP, and K-Means GRI.

Keywords: Deep Learning; Recurrent Neural Networks; Classification; Machine Learning; Breast Cancer Prediction; RNN Classifier

1 Introduction

Breast cancer is one of the most life-threatening diseases for women, with numerous risk factors that originate in the cells of breast tissue and lead to cancer-related deaths worldwide. Early identification of breast cancer is highly essential to the fight against the disease. As the technology evolves, various computational ML and DL classifier models are introduced to increase the treatment options by utilizing the mammogram digital images for predicting disease and making treatment easier. An increase in error rate, non-identification of significant spots in MDI, high training time, and low accuracy are the key research gaps identified in the prevailing models. In addition, current approaches have minimal drawbacks like the non-deduction of complete morphology and texture changes, high false rates, etc. To overcome the drawbacks, this research study mainly focused on enhancing the accuracy of detection & classification in an optimized manner, minimizing error rate and training time by employing a new DL approach called GRU-LMA. GRU-LMA (Gated Recurrent Unit, Levenberg-Marquardt algorithm) is an RNN-based classifier employed along with AMF, TTM, and AHE. Here, the GRU-LMA model identifies deep spots in the breast tissue and optimizes after MDI enhancement (FS & FE). In addition to this, the features are separated into two categories, which are used for training and testing purposes. The reconstruction of deep spots in breast tissue tested for classification and GRU-LMA produces results in five classes, which include negative (1), benign (2), malignant (5), and two subclasses, including mammography-incomplete (0) and suspicious malignant (4).

Breast cancer diagnosis [(1)] by employing DL-based digital computational pathology classifiers specially focused on machine-to-machine learning processes. Also, existing DL models are reviewed, and results are compared. The authors highlight all the MDI features and challenges of DL-based diagnosis in a real-time scenario. The formalin-fixed tissues are spotted by deep learning models for efficient detection, but the model slowdown when using complex datasets is one of the major drawbacks. Radiographic digital images and ultrasound clinical reports are utilized by the authors [(2)] to identify breast cancer detection at an early occurrence. Deep RL concepts were reviewed with the publicly available BC dataset, which includes image modalities and USM patterns to detect cancer. Difficulties and risks were discussed, and the results were demonstrated. It is observed that the false rate is high, which has to be reduced in future enhancements. CAD tools [(3)] are used for implementation purposes in early diagnosis to classify different types of abnormalities. The Softmax classifier is employed to classify malignant and normal cases, and the HMF classifier is employed to detect sub-classes like negative or incomplete cases. 87% accuracy is achieved by DCNN-Softmax with limited drawbacks, which include high AUC-ROC and complex data handling. Along with DL and ML, transfer learning [(4)] models are employed to identify cancer tumors automatically with the use of three mammogram datasets. UI and HI images are used additionally to mark the ROI for efficient detection and classification. The model resulted in a high accuracy of 92%. Here, complex data handling is made easy, but deep spot identification is not achieved, which is considered a shortcoming. Dynamic link failure detection was introduced with a science-based algorithm called QoS-IWD-ARP [(5)], where the model utilized Alex Net and produced efficient feature selection from huge datasets like BCMD, FOGG-BC, etc. The model helps the deep learning approach optimize the image or links in a dynamic way by extracting necessary features from datasets. A fully automated BC detection model using thermo-grams [(6)] and AI-based DL and ML models [(7)] are utilized for BC detection, where a U-Net network is used to isolate the significant spots in breast tissue from the rest of the region. A benchmark

has been created in image segmentation and masking with the help of AI models that are trained for automatic detection and to handle complex data. A few limitations were discussed by the authors: masking the thermo-grams is quite a challenge compared to annotating the image. MRI images were used for cancer detection [8] by utilizing deep learning models and PRISMA guidelines. The screening of MRI is done with the help of DCNN, and selected features are marked as patterns, which are then used for training and testing purposes. Knowledge-driven feature learning and IM were used to boost the accuracy levels where pattern matching is accurately done in the real-time process. RVSRP-ML [9] model was introduced by the author, where the bioinspired method is used for optimization of dynamic links and RF, MLP, and GBT [10] are employed for efficient detection of cancer tissues in real-time by training and testing the 5179 independent clinical records. These models were trained initially with all features, and after the decoding process, accuracy was achieved. The drawbacks noted here are that the false alarm rate during complex data handling is high. A gated attentive MM-DL model was introduced [11] to predict cancer in two phases. The first phase generates the stacked F-model, and the second phase uses dropout features for bi-model attention. All the drop-out features are gated in the repository to identify the morphology changes to train and test the model to boost the accuracy level. The drawback of this model is the non-identification of deep tissue spots, where the classification accuracy might differ in the prognosis stage. InceptionResNetV3 [12] DL-based model is used for risk prediction of breast cancer, where the breast tissue texture is used for comparative analysis and feature extraction. The model is actually based on transfer learning to enhance the risk assessment scores in terms of boosting the accuracy level of prediction at an early stage. RAB-CRP model [13] was introduced by the authors, where a novel bio-inspired population-based technique is employed for robust optimization to detect link failure in search space, and it is used in some of the image annotation in disease prediction methods to enhance optimization. Breast tumors are identified by the CGRS system [14], where cluster methods are used and cluster tumors are segmented based on features extracted from the mammogram digital image dataset. E-CSDNN is employed to classify the cancer categories into six. The model has minor shortcomings where it cannot process complex data of more than 3000 MDI. Extreme learning machine-based DenseNet121, CNN, and TL models [15–17] wereintroduced to detect cancer tissue at the prognosis stage. The model works on 2-LC, where the target classes are classified into 2 levels, which include feature extraction and spot identification of BC tissues at the initial stage. The accuracy has increased to 92%. A D-CNN-based approach was employed by the authors for efficient feature selection and extraction to boost the classification level. A homo-morphed adaptive model is employed along with enhanced D-CNN to select all the global and local features to reduce the false alarm rate in cancer prediction. These models set a benchmark for deep learning disease prediction algorithms in terms of boosting precision and recall performance metrics. Linear-SVM [18] model was proposed for detection and classification, with LDA employed for FS and FE. PCA was employed to extract the deep spots in breast tissue and find out the missing values in the mammogram images. The accuracy is achieved up to 89% with an increased false rate. The RFC model was also deployed along with L-SVM to identify the morphology changes in MDI. The KNN-HP machine learning model [19] was proposed as a problem-solving technique where the nearest neighbor states are identified in the mammogram dataset and features are trained and tested to enhance the robustness of the model. The drawbacks of KNN-HP are that it could not handle large datasets where ROI is measured for efficient classification. K-Means GRI [20] method was developed with the objective of minimizing the error rate and training time, but the levels were not achieved due to high-order datasets. The 2C-Multiclass SVM [21] model is trained for tumor tissue detection within the MRI dataset for efficient classification and to enhance images in two classes. Both classes are extracted, and all the histogram features are selected. The KAUBCMD [22] clinical dataset was released by the clinical experts for breast cancer classification purposes in six categories. The improvised Firefly model and the Deep CNN [23,24] model were introduced for optimization and image classification to overcome the drawbacks of feature selection and extraction. These models have minimal drawbacks in terms of ROI levels, TPR, and TNR rates. To overcome the shortcomings of the prevailing BC classifier models, the new RNN-based GRU-LMA classifier is proposed as a problem-solving model to detect breast tissue and classify it in a robust manner. The core objectives of new classifier model GRU-LMA are: i) breast cancer detection; ii) classification & optimization of MDI; iii) identification of deep spots in breast tissue; iv) MDI enhancement using AHE; v) efficient FS & FE; vi) enhancing classification accuracy rate, etc. The objectives are achieved by GRU-LMA with the following steps,

