International Journal for Research in Science Engineering & Technology (IJRSET)

DEEP LEARNING BASED BREAST CANCER DETECTION USING ULTRASOUND IMAGE

[1] **Heet Vora**, [2] **Manav Malavia**, [3] **Manav Hirey**
[1, 2, 3] **Student,**
[1, 2, 3] **K.J. Somaiya College of Engineering,**
[1, 2, 3] **Mumbai India.**

ABSTRACT - In the fields of healthcare and bioinformatics, the definition of breast cancer has been the topic of concern, since this is the second principal explanation for death from cancer in women. This form of study will only be achieved if a sample of breast tissue is extracted from the breast, evaluated, and analyzed under a microscope. In the histopathology lab, problems are detected by examining the specimens using trained and qualified pathologists, and identified with further investigation using special techniques. However, the ultrasound can incorrectly identify pathological changes or conditions because they have practiced in this specialty before. Several pattern recognition studies have recently identified a lot of areas for potential for improvement, and, hence there is now an increased focus on constructing powerful image processing experiments to create a highly-and improving existing diagnoses. Let's use the histology and image recognition techniques to identify the disease types of breast through the use of the image feature extraction technique and deep learning approaches for the image feature. This picture can be expanded using ultrasound processing and CNN technique is available before ultrasound feature extraction is applied and the final classification in ultrasound feature extraction.

Keyword - [image classification, breast cancer diagnose, feature extraction, classification, deep Learning.]

1. INTRODUCTION

The major consequence of death from cancer in women is intrusive breast cancer, after lung cancer. The World Health Organization (WHO) International Agency for Research on Cancer (IARC) is only responsible for about 8.2 million deaths in 2012 due to cancer. By 2030, the number of cases new is expected to increase to over 27 million.

The main course of action to reduce breast cancer deaths is the fast detection of breast cancer and state-of-the-art cancer care. Hematoxylene and eosin (HE) stained histological diagrams which are examined under a high power microscope in a modified region of the breast are a typical method of detecting breast cancer. In surgical practice it is experienced pathologists who are manually responsible for classifying breast cancer biopsy into various planes(e.g. cancerous and noncancerous). Come out machine learning approaches and enlarging image volume developed automatic system for breast cancer classification possible and can help

pathologists to obtain precise identification of problem more efficient. Health photographs testing for histology and images in radiology may be used to detect or identify breast cancer. The quest of radiology pictures will help to identify the location of the gap. They cannot however be used to find or to decide whether the region is cancerous. The most sure means of seeing whether a region is cancerous is a biopsy in which a tissue is used as an input and processed under a microscope to see if cancer is present. Only the histologists who test a tissue under the microscope, searching for exceptional or cancer cells, will be able to detect the problem after the biopsy has been completed. The histological images enable us, through a particular pattern, to distinguish between cell nuclei types and their flowchart. The quality of cell forms and tissue distributions was investigated by ultrasound and the cancerous regions and degree of malignancies is determined. In addition, the experts who kept the tissue sample up to two months are lacking. The reproducibility problem still exists, as histopathology is a subjective discipline. This is particularly true for non-specialist pathologists, where on the same sample we can identify a different concern. Consequently, there is an ongoing request to identify a challenge using computers.

2. REVIEW OF LITERATURE

Breast cancer (BC) is a vicious disease that kills a large number of people every year. Developing a robotized harmful BC recognition framework based on the patient's symbols can aid in more successfully controlling this issue, making diagnosis more adaptable and less prone to errors. DeCAF (or deep) highlights are an in-between arrangement that relies on reusing previously trained CNNs as highlight vectors, which are subsequently used as input for a classifier created specifically for the new order assignment. In light of this, they present a comparison of DeCaf highlights for BC recognition, with the purpose of better understanding how they compare to alternative techniques [1].

Using convolutional neural networks, this study proposes to classify breast cancer histopathology photos regardless of magnification (CNNs). They suggest two architectures: a single task CNN for predicting malignancy and a multi-task CNN for predicting both malignancy and picture magnification level at the same time. The BreaKHis dataset [2] is used for evaluations and comparisons with earlier results.

The goal of this project is to develop a useful remote discovery and detection technique for breast disease based on cytological images. Initially, this work shows a fully automated method for recognising and dividing cell nuclei in bosom cytological images. With a roundabout Hough modification, the areas of the cell cores in the image were detected. Otsu's thresholding process and fluffy c-implies grouping strategy were used to eliminate false-positive (FP) discoveries (loud circles and platelets). With the help of the marker-controlled watershed alteration, the nuclei limits were successfully divided. Following that, an intelligent breast malignancy categorization framework was developed [3].

The success of breast cancer treatment is dependent on early detection. The cytological study of breast material obtained directly from the tumour is an early milestone in the discovery. In view of the analysis of cytological pictures of tiny needle biopsies to recognise this biopsy as either benign or detrimental, this work provides identifiable proof of issue in PC aided breast growth. Instead of conferring on the precise division of cell nuclei, the nuclei are found by circles using the roundabout Hough change approach. The result circles are then sorted to maintain only the most amazing estimations for further consideration by an assistance vector machine, which classifies identified circles as correct or incorrect based on surface highlights and the amount of cores pixels as determined by Otsu's thresholding approach [4].

This work directs some basic investigations employing the deep learning approach to deal with organise breast cancer histological photographs from BreaKHis, a publicly dataset accessible at http://webinf.ufpr.brivri/bosom malignancy database. They suggest an approach for preparing the CNN and mixing the picture patches for definite grouping based on the extraction of picture patches. This method enables the use of BreaKHis high-goals histopathology images as a contribution to existing CNN while keeping a strategic distance from model alterations that could lead to more unpredictable and computationally expensive engineering [5].

In arranging two regions, current techniques rely on handcrafted highlight representation, such as shade, surface, and Local Binary Patterns (LBP). In contrast to meticulously prepared include-based methods that require subordinate portrayal, DCNN is an end-to-end highlight extractor that may be easily obtained from crude pixel force estimate of EP and ST tissues in an information-driven mould. The creation of a directed classifier for differentiating the two types of tissues is aided by these abnormal condition highlights [6].

The test proves to be a useful tool for combining fix level arranging results and modelling how not all patches will be discriminative. They propose developing a choice combination model for total fix level forecasts generated by fix level CNNs, which, to our knowledge, has never been done before. They use the method to divide cases of glioma and non-little cell lung cancer into subtypes [7].

Computerized atomic identification is a fundamental advancement for a variety of PC-assisted pathology-related image examination calculations, such as the automated evaluation of breast cancer tissue samples. However, computerised core location is complicated by (1) the large number of nuclei and the size of high-resolution digitised pathology images, and (2) the inconsistency in nuclei size, shape, appearance, and surface. Recently, there has been a lot of buzz about using "Profound Learning" approaches to organise and investigate massive amounts of image data [8].

This study presents a dataset of 7,909 breast cancer (BC) histopathology images obtained from 82 patients, which is now freely available at http://web.inf.ufpr.br/vri breast-cancer-database. Both benign and cancerous images are included in the dataset. The project using this dataset is the automated classification of these images into two classifications, which would be a useful PC-assisted finding tool for clinicians. author provide some preliminary results obtained with state-of-the-art image categorization methods [9] to assess the difficulty of this project.

There are still certain difficulties with traditional individual Breast Cancer Diagnosis. An individual credit appraisal display based on the assist vector order technique is offered to address the concerns. Individual credit information is bunching investigated by Support Vector Machine using the SPSS Clementine information mining device. It is thoroughly explored with respect to the various part capacities and characteristics of the Support vector machine. The boost vector machine could be used to improve the methods used by medical professionals to determine breast growth [10].

Using algorithms, people can make use of the information contained in this data to predict the risk of diabetes in the human body. This system uses a number of classification models, including Decision Tree, Artificial Neural Networks, Logistic Regression, Association rules, and Naive Bayes. The Random Forest approach is then used to find the overall project accuracy. This project uses the Pima Indians Diabetes Data Set, which contains patient information. Some of the patients have diabetes, which means they're at risk of developing the condition, so this project is made to create a mobile application that can assess their risk of developing diabetes and cancer. The classifiers were optimized for cancer and diabetes prediction using several classification algorithms [11]

These systems enable the user to make use of algorithms in order to predict the likelihood of developing diabetes in the human body. These models, which include Decision Trees, artificial Neural Networks, Logistic Regression (with association rules), and Naive Bayes (with no association rules), are utilized in this system. Once each model in the system has been identified, the Random Forest technique is used to determine its accuracy. The dataset used in this project is the Pima Indians Diabetes Data Set, which contains information about patients, some of whom are developing diabetes. As a result, this project aims to develop a mobile application for predicting whether a person is at risk for diabetes and cancer or not. In this study, author analyzed medical data using a variety of classification algorithms in order to improve the performance of classifiers for cancer and diabetes prediction, among other things [12].

3. PROPOSED METHODOLOGY

Classifying breast cancer ultrasound images automatically is an important task in computer assisted pathology analysis. However, extracting informative and non-redundant features for ultrasound image classification is challenging.

In this work using ultrasound image, firstly apply image pre-processing technique to remove the noise of an image. After that apply the feature extraction process. The feature-based approaches consist of the features extraction phase and then classification phase. This approach focuses on extracting the feature of image and classify them using machine learning classification technique. The extracted features are trained using CNN Classification technique. Finally, compared the performance using the existing classification methods.

Advantages of Proposed System:
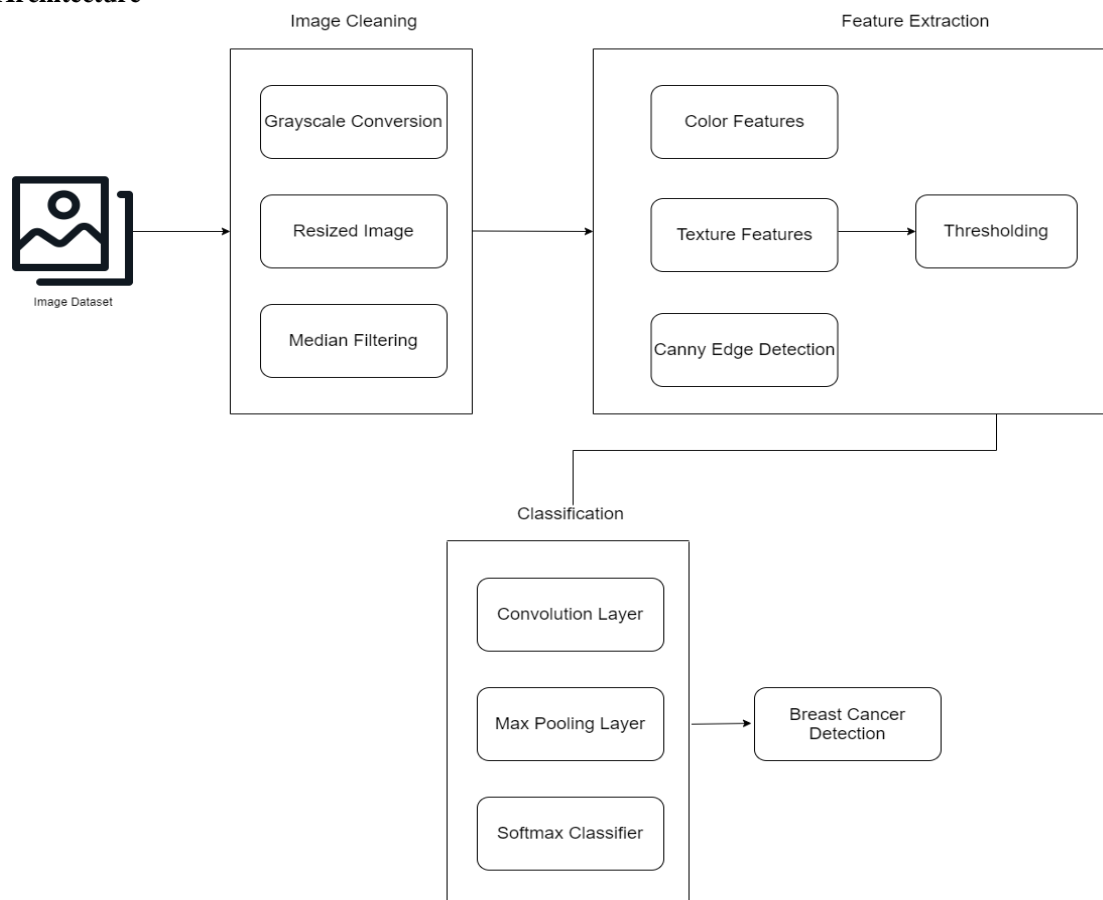
A. System Architecture

1.   Work could be beneficial to obtain fast and precise quantification, reduce observer variability, and increase objectivity.
2.   Cell nuclei detection using image thresholding and image edge detection.
3.   In this work measure accurate cell features.
4.   This application can be used by physicians from their homes or any other place.
5.   This work will be suitable for images with a high degree of noise and blood cells and cell overlapping, as it can successfully detect the cell nuclei.

Figure 1. System Architecture

B.   Algorithms

1. Convolution Neural Network (CNN)

The construction of CNN is divided into two tiers. The first is the feature extract layer, which connects each neuron's input to the preceding layer's local ready fields and removes the local feature. When the local characteristics are extracted, the location link between local functions and other features will also be revealed. The other is a map layer with a benefit for each network computer layer. Every feature map has the same weight as the neurons in the plane. The sigmoid function is used as the network's activation function in the feature plan structure, which gives the map a unique look. Furthermore, because neurons in the same mapping plane share weight, the network's number of free parameters is reduced. After each layer of convolution in the convolution neural network, which is used to discover the local mean and the second extract, the two distinct features extracting structure reduces resolution.

Step 1: Choose a dataset.
Step 2: Feature selection using information gain and ranking
Step 3: Classification algorithm
Step 4: Each Feature calculate fx value of input layer
Step 5: bias class of each feature calculate
Step 6: Next produce the feature map it goes to forward pass input layer
Step 7: Calculate the convolution cores in a feature pattern
Step 8: Produce sub sample layer and feature value.
Step 9: Back propagation input deviation of the kth neuron in output layer.

Step 10: Finally give the selected feature and classification results.

C. Mathematical Model

In thi paper have k sub-spaces so that there are k classification results of sub-space, called CL_SS1, CL_SS2,..., CL_SSk. Thus the problem is how to integrate all of those results. The simple integrating way is to calculate the mean value:

$$CL = \frac{1}{k} \sum_{i=1}^{k} CL\_SS_i$$

Or weighted mean value:

$$CL = \frac{1}{k} \sum_{i=1}^{k} w_i CL\_SS_i$$

Where w_i is the weight of classification result of subspace SSi, and satisfies:

$$\sum_{i=1}^{R} w_i = 1$$

The centroid of a hand is calculated as follows:

$$\underline{X} = \frac{\sum_{i=0}^{k} xi}{k}, \underline{Y} = \frac{\sum_{i=0}^{k} yi}{k}$$

Where $(\underline{X}, \underline{Y})$ represents the centroid of the hand, xi and yi are x and y coordinates of the ith pixel in the hand region and k denotes the number of pixels that represent only the hand portion.

In the next step, the distance between the centroid and the fingertip was calculated. For distance, the following Euclidean distance was used:

$$Distance = \sqrt{(x2 - x1)^2 - (y2 - y1)^2}$$

Where (x1, x2) and (y1, y2) represent the two co-ordinate values.

CONCLUSION

In this paper, use CNN Classification with various configurations to classify ultrasound images of breast cancer into healthy and infected. In classification tests, the developed CNN Classification performed well on ultrasound image characteristics. However, when compared to one of the existing classification algorithms, the CNN Classification performs better. Deep and machine learning have advanced to the point where they can perform difficult categorization jobs. Breast cancer ultrasound images are successfully classified into healthy and infected in this proposed system. In the future, we will be able to expand this work by utilizing 3D breast images assessment and working on them and also including task decomposition and synchronization, can be further explored especially for a large amount of training dataset.

REFERENCE

[1]. F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," 2017 IEEE International Conference

[2]. N. Bayramoglu, J. Kannala, and J. Heikkila, "Deep learning for magnification independent breast cancer histopathology image classification," 2016 23rd International Conference on Pattern Recognition (ICPR), 2016.

[3] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," IEEE Systems Journal, vol. 8, no. 3, pp. 949–964, Sept 2014.

[4] P. Filipczuk, T. Fevens, A. Krzyzak, and R. Monczak, "Computeraided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," IEEE Transactions on Medical Imaging, vol. 32, no. 12, pp. 2169–2178, 2013.

[5] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in International Joint Conference on Neural Networks, Vancouver, BC, Canada, July 2016, pp. 2560–2567.

[6] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," Neurocomputing, vol. 191, pp. 214–223, 2016.

[7] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," in IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, June 2016, pp. 2424–2433.

[8] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, and A. Madabhushi, "Stacked sparse autoencoder (SSAE) for nuclei detection of breast cancer histopathology images," IEEE transactions on medical imaging, vol. 35, no. 1, pp. 119–130, 2016.

[9] F. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455–1462, 2016.

[10] Shang Gao Hongmei Li, "Breast Cancer Diagnosis Based on Support Vector Machine". 2012 International Conference on Uncertainty Reasoning and Knowledge Engineering.

[11] Samrudhi Kaware, Dr. V. S. Wadne" Novel approach to improve accuracy of cancer and diabetes detection by using Machine Learning" IJASRW 2020.

[12] Samrudhi Kaware, Dr. V. S. Wadne" Enhance Cancer and Diabetes Detection by using Machine Learning Techniques" © 2020 IJRAR March 2020.