iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VI **Month of publication:** June 2020

DOI: http://doi.org/10.22214/ijraset.2020.6180

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: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

Deep Learning for Early Detection of Breast Cancer using Histopathological Images

Prof. D. D. Pukale[1], Diksha Kale[2], Rutuja Jadhav[3], Akshada Gaikwad[4], Arati Hadke[5]
[1, 2, 3, 4, 5]*Bharati Vidyapeeth's College of Engineering for women*

Abstract: ***Breast Cancer is one of the uncontrollable growth of malignant cells in the chest. The most common treatments used are chemotherapy or hormonal therapy using mammographic images, but the problem with mammographic images is complex and low contrast noisy. Deep Learning can extract high level abstract features from images automatically. Therefore, we have used histopathological images of breast cancer via supervised and unsupervised deep convolutional neural networks. The Pre-Processing process is followed by Image Segmentation where we have implemented Watershed Algorithm. The System has been developed using python Programming Language. In this paper, we have used convolutional neural network(CNN) for Classification purpose which is one of the Deep learning techniques.***
Keywords: Deep Learning, Convolutional Neural network, Histopathological images, classification, Breast Cancer.

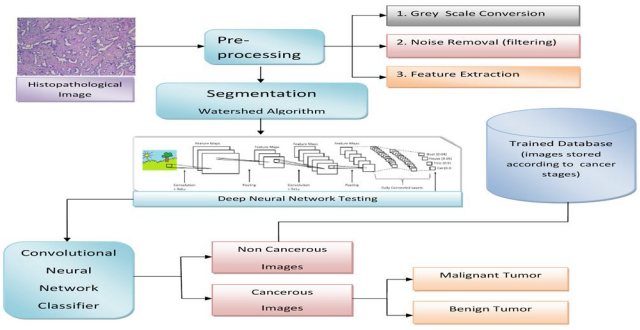
I. INTRODUCTION

Breast Cancer has become one of the most common types of cancer among women. As per clinical statistics, every 1 out of 8 women is diagnosed with breast cancer in their lifetime. Early detection of breast cancer can offer the best treatment outlook. In order to detect signs of breast cancer, breast tissue from biopsies is stained to enhance the nuclei and cytoplasm for microscopic examination. To correctly detect and to diagnose breast cancer ,radiologists face challenges due to the large amount of breast images they have to examine daily and the difficulty of reading the images .Thus, computer-aided detection and diagnosis (CAD) are essential through which a second opinion can be provided to physicians to aid and to support their decisions.
Deep learning with Convolutional Neural Networks has emerged to be one of the most powerful machine-learning tools in Image classification, surpassing the accuracy of almost all other traditional classification methods and human ability. Also the convolutional process can simplify an image which contains millions of pixels to a set of small feature maps,which thereby reduces the dimension of input data while retaining the most important differential features.

A. Objective

Researchers in the image analysis and the pathology fields have recognized the importance of quantitative analysis of pathology images. Since the most current pathology diagnosis is based on the subjective opinion of pathologists, there is clearly a need for quantitative image-based assessment of digital pathology slides.
In this Paper,we propose a system using deep learning with Convolutional Neural Network for the Early Detection of Breast Cancer which will help in better treatment outlook.

II. PROPOSED SYSTEM



III. DESCRIPTION

A. *Input Histopathological Image*

Computer assisted Diagnosis has become one of the most widely researched subjects in medical imaging and diagnostic radiology. Given recent advances in high-throughput tissue bank and archiving of digitized histological studies, it is now possible to use histological tissue patterns with computer-aided image analysis to facilitate disease classification.

In this proposed System we have used a dataset of histopathological dataset which we have refereed from kaggle.

Given is the link for the same:

https://www.kaggle.com/paultimothymooney/breast-histopathology-images/kernels

B. *Pre-Processing*

1) *Gray Scale:* Grayscale is a range of shades of gray without the apparent color. The intensity of pixels is expressed within a given range , between a minimum and a maximum. This range is represented in an abstract way as a range from 0 (or 0%) (total absence, black) and 1 (or 100%) (total presence, white), with any fractional values in between them.

Converting a color image into gray scale image following 3 steps:

a) Take the RGB value of the pixel.

b) Find the average of RGB Avg= (R+G+B)/3

c) Replace the R,G and B value of the pixel with Avg.

2) *Noise Removal:* Noise reduction is the process of removing noise from a signal.

