



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: V Month of publication: May 2019

DOI: <https://doi.org/10.22214/ijraset.2019.5189>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Detecting Breast Cancer by using Mammography Microcalcification

Darshana. D. Tupkar¹, Madhura. M. Jadhav², Khushboo. A. Tandel³, Piyush Gawali⁴

Abstract: Breast cancer is detected based on the presence of micro calcifications in mammographic region. Mammography is one of the most effective approach in identifying the breast cancer.

However, radiologists find difficult to detect breast cancer and also it is complicated to detect it by using the existing methods such as Region of Interest(ROI), Digital Mammography techniques and many more, therefore the breast cancer remains ambiguous and becomes challenging to prevent the tumor.

Many researchers are trying to find the superior methods in order to detect the tumor region of the breast tissue. It is complicated to classify the cancer signs present in the breast tissues. Due to certain reasons the extraction of the features has become difficult by the morphological operations. In this paper, feature-based method is initiated for the purpose of the analysis and classification of the malignancy. The features will be extracted to detect the breast cancer. Machine learning algorithms consists of several classifiers to classify the images will be carried out by many classifiers such as support vector machine, naïve bayes, etc. This paper deals with the popular machine learning classification algorithm that is Naïve Bayes algorithm.

Keywords: Breast Cancer detection, Feature Extraction, GLCM, Naive Bayes algorithm

I. INTRODUCTION

Now-a-days, breast cancer is one of the most leading cause for the women death. It is necessary to detect the breast cancer [1]. Several methods are used but it is very complicated to diagnose cancer accurately. There are many complications for detecting the breast cancer due to architectural contortion, structures that are asymmetry and breast contortion patterns. [2]. For detecting the breast cancer the mammograms are widely used by radiologist. [2]

Mammography technique is one of the most popular method which is used by most of the researchers. Mammograms are classified according to the tissues that are fatty glandular and dense regions. [2] Mammography is the method that uses X-rays of low energy to diagnose the human breast. [1].

Mammography technique is standard procedure for diagnosing the breast cancer. [3] The results of the diagnosis are classified into three classes and they are normal, benign and malignant, where normal breast tissues represent the mammograms with no cells of cancerous whereas benign class represents the tumour but it is not generated by cancerous cells and malignant class represents the cancer produced by the cancerous cells. [3] Benign is not dangerous whereas malignant is the type of the tumour which causes danger to a person. In a malignant tumour, the cancer cells have the tendency to spread beyond the tissues of the breast if they are not treated at early stage. For an example, if a malignant tumour in the breast is not prevented, it may rise into the muscles that will lie under the breast. It can also lead into the skin that will cover the breast. [6] Microcalcification is the primary sign for the breast cancer detection. [1] Detecting the microcalcifications is somewhat a challenging task considered by the radiologists. [1]

Breasts that are dense, improper factors, or simple oversight by radiologists can result to the failure of microcalcification detection. [1]

In case of an image, texture analysis plays an important role in extracting the relevant information and also in classifying the image as a benign or malignant. [3]. Radiologists have independently studied about CADx schemes that employ artificial neural networks where the image features were extracted by the radiologist. [4] There are several approaches for extracting the features. This paper deals with the grey level co-occurrence matrix (GLCM) for extracting the texture from an image. [5].

This paper aims at investigating the textural properties of the breast tissues around the microcalcifications for the purpose of diagnosing the breast cancer, whether the cancer is of benign class type or malignant class type.

The set of features are compared by means of their ability in differentiating malignant from benign tissue using a Naïve Bayes classifier. Naive Bayes classifier is responsible for calculating the accuracy of the system.

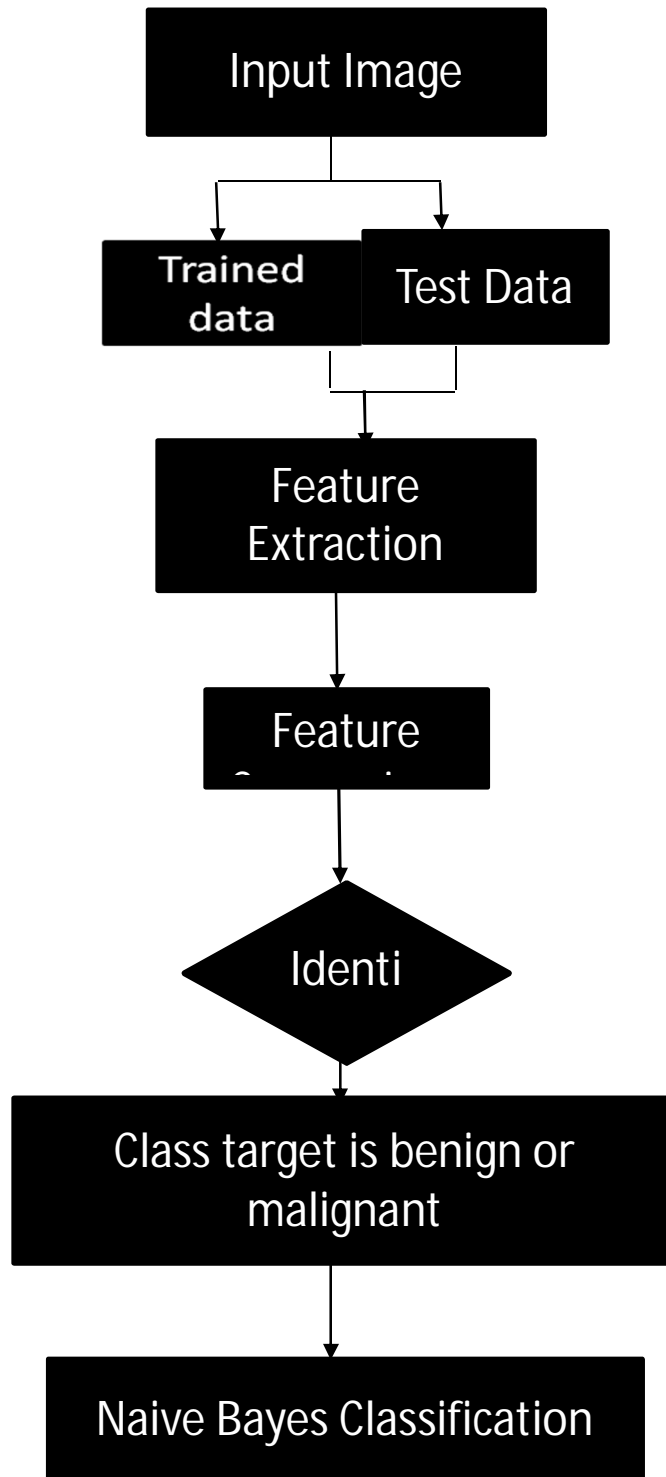


Fig1.1 Flowchart for detecting breast cancer.

II. METHODOLOGY

1) *GLCM* refers to Gray level co-occurrence matrix. Technically, a co-occurrence matrix C is defined as an $m \times n$ image I , which is parameterized by an offset $(\Delta x, \Delta y)$, as

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{otherwise} \end{cases}$$

[4].The co-occurrence matrix is often produced by using a set of offsets extending through 180 degrees (i.e. 0, 45, 90 and 135 degrees) at the same distance to get a degree of rotational invariance. After creating the GLCM symmetrical, there is one step to be taken before texture measures can be evaluated. The measures need that each GLCM should not only contain a count, but also a probability. The equation is ,

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}$$

[4] where i is the number of row and j is the number of column, i and j keep record of cells by their horizontal and vertical coordinates.

- 2) *Evaluation Of Texture Measures:* Generally texture evaluations are calculated as averages of the normalized GLCM cell contents. A weighted average is multiplied to each value to be used by a factor before dividing and summing by the number of values.
- 3) *Producing A Texture Image:* The result of a texture evaluation is a single number showing the whole window. This number will be placed in the centre pixel of the window, then this window is moved by one pixel and then this process is repeated of evaluating a new GLCM and a new textural measure. In this way an entire image is created of the textural measure.
- 4) *Contrast Feature:* GLCM diagonal consists of the values that displays no contrast, and results to the increase of the contrast that is away from the diagonal. So, create a weight that increases as the distance from the diagonal increases. Contrast is also known as sum of squares variance.

$$\sum_{i,j=0}^{N-1} P_{i,j} (i-j)^2$$

- 5) *Dissimilarity:* In case of the contrast measure, weights increases by (0, 1, 4, 9, etc.) as it moves away from the diagonal. However, in case of dissimilarity measure there is linear increase in the weights (0, 1, 2, 3 etc.). Equation for Dissimilarity is,

$$\sum_{i,j=0}^{N-1} P_{i,j} |i-j|$$

- 6) *Homogeneity:* It is also known as Inverse Difference Moment. In this case, weights decrease away from the diagonal, then the result will be greater for windows with showing little contrast. The equation for Homogeneity is ,

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$$

- 7) *Energy ASM:* Energy uses each Pij .It uses as a weight for itself. Maximum value of Energy occur only when the window is placed orderly. Energy equation is,

$$\sum_{i,j=0}^{N-1} P_{i,j}^2$$

The square root of the energy is sometimes referred as a texture measure, and is called ASM i.e Angular second moment. ASM equation is,

$$Energy = \sqrt{ASM}$$

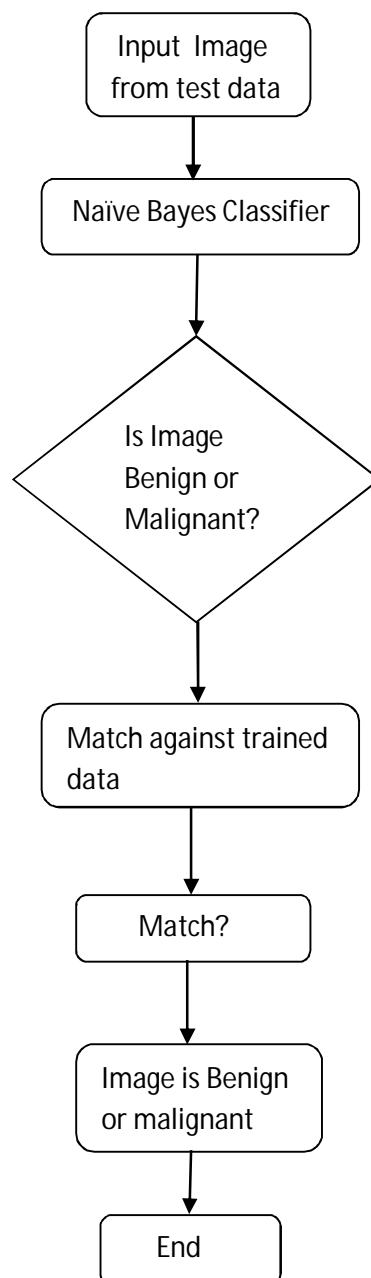
- 8) *Entropy Measure:* Entropy is considered as a measure of the suspicion that is associated with a random variable. Entropy is a measure of unpredictability. It is given by,

$$f_8 = -\sum_i \sum_j p(i,j) \log(p(i,j))$$

III. NAÏVE BAYES

Naïve Bayes is a machine learning algorithm for classification purpose. It is on the basis of the Thomas Bayes's probability theorem. Naïve Bayes is basically used for the purpose of the classification of text which involves training datasets with high-dimensional. For example, it is used for spam filtration, sentimental analysis and also for classifying news articles. This algorithm is not only popular for its simplicity but also for its effectiveness.[6] This algorithm is fast to create models and make predictions. Naïve Bayes algorithm is the algorithm that learns the probability of an object with certain features belonging to a particular group/class. The Naïve Bayes algorithm is called "naïve" because it makes the assumption that the occurrence of a certain feature is not dependent on the occurrence of other features. The "Bayes" part is known for the statistician and also for the philosopher called Thomas Bayes and the theorem was introduced by him, It gives us a procedure to evaluate the conditional probability, i.e. the probability of an event based on previous knowledge available on the events. More formally, Baye's theorem is given by the equation, $P(X|Y)=P(Y|X)P(X)P(Y)$

$P(X|Y)$: Probability (conditional probability) of occurrence of event X given the event Y is true. $P(X)$ and $P(Y)$: Probabilities of the occurrence of event X and Y, respectively





IV. CONCLUSION

A System for detecting breast cancer by extracting features and by using the naïve bayes algorithm is presented in this paper. Feature extraction of an image provide an useful information for identifying microcalcifications. Naive Bayes algorithm is used for classifying whether the image contains benign class type or malignant. The texture features are extracted using grey level method. Hence, this method detects the breast cancer.

REFERENCES

- [1] D.A.Kulkarni, G.R.Udupi Bhagyashree.S.M, , "Texture Analysis of Mammographic images", August 2010.
- [2] K. Rajendra Prasad, C. Raghavendra, K Sai Saranya, "A review of classification of breast cancer detection using combination of feature extraction models", 2017, 203-208 .
- [3] Hashem B. Jehlol, Zainab khyioon ,Anwer Subhi Abdulhussein Oleiwi, "Classification of Mammography Image Using Machine Learning Classifiers and Texture Features", September 2015.
- [4] Monika Sharma, R. B. Dubey, Sujata, S. K. Gupta, "Feature Extraction of Mammograms", September 2012.
- [5] T. Suneetha Rani, S J Soujanya, Pole Anjaiah, "Breast Cancer Detection Using Combination of Feature Extraction Models".
- [6] Vikas Chaurasia, Saurabh Pal, BB Tiwari, "Prediction of benign and malignant breast cancer using data mining techniques", February 2012.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)