Breast Cancer Detection and Classification: A Comparative Analysis Using Machine Learning Algorithms

Chapter · March 2022

DOI: 10.1007/978-981-16-8862-1_46

6 authors, including:

Shadman Sakib
Leading University
49 PUBLICATIONS **165** CITATIONS
SEE PROFILE

Nowrin Yasmin
Ahsanullah University of Science & Tech
6 PUBLICATIONS **33** CITATIONS
SEE PROFILE

Abyaz Kader Tanzeem
BRAC University
7 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Khan Md. Hasib
Ahsanullah University of Science & Tech
32 PUBLICATIONS **177** CITATIONS
SEE PROFILE

Breast Cancer Detection and Classification: A Comparative Analysis Using Machine Learning Algorithms

*Shadman Sakib[1], Nowrin Yasmin[2], Abyaz Kader Tanzeem[3], Fatema Shorna[4], Khan Md. Hasib[5] and Saadia Binte Alam[6]

[1]Department of CSE, Leading University, Sylhet 3110, Bangladesh

[2]Samsung R&D Institute, Dhaka 1205, Bangladesh

[3]Department of EEE, BRAC University, Dhaka 1212, Bangladesh

[4]Department of Pharmacy, BRAC University, Dhaka 1212, Bangladesh

[5]Department of CSE, Ahsanullah University of Science and Technology, Dhaka 1208, Bangladesh

[6]Department of EEE, International University of Business Agriculture and Technology, Dhaka 1230, Bangladesh

* sakibshadman15@gmail.com

Abstract. The risk of fatality from breast cancer is increasing exponentially as the population rises. It is the most typical kind of cancer and the major cause of death in women throughout the world. Early detection and treatment of breast cancer could significantly improve the prognosis and survival rate. It is an important phase in rehabilitation and medication since it can help patients receive prompt medical services. An automated disease detection technique that employs machine learning (ML) and deep learning (DL) techniques assist medical professionals in the diagnosis of diseases and provide a reliable, efficient, and faster response thereby minimizing the chance of death. This research aims to perform a comparison among ML and DL methods for breast cancer detection and diagnosis. The five most popular supervised ML techniques named Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbor (KNN), and a DL technique were used for classification using cross-validation technique. The Breast Cancer Wisconsin (Diagnostic) data set has been used as a training set to evaluate and compare each algorithm's effectiveness and efficiency through classification accuracy, recall, specificity, precision, false-negative rate (FNR), false-positive rate (FPR), F1-score, and Matthews Correlation Coefficient (MCC). Experimental results show that Random Forest (tuned) outperformed all the other models with accuracy and F1-score of 96.66% and 0.963 respectively.

Keywords: Breast cancer classification, Machine learning, Early diagnosis, Disease prediction, Medical data mining.

1 Introduction

Among women, breast cancer is known to be the second-most significant reason of death worldwide [1]. According to data of 2019, around 268,600 invasive breast cancer cases and 62,930 non-invasive breast cancer cases were said to be found among women in the United States [2]. Early Stage testing and detection is found to be the most appropriate approach to enhance the probability of quick treatment and reduce the mortality rate among women. Mammogram is widely used as an inexpensive technique to detect any lesions, tumors or cancerous cells in breast region [3]. Many underlying factors can bring about a negative effect on the accuracy of the results. A few and most common errors can be due to distraction or fatigue of the radiologist himself/herself, the breast structure itself can be complex so the radiologist might find it difficult to locate the cancerous cells and also the early phase characteristics of the disease can be hard to analyze too [4] [5].

Computer-aided diagnosis also known as (CAD) comes into play for diagnosing breast cancer in an early stage. The whole process comprises of three steps all together: firstly, finding and locating the area of possible tumor or cancerous cells by the aid of preprocessed mammogram, secondly getting a hold of the characteristics of the tumor including its shape, size, density, possible weight, texture and finally stating whether it is a benign or malignant version of a tumor following the noticeable features obtained from the second step [6]. X-ray was the only known screening method used earlier to detect any sort of malignancy including cancer of breast [7]. And since then various new and advanced detection technologies have been introduced which have proven to be more efficient compared to x-ray or mammography detection methods. These technologies are named as neural networks, data mining and artificial intelligence.

In cancer research, machine learning (ML) techniques can be successfully applied to come across various patterns within a set of data and as a result we can also identify and state about the malignancy or benignancy of cancerous cells. One of the ways to expose the ML into supervised learning is by classification of the data sets which will be automatically used for future predictions due to its set algorithms. As such techniques are putting the clinics, hospitals and other medical fields into huge advantageous situation by saving time and also aiding for early detection by the help of classification models created by previous cases of cancer [8]. In this study, five popular ML methods which included Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbor (KNN), and a DL technique applied to a Breast Cancer Wisconsin (Diagnostic) data set are investigated and compared to analyze the classification performance of different classifying models.

The rest of the paper is structured as follows: In Section 2, previous work has been discussed. Section 3 comprises a detailed methodology describing the dataset used, the proposed methodology, and the algorithms used. Section 4 represents the experimental results and the discussion including evaluation metrics, performance

evaluation, and the comparison of performance. Finally, Section 5 concludes the paper with future work.

2 Literature Review

Since decades, more and more researchers are putting their efforts to carry out studies through which they are trying to save time and bring the best output in cancer research. They are doing this by incorporating data mining algorithms in various available patient datasets. In addition to these, they have also been successful to classify stages and types of cancers including breast cancer. With the help of existing datasets, doctors and researchers are being able to solve challenging cases of cancer in patients within a short time frame as the algorithm plays its role and give results according to the symptoms and conditions of the patient which are further to be applied in patients for mitigation of the cancer. However, the results obtained by applying data mining in various categories for detection of breast cancer were found to very favorable. Table 1 portrays the different algorithms used till date and their efficiency in each case.

With reference to the comparative analysis constituted by Table 1, we can observe that Khuriwal *et al*. [9] obtained the best performing CNN model on the 'Mammogram MIAS' database; achieved an accuracy of 98%. In addition, Fathy *et al*. [10] also obtained a similar accuracy on the 'Screening Mammography' dataset with an accuracy of 96%. In addition, Saabith *et al*. [11] implemented an ANN model called MLP which had an accuracy of 97.28 on the 'Breast Cancer' dataset. Among all the different ML models, Amrane *et al*. [12]'s KNN model had demonstrated the highest performance accuracy with a figure of 97.51%; trained and tested on the 'Wisconsin Breast Cancer' dataset. Moreover, for the Naïve Bayes classifier, Amrane *et al*. [12]'s model achieved the highest performance accuracy with a value of 96.19%, which is quiet close to the accuracy obtained by Amrane *et al*. [12]'s KNN model.

Table 1. Existing works using diverse algorithms and data set on breast cancer detection

Author	Dataset	Efficiency	Algorithm	Findings
Nawaz et al. [13]	BreakHis Dataset	95.4%	Deep learning (DL) Convolutional neural network (CNN)	In comparison with DenseCNN model and state-of-art models, it showed more efficiency.
Khuriwal et al. [9]	Mammogram MIAS database.	98%	CNN	Around 98% efficiency was found with CNN

Fathy et al. [10]	Screening Mammography dataset.	96%	DL	This method was found to have a sensitivity of 99.8% sensitivity.
Amrane et al. [12]	Wisconsin Breast Cancer	97.51%,	KNN (k-Nearest Neighbors)	More accuracy was found with KNN compared to NB
		96.19 %,	NB (Naive Bayes classifier)	
		79.97%	J48	
		75.35%	MLP	
Saabith et al. [11]	Wisconsin Breast Cancer	97.2818%	J48	In this case, exactly same accuracy was achieved by both algorithms.
		97.28%	MLP	
		95.61%	SVM	
		95.75%	Bi-clustering and Ada boost techniques	
		91.3%	RCNN	
		82.50%	Bidirectional Recurrent Neural Networks (HA-BiRNN)	

All classifiers considered, it was observed that CNN-based classifiers demonstrated better performance in terms of efficiency and classification accuracy, as opposed to other families of classifiers. ANN and ML models also had, more or less, similar performances.

3 Methodology

3.1 Dataset

The Breast Cancer Wisconsin (Diagnostic) Dataset [14] has been used in this study. In this dataset, there are 569 instances of computed features and classification of a digitized image of Fine Needle Aspirate (FNA) of a breast mass. There are 32 features of the dataset – two of them being the ID number and the diagnosis. The instances are diagnosed as either malignant or benign. There are 63% of instances belonging to the "benign" class and the other 37% belong to the "malignant" class. The other 30 features denote the mean, standard error, and the

worst (mean of the three largest) values of 10 real-valued features that describe the characteristics of the cell nuclei present in the image.

In this study, 20% of the total dataset was used as test data, while the rest was used as training data. To prepare the data, the 'diagnosis' column was converted into int64 data type. Furthermore, the correlation between the features was analyzed and upon finding strong correlation, all of them were decided to be used.

3.2 Proposed Methodology

The dataset has been used to analyze the classification performance of six models, of which five are ML models and the other is a DL model. In essence, we obtain a comparison between several popular machine learning models as well as between machine learning and deep learning techniques. The models used as classifiers are – Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbor (KNN), and Neural Network (NN).

For each of the machine learning models, 3 instances were analyzed – using default parameters, using tuned parameters, and using 10-fold cross-validation. The parameter tuning is done using GridSearchCV function of the scikit-learn library where an exhaustive search is performed over the provided parameter values. It performs the default 5-fold cross-validation using the dictionary of parameters provided for each model. Additionally, the DL model is analyzed using default parameters and using 10-fold cross-validation. Fig. 1 portrays the overall process for evaluating the model in order to predict and classify breast cancer.

For the ML models, first their performance is analyzed using default parameters. Secondly, tuned parameters are used and their performances are analyzed again. The parameters specified for tuning and the best set of parameters found for each model are listed in Table 2.

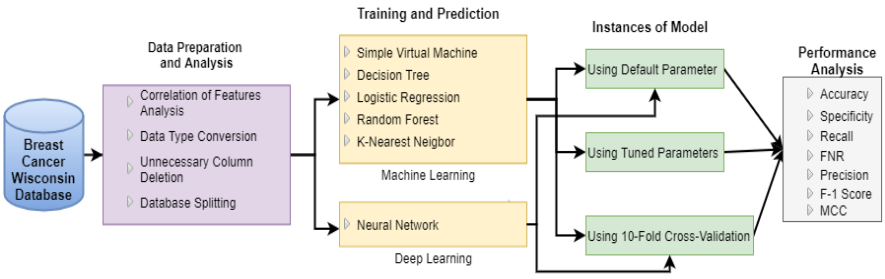

Fig. 1. Flowchart of proposed methodology

Whereas, for the DL model, two instances were analyzed – the first using a simple neural network (NN) and the second using 10-fold cross validation.

Table 2. Tuned and best set of parameters

Machine learning models	Tuned and best set of parameters
Support Vector Machine (SVM)	Gamma: 'scale' (*1 / (n_features * X.var())*), 'auto'(*1 / n_features*) Kernel: linear, radial basis function (RBF) Best parameters: gamma: 'scale', kernel: linear
Decision Tree (DT)	Criterion: Gini, entropy Max depth: Between 1-10, excluding 10 Best parameters: criterion: entropy, max depth: 4
Logistic Regression (LR)	Solver: 'newton-cg' (Newton method), 'liblinear' (Library for Large Linear Classification), 'sag' (Stochastic Average Gradient descent), 'saga' (extension of 'sag') Max iteration: 10,000 Best parameters: solver: 'liblinear', max iteration: 10,000
Random Forest (RF)	Number of estimators: 100, 200, 500 Max features: 'auto', 'sqrt', 'log2' Criterion: Gini, entropy Best parameters: criterion: entropy, max features: 'sqrt', number of estimators: 200
K-Nearest Neighbor (KNN)	Number of neighbors: 5, 7, 9, 11, 13, 17, 19, 21 Algorithm: 'auto', 'ball_tree', 'kd_tree', 'brute' Weights: uniform, distance Best parameters: algorithm: 'auto', number of neighbors: 13, weights: uniform

In the first instance, there are 3 layers to the network. The first layer is the input layer having input size of 29 (since we have 30 features), an output size of 8, and ReLU as the activation function. The second layer also has an output size of 8 and ReLU activation function. And, the last layer is the output layer with output size of 1 and sigmoid activation function. The model is trained using an Adam optimizer, with a learning rate of 0.001, decay of 0.0001, and a clip value of 0.5. The model is fitted in batch sizes of 10 for 100 epochs. Prediction values above 0.5 are considered finally.

In the second instance, a NN of 5 layers is used. The first 4 layers are a repetition of the following sequence – a layer with output size of 8 and ReLU activation function and a dropout layer with dropout value of 0.2. The last layer is the output layer with output size of 1 and sigmoid activation function. It is trained and fitted using the same metrics as the first instance. Finally, the cross-validation scores are obtained using 10-fold cross-validation. Fig. 2 illustrates the layers of NN used in both case.

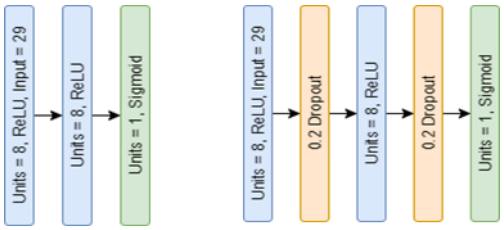

Fig. 2. Layers of neural network used: in the first instance (left) in the second instance (right)

3.3 Classifier Description

Support Vector Machine (SVM)
Support Vector Machine (SVM) [15] is one of the most popular machine learning algorithms that provides sufficient accuracy using less computational power. This algorithm finds a hyper plane to clearly classify data points in an N-dimensional space with maximum margin i.e., space between data points of the different classes [16]. The hyper plane acts as the decision boundary of the classification. The hyper-plane derived from SVM is depicted in Fig. 3 (left).

Decision Tree
Decision Tree (DT) [17] is another popular supervised learning algorithm that can be applied and interpreted easily. It is essentially a tree diagram that contains internal nodes, leaf nodes, and branches. The internal nodes represent the independent variables of a classification or regression problem. The leaf nodes represent the outcome or the dependent variable of the problem [18]. And the branches represent the decision rules to move down the nodes. The decision making starts at the root of the tree moving downwards. Each internal node has a decision rule that is followed based on the comparison between the dataset's feature and the feature on the tree until a decision is reached i.e., until a leaf node is reached. The general representation of a DT is shown in Fig. 3 (right).

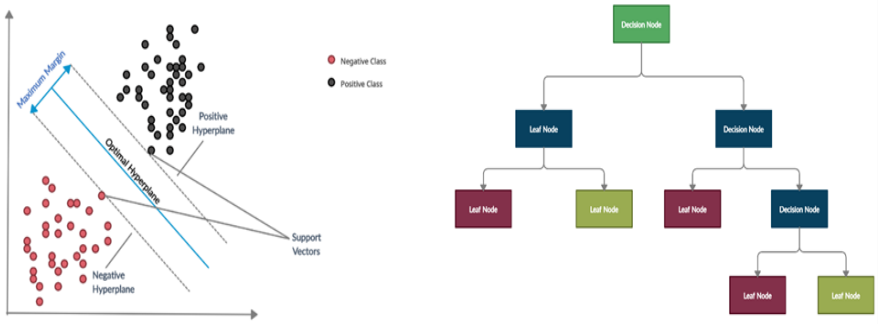

Fig. 3. General representation: hyper plane derived using SVM (left) decision tree (right)

Logistic Regression

Regression is a statistical process to predict the relation between a dependent variable and one or more independent variable. It is one of the categories of supervised learning. Logistic regression (LR) [19] is a class of regression thatpredicts categorical variables using one or more independent variables. It is also known as logit regression or logit model. It provides the probability of the "positive/true" class being decided. Based on that probability and a set threshold, the class can be categorized as 1 or 0. An illustration of logistic regression is shown in Fig. 4 (left). LR can be represented by the following equation,

$$\log\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta_1 \mathrm{X}_1 + \beta_2 \mathrm{X}_2 + \cdots$$

Random Forest

Random Forest (RF) is a decision-tree based ensemble method that is simple yet very powerful. It was first created by Tin Kam Ho in 1995 [20]. It creates several decision trees and classifies the data samples individually. Their decisions are then summed to find the greater number of votes to a class and thus reducing the overall error. This combining of decision trees is known as 'bagging'. In essence, if there is some correlation between the data samples and low correlation between predictions of individual trees, then the algorithm will converge to the right direction.

K-Nearest Neighbor

K-Nearest Neighbor [21] is another simple supervised learning algorithm used to model both classifications and regressions. It works on the basic idea that data points of the same kinds are in the same vicinity as each other and uses this idea to classify any given data point. For any given data point, K stands for the number of neighbors chosen to take the votes from i.e., the closest K neighbors are enclosed in a perimeter and the data sample is classified depending on the class of maximum number of neighbors. The parameter weight can be set to either uniform or w.r.t. distance. Fig. 4 (right) demonstrates the decision boundary of two classes generated by KNN.

Neural Network

A neural network (NN) is composed of an input layer, an output layer and layers in between called hidden layers [22]. The input layer consists of the individual values denoting the smallest unit of the input. The output layer consists of as many outputs as there are classes in the specific classification problem. Each hidden layer is responsible for recognizing a specific pattern. Each of the connections from all the neurons of one layer to the neurons of another layer has a weight assigned to them. The activation of a neuron can be found by calculating the weighted average of all the neurons connected to it from the previous layer along with a bias; this weighted average is then brought between 0 and 1 using a neural activation function such as the sigmoid function. The activation of a neuron can be found by calculating the weighted average of all the neurons connected to it from the previous layer along with a bias; this weighted average is then brought

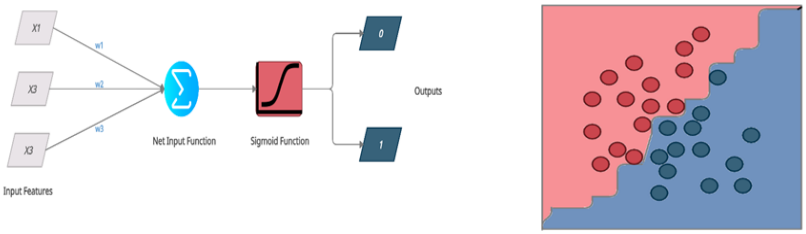

Fig. 4. Logistic regression (left) Decision boundary of two classes generated by KNN (right)

0 and 1 using a neural activation function such as the sigmoid function.

4 Experimental Results and Discussion

In our work, we have employed the Breast Cancer Wisconsin Dataset that contains digitized FNA images of breast masses, that depicts the characteristics of cell nuclei. The dataset constitutes a total of 569 images; 357 benign and 212 malignant. 30 different features of each image were used to train our classifiers.

4.1 Performance Evaluation

In order to analyze the classification performance of different classifying models, we have employed five different models of ML and one DL model. The models were studied in order to see how accurately they were able to distinguish between benign tumors and malignant tumors (cancerous); malignant being the positive class and benign being the negative.

Moreover, we have made use of different performance measure parameters such as Specificity/TNR, Recall/TPR, FNR, FPR, Precision, F1-Score and MCC, using the Confusion matrices in order to provide more evaluations. The following segment discusses all the results obtained.

Performance Measure Parameters

Recall/True Positive Rate – Recall indicates what proportion of the total Malignant samples were correctly predicted as Malignant.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

Specificity/True Negative Rate – Specificity indicates what proportion of the total Benign samples were correctly predicted as Benign.

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

False Negative Rate – It indicates what proportion of the total Malignant samples were incorrectly classified as Benign.

$$\text{FNR} = \text{FN} / (\text{TP} + \text{FN})$$

False Positive Rate – It indicates what proportion of the total Benign samples were incorrectly classified as Malignant.

$$\text{FPR} = \text{FP} / (\text{TN} + \text{FP})$$

Precision – It indicates what proportion of the total positive predictions (all samples there were predicted as Malignant) were actually positive.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

F1-Measure – The F1-Measure represents an overview of the overall performance of the model on the test dataset.

$$\text{F1} - \text{Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

MCC (Matthews Correlation Coefficient) – The MCC is suitable for use in binary classification tasks. It's value ranges from +1 to -1. Closer the value of MCC is to +1, better the performance of the model, and further away it is, poorer the performance.

$$\text{MCC} = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)\ (TP + FN)\ (TN + FP)\ (TN + FN)}}$$

Confusion Matrices

For determining the different performance measure parameters, we will be using the confusion matrix of each of the classifier models in order to derive TP (True Positive), FP (False Positive), True Negative (TN) and False Negative (FN). Malignant is the positive class while Benign has been assumed to be the negative class. Table 3 summarizes the parameters that can be derived using the confusion matrices of each of the classifiers.

True Positive (TP) – Indicates number of images that were predicted as Malignant and were correctly predicted.

False Positive (FP) – Indicates number of images that were predicted as Malignant but were in-fact Benign.

True Negative (TN) – Indicates number of images that were Benign and were correctly predicted as being Benign.

Table 3. Confusion matrix parameter values for different models

Classifiers	**TN**	**FP**	**TP**	**FN**
SVM	70	2	36	6
Decision Tree	70	2	37	5
Logistic Regression	71	1	36	6
Random Forest	72	0	39	3
KNN	72	0	33	9
Neural Network	71	1	32	10

False Negative (FN) – Indicates number of images that were Malignant but were predicted as Benign.

4.2 Performance Comparative Analysis

The performances of the ML and NN models were measured by determining the accuracy of the total correct predictions made with respect to the total number of predictions made. Fig. 5 presents the visual representation of the overall accuracies of the different models that were trained and tested on the dataset. The first conclusion that can be drawn from these values is that the Random Forest classifier had the best overall performance compared to all other models. It had achieved the highest performance accuracy of 97.37% when implemented with default parameters and the second-best performance accuracy of 96.66% when implemented with tuned parameters. Additionally, after implementing 10-fold cross validation on all the models, Random Forest had the highest cross validation accuracy of 96.84% among all the models, which proves that the model can generalize most effectively on new dataset. On the other hand, the NN classifier had the minimal overall performance accuracy of 90.35% with split and also showed very low cross validation accuracy of 85.61% in regards to the machine learning classifiers. Finally, we can also deduce that the performances of the ML models improved slightly after tuning the parameters.

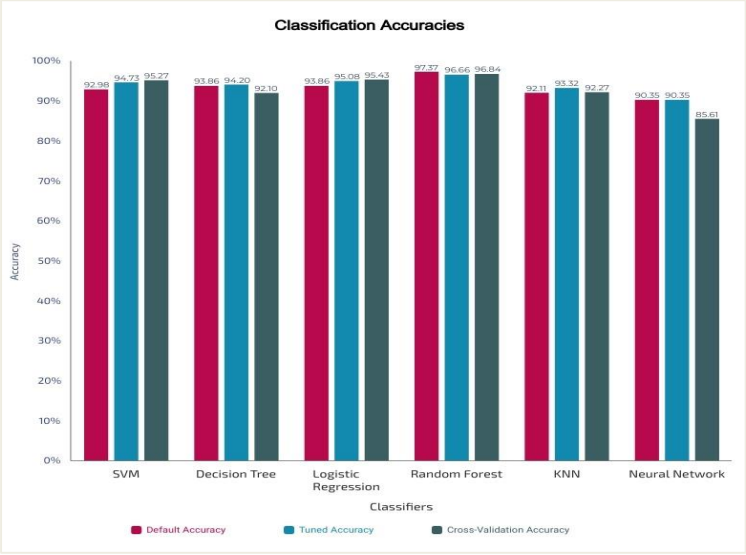

Fig. 5. Classification Accuracies of the different machine learning models

From Table 4 and Fig. 6, we can observe that Random Forest (tuned) outperformed all the other models with an accuracy of 96.66%, outperforming all the other models. Logistic Regression (tuned) was the second-best performing model on our dataset, with an accuracy of 95.08%. The DL network had the least classification accuracy – 90.35%. In addition, Random Forest had the highest F1 -

Table 4. Performance Measure Parameter Values for the different Classifiers

Parameters	SVM	Decision Tree	Logistic Regression	Random Forest	KNN	Neural Network
Tuned Accuracy (%)	94.73	94.20	95.08	96.66	93.32	90.35
Recall/TPR	0.857	0.881	0.857	0.929	0.786	0.762
Specificity/TNR	0.972	0.972	0.986	1	1	0.986
FNR	0.143	0.119	0.143	0.071	0.214	0.238
FPR	0.028	0.028	0.014	0	0	0.014
Precision	0.947	0.949	0.973	1	1	0.970
F1-Measure	0.900	0.914	0.911	0.963	0.880	0.854
MCC	0.849	0.900	0.869	0.944	0.836	0.796

Measure. All parameters considered, Random Forest Classifier has proven to be the best performing model.

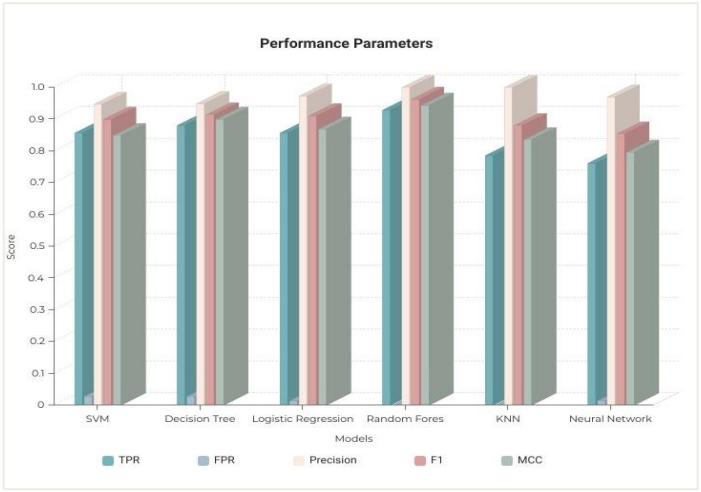

Fig. 6. Performance measure summary

4.3 Comparative Analysis of the proposed Classifier with existing models

In order to measure the robustness and effectiveness of the proposed methodology, the best performing model has been compared with other state-of-the-art models

which similarly, have been trained on the Wisconsin Diagnosis Breast Cancer (WDBC) dataset. In the proposed methodology, the Random Forest classifier had the best performance accuracy of 97.37% and from Table 5, we can find that it outperforms classifiers such as Naïve Bayes [23], Classification & Regression Trees (CART) [24] and Extra Trees Classifier [25] by a notable margin. On the other hand, classifiers such as Artificial Neural Network (ANN) [26] and Probabilistic Neural Network [7] outperformed the Random Forest classifier by a very small margin, having achieved 98.57% and 98.15% performance accuracies respectively. The better performance can be associated with the fact that these classifiers had been trained on a larger set of images than the Random Forest model. It is therefore, safe to state that the proposed Random Forest classifier is highly robust and is on par with the performances of most of the state-of-the-art models. However, it is noteworthy to mention that the AdaBoost Classifier [25] can be said to have achieved the soundest performance accuracy when taking into the account the number of images it has been trained on. The AdaBoost Classifier had been trained on 569 images, same as that of the Random Forest classifier and lower than other models; it achieved a performance accuracy of 98.23% which is higher but very marginal compared to the proposed classifier.

Table 5. Performance Comparison of the proposed classifiers with other existing models

Classifiers	Performance Accuracy	Dataset Implemented
Random Forest [Proposed]	97.37%	Wisconsin Diagnosis Breast Cancer (WDBC) 569 images; 357 benign and 212 malignant
Artifical Neural Network [26]	98.57%	√
Probabilistic Neural Network [7]	98.15%	√
Naive Bayes [23]	94.47%	√
CART [24]	92.00%	√
AdaBoost Classifier [25]	98.23%	√
Extra Trees Classifier [25]	97.34%	√

5 Conclusion and Future Work

Medical technology is constantly looking for ways to automate steps by either

eliminating tedious steps or by providing a diagnosis to confirm. This can help the medical industry to provide healthcare efficiently and spontaneously. This literature is an attempt to contribute to this notion. In this study, we have used digitized images of FNA of breast masses to classify them as either malignant or benign. There were comparisons made using both machine learning and deep learning techniques. Furthermore, stability of the models was verified using cross-validation technique. The best performance was obtained by tuning the parameters of a random forest model with an accuracy of 96.66%. Logistic regression (with tuning) followed closely with an accuracy of 95.08%. The lowest performing model was the deep neural network (with cross-validation) with an accuracy of 85.61%. In future, we will be implementing models of DNN and transfer learning to increase the model's efficiency.

References

1. U. S. C. S. W. Group and others, "United States cancer statistics: 1999--2012 incidence and mortality web-based report," *Atlanta Dep. Heal. Hum. Serv. centers Dis. Control Prev. Natl. cancer Inst.*, 2015.
2. H. Zhu and B. E. Do\ugan, "American Joint Committee on Cancer's Staging System for Breast Cancer: Summary for Clinicians," *Eur. J. Breast Heal.*, vol. 17, no. 3, p. 234, 2021.
3. C. Lerman *et al.*, "Mammography adherence and psychological distress among women at risk for breast cancer," *JNCI J. Natl. Cancer Inst.*, vol. 85, no. 13, pp. 1074–1080, 1993.
4. P. T. Huynh, A. M. Jarolimek, and S. Daye, "The false-negative mammogram.," *Radiographics*, vol. 18, no. 5, pp. 1137–1154, 1998.
5. M. G. Ertosun and D. L. Rubin, "Probabilistic visual search for masses within mammography images using deep learning," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2015, pp. 1310–1315.
6. D. C. Pereira, R. P. Ramos, and M. Z. Do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Comput. Methods Programs Biomed.*, vol. 114, no. 1, pp. 88–101, 2014.
7. E. D. Übeyli, "Implementing automated diagnostic systems for breast cancer detection," *Expert Syst. Appl.*, vol. 33, no. 4, pp. 1054–1062, 2007.
8. D. Bazazeh and R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis," in *2016 5th international conference on electronic devices, systems and applications (ICEDSA)*, 2016, pp. 1–4.
9. N. Khuriwal and N. Mishra, "Breast cancer detection from histopathological images using deep learning," in *2018 3rd international conference and workshops on recent advances and innovations in engineering (ICRAIE)*, 2018, pp. 1–4.
10. W. E. Fathy and A. S. Ghoneim, "A deep learning approach for breast cancer mass detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 175–182, 2019.
11. A. L. S. Saabith, E. Sundararajan, and A. A. Bakar, "Comparative study on different classification techniques for breast cancer dataset," *Int. J. Comput. Sc.*

Mob. Comput, vol. 3, no. 10, pp. 185–191, 2014.
12. M. Amrane, S. Oukid, I. Gagaoua, and T. Ensari, "Breast cancer classification using machine learning," in *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, 2018, pp. 1–4.
13. M. Nawaz, A. A. Sewissy, and T. H. A. Soliman, "Multi-class breast cancer classification using deep learning convolutional neural network," *Int. J. Adv. Comput. Sci. Appl*, vol. 9, no. 6, pp. 316–332, 2018.
14. "Breast Cancer Wisconsin (Diagnostic) Data Set," UCI machine Learning Repository: Breast Cancer Wisconsin (diagnostic) data set. [Online]." https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29 (accessed Jul. 05, 2021).
15. S. Sakib, M. A. B. Siddique, and M. A. Rahman, "Performance Evaluation of t-SNE and MDS Dimensionality Reduction Techniques with KNN, ENN and SVM Classifiers," in *2020 IEEE Region 10 Symposium (TENSYMP)*, 2020, pp. 5–8.
16. K. M. Hasib *et al.*, "A survey of methods for managing the classification and solution of data imbalance problem," *arXiv Prepr. arXiv2012.11870*, 2020.
17. H. Cho *et al.*, "A similarity study of content-based image retrieval system for breast cancer using decision tree," *Med. Phys.*, vol. 40, no. 1, p. 12901, 2013.
18. K. M. Hasib, M. I. H. Showrov, and A. Das, "Accidental prone area detection in bangladesh using machine learning model," in *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*, 2020, pp. 58–62.
19. H.-L. Hwa *et al.*, "Prediction of breast cancer and lymph node metastatic status with tumour markers using logistic regression models," *J. Eval. Clin. Pract.*, vol. 14, no. 2, pp. 275–280, 2008.
20. H. T. Kam and others, "Random decision forest," in *Proceedings of the 3rd international conference on document analysis and recognition*, 1995, vol. 1416, p. 278282.
21. M. Siddique, A. Bakr, S. Sakib, M. Rahman, and others, "Performance Analysis of Deep Autoencoder and NCA Dimensionality Reduction Techniques with KNN, ENN and SVM Classifiers," *arXiv Prepr. arXiv1912.05912*, 2019.
22. M. A. B. Siddiaue, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, "Deep Convolutional Neural Networks Model-based Brain Tumor Detection in Brain MRI Images," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, 2020, pp. 909–914.
23. S. Sharma, A. Aggarwal, and T. Choudhury, "Breast cancer detection using machine learning algorithms," in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 2018, pp. 114–118.
24. H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "Using machine learning algorithms for breast cancer risk prediction and diagnosis," *Procedia Comput. Sci.*, vol. 83, pp. 1064–1069, 2016.
25. H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, and M. Faisal Nagi, "Automated breast cancer diagnosis based on machine learning algorithms," *J. Healthc. Eng.*, vol. 2019, 2019.
26. M. M. Islam, M. R. Haque, H. Iqbal, M. M. Hasan, M. Hasan, and M. N. Kabir, "Breast cancer prediction: a comparative study using machine learning techniques," *SN Comput. Sci.*, vol. 1, no. 5, pp. 1–14, 2020.