International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Breast Cancer Detection using Deep Learning

Mr. NOOR AHAMED J[1] M.C.A., M.Phil., Mythili K[2]

[1]Assistant Professor(SG) , Department of Computer Applications, Nehru college of management, Coimbatore, Tamilnadu, India.
jnamca@gmail.com

[2] II MCA, Department of Computer Applications, Nehru college of management, Coimbatore,

Tamilnadu, India.

Abstract: *One of the most researched issues in the medical field is the diagnosis of breast cancer. Numerous studies have been conducted on cancer diagnosis, highlighting the necessity of early cancer illness prediction. Health records are used as input to an automated system in order to provide advance prediction. The main goal of the paper is to build an automated system using recurrent neural network models based on deep learning. This research proposes an LSTM-BRNN-GRU that uses a patient's medical records to assess the likelihood that the patient will get breast cancer. The suggested model is contrasted with different baseline classifiers, including the RNN, LSTM, and GRU models. According to the study's comparative findings, the LSTM-BRNN-GRU model performs better in classification for breast cancer predictions.*

.

Keywords: Breast cancer, long short-term memory (LSTM), bidirectional recurrent neural network (BRNN), gated recurrent unit (GRU), predictive model, mammography.

1. INTRODUCTION:

Breast cancer (BC), one of the more prevalent cancer kinds, frequently affects the adult female population. Malignancies of various stages and growth rates are included in the community of patients with breast cancer. Cells lining the milk ducts are the starting point for breast cancer, which gradually develops into a lump or tumor [1]. An average of 50% of instances of breast cancer in India are discovered in stages III and IV. In the case of wealthy nations like the United States, this diagnostic rate rises to 12% [2]. to another study, about 7% of women are diagnosed with breast cancer before the age of 40, and the disease accounts for almost 40% of all cancers in women in this age group [3]. In the realm of medicine, computer-aided diagnostic technologies were used frequently investigated and supported. This research proposes an automated technique that provides early BC prediction to help in the clinical care unit. A number of conversations with the medical professionals led to this study. Numerous combinations of chemotherapy, surgery, radiation therapy, hormone therapy, and targeted therapy using a multimodality approach are possible treatments for breast cancer. Consequently, early discovery of this illness will let physicians recommend likely remedies. Data mining and knowledge discovery techniques that automatically identify patterns and connections among the vast amount of data are being investigated to support the medical industry. In the context of big databases, knowledge discovery (KDD) procedures are used to extract knowledge from data [4]. Data mining techniques can be used to improve diagnosis, prognosis, and treatment, among other fields of medicine. The technology suggested in this research automatically gathers patients' past medical records and determines whether the patient may have breast cancer. Given that cancer is frequently referred to as a silent killer that manifests without any signs, early detection of this illness is necessary.

A type of machine learning called deep learning techniques are frequently beneficial due to their self-adaptive structure, which can handle data with little processing. By assigning the feature engineering stage to computers rather than a human, non-experts can participate in the analysis portion of the process. Since deep learning makes it easier to build networks by combining more than two layers, it is an improvement over traditional artificial neural networks [5]. This research aims to increase the accuracy of breast cancer disease prediction utilizing medical data by implementing a deep learning-based methodology. One kind of deep learning model with a feedback loop structure that is frequently useful for predicting is a Recurrent Neural Network (RNN) [6].

This study proposes a stacked GRU- LSTM model that predicts the diagnosis of this illness based on historical medical records as input. Two RNN variations that are utilized for predicting are the Gated Recurrent Unit (GRU)
[7] and Long-short Term Memory (LSTM) [8]. The suggested approach is assessed and contrasted with alternative baseline models, including the stacked LSTM model, the stacked GRU model, and the basic RNN model. In fact, the results of the suggested model maximize the efficiency of early identification of breast cancer disease. Analyzing the suggested algorithms entails figuring out complexity, comparative, qualitative, and quantitative metrics. A dataset has been used to thoroughly test the suggested approaches. A dataset has been used to thoroughly test the suggested approaches. The study is highly motivated to predict breast cancer because:

1) pre-processing is done to get a balanced dataset.
2) models based on GRU LSTM and BRNN are used for higher accuracy.
3) the method is compared to existing methods to demonstrate that it performs better than the latest models in every way.