$D=\arg\min||I-d||2+\lambda||f(d)||1$

Where,

DD is de noised image,

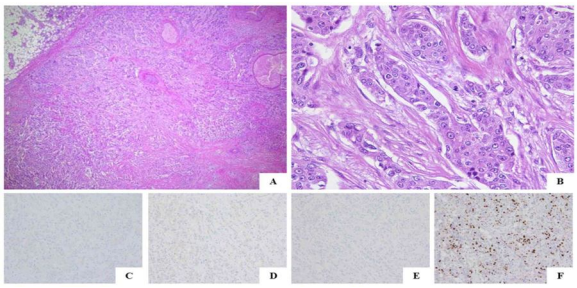
II is noisy image,

ff is transformation operator and

λλis regularization factor.

The first term tries to make DD closer and closer to the original image, and the second term tries to keep DD as sparse as possible. Through solving the above problem, we would remove noise from the image.

3) *Feature Extraction:* Feature Extraction aims to reduce the number of features in a dataset by creating the new features from the existing ones. These new reduced set of features should then be able to summarize most of the information contained in the original set of image features.



Feature Extraction of Histopathological Images

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

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

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

$$\text{Correlation} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(i-m_r)(j-m_c)P_{ij}}{\sigma_r \sigma_c}$$

$$m_r = \sum_{i=1}^{N} i \sum_{j=1}^{N} P_{ij} \qquad m_c = \sum_{j=1}^{N} j \sum_{i=1}^{N} P_{ij}$$

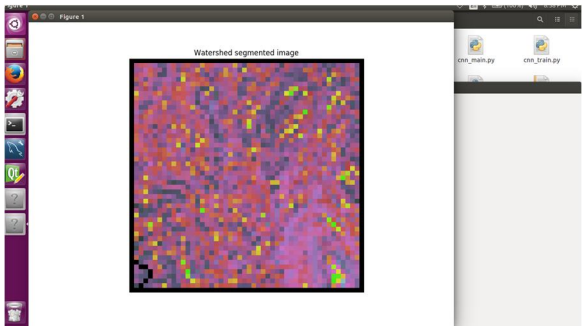
The term “Feature Extraction” deals with extracting feature such as energy, entropy, randomness, correlation, homogeneity.

C. *Segmentation*

Image segmentation is the process of partitioning Digital images into multiple segments.The goal of segmentation is that it simplifies and/or changes the representations of an images into something that is more meaningful and easier to analyze.
We have used the Watershed Algorithm for segmentation purposes.

D. *Watershed Algorithm*

Watershed analysis refers to the process of using DEM and raster data operations to delineate the watersheds and to derive different features such as streams, stream network, catchment areas, basin etc.



IV. CONVOLUTIONAL NEURAL NETWORK

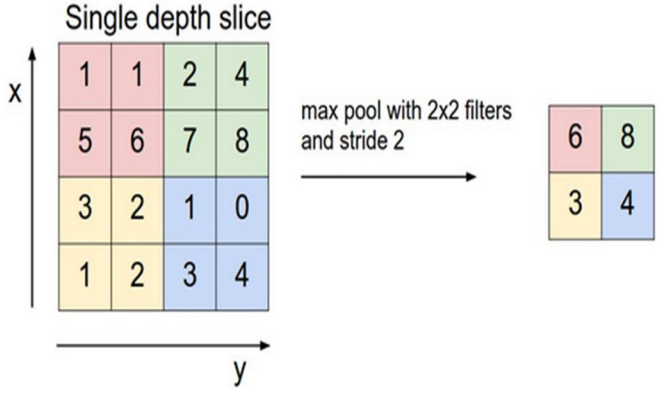
Convolutional neural networks are neural networks used primarily for the classification images, cluster images by similarity, and perform object recognition within scenes. For example, convolutional neural networks are used to identify faces, individuals, street signs, tumors, platypuses and many other aspects of visual data.
Convolutional neural networks are the leading architecture in deep learning that are used to solve an image classification problem. The goal of this paper is to tell which class the input image belongs to the process of building a convolutional neural network always involves 4 major steps:

A. *Convolution Layer*

Convolution is the mathematical operation that is used in image processing to filter signals, find patterns in signals etc. All neurons in this layer perform convolution on inputs. The most important parameter in a convolutional neuron is the filter size.
We shall slide convolution the filter over the whole input image to calculate this output across the image and here we slide our window by 1 pixel at time this number is called Stride. Typically we use more than 1 filter in one convolution layer.

