Biomedical Signal Processing and Control 71 (2022) 103141

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Computer-aided detection of breast cancer on the Wisconsin dataset: An artificial neural networks approach

Mohammad H. Alshayeji [a,*], Hanem Ellethy [b], Sa'ed Abed [a], Renu Gupta [c]

[a] Computer Engineering Department, College Engineering and Petroleum, Kuwait University, Kuwait
[b] Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia
[c] Department of Radiology, Faculty of Medicine, Kuwait University, Kuwait

ARTICLE INFO

Keywords:
Artificial neural networks (ANN)
Breast cancer classification
Wisconsin breast cancer dataset (WBCD)
Wisconsin diagnostic breast cancer (WDBC)

ABSTRACT

The early detection of breast cancer (BC) has a significant impact on reducing the disease's mortality rate. As an effective cost- and time-saving tool, computer-aided diagnosis (CAD) systems have been developed in this research field, aiding clinicians and radiologists' decision-making process by offering highly accurate information. In this study, a shallow artificial neural network (ANN) model with one hidden layer is used to diagnose and predict BC using the Wisconsin breast cancer dataset (WBCD) and the Wisconsin diagnostic breast cancer (WDBC) dataset without employing feature optimization or selection algorithms. The datasets are divided into 80% for training and 20% for testing using five-fold cross-validation. The model's effectiveness and efficiency are evaluated based on sensitivity, specificity, precision, accuracy, and F_1 score, along with the area under the receiver-operating characteristic curve (AUC). In the task of classifying benign and malignant tumours using the WBCD, the shallow ANN model showed promising performance with an average accuracy of 99.85%, specificity of 99.72%, sensitivity of 100%, precision of 99.69%, and F_1 score of 99.84%. For BC detection using WDBC, it achieved an average accuracy of 99.47%, specificity of 99.53%, sensitivity of 99.59%, precision of 98.71%, and F_1 score of 99.13%. The AUC of the proposed model was 99.86% and 99.56% for the WBCD and the WDBC dataset, respectively, illustrating the model's discrimination capacity. Moreover, the ANN model outperformed state-of-the-art models that integrate feature optimization and selection algorithms to classify BC tumours using WBCD and WDBC. Hence, the shallow ANN model presented here demonstrates significant potential for diagnosing BC using WBCD and WDBC without the need for feature optimization or selection algorithms.

1. Introduction

Breast cancer (BC) is the second-largest killer of women globally, and the early detection of the disease is key to tackling this high mortality rate. The causes of BC are unknown, and many women develop the disease without showing any symptoms. Early detection, facilitated by regular breast screening, can enable early treatment, especially for women with a high or average BC risk [1,2].

Computer-aided diagnosis (CAD) systems are widely used throughout the medical field to reduce costs, aid clinicians and radiologists in detecting disease, and differentiate between disease courses. Many studies in the field of BC use CAD systems based on thermography [3], ultrasound [4–6], mammogram [7–10], magnetic resonance imaging (MRI) [11,12], histopathological biopsy images [13,14] or combinations thereof [15]. CAD systems are used to both detect BC and classify the detected cancer as benign or malignant. CAD systems usually

Abbreviations: AMMLP, Artificial Metaplasticity Multi Layers Perceptron; ANN, Artificial Neural Networks; AR-ANN, Association Rules and Artificial Neural Networks; AUC, Area Under the receiver-operating characteristic Curve; BC, Breast Cancer; BPNN, Back-Propagation Neural Networks; CAD, Computer-Aided Diagnosis; CI, Confidence Interval; CNN, Convolutional Neural Networks; DBN, Deep Belief Networks; DM, Digital Mammography; DT, Decision Tree; EANN, Evolutionary Artificial Neural Networks; FNA, Fine-Needle Aspirate; GOANN, Genetically Optimized Neural Network; GRNN, General Regression Neural Network; GS, Grid Search; K-NN, k-Nearest Neighbours; LDA, Linear Discriminant Analysis; ML, Machine Learning; MLP, Multilayer Perceptron; MRI, Magnetic Resonance Imaging; OGCNN, One-pass Generalized Classifier Neural Network; PSO, Particle Swarm Optimization; RF-ANN, Rotation Forest Artificial Neural Networks; ROC, Receiver Operating characteristic Curve; RS-BPANN, Rough Sets and Back-Propagation Neural Network; SVM, Support Vector Machine; WBCD, Wisconsin Breast Cancer Dataset; WDBC, Wisconsin Diagnostic Breast Cancer; WPBC, Wisconsin Prognostic Breast Cancer.

* Corresponding author at: Computer Engineering Department, College of Engineering and Petroleum, Kuwait University, P.O. Box: 5969, Safat 13060, Kuwait.
E-mail address: m.alshayeji@ku.edu.kw (M.H. Alshayeji).

https://doi.org/10.1016/j.bspc.2021.103141
Received 23 January 2021; Received in revised form 30 August 2021; Accepted 2 September 2021
Available online 15 September 2021
1746-8094/

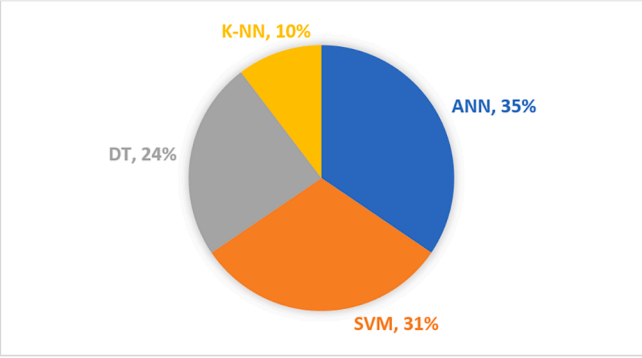

Fig. 1. Pie chart of the ML models used in BC-CADs based on WBCD in [20]. K-NN: k-nearest neighbours; SVM: Support vector machine; DT: Decision tree.

involve five basic steps following the data acquisition: image pre-processing, image segmentation, feature extraction, feature selection and optimization, and classification. However, these steps may vary between systems due to the different algorithms and techniques involved.

Dataset acquisition is a difficult task when building a machine learning (ML) based system, especially in the medical field as issues arise regarding privacy, confidentiality, ethics, and security [16]. Fortunately, the Wisconsin breast cancer dataset (WBCD) recently emerged as a benchmark dataset for BC detection and classification [17] and is available online at the UCI machine learning repository [18], along with other Wisconsin BC datasets. Moreover, this dataset contains the extracted features of the acquired images after the image pre-processing, image segmentation, and feature extraction steps have been conducted.

Feature selection and/or optimization is a technique used to improve the classification accuracy and prediction outcomes by removing redundant and irrelevant features to reduce the computational complexity. Consequently, as an optimization step, feature selection can reduce the amount of training phase data required, thereby speeding up the learning process and enhancing the ML model's outcomes [19]. However, this step requires additional time and computation costs, and it may affect the data quality, especially when a small number of features are applied. Meanwhile, although feature transformation and dimensionality reduction are used for the same reason, these are mainly applied to high-dimensional data containing extensive features.

Classification is a complex step because it depends on the extracted and selected features as inputs as well as the used ML model and its parameters. A variety of ML models have been applied to WBCD, using nine key features to detect BC and classify it as either benign or malignant. According to a review conducted in 2018 [20], ANN is the most widely used model for building BC-CAD systems, as shown in Fig. 1. However, different optimization and feature selection algorithms have been used to build BC-CAD systems that utilize ANN models, namely deep belief networks [21], pruning processes [15,16], evolutionary algorithms [24], association rules [25], metaplasticity processes [26], rotation forest algorithms [27], genetic programming [28], and rough-set techniques [29]. However, even though most models using feature selection and/or optimization techniques to enhance the classification accuracy [19], a shallow ANN model can achieve comparable performance if the combined parameters and hyperparameters are carefully selected and tuned.

