Artificial neural network with Taguchi method for robust classification model to improve classification accuracy of breast cancer

Md Akizur Rahman [Corresp., 1] , **Ravie chandren Muniyandi** [1] , **Dheeb Albashish** [2] , **Md. Mokhlesur Rahman** [1] , **Opeyemi Lateef Usman** [1]

[1] Research Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

[2] Computer Science Department, Prince Abdullah Bin Ghazi Faculty of Information and Communication Technology,, Al-Balqa Applied University,, As-Salt, Salt, Jordan

Corresponding Author: Md Akizur Rahman
Email address: akiz.akizur@gmail.com

Artificial neural networks (ANN) perform well in real-world classification problems. In this paper, a robust classification model using ANN was constructed to enhance the accuracy of breast cancer classification. Taguchi method was used to determine the suitable number of neurons in a single hidden layer of the ANN. The selection of a suitable number of neurons helps to solve the overfitting problem by affecting the classification performance of an ANN. With this, a robust classification model was then built for breast cancer classification. Based on the Taguchi method results, suitable number of neurons selected for hidden layer in this study is 15, which was used for the training of the proposed ANN model. The developed model was benchmarked upon the Wisconsin Diagnosis Breast Cancer Dataset, popularly known as the UCI dataset. Finally, the proposed model was compared with seven other existing classification models, and it was confirmed that the model in this study had the best accuracy at breast cancer classification, at 98.8%. This confirmed that the proposed model significantly improved performance.

Artificial Neural Network with Taguchi Method for Robust Classification Model to Improve Classification Accuracy of Breast Cancer

Md Akizur Rahman[1*], Ravie Chandren Muniyandi[1], Dheeb Albashish[2], Md. Mokhlesur Rahman[1], and Opeyemi Lateef Usman[1]

[1] Research Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

[2] Computer Science Department, Prince Abdullah Bin Ghazi Faculty of Information and Communication Technology, Al-Balqa Applied University, Salt, Jordan

Corresponding Author:
Md Akizur Rahman[1*]
Research Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
Email address: akiz.akizur@gmail.com

Abstract

Artificial neural networks (ANN) perform well in real-world classification problems. In this paper, a robust classification model using ANN was constructed to enhance the accuracy of breast cancer classification. The Taguchi method was used to determine the suitable number of neurons in a single hidden layer of the ANN. The selection of a suitable number of neurons helps to solve the overfitting problem by affecting the classification performance of an ANN. With this, a robust classification model was then built for breast cancer classification. Based on the Taguchi method results, suitable number of neurons selected for hidden layer in this study is 15, which was used for the training of the proposed ANN model. The developed model was benchmarked upon the Wisconsin Diagnosis Breast Cancer Dataset, popularly known as the UCI dataset. Finally, the proposed model was compared to seven other existing classification models, and it was confirmed that the model in this study had the best accuracy at breast cancer classification, at 98.8%. This confirmed that the proposed model significantly improved performance.

Introduction

Despite current technological advances, medical sciences remain limited in their ability to contain and treat cancer diseases. The containment and treatment of cancer diseases form the crux of the medical science community's efforts at technological advancements. Cancer is known to be the

39 most severe complex of diseases when it comes to mortality rates, and breast cancer is the most
40 common leading cause of cancer death in women. Many women above 40 years old suffer from
41 breast cancer. It is prevalent to the point that it has been identified as the second most deadly
42 unavoidable disease for women in this age bracket (>40 years old) [1].

43 In order to properly treat breast cancer, its identification and diagnosis during the early stages
44 are crucial. The traditional approach to diagnosis is highly reliant upon the experience of the
45 attending physician(s). The reliability of physicians' experience and visual inspections is
46 questionable due to the high probability of human error. There is also an extremely large volume
47 of datasets (Big data) with poor quality and redundant information, making the diagnosis of cancer
48 at an acceptable level of accuracy a complex affair despite physicians' vast experience. In an
49 attempt to improve the accuracy of cancer classification, a computer-aided diagnostic (CAD)
50 system has been used to assist physicians [2][3][4][5][6][7][8][9][10]. The use of the CAD system
51 for classification is improving the medical diagnosis process. Classification systems can help not
52 only to minimize possible mistakes associated with a lack of experience among physicians, but
53 also to provide accurate information for the examination(s) of medical datasets [2][11].

54 Artificial neural network (ANN) models, which are inspired by the complex, interconnected
55 neural structure of the brain, have been proposed for classification tasks. In ANN, learning is
56 realized via experience, and behavior modifications are in response to environmental stimuli. The
57 model also generalizes from previous examples to address new problems. Figure 1 shows the three
58 layers of an ANN: input, hidden, and output. Two layers are in communication with the external
59 environment. The input layer receives input signals directly from the environment, which are then
60 processed by the network, while the output layer of the network delivers the processed results to
61 the (outward) environment. The number of neurons in the output layer is directly linked to a
62 particular number of tasks the neural network was designed to carry out. The intermediate layer
63 linking the input and output layers is called the hidden layer. This layer contains a function known
64 as the activation or transfer function, which performs nonlinear activation on the sum of the
65 weighted inputs from the preceding layer. It is associated with the hidden neurons, which are
66 neurons that are absent from both the input and output layers [12]. Hidden neurons result in two
67 outcomes: overfitting and underfitting. The large number of hidden neurons can be a potential
68 cause of the latter, while overfitting occurs when there are multiple unnecessary neurons present
69 in the hidden layer [12]. Underfitting occurs when the number of hidden neurons is lower than
70 what is required to model a problem dataset. The small number of neurons in the hidden layers are
71 pressured to properly detect highly complex signals [12]. A neural network architecture is dictated
72 by the number of its hidden layers, due to the fact that it is directly linked to the external
73 environment [13]. The determination of the accurate number of hidden layers is crucial for
74 increasing the accuracy of cancer classification. One method that is effective for the determination
75 of the suitable number of hidden layers is the Taguchi method [14].

76 The Taguchi method [14] has gained prominence in several research works due to its focus
77 on optimization problems. It performs well despite uncertain conditions, producing low-cost
78 outputs and robust parameter design via the integration of traditional engineering with statistics

79 for approximation and performance enhancement in multiple cases. For processes and parameter
80 arrangement, the Taguchi method utilizes the statistical investigational architecture. In a robust
81 experimental setup and design, processes or products can be investigated by altering design-related
82 factors. This experimental design helps to proficiently and consistently analyze outputs [14][15].
83   Therefore, this study proposes the Taguchi method for the parameter optimization for an ANN
84 algorithm, specifically, for the determination of the optimal number of neurons in a single hidden
85 layer. This helps to increase the accuracy of cancer classification.
86   The remainder of this paper is organized as follows: Section 2 introduces related works, while
87 Section 3 details the experimental dataset, the Taguchi method, and ANN. Section 4 discusses the
88 proposed algorithm for breast cancer classification. Section 5 presents the experimental results and
89 discussion. Finally, Section 6 concludes the work and suggests future research possibilities.
90

Related Studies

92 The most important requirements for machine learning techniques in medical diagnosis and cancer
93 classification are accuracy and reliability. This section presents related studies that were previously
94 conducted in this field, emphasizing ANN performance improvement and cancerous dataset
95 classifications. The literature offers many examples of research works employing an experimental
96 design to determine the suitable number of parameters, which could influence the performance of
97 the ANN. To efficiently establish an ANN's parameters, Khaw et al. [16] proposed the Taguchi
98 method using two sets of simulated datasets in order to increase the accuracy and convergence
99 speed of the back-propagation network (BPN). Peterson et al. [17] utilized the Taguchi method to
100 determine the causes of faults in the BPN, while Yang and Lee [18] minimized the ANN training
101 duration using the Taguchi method. Packianather et al. [19] reported the outcome of parametric
102 design on the performance of a neural network for a wood veneer examination using the Taguchi
103 method. For the purpose of designing a multilayer feed-forward neural network, Yum and Kim
104 [20] used Taguchi's active method to account for noise. Tortum et al. [21] utilized the Taguchi
105 method to determine the suitable combination of effectual parameters, and reported the
106 consequences of the performance criteria of every parameter in the neural network. Sukthomya
107 and Tannock [22] utilized the Taguchi method to determine the suitable combination of effectual
108 parameters in a neural network. Jung et al. [23] employed the Taguchi method to develop a
109 dynamic parameter design that relies on an ANN. Becherer et al. [24] used the parametric fine-
110 tuning technique with a convolutional neural network (CNN) for the purpose of image
111 classification.
112   On the topic of breast cancer diagnosis and classification, Zheng et al. [25] proposed a
113 hybridization of the k-means algorithm and support vector machine (K-SVM) for breast cancer
114 diagnosis. The results achieved 97.38% accuracy using the Wisconsin Diagnostic Breast Cancer
115 (WDBC) dataset. Orkcu et al. [26] compared the performance of a real-coded genetic algorithm,
116 back-propagation neural network (BPNN), and binary coded genetic algorithm models using the
117 breast cancer datasets, and reported accuracies of 96.50%, 93.10%, and 94.00%, respectively.
118 Salama et al. [27] used the classifiers Naive Bayes (NB), sequential minimal optimization (SMO),

119 decision tree (J48), multi-layer perception (MLP), and instance-based for k-nearest neighbor (IBK-
120 NN) for breast cancer classifications. The experiment, which adopted a confusion matrix based on
121 the 10-fold cross-validation method, using datasets from three distinct databases, showed that the
122 highest classification accuracy of 97.70% was realized using the sequential minimal optimization
123 (SMO) model.

124   On the other hand, Malmir et al. [28] used an imperialist competitive algorithm (ICA) with
125 multilayer perceptron (MLP) network and particle swarm optimization (PSO) for breast cancer
126 classification. This study achieved 97.75% and 97.63% of classification accuracies for MLP and
127 PSO, respectively. Koyuncu et al. [29] attempted to achieve higher classification accuracy using a
128 breast cancer dataset and Rotation Forest artificial neural network (RF-ANN) classifier. The result
129 from the analysis of the above classifier was 98.05% accurate. Aalaei et al. [5] compared the
130 performance of a genetic algorithm with two similar classifiers, namely a particle swarm classifier
131 and ANN, using the WDBC datasets without feature selection. The experiment reported accuracies
132 of 96.40%, 96.50%, and 96.10%, respectively.

133   Osman et al. [30] proposed Ensemble Boosting Learning (EBL) method for breast cancer
134 classification. The author used UCI different breast cancer dataset including Wisconsin Breast
135 Cancer Diagnostic (WBCD) dataset. The proposed method able to achieve the classification
136 accuracy of 97.0% using 10-fold cross-validation with Radial Based Function Neural Network
137 (RBFNN) classifier. Xue et al. [31] proposed the PSO technique with novel initialization and
138 updated mechanisms hybridized with a KNN classifier for breast cancer classification. Author
139 performed on WDBC dataset by using for training 70% and 30% for testing with 10-fold cross-
140 validation and able to achieve the classification accuracy of 92.98%. Nekkaa et al. [32] proposed
141 a classification model by using Memetic algorithm (MA) with Support vector machine (SVM) to
142 address the classification problem. The authors used particular datasets to compare certain popular
143 classifiers including WDBC dataset. The model achieved 97.85% of classification accuracy.

