International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-4, April 2019

GPU Based Digital Histopathology and Diagnostic Support System for Breast Cancer Detection: A Comparison of CNN Models and Machine Learning Models

Mradul Kumar Jain, Nirvikar, Amit Kumar Agarwal

Abstract: ***One of the decisive reasons of cancer is uncontrolled augmentation of cancerous cells, malignant cells, or tumor cells in any living organism life at any stage. The scientific role of pathology is to diagnosis and prognosis of diseases to find changes at the level of cell structures with cell components including nucleolus and cytoplasm, tissues (i.e. grouped cell with complex structures) and organs which in turn give rise to the presenting signs and symptoms of the patient. It has been observed by clinical pathology system and histopathology methods that damaged or unrepaired cells do not die and show uncontrolled growth - a reason to mass development of cancerous cells. Frequently, cancerous cell travel through the blood and lymph systems, and cross the effected boundary organs to other body region where they repeats the process of uncontrolled growth cycle. This process of cancer cells leaving one region and growing in other part of body system is termed as metastatic spread or metastasis. Histopathological methodology can detect breast cancer. This diagnosis can be done with various Machine Learning Models and Deep Learning based Convolutional Neural Networks Models. The analysis shows convolutional neural networks models provide significant accurate results in comparison machine learning based models.***

Keywords: Digital histopathology images processing system (DHIPS), nonlinear mapping, principal component Analysis, CNN (Convolutional Neural Networks),Visual Geometry Group (VGG), AlexNet, GoogLeNet

I. INTRODUCTION

Histopathology means microscopic examination of cell's structure and tissues of infected organs by cancer. Histopathological microscopic image can be analyzed at cellular, sub-cellular, tissue and organs levels for the purpose to diagnosis and prognosis of disease. A medical technician or the pathologists may have the capability to identify the morphological characteristics of tissues to indicate the sign of carcinoma. The glass slides will be prepared by the biopsy sample and processed using microscope for image parameter and biological relevant parameter analysis. Viewing microscopic image, it is possible to detect cells, glands, nucleus and also able to discriminate [15] these shapes with normal vs. diseased tissue. This exercise is sufficient to detect the grading of infected cells or presence of carcinoma in tissues. On the basis of this prognosis patient treatment can be planned by considering the grading of breast cancer disease. The main aim of the author is to result out correct image analysis using machine learning algorithms and deep learning algorithms for digital histopathology imagery. A number of machine learning techniques can be used for breast histopathological image to find [7] image texture classification, feature extraction and segmentation, cell type identification, cell counting, cell classification and gland & nuclei segmentation. It will be used to calculate quantitative measurements of breast cancer disease parameters from histopathological images to determine the presence of carcinoma. This investigation can help by examining the [10] grading of breast cancer or severity of breast cancer in the biopsy sample. Digital histopathology can be used to detect different types of cancer including with their [31] grading in breast, pediatric tumor neuroblastoma, prostate, renal cell carcinoma and lung cancer. In this paper the author is analyzing and reviewing histopathological images of breast cancer. This analysis and methods are also applicable to all medical image modalities Computed Tomography, MRI, PET, X-ray, ultrasound and biomarkers. For Histopathology microscopic image analysis a good quality microscope for image magnification/resolution is needed. In our literature Machine Learning and Deep Learning Classification techniques for Histopathological images has been reviewed.

Manuscript published on 30 April 2019.
* Correspondence Author (s)
Mradul Kumar Jain, Dept. of Computer Science & Engineering, ABES Engineering College, Ghaziabad (U.P), India.
Dr. Nirvikar, College of Engineering Roorkee, Roorkee (Uttarakhand), India.
Amit Kumar Agrawal, ABES Engineering College, Ghaziabad (U.P), India.

II. BREAST CANCER SIGN AND SYMPTOMS

A new lump or a painless hard mass with irregular boundary over the breast is a most common type of symptoms. Breast can also be round in shape, tender in nature, soft in physical touch and painful. The other symptoms are [24] swelling a part or whole breast, [31] irritation over breast skin and nipple, change in nipple and breast shape, leakage of fluid other than milk from nipples, redness over breast skin. Sometime it has been observed these are symptom are present due to other medical issues even though they are not related to breast.

III. MEDICAL IMAGING MODALITIES

In the current era the following medical imaging methods are used to diagnose the diseases.

Computed Tomography (CT)

This is a medical imaging technology that takes numerous X-ray projections from different positions of angles and combines to create and produce comprehensive [28] cross-sectional images of inner body parts area. It provides very accurate images, three dimension views of particular body part. Generally it is used to take CT images of [1] soft tissues, bones, blood vessels, the heart, the lungs, the brain, the pelvis and abdomen. This method is used to diagnose different types of cancer in different body parts mainly in lung, breast, liver and pancreatic cancers. CT scan is used to examine [32] Bone injuries, Pulmonary embolism (CT angiography), Cardiac tissue, Vascular condition/blood flow, Organs in the pelvis, Chest, abdomen, traumatic injuries, presence, size and location of tumors, Cardiovascular disease and Colon health (CT colongraphy).

Magnetic Resonance Imaging (MRI)

It is a technology that uses radio waves and magnetic field to produce comprehensive and exhaustive images of tissues and organs. MRI is highly efficient to discriminate prognosis and diagnosing a lot of circumstances by presenting the [12] differentiation between normal soft tissues and diseased soft tissues of the inner body parts. MRI is used to examine Tendon and ligament tears, bones and joints, Organs in the male and female [30] pelvis, chest, heart, liver, kidney, spleen, abnormal tissue, blood vessels, Breasts and Spinal injuries

Positron Emission Tomography (PET)

A nuclear based imaging technology that has the capability to providing formation to physicians [23] or other medical technician regarding functioning of tissues and organs.

PET-CT

A coupling of CT and PET is medical imaging method which is used by technician to increase the precision. Positron Emission Tomography is used in combination with CT imaging, scanner and a small amount of radiopharmaceuticals is inserted into the [2] vein of a patient to build meticulous digital images of inside organs of the body. PET is used to examine Cancer, Heart conditions and Neurological diseases. This combination used to improve care of tumor with the help of positive and active treatment decisions, [17] monitoring of disease which occurs recursively and patient outcomes in terms of improvement and in terms of disease-free progression.

Ultrasound

An image based diagnostic method that uses high frequency sound waves which helps to create image of inner body parts is called medical sonography or ultra-sonography. This ultrasound machine transfers high frequency sound waves into [25] the inner body parts and it has the ability to convert the returning sound echoes into the form of a picture. On the basis of these images the practitioner may give the answer of query asked by patient. This technology has a capability to make capable of being heard sounds of blood flow to assess and examine patient's health, swelling, infection, pregnancy, symptoms of pain and abnormalities of heart and blood vessels.

X-Ray

An X-ray is a oldest imaging technology which provides an image in different shades of black and white color. X-rays uses ionizing radiation to produce internal structure images of a person by sending X-ray (a form of electronic radiation) beams through the body, which are absorbed in different amounts of radiation depending on the density of the material. Radiation Therapy [14] is a type of device which also utilizes x-rays, gamma rays, electron beams or protons to treat cancer. It is used to examine Lungs, Blood vessels, Broken bones, Cavities, Swallowed objects, and Breast (mammography)

Cytology

Cytology is the body cell's examination through microscope. Cytology is related to Cell biology [16] which is associated with cell's functions, structures and related chemistry. The purpose of cytology is biopsy and disease screening. Cytology image set can also be used to locate isolated cells and cell structures.

Histopathology

Histopathology processing is a microscopic examination making biopsy glass slide to visualize all the component of tissues to investigate the symptoms of diseases. The tissue undergoes a series of steps which are [3, 6] tissue handling, tissue processing & tissue staining. Staining procedure is used to provide more clarity in its cell and its objects like cell nucleolus and cell boundary under microscopic examination For this purpose we need to dye or chemical sections with one or more stains. According to a large number of researchers and pathologists used Hematoxylin to stain cell [22] nucleolus blue and Erosin stains cytoplasm and connective tissue pink.

