Chapter 5: Classification and analysis of histopathological images of breast cancer

Abstract:

The objective of this chapter is to provide a robust and reliable computational method for the classification of benign and malignant cells using the features extracted from histopathology images of breast cancer through image processing. Morphological features of malignant cells show changes in patterns as compared to that of benign cells. However, manual analysis is time-consuming and varies with perception level of the expert pathologist. To assist the pathologists in analyzing, morphological features are extracted, and two datasets are prepared from the group cells and single cells images for benign and malignant categories. Finally, classification is performed using supervised classifiers. Morphological feature analysis is always considered as an important tool to analyze the abnormality in cellular organization and changes in the state of physiology. In the present investigation three classifiers such as (1) Artificial Neural Network (ANN), (2) k-Nearest Neighbour (k-NN) and (3) Support Vector Machine (SVM)) are trained using publically available breast cancer datasets. The result of performance indication for benign and malignant images was measured by True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) respectively. By utilizing the number of samples that fall into these categories, the performance parameters accuracy, sensitivity and specificity, balanced classification rate (BCR), F-measure and Matthews' correlation coefficient (MCC) are calculated. The statistical measure regarding sensitivity and specificity has been obtained by calculating the area under the Receiver Operating Characteristic (ROC) curve. It is found that the classification accuracy achieved by the single cells dataset is better than the group cells. Furthermore, it is established that ANN provides a better result for both datasets than the other two (k-NN and SVM). The proposed method of the computer-aided diagnosis system for the classification of benign and malignant cells provides better accuracy than the other existing methods.

Keywords: Segmentation, Cancer, Morphological features, Histopathology, Classification.

5.1 Introduction:

Classification refers to the problem of categorizing objects, such as images, regions or pixels, into appropriate classes based on contextual information. An identity from one of the known class is used to assign the unknown test parameters; this task is known as Classification. All the pixels of a digital image are categorized in one of the several classes that are the primary intent of classification. The features are extracted from the image and are used to identify cancer. A classifier refers to an algorithm, sometimes a mathematical function that implements a classification (Chen *et al.,* 2012; Veta *et al.,* 2014). The contextual information of a sample object is often represented by a feature vector that ideally contains discriminative information of the object (Demir *et al.,* 2005; Gonzalez *et al.,* 2004; Belsare *et al.,* 2012; Bhattacharjee *et al.,* 2015; Chankong *et al.,* 2014; Shazia *et al.,* 2016). After defining an appropriate set of features, the extracted features are then passed to a classifier which analyses the quantified characteristics and determines the class of each object indicated (Bhattacharjee *et al.,* 2015; Zhang *et al.,* 2015; Zhao *et al.,* 2015). The classification algorithm is decided based on the assumption that one or more features belonging to a distinct class are sufficiently distinguishable. The goals of classification are to separate a breast histopathology image into appropriate classes based on the available features. A further classification algorithm sorts an image into several classes, comprising normal and abnormal categories (Chankong *et al.,* 2014; Shazia *et al.,* 2016; Wang *et al.,* 2016; Veta *et al.,* 2014; Dheeba *et al.,* 2014)

5.2 Validation of the Classifier (Training and Testing):

There are two processes involved in defining the method for classification, and these include training and testing of the proposed classifier. During the training process, the algorithm, the parameter and features boundaries of the classifier are determined, whereas the testing process is used to classify new cases based on the previously determined training parameters and feature boundaries. In the training process, the feature properties of a current image are determined based on the rules defined for each constructed classification. The training process is crucial in the classification stage; its

purpose is to express the characteristics of the classification categories (Fisher *et al.,* 2000). The criteria for constructing the training process are described below.

Consistency: the classification system should provide the same result if given the same data.

Discriminatory: the different class or image features should have significantly different characteristics.

Reliability: in the training process, the standard description of a current group should represent all image features within the group.

For accuracy of results, it is necessary to use other independent data for testing so that the system is not biased towards the training data set (Demir *et al.,* 2005). Henceforth, the data used for testing should be different from the data used for training (Diamond *et al.,* 2004; Blekas *et al.,* 1998; Esgiar *et al.,* 1998; Gunduz *et al.,* 2004).

5.3 Types of Classification:

To optimize the classifier, there are two models for learning: **(a)** unsupervised classification and **(b)** supervised classification.

5.3.1 Unsupervised classification:

Unsupervised classification is a heuristic approach that tries to find a hidden structure of unlabeled data, and thus it is a clustering method of data into groups. In unsupervised classification, the system is trained with unknown samples. To achieve accuracy during the test process the training samples must be representative of test data. Training data is not utilized by unsupervised classifiers as it forms the basis for classification. Rather, this family of classifiers includes algorithms that observe the unknown pixels in an image and on the basis of natural groupings or clusters present in the image values aggregate them into some classes. It performs good in cases where the values within a given cover type are close together in the measurement space, data in different classes are comparatively well separated. Unsupervised algorithms include K-means clustering (Wu *et al.,* 2007; Ilea and Whelan, 2006) and Fuzzy c-means (FCM) (Hafiane *et al.,* 2008; Rezaee *et al.,*1998; Belsare *et al.,* 2012).

5.3.2 Supervised classification:

In supervised classification, a sample is taken from the known label, and then the classifier is trained. Supervised Learning is a heuristic method in which the input and output data are known, and a priori including training data is provided to guide the classification. Supervised classification consists of machine learning task, where, from a set of labelled training data, a particular function is to be inferred. A set of training examples is provided in training data. Each of the examples has an input and a predetermined output value. In this type of classification technique, the learning algorithm is utilized to analyze the training data by which it infers a function for predicting new examples.

Three types of classifiers are chosen to test their classification accuracy with the extracted features. The three classifiers such as (1) Artificial Neural Network (ANN) (2) k-Nearest Neighbour (k-NN) and (3) Support Vector Machine (SVM) are utilized to classify the cells into two sets, normal and diseased. A brief description of each classifier was explained as below.

5.3.2.1. Artificial Neural Network (ANN):

Artificial Neural Network (ANN) is a statistical learning algorithm inspired by a human logic neural network that is defined by an interconnected set of nodes, as similar to the network of neurons found in the brain to predict the outcome. Each node of the network is the representation of a neuron, which are working on simultaneously to solve a specific task in machine learning and pattern recognition. ANN is composed of large number of interconnected processing elements (artificial neurons) working in unison to solve specific problems. It is mainly used to generate or estimate the output from a large input data set (Theodoridis *et al.,* 2010). Like human being the ANN is also learning from the instances or examples (Mohanty and Ghosh, 2010). In case of ANN learning is achieved by the adaptation of weights and bias of the neurons with respect to the training procedure and training data. Figure 5.1 depicts the example of ANN classification.
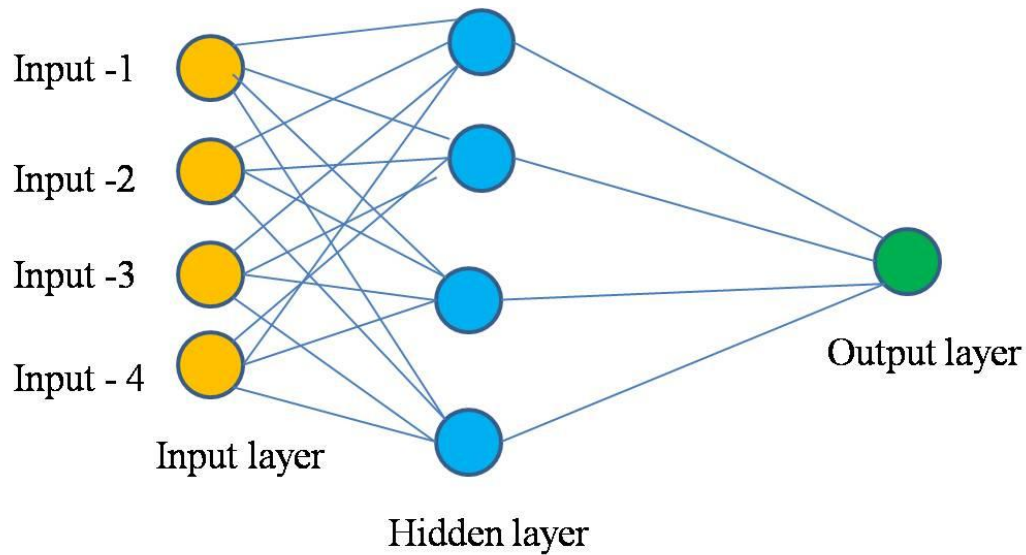
Figure 5.1: Example of ANN classification.

Suppose a neural network is designed for character recognition. The input is designed in this fashion, when it gets the pixel of an input image, the set of input neurons in the developed neural network get activated. The weighted activation and transformation are entirely dependent on the designer. After activation and transformation of the input neurons, it passes through the other neurons and repeats the process until the output neuron is activated. This process can determine the input character from the images (Theodoridis *et al.,* 2010).The weights are adjusted by propagating the discrepancy in a backward direction to minimize the discrepancy between real and predicted outcomes (Ashwin *et al.,* 2012; Han, 2006; Svozil, 1997).

