iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** III **Month of publication:** March 2022

DOI: https://doi.org/10.22214/ijraset.2022.41149

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Breast Cancer Detection System Using Deep Learning

Dr. Sharada K. A[1], Samreen Fathima S Md[3], Sayeda Fathimunissa[5], Vishnu Raj S[2], Sonu Surendran[4]
[1]*Associate Professor, Dept. of computer scienceEngineering, HKBK College of Engineering*
[2, 3, 4, 5]*Dept. of Computer Science Engineering, HKBK College of Engineering*

Abstract: ***Breast Cancer detection and classification is very hard. In fact, tumor or cancer is a complex process during which mammogram images undergoes various transformations. In addition, different areas of image which exhibit variable and high appearance are characterized by various tissues. Our main benefaction of this process is image classification for cancer prediction and improvise its performance. We have trained and tested the implementation of our work on an opensource dataset. This project will be developed using python3. The project will be deployed on Jupyter IDE. Overall this project focuses on providing maximum performance and efficiency.***

Keywords: Mammogram, Python, Jupiter,Deep Learning, Artificial Intelligence, Cancer, Image Classification.

I. INTRODUCTION

As we all know that cancer emerged as one of the most dangerous and life threatening diseases in the world, it mostly effects women. So, our prime target is to use deep learning technique for diagnosis of cancer in patients especially women.

The cancer or tumor in breast or any other body part starts when abnormal cells in effected body part begin to grow uncontrollably. A combination of genetic and environmental factors can be termed as the main factors for breast cancer. The genetic factors such as family history are the factors that cannot be controlled or changed. Our main contribution of this process is image classification, tumor prediction and improvise in performance and analysis. In addition, different areas of image whichexhibit variable and high appearance are characterized by various tissues. In order to extract information and also to get enhanced image the method used is commonly referred to as Image processing. Here we have used signal processing to get the required output image from an input image. In today's rapidly growing technologies image processing is one of the technologies that is used in research, engineering and computer science too.

Image processing basically includes the following three steps:

1) Importing the image via imageacquisition tools
2) Analyze and manipulate the image
3) The result is the altered image which isnothing but the output.

II. LITERATURE SURVEY

1) Paper titled "Breast Cancer Detection Based on Deep Learning Technique" proposed a method where the two models of deep learning technique are compared with the breast cancer detection system. The process that were used by them were methods like pre-processing, classification and evaluation of performance. In this paper the majorly used models are VGG16 and ResNet50 and these two models are used to differentiate the difference between the abnormal tumor and normal tumor with the help of IRMA dataset. When ResNet50 and VGG16 were compared, ResNet50 produced an accuracy of 91.7% were as the result using VGG16 was found to be around 94%. The raw images collected from image retrieval of medical application (IRMA) datasets are pre processed using the resized image and its is converted to fit the network system.

2) Paper titled "Breast Cancer Detection FromHistopathological Images Using Deep Learning" focuses on using the MIAS dataset that is available for the diagnosis of breast cancer. The other methods used here were pre processing of image, neural language processing and the diagnosis of medical patients. The paper was mainly divided into three parts first the dataset collection and then the preprocessing algorithmwas applied and then after scaling and filtering thedata the dataset was split into testing as well as forthe training purpose and the output was visualized using a data graph. Therefore we have seen that the deep learning method used here for the diagnosis of cancer works well only with the MIS dataset. This dataset has 12 features for breast cancer and diagnosis. In the paper they have also compared the proposed deep learning algorithmwith other machine learning models as well.

©IJRASET: All Rights are Reserved

3) Paper titled "Research on the Detection Method of Breast Cancer Deep Convolutional Neural Network Based on Computer Aid" uses the Convolutional Neural Network method to classify the image and also for method detection with the help of computer aided features. Here the Convolutional Neural Networks with different structures were pre trained and then it was further used to automatically extract the characteristic features and also fuse the extracted features from the two structures, they have compared their method with traditional methods and tried to improve the accuracy and the accuracy was found to be around 89% and hence the accuracy of the classification of images of breast cancer are improved significantly when compared with other traditional methods. They focused on adjusting the impact of CNN structure so that they can integrate various information for the classificationperformance.