This study aims to test a shallow ANN model to detect breast cancer and classify malignant and benign tumours using two public datasets without the integration of a feature optimization or selection approach. As reported in the literature, ANN is a powerful classifier and one of the most commonly used to detect BC using the publicly available WBCD [20,30]. In this study, we hypothesize that an ANN with a shallow design – one hidden layer – can be effectively used to diagnose and classify BC using the WBCD and the Wisconsin diagnostic breast cancer (WDBC) datasets. This method requires no feature optimization or selection techniques to build this system. The block diagram of the proposed BC-CAD system is shown in Fig. 2. First, the required step of dataset pre-processing prepares and cleans the data, i.e., any rows with missing features are deleted. Next, the model is set up and the hyperparameters are tuned to select the best combination for training the model. Finally, the model is tested to measure its performance.

The contributions of this study include:

- Using the WBCD and the WDBC dataset to implement a CAD system to detect BC.
- Testing the ability of a shallow ANN model to classify and predict BC tumours as malignant or benign without applying feature optimization or selection algorithms.
- Tuning the hyperparameters of the ANN model to achieve a comparable performance, which is achieved using one hidden layer with 100 neurons, a rectified linear unit (ReLU) activation function in the hidden layer, a sigmoid activation function in the output layer, and an Adam algorithm to update weights and biases.
- Achieving high performance in terms of sensitivity, specificity, precision, and F_1 score to demonstrate system efficiency while maintaining system effectiveness regarding the accuracy of the classified instances.
- Comparing the model's performance with state-of-the-art alternatives designed to utilize the WBCD and the WDBC dataset, resulting in superior performance in terms of sensitivity, specificity, and accuracy.

This paper is organized as follows: Section 2 describes related work. Section 3 discusses the materials and methodology. The results and discussion are presented in Section 4. Finally, the conclusion and future trends are presented in Section 5.

2. Related work

Various surveys have been conducted to review and summarize the CAD systems used in BC detection and classification [30–34]. A study of classical CAD systems for BC detection was conducted in [32] to analyse various CAD systems based on mammogram, breast ultrasound, MRI, and biopsy images. Different approaches were explored according to the four main steps of CAD systems, i.e., namely image pre-processing, image segmentation, feature extraction, and classification. Image pre-processing is the basic, preliminary step taken to remove noise from the images while also normalizing and enhancing them. Segmentation techniques can be divided into three main categories: discontinuity-based approaches, similarity-based approaches, and others. Feature extraction can also be classified into three groups: shape-based descriptors, textural descriptors, and colour-based descriptors. Meanwhile,

Fig. 2. Block diagram of the proposed breast cancer CAD system. ANN: Artificial neural networks.

classification can be broadly categorised into supervised and unsupervised techniques. At the end of each section, the authors presented a comparison table containing the name, definition, reference, and pros and cons of each approach. In addition, six recent classification techniques were compared based on accuracy, sensitivity, specificity, positive predictive value, and negative predictive value, while the segmentation techniques were compared based on accuracy alone.

A systematic review was conducted to investigate all BC-CAD systems published in academic journals or presented at institute of electrical and electronics engineers (IEEE) conferences between 2012 and 2017 [30]. According to the survey, the most widely used ML models are the ANN and the support vector machine (SVM), while digital mammography (DM) is the most commonly used modality in BC-CAD systems.

ANNs have been used effectively in different BC-CAD systems [7–9,35–37], demonstrating good performance in terms of accuracy, sensitivity, and specificity in identifying benign and malignant breast tumours through ultrasound imaging by utilizing biclustering mining algorithms [4,5]. ANN was used to improve a CAD system for BC malignancy grading using fine-needle aspirate (FNA) cytological slides [31], before being integrated with various optimization algorithms for the detection of BC in DM, such as swarm optimization [7], two-dimensional wavelet transforms [30], wavelet transformation and swarm optimization [6], and weighted-gray levels and local difference features [38].

Several ML models developed to detect BC using WBCD were presented and evaluated in terms of accuracy in [20], focusing on the 1996–2017 period. Out of 29 models, ten applied ANN, nine used SVM, seven applied decision tree (DT), and three used k-nearest neighbours (k-NN). All models showed high performance in terms of accuracy but applied different algorithms to select, reduce, and optimize the features before feeding them into the ML model. In addition, ANN was reported to show superior pattern recognition performance compared to other models in [39]. In detecting BC using the WBCD via other ML models, SVM outperformed other classifiers in terms of accuracy and sensitivity [40], while decision tree forest achieved the best accuracy, specificity, and receiver operating characteristic curve (ROC) compared to the other reported models [41]. Moreover, quadratic SVM was superior to k-NN and DT in classifying BC using WDBC [42].

Rough-set feature selection based on k-means clustering was applied in [43] to select the most effective WDBC and Wisconsin prognostic breast cancer (WPBC) dataset features using an SVM classifier. The performance achieved by integrating grid search (GS), particle swarm optimization (PSO), and the uniform configuration Weka tool with SVM, DT, and ANN was tested for the classification of BC data in the WDBC, WPBC, and WBCD [44]. Although GS and PSO affected the three classifiers positively, ANN and SVM outperformed DT in terms of accuracy, recall, and precision.

In another study, a combination of feature reduction approaches and ML classifiers was analysed to select the best integration model for BC detection using the WDBC dataset [45]. Linear discriminant analysis (LDA) was found to be the most effective for diminishing a feature's dimensionality and positively influencing SVM, ANN, and DT, while SVM-LDA was selected due to the shorter computational time. In addition, other studies have applied the genetic algorithm [46] and the wrapper method [47–49] to select and optimize the features used to train the selected classifier.

Yue et al. confirmed ANN as the most commonly used ML model for classifying and detecting BC using the WBCD [20], while ANN is reported to have the same level of efficiency as deep convolutional neural networks (CNN) in BC diagnosis [50]; however, most of the reported models applied feature selection and/or optimization algorithms to boost the classifier's performance. Despite the absence of feature optimization or reduction approaches in our proposed model, ANN shows its capacity to detect BC using the complete features of the WBCD and the WDBC dataset. In addition, the hyperparameters of the proposed model are carefully selected to achieve a promising performance while maintaining a fast model in terms of processing time. The proposed model has been speedily trained and tested, achieving superior BC detection performance.

Table 1
WBCD and WDBC dataset details.

Instances	Features	Malignant	Benign
WBCD			
699	9	241 (34.5%)	458 (65.5%)
WDBC			
569	30	212 (37.2%)	357 (62.8%)

WBCD: Wisconsin breast cancer dataset; WDBC: Wisconsin diagnostic breast cancer.

3. Materials and methodology

Two BC-CAD systems are introduced in this study: an ANN-WBCD system using nine features for each patient, and an ANN-WDBC system employing 30 features for each patient – without applying a feature selection or an optimization algorithm. By evaluating the two systems and comparing them to state-of-the-art models using the same dataset, we successfully prove our hypotheses. Moreover, the two systems are accurate and offer swift convergence.

3.1. Materials

The WBCD and the WDBC dataset are used to evaluate the ability of the ANN to classify BC tumorous using all the given features. Acquired from University of Wisconsin hospitals in the U.S., the datasets were created by Dr. William Wolberg. Both the WBCD and the WDBC dataset are available online at the UCI Machine Learning Repository [18].

The WBCD contains 699 instances with a nine-feature vector size; the nuclear features are assigned via a visual assessment of an FNA sample taken from a suspicious breast mass. Each feature has a value between 1 and 10 based on the diagnosis, with 1 referring to normal or benign cases and 10 indicating the most abnormal. Malignant cases are confirmed by biopsy examination, and benign cases are identified based on biopsy or periodic examination. Feature vectors contain the assigned values of clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nuclei, and mitoses [17].

The WDBC dataset contains 569 instances with a feature vector size of 30; features are extracted from a digital image of an FNA sample taken from a patient's breast mass. The image pre-processing and analysis steps are conducted as follows: 1) the cellular nuclei of the FNA slide are highlighted using a microscope, 2) a digital camera and a frame-grabber board are used to scan the well-identified slide portion taken from the FNA sample, 3) Xcyt software is used to isolate the individual nuclei interactively, using a computer mouse to identify the approximate boundary of the nucleus (subsequently, the exact boundaries are identified automatically), and 4) 30 features of each nucleus are extracted and calculated using Xcyt. The features include the mean, standard error, and extreme values of the real-valued features. The real-valued features contain the radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimensions of each nucleus [51]. The details of the WBCD and the WDBC dataset are presented in Table 1.