B. *Pooling*

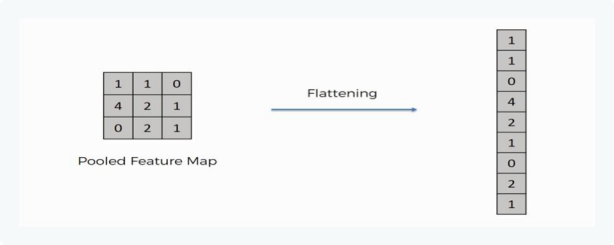
The next layer in the convolutional network has three names: Max pooling, Downsampling and Subsampling. The activation maps are fed into a downsampling layer, and like convolutions , this method is applied one patch at a time. In this case, a max pooling simply takes the largest value from one patch of an image and places it in a new matrix next to the max values from other patches, and discards the rest of the information contained in the activation maps.



C. Flattening

After the process of pooling the next step to do is flattening of images.

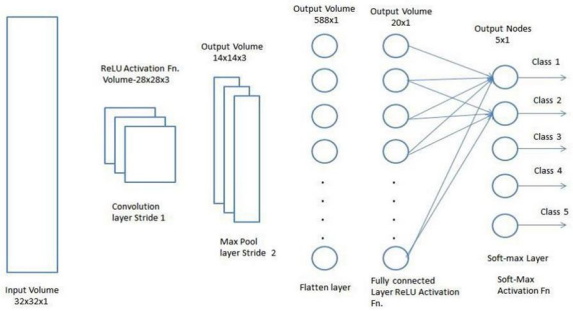
Flattening is the function that converts the pooled feature map to a single column that is passed to the fully connected layer. Dense adds the fully connected layer to the neural network to initialize the neural network we have created an object of the Sequential class.



Process of Flattening

D. Fully Connected Network

A fully connected neural network consists of a series of fully connected layers. A fully connected layer is a function from $\mathbb{R}$ m to $\mathbb{R}$ n . Each output dimension depends on each input dimensions.

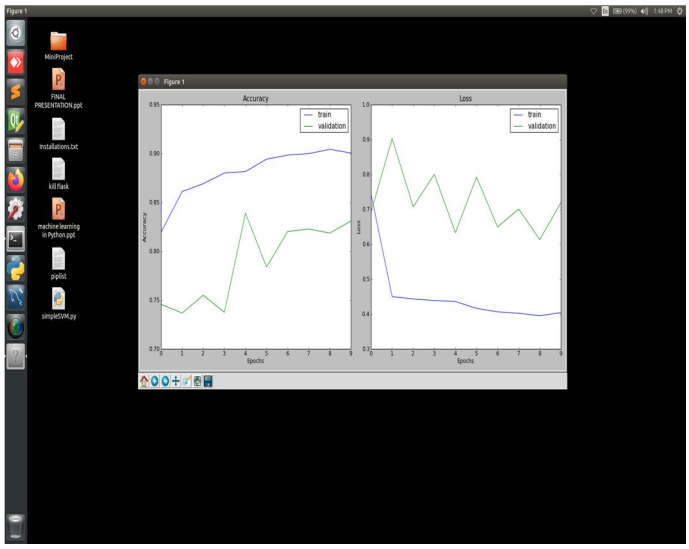


Fully Connected Network

V. RESULTS

In this proposed system, We obtain results with the presence of Cancer in the patient or not. The Dataset used is of Histopathological Images which have images of two types namely malignant and Benign Tumour.

The System at the end generates Report with the patients details and result of the cancer presence or absence. Accuracy of the raw images obtained by using different filter sizes in CNNs. Accuracy is determined when the model parameters are learned and fixed and no further learning takes place. The Accuracy for the system developed using CNN comes out to be 90%.



REFERENCES

[1] Prof. Kamalakannan J, AbinayaVaidhyanathan , TamilarasiThirumal, Kansagara Deep MukeshBhai "Study On Different Classification Technique for Mammogram Image.", Vol.5, No.6 (2013), pp.171-180, 2015 IEEE.