- **Image Annotation:** Adding metadata to mammogram images, which includes classification and categorization into groups, will train the MDI efficiently.
- **Deep Filtering and Noise Removal:** Removal of unwanted artifacts in the KAUBCMD mammogram images to efficiently identify the patterns and anomalies within the MDI during testing process.
- **MDI Enhancement:** Enhancing MDI to highlight specific features and spots for high TPR & TNR.
- **Spotting of significant spots :** To reduce false rates, specific cancer tissues are masked and recorded as a pattern during the training process. In the follow-up tests, deep spots are masked for pattern matching to boost the accuracy level in classification and detection.

2 Methodology

The proposed deep learning-based RNN classifier, GRU-LMA, focuses on breast cancer detection and classification in a robust way by employing mammogram digital images. Complex features like micro-calcifications, masses and tumors, the structure of breast tissue, subtle changes in size, shape, and texture, hidden lesions, and some of the local handcrafted features are extracted from the KAUBCMD digital dataset for training and testing purposes. The GRU-LMA is employed to identify the deep spots in breast tissue, and ROI is measured. ROI separates the suspicious cancer tissues from the normal. Morphology and texture features are divided into two categories to test their similarity during classification. The deep segmentation process is carried out, where LMA optimizes the ROI efficiently to minimize the loss functions and improve accuracy. The training and testing process is done in both raw and contour images to record the deep significant spots. This deep learning GRU-LMA approach boots detection and classification accuracy in a robust manner, which is compared against existing models such as linear SVM[18], KNN-HP[19], and K-Means GRI[20].

2.1 Proposed Methodology

The new GRU-LMA RNN-based classifier model is trained and tested for breast cancer classification and detection for accurate diagnosis. Data selection and pre-processing is the first phase of GRU-LMA, where AMF is utilized for noise removal in the mammogram images (MDI). The FS and FE processes are carried out, and two categories are bifurcated, which include morphology and texture. Morphology features include the shape and inner structure of the breast tissue, and texture features include patterns and textures in the tissue. Global and local features are extracted from the KAUBCMD dataset to compare the training and testing data and identify the most significant ones for classification. The Gated Recurrent Unit (GRU) is employed to speed up the learning phase, identify the deep, significant spots in the breast tissue, and record the measurements. The breast tissue was then divided into two categories: suspicious and normal. The Levenberg-Marquardt algorithm (LMA) is employed to optimize the neural parameters and reduce the error rate during the classification. MDI image enhancement is done using adaptive histogram equalization. Then the optimization and classification process is carried out to identify the classes and sub-classes in a robust manner. Assume that *mi* represents the mammogram digital images and *image* $i = (1,2,3,\ldots..n)$ is the grid of pixels and each pixel is considered as feature. Here sigmoid is the activation function to record the height and width if the *mi*. Here H,W is the height and width. The GRU consists of 2 gates, update and reset gate. These gates control the mammogram refracted information flow through the network. Train the model $train\_gru\_lma\_model(model,\ training\_data,\ validation\_data,\ LMA)$. The equation for BC tissue bifurcation using GRU is as follows,

$$GRUBC_{Model} = Build\ (Input\,(SHAPE)\,,\ Num\,(CLASSES) + Calc\,[\,mi\ (image\ 1,\ldots n)] \tag{1}$$

where, $GRUBC_{Model}$ is the trained information of the KAUBCMD mammogram images. The result of GRU-LMA classifier is *Shape and Classes*. The significant deep spots are marked by AMF and the selected abnormal malignant tissues are separated from normal ones. Here the loss function is carried out by the following equation,

$$AMF\_loss\ =\ MDI\ (features,\ significant\_features) \tag{2}$$

$$Significant_{features} =\ GRU - LMA(selected\_features) \tag{3}$$

2.2 Data Attainment, Acquisition and De-Noising

This GRU-LMA model used the KAUBCMD mammogram digital image dataset, which consists of clinical and contour results from 1416 instances. 5662 images are used, which have both CC and MLO views of the left and right breasts. 6 categories are classified here, where the study uses only 3 target classes and 2 subclasses, which include negative (1), benign (2), malignant (5), and two subclasses: mammography-incomplete (0) and suspicious malignant (4). The images are loaded in DICOM-F format, which is a digital format of mammogram images. Table 1 shows the dataset with expert results with sample data of 3 classes and 2 sub-classes, as mentioned above. In this case, 65% of the data is used for training and 35% for testing and validation. Figure 1 shows the MDI with normal and deeply significant spots.