IV. COMPARATIVE STUDY OF CNN MODELS AND ML MODELS

The architecture of AlexNet which is combination of 5 convolutional layers and 3 fully connected layers uses ReLu function (f(x) = max (0, x)) instead of Hyperbolic tangent function or Sigmoid function. The major advantage of ReLu function is much [4] faster training than traditional neural networks over Sigmoid or hyperbolic function.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

It happens because the derivation of these functions is very small in the saturating reason and became the reason for vanishing of weight. VGG architecture as shown in figure1 is consists of 16 convolutional layers as a combination of Convolutions layers (used only 3*3 size), Max pooling layers (used only 2*2 size), fully connected layers at end. It is very similar to AlexNet which uses 3x3 convolutions. It is one of the preferred artificial intelligence machines to extract features from the histopathological images. It takes input image of size 224 * 224 * 3 (RGB image).

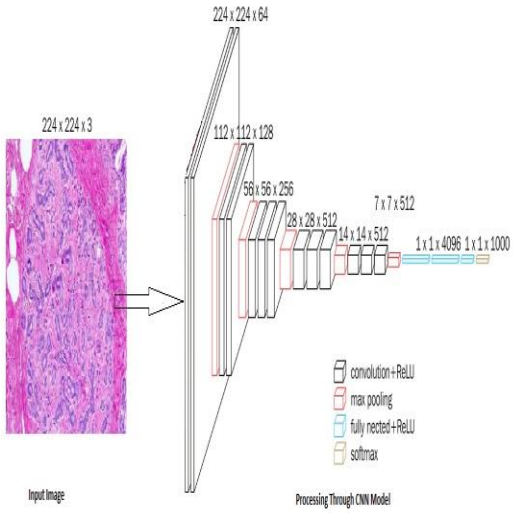


Figure1. CNN Model Processing

GoogLeNet or Inception model is a 22 Deep Layer Architecture. This model has the capability to classify the image set in to 1000 objects. It takes input image same as VGG which is of size 224x224. An inception module is the basic building block of the network and one inception module contains multiple convolutions [29] of different filter sizes with the use pooling only on single layer. The approach of GoogleNet is exclusively different from AlexNet and VGG because it is a combination of inception modules (each inception module contains some pooling, convolutions of dissimilar scales) and concatenation operations. GoogLeNet uses 1x1 feature convolutions which is used for feature selectors. This model is intelligent enough to figure out which layer is best suitable to give its best result after training. SVM is a method based on hyper plane to classify the objects in two classes or more classes in a multidimensional data space. It is very obvious that there will be a difference in skill set, knowledge and experience in different-different [8] pathologists, clinicians and other medical technician to interpret clinical data. As a result it may be misleading to patient or undirected disease prognosis. It is because of linked biology behind all process. In many cases it has been find a combination of PCA with SVM provides better results or increases the performance of SVM sometimes it degrades the SVM performance depending on size of training and testing data.PCA can be used when you do not have well defined data as well as unable to find [8] patterns and correlation between them. In such a case it is required data dimensionally reduction by the classification of correlated data. In many cases it is fundamentally difficult to understand the cases like satellite images, in such cases PCA should not be useful. It may be good for analysis of cancer images related tissues or organ's cell. KNN is one of the best classification technique based on data mining. It is identified as one of the top ten classifications by data mining method research community. KNN is non-parametric based classification technique. In this technique the [22] Euclidean distance between features vector of training images and feature vector of testing image is calculated. The Accuracy of KNN technique is depending on noise and unnecessary features. For this purpose it is required to give efforts during selection of features.