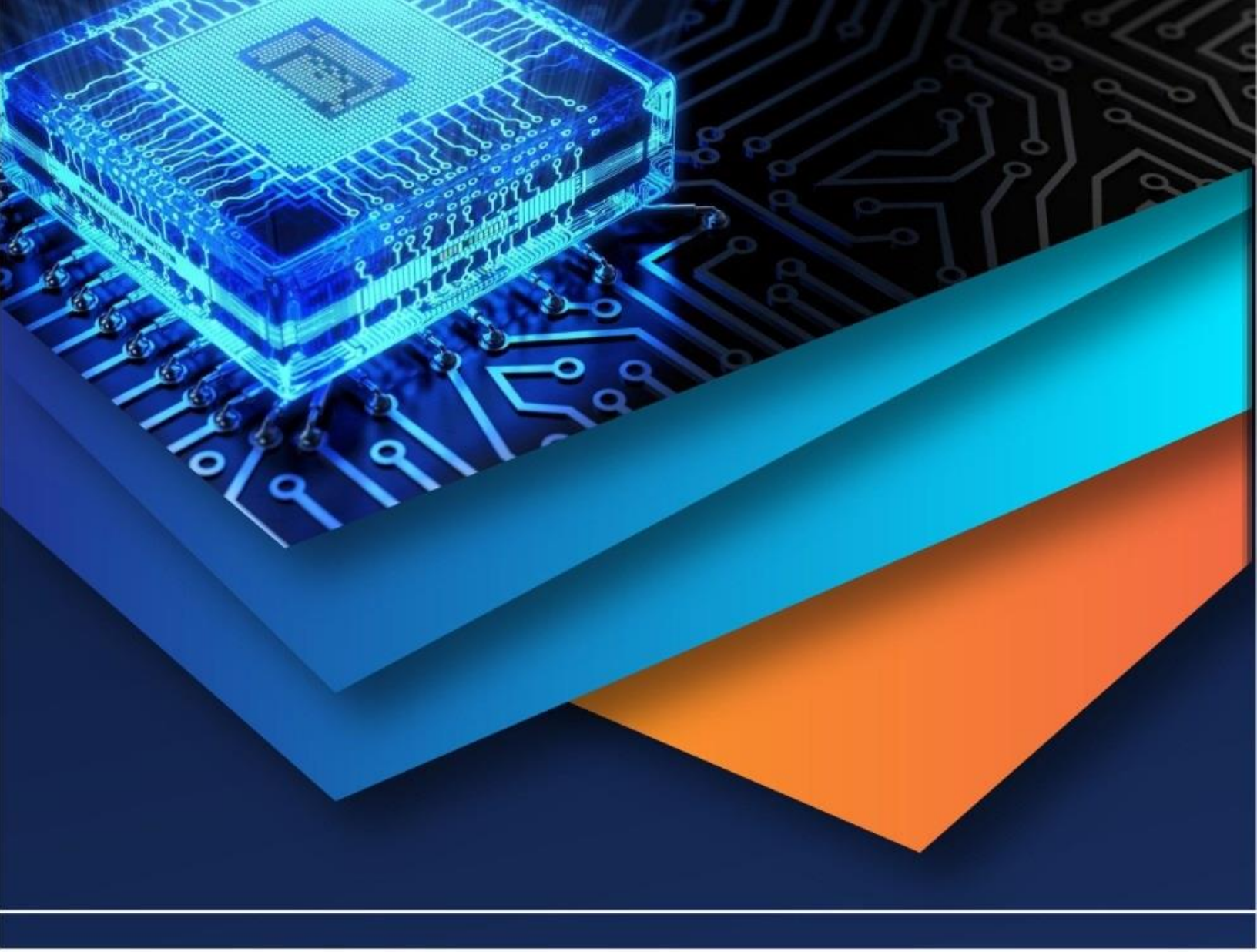
[2] JD. Selvathi and A. Aarthy Poornila,"Breast Cancer Detection In Mammogram Images Using Deep Learning Technique," Middle-East Journal of Scientific Research 25 (2): 417-426, 2017 IEEE

[3] Ketan Sharma; Bobbin Preet " Classification of Mammogram Images by using CNN Classifier", vol. 5032, pp. 882–889, 2016 IEEE.

[4] J. Sceneay, T. Marsh, I. Wong, Y. Qin, A. Sjodin, B. Nilsson, S. A. ¨ Stewart, and S. S. McAllister, "Hematopoietic age at onset of breast cancer dictates disease aggressiveness and progression," Cancer Research, vol. 76, no. 14 Supplement, pp. 4400–4400, 2016.

[5] C. Dromain, B. Boyer, R. Ferre, S. Canale, S. Delaloge, and C. Balleyguier, "Computed-aided diagnosis (cad) in the detection of breast cancer," European Journal of Radiology, vol. 82, no. 3, pp. 417–423,2013.

[6] J. Sceneay, T. Marsh, I. Wong, Y. Qin, A. Sjodin, B. Nilsson, S. A. ¨ Stewart, and S. S. McAllister, "Hematopoietic age at onset of breast cancer dictates disease aggressiveness and progression," Cancer Research, vol. 76, no. 14 Supplement, pp. 4400–4400, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY