Breast Cancer Histology Image Classification Using Deep Neural Networks

A report submitted in partial fulfilment of the requirements for the degree of

Bachelor of Technology

in

Computer Science & Engineering

By

SHUBHAM KUMAR
(Roll No.:11000215023)

SAYANTANI GHOSH

(Roll No.: 11000115021)

MANAS KUMAR MISHRA

(Roll No.: 11000115011)

PRANAB NANDY

(Roll No.: 11000115015)

SOMOSHREE DATTA

(Roll No.: 11005215022)

Under the Guidance of:

Mr. SOURAV CHATTERJEE

Assistant Professor

Department of Computer Science and Engineering

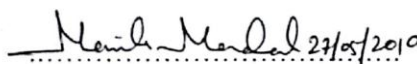For the Academic Year 2015-2019

MAY 2019

<u>Abstract</u>

Breast cancer is one of the largest causes of women's death in the world today. It is also ranked the number one cancer among Indian females with a rate of 25.8 per 100,000 women and a mortality rate of 12.7 per 100,000 women. It is also one of the most common causes of cancer worldwide. There have been many biological and non-biological research in the past and present to be able to prematurely detect breast cancer. Advance engineering of natural image classification techniques and Artificial Intelligence methods has largely been used for the breast-image classification task.

In this paper, we have put a special emphasis on the Convolutional Neural Network (CNN) method for breast image classification. We have taken an approach using Deep Learning to try to predict whether a breast cancer tumor is noncancerous or is benign, in situ or invasive stage from high quality histopathological images. The experimental results show higher test accuracy than the most state-of-the-art methods in this field. The involvement of digital image classification allows the doctor and the physicians a second opinion, and it saves the doctors' and physicians' time. Despite the various publications on breast image classification, very few review papers are available which provide a detailed description of breast cancer image classification techniques, feature extraction and selection procedures, classification measuring parameterizations, and image classification findings.

BONAFIDE CERTIFICATE

This is to certify that this project report "**Breast Cancer   Histology Image Classification Using Deep Neural Networks** " is the bonafide work of **Shubham Kumar, Manas Kumar Mishra, Sayantani Ghosh, Pranab Nandy** and **Somoshree Datta** who carried out the project work under my supervision.
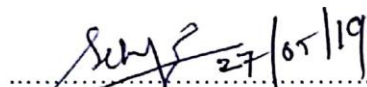
Signature with Seal 27/05/2019

HOD in CSE
Govt. College of Engg. & Textile
Technology, Serampore, Hooghly

Mr. Manik Mondal

Head of the Department

Department of CSE

GCETTS Serampore

Signature with Seal 27/05/19

Mr. Sourav Chatterjee

Assistant Professor

Department of CSE

GCETTS Serampore

<u>Acknowledgements</u>

We would like to take this opportunity to thank everyone whose cooperation and encouragement throughout the ongoing course of this project remains invaluable to us.

We are sincerely grateful to our guide and mentor **Mr. Sourav Chatterjee** of the Department of Computer Science & Engineering, GCETTS Serampore, for his wisdom, guidance and inspiration that helped us go through with this project and take it to where it stands now.

Last but not the least, we would like to extend our warm regards to our families and Peers, especially Gitesh Jain, who have kept supporting us and always had faith in our work.

<u>Contents</u>   <u>Page No</u>

References

List of Figures:

List of Tables:

Chapter 1: Introduction

Motivation

The cell of the body maintains a cycle of regeneration processes. The balanced growth and death rate of the cells normally maintain the natural working mechanism of the body, but this is not always the case. Sometimes an abnormal situation occurs, where a few cells may start growing aberrantly. This abnormal growth of cells creates cancer, which can start from any part of the body and be distributed to any other part. Different types of cancer can be formed in human body, among them breast cancer creates a serious health concern. Due to the anatomy of the human body, women are more vulnerable to breast cancer than men. Among the different reasons for breast cancer, age, family history, breast density, obesity, and alcohol intake are reasons for breast cancer.

Statistics reveal that in the recent past the situation has become worse. As a case study, shows the breast cancer situation in Australia for the last 12 years [2]. In 2007, the number of new cases for breast cancer was 12775, while the expected number of new cancer patients in 2018 will be 18235 [2]. Statistics show that, in the last decade, the number of new cancer disease patients increased every year at an alarming rate.

Breast cancer tumors can be categorized into two broad scenarios.

(i) Benign (Noncancerous)

(ii) Malignant (Cancerous)

Identification of the normal, benign, and malignant tissues is a very important step for further treatment of cancer. Based on the penetration of the skin and damage of the tissue medical photography techniques can be classified into two groups:

(i) Noninvasive.

(ii) Invasive.

Histopathology refers to the microscopic examination of tissue in order to study the manifestations of disease. Specifically, in clinical medicine, histopathology refers to the

examination of a biopsy or surgical specimen by a pathologist, after the specimen has been processed and histological sections have been placed onto a glass slide [11]. Histopathology slides, on the other hand, provide a more comprehensive view of disease and its effect on tissues, since the preparation process preserves the underlying tissue architecture. As such, some disease characteristics, e.g., lymphocytic infiltration of cancer, may be deduced only from a histopathology image. The diagnosis derived from a histopathology image remains the gold standard in diagnosing a substantial number of diseases including almost all types of cancer.

Objective

The aim of this project is to build a classifier that can classify histology images into different classes based on the phase of cancer they are in. Four classes are used:

(I)   Benign

(II)   In-Situ

(III)   Invasive

(IV)   Normal

However, the main challenge is not only to classify the images but to perform the classification in with higher accuracy and precision. There is extensive work in the field of in health care and also cancer phase detection in particular, and a number of reviews exits. One of the most important aspects of Machine Learning in health care is the accuracy we obtain. Our aim is to classify histology images into different classes with acceptable accuracy and precision.