2.3 Adaptive Median Filtering and Tetrolet Transform for Denoising & Filtering

To filter and denoise the mammogram images AMF and TTM methods are employed, where the AMF analyzes the pixel values to extract all the significant features based on the characteristics of the KAUBCMD dataset. In the initial phase, define the

Table 1. KAUBCMD Clinical Dataset (Samples)

P.No	Age	Type	View	Diagnosis	% of Granular Tissue	Ultrasound	Proven Results
BC007741	51	R	CC	1	0%-25%	BIRAD 3	
BC007741	51	R	MLO	1	0%-25%	BIRAD 2	
BC007741	51	L	CC	1	0%-25%	BIRAD 3	
BC007741	51	L	MLO	1	0%-25%	BIRAD 1	
BC005401	58	R	CC	0	26%-50%	BIRAD 2	**Negative (1)**
BC005401	58	R	MLO	0	26%-50%	BIRAD 3	
BC005401	58	L	CC	4	26%-50%	BIRAD 2	**2982**
BC005401	58	L	MLO	4	26%-50%	BIRAD 2	**Benign (2)**
BC0026041	50	R	CC	1	26%-50%	BIRAD 2	**3069**
BC0026041	50	R	MLO	1	26%-50%	BIRAD 2	**Malignant (5)**
BC0026041	50	L	CC	1	26%-50%	BIRAD 1	**58**
BC0026041	50	L	MLO	1	26%-50%	BIRAD 2	**Suspicious Malignant (4)**
BC0026062	40	R	CC	5	26%-50%	BIRAD 4	**31**
BC0026062	40	R	MLO	2	26%-50%	BIRAD 1	**Incomplete (0)**
BC0026062	40	L	CC	5	26%-50%	BIRAD 2	**17**
BC0026062	40	L	MLO	2	26%-50%	BIRAD 3	
BC0022845	44	L	MLO	2	0%-25%	BIRAD 2	
BC014201	58	R	CC	0	0%-25%	BIRAD 1	
BC014201	58	L	MLO	1	0%-25%	BIRAD 1	

(Out of 1416 instances, only few listed above for sample reading of both CC & MLO of R&L)

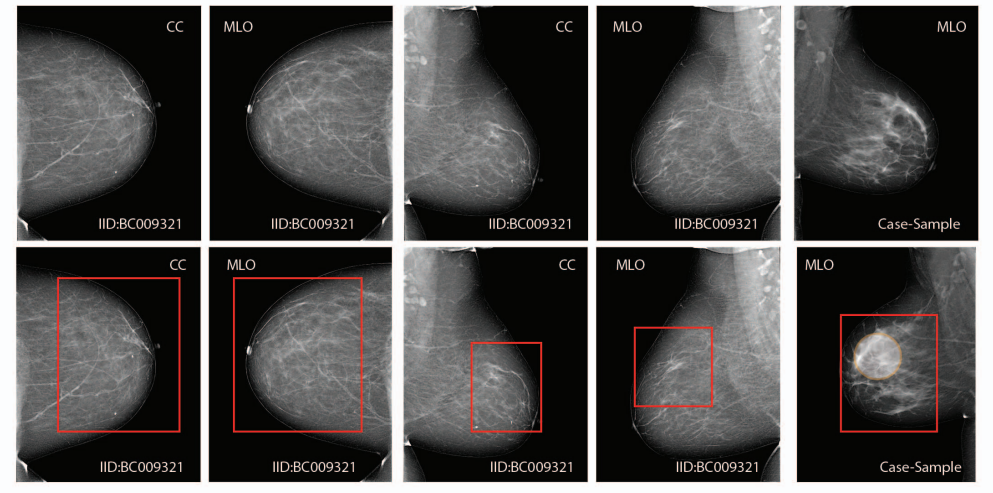

Fig 1. MDI of 2 patients with sample contour cases of negative, benign and malignant

local neighborhood and coordinate the pixel x, y. Median filtering is applied as the window size increases until the noisy pixels are found. Calculate the intensity value of mammogram images and repeat the process until all the noisy pixels are identified. Once noisy pixels are found, remove and reconstruct the mammogram images for classification. Here, window size can be increased any time for noisy removal to find the median value of MDI. AMF helps denoising and enhancement of MDI. After enhancement, the deep, significant spots are identified in contour MDI, where the features are extracted for detection and classification. After processing the 5662 images, the median value is recorded, and the final hidden state is used as a feature representation that encodes contextual features of mammogram images. The Tetrolet compression model is employed for decomposition and denoising by using the below equation,

$$T(x,y) = \in c(i,j) *' \Omega\ (i,j,x,y) + Tetrolet\ Coeff_{position} \tag{4}$$

where, $T(x,y)$ is the Tetrolet coefficient position, i, j represents the basic functions and $\in c(i,j)$ is the position of spots within the masked image. Once the image is denoised and decomposed, a threshold limit is set to retain the significant spots needed to identify the classification. The value of the threshold is between 0 and 7. Then the reconstruction process is carried out, where the denoised image is reconstructed from tetrolet coefficients, which helps to classify the states easily at an early stage. Various iterations are carried out with different threshold limits. Features like micro-calcifications, masses, and tumors, the structure of breast tissue, subtle changes, shape, and texture, hidden lesions, skeweness, GLCM, G-texture, area, solidity, wavelet, fractal, tumor shapes (speculated and lobulated), etc. are extracted & used as input for GRU-LMA for detection and classification of BC.