To detect breast cancer early, experiments were conducted on a dataset. As stated in the connected works, a number of investigations were conducted to guarantee the construction of a diagnosis system and to enhance the effectiveness of computational techniques for diagnosing BC. This research proposes a fresh system to go along with this diagnosis approach. Layers are put together within the same platform by combining GRU and LSTM to enable early breast cancer prediction.

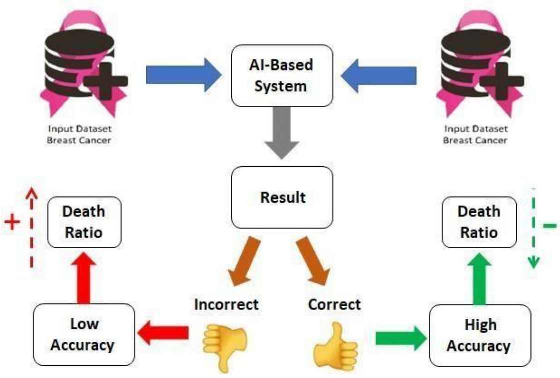

Figure:1 Problem of AI-based breast cancer detection systems.

2. RELATED WORKS:

To distinguish between benign and malignant breast tumors, neural networks and ultrasound pictures of multi-fractal dimensional features were assessed. Classification findings with the highest precision of 82.04% were obtained in this study [9]. A comparative analysis of clustering techniques, including hierarchical clustering, the farthest first, LVQ, canopy, and DBSCAN in the Weka tool for breast tumor identification, was conducted in [10]. It was determined that the farthest first clustering

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

technique had the highest prediction accuracy of 72% based on the results that were presented [10]. A deep classification system for mammography pictures was proposed in a study. Using mammography pictures, an algorithm called the Convolutional Neural Network Improvement for Breast Cancer Classification (CNN-BCC) system was presented. The Mammographic Image Analysis Society (MIAS) offered 183 normal instances, 17 malignant cases, and 21 benign cases for which the algorithm was used. The accuracy of the model was 91.50 percent. For BC classification, CNN and Multiple Instance Learning (MIL) were coupled and presented. The Break his dataset, which included 8,000 microscopic biopsy pictures of benign and malignant breast tumors, was used for the tests. With 40x magnification factor rate, the classification rate was found to be 92.1%. Using the VGG-16, VGG-19, and ResNet50 deep architectures for the BC histological image classification problem, a transfer learning analysis was conducted. With 92.60 percent accuracy, the greatest results were obtained when VGG-16 and logistic regression (LR) were combined. Radial Basis Neural Networks (RBFNs) and Back Propagation Neural Networks (BPPNs) were used to automatically classify photos for the diagnosis of breast cancer. Additionally, the reported BPNN and RBFN accuracies were 60.0% and 71.4%, respectively.

A diagnosis method that uses RepTree, RBF Network, and Simple Logistics was proposed for the detection of breast cancer. A 10-fold cross-validation approach was used to assess the suggested system's performance throughout the testing phase. The efficiency of the suggested system's correct categorization rate reached 75.5% [15].To detect breast cancer early, experiments were conducted on a dataset. Delen et al. looked at creating prediction models for breast cancer survival using artificial neural networks, decision trees, and logistic regression. In order to measure the unbiased estimate of the three prediction models for performance comparison, 10-fold cross-validation techniques were described. The decision tree proved to be the most accurate predictor, according to the results, with 95.6% accuracy. In order to determine classification accuracy in a dataset related to breast cancer, another study examined the application of three algorithms: Decision Tree (C4.5), Artificial Neural Networks (ANN), and Support Vector Machine (SVM). According to the comparative analysis, SVM provided greater classification accuracy.