Organization

In our work, a densely connected CNN designed for the analysis of breast cancer histology images is proposed. Unlike previous approaches we perform image wise classification in four classes of medical relevance:

i) Normal tissue,

ii) Benign lesion,

iii) In situ carcinoma

iv) Invasive carcinoma.

For this, a new breast cancer image dataset is presented taken from Bioimaging website. In addition, the proposed CNN architecture is designed to integrate information from multiple histological scales, including nuclei, nuclei organization and overall structure organization. By considering scale information, the CNN can also be used for patch wise classification of whole-slide histology images. A data augmentation method is adopted to increase the number of cases in the training set.

The convolutional neural network architecture used to train the datasets is Dense Neural Network (DNN). DenseNet is being used instead of other convolutional networks because of many advantages of it:

- DenseNet combines features by concatenating them instead of summing them.
- DenseNet requires fewer parameters to make accurate predictions.
- The layers of DenseNet architecture are very narrow.
- Besides better parameter efficiency, one big advantage of DenseNets is that they are easy to train due to their improved flow of gradients and information throughout the network.
- Each layer has direct access to the gradients from the original input signal and the loss function, which leads to an implicit deep supervision. This is massively helpful in training very deep network architectures.
- Further, we also observe that dense connections have a regularizing effect, which reduces overfitting on tasks with smaller training set sizes.

The report is organized as below:

Chapter 1 Introduction:

We discuss about the motivation, objective and organization of this report and project.

Chapter 2 Dataset:

We discuss about the data we used for the project.

Chapter 3 Methodology:

We discuss about our algorithm and approach for the project.

Chapter 4 Data Preprocessing:

We discuss about the preprocessing steps required for the data before augmentation.

Chapter 5 Data Augmentation:

We discuss about our approach for data augmentation.

Chapter 6 Deep Learning Model:

We discuss about Deep Learning Model architecture which resides at the heart of our project.

Chapter 7 Training:

We discuss the training process and frameworks used for the project.

Chapter 8 Results:

We discuss about the results and observations of our project.

Chapter 9 Conclusion:

We discuss about the performance and further improvements that can be made.

Chapter 2: Dataset

The image dataset contains high-resolution (2048 x 1536 pixels), uncompressed, and annotated images from the Bioimaging 2015 breast histology classification challenge. The images in the dataset are then digitized by using the same acquisition conditions, with magnification of 200 and pixel size of (0:42 micron x 0:42 micron). Each image is labeled with one of four classes:

i) In situ carcinoma

ii) Invasive carcinoma

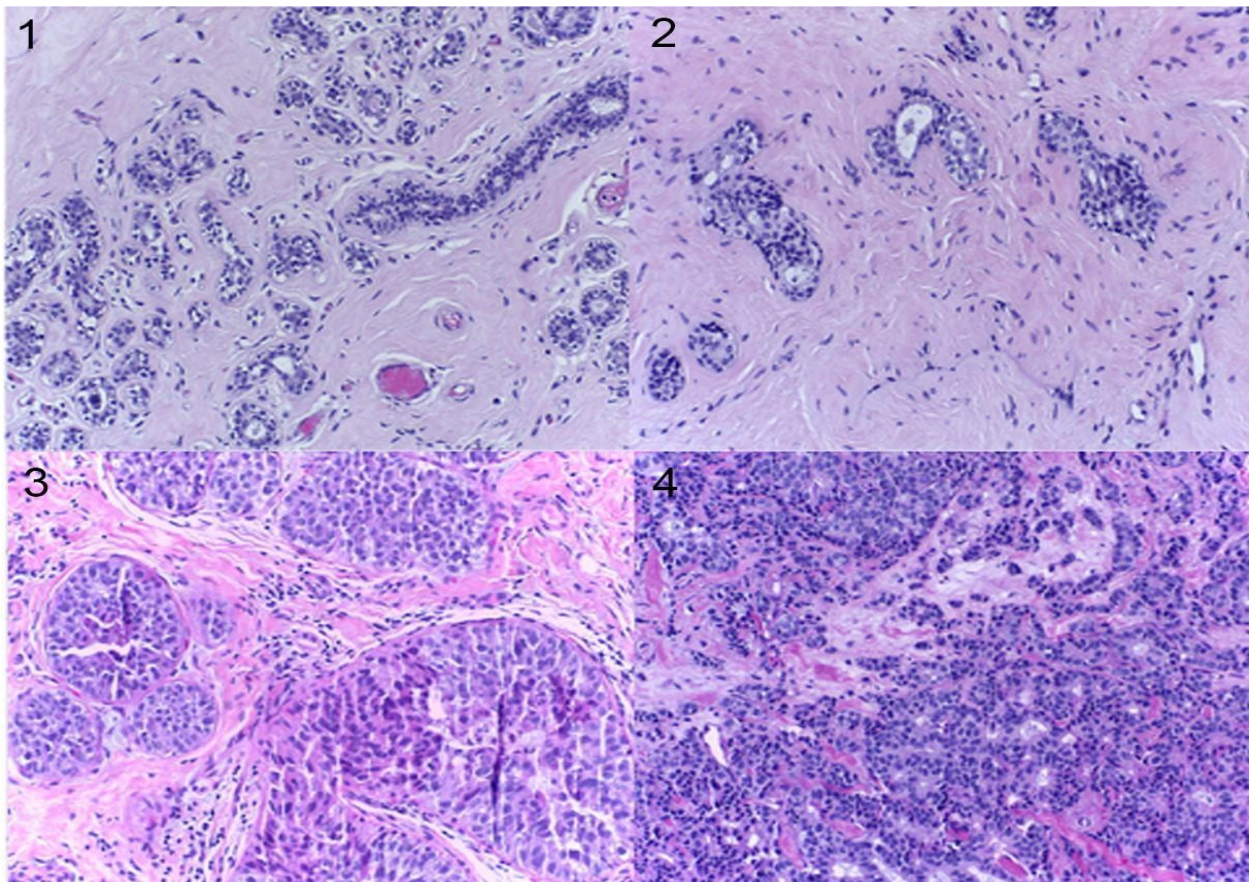iii) Normal tissue

iv) Benign lesion

Fig. 1: Examples of Histopathological images. Clockwise from left -(1) Normal (2) Benign (3) In Situ (4)Invasive Carcinoma

According to the dataset website, the labeling of the dataset was performed by two pathologists. Without specifying the area of interest for the classification, they only provided a diagnostic from the image contents. In the cases where the pathologists disagreed with each other, those images where discarded. The aim of the project is to output an automatic classification for any histopathological image provided as input. The dataset is made of an extended training set of 249 images, and a separate test set of 20 images. In these datasets, the four classes are balanced. The images were selected so that the pathology classification can be objectively determined from the image contents. We mark as *"extended"* dataset a set of 16 additional images provided which have increased ambiguity.