144   Idri et al. [33] were developed heterogeneous ensembles-based classification technique for
145 breast cancer classification. The heterogeneous ensembles included support vector machines
146 (SVM), multilayer perceptron (MLP), and decision trees (DTs) to evaluate the classification
147 performance. The author was built three groups of heterogeneous ensembles using three single
148 classifiers optimized by GS, PSO, and UC Weka. From the technique, the author performed by
149 using UCI breast cancer dataset including WDBC and achieved the highest classification accuracy
150 of 98.07% by GSVM. Jijitha et al. [34] performed on six types of different breast cancer dataset
151 including BCWD (Breast Cancer Wisconsin Diagnostic) dataset using machine learning
152 techniques. For breast cancer classification the author used K-Nearest Neighbor (K-NN) and
153 Logistic Regression (LR) technique. From BCWD dataset achieved a classification accuracy of
154 96.5% and 97.02% by LR and K-NN. Thiyagarajan et al. [35] investigated the application of
155 machine learning methods for classification of breast cancer. For an investigation of breast cancer
156 classification performance was used different machine learning methods which are included as
157 follows: ANN, Decision Tree, KNN and SVM. The experiment was completed by using WBCD
158 dataset and the highest classification accuracy of 96.2% from the ANN method. Quy et al. [36]

159 applied Machine Learning-Based Evolutionary Neural Network Approach for breast cancer
160 classification. The author focused on parameter optimize of neural network (NN) model using
161 Adaptive Particle Swarm Optimization (APSO) algorithm. For the NN used 20 neurons in a single
162 hidden layer. The WDBC dataset was used and partitioned 70% for training and 30% for testing.
163 The model was trained by Back-propagation (BP), classical PSO and APSO respectively. From
164 the experimental result achieved the highest classification accuracy of 98.24% by APSO-NN.

165   From the literature, it can be concluded that ANN is utilized since it involves pattern
166 recognition and data classification. The most important advantage of ANN with regard to the
167 classification problem in multisource databases has been solved. ANN is an established tool in
168 data classification that is easy to utilize and implement. In previous research, two limitations were
169 identified: (1) low classification accuracy [27][30][31][32], and (2) neuron selection in the hidden
170 layer [5][36]. This study proposes 15neurons in single hidden layer of ANN which can assist to
171 improve the accuracy of breast cancer classification.

172
173

Materials & Methods

175 *Experimental Dataset*
176 This research used the Wisconsin Diagnosis Breast Cancer (WDBC) dataset from the UCI
177 Machine   Learning   Repository   [available   at:
178 http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)] to differentiate
179 malignant tumors from normal tumor samples. The dataset was used to compare normal tumors
180 with cancerous (malignant) tumors [37]. The dataset contains 32 features, namely ID, diagnosis,
181 and 30 real-valued input features, followed by 569 samples, of which 357 are normal and 212 are
182 cancerous, with zero missing attribute values. The dataset features were originally computed from
183 a digitized image of a fine needle aspirate (FNA) of a breast mass by the first user and have become
184 a reference dataset for many recent studies on breast cancer classification. The attribute
185 information is as follows:
186 **Attribute Information:**
187 *1   ID number*
188 *2   Diagnosis (M = Malignant, N = Normal)*
189 *3 to 32   Ten real-valued features are computed for each cell nucleus:*
190
191 *i   Radius (mean of distances from center to points on the perimeter)*
192 *ii   Texture (Standard deviation of gray-scale values)*
193 *iii   Perimeter*
194 *iv   Area*
195 *v   Smoothness (Local variation in radius lengths)*
196 *vi   Compactness (Perimeter^2 / area - 1.0)*
197 *vii   Concavity (Severity of concave portions of the contour)*
198 *viii   Concave points (Number of concave portions of the contour)*

199   *ix   Symmetry*

200   *x   Fractal dimension ("Coastline approximation" - 1)*

201   For this study, our experimental dataset was divided into training, validation, and testing datasets

202   using four different partitions based on the Taguchi method (Followed by Equation 1 and Table

203   1). The partitions were chosen based on the experimental performance. The dataset partitions are

204   as follows: The first partition contains 50, 25, and 25; the second partition contains 60, 20, and 20;

205   the third partition contains 70, 15, and 15; and, finally, the fourth partition contains 80, 10, and 10,

206   where the first number represents the training set, the second number represents the validation set,

207   and the third number represents the testing set for each partition, respectively. Section 3.2 presents

208   details of how the Taguchi method was used. In Section 4, we discuss the performance of the

209   different dataset partitions used (Table 2).

210

211   ***Taguchi Method***

212   The Taguchi method is a robust experimental design [14] process that can be analyzed and

213   improved upon by altering the relative design factors. It is also called the statistical method and

214   can be used to realize the highest product quality. The Taguchi method utilizes a three-stage

215   method: system design, parameter design, and tolerance design. In the system design, suitable

216   working levels of design factors are accounted for. For design and testing, a system needs to be

217   based on the designers' judgement of factors such as materials, parts, nominal products, processes,

218   or parameters, based on the latest technology. The parameter design is used to determine

219   parametric levels in order to enhance the accuracy of the process being considered. Tolerance

220   design is used to fine-tune the results. As a commonly used robust design approach, the Taguchi

221   method has two mechanisms: Orthogonal array and signal-to-noise ratio (SNR) [38], for

222   improvement and analysis. To minimize experimental efforts, Orthogonal array is mainly utilized,

223   using N number of design parameters.

224   An Orthogonal array is very useful, as it gives an extensive investigation of associations

225   among all design factors and reasonably adjusts the methodical correlations of various dimensions

226   of each factor. An orthogonal array is a two-dimensional array. Each column represents a certain

227   design parameter, while each row denotes an experimental test with an actual arrangement of

228   various levels for all of the design factors. In this research, the two-level Orthogonal array for

229   determining optimal neurons is shown in Equation (1):

$$L_M(2^N) , \tag{1}$$

230   where N is the number of columns in the Orthogonal matrix. $M = 2K$ $(M > N, K > \log_2(N))$; M

231   is the number of expected experimental trials, and K is an integer. Base 2 is the number of levels

232   of every design parameter. In this study, there are 20 numbers of columns in the Orthogonal matrix

233   (where F1 to F20 indicate the design factor the number of neurons in the hidden layer), and a two-

234   level Orthogonal array was used for selecting appropriate number of neurons. The two-level

235   Orthogonal array was created using $L_{21}(2^{20})$, as shown in Table 1.

236   Only 21 experimental trials are required for evaluation, analysis, and improvement.

237   Conversely, all possible combinations of 20 design factors (i.e., $2^{20} = 1,048,576$) should be

238 accounted for in the full factorial experimental design, which is frequently inapplicable in practice
239 [39] (for neuron selection: 1: selected, 0: not selected).
240 This research used an Orthogonal array mechanism to analyze and enhance the ANN algorithm
241 performance by determining the optimal number of neurons in hidden layer. If a particular target
242 has N different design factors, 2^N possible experimental trials will be considered in the full factorial
243 experimental design.
244
245   **Table 1.** Orthogonal array (OA).
246
247 ***Artificial Neural Network:***
248 An artificial neural network (ANN) [11] is a machine learning approach, which models the human
249 brain with a number of artificial neurons and interconnected associations. The neurons in ANNs
250 tend to have fewer connections relative to a biological neural system. Figure 1 illustrates the basics
251 of the ANN architecture.
252
253   **Figure 1.** The basic architecture of an ANN.
254 Neurons are highly interconnected computational units inspired by the mammalian brain. The
255 ANN system consists of the smallest processing nodes called neurons, or processing elements. To
256 obtain outputs, a function called Sigmoid activation works on inputs and connects the weights with
257 the help of neurons. The connections between weight (w) values and single nodes are called biases
258 (b). The iterative flow of training data determines weight values throughout the network. During
259 the training phase, weight values are verified until the network is able to detect a particular cluster
260 using criteria for typical input data. Also, these weights assign the link relating one layer of neurons
261 to another. Hence, changes in the relationship between input and output take place repeatedly,
262 along with changes in weight values. The method of balancing the links' weight values by
263 repeatedly exposing the network to the input–output dataset for learning is called training.
264
265   **Figure 2.** The architecture of a Multi-layer feed-forward neural network with four inputs.
266 ANN has various types of architectures. Multi-layer feed-forward neural networks have been
267 widely used in cancer classification [40][41][42]. The architecture of a Multi-layer feed-forward
268 neural network system is shown in Figure 2, consisting of one or more input layer and one or more
269 hidden layer along with an output layer.
270

The Proposed Artificial Neural Network Model with 15 Neurons in a Single Hidden Layer for Breast Cancer Classification

273 In this section, an ANN model with 15 neurons in a single hidden layer is presented for breast
274 cancer classification. The ANN model, which applied the Taguchi method [14] to determine the
275 optimal number of neurons in the hidden layer for breast cancer classification, is called the
276 Improved-ANN (IANN). The design of the proposed IANN model consists of six phases: data
277 collection, data preprocessing, data loading using MATLAB simulation tool, data partitioning,

278 ANN training and validation, and the application of the Taguchi method to determine the optimal
279 number of neurons in the hidden layer of the ANN.
280   The first step is data collection. It involves the collection of the appropriate dataset from the
281 data source and center for the purpose of research. In the second step (data preprocessing), all of
282 the data features were prepared and filtered to remove noise in order to enhance the quality of the
283 features that will be selected for the classification exercise. In the third stage, the Neural Network
284 Pattern Recognition Tool of MATLAB (R2017b) simulation software [43] was used to load the
285 dataset. Then, the dataset was divided into three categories: a training dataset, a validation dataset,
286 and a testing dataset. A two-layer feed-forward neural network, Gradient Descent withSigmoid
287 activation function and Softmax Output Neurons (patternnet), was used to classify the dataset with
288 adequate neurons in the hidden layer. Multi-layer feed-forward neural network widely used for
289 classification problem and it's able to achieved high classification performance. The most
290 algorithm to train neural network is used gradient descent. Gradient descent is a way to minimize
291 an objective function which assist to adapt the learning rate. The sigmoid activation function is a
292 nonlinear function which can help the network learn complex data in hidden layer. The softmax
293 function can able to handle multiple classes output. Basically, Softmax function is used for output
294 layer. The number of neurons for the experiments was at least 10 (default value) because below
295 that the performance was too low. It could be seen that, in the case of each training phase, the
296 performance of the model increased gradually and peaked when the number of neurons in the
297 hidden layer hit 15. Further increasing the number of hidden layer neurons, from 16 to 20 neurons,
298 the performance was observed to decrease. The number of neurons determined using the Taguchi
299 method is detailed in Section 5.1 and Table 2. By this method, the number of hidden layer neurons
300 was selected randomly and repeatedly until excellent accuracy was obtained for the validation set.
301 The excellent performance of the validation set in this study was achieved via 15 neurons in the
302 hidden layer, which was sufficient to construct a suitable ANN model for breast cancer
303 classification when the number of output layer neurons was two, as shown in Figure 3.
304
305 **Figure 3.** Architecture of the proposed 15 neurons ANN classifier and two output layers.
306
307   **Figure 4.** The proposed Improved-ANN (IANN) cancer classification model.
308 Understanding the classification of breast cancer forms the first step in its containment and
309 eradication worldwide. Cancer researchers have utilized various machine learning algorithms for
310 cancer classification. The most significant benefit of ANN for the classification of the problem in
311 multisource databases has been solved [45][46][47]. This study employs an ANN with 15 neurons
312 in hidden layer for cancer classification. Figure 4 shows the conceptual framework for breast
313 cancer classification using an ANN model with 15 neurons in its hidden layer.
314

315 ## Experimental Results and Discussion
316 *Experimental Setup and Evolution Methods*

317 In this research, the experiment was conducted by using MATLAB Tool, WDBC dataset, Taguchi
318 Method and ANN Algorithm, implemented on a system with the following configuration: Core i7
319 GPU, Windows 10 with 8GB RAM and 1TB HDD. The proposed ANN model comprises a single
320 hidden layer with 15 neurons carefully selected using Taguchi method. Accuracy of the proposed
321 model was obtained directly from the confusion matrix using the formula:

322 $$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100$$

323

324 ***ANN Parameter Optimization Based on the Taguchi Method***
325 To build an appropriate classification model using ANN, parameter selection is one of the most
326 significant steps. The performance and stability of an ANN is dictated by its selected parameters.
327 The Taguchi method is mainly used for experimental parametric selection, owing to the concept
328 of engineering and technology. For instance, this method can decrease the vast number of
329 experiments and simultaneously analyze several parameters. For excellent performance and low-
330 cost computations, this method performs well in the context of systematic and efficient designs
331 [47]. A different partition for a dataset can be created to simplify the selection of a suitable dataset.
332   As per Section 3.1, the L_{21} Orthogonal array was used to determine the appropriate number of
333 neurons in the hidden layer of the proposed IANN model. The number of neurons for the
334 experiments ranges from 10 (the default value) to 20, as shown in Table 2. Also, different partitions
335 of the dataset were utilized in order to determine a suitable partition for the breast cancer dataset.
336 For the different partitions of the dataset, Taguchi Orthogonal array mechanism was implemented
337 in the same way, where N=3 is the number of design factors of dataset partitions. Therefore, the
338 L_4 Orthogonal array was used for data partition (for more details, see section 3.1, equation 1). The
339 dataset was randomly divided into three partitions: training, validation, and testing. These
340 partitions were utilized in different groups: first, 50%, 25%, and 25%; second, 60%, 20%, and
341 20%; third, 70%, 15%, and 15%; and fourth, 80%, 10%, and 10%.

342   **Table 2.** The Taguchi method for ANN to select the suitable number of neurons for
343   breast cancer datasets.

344
345 It can be observed in Table 2 that the hidden layer parameter is crucial to the performance of the
346 ANN. As established earlier, the experiments were conducted in four phases using different dataset
347 partitions for each phase. For the first experiment, we partitioned the entire dataset into 50%
348 training, 25% validation, and the remaining 25% for testing. Based on the three partitions used,
349 the highest accuracy for the validation dataset was realized when the number of neurons in a single
350 hidden layer is 15. In this experiment, the accuracy was found to be 93.90% for the validation
351 dataset and 93.50% for the training dataset. However, when the number of neurons increased, the
352 performance diminished significantly. This seemingly implied that when the number of neurons
353 in the hidden layer increases further, the performance of the ANN model is negatively affected.
354   In the second experiment, when the partitioning of the dataset was changed to 60%, 20%, and
355 20% for training, validation, and testing, respectively, the Taguchi method confirmed that the
356 highest performance was maintained at 15 neurons in a single hidden layer. It is therefore clear
357 from Table II that utilizing the aforementioned partitions significantly improves the performance
358 of IANN model by ~3.0%.

359   In the third experiment, a new partition for the dataset was introduced, with 70% data for
360   training, 15% for validation, and 15% for testing. In this scenario, the highest accuracy for the
361   validation dataset was determined to be 98.30%, and that of the training dataset was 98.50% when
362   the number of neurons in a single hidden layer was 15 using the Taguchi method.
363   Finally, in the fourth experiment, the partitioning of the dataset was 80%, 10%, and 10% for
364   training, validation, and testing, respectively. This study confirmed that the results declined from
365   98.30% for the validation dataset to 93.80% with 4.5% when compared with the third experiment.
366   All of the experiments proved that, 15 neurons are sufficient for modeling a single hidden layer in
367   the proposed IANN for breast cancer classification.
368   From the empirical results, this study posits that the optimum conditions for breast cancer
369   classification are 15 neurons in a single hidden layer and the partitioning of the data into 70%,
370   15%, and 15% for training, validation, and testing sets, respectively. Thus, we built an IANN
371   model with three layers: input, hidden, and output. In summary, there were 31 input neurons at the
372   input layer and 2 output neurons at the output layer, while the number of neurons in a single hidden
373   layer was 15.

374   ***Breast Cancer Classification Results***

375   As established in Section 4, the suitable number of neurons for a single hidden layer based on the
376   Taguchi is 15. In addition, 70% of the dataset for training, 15% for validation, and 15% for testing
377   is the optimal partitioning. The experiments were simulated on MATLAB (R2017b) software
378   using Neural Network Toolbox for the implementation of the proposed model. The proposed
379   IANN model has been tested using the WDBC dataset.
380   Table 3 shows the results from 30 different classification simulations using the WDBC dataset
381   with their corresponding percentage errors. The average of the tests' accuracy was 98.80%. This
382   study confirmed that the results of the simulations were relatively similar, which could be due to
383   the high stability resulting from the selection of suitable parameters for the IANN classifier.

384   **Table 3**. Classification accuracy of the proposed method based on 30 different runs.

385

386   This study used 569-sample datasets. Each data sample consists of 32 features, with a total of
387   18,208 data points distributed as follows: 11,424 data points for normal and 6784 data points for
388   cancerous. These features were used to train and simulate the proposed IANN model. A hidden-
389   layer simple feed-forward neural network architecture is considered because our aim is to enhance
390   the classification of breast cancer through optimal neuron selection in the hidden layer. The
391   proposed IANN was tested with training, validation, and testing sets from the dataset. The network
392   was adjusted based on the reported error during testing. Validation was used to simplify the
393   network and halt training. Testing is ineffective in training; therefore, the performance of the
394   network provides independent measures during and after training. A breast cancer dataset
395   consisting of 569 samples was divided randomly into two groups: 399 samples (70%) for training
396   and 170 samples (30%) for testing. The dataset for training was randomly divided into three

397 groups: out of 399 samples 279 (70%) samples for training, and 60 (15%) samples each for testing
398 and validation respectively.
399   Figures 5–11 show the IANN training performance, training state performance, error
400 histogram, performance confusion matrix, ROC curve and testing dataset results for breast cancer,
401 respectively.
402
403   **Figure 5.** IANN training performance.

404 It can be seen in Figure 5 that the IANN training performance plot at the beginning of the training
405 of cross-entropy resulted in the maximum error. The proposed system reported the best
406 performance at epoch 20 iterations and an exact cross-entropy of 0.031613.
407   Figure 6 shows the network training state performance at epoch 26, when the gradient is
408 0.032875. The network halts the training session because its generalization stops improving.
409   Error histograms for training, validation, and testing data are shown with 20 bins in Figure 7.
410 The experimental results shown in error histogram of Figure 7 indicate that, the proposed system
411 can handle the dataset used successfully, since the error is close to zero.
412
413   **Figure 6.** Training state performance.

414
415   **Figure 7.** Error histogram.

416 To evaluate the accuracy, a confusion matrix was used for all partitions of the dataset. A confusion
417 matrix is a two-dimensional array, $r \times r$ (where r is the number of classes). Figure 8 shows the
418 MATLAB output representation of training dataset confusion matrix performance after 10-fold
419 cross-validation. In Figure 8, the first two rows and first two columns represent the actual
420 confusion matrix. The third row and third column are the summary of percentage accuracy, and
421 sensitivity and specificity, respectively. Row one, column one is the true positives (TP); row one,
422 column two is the false negatives (FN); row two, column one is the false positives (FP); and row
423 two, column two is the true negatives, respectively. All false positives (FP) are Type I errors, while
424 all false negative (FN) are Type II errors generated by different classifiers.
425   The first two diagonal cells of the confusion matrix demonstrate the correct classification
426 number and percentage accuracy of the trained network. The normal biopsy results are 247, which
427 were correctly classified to represent 61.90% of the total 399 biopsies. On the other hand, the
428 cancerous (malignant) biopsy results show 146 correctly classified tumors, representing 36.60%
429 of the total biopsies. The experiment also revealed that three of the cancerous biopsies were
430 incorrectly classified as normal, representing 0.8% of the total biopsy dataset, while three of the
431 normal biopsies were incorrectly classified as cancerous, also representing 0.8% of the total biopsy
432 data. The total result of 250 normal revealed that 98.80% was correct, and 1.20% incorrect. The
433 total result of 149 cancerous was 98.00% correct and 2.00% incorrect. The total result of 250
434 normal cases was 98.80% correctly classified as normal, and 1.2% classified as cancerous. Out of

435   149 cancerous cases, 98.00% were correctly classified as cancer, and 2.00% classified as normal.
436   The confusion matrix plots with 98.50% accuracy show that this system performed well and had
437   1.40% misclassification during its training stage from the proposed IANN.
438

439   **Figure 8.** Performance confusion matrix.

440

441   **Figure 9.** Receiver operating characteristic (ROC) curve.

442   The neural network training performance with receiver operating characteristic (ROC) plot is
443   shown in Figure 9. The ROC plot represents the performance of the binary classification system
444   when the discrimination threshold fluctuates. The graph is formed by plotting the true positive rate
445   (TPR) against the false positive rate (FPR). From ROC plot, it can be seen that the NN's
446   performance increases in response to the number of iterations. A perfect classification result was
447   evident at 26 iterations. This shows that every class achieved perfect classification accuracy.
448   Iteration 26 is the optimal iteration for the proposed model where the IANN performed at its peak.
449   After completing the training session, we tested the network for accurate classification, using
450   30% of the test dataset of breast cancer. Figure 10 shows the testing accuracy, otherwise known as
451   the classification accuracy from the proposed IANN, to be rather high.
452

453   **Figure 10.** Test dataset result of breast cancer.

454

455   **Figure 11.** ROC curve with test dataset of breast cancer.

456   Figure 10 shows the experimental results with a test dataset of breast cancer: normal versus
457   malignant tumor is correctly classified at 98.8% accuracy for the test case.
458   This study computed the receiver operating curve (ROC) and the results are shown in Figure
459   11. The ROC curve indicates TPR and FPR at different edge settings of the network, with better
460   results arising from the proposed system. The IANN, after training, validation, and testing,
461   achieved 98.8% correct classification for two classes: normal and cancerous (malignant). The area
462   under the curve was large.
463   Table 4 shows the model performance obtained from the confusion matrix shown in Figures
464   8 and 10. In Table 4, the F-Score for the training dataset is 98.4%, while the F-Score for the testing
465   dataset is 98.8%.

466   **Table 4.** Model performance based on confusion matrix after 10-fold cross-validation.

467   ***Discussion of Findings and Implications***

468   Table 5 shows the performance of multiple machine learning algorithms relative to the proposed
469   IANN model. Performance comparison was realized via MATLAB Classification learner tool-

470 staking for the breast cancer dataset. Although the results obtained with other machine learning
471 algorithms are equally good, the result reported by our model is optimal. It can be surmised that
472 the proposed IANN outperforms the existing machine learning models in classifying breast cancer.

473   **Table 5**. Performance comparison between different classifiers and the proposed
474   method in breast cancer classification.

475   **Table 6**. Comparison of breast cancer classification performance between existing
476   methods in the literature and the proposed method.

477
478 Table 6 indicates the superiority of this study in the context of cancer classification based on the
479 same dataset but different methods for classification. Among the previous studies tabulated in
480 Table 6, a lower classification accuracy of 92.98% was achieved by Xue et al. [31]. These authors
481 proposed the PSO technique with novel initialization and updated mechanisms hybridized with a
482 KNN classifier and 10-fold cross-validation to maximize the classification performance of the
483 WDBC datasets and dataset were partitioned for training 70% and 30% for testing. Quy et al. [36]
484 achieved the best classification accuracy of 98.24% using APSO-NN. For the NN used 20 neurons
485 in a single hidden layer. The WDBC dataset was used and partitioned 70% for training and 30%
486 for testing. Abdar et al. [49] used the WDBC dataset for the experiment and achieved a
487 classification accuracy of 98.07% using SV-NB-3-MetaClassifiers with a K-fold cross-validation
488 technique. The performance is quite good. Nekkaa et al. [32] used a memetic algorithm (MA) with
489 support vector machine (SVM) to address the classification problem. The authors used particular
490 datasets to compare certain popular classifiers for the data classification task. The model achieved
491 97.85% accuracy using WDBC datasets. Beside, Salama et al. [27] achieved a classification
492 accuracy of 97.70% using sequential minimal optimization (SMO) technique with a confusion
493 matrix based on the 10-fold cross-validation method. Amrane et al. [48] used the KNN and Naive
494 Bayes (NB) classifier machine learning technique for breast cancer classification. The author
495 partitioned the dataset 60% for training and 40% for tesing. From KNN classifier achieved the
496 highest classification accuracy of 97.51% which is not so high. Aalaei et al. [5] used a GA-based
497 classifier with ANN and reported a classification accuracy of 96.50% by making use of a single
498 hidden layer with 5-neurons and 2-neurons in output layer, and the dataset partitioned of 80% for
499 training and 20% for test data.
500 From the analysis of the results presented in Tables 5 and 6, it is obvious that the proposed method
501 outperformed the other methods due to it realizing a classification accuracy of 98.80% —an
502 improvement of 0.56% compared to Quy et al. [36], which has the best performance among the
503 existing methods. This could be due to the use of the Taguchi method for selecting the optimal
504 number of neurons in a single hidden layer of ANN and suitable percentage data partition. Our
505 results also show that the proposed method is more stable and reliable than existing classification
506 models. By implication, incorporating IANN into breast cancer prediction could enable timely and
507 accurate prediction of breast cancer, thereby helping medical practitioners to make the most
508 appropriate decisions on breast cancer treatment. Table 7 (where K=10) shows the 10-fold cross

509 validation experimental performance which was conducted to validate the results shown in Table
510 5 and 6.

511 **Table 7.** 10-fold cross-validation experimental performance.

Conclusions

513 In breast cancer diagnosis, differentiating between normal and malignant tumors is one of the
514 challenges faced by physicians. In order to tackle this problem, ANN is incorporated in the CAD
515 system for the binary classification of breast cancer datasets into normal/malignant tumors. The
516 main purpose of this research is to enhance the classification accuracy to improve image-based
517 diagnosis. The proposed IANN utilized the Taguchi method to optimize the hidden layer parameter
518 of the ANN model. The experimental results proved that 15 is the optimal number of neurons for
519 a single hidden layer of ANN, and can enhance the classification accuracy for the training, testing,
520 and validation of selected datasets.

521 In this paper, the proposed IANN method reported 98.80% classification accuracy for breast
522 cancer dataset classification. Additionally, 10-folds cross validation experiments performed
523 (shown in Table 7) also confirmed the efficiency of the proposed IANN model with an average
524 accuracy of 98.7% on testing dataset. The empirical results, presented in both tabular and graphical
525 form, proved that the proposed IANN greatly enhanced the overall classification performance by
526 differentiating between normal and malignant tumors for superior breast cancer diagnosis. This
527 study confirmed that the use of the Taguchi method with ANN improved classification
528 performance relative to existing breast cancer classification methods. While the results are
529 promising, future research can focus on feature selection to reduce the computational discrepancy.
530 Also, future studies should focus on the practical implications of Type I (FP) and Type II (FN)
531 error in breast cancer classification. Meanwhile, an express medical test for breast cancer
532 classification has often been designed in such a way that Type I error is given precedence. This is
533 done to prevent a breast cancer patient from going about without the knowledge of having a
534 malignant tumor.

Acknowledgments

536 The grant for this research (Universiti Kebangsaan Malaysia (UKM), UKM Grant Code: GGP-
537 2019-023) is acknowledged, as this support played a vital role in the successful accomplishment
538 of the research.

539

References

541 1. "Breast Cancer - Symptoms & Treatment Resources - Imaginis." [Online]. Available:
542   http://www.imaginis.com/breast-cancer-resource-center. [Accessed: 24-Apr-2019].
543 2. S. Sahran, D. Albashish, A. Abdullah, N. A. Shukor, and S. Hayati Md Pauzi, "Absolute
544   cosine-based SVM-RFE feature selection method for prostate histopathological grading,"
545   Artif. Intell. Med., Apr. 2018.

546   3.   D. Albashish, S. Sahran, A. Abdullah, N. AbdShukor, and S. Hayati Md Pauz, "Ensemble
547   Learning of Tissue Components for Prostate Histopathology Image Grading," Eng. Inf.
548   Technol. (IJASEIT)., vol. 6, no. 6, pp. 1132–1138, 2016.

549   4.   M. Xi, J. Sun, L. Liu, F. Fan, and X. Wu, "Cancer Feature Selection and Classification Using
550   a Binary Quantum-Behaved Particle Swarm Optimization and Support Vector Machine,"
551   Comput. Math. Methods Med., vol. 2016, pp. 1–9, 2016.

552   5.   S. Aalaei, H. Shahraki, A. Rowhanimanesh, and S. Eslami, "Feature selection using genetic
553   algorithm for breast cancer diagnosis: experiment on three different datasets.," Iran. J. Basic
554   Med. Sci., vol. 19, no. 5, pp. 476–82, May 2016.

555   6.   M. A. Rahman and R. C. Muniyandi, "Review of GPU implementation to process of RNA
556   sequence on cancer," Informatics Med. Unlocked, vol. 10, pp. 17–26, Jan. 2018.

557   7.   D. Tomar and S. Agarwal, "Hybrid feature selection based weighted least squares twin support
558   vector machine approach for diagnosing breast cancer, hepatitis, and diabetes," Adv. Artif.
559   Neural Syst., vol. 2015, no. article.1, 2015.

560   8.   N. Elkhani and R. C. Muniyandi, "Membrane computing inspired feature selection model for
561   microarray cancer data," Intell. Data Anal., vol. 21, no. S1, pp. S137–S157, Apr. 2017.

562   9.   J. Dheeba, N. Albert Singh, and S. Tamil Selvi, "Computer-aided detection of breast cancer
563   on mammograms: A swarm intelligence optimized wavelet neural network approach," J.
564   Biomed. Inform., vol. 49, pp. 45–52, Jun. 2014.

565   10.   Jafari-Marandi, Ruholla, Samaneh Davarzani, Maryam Soltanpour Gharibdousti, and Brian
566   K. Smith. "An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN
567   learning and decision-making goals." *Applied Soft Computing* 72 (2018): 108-120.

568   11.   M. A. Rahman and R. C. Muniyandi, "Feature selection from colon cancer dataset for cancer
569   classification using Artificial Neural Network," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 8, no.
570   4–2, p. 1387, Sep. 2018.

571   12.   S. Karsoliya, "Approximating number of hidden layer neurons in multiple hidden layer BPNN
572   architecture," Int. J. Eng. Trends Technol., vol. 3, no. 6, pp. 714–717, 2012.

573   13.   G. Panchal, A. Ganatra, … Y. K.-I. J. of, and U. 2011, "Behaviour analysis of multilayer
574   perceptronswith multiple hidden neurons and hidden layers," Int. J. Comput. Theory Eng.,
575   vol. 3, no. 2, pp. 332–337, 2011.

576   14.   Y. Wu and A. Wu, Taguchi Methods for robust design. ASME Press, 2000.

577   15.   N. S. Jaddi, S. Abdullah, and A. R. Hamdan, "Taguchi-Based Parameter Designing of Genetic
578   Algorithm for Artificial Neural Network Training," in 2013 International Conference on
579   Informatics and Creative Multimedia, 2013, pp. 278–281.

580   16.   J. F. C. Khaw, B. S. Lim, and L. E. N. Lim, "Optimal design of neural networks using the
581   Taguchi method," Neurocomputing, vol. 7, no. 3, pp. 225–245, Apr. 1995.

582   17.   G. E. Peterson, D. C. St. Clair, S. R. Aylward, and W. E. Bond, "Using Taguchi's method of
583   experimental design to control errors in layered perceptrons," IEEE Trans. Neural Networks,
584   vol. 6, no. 4, pp. 949–961, Jul. 1995.

585   18.   S. M. Yang and G. S. Lee, "Neural Network Design by Using Taguchi Method," J. Dyn. Syst.
586   Meas. Control, vol. 121, no. 3, p. 560, Sep. 1999.

587   19.   M. S. Packianather, P. R. Drake, and H. Rowlands, "Optimizing the parameters of
588   multilayered feedforward neural networks through Taguchi design of experiments," Qual.
589   Reliab. Eng. Int., vol. 16, no. 6, pp. 461–473, Nov. 2000.

590   20.   Y.-S. Kim and B.-J. Yum, "Robust design of multilayer feedforward neural networks: an
591   experimental approach," Eng. Appl. Artif. Intell., vol. 17, no. 3, pp. 249–263, Apr. 2004.

592   21. A. Tortum, N. Yayla, C. Çelik, and M. Gökdağ, "The investigation of model selection criteria
593   in artificial neural networks by the Taguchi method," Phys. A Stat. Mech. its Appl., vol. 386,
594   no. 1, pp. 446–468, Dec. 2007.
595   22. M. Bashiri and A. Farshbaf Geranmayeh, "Tuning the parameters of an artificial neural
596   network using central composite design and genetic algorithm," Sci. Iran., vol. 18, no. 6, pp.
597   1600–1608, Dec. 2011.
598   23. J.-R. Jung and B.-J. Yum, "Artificial neural network-based approach for dynamic parameter
599   design," Expert Syst. Appl., vol. 38, no. 1, pp. 504–510, Jan. 2011.
600   24. N. Becherer, J. Pecarina, S. Nykl, and K. Hopkinson, "Improving optimization of
601   convolutional neural networks through parameter fine-tuning," Neural Comput. Appl., pp. 1–
602   11, Nov. 2017.
603   25. B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction
604   using a hybrid of K-means and support vector machine algorithms," Expert Syst. Appl., vol.
605   41, no. 4, pp. 1476–1482, Mar. 2014.
606   26. H. H. Örkcü and H. Bal, "Comparing performances of backpropagation and genetic
607   algorithms in the data classification," Expert Syst. Appl., vol. 38, no. 4, pp. 3703–3709, Apr.
608   2011.
609   27. G. Salama and M. Z. Abdelhalim, M, "Breast cancer diagnosis on three different datasets
610   using multi-classifiers," Int. J. Comput. Inf. Technol., vol. 1, no. 1, 2012.
611   28. H. Malmir, F. Farokhi, and R. Sabbaghi-Nadooshan, "Optimization of data mining with
612   evolutionary algorithms for cloud computing application," in ICCKE 2013, 2013, pp. 343–
613   347.
614   29. H. Koyuncu and R. Ceylan, "Artificial neural network based on rotation forest for biomedical
615   pattern classification," in 2013 36th International Conference on Telecommunications and
616   Signal Processing (TSP), 2013, pp. 581–585.
617   30. V. Chaurasia and S. Pal, "A Novel Approach for Breast Cancer Detection Using Data Mining
618   Techniques," Int. J. Innov. Res. Comput. Commun. Eng., vol. 2, no. 1, p. pp.17, Jun. 2017.
619   31. B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimisation for feature selection in
620   classification: Novel initialisation and updating mechanisms," Appl. Soft Comput., vol. 18,
621   pp. 261–276, May 2014.
622   32. M. Nekkaa and D. Boughaci, "A memetic algorithm with support vector machine for feature
623   selection and classification," Memetic Comput., vol. 7, no. 1, pp. 59–73, Mar. 2015.
624   33. Idri, Ali, Mohamed Hosni, and Ibtissam Abnane. "Assessing the impact of parameters tuning
625   in ensemble based breast Cancer classification." *Health and Technology* 10, no. 5 (2020):
626   1239-1255.
627   34. Jijitha, S., and Thangavel Amudha. "Breast Cancer Prognosis Using Machine Learning
628   Techniques and Genetic Algorithm: Experiment on Six Different Datasets." In *Evolutionary
629   Computing and Mobile Sustainable Networks*, pp. 703-711. Springer, Singapore. 01 August
630   2020.
631   35. Thiyagarajan,S., T. Chakravarthy, and PV Arivoli. "Diagnosing Breast Cancer with Machine
632   Learning Algorithms".Int J Life Sci Pharma Res. ISSN 2250 – 0480; SP-07; *23rd January
633   2020*: 42. "International Conference on Cancer Research" 2020**.**
634   36. Quy, Hoang Duc, Cao Van Kien, Ho Pham Huy Anh, and Nguyen Ngoc Son. "Machine
635   Learning-Based Evolutionary Neural Network Approach Applied in Breast Cancer Tumor
636   Classification." In *International Conference on Green Technology and Sustainable
637   Development*, pp. 72-83. Springer, Cham, 2020.

638 37. M. Pobiruchin, S. Bochum, U. M. Martens, M. Kieser, and W. Schramm, "A method for using
639   real world data in breast cancer modeling," J. Biomed. Inform., vol. 60, pp. 385–394, Apr.
640   2016
641 38. L.-Y. Chuang, C.-S. Yang, K.-C. Wu, and C.-H. Yang, "Correlation-based Gene Selection
642   and Classification Using Taguchi-BPSO," Methods Inf. Med., vol. 49, no. 3, pp. 254–268,
643   Jan. 2010.
644 39. C.-H. Yang, C.-C. Huang, K.-C. Wu, and H.-Y. Chang, "A Novel GA-Taguchi-Based Feature
645   Selection Method," Springer, Berlin, Heidelberg, 2008, pp. 112–119.
646 40. O. Alesawy and R. C. Muniyandi, "Elliptic Curve Diffie-Hellman Random Keys Using
647   Artificial Neural Network and Genetic Algorithm for Secure Data over Private Cloud," Inf.
648   Technol. J., vol. 15, no. 3, pp. 77–83, Jun. 2016.
649 41. J. Hopfield, "Artificial neural networks," IEEE Circuits Devices Mag., vol. 4, no. 5, 1988.
650 42. G. Bebis and M. Georgiopoulos, "Feed-forward neural networks," IEEE Potentials, vol. 13,
651   no. 4, 1994.
652 43. The Math Works Inc, "MATLAB 2017b," Natick, Massachusetts United States.
653 44. B. Cai and X. Jiang, "A novel artificial neural network method for biomedical prediction based
654   on matrix pseudo-inversion," J. Biomed. Inform., vol. 48, pp. 114–121, Apr. 2014.
655 45. M. Wadhonkar, P. Tijare, S. S.- Integration, and U. 2014, "Artificial Neural Network
656   Approach for Classification of Heart Disease Dataset," ijaiem.org, vol. 3, no. 4, pp. 388–392,
657   2014.
658 46. Rahman, Md Akizur, and Ravie Chandren Muniyandi. "An Enhancement in Cancer
659   Classification Accuracy Using a Two-Step Feature Selection Method Based on Artificial
660   Neural Networks with 15 Neurons." Symmetry 12, no. 2 (2020): 271.
661 47. M.-L. Huang, Y.-H. Hung, W. M. Lee, R. K. Li, and B.-R. Jiang, "SVM-RFE based feature
662   selection and Taguchi parameters optimization for multiclass SVM classifier,"
663   ScientificWorldJournal., vol. 2014, p. 795624, Jan. 2014.
664 48. M. Amrane, S. Oukid, I. Gagaoua, and T. Ensari, "Breast cancer classification using machine
665   learning," in 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting
666   (EBBT), 2018, pp. 1–4.
667 49. Abdar, Moloud, Mariam Zomorodi-Moghadam, Xujuan Zhou, Raj Gururajan, Xiaohui Tao,
668   Prabal D. Barua, and Rashmi Gururajan. "A new nested ensemble technique for automated
669   diagnosis of breast cancer." Pattern Recognition Letters 132 (2020): 123-131.

Table I. Orthogonal array (OA)

Table I. Orthogonal array (OA)

RUN	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Table II. The Taguchi method for ANN to select the suitable number of neurons for breast cancer datasets

Table II. The Taguchi method for ANN classification to select the suitable number of neurons for breast cancer datasets.

First Number of neurons selected by Taguchi method	Training Accuracy (%)	Validation Accuracy (%)	Data Partitioning		
			Training (%)	Validation (%)	Testing (%)
10	89.7	89.4			
11	89.9	90.8			
12	90.8	92.2			
13	91.5	92.5			
14	92.7	92.8			
15	**93.5**	**93.9**			
16	93.0	93.4			
17	92.8	91.9	50	25	25
18	92.1	92.3			
19	91.2	91.7			
20	90.9	91.3			
Second Number of neurons selected by Taguchi method	Training Accuracy (%)	Validation Accuracy (%)	Data Partition		
			Training (%)	Validation (%)	Testing (%)
20	90.7	91.1			
19	90.1	90.8			
18	91.9	91.2			
17	93.7	92.8			
16	95.2	94.7			
15	**96.6**	**96.2**			
14	96.1	96.2			
13	95.8	95.6	60	20	20
12	95.4	95.3			
11	93.8	94.2			
10	92.3	93.0			
Third Number of neurons selected by Taguchi method	Training Accuracy (%)	Validation Accuracy (%)	Data Partition		
			Training (%)	Validation (%)	Testing (%)
10	93.6	93.1			
11	94.5	94.3			
12	95.7	94.9			
13	97.2	96.8			
14	98.0	97.5			

Number of neurons selected by Taguchi method	Training Accuracy (%)	Validation Accuracy (%)	Training (%)	Validation (%)	Testing (%)
15	**98.5**	**98.3**	**70**	**15**	**15**
16	98.3	98.2			
17	97.8	97.6			
18	97.1	97.1			
19	96.8	96.3			
20	96.2	96.3			

Fourth

Number of neurons selected by Taguchi method	Training Accuracy (%)	Validation Accuracy (%)	Data Partition		
			Training (%)	Validation (%)	Testing (%)
20	93.4	93.1			
19	92.8	93.2			
18	92.2	92.9			
17	93.1	92.7			
16	93.9	92.4			
15	**94.7**	**93.8**			
14	94.2	93.4			
13	93.9	93.8	80	10	10
12	93.2	93.6			
11	93.0	92.5			
10	92.9	92.4			

Table III . Classification accuracy of the proposed method based on 30 different runs

1   **Table III**. Classification accuracy of the proposed method based on 30 different runs

Number of Experiments	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)	Percent Error
1	97.7	96.5	98.7	1.3
2	98.5	98.7	98.8	1.2
3	98.9	97.9	99.2	0.8
4	98.2	98.7	98.5	1.5
5	97.7	97.2	98.6	1.4
6	98.6	98.6	98.8	1.2
7	98.8	98.5	99.2	0.8
8	98.4	98.2	98.4	1.6
9	98.7	98.7	99.2	0.8
10	98.5	98.5	98.8	1.2
11	98.2	98.2	98.8	1.2
12	98.8	98.8	99.2	0.8
13	98.3	98.0	98.6	1.4
14	98.5	98.3	98.8	1.2
15	98.3	98.5	98.7	1.3
16	98.5	98.5	98.9	1.1
17	98.6	98.5	99.2	0.8
18	98.8	98.5	99.2	0.8
19	98.9	98.7	98.8	1.2
20	98.8	98.5	98.9	1.1
21	98.2	98.2	98.5	1.5
22	98.5	98.5	98.7	1.3
23	98.9	98.8	98.8	1.2
24	98.6	98.5	98.3	1.7
25	98.5	98.5	98.8	1.2
26	98.4	98.4	99.0	1.0
27	98.9	98.5	99.1	0.9
28	98.6	98.1	98.4	1.6
29	98.2	98.3	98.9	1.1
30	98.6	97.5	98.6	1.4
Average	**98.5**	**98.3**	**98.8**	**1.2**

Table IV. Model performance based on confusion matrix after 10-fold cross-validation

1 **Table IV.** Model performance based on confusion matrix after 10-fold cross-validation

Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	F-Score (%)
Training Dataset	98.5	98.8	98.0	98.4
Testing Dataset	98.8	99.1	98.4	98.8

Table V. Performance comparison between different classifiers and the proposed method in breast cancer classification

1 **Table V**. Performance comparison between different classifiers and the proposed
2 method in breast cancer classification.

Model	Name of the classifier	Accuracy
1.1	Fine Tree	97.20%
1.2	Medium Tree	96.90%
1.3	Coarse Tree	97.50%
1.4	Linear SVM	98.00%
1.5	Quadratic SVM	97.70%
1.6	Cubic SVM	97.00%
1.7	Fine Gaussian SVM	80.70%
1.8	Medium Gaussian SVM	97.90%
1.9	Coarse Gaussian SVM	98.10%
1.10	Fine KNN	79.80%
1.11	Medium KNN	97.80%
1.12	Coarse KNN	96.30%
1.13	Cosine KNN	97.10%
1.14	Cubic KNN	97.60%
1.15	Weighted KNN	97.80%
1.16	Boosted Tree	62.70%
1.17	Bagged Tree	97.30%
1.18	Subspace Discriminate	95.80%
1.19	Subspace KNN	85.40%
1.20	RUSBoosted Tree	62.70%
1.21	**Proposed IANN**	**98.80%**

Table VI. Comparison of breast cancer classification performance between existing methods in the literature and the proposed method

1   **Table VI**. Comparison of breast cancer classification performance between existing
2   methods in the literature and the proposed method

Author Name and Reference	Method	Accuracy
Aalaei et al. [5]	GA- Based ANN	96.50%
Salama et al. [27]	SMO	97.70%
Xue et al. [31]	PSO	92.98%
Nekkaa et al. [32]	MA+SVM	97.88%
Ouy et al. [36]	APSO-NN	98.24%
Amrane et al. [48]	KNN	97.51%
Abdar et al. [49]	SV-NB-3-MetaClassifier	98.07%
The proposed	**IANN**	**98.80%**

Table VII. 10-fold cross-validation experimental performance

Table VII. 10-fold cross-validation experimental performance.