3.2. Methodology

The main goal of this study is to investigate the ability of a shallow ANN model to classify BC using the WBCD and the WDBC dataset without applying a feature optimization or selection algorithm. The ANN model represents the backbone of neural networks and is

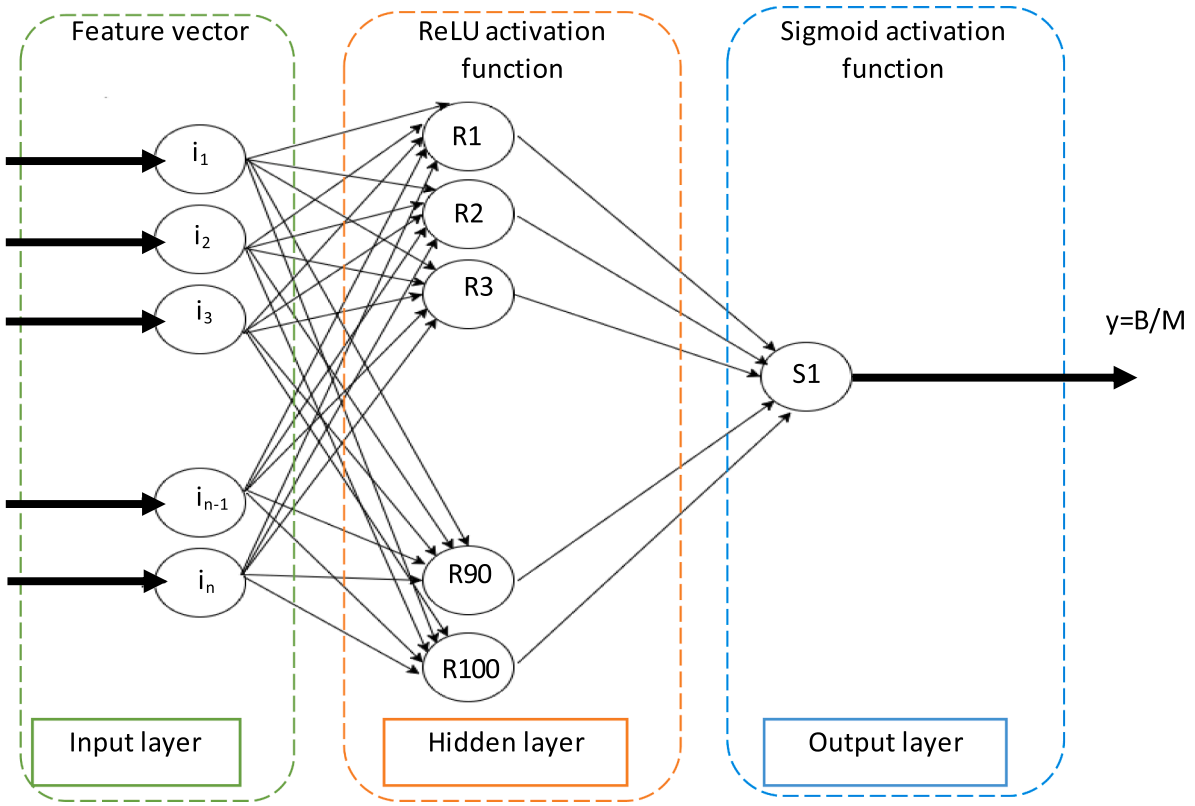

Fig. 3. The ANN model used, with 100 neurons in the hidden layer.

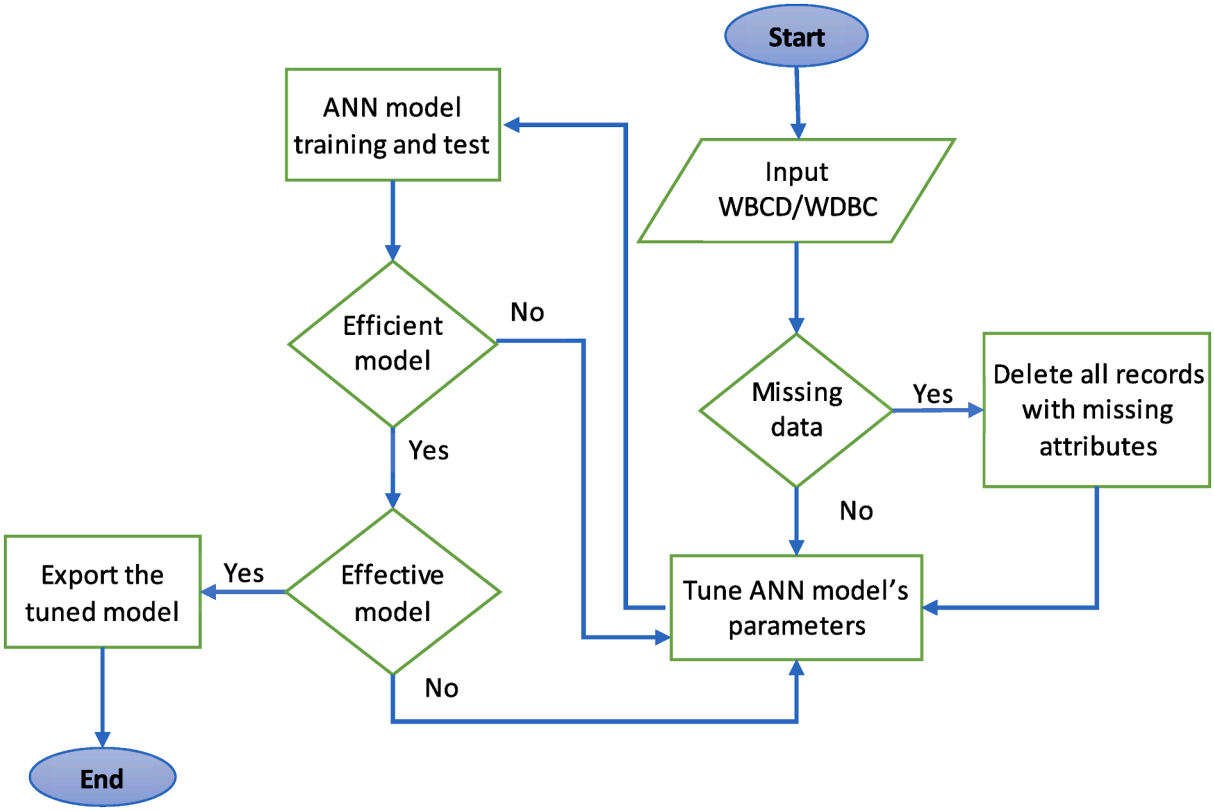

Fig. 4. Flowchart of the proposed BC-CAD system using ANN.