Class wise distribution of images:

CLASS	NUMBER OF IMAGES
Benign	78
Invasive	71
In Situ	72
Normal	65

Table 1: Class wise distribution of images

Chapter 3: Methodology

A number of steps are involved in this project starting from raw images to the trained model. The methodology used for this project can be divided into three parts:

    i)   Data Preprocessing
    ii)   Data Augmentation
    iii)   Model Training

3.1 Data Preprocessing

The histology images from the laboratory are stained with hematoxylins and eosins, thus in order to proceed forward we need to remove these stains from the images. For the preprocessing, we used Macenko Stain Normalization technique for getting normalized histopathological images. The normalized images in the training set are used to create an augmented dataset after the stain normalization process.

3.2 Data Augmentation

The network may suffer from the problem of overfitting as the dataset has a low number of samples when compared to other convolutional neural network classification problems. The complexity and dimension of the dataset is increased by dividing the images into patches. The dataset is further improved by data augmentation through mirroring and patch rotation. Physicians can study the histopathological images for breast cancer from different orientation without changing the diagnosis. This is what makes the patch rotation possible. This also helps increase the size. Briefly, the data augmentation part involved three steps:

   I)   Patching

   II)   Rotation

   III)   Mirroring

The original 250 images are thus converted into 70000 images with each of the patches having the same class label as the original image.

3.3 Model Training

The training part uses Convolution Neural Networks. We are using the DenseNet-121 architecture for our prediction model. We resize our 512 x 512 pixel images to 224 x 224 after which we pass them to a convolutional layer which gives it a size of 112 x 112 pixels. The images are then transferred through the pooling layer which resizes the images to a size of 56 x 56. This architecture has four Dense blocks and subsequently three transition layers. To further improve model compactness, we reduce the number of feature-maps at transition layers. The Dense blocks applies the filters of different sizes in each block to the image for training and reduces the output image size by half each time. The final dense block outputs a 7 x 7 pixel image which is vectorized into a 1 x 1 pixel output by the classification layer. The weights are assigned to the nodes by using Glorot method and Batch Normalization, to prevent overfitting, Dropout technique was used by dropping out units in random with a dropout rate of 0.8.

Chapter 4: Data Preprocessing

The histology images from the laboratory are stained with hematoxylins and eosins, thus in order to proceed forward we need to remove these stains from the images. For the preprocessing, we used Macenko Stain Normalization technique for getting normalized histopathological images. The normalized images in the training set are used to create an augmented dataset after the stain normalization process.

Macenko Stain Normalization:

Marc Macenko and others deduced an algorithm that automatically finds the correct stain vectors for the image and then performs the color deconvolution. This is a fully automatic method that is suitable for rapid analysis of multiple slides. It is also a direct method with very few parameters and no optimizations needed.
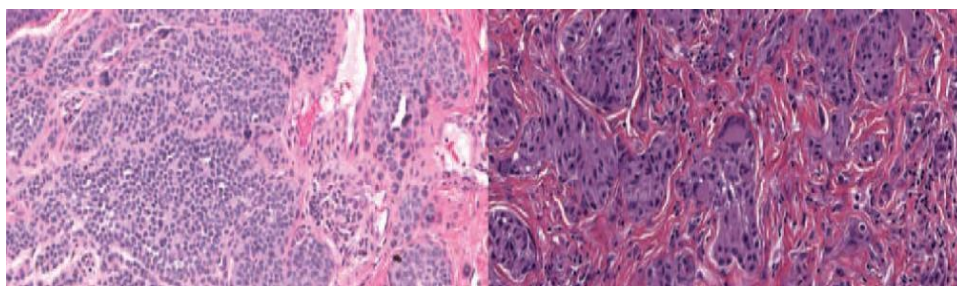

Fig. 2: Example of two histopathology images, both stained with hematoxylin and eosin, but with drastically different appearances.

STAIN VECTOR VARIATION AND CORRECTION

An algorithm is presented to find these particular stain vectors for each image based on the colors that are present. The first step in this process is to calculate the plane that the vectors form. For stability reasons, the pixels with nearly no stain (low OD) were thresholder. After much empirical analysis, a threshold value of β = 0.15 was found to provide the most robust results while removing as little data as possible. Acceptable results are achieved for a wide range of both α and β.

Algorithm 1: SVD-geodesic method for obtaining stain vectors.