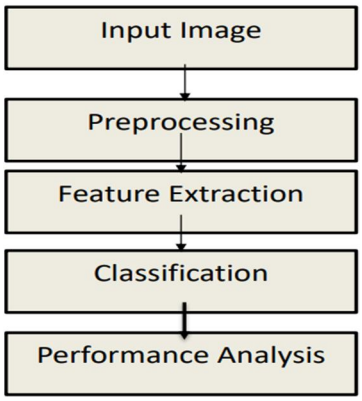
4) Paper titled "Self-Supervised Learning For Detection Of Breast Cancer In Surgical MarginsWith Limited Data" has used the self supervised learning so that the performance of tumor detection can be improved at surgical margins, the labeled data samples used here are limited in number. The model divides the image spectra into batches of smaller size there by shuffling their order and hence new instances are generated. The characteristics of the model is captured by learning the order of the patches and also by interrogating the shuffled data. The weights are then fine tuned for detection of cancer. The self supervised modelis applied on the REIMS dataset. 144 sample cancer data. These features are then applied to categorize the cancer as malignant and benign.

5) Paper titled "Breast Cancer Malignancy Prediction Using Deep Learning Neural Networks" focuses on predicting the breast cancer's malignancy with the help of deep neural networks. The Wisconsin breast cancer dataset is used here. The overfitting of the neural networks model is optimized to proceed with the early stopping mechanism and a F1 score of over 98 is obtained, if the cancer is benign then the value indicated is zero and if it is malignant the value indicated is one. Since in the deep learning technique the machine tires to copy or mimic the human behavior. Hence it is used in computer aided diagnosis.

III. PROPOSED SOLUTION

The model is implemented using the patient dataset and the archived reports from the data warehouse. The Deep Learning model used for analytic purpose and assessment can help to identify more subtle patterns accurately. By doing so we will be able to generate an optimal threshold and hence for further diagnosis support improve the accuracy. We have also used the image processing as a result which can be used to perform various operations on the image and hence we'll be able to extract some features that are useful, in order to get an enhanced image. A method called signal processing is used when the input and the output both are images. The output can also be a characteristics/features associated with that particular image. As we all know image processing forms the core for research areas within the various disciplines of computer scienceengineering.

Fig:- Architecture Diagram of overall breast cancer detection system.



A. Image preprocessing

1) A resizing operation is performed in order to fit the images into the network of the input layer.The size of the images should be 224-by-224. Hence the images are resized to this particular format for an uniform dimension.

2) Next the mammogram images are converted into a three channel input image and the grey scale images are also converted to RGB images as this can be fed into the three channel input.

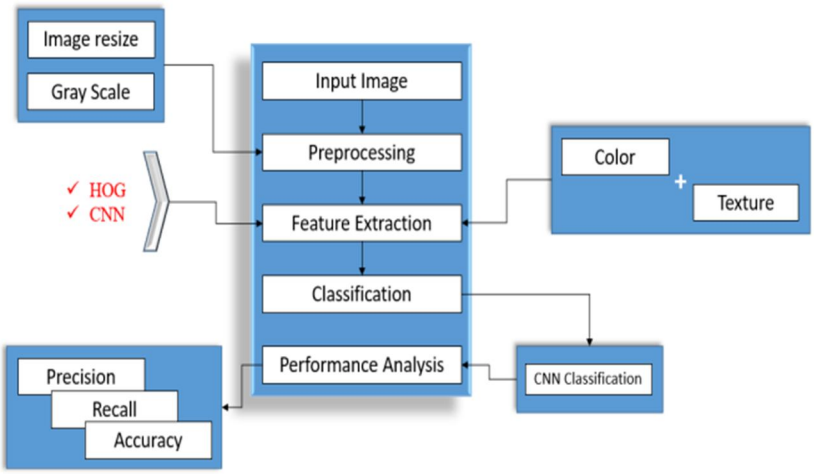
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

B. Models used

The main model used is Convolutional Neural Networks(CNN). Various application such as image recognition and computer vision make use of the Convolutional Neural Networks(CNN). Hence it is one of the popular models that is used. Even areas such as image classification andrecognition make use of this model. The CNN consist of the input layer along with multiple hidden layers and one output layer. There are two important layers in CNN which are the convolutional layers and the pooling layers. The pooling layers often lie between the convolutional layers hence it can be used to perform subsampling of the data in order to reduce the overheads during the training.

