International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

BREAST CANCER DETECTION USING MACHINE LEARNING

Geetha P, Sneha R, Sneka.S, Subhiksha.R

Electronics and Communication Engineering, Hindusthan College of Engineering And Technology, Coimbatore , Tamil Nadu , India

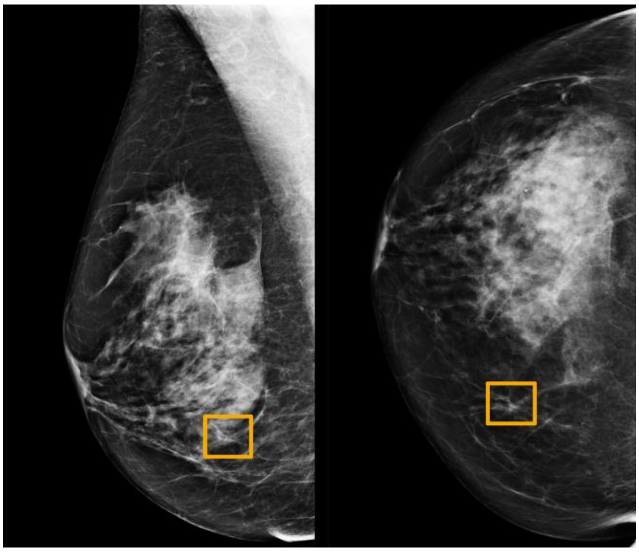
---***---

ABSTRACT

According to global statistics, breast cancer (BC) is one of the most frequent diseases among women globally, accounting for the majority of new cancer cases and cancer-related deaths, making it a serious public health issue in today's society. Early detection of BC improves the prognosis and chances of survival by allowing patients to receive timely clinical treatment. Patients may avoid unneeded therapies if benign tumours are classified more precisely. The classification method utilised in this paper is Modified CNN based feature extraction with transfer learning (ANN). A gaussian kernel approach is utilised for segmentation expectation maximisation .

INTRODUCTION

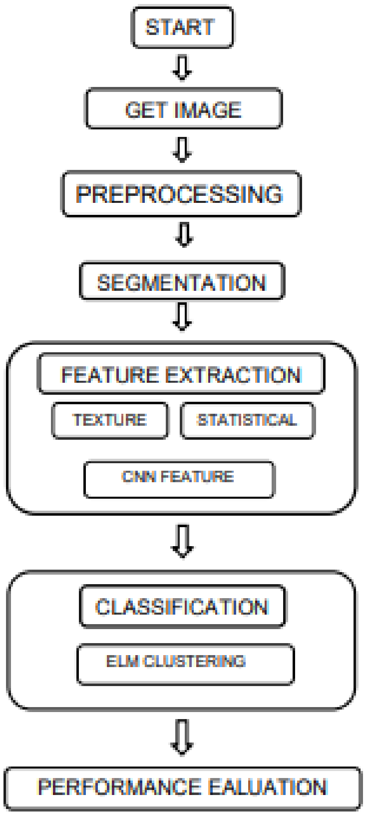
Human cancer is a multifaceted illness characterised by genetic instability and the accumulation of numerous alternative molecules. Current diagnostic categories do not account for the whole clinical heterogeneity of malignancies and are insufficient to predict treatment efficacy and patient outcomes. The majority of currently used anti-cancer drugs do not distinguish between malignant and normal cells. Furthermore, cancer is frequently discovered and treated too late Cancer cells have already spread throughout the body. A large majority of individuals with breast, lung, colon, prostate, and ovarian cancer have hidden and over metastatic colonies at the time of clinical presentation. The effectiveness of therapy techniques is currently restricted. Use deep learning algorithms for breast cancer detection with mammography images in our planned effort.



EXISTING METHODOLOGY

Breast image preprocessing, mass detection, feature extraction, training data creation, and classifier training are the five phases in breast cancer detection covered in this work. To raise the contrast between the masses and the surrounding tissues, de-noising and boosting contrast techniques on the original mammography were used in the breast image pre-processing. After that, mass detection is used to locate the ROI. The ROI is then used to extract characteristics such as deep features, morphological features, texture features, and density features. The classifiers were trained with every image from the breast image dataset during the training phase, using their extracted features and labels. The well-trained classifier can so identify the mammography under diagnosis. A contrast enhancement method was utilised in this study to improve the contrast between the suspicious masses and the surrounding tissues. The fundamental concept is to make the original image's histogram equally distributed. The image's grey scale is expanded as a result of this procedure, which improves contrast and makes image details more visible.