Input: RGB Slide

1 Convert RGB to OD

2 Remove data with OD intensity less than β

3 Calculate SVD on the OD tuples

4 Create a plane from the SVD directions corresponding to the two largest singular values

5 Project data onto the plane, and normalize to unit length

6 Calculate angle of each point wrt the first SVD direction

7 Find robust extremes (ath and (100−a)th percentiles) of the angle

8 Convert extreme values back to OD space

Output: Optimal Stain Vectors

INTENSITY VARIATION AND CORRECTION

The intensity of a particular stain depends on the original strength of the stain, the staining procedure, how much fading has occurred since the sample was originally processed, and finally how much of the cellular substance of interest is present in the material. The last quantity is what we actually want to measure. Removing the confounding factors that degrade the signal is necessary for direct analytical analysis of these samples.
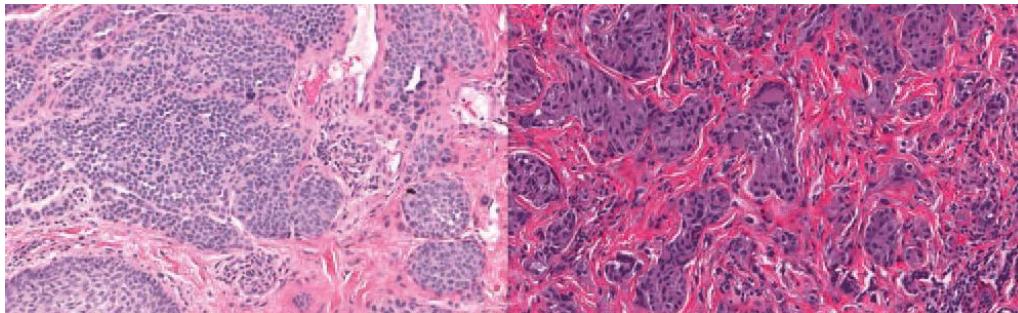

Fig. 3: Same images as Fig. 2, but the second slide has been transformed into the same color-space as the first slide

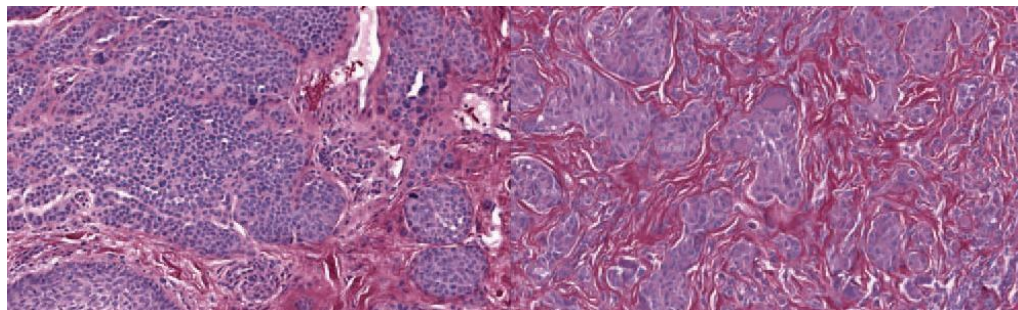
Fig. 4: Same images as Fig. 3, but both are now at the same average intensity level

17

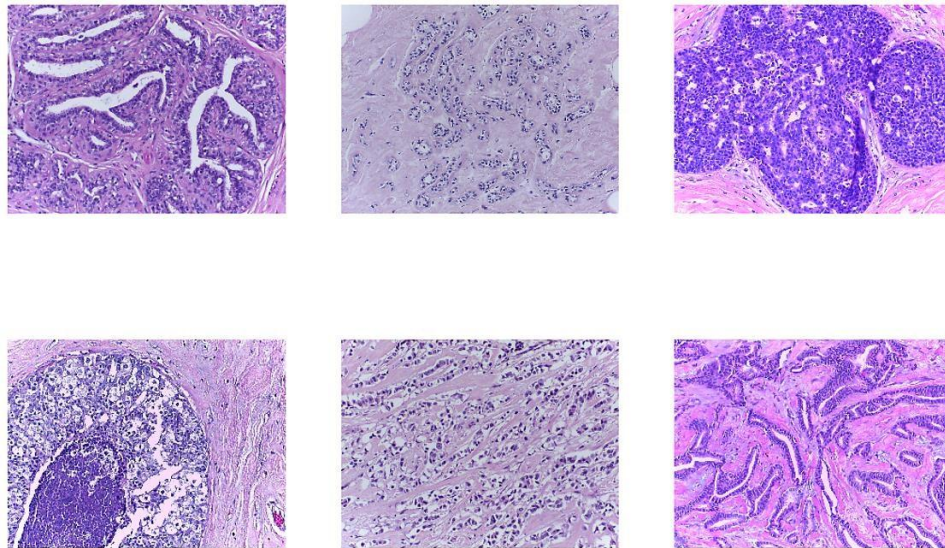**Macenko Stain Normalization on our dataset:**

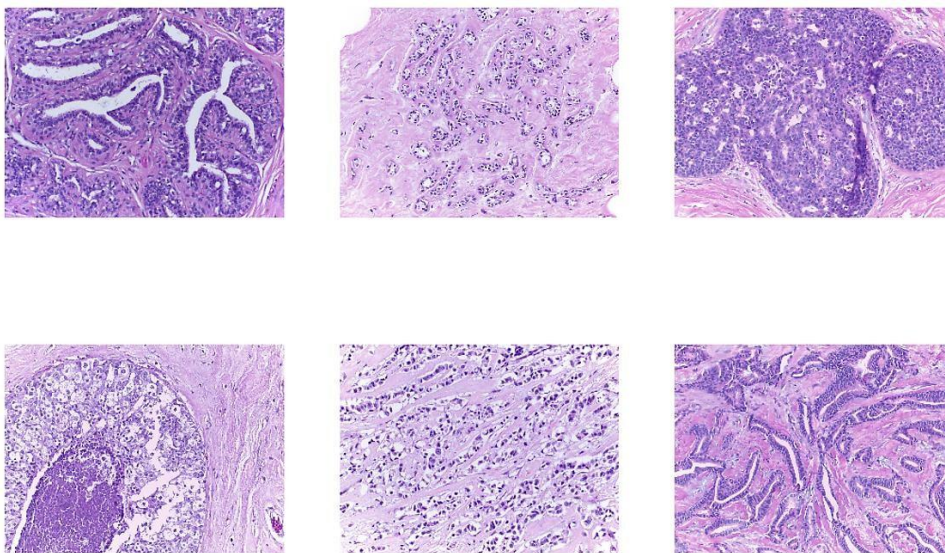Fig. 5: This is how the original image looks like