In order to extract features from the image data the HOG or (Histogram Of Oriented Gradients) descriptor is used. Lets look at some of the main advantages of HOG that makes this particular descriptor stand out from other feature descriptors, the main focus of HOG is regarding the object's shape rather than just identifying whether the pixel is an edge or not. HOG can be used to provide the direction for the edge as well,this is done by calculating the gradients, where the gradients are broken down into smaller regions and the orientation along with thegradients are calculated for each of the region and finally an histogram is generated for each of these regions separately and therefore it is knownas "Histogram Of Oriented Gradients".

C. The flow Diagram



D. Some of Methods Applied

1) Image Aquistation
2) Image Preprocessing
3) Extraction Of Components
4) Classify the Images
5) Analyzing the Performance

IV. CONCLUSION

Idea of harnessing the power of Deep Learning to make an efficient Breast Cancer Detection System has merit to it. Based on the training and testing of datasets, the model is able to predict an accuracy that varies from 85% to 95%. The main models we used were HOG(Histogram Of Oriented Gradients), CNN(Convolutional Neural Networks).In Deep Learning, a Convolutional Neural Network (CNN/ConvNet) consist of deep neural networks, most commonly used to analyze the visual imagery. After considerable analysis, comparison and prediction it was found that the cancer detection system that was put forward by us mainly using the CNN algorithm was able to accurately and efficiently display the results.

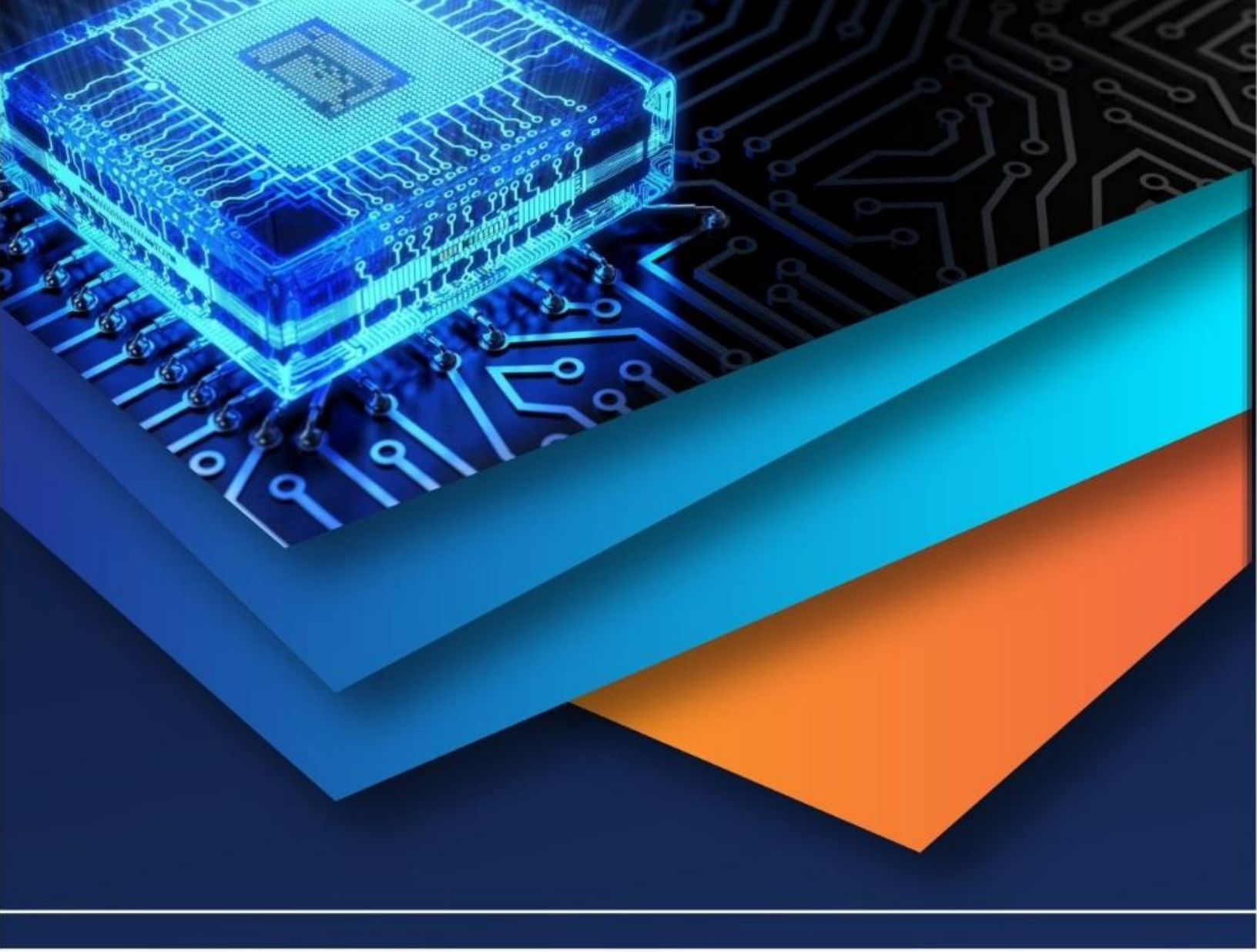
V. ACKNOWLEDGMENT

We would like to express our gratitude to our guide Prof. Dr. Sharada K.A and our Head Of Department Dr. R Loganathan for giving us agreat opportunity to excel in our learning through this project. We would also like to thank our families and friends for their consistent encouragement throughout the project. This project has helped us to expand our knowledge to a great extend.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] J. Ferlay e t a l ., Globocan 2012 v1.0, Cancer Incidence and Mortality Worldwide: Iarc Cancerbase no. 11 2014 [Online]. Available: http:// globocan.iarc.fr
[2] M. Veta, J. P. Pluim, P. J. van Diest, and M. A. algorithms to the segmentation of clustered nuclei," J. Viergever, "Breast cancer histopathology analysis: Cytometry, vol. 28, pp. 289–297, 1997.
[3] A review," IEEE Trans. Biomed. Eng., vol. 61, no.5, pp. 1400–1411, May 2014.
[4] F. Ghaznavi, A. Evans, A. Madabhushi, and M. Feldman, "Digital imaging in pathology: Whole-slide imaging and beyond," Annu. Rev. Pathol., Mechanisms Disease, vol. 8, pp. 331–359, 2013.
[5] C. F. Lucchinetti et al., "Inflammatory cortical demyelination in early multiple sclerosis," N. Eng. J. Med., vol. 365, no. 23, pp. 2188–2197, 2011.
[6] L. S. Hu e t a l ., "Relative cerebral blood volume values to differentiate high-grade gliomarecurrence from posttreatment radiation effect: Direct correlation between image-guided tissue histopathology and localized dynamic susceptibility-weighted contrast-enhanced perfusionMR imaging measurements," A m . J . N e u ro r a di o l ., vol. 30, no. 3, pp. 552–558, 2009.
[7] D. Scheie et al., "Can morphology predict 1p/19q loss in oligodendroglial tumours?," Histopathology, vol. 53, no. 5, pp. 578–587, 2008.
[8] S. B. Wharton e t a l ., "Subtypes ofoligodendroglioma defined by 1p, 19q deletions, differ in the proportion of apoptotic cells but not in replication-licensed non-proliferating cells," A c t a Neuropathology, vol. 113, no. 2, pp. 119–127, 2007.
[9] V. Grau, A. U. J. Mewes, M. Alcaniz, R. Kikinis,and S. Warfield, "Improved watershed transform for medical image segmentation using prior information," IEEE Trans. Med. Imag., vol. 23, no. 4, pp. 447–458, Apr. 2004.
[10] N. Malpica et al., "Applying watershed



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)