EXISTING FLOW CHART



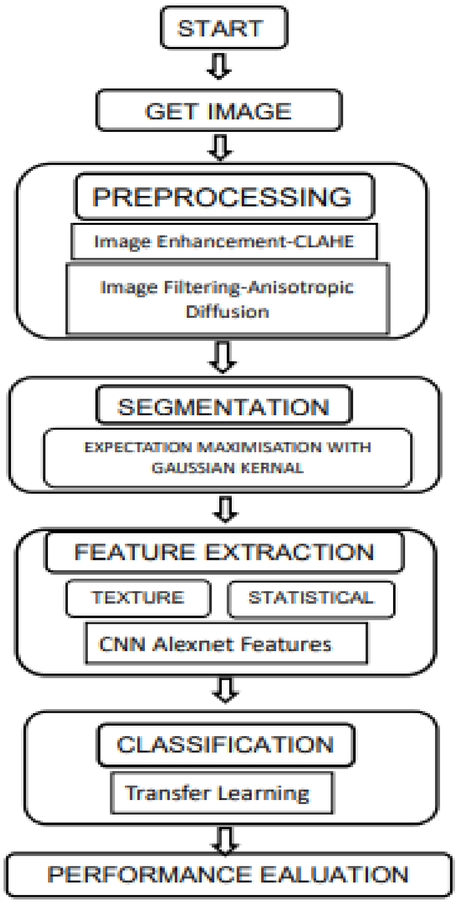
DISADVANTAGES

1. Cancer detection accuracy is lower.

2. Time it took to extract the mass's form features.

PROPOSED METHODOLOGY

PROPOSED FLOW CHART



PROPOSED WORKFLOW

Step1: Image for Input (Training and Testing)

Step 2: Reduce the input size Image

Step 3: Filtered using Anisotropic Diffusion Filter

Step 4: Use Contrast Limited Adaptive Histogram Equalization to improve the images.

Step 5: Expectation Maximization with Gaussian Kernels is used to segment the tissue component.

Step 6: Extract Statistical and Texture Features

Step 7: Load Training Images to CNN and retrieve features from the Fully Connected Layer before storing them.

Step 8: For both training and testing images, combine the features acquired from the segmented picture with CNN Layer Features.

Step 9: For those features with Appropriate Targets, use Transfer Learning.

Step 10: Get Result from Transfer Learning.

PREPROCESSING

We must improve and denoise our input image after reading it. To raise the standard. Here, we're employing algorithms like,

Image Enhancement: Equalization of contrast-limited adaptive histograms

Image Filtering: Anisotropic Diffusion Filter

CLAHE

CLAHE is a variation of adaptive histogram equalisation (AHE) that corrects for contrast overamplification. CLAHE works on tiles, which are small areas of an image rather than the complete image. To remove the false boundaries, the surrounding tiles are blended using bilinear interpolation. The contrast of photographs can be improved with this approach. Ordinary histogram equalisation transforms all pixels with the same transformation determined from the image histogram.

SEGMENTATION

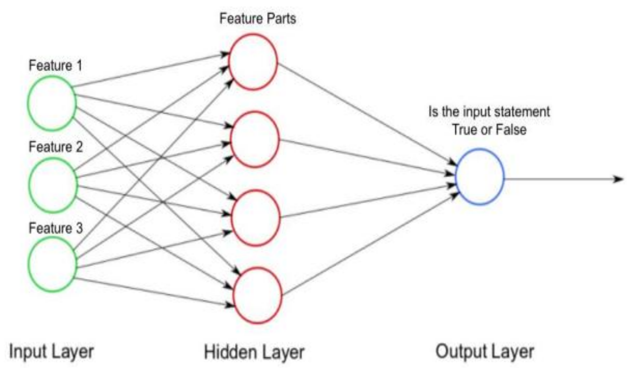
Segmentation is the process of splitting a picture into regions that are homogeneous in terms of specific characteristics such as colour, intensity, and so on. It is an important stage in picture analysis since it locates objects and lines (lines, curves etc). The cancer component of the Breast is extracted using expectation maximisation with Gaussian kernels.

FEATURE EXTRACTION

In many image processing applications, feature extraction is the first step. The main tool for image texture analysis is GLCM. To extract textural properties including entropy, contrast, correlation, and homogeneity, we use the Gray Level Co-occurrence Matrix (GLCM) approach. For classification, we additionally extract statistical and form features.