Fig. 6: This is how the preprocessed image looks like

Chapter 5: Data Augmentation

The network may suffer from the problem of overfitting as the dataset has a low number of samples when compared to other convolutional neural network classification problems. The complexity and dimension of the dataset is increased by dividing the images into patches. The dataset is further improved by data augmentation through mirroring and patch rotation. Physicians can study the histopathological images for breast cancer from different orientation without changing the diagnosis. This is what makes the patch rotation possible. This also helps increase the size. Briefly, the data augmentation part involved three steps:

I.   Patching

II.   Rotation

III.   Mirroring

Fig. 7: Patching of Histopathological Images

The process is as follows:

1. First, the image is divided in patches of 512 x 512 pixels size, with 50 percent overlap. This results in 35 patches per image.
2. Eight different patches are formed from each patch by combining k*90 degrees rotations, with k values being 0, 1, 2, 3. This results in 35*4 = 140 patches per image.
3. Each patch from the last step is then reflected vertically. This results in 140*2 = 280 patches from each image

The original 250 images are thus converted into 250*280 = 70000 images with each of the patches having the same class label as the original image. This forms our final data upon which we will be performing the classification task.

Chapter 6: Deep Learning Model

The training part uses Convolution Neural Networks. We are using the DenseNet-121 architecture for our prediction model. We resize our 512 x 512 pixel images to 224 x 224 after which we pass them to a convolutional layer which gives it a size of 112 x 112 pixels. The images are then transferred through the pooling layer which resizes the images to a size of 56 x 56. This architecture has four Dense blocks and subsequently three transition layers. To further improve model compactness, we reduce the number of feature-maps at transition layers. The Dense blocks applies the filters of different sizes in each block to the image for training and reduces the output image size by half each time. The final dense block outputs a 7 x 7 pixel image which is vectorized into a 1 x 1 pixel output by the classification layer. The weights are assigned to the nodes by using Glorot method and Batch Normalization, to prevent overfitting, Dropout technique was used by dropping out units in random with a dropout rate of 0.8.
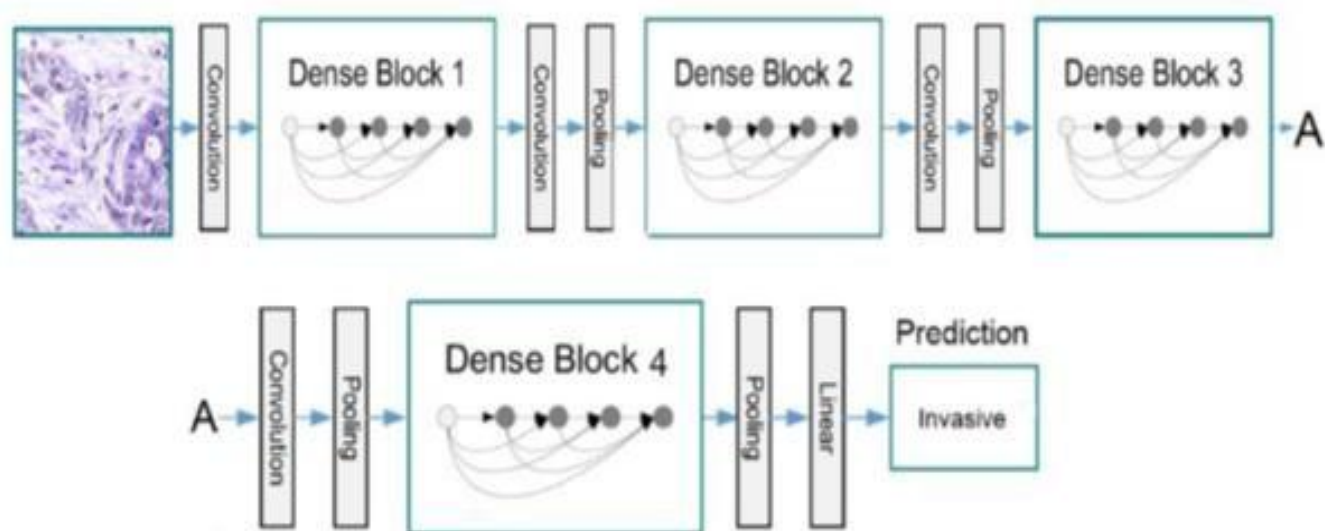

Fig. 8: The architecture pipeline after the image patches are generated.

Dense Net:

Dense Convolutional Network (DenseNet), which connects each layer to every other layer in a feed-forward fashion. Whereas traditional convolutional networks with L layers have L connections—one between each layer and its subsequent layer—our network has L(L+1)/2 direct connections. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers. DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

DenseNet is being used instead of other convolutional networks because of many advantages of it:

- DenseNet combines features by concatenating them instead of summing them.
- DenseNet requires fewer parameters to make accurate predictions.
- The layers of DenseNet architecture are very narrow.
- Besides better parameter efficiency, one big advantage of DenseNets is that they are easy to train due to their improved flow of gradients and information throughout the network.
- Each layer has direct access to the gradients from the original input signal and the loss function, which leads to an implicit deep supervision. This is massively helpful in training very deep network architectures.
- Further, we also observe that dense connections have a regularizing effect, which reduces overfitting on tasks with smaller training set sizes.