2.4 Proposed GRU-LMA Architecture Diagram

The GRU-MLA classifier model architecture initially takes input as MDI images from the KAUBCMD dataset and preprocesses for FS and FE. AMF and TTM are applied for denoising and identification of deep spots, and ROI is detected to divide normal and suspicious cancer tissue. Classification is done with the help of training and testing values in KAUBCMD.

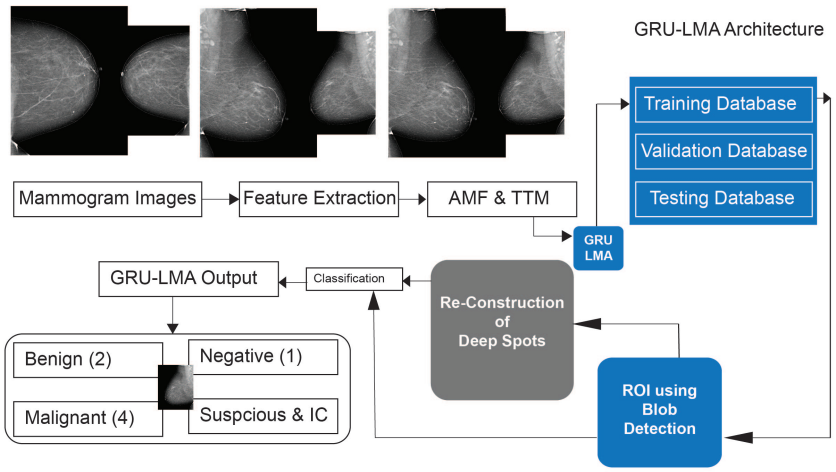

Fig 2. Architecture Diagram of RNN based GRU-LMA classifier

GRU-LMA Classification Algorithm

1. **Input:** MATLAB settings - 5662 Mammogram Images (KAUBCMD Dataset)
2. **Begin**:
3. Load the MDI (Digital Images) from KAUBCMD dataset$data.kaubcmd$
4. Execute pre-processing

for image in KAUBCMD dataset:

$segmentation = denoise(di) + size$

$resized_{image} = resize\ (mammogram\ images)$

$reconstruct_{MDI} = denoise(di) + resized(mi)$

$return\ (pre - processed\ image)$

Apply Adaptive Median Filter $AMF = MDI\ (features, significant\_features)$
Apply TTM $TTM = Coeff_{position} + \ Significant\ deep\ spots\ (SS)$
5. **Train GRU-LMA Classifier** using $New\ Features\ (FS)$
6. **Calculate ROI** using $Selected\ Features\ \ (SF)$
7. Reconstruct Deep Spots with $Segmented\ Image\ Values$
8. Apply Adaptive Histogram Equalization $AHE = MDI\ (Enhancement) + \ Significant\ Spots\ (SS)$
9. Breast Cancer Class Categorization $Morpology\ or\ Texture$
10. Generate **GRU-LMA output (0,1,2,4,5)**
11. **Output:** BC detection and classification
12. **End**

2.5 Adaptive Histogram Equalization for MDI enhancement

The adaptive histogram method is employed to identify the local characteristics of the marked region, which means the abnormal intensities such as breast tissue variations, abnormal sizes, micro-calcifications, etc. The MDI is enhanced optimally based on the pixel values of mammogram images and improved boundary lesions. AHE reduces the variations in image quality and retains the same pixel value, which helps potentially classify the abnormalities. The following steps were used for MDI enhancement during the classification process.

- Identify the unique regions in the MDI images
- Mark the regions for pixel values and enhancements
- Enhance all the deep spots and features extracted by AMF and TTM
- Enhance the contrast level to discover morphology and texture changes
- Record the pixel value and maintain the contour results
- Create the reconstructed segmented images for classification

2.6 ROI using Blob Detection

To reduce false rates of detection and classification, ROI is measured. It refers to the significant regions within the mammogram images and classifies them during the training process. The potential abnormality regions are detected and trained by GRU-LMA. Those regions are isolated and marked for testing purposes for efficient detection and classification. Accurate diagnosis can be done by using marked significant ROI regions. Once significant regions are identified, specific features are extracted from MDI to train the GRU-LMA model to spot the tumor and classify it in a robust manner. GRU-LMA focuses only on the areas of interest for classification. Reliable classification results are acquired. Color detection and blob analysis are also used to identify the different regions of varying pixel intensities within the MDI. All the potential abnormalities were marked to improve efficiency.

2.7 Comparative analysis using MATLAB R2020a

The deep learning-based new RNN classifier GRU-LMA is compared against the prevailing approaches of linear SVM [18], KNN-HP [19], and K-Means GRI [20]. The MATLAB R2020a tool is used to assess the performance of the proposed neural network classifier. MATLAB provides functions to perform cross-validation to assess the GRU-LMA model by utilizing the KAUBCMD dataset. The segmentation is simplified in MATLAB while using MDI patterns, which helps to assess and train the DL models. The GRU-LMA model is tested in different iterations by comparing it with existing models. As the tool supports parallel GPU computing, the training phase is significantly speeding up to work with large datasets. The PEM values of GRU-LMA are tested in the dataset, and the same are used for comparative analysis. The following is the evaluation metric equation used to assess the performance of GRU-LMA during the execution phase.

$$BC_{Accuracy} = \frac{(TPR+TNR)}{(TPR+TNR+FPR+FNR)} \times 100 \tag{5}$$

$$BC_{Sensitivity} = \frac{TPR}{(TPR+FNR)} \times 100 \tag{6}$$

$$BC_{Specificity} = \frac{TNR}{(TNR+FPR)} \times 100 \tag{7}$$

$$BC_{Precision} = \frac{TPR}{(TPR+FPR)} \times 100 \tag{8}$$

$$BC_{Errorrate} = \frac{2xPrecision * Recall}{(Precision+Recall)} \; and \; BC_{Recall} = \frac{TPR}{(TPR+FNR)} \tag{9}$$

$$MCC = \frac{T_1}{\sqrt{T_2 \times T_3 \times T_{4\times} T_5}} \times 100 \tag{10}$$

where, BC denotes breast cancer detection and the evaluation metrics is derived using the equation, $T_1 = (TPR \times TNR - FPR \times FNR)$, $T_2 = (TPR + FPR)$, $T_3 = (TPR + FNR)$, $T_4 = (TNR + FPR)$, and $T_5 = (TNR + FNR)$.