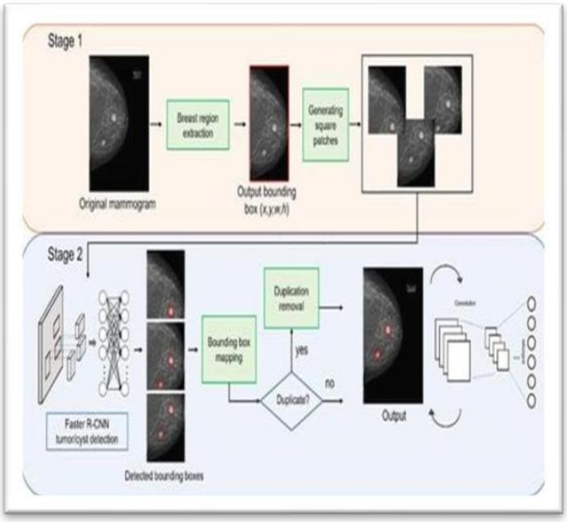

Figure:2 Related work based on breast cancer.

As stated in the connected works, a number of investigations were conducted to guarantee the construction of a diagnosis system and to enhance the effectiveness of computational techniques for diagnosing BC. This research proposes a fresh system to go along with this diagnosis approach. Layers are put together within the same platform by combining GRU and LSTM to enable early breast cancer prediction.

3. BACKGROUND:

By putting together more than two layers, deep learning offers a multi-layered hierarchical data representation, usually in the form of a neural network. Multiple layers with linear or non- linear activation functions are combined to create neural network models, which are then trained collectively to achieve sophisticated methods for problem resolution. Activation functions can perform a variety of calculations and generate results that fall into a specific range. Put otherwise, a step that converts an input signal into an output signal is called an activation function. ReLu and Sigmoid are two well-liked activation functions. One kind of neural network architecture that can handle both sequential and parallel data processing is the recurrent neural network (RNN). By adding memory cells to the neural network, similar functions to those of the human brain can be replicated. Bidirectional RNNs (BRNNs) are another type of RNN that may access input sequences with predetermined beginnings and endings. Since RNN can only use data from the prior context, Bi-RNN can be used to further enhance the system. Two sources of information can be handled by the Bi-RNN. One RNN analyzes the sequence from beginning to end, while the other works backwards from end to beginning, taking into account the past and future context of each sequence element in the justification. Depending on the gating units, RNN options include LSTM-RNN and GRU-RNN. One type of RNN that uses context-based prediction something that regular RNNs do not take into account is the Long Short-term Memory (LSTM) neural network. To put it another way, by training RNN, LSTM can solve the vanishing gradients issue. Long-range dependencies can be better preserved, and a gradient flow can be controlled with the help of LSTM.

Each LSTM cell is made up of gates that control when the input should be remembered, when the value should be remembered or forgotten, and when the value should be output. There are several types of gates, including input, output, and forgotten gates, depending on performance. When the input gate generates a value that is nearly equal to zero, it prevents a value from moving on to the next layer. The value of the net input is simply removed by this input gate. Until the forget gate generates a value larger than zero, the forgotten gate retains the value. When a value that is near zero is generated, the block essentially forgets the value it has been remembering. When the unit should output, the value stored in its memory is determined by the output gate. The Gated Recurrent Unit (GRU) and LSTM are quite similar in that the GRU's gating units regulate the information flow inside the unit without taking into account individual memory cells.

Similar to LSTM, GRU is devoid of memory cells and has fewer gates that are triggered by both the previous output and the current input. When calculating the new candidate activation, GRU regulates the flow of information from the prior activation, but it has no independent control over how much candidate activation is added. Due to parameter reduction, GRU has a higher rate of convergence than LSTM, and in certain situations, it performs better than LSTM models. To reduce an over-fitting issue, it is frequently helpful to design deep models with dropout layers. During each training iteration, dropout layers randomly disable a portion of the network's units or links [5]. The training procedure is carried out once neural network models have been configured. The dataset is divided into smaller portions throughout a single cycle of the training process called an epoch. In order to complete epoch execution, an iterative process is carried out using a few batch sizes that take into account subsections of a training dataset. The binary cross-entropy function serves as a training criterion because the entire framework is geared at resolving a binary classification problem. For each class, binary cross-entropy calculates the difference between the true value (which is either 0 or 1) and the forecast.

The final loss is then calculated by averaging these class-wise mistakes. An optimizer must be used when stacking RNN- based layers into a single framework. Adam is one of the well-known optimizers that is simple to use, computationally effective, and requires less memory. Based on adaptive estimations of lower-order moments, this approach can be used for first-order gradient-based optimization of stochastic objective functions. Because it may be used for non-stationary targets a issues

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

with extremely noisy and/or sparse gradients, this approach is widely used.

4. DATASET USED:

The Breast Cancer Wisconsin (Diagnostic) Dataset is gathered from UCI in this regard [26]. Each of the 569 sample records in the dataset can be viewed as a collection of variables that include many criteria for identifying people exhibiting signs of breast cancer. The dataset includes the attribute "diagnosis," which is used as a prediction's output class. It includes both benign and malignant classes. Figure 3 displays the MIAS dataset.

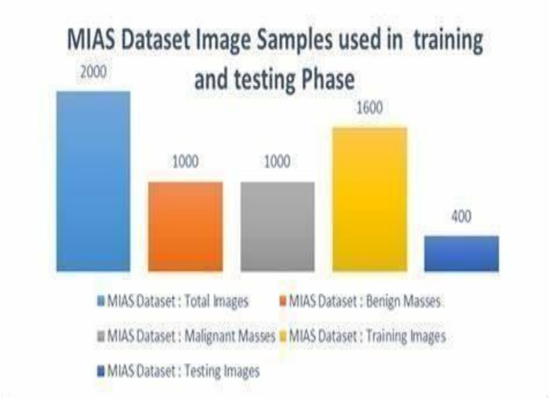

Figure:3 MIAS Dataset.

Figure 4 displays the precise distribution of patients who fall into the "benign" or "malignant" categories, while Figure 5 illustrates the general comprehension of the dataset.

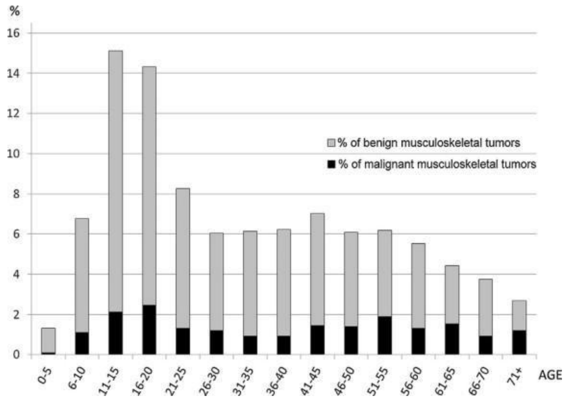

Figure:4 & 5 Distribution of benign and malignant.

5. THE PROPOSED METHODOLOGY:

In order to forecast breast cancer, data mining techniques are used in this work. Data pre-processing and classification/clustering are two key procedures that are crucial to the data mining process. The implementation of data pre-processing, which removes unnecessary or redundant information from the original, comes after any classification or clustering phase. Steps of classification or clustering are carried out to obtain the task of estimation, prediction, etc. The use of data mining techniques for breast cancer prediction is explained in the section that follows. Techniques for data pre- processing are explained in depth. Then, in order to find noticeably higher predicted performance, the combinations of several RNN-based architectures are explained and contrasted.

A. Dataset Pre-Processing:

To create a balanced dataset, pre-processing methods are used after the data is gathered. Pre- processing methods include scaling certain properties and examining and managing missing data. All attributes are examined for existing "nan" values. The mean values for the associated attribute are used in place of these values. Certain properties, such as "id" and "Unnamed:32," are removed because they don't help with prediction. In order to improve efficiency while fitting to a classifier, feature scaling of pertinent attributes is then carried out. In order to improve efficiency while fitting to a classifier, feature scaling of pertinent attributes is then carried out. Ten real-valued attributes are included in the dataset: compactness (perimeter2 / area – 1.0), smoothness (local variation in radius lengths), symmetry, fractal dimension ("coastline approximation"), radius (mean of distances from the center to points on the perimeter), texture (standard deviation of grey-scale values), perimeter, area, and concavity (severity of concave portions of the contour). These attribute values are scaled down into a predetermined range of 0 to 1 in order to conduct feature scaling operations. A modified dataset that can be fitted to the classifier will be produced by using these pre-processing approaches. The converted dataset is divided into two parts: a training set and a testing set. This is accomplished by dividing the transformed dataset into a 7:3 ratio. RNN models are fitted to training data, and predictions for the test dataset are obtained. sequence of GRU and LSTM layers.

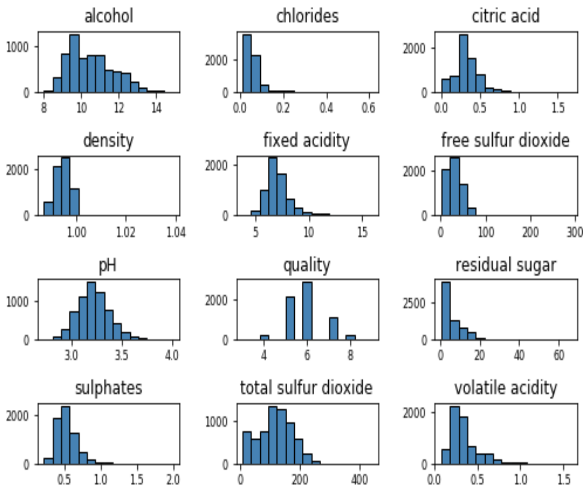

Figure:6 Visualizing attributes as one dimensional data.

B. Methodology and Implementation: The primary goal of the suggested classifier is to use deep learning techniques to determine whether a patient has breast cancer. Using an end-to-end training process and a supervised learning paradigm, deep learning approaches help automatically identify characteristics from raw data. The supervised learning technique used in the classification process seeks to identify patients with benign or malignant cancer. For such predictions, the suggested approach uses a framework based on LSTM- BRNN and GRU. In addition to four thick layers, the current paper suggests a stacked GRU-LSTM based model with an alternate sequence of GRU and LSTM layers.

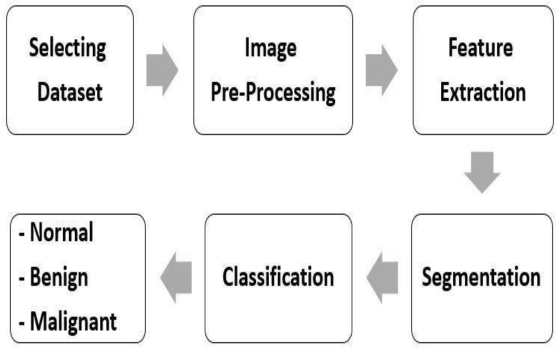

Figure:7 Flowchart of the methodology proposed approaches.

C. Selecting the Dataset:

The mammographic Image Analysis Society (MIAS) dataset from the University of Essex was used.

Description: Contains mammography pictures and labels for abnormalities such as calcifications, restricted masses, and spiculated lesions.

Use: Applied to image analysis for the identification of breast cancer via mammography.

Size: The paper employed 322 images that showed different abnormality groups.

Link: Upon request, MIAS can supply it.

D. Image Pre-processing:

The goal of this phase is to improve the quality of the mammography images. This leads to improved categorization accuracy. The reasoning for this is that the more distinct the borders and features, the more accurate the diagnosis can be, which in turn improves the ability to locate the tumor.

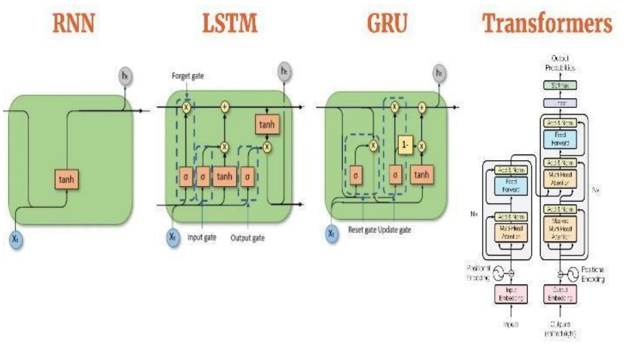

Figure:8 Details of the pre-processing step for the BRNN LSTM and GRU approaches.

E. Feature Extraction:

A list of the general advantages of feature extracting approaches is as follows:

1. Improvements in data accuracy were made.
2. The overfitting risk was reduced.
3. Instructions were slowed down.
4. Improved data representation was examined.
5. The explainability of our model has improved.

In CNN architecture, feature extraction is achieved by the creation of conventional layers. Specifically, the typical layers are created by scanning the mammography input image with a certain kernel (filter).

F. Segmentation:

In the context of breast cancer, segmentation is the process of identifying and separating the region of interest, such as a tumor, from medical imaging data (such as mammograms, MRIs, or ultrasound images). Segmentation is a method for accurate breast cancer identification, classification, and treatment planning by precisely identifying the tumor's boundaries, size, and location.

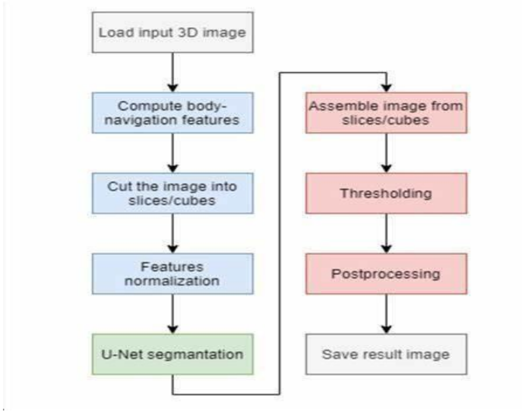

Figure:9 Output of the segmentation step.

G. Classification:

Classification is used to predicate procedure is finished. The output of the classifier (LSTM/GRU) is intended to be shown at this stage. The output of the classifier belongs to one of the three classes: normal, benign, or malignant. In terms of numbers, a mathematical function must be used in the prediction process. In this case, the soft max function is used for multi-classification. Neural networks are generally used to translate the non-normalized output to a probability distribution over expected output classes.

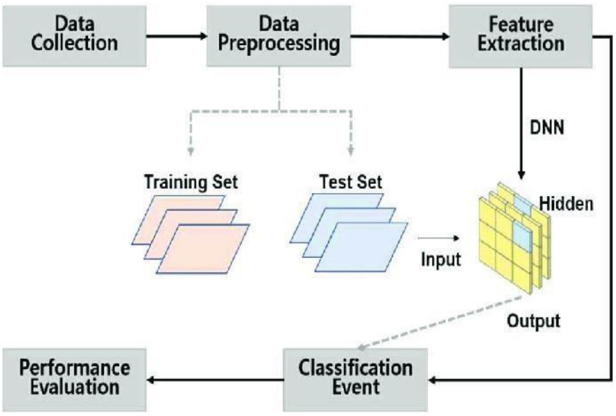

Figure:10 output classes of the prediction process.

6. RESULT:

Accuracy, F1-Score, MSE, and Cohen-Kappa Score are assessed after the suggested GRU- LSTM BRNN model is put into practice. Testing-related losses are also quantified. Later, this model is contrasted with various baseline classifiers, such as the stacked LSTM model, stacked GRU model, and simple RNN-based models. When the 100th period, test accuracy is measured when the training process is over. In other words, following the completion of 100 training epochs, the performance of each constructed model is shown using evaluation measures. Predictions are made for the test dataset following this training session with training data. The evaluation of deep models in relation to the used metrics is instantiated by comparing these predictions with actual observed data. In comparison to existing classifiers, the suggested model shows far more promising results.

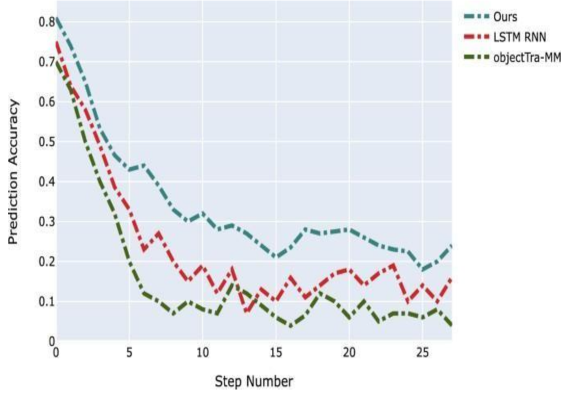

Figure:10 Result predicted.

7. CONCLUSION:

Breast Cancer is a serious condition that requires cautious treatment. Early detection of this illness can significantly improve patient outcomes. This study's goals are to ascertain the likelihood of developing breast cancer and to assess the viability of using prior medical records. This study proposes and implements a stacked GRU-LSTM layer based model using deep learning techniques. When constructing the model with the required parameter adjustment, interfering attributes that affect this disease were taken into account. The stacked GRU-LSTM model has achieved the highest accuracy, F1-score, Cohen-kappa score, and the lowest test loss and MSE values, indicating that it is superior to other baseline models. At 98.34% accuracy, 0.96 F1-score,
0.95 Cohen-kappa score, and 0.04 MSE, the suggested approach shows encouraging results.

7. REFERENCES:

[1] Wu, J., Hicks, C. (2021). Types of breast cancer can be classified using machine learning.

[2] Maabreh, R. S. A., AlGhamdi, A. S., and Alazzam, M. B. (2021). Machine learning methods for forecasting the survival curves of patients with breast cancer. Applied Bionics and Biomechanics, 2021.

[3] Arowolo, M. O., Adebiyi, M. O., Mshelia, M. D., and Olugbara, O. O. (2022). A Linear Discriminant Analysis and Classification Model for the Diagnosis of Breast Cancer. 11455; Applied Sciences, 12(22).

[4] M. Hasan, M. Islam, M. Haque, M. Iqbal, and M. N. Kabir (2020). A comparative analysis that uses machine learning techniques to predict breast cancer. SN Computer Science, 1(5), 1–14.

[5] Hu, Zilong, and colleagues' paper, "Deep learning for image-based cancer detection and diagnosis− a survey," Pattern Recognition 83 (2018), 134–149.

[6] Van Essen, B.C., Awwal, A.A.S., Asari, V.K., Westberg, S., Sidike, P., Alom, M.Z., Taha, T.M., Yakopcic, C., Nasrin, M.S., Hasan, M. An Extensive Review of Deep Learning Advanced Theory and Architectures. (2019) Electronics 8:292.

[7] Arya, N., and Saha, S. (2021), Sophisticated deep learning models that forecast breast cancer survival by utilizing a variety of modalities. Knowledge-Based Systems, 221, 106965.

[8] Sun (2021), Yuan (2021), Wu (2021), and Hu (2021) [6]. Breast cancer histological picture classification using deep residual learning. International Journal of Imaging Systems and Technology, 31(3), 1583-1594.

[9] The study was carried out by Bharati S., Podder P., and Mondal M.R.H. 2020;12:125– 137; Int. J. Comput. Inf. Syst. Ind. Manage. Appl. [Google Scholar]

[10] "Empirical evaluation of gated recurrent neural networks on sequence modeling," by J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, 2014 [Online].

[11] "Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images," by M. A. Mohammed, B. Al-Khateeb, A. N. Rashid, D. A. Ibrahim, M. K. Abd Ghani, and S. A. Mostafa, appeared in Computers & Electrical Engineering, vol. 70, pp. 871–882, August 2018.

[12] "Large-batch training for LSTM and beyond," by Y. You, J. Hseu, C. Ying, J. Demmel, K. Keutzer, and C.-J. Hsieh, in the International Conference on High Performance Computing, Networking, Storage, and Analysis, 2019.

[13] "Improving convolutional neural networks for breast cancer classification," Expert Systems with Applications, vol. 120, pp. 103– 115, April 2019. F. F. Ting, Y. J. Tan, and K. S. Sim.