Dense connectivity
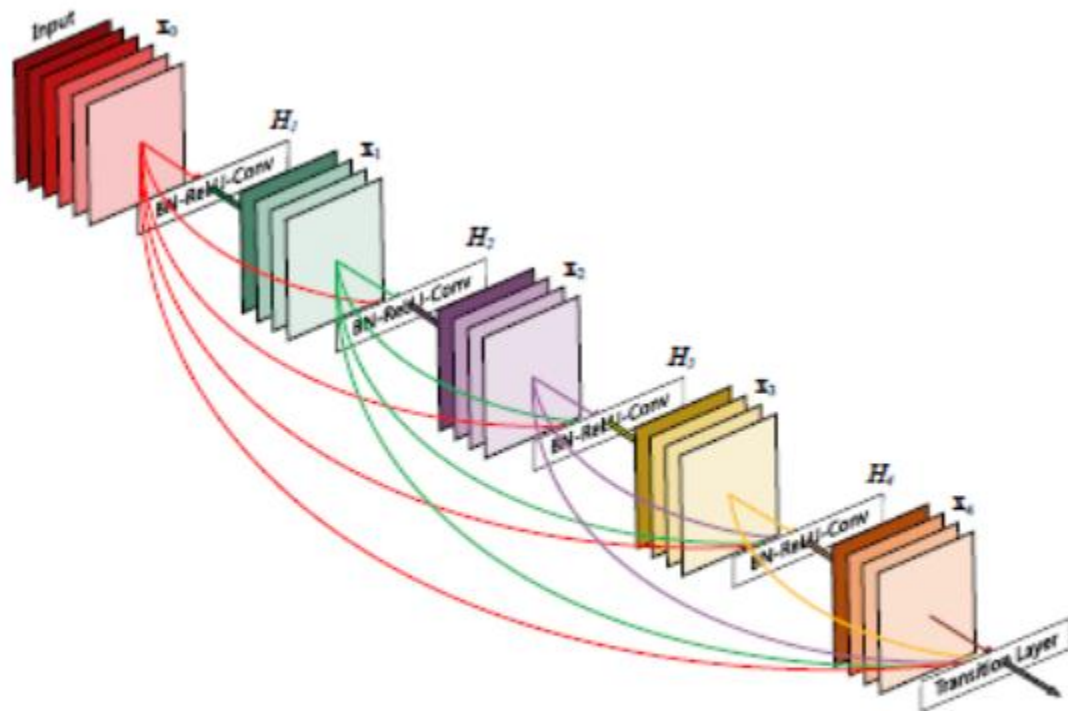
Fig. 9: A 5-layer dense block with a growth rate of k = 4. Each layer takes all preceding feature-maps as input.

Dense Net – 121 layers:

Layers	Output Size	DenseNet-121	
Convolution	112 × 112		
Pooling	56 × 56		
Dense Block (1)	56 × 56	1 × 1 conv 3 × 3 conv	× 6
Transition Layer (1)	56 × 56 28 × 28		
Dense Block (2)	28 × 28	1 × 1 conv 3 × 3 conv	× 12
Transition Layer (2)	28 × 28 14 × 14		
Dense Block (3)	14 × 14	1 × 1 conv 3 × 3 conv	× 24
Transition Layer (3)	14 × 14 7 × 7		
Dense Block (4)	7 × 7	1 × 1 conv 3 × 3 conv	× 16
Classification Layer	1 × 1	1000D fully-connected, softmax	

Fig. 10: Dense Net 121 layers architecture

Chapter 7: Training

A virtual machine with the following specifications was created on Google Cloud Platform and used to train the model:

I.   8GB RAM

II.   220GB SSD

III.   8 Core Processor

IV.   NVIDIA Tesla K80 Graphics Card

The dataset was divided into training, validation and test set. 68880 images were used as training set, 5600 images as validation set and 5600 images as test set. A batch size of 32 was used. We trained a densenet model with 121 layers using Adams Optimizer for 25 epochs.

Total params: 4,226,224

Trainable params: 4,180,330

Nontrainable params: 45,894

A dropout rate of 0.8 was used to avoid overfitting. As we see in Figure 4, both training and validation losses are continuously decreasing with each epoch. After 25 epochs we get a train loss of 0.0852 and validation loss of 0.0822. The model accuracy, shown in Figure 5, also increases for both training and validation set with every passing epoch and reaches a maximum of 0.974 and 0.973 respectively on the 25th epoch.

Training Details:

Found 68880 images belonging to 4 classes.

Found 5600 images belonging to 4 classes.

Found 5600 images belonging to 4 classes.

{'Normal': 3, 'Invasive': 2, 'In Situ': 1, 'Benign': 0}

Epoch 1/25

2153/2153 [==============================] - 1918s 891ms/step - loss: 0.6007 - acc: 0.7692 - val_loss: 0.2817 - val_acc: 0.9407

Epoch 2/25

2153/2153 [==============================] - 1894s 880ms/step - loss: 0.2564 - acc: 0.9171 - val_loss: 0.1545 - val_acc: 0.9414

Epoch 3/25

2153/2153 [==============================] - 1895s 880ms/step - loss: 0.1600 - acc: 0.9484 - val_loss: 0.0718 - val_acc: 0.9786

Epoch 4/25

2153/2153 [==============================] - 1895s 880ms/step - loss: 0.1129 - acc: 0.9635 - val_loss: 0.0765 - val_acc: 0.9711

Epoch 5/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0870 - acc: 0.9728 - val_loss: 0.1333 - val_acc: 0.9587

Epoch 6/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0595 - acc: 0.9809 - val_loss: 0.0774 - val_acc: 0.9705

Epoch 7/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0579 - acc: 0.9817 - val_loss: 0.0387 - val_acc: 0.9879

Epoch 8/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0415 - acc: 0.9873 - val_loss: 0.0325 - val_acc: 0.9902

Epoch 9/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0384 - acc: 0.9885 - val_loss: 0.0681 - val_acc: 0.9827

Epoch 10/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0380 - acc: 0.9887 - val_loss: 0.0235 - val_acc: 0.9932

Epoch 11/25

2153/2153 [==============================] - 1898s 881ms/step - loss: 0.0315 - acc: 0.9907 - val_loss: 0.0269 - val_acc: 0.9907

Epoch 12/25

2153/2153 [==============================] - 1898s 881ms/step - loss: 0.0264 - acc: 0.9919 - val_loss: 0.0426 - val_acc: 0.9862

Epoch 13/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0235 - acc: 0.9926 - val_loss: 0.0253 - val_acc: 0.9929

Epoch 14/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0240 - acc: 0.9931 - val_loss: 0.0263 - val_acc: 0.9918

Epoch 15/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0199 - acc: 0.9938 - val_loss: 0.0125 - val_acc: 0.9961

Epoch 16/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0187 - acc: 0.9947 - val_loss: 0.0190 - val_acc: 0.9948

Epoch 17/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0156 - acc: 0.9955 - val_loss: 0.0530 - val_acc: 0.9827

Epoch 18/25

2153/2153 [==============================] - 1898s 881ms/step - loss: 0.0154 - acc: 0.9957 - val_loss: 0.0258 - val_acc: 0.9907

Epoch 19/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0154 - acc: 0.9952 - val_loss: 0.0388 - val_acc: 0.9925

Epoch 20/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0177 - acc: 0.9949 - val_loss: 0.0805 - val_acc: 0.9686

Epoch 21/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0115 - acc: 0.9966 - val_loss: 0.0132 - val_acc: 0.9959

Epoch 22/25

2153/2153 [==============================] - 1896s 881ms/step - loss: 0.0131 - acc: 0.9964 - val_loss: 0.0156 - val_acc: 0.9961

Epoch 23/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0133 - acc: 0.9960 - val_loss: 0.0147 - val_acc: 0.9950

Epoch 24/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0101 - acc: 0.9970 - val_loss: 0.0127 - val_acc: 0.9952

Epoch 25/25

2153/2153 [==============================] - 1897s 881ms/step - loss: 0.0113 - acc: 0.9967 - val_loss: 0.0160 - val_acc: 0.9957 Found

68880 images belonging to 4 classes.

Found 5600 images belonging to 4 classes.

Found 5600 images belonging to 4 classes.

{'Normal': 3, 'Invasive': 2, 'In Situ': 1, 'Benign': 0}

{'Normal': 3, 'Invasive': 2, 'In Situ': 1, 'Benign': 0} {'Normal':

3, 'Invasive': 2, 'In Situ': 1, 'Benign': 0}

Loss = 0.744450330668

Test Accuracy = 0.90803571428

Chapter 8: Results
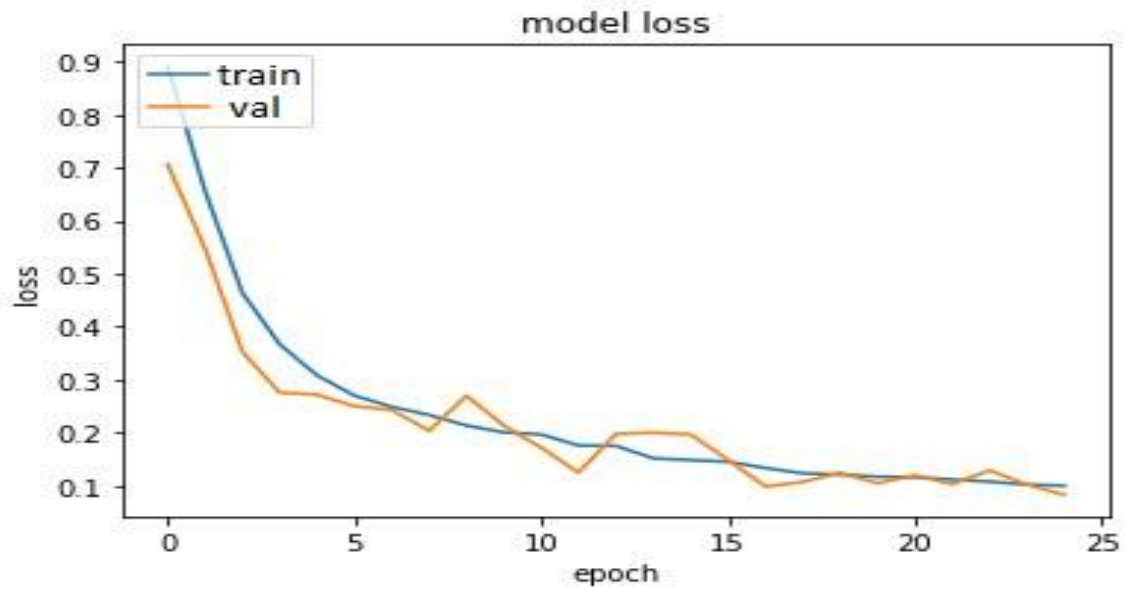
Model Loss:
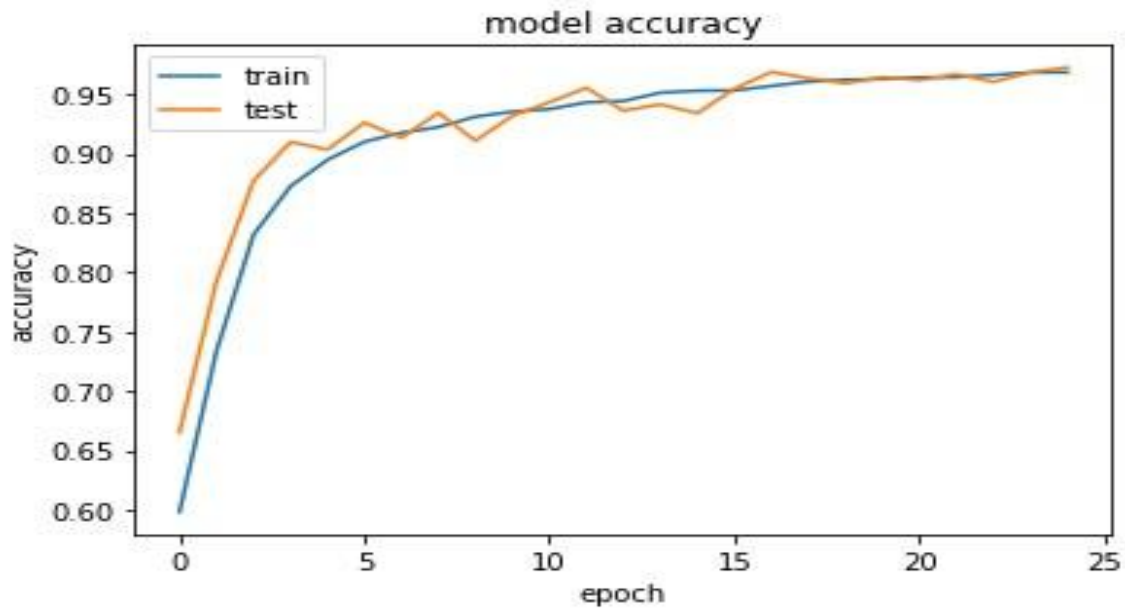