- **Sensitivity and Specificity:** Sensitivity measures the ability of the GRU-LMA model to correctly identify true positive cases of BC. Specificity measures the true negative cases out of the testing data.
- **Accuracy:** It calculates the proportion of the accuracy rate where the GRU-LMA model correctly identifies both true positive and negative cases out of all the cases evaluated.
- **Error Rate:** Used to assess the error rate, i.e., incorrect classification, to avoid an increase in false rates.
- **TPR and TNR :** TPR quantifies the percentage of malignant cases identified by GRU-LMA, and TNR quantifies the normal cases that show negative results.
- **FPR and FNR:** Assess the performance of FPR and FNR by dealing with imbalanced datasets and false classification rates in the GRU-LMA DL classifier model.
- **Precision:** Evaluates the classifier model on how well the accurate predictions and classifications are done in order to minimize the false rate.
- **Training Time:** It minimizes batch processing time by allocating it in a distributed environment and also by utilizing pre-trained models.

3 Results and Discussions

This chapter shows the comparative analysis of the new DL-based classifier to boost the accuracy of breast cancer detection & classification and to minimize the error rate and training time. The model GRU-LMA was compared against existing models such as linear SVM [18], KNN-HP [19], and K-Means GRI [20]. Promising results have been shown in terms of classification, detection, and speeding up the training process, which helps the radiologists make easy diagnoses. The GRU-LMA model overcomes all the shortcomings of the prevailing unsupervised approaches. The identification of MDI deep spots and significant features and AMF MDI enhancements helps boost accuracy during comparative analysis. The DL-EAEN overcomes the drawbacks of existing unsupervised learning methods. Figure 2-8 showcases the graphical representation of GRU-LMA analysis with the X axis and the Y axis.

3.1 Sensitivity and Specificity

Figure 3 showcases the comparative analysis of the proposed novel DL-based RNN classifier GRU-LMA for BC detection and classification. The results are compared against the existing automated models which classifies the cancer spots. As the new model detects subtle changes in the tumor characteristics, the rate of detection is high with a minimal false rate. GRU-LMA handles the class imbalance data to maintain high specificity. As the model bifurcates the classes into three during the initial stage, it achieves enhanced separation between benign, malignant, and negative cases, leading to improved sensitivity and specificity of 95.07% and 97.81%, which is comparatively higher than the baseline models.

Table 2. Sensitivity and Specificity Analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
Sensitivity	70.91	79.65	82.60	95.07
Specificity	70.45	76.55	81.25	97.81

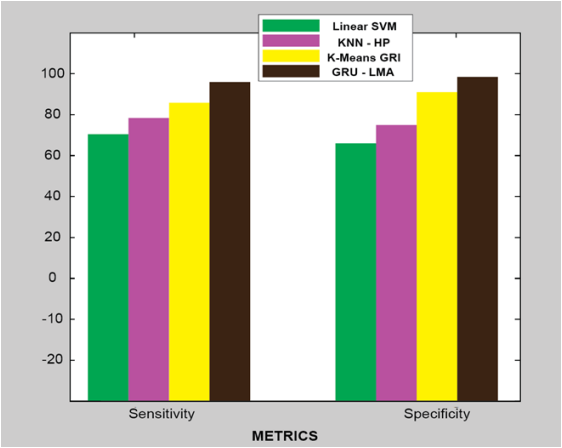

Fig 3. Sensitivity and Specificity

3.2 Accuracy Analysis

The detection and classification accuracy of the proposed GRU-LMA classifier model are measured and portrayed in Figure 4. Due to rigorous preprocessing and feature selection in the KAUBCMD dataset, well-labeled data with minimal noise is being used, leading to high accuracy. Hyper-tuning parameters such as learning and dropout rates and batch sizes are optimized by LMA. A 3x3 confusion matrix is performed to ensure high augmentation. Through ensemble techniques like AMF and TT methods, the model outperforms the prevailing machine learning methods which has drawbacks of accuracy and classification error. 95.19% accuracy is attained in GRU-LMA, which is higher than other models.

Table 3. Accuracy analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
Accuracy (It-1)	74.56	78.90	85.66	94.05
Accuracy (It-N)	79.90	81.60	88.09	95.19

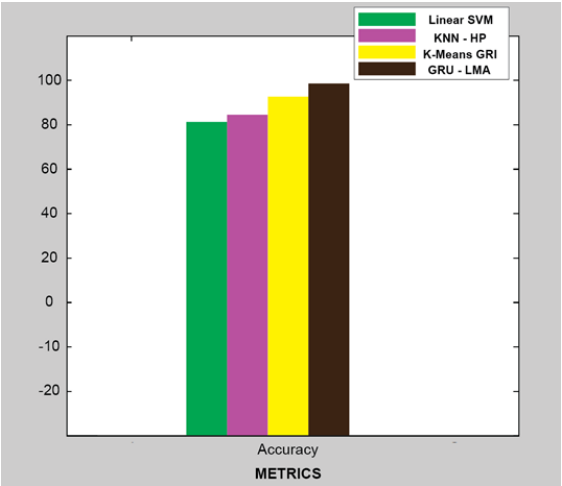

Fig 4. Accuracy

3.3 Error Rate Analysis

Figure 5 presents the error rate analysis of the proposed deep learning breast cancer classifier model. The GRU-LMA is compared against the existing methods to assess the training and error rate. As the local features are extracted in this model, all the significant spots are identified along with tumor characteristics. High accuracy is attained by distinguishing the BC classes, which leads to an impressively low error rate of 3% which is comparatively best than any other DL classifier models. The fine-tuning of image segmentation takes place in GRU-LMA, which correctly identifies the deep spots of tissue and classifies the tumor accurately. Even though the model has been tuned for more than 600 iterations, remarkable performance has been achieved.

Table 4. Error Rate Analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
Error Rate (It-1)	7.81	6.75	6.00	4.6
Error Rate (It-N)	7.01	6.00	5.25	3

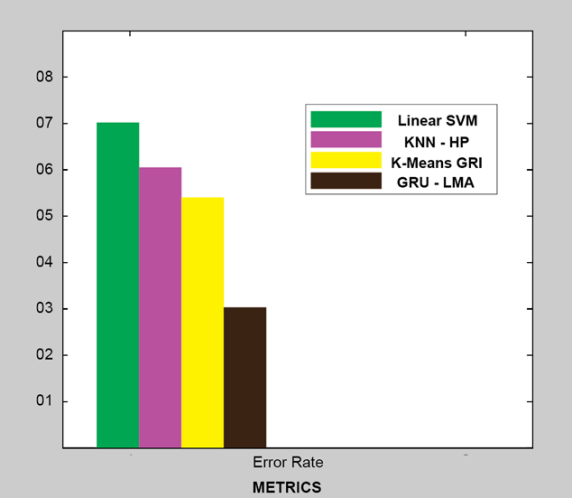

Fig 5. Error Rate

3.4 TPR & TNR Analysis

Figure 6 presents the TPR and TNR analysis of the novel GRU-LMA model under various iterations and threshold settings. The comparative analysis is done to assess the true positive and false positive out of all cases against the prevailing deep learning breast cancer classification models. To balance sensitivity and specificity, a threshold value is set. The KAUBCMD dataset is trained and tested after the filtering and segmentation process for effective spotting of tumor tissues found in breasts in MLO and CC views. Remarkable output is achieved in GRU-LMA with 96.03% TPR and 95.09% TNR, which is higher than existing classifiers.

Table 5. TPR & TNR Analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
TPR	59.87	63.45	80.70	96.03
TNR	46.58	59.58	71.20	95.09

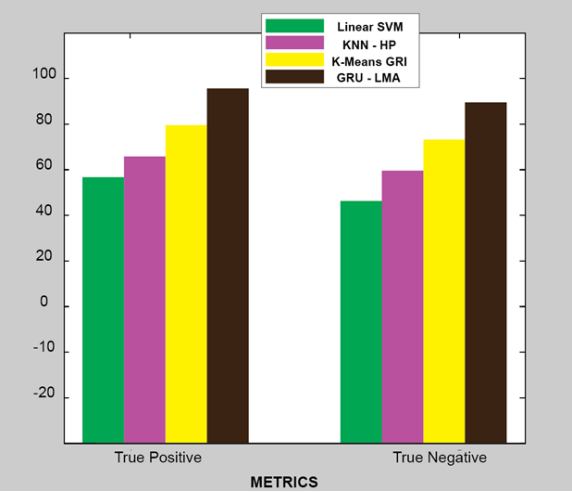

Fig 6. TPR & TNR

3.5 Precision and Recall Analysis

The precision and recall analysis is done for the proposed DL-based RNN classifier GRU-LMA for BC detection and classification is shown in Figure 7. GRU-LMA is compared against linear and unsupervised learning models where it has limitations in identifying cancer spots. Due to vibrant FE and FS with optimization, promising results have been achieved. False alarms are reduced, which shows the ability of the GRU-LMA classifier model, which detects and classifies breast cancer in a dynamic manner. LMA regularizes and prevents over fitting as it does not classify the target values only based on noise data; it classifies based on deep spots that are identified. 96% precision and 97.03% recall are achieved.

Table 6. Precision and Recall Analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
Precision	62.15	80.55	84.56	96.00
Recall	64.30	81.25	86.93	97.03

3.6 FPR & FNR Analysis

The FPR and FNR comparative analysis are done, and the results are showcased in Figure 8. The GRU-LMA classifier model is evaluated against existing ML unsupervised models and results are presented below. As the dataset is already annotated by experts, the images are segmented through a filtering and transformation process where the edges are measured and tumor areas are marked in a rectangle. The malignant cases are identified by deep spotting in GRU-LMA, and classes are defined. 6.14% FPR and 4.20% FNR are achieved, which is relatively good compared to baseline versions.

Table 7. FPR & FNR analysis

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
FPR	31.12	20.35	18.58	6.14
FNR	29.45	19.68	16.80	4.20

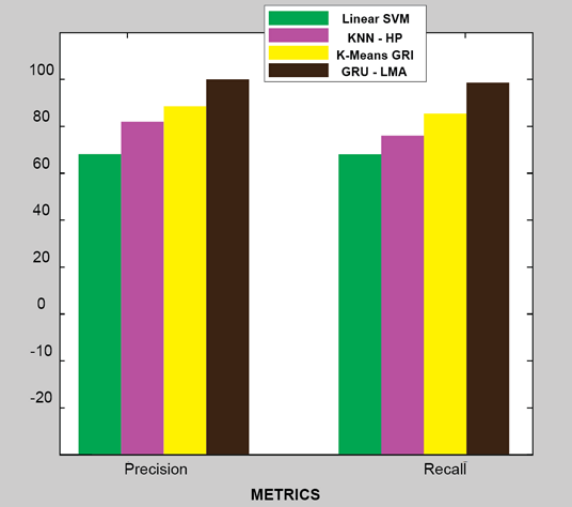

Fig 7. Precision and Recall

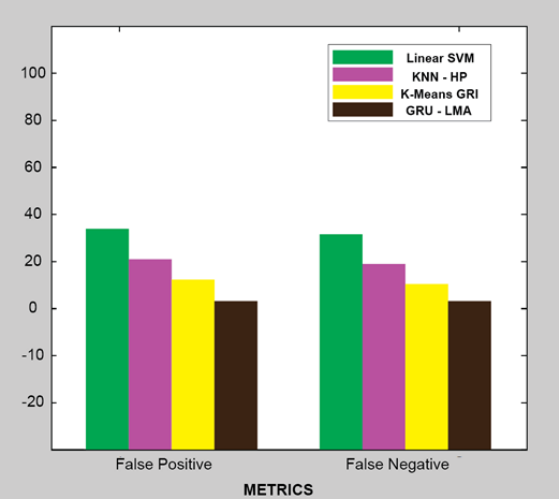

Fig 8. FPR & FNR

3 .7 Training Time / Detection Speed Analysis

Figure 9 displays the comparative analysis of the training time and detection speed of the proposed RNN classifier, GRU-LMA. The new classifier is compared against the existing classifier models to assess the level of detection speed. The training time is remarkable, reduced to 17 minutes per batch, which is lower than the baseline classifier models. The GRU-LMA optimizes the image and adjusts efficiently to minimize the loss function, leading to faster convergence. Also, it strikes the balance between each batch and maintains the threshold limits. Gradient clipping and adaptive learning rates help the GRU-LMA classifier stand out from the rest.

Table 8. Detection Speed Analysis (In Minutes / Per Batch Training)

Metrics / Schemes	L-SVM [18]	KNN-HP [19]	K-Means GRI [20]	GRU-LMA (Proposed)
Training Time (It-1)	40	34	31	20 M
Training Time (It-N)	36	30	27	17 M

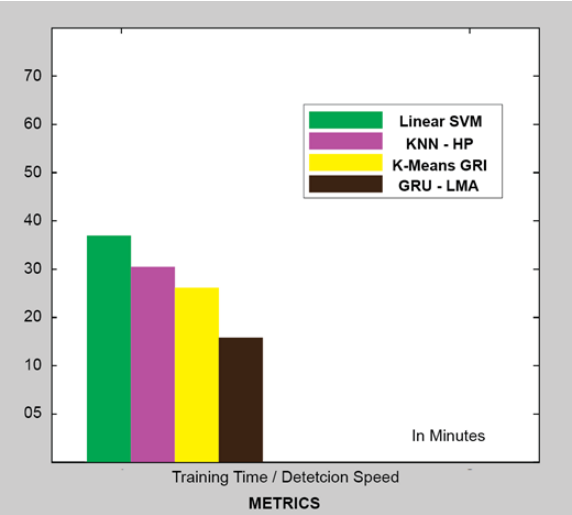

Fig 9. Training Time & Detection Speed

4 Conclusion

The newly proposed deep learning GRU-LMA classifier model can be very well-used to boost accuracy in breast cancer detection and classification and to reduce the error rate while handling large datasets. The KAUBCMD breast cancer MDI dataset is used to extract relevant features. 5662 mammogram images were taken. 65% are used for training, and 35% are used for testing and validation purposes. The GRU-LMA classifier identifies all the deep spots in the MDI by extracting relevant features for BC classification. Adaptive histogram equalization is employed additionally for MDI enhancement. ROI is detected using blob detection to delineate breast tissue from the background, and suspicious tissue regions are separated from normal. The AMF and TTM algorithms are used to extract features into two categories, which are morphology and texture. The new classifier is classified into three target classes, which include negative (1), benign (2), malignant (5), and two subclasses: mammography-incomplete (0) and suspicious malignant (4). The GRU-LMA shows remarkable results in the detection and classification of breast cancer using MDI images. Fractal dimension and pixel intensity are measured to boost the accuracy level. 95.19% accuracy is achieved, which is comparatively higher than any other existing model that trains and validates the KAUBCMD dataset. Training time is reduced to 17 minutes for one batch. The error rate is minimized to 6%. 96.03% TPR and 95.09% TNR, 95.07% sensitivity, 97.81% specificity, 96% precision, 97.03% recall, 3% FPR, and 4% FNR are achieved during comparative analysis against L-SVM, KNN-HP, and K-Means GRI. Some of the noted limitations and weakness of the GRU-LMA model are that it cannot process and handle high regions as this is an RNN classifier and works with only sequential data and not spatial patterns. GRU-LMA model could not handle sequential pattern of data in robust way. Recognizing complex patterns is also one of the challenges in RNN-based classifiers, which requires more training time. In future study, this model can be enhanced by incorporating 2D and 3D image handlers with the use of artificial intelligence techniques for high-end image translations and robustness.

References

1) Iqbal MS, Ahmad W, Alizadehsani R, Hussain S, Rehman R. Breast Cancer Dataset, Classification and Detection Using Deep Learning. *Healthcare*. 2022;10(12):1–19. Available from: https://doi.org/10.3390/healthcare10122395.
2) Din NMU, Dar RA, Rasool M, Assad A. Breast cancer detection using deep learning: Datasets, methods, and challenges ahead. *Computers in Biology and Medicine*. 2022;149:106073. Available from: https://doi.org/10.1016/j.compbiomed.2022.106073.
3) Wang X, Ahmad I, Javeed D, Zaidi SA, Alotaibi FM, Ghoneim ME, et al. Intelligent Hybrid Deep Learning Model for Breast Cancer Detection. *Electronics*. 2022;11(17):1–18. Available from: https://doi.org/10.3390/electronics11172767.

4) Arooj S, Atta-Ur-Rahman, Zubair M, Khan MF, Alissa K, Khan MA, et al. Breast Cancer Detection and Classification Empowered With Transfer Learning. *Frontiers in Public Health*. 2022;10:1–18. Available from: https://doi.org/10.3389/fpubh.2022.924432.
5) Nithyanandh S, Jaiganesh V. Quality of service enabled intelligent water drop algorithm based routing protocol for dynamic link failure detection in wireless sensor network. *Indian Journal of Science and Technology*. 2020;13(16):1641–1647. Available from: https://doi.org/10.17485/IJST/v13i16.19.
6) Mohamed EA, Rashed EA, Gaber T, Karam O. Deep learning model for fully automated breast cancer detection system from thermograms. *PLOS ONE*. 2022;17(1):1–20. Available from: https://doi.org/10.1371/journal.pone.0262349.
7) Hamed G, Marey MAERAE, Amin SESE, Tolba MF. Deep Learning in Breast Cancer Detection and Classification. In: The International Conference on Artificial Intelligence and Computer Vision: Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020);vol. 1153 of Advances in Intelligent Systems and Computing. Springer International Publishing. 2020;p. 322–333. Available from: https://doi.org/10.1007/978-3-030-44289-7_30.
8) Adam R, Dell'aquila K, Hodges L, Maldjian T, Duong TQ. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review. *Breast Cancer Research*. 2023;25(1):1–12. Available from: https://doi.org/10.1186/s13058-023-01687-4.
9) Nithyanandh S, Jaiganesh V. Dynamic Link Failure Detection using Robust Virus Swarm Routing Protocol in Wireless Sensor Network. *International Journal of Recent Technology and Engineering*. 2019;8(2):1574–1579. Available from: https://www.ijrte.org/wp-content/uploads/papers/v8i2/B2271078219.pdf.
10) Sandeep D, Bethel GNB. Accurate Breast Cancer Detection and Classification by Machine Learning Approach. In: 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 11-13 November 2021, Palladam, India;vol. 2021. IEEE. 2021;p. 366–371. Available from: https://doi.org/10.1109/I-SMAC52330.2021.9640710.
11) Arooj S, Atta-Ur-Rahman, Zubair M, Khan MF, Alissa K, Khan MA, et al. Breast Cancer Detection and Classification Empowered With Transfer Learning. *Frontiers in Public Health*. 2022;10:1–18. Available from: https://doi.org/10.3389/fpubh.2022.924432.
12) Humayun M, Khalil MI, Almuayqil SN, Jhanjhi NZ. Framework for Detecting Breast Cancer Risk Presence Using Deep Learning. *Electronics*. 2023;12(2):1–16. Available from: https://doi.org/10.3390/electronics12020403.
13) Nithyanandh S, Jaiganesh V. Reconnaissance Artificial Bee Colony Routing Protocol to Detect Dynamic Link Failure in Wireless Sensor Network. *International Journal of Scientific & Technology Research*. 2020;10(10):3244–3251. Available from: https://www.ijstr.org/paper-references.php?ref=IJSTR-1019-23039.
14) Behar N, Shrivastava M. A Novel Model for Breast Cancer Detection and Classification. *Engineering, Technology and Applied Science Research*. 2022;12(6):9496–9502. Available from: https://etasr.com/index.php/ETASR/article/view/5115.
15) Botlagunta M, Botlagunta MD, Myneni MB, Lakshmi D, Nayyar A, Gullapalli JS, et al. Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms. *Scientific Reports*. 2023;13(1):1–17. Available from: https://doi.org/10.1038/s41598-023-27548-w.
16) Jafari Z, Karami E. Breast Cancer Detection in Mammography Images: A CNN-Based Approach with Feature Selection. *Information*. 2023;14(7):1–14. Available from: https://doi.org/10.3390/info14070410.
17) Arooj S, Atta-Ur-Rahman, Zubair M, Khan MF, Alissa K, Khan MA, et al. Breast Cancer Detection and Classification Empowered With Transfer Learning. *Frontiers in Public Health*. 2022;10:1–18. Available from: https://doi.org/10.3389/fpubh.2022.924432.
18) Egwom OJ, Hassan M, Tanimu JJ, Hamada M, Ogar OM. An LDA–SVM Machine Learning Model for Breast Cancer Classification. *BioMedInformatics*. 2022;2(3):345–358. Available from: https://doi.org/10.3390/biomedinformatics2030022.
19) Chen H, Wang N, Du X, Mei K, Zhou Y, Cai G. Classification Prediction of Breast Cancer Based on Machine Learning. *Computational Intelligence and Neuroscience*. 2023;2023:1–9. Available from: https://doi.org/10.1155/2023/6530719.
20) Niknam A, Ghousi F, Masoumi R, Atashi A, Makui A, A. Hybrid Medical Data Mining Model for Identifying Tumor Severity in Breast Cancer Diagnosis. *Advances in Industrial Engineering*. 2021;55(2):151–164. Available from: https://aie.ut.ac.ir/article_84377_3991e2d869817d530418c7e80cec346d.pdf.
21) Wajeed MA, Tiwari S, Gupta R, Ahmad AJ, Agarwal S, Jamal SS, et al. A Breast Cancer Image Classification Algorithm with 2c Multiclass Support Vector Machine. *Journal of Healthcare Engineering*. 2023;2023:1–12. Available from: https://doi.org/10.1155/2023/3875525.
22) Alsolami AS, Shalash W, Alsaggaf W, Ashoor S, Refaat H, Elmogy M. King Abdulaziz University Breast Cancer Mammogram Dataset (KAU-BCMD). *Data*. 2021;6(11):1–15. Available from: https://doi.org/10.3390/data6110111.
23) Nithyanandh S, Omprakash S, Megala D, Karthikeyan MP. Energy Aware Adaptive Sleep Scheduling and Secured Data Transmission Protocol to enhance QoS in IoT Networks using Improvised Firefly Bio-Inspired Algorithm (EAP-IFBA). *Indian Journal Of Science And Technology*. 2023;16(34):2753–2766. Available from: https://doi.org/10.17485/IJST/v16i34.1706.
24) Nasser M, Yusof UK. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. *Diagnostics*. 2023;13(1):1–26. Available from: https://doi.org/10.3390/diagnostics13010161.