CLASSIFICATION

Transfer Learning is used for classification. First, features from the CNN layer are retrieved and hybridised with GLCM and statistical features before being put into transfer learning to classify normal and abnormal breast tissues. Artificial neural networks (ANNs) or connectionist systems are computing systems based on biological neural networks seen in animal brains. The neural network provides a framework for several distinct machine learning algorithms to operate together and process complicated data inputs, rather than an algorithm itself. Such systems "learn" to execute tasks by considering examples, usually without any task-specific rules being coded. /They might learn to identify photographs that contain cats, for example, by evaluating sample images that have been manually classified as "cat" or "no cat" and then utilising the results to detect cats in other images. They do so without knowing anything about cats, such as the fact that they have fur, tails, whiskers, and cat-like faces. Instead, they generate identifiable traits from the learning content they analyse automatically.



ADVANTAGES

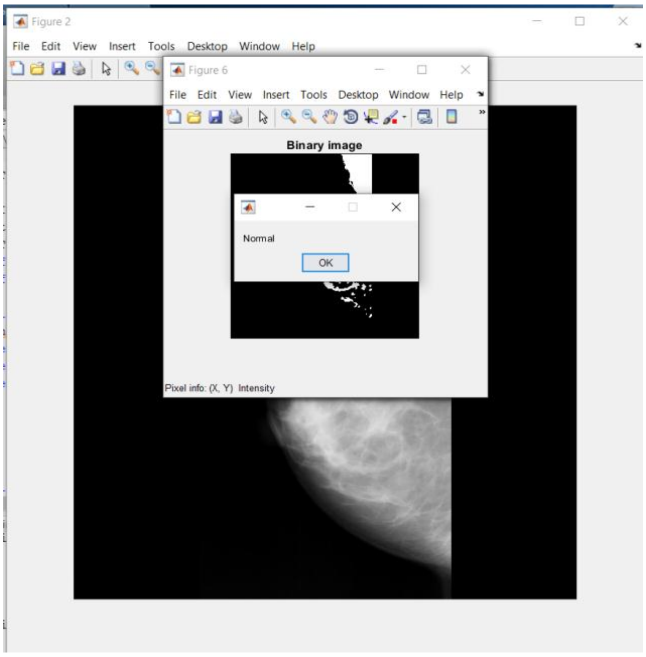
1. High Precision

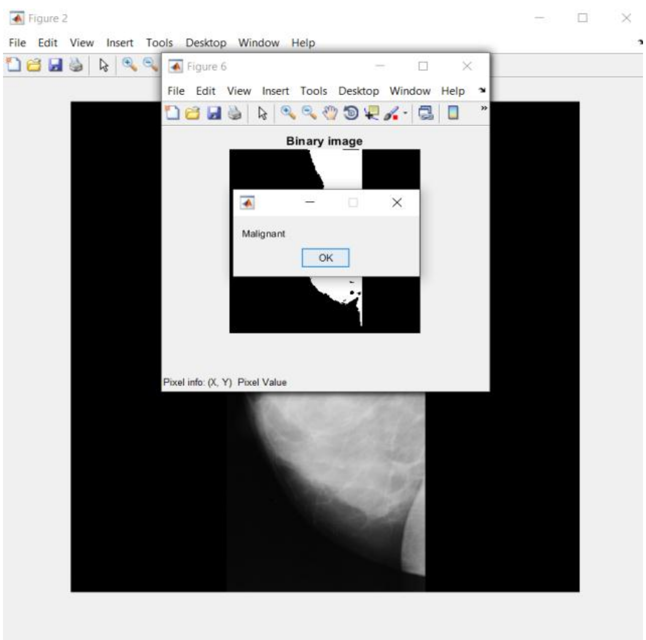
2. The EM-based segmentation Image yields precise texture details.

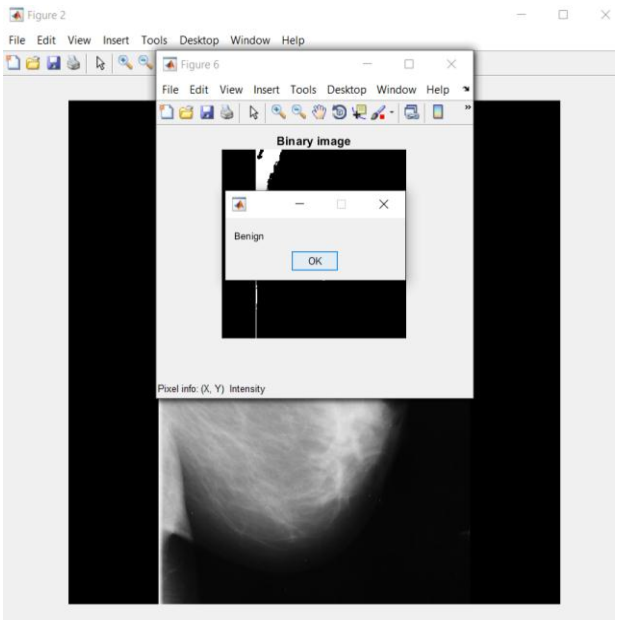
3. The greatest result for detecting malignant illness is Transfer Learning.