5.3.2.2 k-Nearest Neighbour (k-NN):

k-NN method was firstly described in 1950 (Fix and Hodges, 1951; Silverman *et al.,* 1951). Due to its simplicity, efficiency, and application in a few of educational patterns or the same samples, it has been used in recognizing pattern (Phyu, 2009; He *et al.,*2000). k-NN is considered as one of the best classifiers family of the non parametric method with a good scalability. The classifier output is the assignment of sample data to a particular group of the predefined groups. Specific data is assigned to a particular class by considering the higher number votes of its neighbors.

A few numbers (defined by k) of data are considered as neighbors that are taken from a set of data for which the class is predefined to the classifier. The function which has to be approximate dis predicted locally. Furthermore, all the computation will be blocked while classification process is running. The learning process of this classifier is called as lazy learning or termed as instance based learning. The k-NN algorithm is the least complex algorithm among all machine learning algorithms. Here, the value of k-NN=1 has been utilized to classify between benign cells and malignant cells (Wold *et al.,* 1987). Figure 5.2 depicts the illustration of k-NN classification.
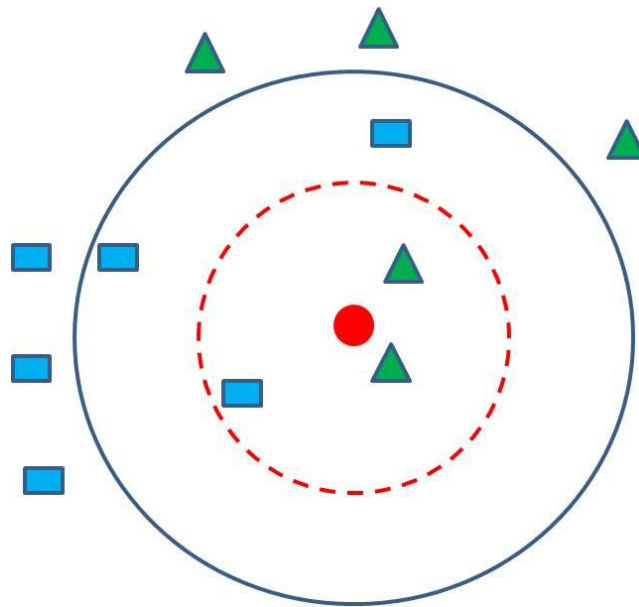

Figure 5.2: Illustration of k-NN classification.

In the image 5.2 shown, there are two classes, the first class is the blue squares and the second class is of green triangles. The red circle is the test sample which is to be classified. As there are 2 green triangles and 1 blue square inside the circle when we take k=3 (3 neighbors only), it will be classified to the class of green triangles. By the same procedure for k = 5 (5 neighbors only), the sample will be classified to the class of blue squares.

5.3.2.3. Support Vector Machine (SVM):

Support vector network is used for the purpose of data analyzing and pattern recognition. Vladimir and Vapnik invented the original SVM algorithm further, Cortes

and Vapnik proposed the current standard incarnation (soft margin) in year1995. Supervised learning algorithms are provided to SVM so that it can be used as a classifier. If a set of training data is given, such that each of them marked as one of the two predefined categories, then a model is built by SVM algorithm where the points are mapped into space and separated by an apparent gap, as broad as possible. It mainly constructs a hyperplane or set of hyperplanes in high dimensional space that is used for classification (Campbell and Ying, 2011). Figure 5.3 depicts the example of SVM classification.
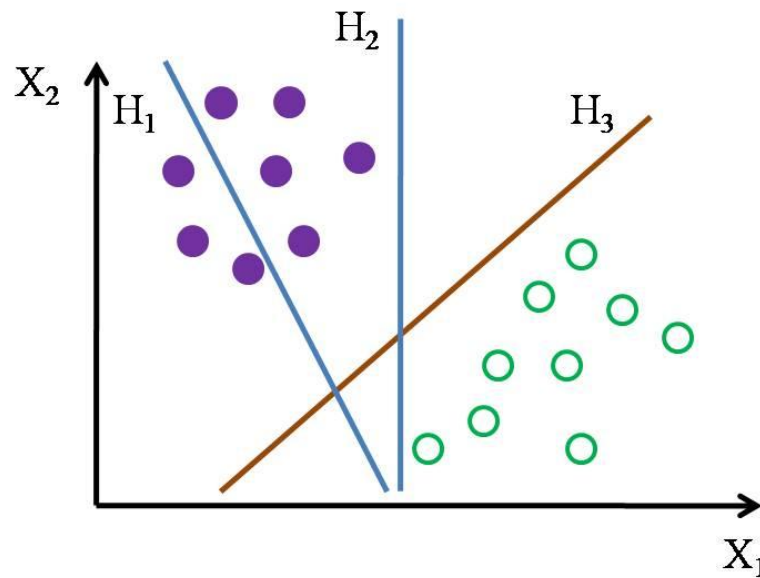

Figure 5.3: Example of SVM classification.

Support Vector Machine is a linear classifier. The following will provide an explanation for this. Suppose, we have to assign a new data point to a particular class, while already we have two reference classes and a few data points, which are already assigned to one of those predefined classes. When SVM classifier is used for classification, a data point is considered as a dimensional vector (a list of numbers). Subsequently, all the reference points for which we have predefined class registered, the task is to separate those points with a dimensional hyperplane. A lot of hyperplanes can be drawn which will be able to classify the data. The best choice will be that one, which will have the largest distance from the nearest data points of both classes.

The hyperplane is defined as the maximum margin hyperplane, and the linear classifier is known as a maximum margin classifier. In figure 5.3, H_1 is unable to separate the classes. However, H_2 is able, only with a small margin and not in the best possible way. H_3 isolates them with the maximum margin (Boser *et al.,* 1992; Hsu and Lin ,2002; Crammer *et al*., 2001).

5.4 Materials and Method:

This chapter deal with the classification of the benign and malignant cells based on features extracted from morphological, intensity and texture. We have followed the same methodology for preprocessing, segmentation and features extraction as given in chapter- 4 of this present thesis. Figure 5.4 represents the flow chart of the proposed system (classification) and basic steps involved in the computer aided diagnosis.

5.4.1 Dataset preparation:

For feature extraction, morphology and intensity based 12 features were utilized to form the group cells dataset. Similarly, the morphology and texture based 30 features were used to form the single cells dataset. In excel file (.xls) dataset of order 463×30 for single cells yields 463(benign 223 + malignant 240) instances and 30 attributes. These features are provided in appendix 1(a) for single benign cells and 1(b) for single malignant cells from (F1-F30) of this present thesis. For group cells, the dataset of order 148×12 in excel file (.xls) based on features yields 148(benign 75 + malignant73) instances and 12 attributes. These features are provided in appendix 2(a) for group benign cells and 2(b) for group malignant cells from (F1-F12) of this present thesis. We have discussed classification part here. For classification of images, various steps are involved in supervised machine learning approaches. Based on the extracted features, the suspicious regions have been classified as benign or malignant cells by various classification methods such as (1) ANN, (2) k-NN, and (3) SVM respectively. For this feature data are divided into two parts -training and testing. The computed dataset has been fed to a classifier for final classification after training the network (75% for training purpose and 25% testing purpose). The classification performance of each classifier for testing and training of samples has been evaluated by 10 fold cross-validation method.
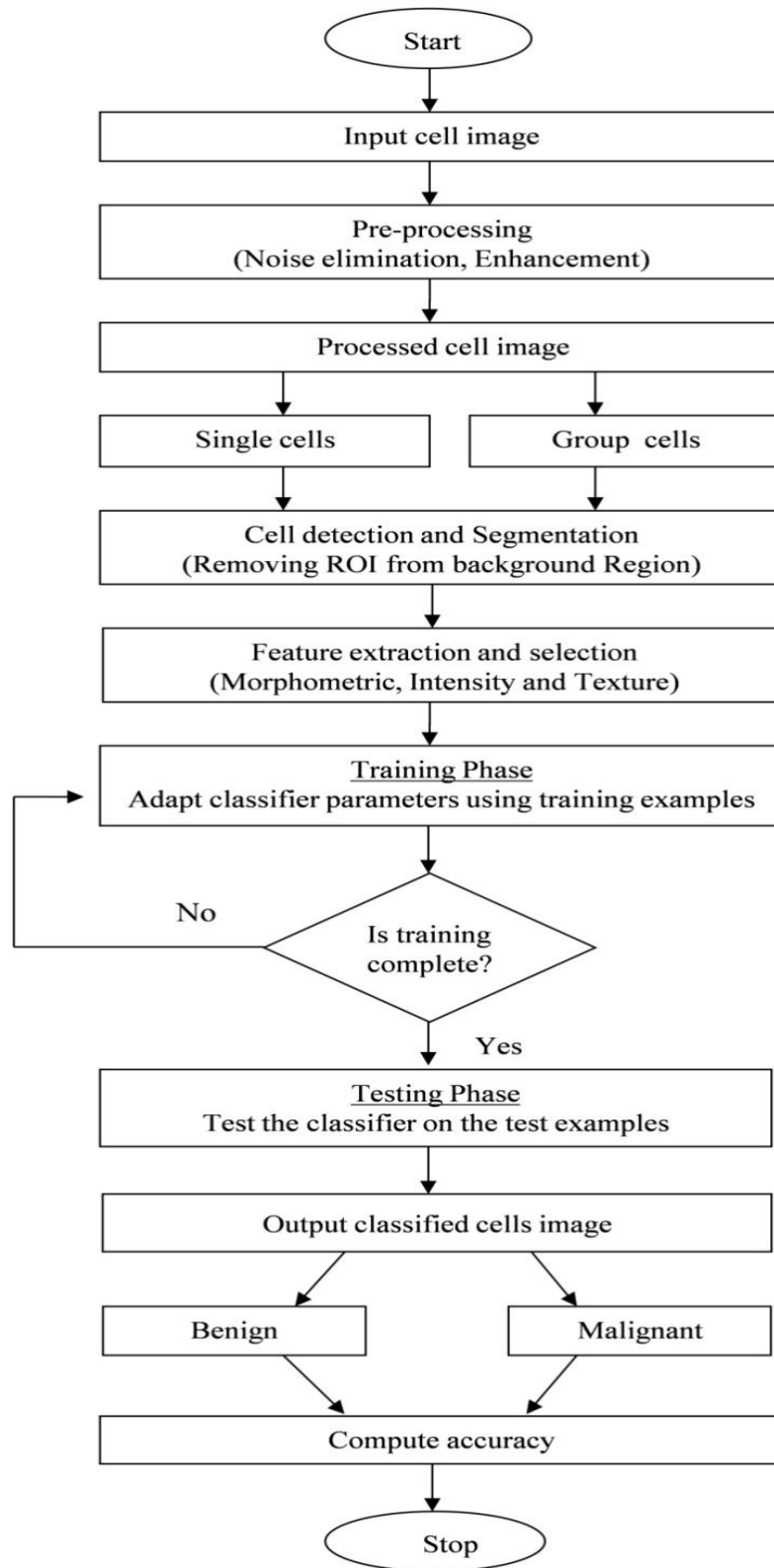
Figure 5.4: Flow chart of the proposed methodology.