V. PREPARATION FOR HISTOPATHOLOGY

By using biopsy sample we can obtain digital histopathology image to examine the images manually for prognosis and diagnosis. Using DHIPS it is possible to overcome the problem leading to our objective to detect the presence of breast cancer with its [13] grading and severity. The DHIPS are capable enough for image segmentation, feature extraction, classification etc. For this purpose DHIPS used image preprocessing, feature selection &extraction and different types of classification techniques such as thersholding, morphological processing, and boundary based, and region based, and supervised classification techniques. The CNN model for DHIPS is shown in figure 2.

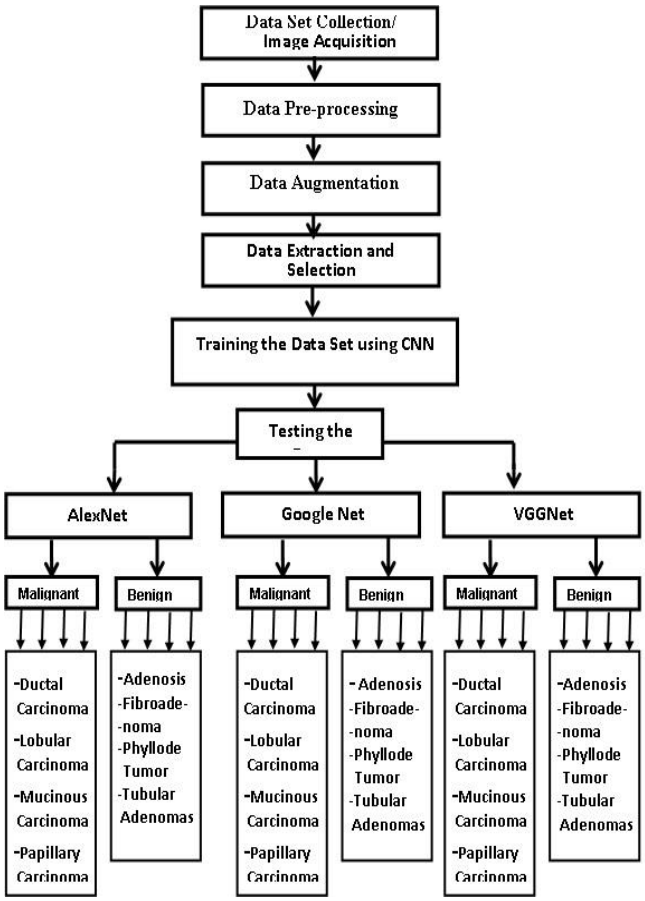


Figure 2. Classification process for Convolution Neural Networks Models

The current improvement of 'digital histopathology' is to desire development of quantitative and automated computerized image analysis [27] algorithms to assists histopathologist in interpreting the large number of digitized histopathological images. Histopathology samples can be prepared in five basic steps which are given below.

Fixing: Samples of The biological tissue will be made "fixed" using chemical fixation to preserve the cells or tissue.

Processing: A dehydration process will be used to remove water from [11] tissue and make the sample in the form of medium type solid state then tissues are available for section cutting.

Embedding sample in wax: Sections will be embedded with hardened wax.

Sectioning: Embedded samples of tissues are used for sectioning to make thin slices to prepare glass slide for taking microscopic image details of cells/tissues.

Staining: In the last, the mounted sections will be by an appropriate Histopathology stain. Staining is a process which is applied on biological tissues to improve the contrast of the tissue and highlighted specific features of interest related to tissues and the stain used.

VI. ABOUT BREAKHIS DATASET

In this paper we have used Breast Cancer Microscopic Image Dataset namely by BreakHis. This dataset of histopathological microscopic images of 82 patients with detail only related to breast cancer is shown in Table1. In the hierarchy [14] of classes at first level it contains two classes named by Benign and Malignant. Number of images in Begin class is 2480 taken from 24 patients and number of images in malignant category is 5429 taken from 58 patients. These all image are further categorized in different subclasses. Each set of sub class is further classified at [4] different magnifying factor 40X, 100X, 200X, and 400X. All the images are of 700X460, 3-channel RGB and 8 bit depth of each channel in PNG format. The entire image set is created by SOB method. Generally the tissue taken in this method is large as compared to other method which uses needle biopsy. BreakHis data set will be divided in to two parts in ratio of 80% and 20% for the preparation of microscopic image training set and testing set. Data set contains images for 82 patients that can be partitioned in two sets 58 and 24 patients for training and testing respectively. Training image set of 58 patients [13] is further categorized with 46 malignant cases and 12 benign cases. Testing image set of 24 patients is further categorized with 19 malignant cases and 5 benign cases. It is also in consideration to classify the image set in the same ratio on the basis of magnification factor of 400x, 200x, 100x & 40x.