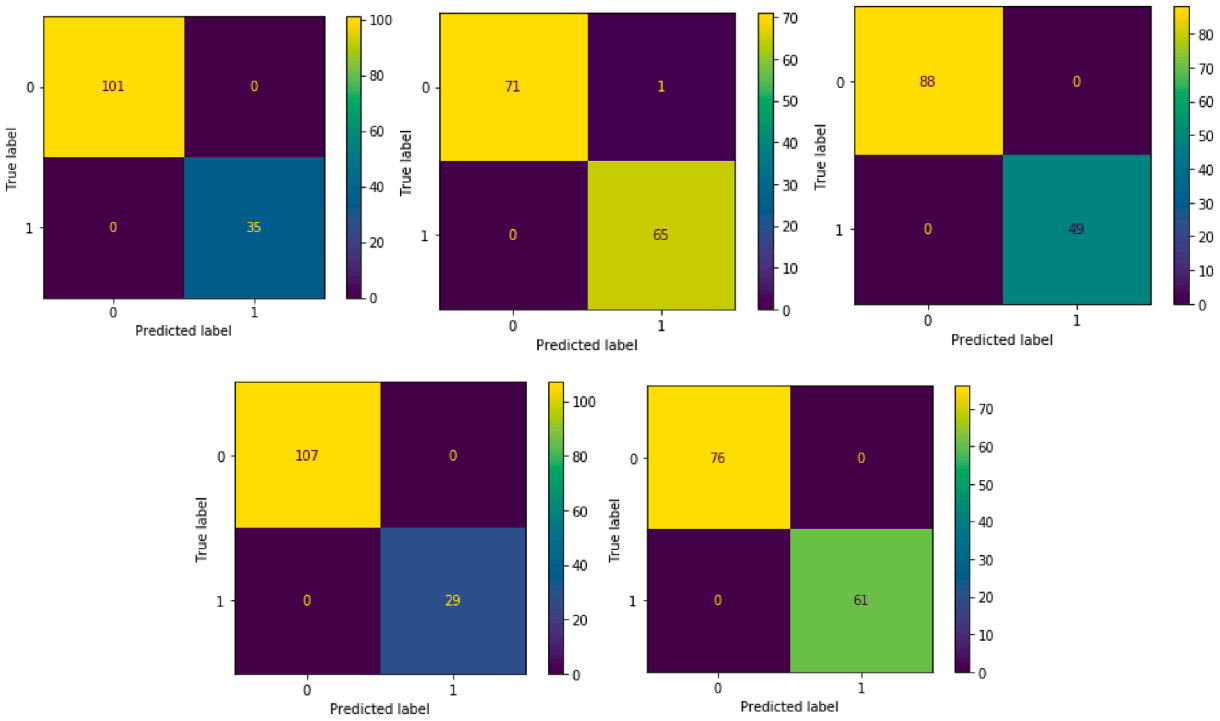

Fig. 5. ANN-WBCD confusion matrices of the testing phase using five-fold cross-validation.

considered to be a general approximator because of its capacity to approximate any problem using a sufficient number of neurons in the hidden layer. In this model, data is forwarded from inputs to outputs via hidden layers containing multiple neurons [52]. An ANN has a high recognition rate in comparison to other models, despite being simple [53,54]. The ANN model used in this study is implemented based on the model developed by Alshayeji et al. [54].

The ANN model is trained and optimized using an Adam algorithm to update the weights and biases, minimize squared errors and weights, and define the optimum combination in the establishment of a well-trained model. Moreover, as shown in Fig. 3, it consists of an input layer, a hidden layer, and an output layer. The input layer is identified based on the feature vector; therefore, the ANN-WBCD input layer has nine inputs and the ANN-WDBC has 30 inputs. The hidden layer has 100 neurons with a ReLU activation function, with the number of neurons defined based on our study's goal of achieving a model that fits the two datasets. The output layer has only one neuron with a sigmoid activation function to classify the input as malignant (M) or benign (B) outputs.

The main goal of the ANN is to map the input (the features vector) to its targeted output (benign/malignant label) by finding the weights via the hidden layer, which can be achieved using the following steps and equations:

- Initialize the weights **W= (w**$_1$, **w**$_2$, **w**$_3$, ….**w**$_n$**)** and bias **b** randomly.
- Calculate the input to the $\mathbf{R}_{th}$ neuron in the hidden layer using Eq. (1):

$$R_{th}input = \sum_{i=1}^{n} w_{iR} i_i + b \tag{1}$$

- Calculate the $\mathbf{R}_{th}$ neuron output using Eqs. (2) and (3):

$$R_{th}output = \mathrm{ReLU}\left(\sum_{i=1}^{n} w_{iR} i_i + b\right) \tag{2}$$

$$ReLU(x) = max(0, x) \tag{3}$$

- Calculate the output of the output layer neuron S1 using Eqs. (4) and (5):

$$S1 = \sigma\left(\sum_{i=1}^{100} R_i output\right) \tag{4}$$

$$\sigma\,(x) = \frac{1}{1 + \ e^{-x}} \tag{5}$$

- Measure the performance of the classifier using the binary cross-entropy as a loss function between the output S1 and the label y using the Eq. (6)

$$loss = \mathrm{y}\ log(S1) + (1 - y) log(1 - S1) \tag{6}$$

- Minimize the loss function by recalibrating the weights of the hidden layer nodes using the back-propagation algorithm until the convergence criteria are met [55].
- Once the loss function is close to zero, the model becomes fully trained and the validation/testing dataset should be used to evaluate its performance.

The flowchart of the used methodology is depicted in Fig. 4. Firstly, the dataset is pre-processed and all records containing missing attributes are deleted. Secondly, the ANN model is initiated using the parameters of Alshayeji et al.'s model [54], and the model's efficiency and effectiveness are tested. Thirdly, if the model is found to be neither efficient nor effective, the ANN's parameters are tuned to retrain and test the model. Fourthly, the evaluation parameters are reported, and the model is finalized and saved for testing a new dataset.

The initial parameters of the three-layer ANN model included 18 identical neurons in the hidden layers, a sigmoid activation function also in the hidden layer, a linear activation function in the output layer, and mean square errors as a loss function. Tuning the hyperparameters to achieve our goal resulted in using an Adam algorithm as a weight optimizer, a binary cross-entropy as a loss function, a ReLU and a

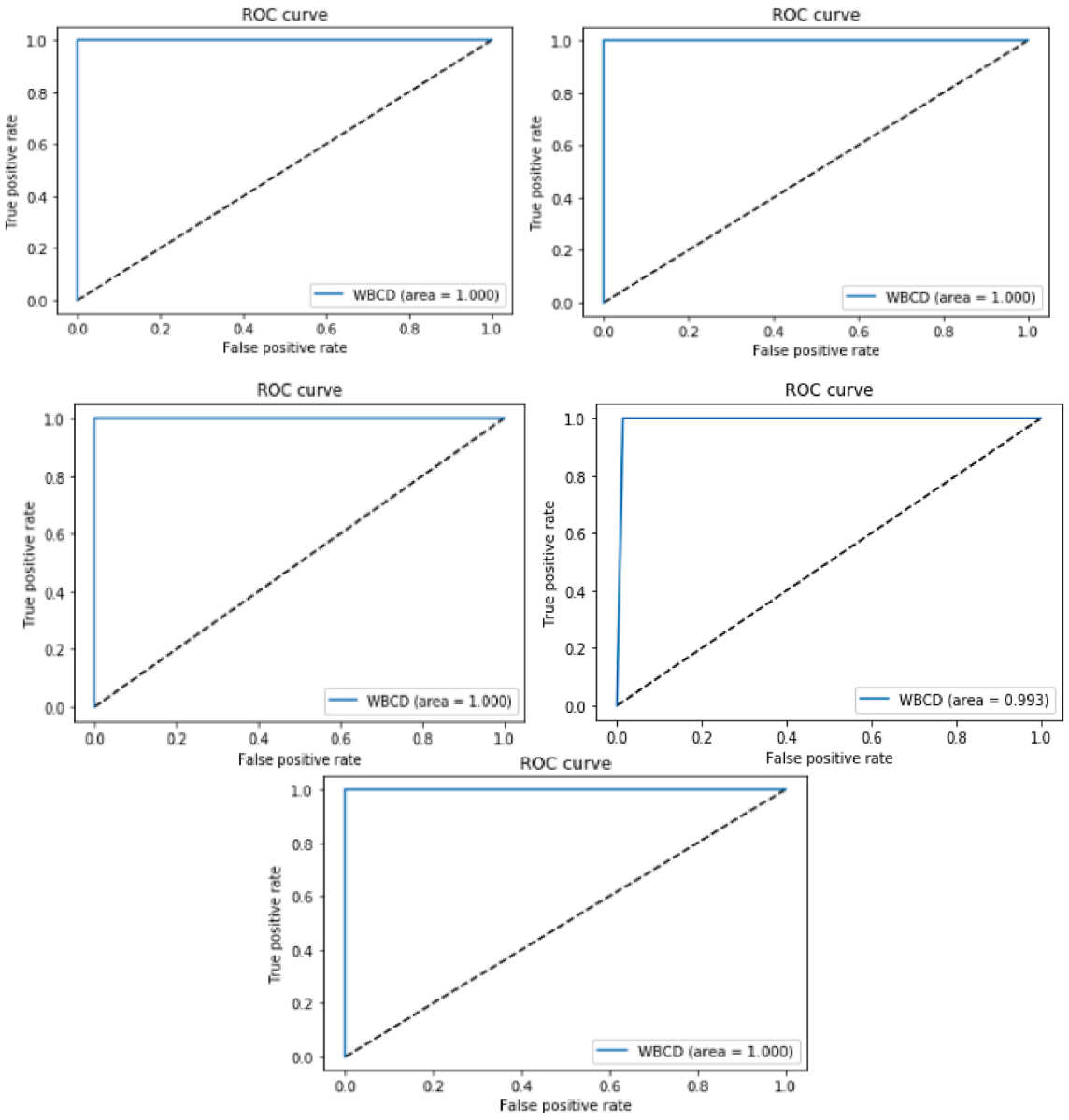

Fig. 6. ANN-WBCD ROC curves of the testing phase using five-fold cross-validation. ROC: Receiver operating characteristic curve.

sigmoid activation function in the hidden and output layers, respectively, and 100 neurons in the hidden layer.

The ANN's effectiveness is evaluated in terms of its accuracy, alongside correctly and incorrectly classified instances, while its efficiency is measured in terms of precision, sensitivity, specificity, and F_1 score (a measure of a test's accuracy). In addition, the ROC and area under ROC (AUC) are used to illustrate the classifier performance.

4. Results and discussion

This section discusses the effectiveness, efficiency, and overall performance of the proposed CAD system in identifying benign and malignant BC tumours in FNA images using the WBCD and the WDBC dataset. Both datasets are divided into two parts: 80% for training and 20% for testing with five-fold cross-validation.

1. ANN-WBCD

The WBCD contains 699 instances, with nine features for each instance, although 16 instances have missing features. Excluding those 16 instances is the preferred option, in line with other state-of-the-art studies. As the first evaluation step, the confusion matrices and ROC curves of the testing phase using five-fold cross-validation are shown in Figs. 5 and 6. The average accuracy of the testing phase is 99.85%, sensitivity is 100%, and specificity is 99.72%. The ROC is a graph depicting the binary classifier's diagnostic ability at various tuning thresholds [56]. As shown in Fig. 6, the ANN-WBCD is a power classifier as it begins from the lower left corner, moving directly to the upper left corner and then straight to the upper right corner. Thus, the ANN-WBCD has an average AUC of 99.86%. Moreover, the proposed model

Table 2
ANN-WBCD evaluation parameters.

Effectiveness evaluation parameters		Efficiency evaluation parameters	
Correctly classified	**682**	Precision, % (95% CI)	**99.7 ± 0.6**
Incorrectly classified	**1**	Sensitivity, % (95% CI)	**100 ± 0**
Accuracy, % (95% CI)	**99.85 ± 0.29**	Specificity, % (95% CI)	**99.72 ± 0.55**
		F_1 score, % (95% CI)	**99.85 ± 0.3**

CI: Confidence interval.

Table 3
Comparison regarding accuracy, sensitivity, and specificity with the state-of-the-art methods reported in [20] using the WBCD.

ANN model	Training/testing	Optimization	Overall evaluation		
			Accuracy	Sensitivity	Specificity
DBN-ANN [21]	54.9/45.1	Deep belief network	99.68%	100%	99.47%
Pruned-ANN [22]	50/50	Pruning process	95%	–	–
Neuro-rule ANN [23]	10-fold cross-validation	Pruning process	98.24%	–	–
EANN [24]	80/20	Evolutionary algorithm	98.01	–	–
AR-ANN [25]	3-fold cross-validation	Association rules	97.4%	–	–
AMMLP [26]	60/40	Metaplasticity process	99.26%	100%	97.89%
RF-ANN [27]	50/50	Rotation forest	97.95%	–	–
GOANN [28]	70/30	Genetic programming	99.21%	99.51%	99.21%
RS-BPANN [29]	80/20	Rough set	98.6%	98.76%	98.57%
GRNN [57]	50/50	–	97.6%	–	–
ANN-WBCD	**5-fold cross-validation**	**–**	**99.85%**	**100%**	**99.72%**

DBN: Deep belief networks; EANN: Evolutionary ANN; AR-ANN: Association rules and neural network; AMMLP: Artificial metaplasticity multilayer perceptron; RF-ANN: Rotation forest ANN; GOANN: Genetically optimized neural network; RS-BPANN: Rough sets and backpropagation neural network; GRNN: General regression neural network.

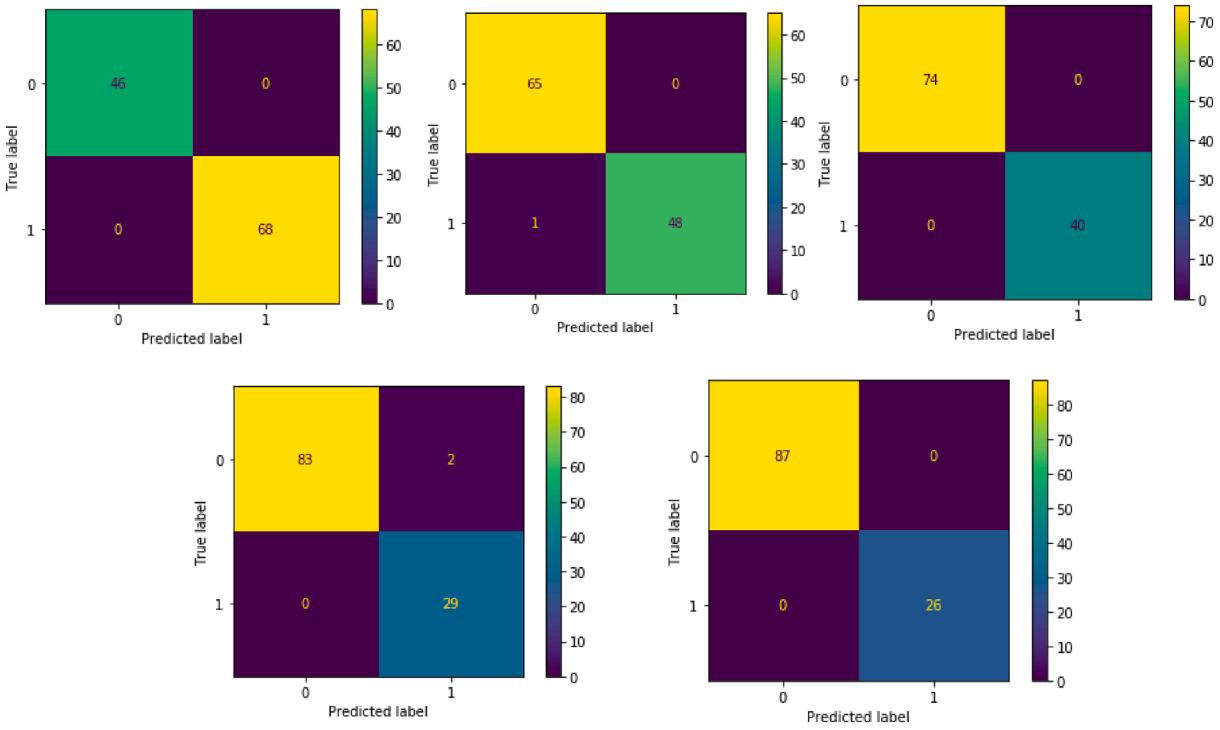

Fig. 7. ANN-WDBC confusion matrices of the testing phase using five-fold cross-validation.

converged in 16 s, which qualifies as a fast response. This result supports our hypothesis that the ANN model is sufficiently powerful to detect BC using the WBCD without applying feature optimization or selection techniques, thereby achieving a high recognition rate.

As Table 2 states, the model's effectiveness is evaluated according to the average number of correctly classified cases, the average number of incorrectly classified cases, and the average accuracy, with a 95% confidence interval (CI). The ANN-WBCD converges swiftly in both the training and testing of WBCD, correctly classifying 682 out of 683 instances, incorrectly classifying 1 out of 683 instances, and achieving a 99.85% average accuracy. Furthermore, the model's efficiency is assessed according to precision, sensitivity, specificity, and F_1-score. The ANN-WBCD has an optimum average sensitivity of 100%, a precision of 99.69%, a specificity of 99.72%, and an F_1-score of 99.84%. Our model has a robust performance, which firmly demonstrates its efficiency and effectiveness.

Comparing our results with those generated by state-of-the-art models is the second evaluation phase for investigating whether the superiority of this model stands up against previously published models that utilized feature optimization or selection techniques. Table 3 includes the results of the ANN models reported in [20], which were applied to the WBCD for BC diagnosis in terms of sensitivity, specificity, and accuracy. The comparisons show that our model's setup outperforms the reported models in terms of accuracy (99.85%), while also being superior to four models [21,26,28,29] in terms of specificity (99.72%) and sensitivity (100%). However, different feature optimization algorithms, such as pruning processes, data pre-processing, evolutionary algorithms, association rules, metaplasticity processes, rotation forest algorithms, genetic programming, and rough sets, have been integrated into ANN models to enhance their accuracy. Consequently, the proposed ANN model is proven to be adequate for detecting BC using the WBCD and achieves superior performance to DBN-ANN models that use deep belief networks (DBN) [21] or general regression neural networks (GRNN) [57], without applying feature optimization or selection algorithms.

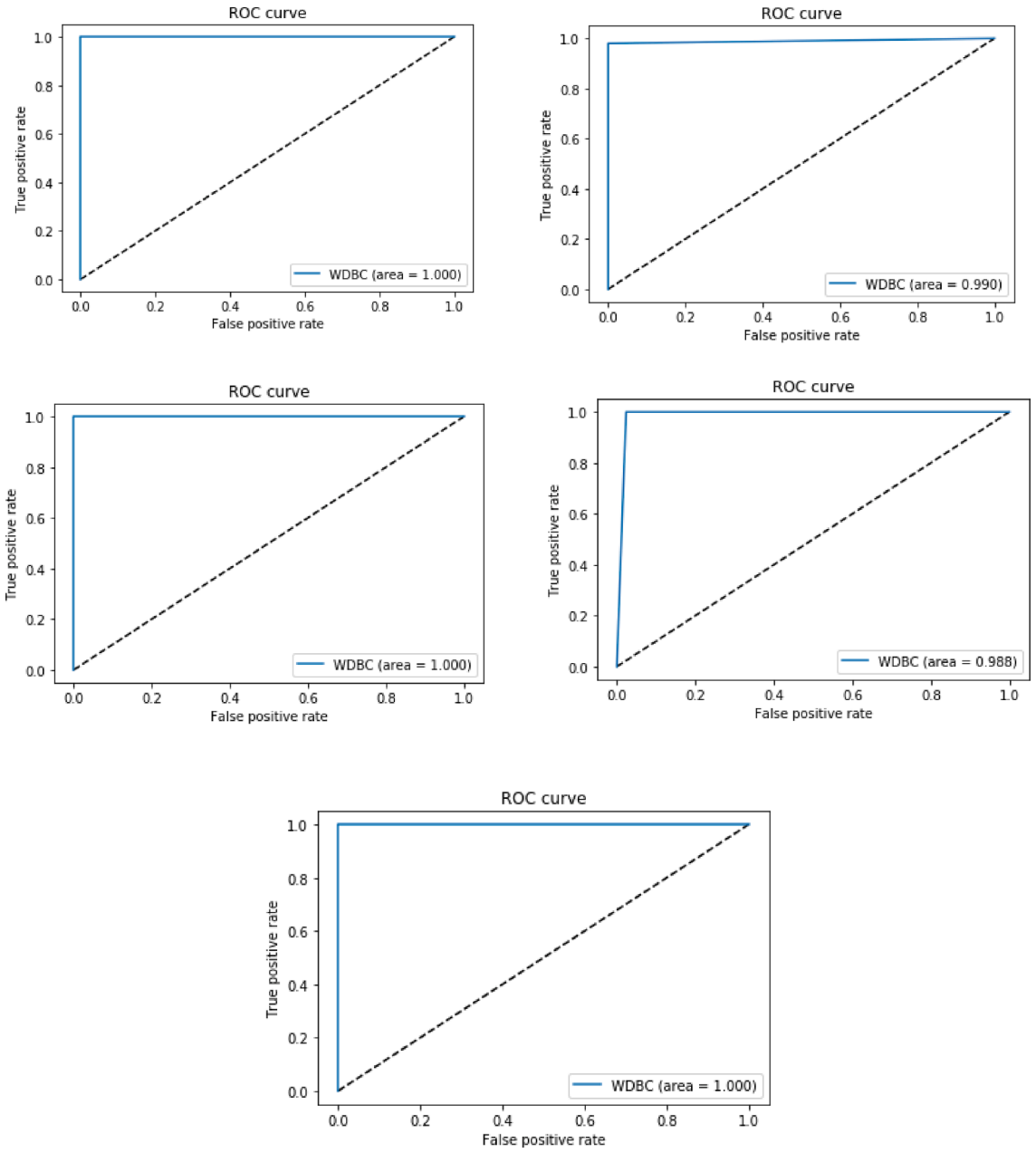

Fig. 8. ANN-WDBC ROC curves of the testing phase using five-fold cross-validation.

Table 4
ANN-WDBC evaluation parameters.

Effectiveness evaluation parameters		Efficiency evaluation parameters	
Correctly classified	**566**	Precision, % (95% CI)	**98.71 ± 2.53**
Incorrectly classified	**3**	Sensitivity, % (95% CI)	**99.59 ± 0.8**
Accuracy, % (95% CI)	**99.47 ± 0.69**	Specificity, % (95% CI)	**99.53 ± 0.92**
		F_1 score, % (95% CI)	**99.13 ± 1.27**

2. ANN-WDBC

The WDBC dataset has also been used to evaluate the proposed ANN model and support our hypothesis. This dataset incorporates 30 features for 569 cases, with no missing attributes. The same model with the same parameters has been examined to classify BC using the WDBC dataset, exhibiting exemplary performance and converging in 10.74 s, thereby proving to be highly efficient and effective, as shown in Figs. 7 and 8 as well as in Table 4. A robust average testing phase accuracy of 99.47% is recorded, with an average sensitivity of 99.59% and specificity of 99.53%. Furthermore, as shown in Fig. 8, the ROC curves demonstrate the efficiency of the ANN-WDBC system, with an average AUC of 99.56%.

As shown in Table 4, the model's effectiveness was assessed according to the number of correctly classified cases, the number of incorrectly classified cases, and the model's accuracy. The ANN-WDBC model converged swiftly to adequately train and test the WDBC data, correctly classifying 566 out of 569 instances, incorrectly classifying 3 out of 569 instances, and yielding an average accuracy of 99.47%. Furthermore, the model's efficiency was evaluated according to precision, sensitivity, specificity, and F_1-score, with 95% CI. The ANN-WDBC scores were as follows: average sensitivity = 99.59%, precision = 98.71%, specificity = 99.53%, and F_1-score = 99.13%. Thus, the ANN model achieves a robust performance, firmly demonstrating its efficiency and effectiveness.

Table 5
Comparison regarding accuracy, sensitivity, and specificity between the proposed ANN model and recent state-of-the-art models using the WDBC dataset.

ANN model	Training/test	Optimization	Overall evaluation		
			Accuracy	Sensitivity	Specificity
NN-LDA [45]	70/30	Linear discriminant analysis	98.82%	98.41%	99.07%
ANN [46]	80/20	Genetic algorithm	97.3%	98.4%	95.1%
ANN [47]	K-fold	Wrapper method	95.6%	–	–
BPNN [48]	10-fold cross-validation	Wrapper method based on ant lion optimization	98.37%	96.43%	99.52%
OGCNN [49]	10-fold cross-validation	Fitness function-based wrapper approach	93.5%	90.7%	95.23%
MLP [58]	70/30	–	99.03%	99.2%	98.7%
ANN-WDBC	**5-fold cross-validation**	–	**99.47%**	**99.59%**	**99.53%**

NN-LDA: Neural networks-linear discriminant analysis; BPNN: Back-propagation neural networks; OGCNN: One-pass generalized classifier neural network; MLP: Multilayer perceptron.

To demonstrate the proposed model's superiority, a comparison is made between the results for the ANN-WDBC and those of the most recent ANN models focused on applying the WDBC dataset for BC diagnosis in terms of accuracy, sensitivity, and specificity – as shown in Table 5. Most of the studies used for comparison applied the feature optimization concept to achieve peak accuracy, incorporating genetic algorithms, linear discriminant analysis, fitness function-based wrapper approaches, and wrapper methods, including one based on ant lion optimization algorithms. Our proposed model outperformed the reported models in terms of sensitivity (99.59%), specificity (99.53%), and accuracy (99.47%). The multilayer perceptron (MLP) [58] had a comparable performance without applying a feature reduction or an optimization algorithm; however, the model had a deep design that included three hidden layers, with 500 neurons in each hidden layer. Hence, the proposed model is considered powerful and sufficient to classify BC using the WDBC dataset without applying feature selection or optimization techniques.