RESULTS

This section displays the results of the mammographic database's experimental results. Because of the fast-paced nature of medical science, research and technology are regularly updated, yet advancement in the diagnosis and treatment of cancer tissue remains basic. Because of the vast range of symptoms associated with breast abnormalities, certain abnormalities may be ignored or misconstrued throughout the diagnosing process. There are also a lot of false positive results, which means a lot of unneeded biopsies may be performed. To assist radiologists in making an accurate diagnosis and reducing the amount of false positives, computer-aided detection and diagnosis algorithms have been created. The typical steps in image processing algorithms have been thoroughly investigated in this paper. Picture pre-processing, image segmentation techniques, feature extraction, feature selection, classification techniques, and mammogram features are some of the techniques used in computer aided mammography. To discriminate between normal and malignant cells, texture features are collected. The confusion matrix is used to calculate the accuracy. To calculate accuracy, use the formula below. With mammography pictures, the updated feature extraction with deep learning algorithm achieves 96 percent accuracy.







CONCLUSION

Early detection of BC improves the prognosis and chances of survival by allowing patients to receive timely clinical treatment. Patients may avoid unneeded therapies if benign tumours are classified more precisely. We learned how to develop a breast cancer tumour predictor using the mammography pictures collection and the findings in this MATLAB project. When it came to distinguishing benign from malignant tumours, Breast-Dense Net offered extremely accurate results. As a result, the predictor could be utilised as a second opinion to help the radiologist make a diagnosis.

Our findings show that multi-stage transfer learning can benefit from information learned through unrelated and related domain source activities. We show that pre-training the CNN with data from similar auxiliary domains can ameliorate the limited data availability in a target domain. We also show that the boost in CNN performance from the additional stage of fine-tuning with auxiliary data is dependent on the relative sizes of the available training samples in the target and auxiliary domains, as well as the transfer learning approach chosen correctly. In addition, when the training sample size is limited, the variance in the trained CNN's performance is high. Exhaustive search utilising a "test" set to report the greatest performance can be excessively optimistic. As a result, it's critical to test the generalizability of the trained CNN using independent unknown cases.

FUTURE SCOPE

Deeper architectures, as well as ultrasound, histopathology, and PET imaging, will be used in the future to address difficulties created by mammography of dense breasts. Other imaging modalities should be used in conjunction with mammography during the learning process to assist develop a robust breast mass predictor.

REFERENCES

[1] Arpita Joshi and Dr. Ashish Mehta "Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer" (2017).

[2] David A. Omondiagbe, Shanmugam Veeramani and Amandeep S. Sidhu "Machine Learning Classification Techniques for Breast CancerDiagnosis" (2019).

[3] Kalyani Wadkar, Prashant Pathak and Nikhil Wagh "Breast Cancer Detection Using ANN Network and Performance Analysis with SVM" (2019).

[4] Anji Reddy Vaka, Badal Soni and Sudheer Reddy "Breast Cancer Detection by Leveraging Machine Learning" (2020).

[5] Monika Tiwari, Rashi Bharuka, Praditi Shah and Reena Lokare "Breast Cancer Prediction using Deep learning and Machine Learning Techniques".

6] Abdullah-Al Nahid and Yinan Kong "Involvement of Machine Learning for Breast Cancer Image Classification: Asurvey" (2017).

[7] K.Anastraj , Dr.T.Chakravarthy and K.Sriram " Breast Cancer detection either Benign Or Malignant Tumor using Deep Convolutional Neural Network With Machine Learning Techniques " (2019).

[8] S.Vasundhara , B.V. Kiranmayee and Chalumuru Suresh "Machine Learning Approach for Breast Cancer Prediction" (2019).

[9] Muhammet Fatih Ak "A Comparative Analysis of Breast Cancer Detection and Diagnosis Using Data Visualization and Machine Learning Applications" (2020).

10] Sivapriya J, Aravind Kumar V, Siddarth Sai S, Sriram S "Breast Cancer Prediction using Machine Learning" (2019).