Table 5.1 shows the division of the dataset based on training the network.

Table 5.1: Division of the single cell and group cells dataset according to the training and testing purpose.

Data type- Breast cancer cellular Dataset	
Total 58 whole-slide images -32 Benign, 26 Malignant	
Single cells	**Group cells**
Total single cells - 463	Total group Cells - 148
Benign single cells - 223	Benign group cells - 75
Training with 75% - 167	Training with 75% - 56
Sample size from benign single cell dataset 25% - 56	Sample size from benign group cell dataset 25% - 19
Malignant single cell - 240	Malignant group cells - 73
Training with 75% - 180	Training with 75% - 55
Sample size from malignant single cell dataset 25% - 60	Sample size from malignant group cell dataset 25% - 18
Total sample size for testing purpose -116 sample	Total sample size for testing purpose - 37 sample

Afterward, the classifier with the best output has been applied to create a Graphical User Interface (GUI) system in MATLAB for easy and fast detection of cancerous cells. It is a lengthy process to take one image and to do the segmentation, feature extraction and classification in step by step using MATLAB scripts. So a compact GUI has been prepared for cancer cells detection. The proposed algorithms have been fed into the GUI. With the help of GUI benign and malignant cell identification has been performed.

5.4.2. Proposed artificial neural network architecture:

The first step incorporates preparation of the dataset based on feature, the second step comprises selection of suitable algorithm, the third step includes fitting a model, the fourth step consist of training the fitted model, and subsequently, and the final step performs calculations using that fitted model. After determining an appropriate set of features, the next step was to differentiate the malignant morphology from their counterparts. In this step, a cancer cell has been assigned to one of the classes of benign or malignant. We have two classes with the labeling value of 0 (zero) for benign and 1 (one) for malignant that indicates the class of corresponding input respectively. Figure 5.5 represents the classification flowchart of neural network toolbox.
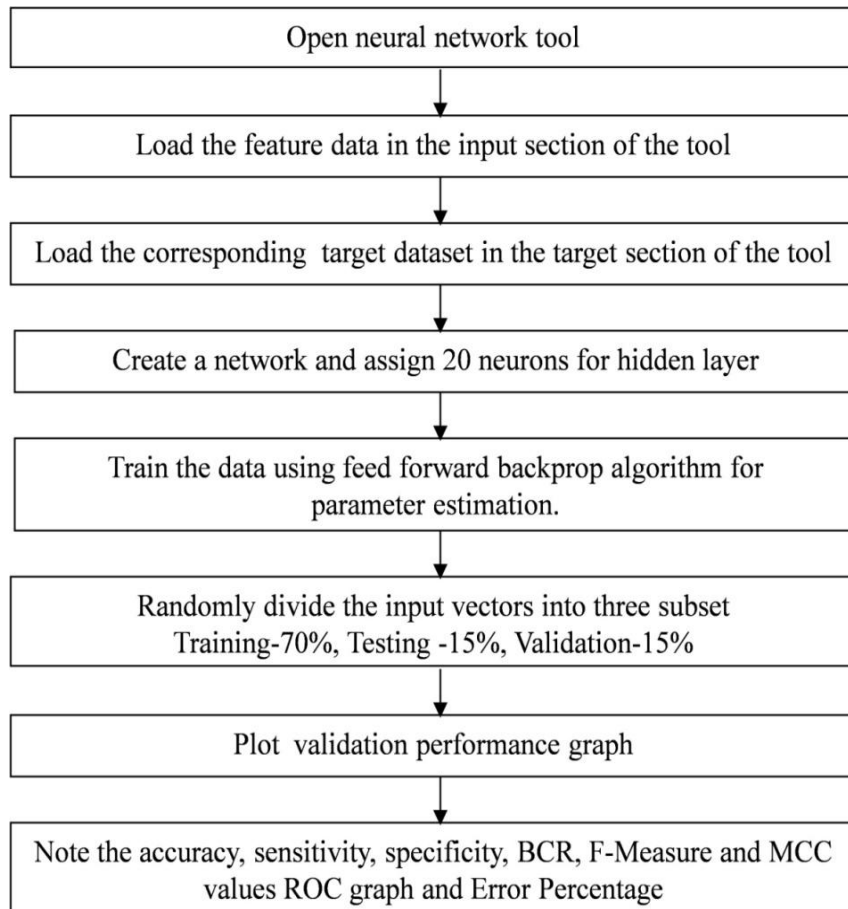
```
┌─────────────────────────────────────────────────────────────┐
│                  Open neural network tool                    │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│        Load the feature data in the input section of the tool │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│   Load the corresponding  target dataset in the target section of the tool │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│      Create a network and assign 20 neurons for hidden layer  │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│     Train the data using feed forward backprop algorithm for  │
│                   parameter estimation.                       │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│     Randomly divide the input vectors into three subset       │
│        Training-70%, Testing -15%, Validation-15%             │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│               Plot  validation performance graph             │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│  Note the accuracy, sensitivity, specificity, BCR, F-Measure and MCC │
│             values ROC graph and Error Percentage             │
└─────────────────────────────────────────────────────────────┘
```

Figure 5.5: Classification flowchart of artificial neural network toolbox.

Load the feature data in the input section of the toolbox and load the target dataset in the target section of the toolbox. The hidden layer taken is about the half of the sum of inputs and outputs and the output consists of one neuron. Network 1 has created. The hidden layer assign of 20 neurons. After that, train the data using feed forward back propagation algorithm. It is trained using algorithm by choosing the training, target and testing sets for evaluation. Figure 5.6 shows the proposed neural network model is a two layer feed forward pattern network with a sigmoid transfer function in both the hidden layer and in the output layer.
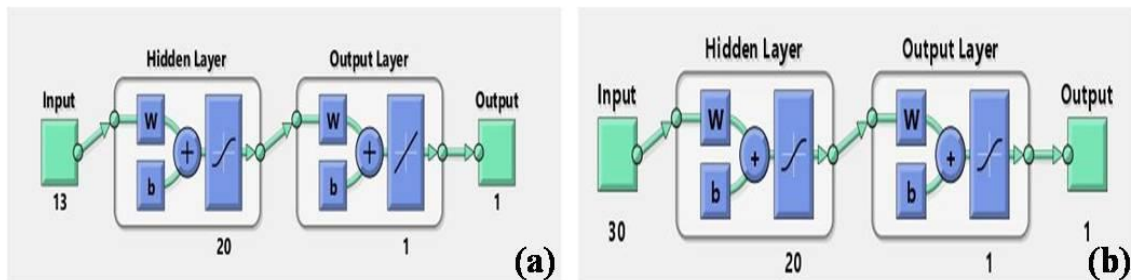
Figure 5.6: Proposed artificial neural network architecture using 20 hidden neurons.
(a) Group cells (b) Single cells

A typical feed forward neural network contains of a number of layers of neural elements. Every neural element in a layer is interconnected with every element in adjacent layers. The strength of every interconnection is characterized by its weight. Information propagates from the input layer (first layer) to the output layer (last layer). Each neural element weights the input sit receives from the elements in the previous layer using the appropriate interconnection weights. Subsequently, the sum of the weighted inputs is filtered through a logistic function to produce an output from the element. The outputs from the output layer represent the final prediction of the neural network. Figure 5.7 depict that data was divided into 70% for training, 15% for validation and 15% for testing. To train a neural network the training set is used and the error of the dataset is minimized during training. To determine the performance of a neural network on patterns that are not trained during learning validation set is used. A test set for lastly examining the overall performance of a neural network.

Results				Results			
	Samples	MSE	R		Samples	MSE	R
Training:	77	1.47777e-2	9.69293e-1	Training:	243	2.91631e-11	9.99999e-1
Validation:	17	5.10259e-2	8.87152e-1	Validation:	52	1.21065e-6	9.99997e-1
Testing:	17	1.07499e-1	7.90945e-1	Testing:	52	2.15218e-6	9.99995e-1
			(a)				**(b)**

Figure 5.7: Train network
(a) Group cells dataset. (b) Single cells dataset.

5.5 Validation of histopathological image:

System validation is an important step to ensure precision in designing medical image processing applications. System validation aims to measure the performance of the medical image processing algorithms and the system analysis. System performance is measured regarding measurement accuracy, image classification, and image segmentation output. Each of these has a different method for measuring the performance of image analysis. Methods for determination of the performance of medical image processing systems are explained in the following subsection.

5.5.1 Performance measurement:

In a medical image processing system, specifically in the feature extraction and classification process, the results are in vector or scalar format (Bowyer, 2000). To calculate the error of the system's performance, a comparison is made of the difference between the values obtained from measured data and those from published literature. It is essential to compare the calculated measurement with some independent measure of the real value (ground truth) (Bowyer, 2000). The average relative error is expressed in equation 5.1 and is an indicator of performance measurement.

$$mean\ relative\ error\ = \frac{(measured\ value) - (true\ value)}{(true\ value)} \qquad (5.1)$$

146

5.5.2 Performance evaluation criteria for classifier:

The performance of classifier depends on various factors like Sensitivity, Specificity and the area under Receiver operating characteristics (ROC) curve. In the context of image analysis, the classification targets can be determined or predicted from the breast histology image. The system should recognize and classify the normality and abnormality of a cell. The CAD then finally identifies the disease (Bowyer, 2000). Performance measurement for image classification has been provided below:

Confusion matrix:

A confusion matrix is used for determining the performance of the classifier. A confusion matrix contains information about actual and predicted classification done by a classification system. The performance of such systems is commonly evaluated using the data in the matrix.

Performance indication for normal (Benign) and abnormal (Malignant) images detection uses True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). Table 5.2 depicts the confusion matrix containing TP, TN, FP, and FN and the terminology for each are described below.

True Positive (TP) is a condition where the system correctly identifies an abnormality.

False Negative (FN) is a condition where the system incorrectly identifies abnormality as normality.

True Negative (TN) is a condition where the system correctly identifies normality

False Positive (FP) is a condition where the system incorrectly identifies normality as an abnormality.

The following table 5.2 represents the confusion matrix for a two class classifier.

Table 5.2: Confusion matrix (FP, FN, TP and TN)

System Decision

		Abnormal (Malignant)	Normal (Benign)
Truth of Clinical Situation	Abnormal	TP	FN
	Normal	FP	TN

Accuracy, Sensitivity, Specificity, balanced classification rate (BCR), F-measure and Matthews' correlation coefficient (MCC):

By using the number of samples that fall into these categories, the performance parameters accuracy, sensitivity, specificity, balanced classification rate (BCR), F-measure and Matthews' correlation coefficient (MCC) are defined to assess the success of the diagnostic system. The definitions of these performance measures have been illustrated as follows from equation 5.2 to 5.7.

Accuracy - The classification accuracy of a technique depends on upon the number of correctly classified samples (i.e., true negative and true positive) and is calculated as follows

$$Accuracy = \frac{number\ of\ TP\ +\ number\ of\ TN}{Total\ number\ of\ N} \times 100 \qquad (5.2)$$

where N is the total number of samples present in the histopathology images for testing.

Sensitivity - Sensitivity is the probability of a positive diagnosis test among persons that have the disease, and it is defined as,

$$Sensitivity = \frac{number\ of\ TP}{number\ of\ TP\ +\ number\ of\ FN} \qquad (5.3)$$

where the value of sensitivity ranges between 0 (mean worst) and 1 (best classification) respectively.

Specificity - Specificity is the probability of a negative diagnosis test among persons that do not have the disease and it is defined as,

$$Specificity = \frac{number\ of\ TN}{number\ of\ TN\ +\ number\ of\ FP} \qquad (5.4)$$

Its value ranges between 0 and 1, where 0 and 1, respectively, mean worst and best classification.

Balanced Classification Rate (BCR) - The geometric mean of sensitivity and specificity is measured as balance classification rate. It is represented by

$$BCR = \sqrt{Sensitivity \times Specificity} \qquad (5.5)$$

F-Measure - The measure is a harmonic mean of precision and recall. It is defined by using

$$Precision = \frac{TP}{TP + FP} Recall = \frac{TP}{TP + FN}$$

$$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (5.6)$$

The value of measure ranges between 0 and 1, where 0 means the worst classification and 1 mean the best classification.

Matthews's Correlation Coefficient (MCC) -MCC is a measure of the distinction of binary class classifications. It can be calculated using the following formula:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{((TP + FN)(TP + FP)(TN + FN)(TN + FP))}} \qquad (5.7)$$

Its value ranges between −1 and +1, where −1, +1, and 0, respectively, correspond to worst, best, at random prediction. The numerical measurement of Sensitivity and Specificity is described as a standard tool called Receiver Operating Characteristic

(ROC) (Charles (1978; Masala (2005; Masala (2006) that is explained in the following subsection.

Receiver Operating Characteristic (ROC):

ROC describes the probability of a correct decision and wrong decision to a varying decision threshold (Masala (2006). True Positive Fraction (sensitivity) is placed along the Y axis, and false positive (1-specificity) is placed along the X axis. The ROC curve describes their relationship. The ideal operating point is achieved in the upper left corner when the TP fraction equals 1, and the FP fraction equals 0. (Masala , 2007). The greater the area under these curves, the better the system (Ojansivu *et al.,* 2013). The performances of each classifier are obtained from ROC curve. The motive has been to choose a classification threshold by random guessing to achieve the business decision of minimizing the false positive rate or maximizing the true positive rate (Spanhol *et al.,* 2016). Receiver operating characteristic (ROC) analysis is an established method of measuring diagnostic performance in medical imaging studies (Sing *et al.,* 2005).

Area under the ROC curve (AUC):

The AUC is a simple measurement metric used to measure the performance of the system. The (AUC) ranges from 0 to 1. The higher value of AUC indicates the greater probability of a true decision and a better performance of the segmentation.

5.6 Results and discussion:

The cell detection test has been carried out with a framed dataset of total 148 group cells consisting of 75 benign and 73 malignant cells. 75% of the dataset (56 benign and 55 malignant cells), for single cells we have used total 463 cells images consisting of 223 benign and 240 malignant cells. All were framed for more accuracy and significant result. 75% of the dataset (167 benign and180 malignant cells) and obtained data has been fed to the classifiers separately for training purpose. Further, for validation of our results, rest 25% of the dataset (19 benign and 18 malignant cells) of group cells and Rest 25% (56 benign and 60 malignant cells) data of single cells were kept for checking purposes. The selected 37 samples in group cells and 116 samples in single cells were used for testing various classification algorithms (e.g. ANN, KNN, and SVM). Figure 5.8 (a) depict the neural network training tool for group cells and figure 5.8 (b) for single cells dataset.
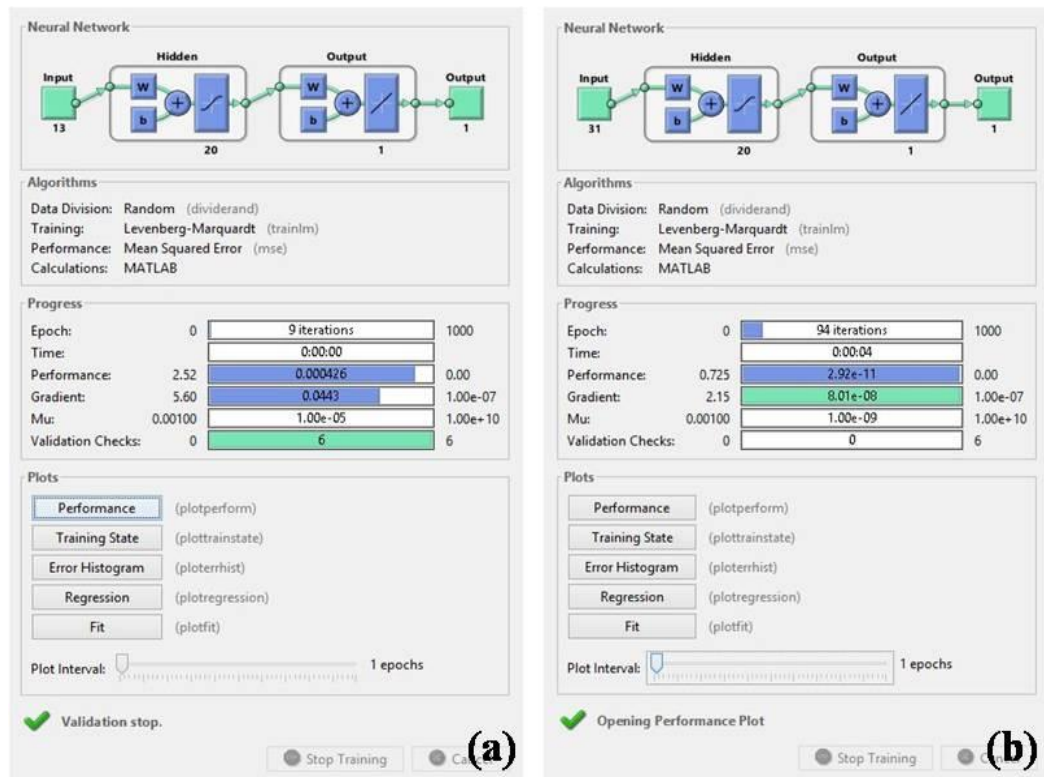
Figure 5.8: Neural network training tool
(a) Group cells dataset. (b) Single cells dataset.

Performance Plot:

The mean squared error is the mean squared normalized error performance function. The error is the difference between the output and the target. If the network performance on the validation vectors fails to improve or remains the same, as indicated by an increase in the mean square error of the validation samples, validation vectors prompt us to stop training early. To see if the network is generalizing well without producing any effect on traininga check based on test vectors was performed.

The performance plot of the training errors, validation errors, and test errors appear, as shown in the training process. Training errors, validation errors, and test errors appear (Ashwin *et al.,* 2012). Figure 5.9 (a) depicts the best validation performance for the group cell dataset is 0.051026 at epoch 3 and Figure 5.9 (b) depicts the single cell dataset is 1.2107e-06 at epoch 94 respectively of breast cancer cells inputs.
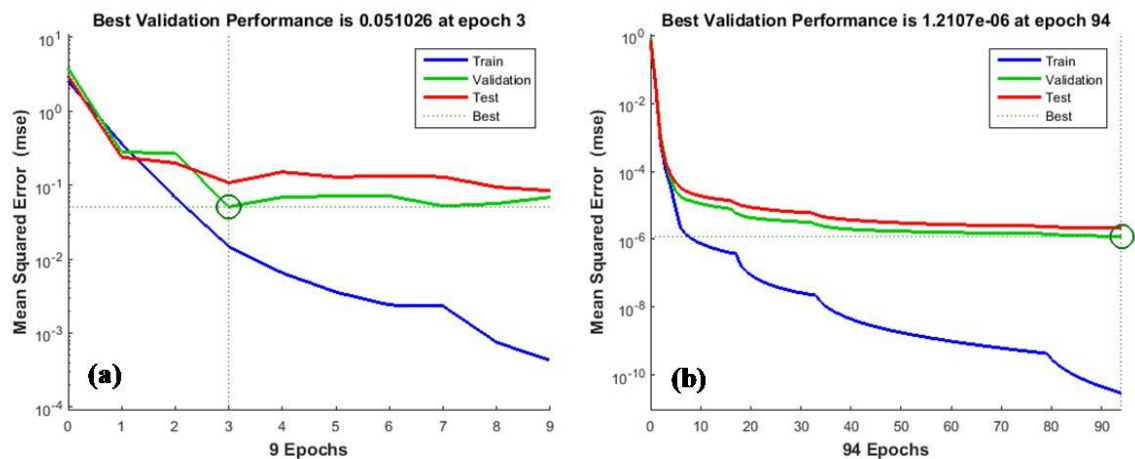
Figure 5.9 Validation performance plot.
(a) Group cells dataset with 13 input. (b) Single cells dataset with 30 input.

The weights are initialized randomly at each execution. The networks continue update during the epochs until the validation reaches the minimum error (MSE between desired and actual outputs) which stops the training Performance plots (Gharehchopogh and Ahmadzadeh, 2012).

Regression Analysis:

The following regression plots as discovered in figure 5.10(a) for group cells and 5.10(b) for single cell dataset reveal the network outputs regarding targets for training, validation, and test sets. The best fit linear regression line is signified by the solid line among outputs and targets. The R value is a sign of the correlation amid the outputs and targets. If R=I, it means that there is an exact linear correlation amid outputs and targets. If R is close to zero, then there is no linear correlation between outputs targets (Ashwin *et al.,* 2012).
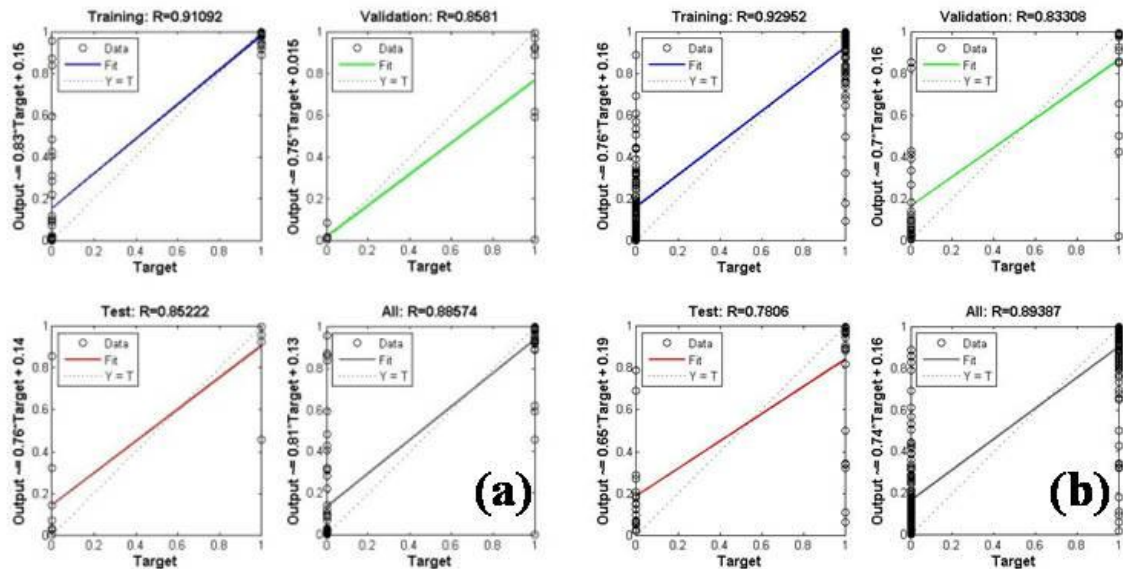
Figure 5.10.Regression plot
(a) Group cells dataset. (b) Single cells dataset.

Error Histogram:

The error histogram of the suggested network that specifies how the error sizes are disseminated. Generally, a maximum number of errors are near zero and very few errors are far from that. The above results suggest that the detection rate with very low false positives are improved. The figure 5.11 (a) shows in group cells and figure 5.11 (b) in single cells the histogram of learning stages errors (Aswin *et al.,* 2012).
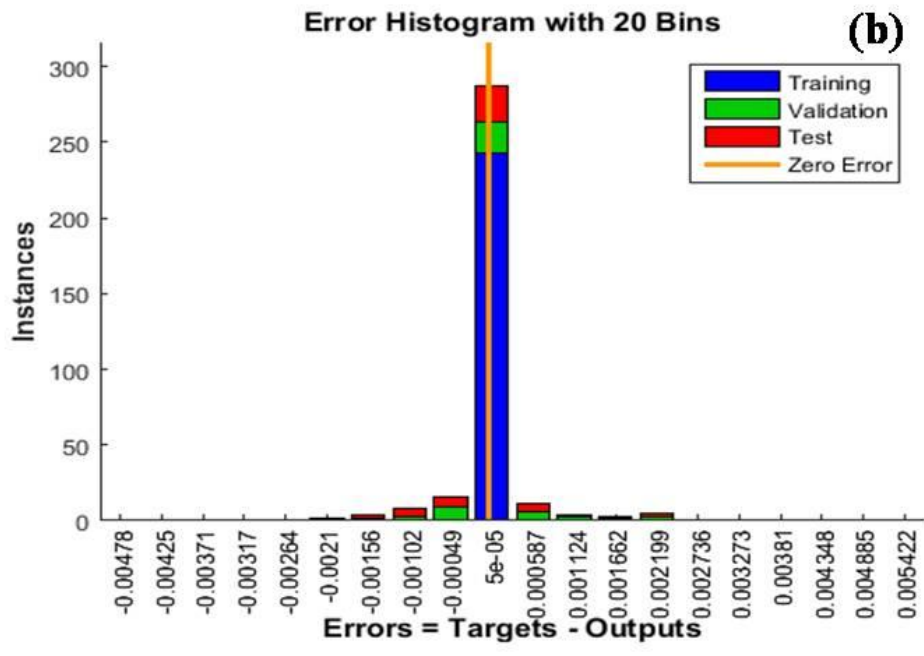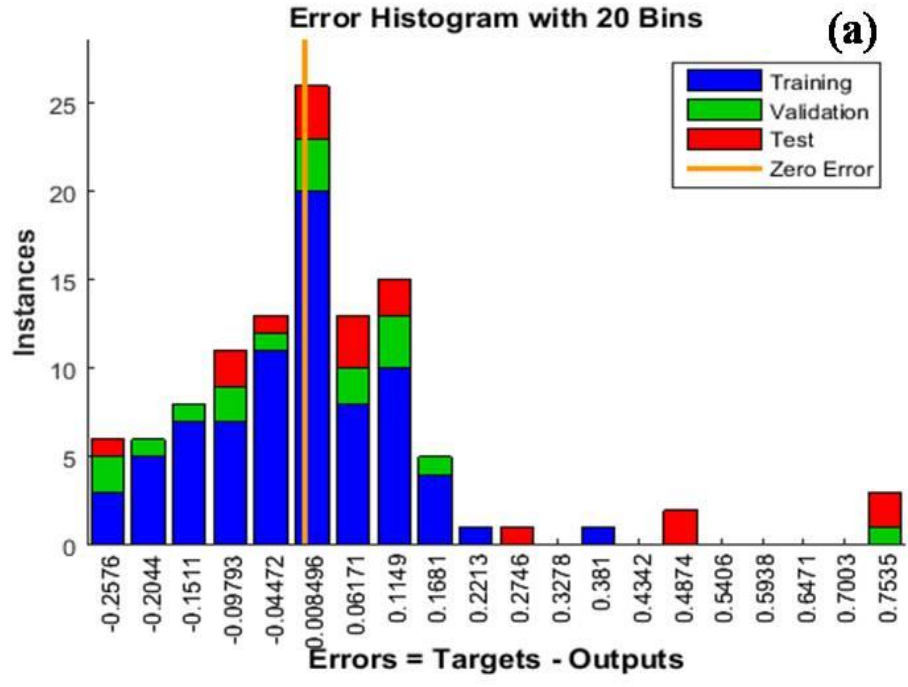
Figure 5.11: Error histogram of learning stages errors of feed forward artificial neural network to classify cancerous cell
(a) Group cells (a) Single cells.

154

Confusion Matrix:

To visualize the performance of the neural network training algorithms a confusion plot is used. The confusion matrix is plotted across all samples. The percentages of correct and incorrect classifications are shown by the confusion matrix. Figure 5.12 depicts network analyze the response of confusion matrix and different errors in final network case. The figure 5.12 (a) and figure 5.12 (b) represents the confusion matrix of 13 and 30 inputs with neural network pattern recognition in group cells and single cells respectively. The result in diagonal green squares shows the correct classification and red squares represent the incorrect classifications. The percentages in the red squares should be very small, indicating few misclassifications, if the network has learned to classify properly. Diagonal elements in each table indicate the number of cases that classifies accurately. The non-diagonal elements are those which classify inaccurately. The blue square cell in the lower right corner of tables indicates total percentage of accurate classification in green colour and the total percentage of inaccurate classification in red color. The proposed neural network system show that results are very good recognition. Figure 5.12 depict the accuracy of the networks shows 98.2% in the blue square in green text colour in group cell and 100% in single cells.
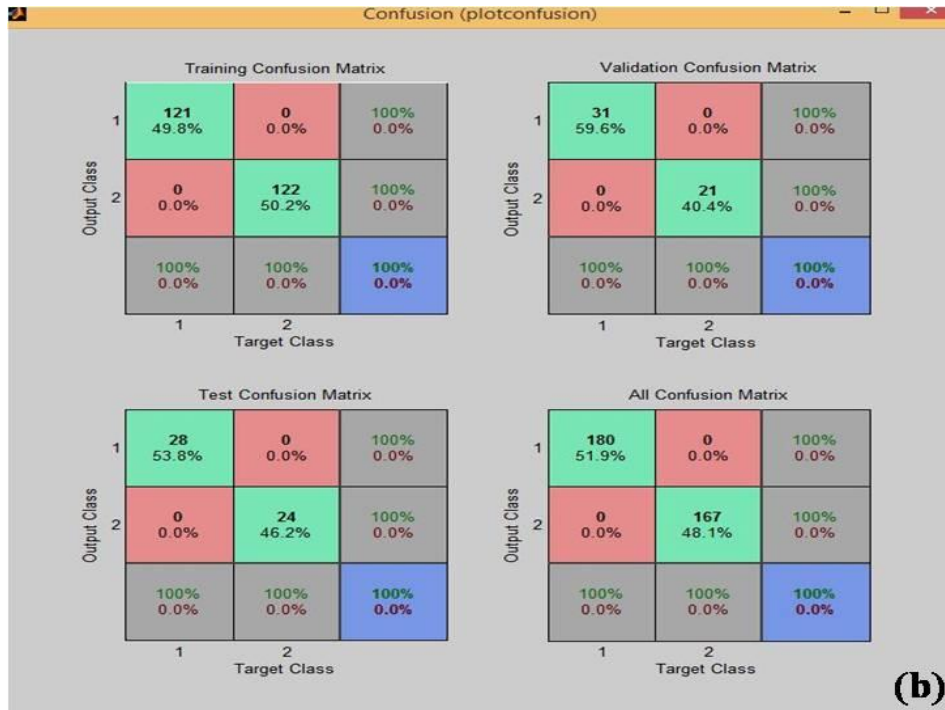

(a)
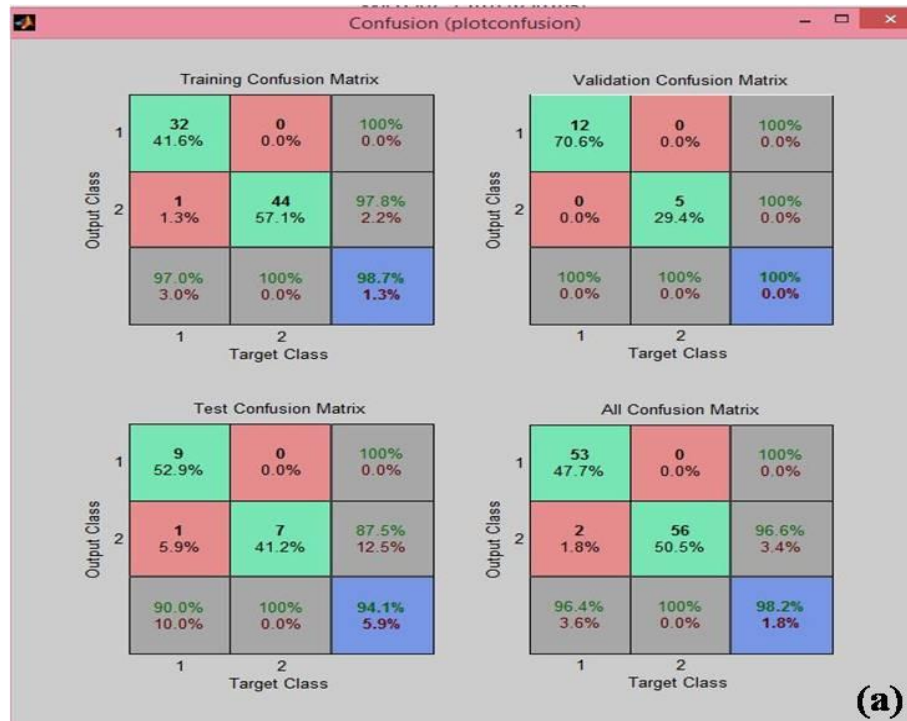
Figure 5.12: Confusion matrix.
(a) Group cells dataset. (b) Single cell dataset.

On the basis of these TP, TN, FP, and FN result accuracy, sensitivity, specificity, BCR, F- measure, and MCC of each classifier has been separately calculated. The table 5.3 and table 5.4 represent performance classification results of the classifiers framework for group cell and single dataset respectively. Furthermore, figure 5.13 and figure 5.14 depicts the performance comparison graph of accuracy and misclassifications for group cell and single dataset respectively. Further, the classification accuracy has been tested with ANN, k-NN, and SVM. The output of k-NN and SVM were compared with the output of ANN.

Table 5.3: Evaluations of proposed method compared with various classifier approaches in group cells.

Dataset (Group cell)	Evaluation Matrices	ANN	k-NN	SVM
Tiff. input histopathological breast images	TP*	12(18)	13	12
	FN*	6	5	6
	TN*	19(19)	16	18
	FP*	0	3	1
	Accuracy (%)	**83.78%**	78.37%	81.08%
	Misclassification (%)	**16.22%**	21.63%	18.92%
	Sensitivity (%)	0.6666	0.7222	0.6666
	Specificity (%)	1	0.8421	0.9473
	BCR	0.8164	0.7798	0.7946
	F- measure	0.7999	0.7646	0.7741
	MCC	0.7118	0.5693	0.6428

Table 5.4: Evaluations of proposed method compared with various classifier approaches in single cells.

Dataset (Single Cell)	Evaluation Matrices	ANN	k-NN	SVM
Tiff. input histopathological breast images	TP*	55 (60)	56	54
	FN*	5	4	6
	TN*	56 (56)	44	43
	FP*	0	12	13
	Accuracy (%)	**95.68%**	86.20%	83.62%
	Misclassification (%)	**4.32%**	13.8%	16.38%
	Sensitivity (%)	0.9166	0.9333	0.9
	Specificity (%)	0.8235	0.7857	0.7678
	BCR	0.8688	0.8563	0.8312
	F- measure	0.956	0.867	0.850
	MCC	0.6769	0.8723	0.6756

TP* True Positive, TN* True Negative, FP* False Positive, FN* False Negative BCR- balanced classification rate, MCC-Matthews' correlation coefficient.
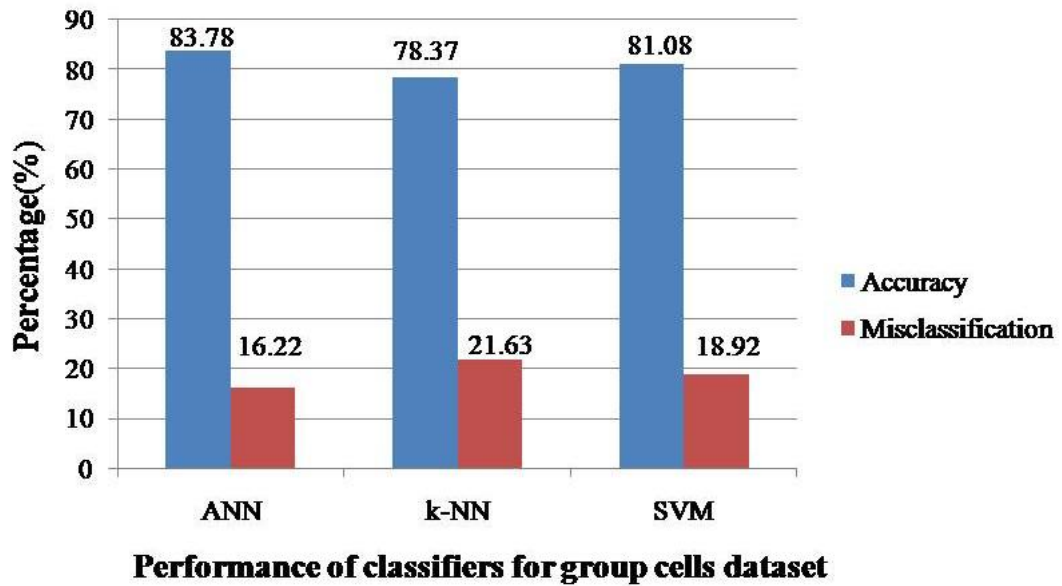
Figure 5.13: Comparison of performance graph of variation of accuracy and misclassification in group cells.
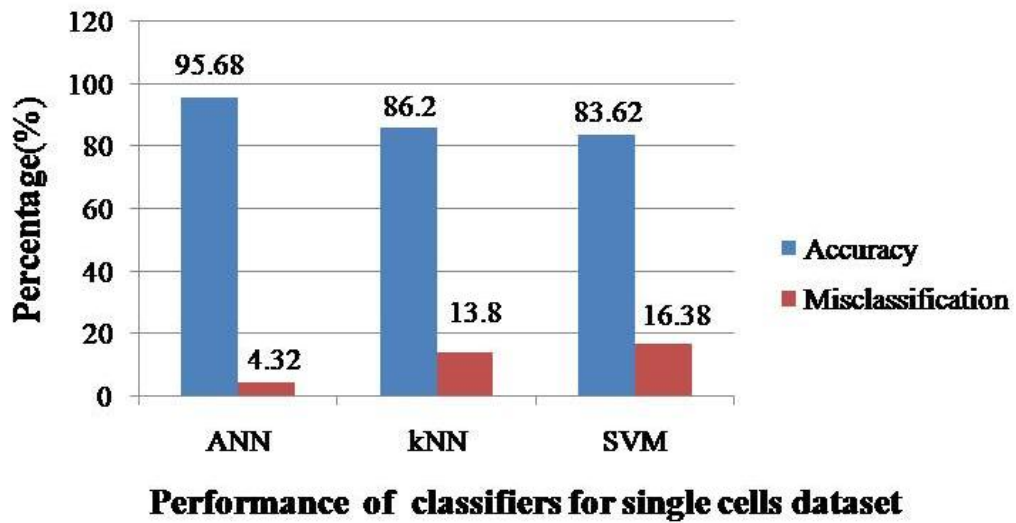

Figure 5.14: Comparison of performance graph for variation of accuracy and misclassification in single cells.

On the basis of the performance parameters, ANN yields the best performance. For group cells, data set and single cell dataset the Accuracy using ANN are 83.78% and 95.68% respectively. Further, for classifying the dataset of a group of cells, SVM classifies better than k-NN, however, for the dataset of a single cell, the classification of k-NN is better than SVM. Considering other parameters, classifying with ANN for both the datasets provides an almost better quantitative result for sensitivity, specificity, BCR, - measure, and MCC when compared to the other classifiers.

The comparative analyses of the proposed framework with other standard methods available in the literature are compared here. The different datasets applied by the researchers proposed methods, and their respective results are shown in table 5.5 Bergmeir *et al.* (2012) accurately identified having true positive rate 88% and false positive rate 19%, respectively, by using manual ground truth dataset of cancerous tissues. Landini *et al.* (2010) presented a method for morphologic characterization and best classification by NN classifier is 83% for dysplastic and neoplastic classes and 58% of correct classification.

It is observed from the table 5.5 that different classifiers have a different level of accuracy as reported by various researchers in their respective results. The accuracy of the proposed method in this chapter (for the singles cells dataset) is comparatively better than the other techniques.

Table 5.5: Comparison of the proposed method with other standard methods.

Authors (year)	Feature set used	Methods of classification	Parameters used (%)	Dataset used
Sinha and Ramkrishan, 2003	Texture and statistical	k-NN	Accuracy = 70.6%	Blood cells histology images
Huang and Lai, 2010	Texture	SVM	Accuracy = 92.8%	1000 × 1000, 4000 ×3000, and 275 ×275 HCC biopsy images
Di Cataldo *et al.,* 2010	Texture and Morphology	SVM	Accuracy = 91.77%	Digitized histology lung cancer IHC tissue images
He *et al.,* 2008	Shape, morphology, and texture	ANN SVM	Accuracy = 90.00%	Digitized histology images
Krishnan *et al.,* 2011	HOG, LBP, and LTE	SVM	Accuracy = 88.38%	Histology images Normal-90 OSFWD-42 OSFD-26
Proposed approach for group cells	Structural, Intensity and Texture	ANN k-NN SVM	Accuracy **ANN = 83.78%** k-NN = 78.37% SVM = 81.08%	Breast cancer dataset
Proposed approach for single cell	Structural, Intensity and Texture	ANN k-NN SVM	Accuracy **ANN = 95.68%** k-NN = 86.20% SVM = 68.10%	Breast cancer dataset

ROC Curve:

Six morphological features have been considered as depicts in figures 5.15 to 5.17 to plot ROC curves such as **(1)** area, **(2)** circularity, **(3)** convex area, **(4)** max intensity, **(5)** minimum intensity and **(6)** perimeter for each of the three classifiers such as ANN figures 5.15 (a) group cell 5.15 (b) single cell, k-NN, figures 5.16 (a) group cell 5.16 (b) single cell and SVM figures 5.17 (a) group cell 5.17 (b) single cell respectively. It is revealed from the curves that the features like **(1)** area, **(2)** convex area, **(3)** perimeter and **(4)** maximum intensity were providing the best Area Under Curve (AUC) values for group cell and single cells. However, the AUC values for **(b)** circularity and **(e)** minimum intensity were not so significant in group cell and single cells.
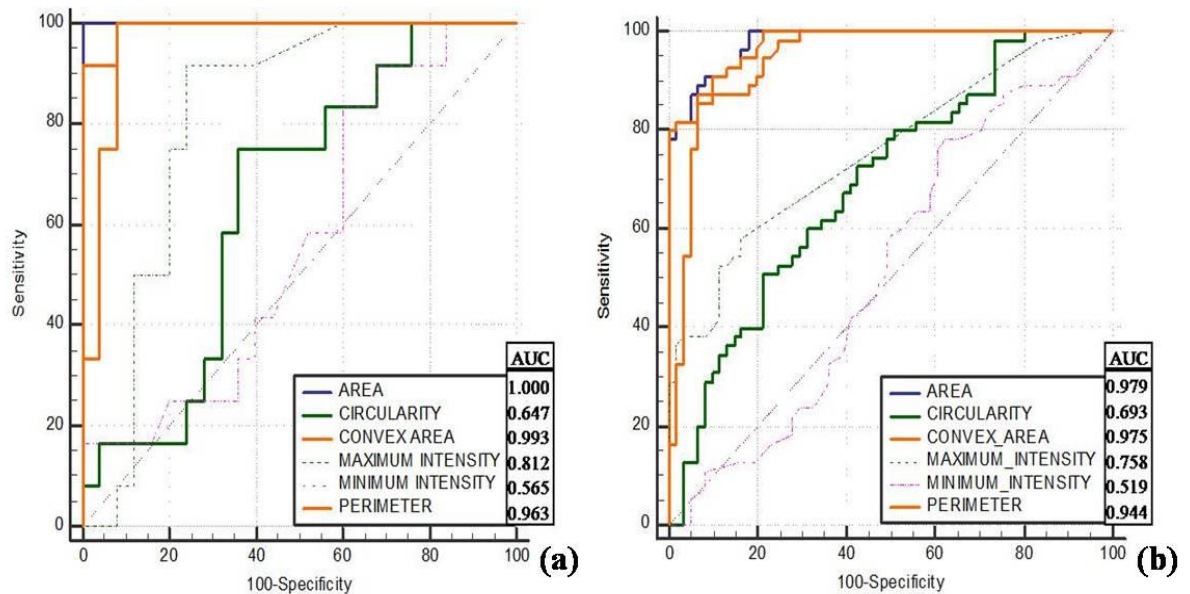

Figure 5.15: ROC curve showing performance classification of ANN.
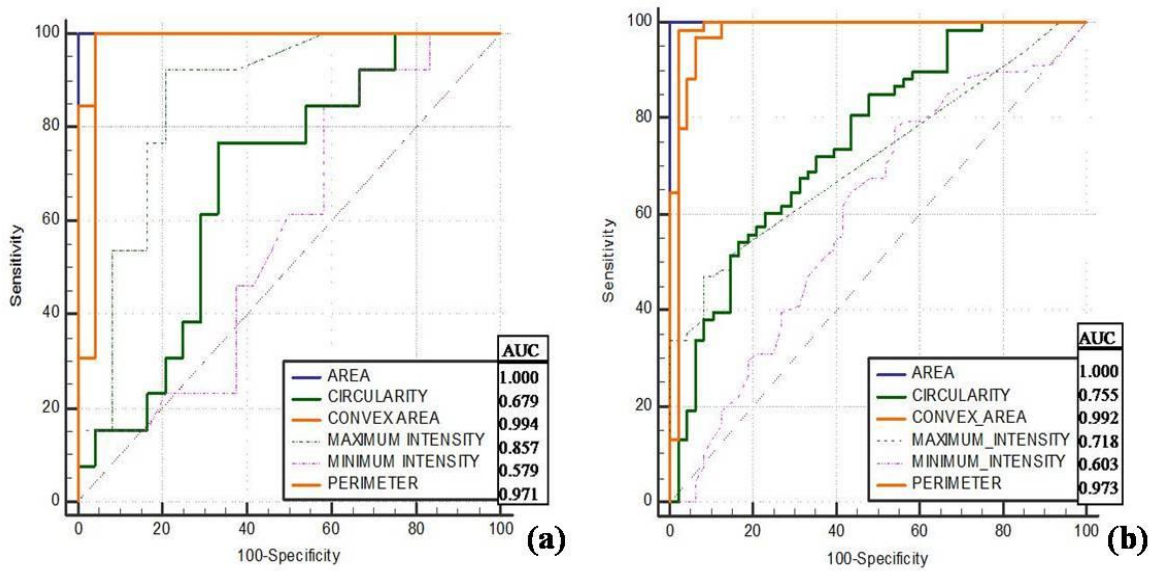(a) Group cells dataset. (b) Single cells dataset.

Figure 5.16: ROC curve showing performance classification of k-NN
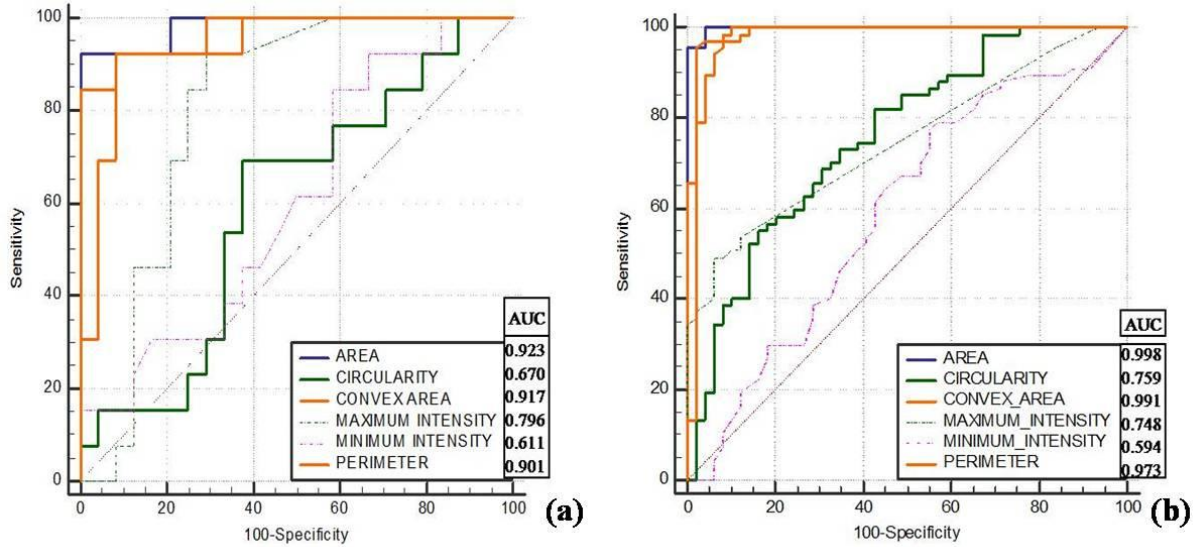(a) Group cells dataset. (b) Single cells dataset.

Figure 5.17: ROC curve showing performance classification of SVM.
(a) Group cells dataset. (b) Single cells dataset.

Graphical User Interface (GUI) Cell Check (v1.0): A research tool for cancer cells quantification

The process of handling MATLAB script for each and every image separately is problematic and hazardous, a user interface, named as CELL CHECK (v1.0) is designed to make the process of identification fast and user-friendly. With the help of the GUI, the operator can not only load, segment and classify the cancerous cell, simultaneously; data for the particular cell can be saved for future reference. Figure 5.18: (a) represents the CELL CHECK (v1.0) and it works.

Figure 5.18: Working of the Cell Check (v1.0) Graphical User Interface (GUI).
(**a**) Cell Check (v1.0) GUI (**b**) Uploading of the image into Cell Check (**c**) Benign cell detected. (**d**) Malignant cell detected.

The Graphical User Interface (GUI) has been prepared using MATLAB–GUI with ANN classifier for easy, fast and robust identification of normal and cancerous cells. The image selected in the GUI image space undergoes segmentation first, and then parameters are calculated. At the end, data will be classified and the result will be displayed. The basic features, that are provided in the GUI, are Browse (for browsing and finding an image), Crop (for cropping the image), and Check (for the segmentation, feature extraction, and classification).

The most important feature that is provided for the classification is the button 'Check'. At the time of initialization, the software automatically fetches the recorded and properly classified data (the research output) and prepares the neural network tool itself. When the software is ready, the next step is to provide cell images and with the help of this developed algorithm, they are classified into the benign cell and malignant cell. Flexibility is provided in the case of choosing any image. The contrast of the image is adjusted automatically.

The buttons which were developed have been described according to the work modules. The procedure of work with the buttons has been described too. The buttons which are provided in the GUI as basic features are described below.

Description General buttons:

Browse: It is used for browsing and selecting the image of interest.

Crop: It is used to crop the area of interest from an image.

Check: It is used to check the benign cell or malignant cell.

Check: It is used for checking the type of the image (RGB or Greyscale or Binary).

Crop: It is used to crop the area of interest from an image.

5.7 Conclusions:

In this chapter, an automatic method was proposed and a GUI is prepared for segmentation and classification of malignant and benign cells of breast cancer. It is used for supporting medical activities in recognition of cancer. The benign and malignant cells were classified on the basis of structure, intensity and texture feature using microscopic images. These features contributed significantly to realize quantification, statistical analysis, and computer aided diagnosis and develop interactive systems to detect benign and malignant tissues. ANN provides an algorithm able to learn to distinguish between classes such as normal and malignant histological images. The results obtained shows that this method could differentiate between benign and malignant cells with a good level of accuracy. In single cells, accuracy was better for classifier as compared to group cells. ANN was better performing than another classifier. The ANN is the more appropriate alternative than the earlier classification methods. It is demonstrated that the ANN is more capable of recognizing patterns than the other methods and an important tool for classification. The major advantages of ANN are that they are capable of using potential data. The process of getting data from potential information is called network learning and assigned to determine weights in ANN. It is also necessary to expand the size of the dataset to utilize the classification model for practical purposes.