This study has some limitations. Firstly, the proposed model was applied to collected and processed data. In this context, acquiring the images, applying the image processing algorithms, and extracting the features may affect the model's performance. Secondly, testing the performance of only one classifier without integrating a feature optimization approach could be restrictive as testing other classifiers like SVM may also give a promising performance. Thirdly, the study did not examine the effects of integrating feature optimization or selection algorithms.

5. Conclusion and future work

In conclusion, a shallow design of an artificial neural network (ANN) model is here proposed to diagnose breast cancer without applying feature optimization or selection algorithms. The model was tested using the Wisconsin breast cancer dataset (WBCD) and the Wisconsin diagnostic breast cancer (WDBC) dataset. The proposed computer-aided diagnosis (CAD) system generated highly accurate predictions in comparison to other state-of-the-art models that were applied to the WBCD and the WDBC dataset. With the first system, the ANN-WBCD model efficiently classified breast cancer as malignant or benign using nine features of 683 instances, resulting in an average sensitivity, specificity, precision, and accuracy of 100%, 99.72%, 99.69%, and 99.85%, respectively. The ANN-WBCD model outperformed nine ANN-based models that incorporated different feature optimization or selection algorithms. With the second system, the ANN-WDBC model classified breast cancer as malignant or benign using 30 features of 569 instances, yielding an average sensitivity, specificity, precision, and accuracy of 99.59%, 99.53%, 98.71%, and 99.47%, respectively. The ANN-WDBC model performed better than six of the models using feature optimization and selection techniques. This simple, robust, and fast model can be used to assist breast cancer diagnosis, while the proposed ANN model may also be applied to the challenge of classifying various other diseases. Future trends can target other ML models (e.g., SVM) used in CAD systems for breast cancer detection and benign or malignant tumour classification, such as segmentation-based systems that use the classical steps of segmentation, feature extraction, and selection. The results can be compared with our proposed approach using a range of metrics, including quality, accuracy, sensitivity, and specificity, and a similarity index of recognition and classification systems. The objective of such future work would be to ascertain the influence of image segmentation on building an accurate scheme in CAD systems to aid with early breast cancer detection and, thus, to save lives.

CRediT authorship contribution statement

Mohammad H. Alshayeji: Conceptualization, Methodology, Validation, Writing – original draft, Supervision, Writing - review & editing. **Hanem Ellethy:** Methodology, Software, Investigation, Writing – original draft. **Sa'ed Abed:** Methodology, Validation, Writing - review & editing. **Renu Gupta:** Conceptualization, Methodology, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by Kuwait University, Research Grant No. [EO02/18].

References

[1] P. Boix-Montesinos, P.M. Soriano-Teruel, A. Armiñán, M. Orzáez, M.J. Vicent, The past, present, and future of breast cancer models for nanomedicine development, Adv. Drug Deliv. Rev. 173 (2021) 306–330.
[2] E.H. Houssein, M.M. Emam, A.A. Ali, P.N. Suganthan, Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review, Expert Syst. Appl. (2020), 114161.
[3] U.R. Acharya, E.Y.K. Ng, J.H. Tan, S.V. Sree, Thermography based breast cancer detection using texture features and support vector machine, J. Med. Syst. 36 (3) (2012) 1503–1510, https://doi.org/10.1007/s10916-010-9611-z.
[4] Y. Chen, Q. Huang, An approach based on biclustering and neural network for classification of lesions in breast ultrasound, in: ICARM 2016–2016 International Conference on Advanced Robotics and Mechatronics, 2016, https://doi.org/10.1109/ICARM.2016.7606988.
[5] Y. Chen, L. Ling, Q. Huang, Classification of breast tumors in ultrasound using biclustering mining and neural network, in; Proceedings - 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2016, 2017, doi: 10.1109/CISP-BMEI.2016.7853007.
[6] X. Qian, et al., Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning, Nat. Biomed. Eng. (2021) 1–11.
[7] M. Pratiwi, Alexander, J. Harefa, S. Nanda, Mammograms classification using Gray-level co-occurrence matrix and radial basis function neural network, Procedia Comput. Sci. 59 (2015) 83–91, https://doi.org/10.1016/j.procs.2015.07.340.
[8] J. Dheeba, N. Albert Singh, S. Tamil Selvi, Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach, J. Biomed. Inform. 49 (2014) 45–52, https://doi.org/10.1016/j.jbi.2014.01.010.

[9] J. Dheeba, S.T. Selvi, A swarm optimized neural network system for classification of microcalcification in mammograms, J. Med. Syst. 36 (2012) 3051–3061, https://doi.org/10.1007/s10916-011-9781-3.

[10] M.A. Al-Antari, S.-M. Han, T.-S. Kim, Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms, Comput. Methods Programs Biomed. 196 (2020), 105584.

[11] F. Soares, F. Janela, M. Pereira, J. Seabra, M.M. Freire, Classification of breast masses on contrast-enhanced magnetic resonance images through log detrended fluctuation cumulant-based multifractal analysis, IEEE Syst. J. 8 (2014) 929–938, https://doi.org/10.1109/JSYST.2013.2284101.

[12] C. Gallego-Ortiz, A.L. Martel, Improving the accuracy of computer-aided diagnosis for breast MR imaging by differentiating between mass and nonmass lesions, Radiology 278 (2016) 679–688, https://doi.org/10.1148/radiol.2015150241.

[13] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, R. Monczak, Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images, Comput. Biol. Med. 43 (2013) 1563–1572, https://doi.org/10.1016/j.compbiomed.2013.08.003.

[14] P. Filipczuk, M. Kowal, A. Obuchowicz, Automatic breast cancer diagnosis based on K-means clustering and adaptive thresholding hybrid segmentation, Adv. Intell. Soft Comput. (2011), https://doi.org/10.1007/978-3-642-23154-4_33.

[15] M. Milosevic, D. Jankovic, A. Peulic, Comparative analysis of breast cancer detection in mammograms and thermograms, Biomed. Technol. (2015), https://doi.org/10.1515/bmt-2014-0047.

[16] K.F. Hollis, To Share or Not to Share: Ethical Acquisition and Use of Medical Data, AMIA Jt. Summits Transl. Sci. ProceedingsAMIA Jt. Summits Transl. Sci. (2016).

[17] O.L. Mangasarian, R. Setiono, W.H. Wolberg, Pattern recognition via linear programming: Theory and application to medical diagnosis, Large-scale Numer. Optim. (1990).

[18] UCI machine learning repository: Breast cncer Wisconsin (diagnostic) data set, 2012. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Breast%2BCancer%2BWisconsin%2B(Diagnostic). [Accessed: 13-Nov-2019].

[19] S. Wu, Y. Hu, W. Wang, X. Feng, W. Shu, Application of global optimization methods for feature selection and machine learning, Math. Probl. Eng. 2013 (2013) 1–8, https://doi.org/10.1155/2013/241517.

[20] W. Yue, Z. Wang, H. Chen, A. Payne, X. Liu, Machine learning with applications in breast cancer diagnosis and prognosis, Designs 2 (2018) 13, https://doi.org/10.3390/designs2020013.

[21] A.M. Abdel-Zaher, A.M. Eldeib, Breast cancer classification using deep belief networks, Expert Syst. Appl. 46 (2016) 139–144, https://doi.org/10.1016/j.eswa.2015.10.015.

[22] R. Setiono, Extracting rules from pruned neural networks for breast cancer diagnosis, Artif. Intell. Med. 8 (1996) 37–51, https://doi.org/10.1016/0933-3657(95)00019-4.

[23] R. Setiono, Generating concise and accurate classification rules for breast cancer diagnosis, Artif. Intell. Med. 18 (2000) 205–219, https://doi.org/10.1016/S0933-3657(99)00041-X.

[24] H.A. Abbass, An evolutionary artificial neural networks approach for breast cancer diagnosis, Artif. Intell. Med. 25 (2002) 265–281, https://doi.org/10.1016/S0933-3657(02)00028-3.