Table 1. Breast Cancer Microscopic Image Dataset

Magnify--ing Factors	Benign Tumor Image Set	Malignant Tumor Image Set	Total No. of Images
Image Set at 40x	625	1370	1995
Image Set at 100x	644	1437	2081
Image Set at 200x	623	1390	2013
Image Set at 400x	588	1232	1820
Total of Images	2480	5429	7909
No. of Patient	24	58	82

The benign tumor is further classified in to 4 types described further. Adenosis- a non-cancerous or benign stage of breast cancer. It can be observed when milk producing glands or lobules size becomes large or increase in number of glands. Another observation is the presence of cysts or fibrosis in the breast of women. Fibroadenoma- This is another category benign breast tumor which is most common in [18] younger women or premenopausal women. Sometime it has also been observed the presence in any age group of female. It is painless as well as at low risk and round or oval in shape.Phyllodes tumor- It is very rare type of breast tumor in benign category. Generally it has been observed in female at the age nearby 40 or premenopausal women. It is difficult to predict the behavior of Phyllodesin female breast tumors. Phyllodes tumors are benign in nature but sometimes it has been observed suspicious to malignancy. Tubular adenomas- It is another common type borderline noncancerous cell but sometime cancer develops in adenomas then it is known as tubular adenomas.These all subclasses are further classified for magnification power of 40x, 100x, 200x and 400x. Malignant tumors are actually a cancer which is a seriously cause to health and lead to death. The major subclasses of malignant cancer are ductal carcinoma, lobularcarcinoma, mucinous carcinomaand papillary carcinoma.

VII. DATA AUGMENTATION AND PREPARATION

The microscopic histopathological image dataset named by BreakHis is freely available on the web which is mainly classified into two groups of benign tumors and malignant tumors present in breast cancer. Benign means slowly growing of cancerous cell or innocent or localized cancerous cell. While the term malignancy is related to cancer or fast growing of cancerous cell marked by [18] cellular atypia, mitosis, destroy of adjacent structure, disruption of basement membranes, spread in nearby area known as metastasize. These all are the cause of [11] major disease and lead to death.The data set has been created by SOB method and also known as partial mastectomy or excisional biopsy. This type of method needs to remove larger tissue sample as compared to needle biopsy. Each image is available at the magnification power 40x, 100x, 200x and 400x. Data is available in [20] multiclasses unbalanced form, which requires multi scale microscopic image data augmentation and over sampling method to ignore the problem of unbalanced multi classes and overfitting.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

In this paper augmentation process has been used by using intensity variation from – 0.1 to +0.1, by rotation of image from −90° to +90°, translation with ±20 pixels and flipping of level, flipping of direction.