Fig. 11: Plot of training loss

Model Accuracy:

Fig. 12: Plot of training accuracy

Precision, Recall, F1 and AUROC scores:

	Precision	Recall	F1-Score	Support	AUROC
Benign	0.98	0.65	0.78	1400	0.9845
In Situ	0.84	0.99	0.91	1400	0.9982
Invasive	0.98	1.00	0.99	1400	1.00
Normal	0.87	1.00	0.93	1400	0.9989

Table 2: Precision, Recall, F1-Score and AUROC-Values
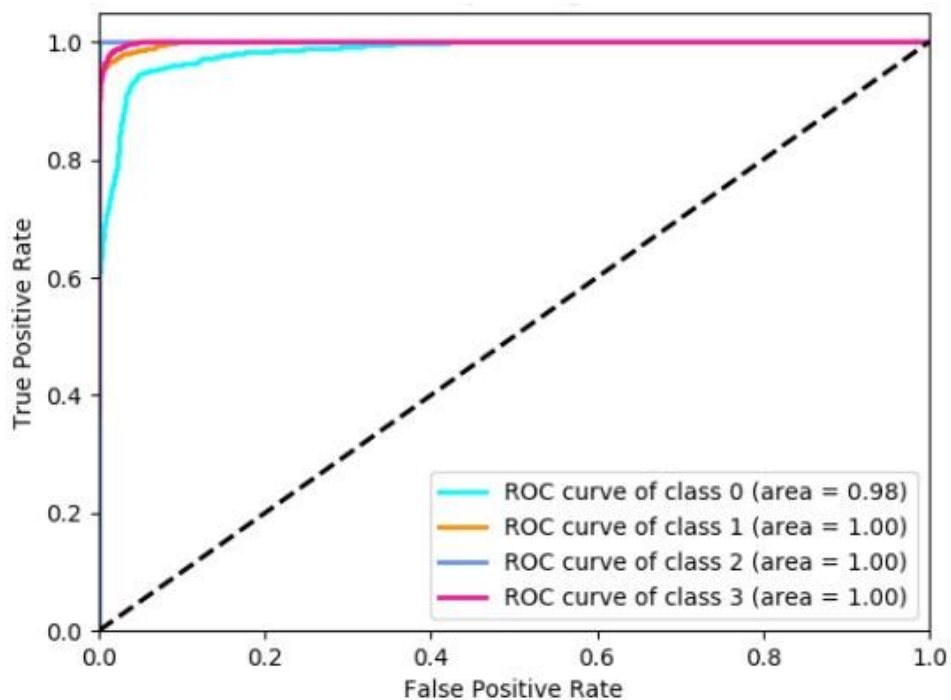
ROC Curve:

Fig. 13: The ROC Curves of the different classes. Class 0 - Benign, Class 1 - In Situ, Class 2 - Invasive, Class 3 - Normal

Comparisions with other works:

In the work of Araujo, Teresa et al , a CNN and SVM combination was used to classify breast cancer histology images. The accuracy achieved was 84% by the CNN+SVM model. The dataset used is the same dataset as the one in the current work. In the current work, we get an 89.5% accuracy in classifying the images by using DenseNet, which is a significant improvement.

In the work of Rakhlin et al, a Deep CNN approach was used to classify breast cancer histology images. The accuracy achieved was 87.22.6% by this process. In the current work, we get an 89.5% accuracy on the classification, which is again an improvement.

Chapter 9: Conclusion:

Here we aimed to predict the class of breast cancer from Macenko stain normalized images using a type of DNN (DenseNet). The major problem while using DNNs for classification is feature loss with increase in number of layers, and even the smallest section of the histopathological images contains immense information. Thus, it gets difficult and at times meaningless to apply DNNs on medical imaging. But, here with the use of DenseNet we were able to overcome this problem.

Further patching and augmentation not only helped us to increase the size of the training set which was and an important issue to resolve if we wanted to avoid overfitting, but also let us focus on every minute details in the image, as we only focus on a very small part of the tissue imaging in every instance. Our results are better than all the state-of-the-art approaches in this field, giving us a higher accuracy. In future we aim to apply this approach to various other types of medical images as well. If achieved the desired results, it can change our way of approaching medical reports.

References

1. Gao Huang, Zhuang, Liu, Laurens Van Der Maaten, Kilian Q. Weinberger, 2016. Densely Connected Convolutional Networks

2. Breast Cancer Statistics in Australia.

3. M. Macenko et al., "A method for normalizing histology slides for quantitative analysis," 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, 2009, pp. 1107-1110. doi: 10.1109/ISBI.2009.5193250

4. ICIAR 2015 Grand Challenge on Breast Cancer Histology Images.

5. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), pp.1929-1958.

6. Glorot, X. and Bengio, Y., 2010, March. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 249-256).

7. Loffe, S. and Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167

8. Rakhlin, A., Shvets, A., Iglovikov, V. and Kalinin, A.A., 2018. Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. arXiv preprint arXiv:1802.00752

9. Arajo T, Aresta G, Castro E, et al. Classification of breast cancer histology images using Convolutional Neural Networks. Sapino A, ed. PLoS ONE. 2017;12(6):e0177544. doi:10.1371/journal.pone.0177544

10. Gurcan MN, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B. Histopathological Image Analysis: A Review. IEEE reviews in biomedical engineering. 2009;2:147-171. doi:10.1109/RBME.2009.2034865

11. Wikipedia page on Histopathology

12. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG]