International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Breast Cancer Detection System

Pranay Shah[1], Rahul Deshpande[2], Nikhil Rao[3]

[1]B.E. student, Dept. of Information and Technology, Atharva college of Engineering, Maharashtra, India
[2]B.E. student, Dept. of Information and Technology, Atharva college of Engineering, Maharashtra, India
[3]B.E. student, Dept. of Information and Technology, Atharva college of Engineering, Maharashtra, India

---

Abstract - *The rapid development in deep learning has being beneficial to me and many others families. Typical breast cancer detection approaches and processes and cumbersome, expensive and quite slow. Due to this valuable lives are lost as people do not get results on time, diagnoses are inaccurate or the price seems to deter them away. We have developed a algorithm which will accurately detect breast cancer results on Cell Biopsy Images This paper aims to serve as a starting point for those who are not acquainted with this growing field.*

Key Words: **Deep Learning**, **Biopsy Images**, **Breast Cancer Detection**, **Disease Detection**, **Image Classification**, **Machine Learning**.

1.INTRODUCTION

Due to various factors such as latency, cost and accuracy, breast cancer detection is not as effective as it can be. Using deep learning and AI we can improve the quality, accuracy and make breast cancer detection cheap. Multiple tests have to be done which are time consuming and expensive. These tests are also uncomfortable and awkward for the women patients as a physical examination is required. In the view of these issues, It is necessary to adopt a better system of detection which is much more reliable, hassle-free and comfortable for the patients. This paper contributes to develop an efficient method to detect breast cancer without involving any pain to women using the power of Artificial Intelligence and Deep Learning.

2. LITERATURE SURVEY

First paper which we have researched was about the method which they are using is LBP that is known as Local Binary Patterns where they are converting RGB Images into Grayscale and later on they are comparing using 3x3 Matrix pixel values by using Mammograms images and there research shown us that using LBP they got accuracy of 84%. So there method was firstly to remove the unwanted background and get breast image from mammograms for this method they have used OTSU method. Second was to calculate Local Binary Patterns for computing feature vectors where the LBP operator uses 3x3 matrix. They assign 0 or 1 values on the basis of the central value of the neighborhood values if center value is greater than neighbor value, then its value is denoted as 0 if counter value is less than neighbors value, then it is denoted as 1. Third step is as featured vector's values has been received then they are inserted into histograms so that same values can be combined and after that it gets results as 59 feature vector size. Final step is that this histogram is being compared with trained model and trying to get match on it. Accuracy received by this was of about 84% [1]. Detecting macrocalcification in dense breast tissue can be a daunting task as both appear to reflect white pixel on the mammogram. The number of false positives on dense breast tissue is greater. The cancer symptom markers are in general, masses and microcalcifications. Mass detection is a more complex job than micro-calcification detection. As their size and shape vary widely but they often show poor contrast picture. In medical diagnosis, the use of grouping structures in the classification and pattern recognition method is increasing rapidly. A main aspect is evaluation and decision-making focused on machine-learning for medical diagnosis. Mainly texture characteristics and statistical features are more relevant in pattern recognition field. Mammography screening is the simplest and most reliable tool for breast cancer diagnosis. Several techniques are used to test the mammography image such as finding edges, smoothing border, finding structures & shapes among matrixes. Finally finding the tissue size distribution in an Image without specifically segmenting each material. Digital mammography is the normal treatment for the diagnosis of breast cancer; on the digital mammography image, various classification problems are applied. Both features are measured from the vulnerable part of the breast to eliminate any unintended features that could impact the question of classification. The Maximum Likelihood Estimation (MLE) measures the region of tumour. All extraction techniques for the functionality are applied to the extracted image of the database. This paper studies mostly the multiple image processing algorithms that can be used extensively to find cancerous cells. The computer-aided mammography techniques include preprocessing of images, segmentation of images, extraction of features, selection and classification of

features. In order to have a better understanding of them, further developments are needed to extract more features to find pattern in tumour. Texture analysis method can be used to classify the micro-calcification in mammography between benign and malignant masses by means of [2]. Modern ML research shows the limited application of the techniques, while the DL methods have tremendous potential for clinical research implementation and improve the diagnostic capability of current CAD systems. In the field of image processing a lot of work has been conducted to identify the cancer. Nevertheless, the accuracy rate is between 75% – 92%. Therefore, an accuracy range of 8 to 25 percent also remains to be reached. New scientific methods and strategies for identifying cancer cells and methodology for eradication to cure every person's cancer. Yet also cancer cells developed to hide from medications and medicines. Because cancer cells are immortal the immune system does not affect them. Mammography images are selected by scale of 1024X1024 and converted into 2D matrix grayscale image. All images in the database are of the same size, if the image size varies from other images then the image enchantment algorithm is applied to match the image resolution. Through Noise removal algorithm these images are filtered. Instead, they are filtered and modified to increase pixel strength. The image goes through the process of thresholding with the intention of segmenting the image ROI. To delete the unwanted part of the image and segment the portion with higher pixel density a Global thresholding value can be applied. Using Histogram to search for distribution of pixels. Thus, the best ROI of the image is acquired using value from all of these fields. Because each image is taken from various angles of the patient, the different characteristics of such images can vary from one another. Also, the patient's past mammography or future mammography image might tell a lot about the details of the tumour. The image's Area of Interest is derived from the image in mammography. The various methods of edge detection Sobel, Canny, Prewitt, Roberts and Fuzzy logic are applied to detect the edges of tumor cells. In terms of imagery, texture can be defined as the arrangement of space and there is a variation of intensities (gray values) within the image. Texture characteristics were very useful in defining the micro-calcification. One may use the image acquisition to create two separate regions. The first region is the central area of the breast, the thickness of which is almost uniform, and is called the constant thickness region. The other consists of tissue near the edge of the breast where, due to the breast geometry, the thickness gradually taperes. Features are to be extracted for region that acquires accurate results of classification from the segmented image. The Gray Level Competition Matrix (GLCM) is generated from the image's ROI, which is a gray-level competition matrix by calculating how often pairs of pixels with particular qualities occur in an image. The Spatial Gray Level Dependence Matrix (SGLD) also known as the ROI competition matrix can be used for the classification of benign and malignant cells. Such functions have different components with different values [3]. Three pre-processing algorithms are initially applied by researching in this paper: First, image inhomogeneities are corrected using the N3 bias field correction algorithm to correct variations in signal intensity within the same structure of a particular case. The N3 is a nonparametric approach intended to be implemented at the early stages of automated data processing, and does not include a tissue intensity model. Second, it detects the sternum which is used in various sections of our algorithm as an important landmark. Third, MR picture intensities are adjusted to account for the variation in interpatient signal strength. Separating the body from the breast begins with segmentation. The body is made up of lungs , heart, pectoral muscle, thoracic area and fat outside breast base. The breast is related to the pectoral muscle and is composed of thick, fatty tissues. A probabilistic atlas containing spatial details about the pectoral muscles, lungs, back, thorax and breast tissue is used to exclude the body from the breast. Finally, the volume of the breast is determined and the dense tissue is delineated using the EM algorithm. The sternum is a bone located between pectoral muscles, and is always visible in axial slices centered. The use of a breast coil makes sure the sternum is roughly in the middle of the picture [4].

3. PROPOSED SYSTEM OF PROJECT

A new approach for classifying texture features through neural networks is in the proposed system. The main purpose of this is to generate highly accurate lump picture texture functionality. Such characteristics are fed into a neural network, which classifies the images as malignant or benign. The neural network is trained using the Keras and Tensorflow algorithms for image adjustment and learning from it. To get at actual results we need to follow some steps which are as follows:

3.1 Image Resizing

Original Image is being passed through the system and after that, the image is being converted into 28*28px that is resizing as it takes much more time to test each and every pixel so resizing helps to reduce time and provide us efficient results without disturbing any pixels of image.

3.2 Feature Extraction and Selection

The extraction of features is a very important process in classifying micro-calcifications for the overall device output. The extracted features are differentiated according to the extraction method and the properties of the image. The features implemented here are texture characteristics and statistical measures such as Mean , Standard deviation, Variation, Smoothness, Skewness, Uniformity, Entropy and Kurtosis.

3.3 Classification and Evaluation

Evaluation is carried out on the basis of the acquired features and to draw the final conclusion, these features are compared with the respective reference.

3.4 Neural Network

Both of these texture feature values are processed and transferred through the Neural Network. Keras and Tensorflow algorithm can be used to identify a trend to automatically locate a cancer within the datasets. The Keras and Tensorflow algorithm can be built to self-learn and produce results as necessary. As more data gets entered in Neural Network the better the pattern recognition and accuracy.

3.5 Block Architecture

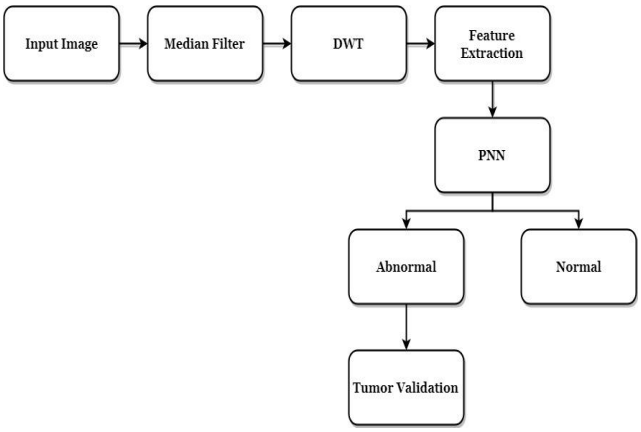

Fig - 3.5: Architecture

4. OUTPUT

Fig - 4.1: Malignant(Cancer Present)

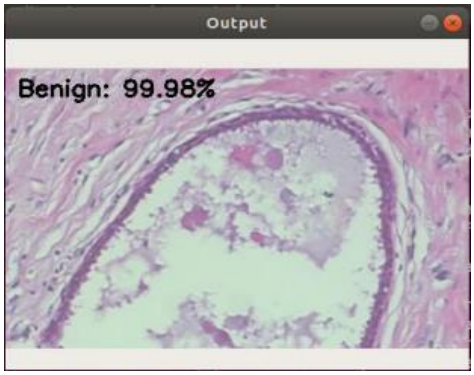

Fig - 4.2: Benignant(Cancer Absent)

5. CONCLUSIONS

Our system is much powerful and accurate as compared to others with the highest accuracy of above 98% and much faster.
The program is proving cost efficient and needs less maintenance. Therefore, many doctors can treat this disease with less equipment, and since the total treatment time is reduced, more people can be treated.
The benefit of this program is that the user does not have to perform several tests but just one accuracy test.
The output of this system is being generated with Malignant and Benignant.
This program is both efficient and user friendly. We learned how to use various techniques such as machine learning , deep learning, image processing and Python to create an impressive, deeply personal project. We have also learned about the description of pictures and their applications in medicine. Last but not least, we've learned how to save lives by using technology for good.

REFERENCES

[1] Pavel Král, Ladislav Lenc. "LBP features for breast cancer detection".

[2] R. Sangeetha Dept. of CSE, PES Institute of Technology, South Campus Bangalore, India, K. Srikanta Murthy. "Novel Approach Of Breast Cancer Detection".

[3] Syed Jamal Safdar Gardezi,Ahmed Elazab, Baiying Lei, and Tianfu Wang. "Breast Cancer Detection and Diagnosis Using Mammographic Data".

[4] Prannoy Giri and K. Saravanakumar . "Breast Cancer Detection using Image Processing Techniques".

[5] Mussarat Yasmin, Muhammad Sharif and Sajjad Mohsin: "Survey Paper on Diagnosis of Breast Cancer Using Image Processing Techniques," Research Journal of Recent Sciences, Vol. 2(10), 88-98, October (2013).

[6] K.Saravana Kumar and A.M. Arthanariee "Evaluate the Multiple Breast Cancer Factors and Calculate the Risk by Software Tool Breast Cancer Risk Evaluator" Indian Journal of Science and Technology, Vol 8(S7),686–691, April 2015