Cross Validation (10-folds)	Training Accuracy (%)	Test Accuracy (%)
K-1	99.2	97.9
K-2	98.6	100
K-3	99.0	100
K-4	98.6	98.2
K-5	97.8	98.5
K-6	98.2	98.2
K-7	98.4	98.5
K-8	98.0	98.6
K-9	98.8	98.8
K-10	98.3	98.2
Average	98.5	98.7

Figure 1

The basic architecture of an ANN
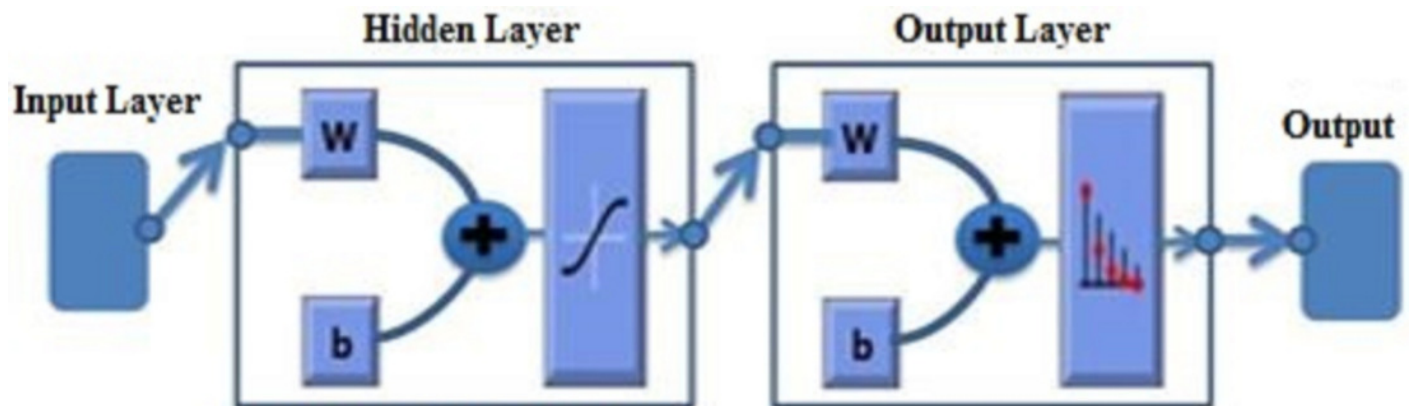
Figure 2

The architecture of a Multi-layer feed-forward neural network with four inputs.
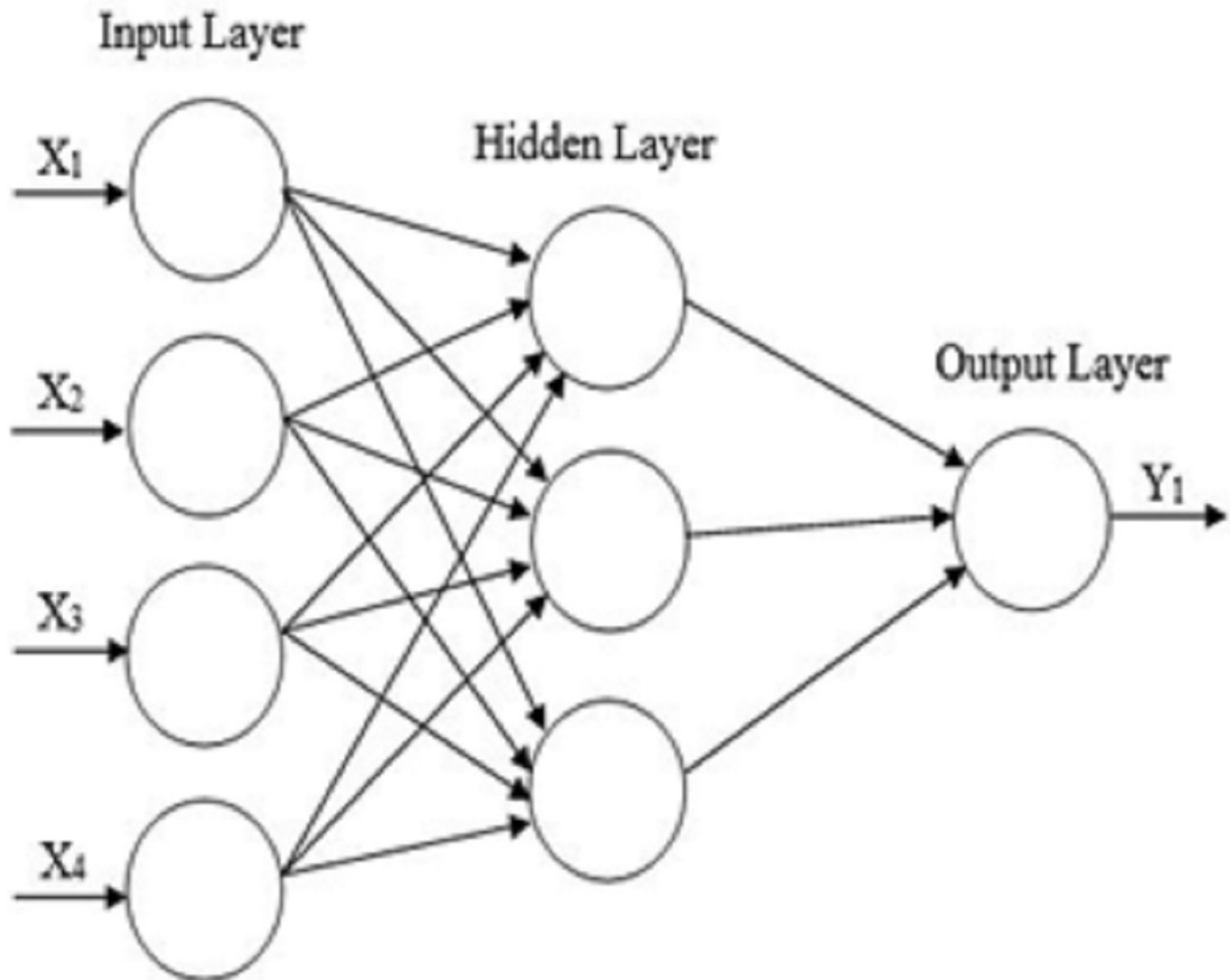
Figure 3

Architecture of the proposed 15 neurons ANN classifier and two output layers
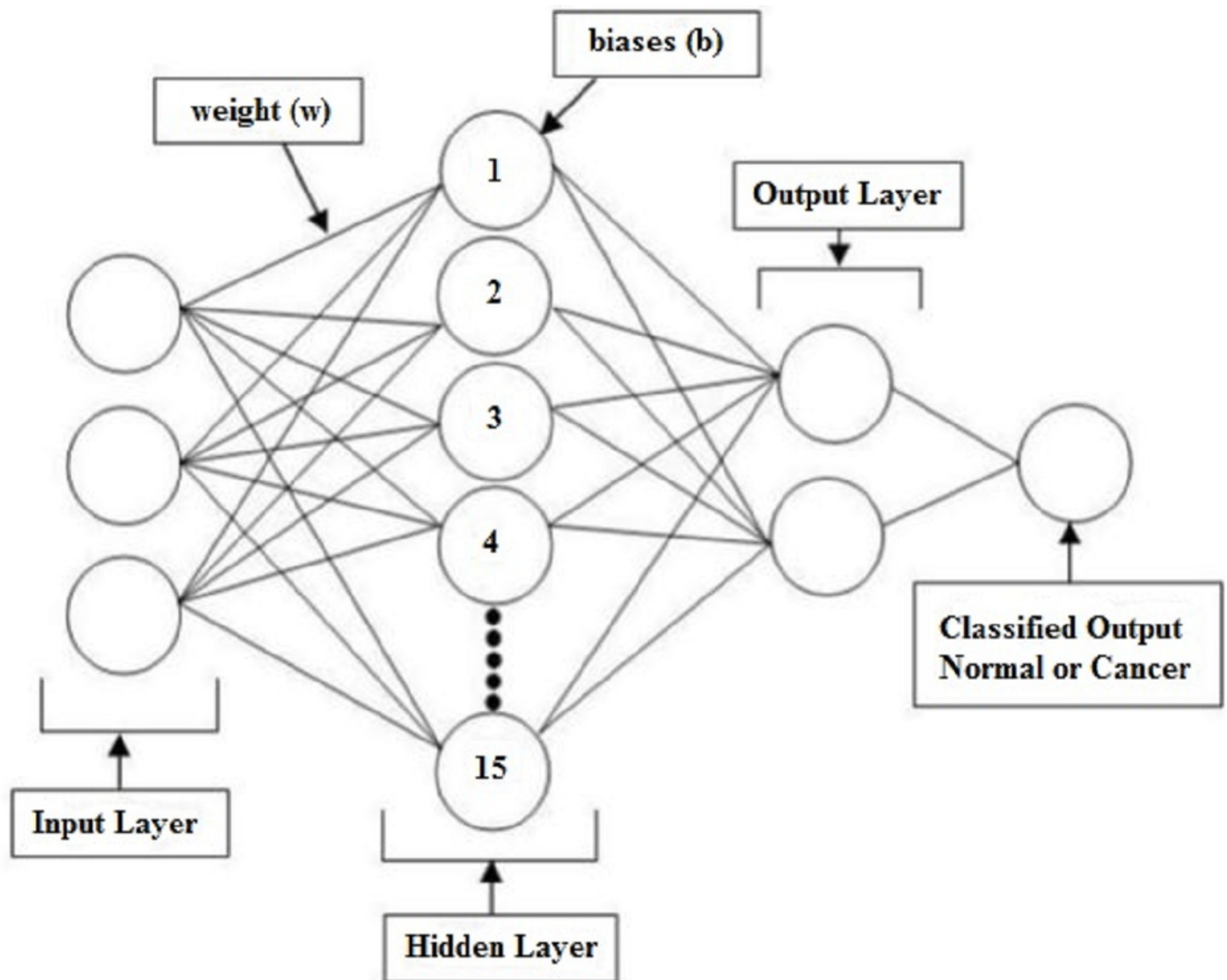
Figure 4

The proposed Improved-ANN (IANN) cancer classification model
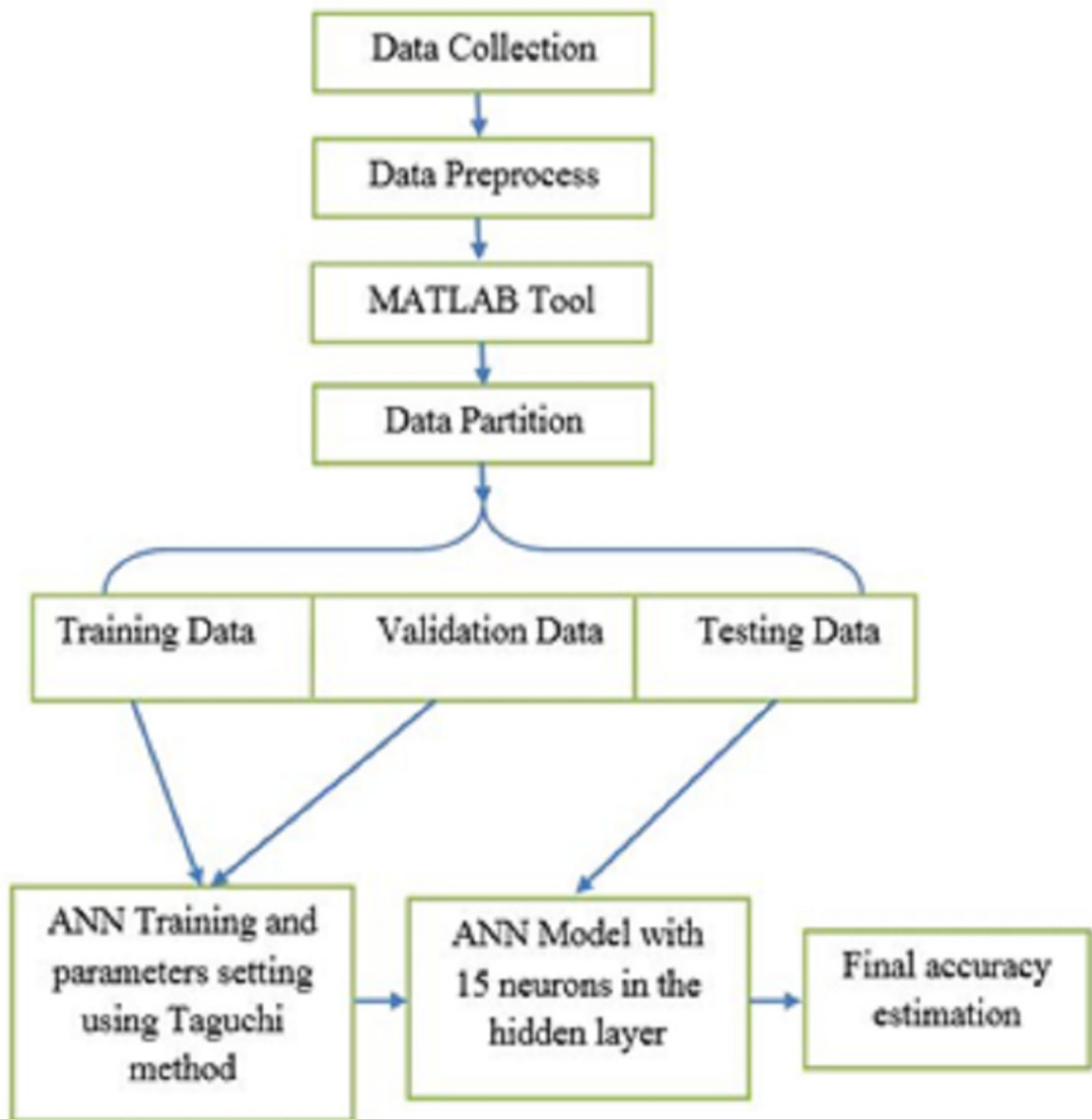
Figure 5

IANN training performance
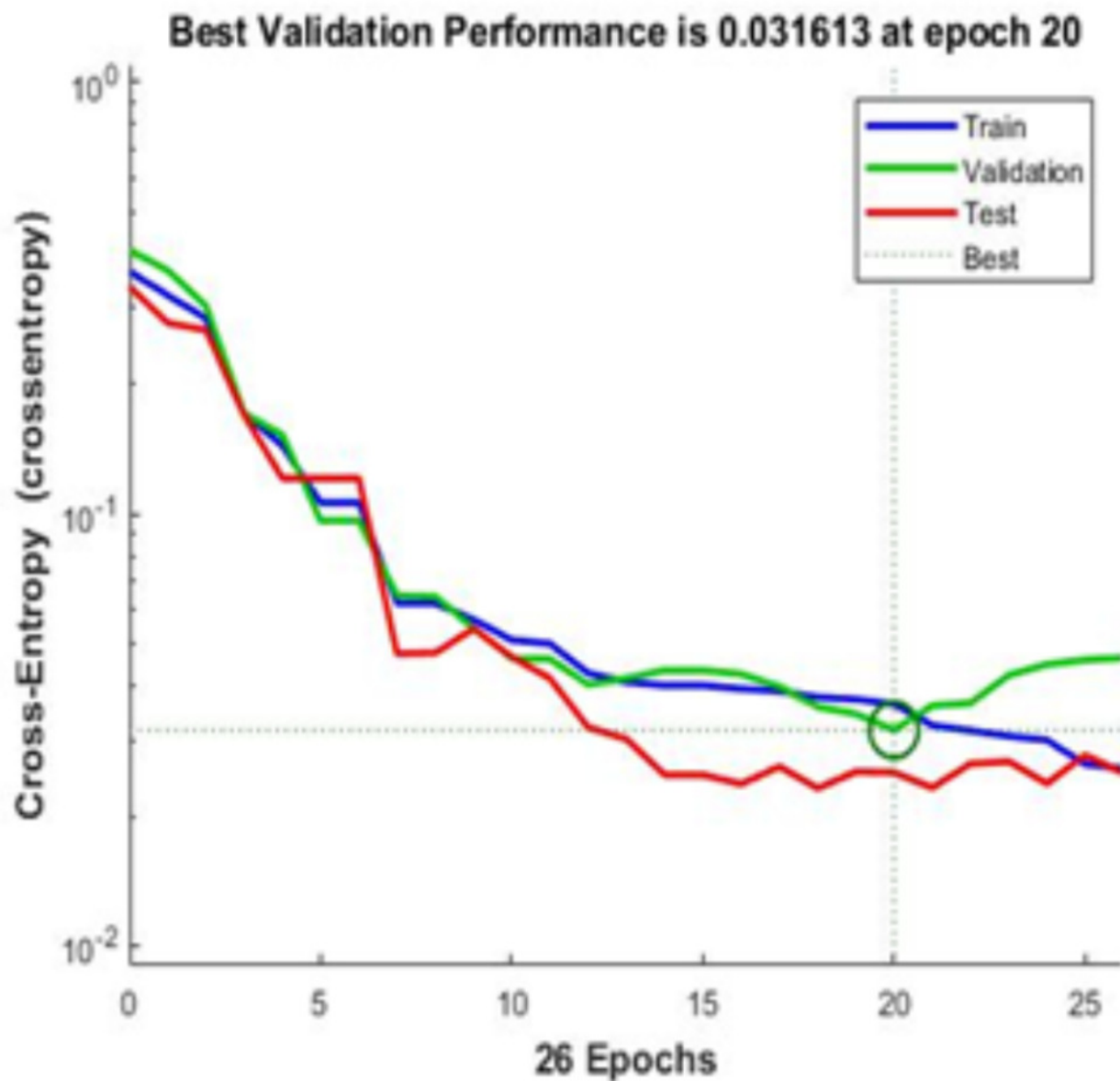
Figure 6

Training state performance
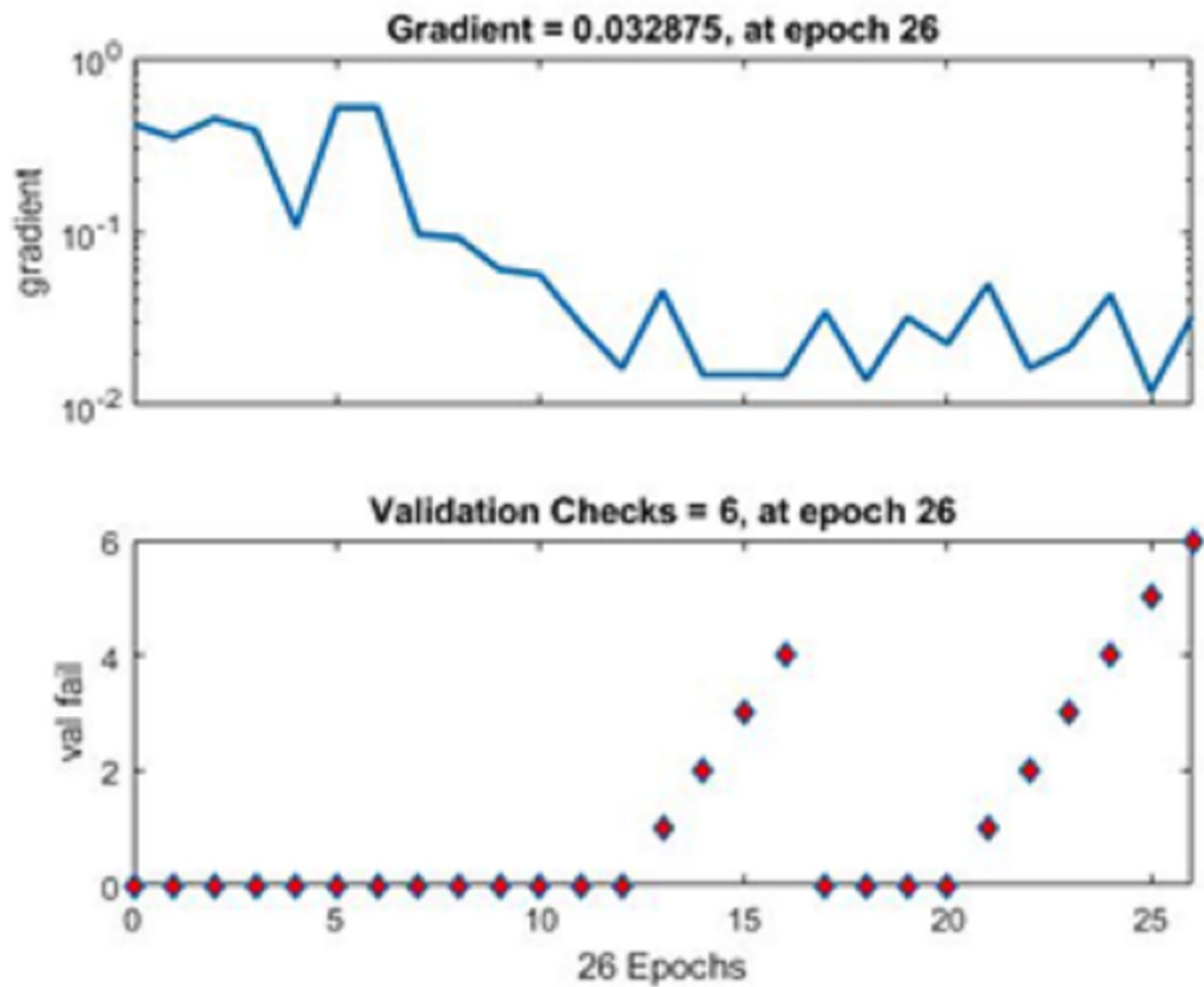
Figure 7

Error histogram
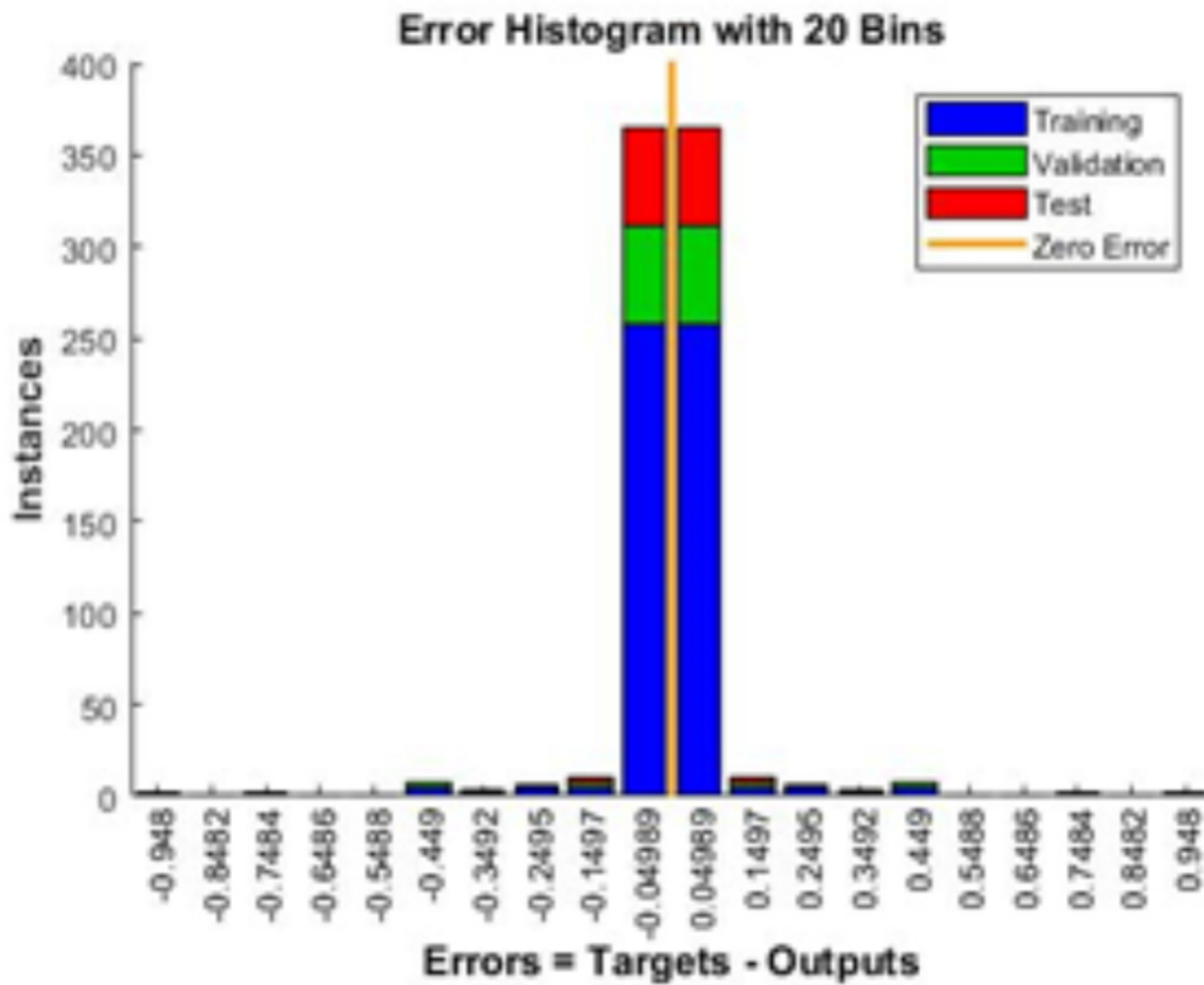
Figure 8

Performance confusion matrix
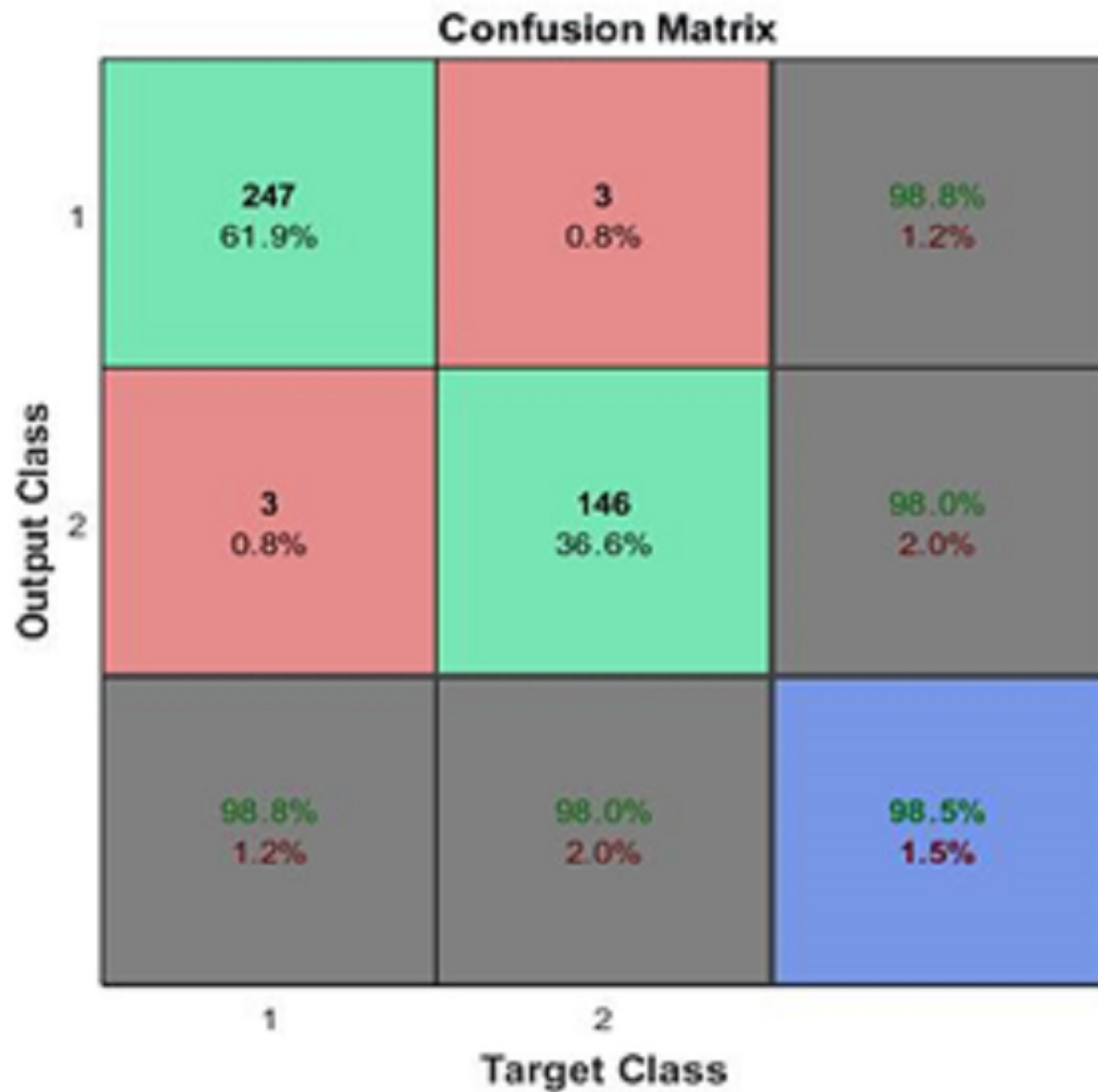
Figure 9

Receiver operating characteristic (ROC) curve
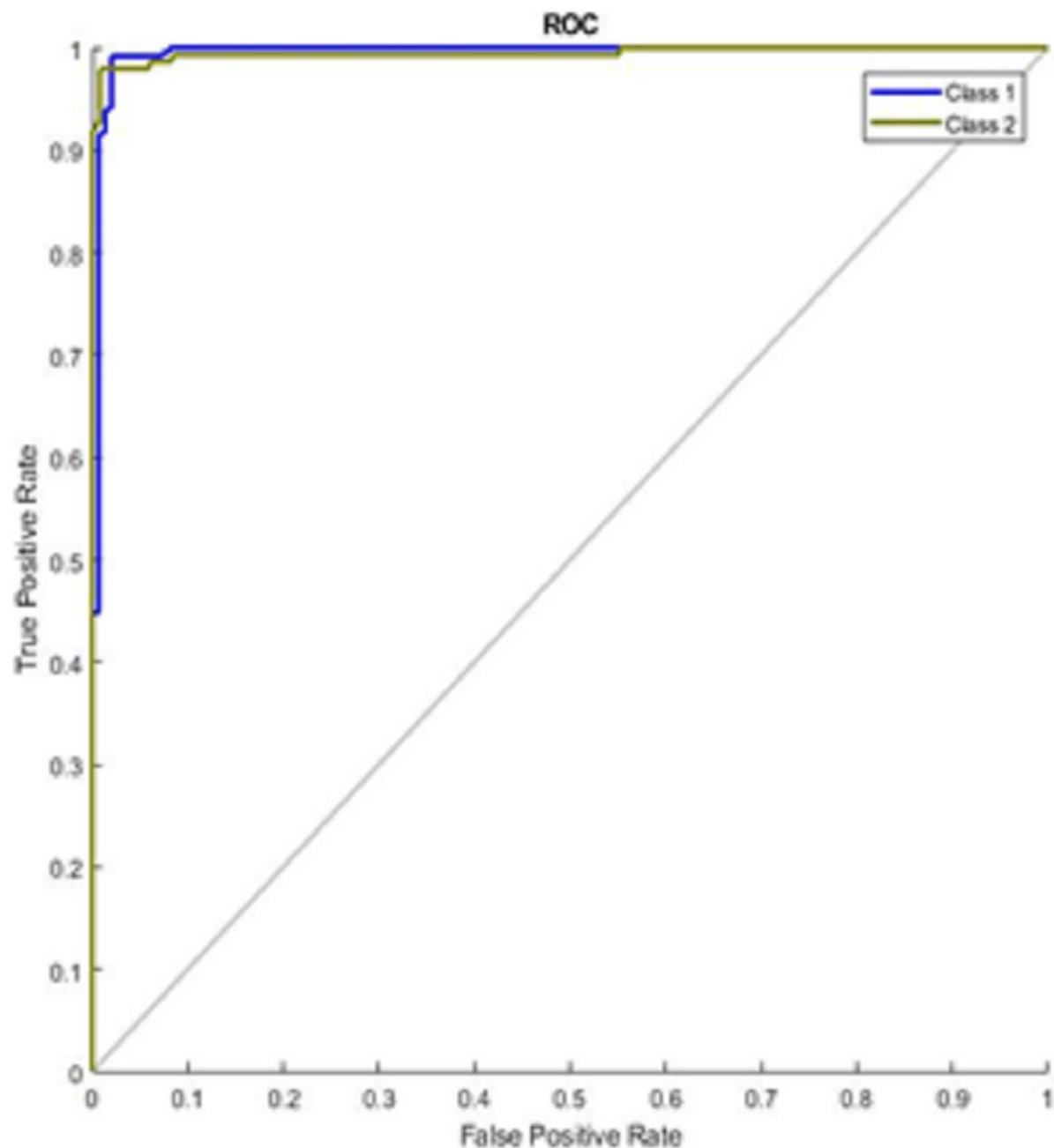
Figure 10

Test dataset result of breast cancer

Figure 11

ROC curve with test dataset of breast cancer