[25] M. Karabatak, M.C. Ince, An expert system for detection of breast cancer based on association rules and neural network, Expert Syst. Appl. 36 (2009) 3465–3469, https://doi.org/10.1016/j.eswa.2008.02.064.

[26] A. Marcano-Cedeño, J. Quintanilla-Domínguez, D. Andina, WBCD breast cancer database classification applying artificial metaplasticity neural network, Expert Syst. Appl. 38 (2011) 9573–9579, https://doi.org/10.1016/j.eswa.2011.01.167.

[27] H. Koyuncu, R. Ceylan, Artificial neural network based on rotation forest for biomedical pattern classification, in: 2013 36th International Conference on Telecommunications and Signal Processing, TSP 2013, 2013, doi: 10.1109/TSP.2013.6614001.

[28] A. Bhardwaj, A. Tiwari, Breast cancer diagnosis using Genetically Optimized Neural Network model, Expert Syst. Appl. 42 (2015) 4611–4620, https://doi.org/10.1016/j.eswa.2015.01.065.

[29] K.B. Nahato, K.N. Harichandran, K. Arputharaj, Knowledge mining from clinical datasets using rough sets and backpropagation neural network, Comput. Math. Methods Med. (2015) 1–13, https://doi.org/10.1155/2015/460189.

[30] N.I.R. Yassin, S. Omran, E.M.F. El Houby, H. Allam, Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review, Comput. Methods Prog. Biomed. 156 (2018) 25–45, https://doi.org/10.1016/j.cmpb.2017.12.012.

[31] M. Xian, Y. Zhang, H.D. Cheng, F. Xu, B. Zhang, J. Ding, Automatic breast ultrasound image segmentation: A survey, Pattern Recognit. 79 (2018) 340–355, https://doi.org/10.1016/j.patcog.2018.02.012.

[32] A. Jalalian, S. Mashohor, R. Mahmud, B. Karasfi, M.I.B. Saripan, A.R.B. Ramli, Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection, EXCLI J. (2017), https://doi.org/10.17179/excli2016-701.

[33] Z. Salod, Y. Singh, A five-year (2015 to 2019) analysis of studies focused on breast cancer prediction using machine learning: A systematic review and bibliometric analysis, J. Public health Res. 9 (1) (2020).

[34] M. Kumar, S.K. Khatri, M. Mohammadian, Breast cancer identification and prognosis with machine learning techniques-An elucidative review, J. Interdiscip. Math. 23 (2) (2020) 503–521.

[35] A.D. Masmoudi, N.G. Ben Ayed, D.S. Masmoudi, R. Abid, LBPV descriptors-based automatic ACR/BIRADS classification approach, EURASIP J. Image Video Process. 2013 (1) (2013) 19, https://doi.org/10.1186/1687-5281-2013-19.

[36] H. Hajiabadi, V. Babaiyan, D. Zabihzadeh, M. Hajiabadi, Combination of loss functions for robust breast cancer prediction, Comput. Electr. Eng. 84 (2020), 106624.

[37] M. Supriya, A.J. Deepa, A novel approach for breast cancer prediction using optimized ANN classifier based on big data environment, Health Care Manage. Sci. 23 (3) (2020) 414–426.

[38] N. Gargouri Ben Ayed, A. Dammak Masmoudi, D. Sellami, R. Abid, New developments in the diagnostic procedures to reduce prospective biopsies breast, in 2015 International Conference on Advances in Biomedical Engineering, ICABME 2015, 2015, doi: 10.1109/ICABME.2015.7323288.

[39] J.K. Basu, D. Bhattacharyya, T. Kim, Use of artificial neural network in pattern recognition, Int. J. Softw. Eng. its Appl. (2010).

[40] H. Asri, H. Mousannif, H.A. Moatassime, T. Noel, Using machine learning algorithms for breast cancer risk prediction and diagnosis, Procedia Comput. Sci. 83 (2016) 1064–1069, https://doi.org/10.1016/j.procs.2016.04.224.

[41] A.T. Azar, S.M. El-Metwally, Decision tree classifiers for automated medical diagnosis, Neural Comput. Appl. 23 (2013) 2387–2403, https://doi.org/10.1007/s00521-012-1196-7.

[42] O.I. Obaid, O.I. Obaid, M.A. Mohammed, M.K.A. Ghani, S.A. Mostafa, F.T. AL-Dhief, Evaluating the performance of machine learning techniques in the classification of Wisconsin breast cancer, Int. J. Eng. Technol., 7(4.36), 160–166, 2018, doi: 10.14419/ijet.v7i4.36.23737.

[43] T. Sridevi, A. Murugan, An intelligent classifier for breast cancer diagnosis based on k-means clustering and rough set, Int. J. Comput. Appl. 85 (11) (2014) 38–42, https://doi.org/10.5120/14889-3336.

[44] A. Idri, M. Hosni, I. Abnane, J.M. Carrillo de Gea, J.L. Fernández Alemán, Impact of parameter tuning on machine learning based breast cancer classification, in: Advances in Intelligent Systems and Computing, 2019, https://doi.org/10.1007/978-3-030-16187-3_12.

[45] D.A. Omondiagbe, S. Veeramani, A.S. Sidhu, Machine Learning classification techniques for breast cancer diagnosis, IOP Conf. Ser. Mater. Sci. Eng., 495(1) (2019), doi: 10.1088/1757-899X/495/1/012033.

[46] S. Aalaei, H. Shahraki, A. Rowhanimanesh, S. Eslami, Feature selection using genetic algorithm for breast cancer diagnosis: Experiment on three different datasets, Iran. J. Basic Med. Sci. (2016), https://doi.org/10.22038/ijbms.2016.6931.

[47] H. Saoud, A. Ghadi, M. Ghailani, B. A. Abdelhakim, Using Feature Selection Techniques to Improve the Accuracy of Breast Cancer Classification, no. November. Springer International Publishing, 2019.

[48] S. Dalwinder, S. Birmohan, K. Manpreet, Simultaneous feature weighting and parameter determination of Neural Networks using Ant Lion Optimization for the classification of breast cancer, Biocybern. Biomed. Eng. 40 (1) (2019) 337–351, https://doi.org/10.1016/j.bbe.2019.12.004.

[49] A.K. Naik, V. Kuppili, D.R. Edla, Efficient feature selection using one-pass generalized classifier neural network and binary bat algorithm with a novel fitness function, Soft Comput. 24 (6) (2020) 4575–4587, https://doi.org/10.1007/s00500-019-04218-6.

[50] M. Desai, M. Shah, An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and convolutional neural network (CNN), Clin. eHealth 4 (2021) 1–11.

[51] W.H. Wolberg, W.N. Street, O.L. Mangasarian, Breast cytology diagnosis with digital image analysis, Anal. Quant. Cytol. Histol. (1993).

[52] G. Dougherty, Pattern Recognition and Classification, Springer New York, New York, NY, 2013.

[53] H. Ellethy, S.S. Chandra, F.A. Nasrallah, The detection of mild traumatic brain injury in paediatrics using artificial neural networks, Comput. Biol. Med. 135 (2021) 104614, https://doi.org/10.1016/j.compbiomed.2021.104614.

[54] M.H. Alshayeji, M.A. Al-Rousan, H. Ellethy, S. Abed, An efficient multiple sclerosis segmentation and detection system using neural networks, Comput. Electr. Eng. 71 (2018) 191–205, https://doi.org/10.1016/j.compeleceng.2018.07.020.

[55] M.W. Ahmad, M. Mourshed, Y. Rezgui, Trees vs neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption, Energy Build. 147 (2017) 77–89.

[56] R.G. Pontius, B. Parmentier, Recommendations for using the relative operating characteristic (ROC), Landsc. Ecol. 29 (3) (2014) 367–382, https://doi.org/10.1007/s10980-013-9984-8.

[57] Y.T. Kiyan Tüba, Breast cancer diagnosis using statistical, J. Electr. Electron. Eng. 4 (1) (2004) 1149–1153.

[58] A.F.M. Agarap, On breast cancer detection: An application of machine learning algorithms on the Wisconsin diagnostic dataset, ACM Int. Conf. Proc. Series (2018), https://doi.org/10.1145/3184066.3184080.