VIII. EXPERIMENTAL FRAMEWORK

First all the images of BreakHis data set has been rescaled to size of 224X224 so that they can be fed to CNN models and ML models smallest side equals 224 before fed into the model. BreakHis data set is created in two sets for the purpose of training and testing in the ratio of 80% and 20% for each cancer category and subcategory of each type of cancer. The author is using ML based models like KNN, SVM and PCA and deep learning based convolutional pretrained neural network models like AlexNet, GoogLeNet and VGG for histopathological [15] breast cancer image classification with a significant improvement in accuracy in comparison to machine learning based methods. The CNN model is composed of 16 layers shown in figure 3 containing with linear and non-linear operators in three category, convolution layers with dimension (number of filters × filter height × filter depth: 64×3×3, 128×3×3, 256×3×3,512×3×3), max pulling layers with dimension size 2x2 and at the end [21] two fully connected layers with hidden node. These sixteen layers used filters as in the sequence. First two layers used 64 filters, next two layers used 128 filters, next three layers used 256 filters and next six layers used 512 filters respectively. Layer no. 2, 4, 7, 10 and 13 used max pooling functions. Layer 14 and 15 is fully connected layers with 4096 nodes. In the end there is only one layer with softmax function called output layer with 1000 nodes. The deep nature of this network, it provides better results. CNN models are fit for localization task and image classification task. Localization [5] means to find the certain object in the image described by boundary box and the classification means to find what the object is in the input image. The author have used two category of models based on entirely different approach, methodology and technique to calculate accuracy for the purpose of image classification of microscopic histopathological image set with magnification factor of 40X in benign cancer category. There are four types of benign cancer which are Adenosis, Fibroadenoma, Phyllodes Tumor, and Tubular Adenomas.

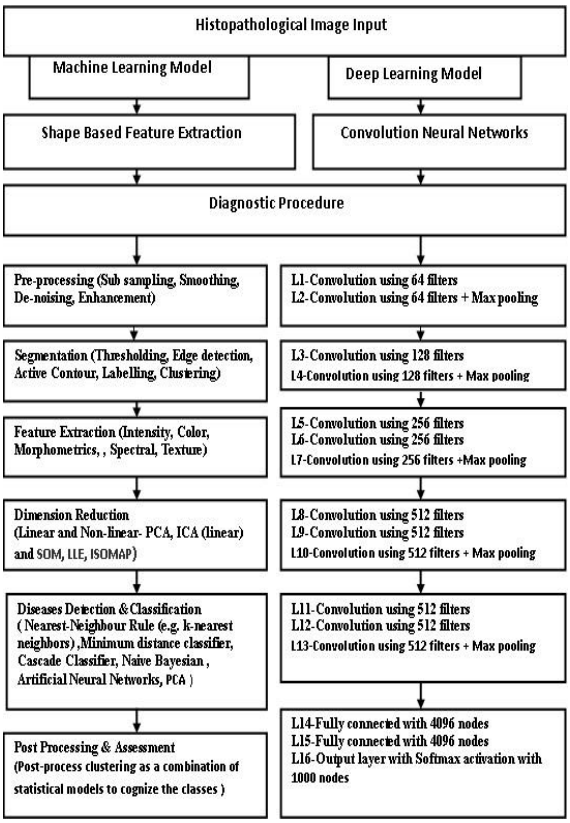


Figure 3. Comparative frame work of convolutional Neural Network and Machine Learning

Table2. Classification on the Basis of Accuracy for 40X Benign Cancer Image Set

Test For Benign	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGGm**	**AlexNet**	**SVM**	**PCA**	**KNN**
Adenosis	**89.6**	**86.2**	**83.4**	**72.7**	**80.7**	**76.8**
Fibroadenoma	**81.7**	**83.7**	**87.3**	**77.8**	**76.9**	**79.9**
Phyllodes Tumor	**90.3**	**91.7**	**91.1**	**78.4**	**81.4**	**78.3**
Tubular Adenomas	**90.2**	**87.5**	**92.3**	**88.5**	**77.8**	**71.9**
Average	**87.95**	**87.28**	**88.53**	**79.35**	**79.2**	**76.73**

Table3. Classification on the Basis of Accuracy for 100X Benign Cancer Image Set

Test For Benign	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Adenosis	**86.7**	**78.3**	**84.7**	**78.4**	**73.6**	**82.8**
Fibroadenoma	**86.9**	**90.7**	**94.7**	**82.8**	**85.7**	**74.7**
Phyllodes Tumor	**82.6**	**93.3**	**79.2**	**84.6**	**78.7**	**78.5**
Tubular Adenomas	**92.3**	**91.9**	**92.3**	**72.8**	**76**	**73.7**
Average	**87.13**	**88.55**	**87.73**	**79.65**	**78.5**	**77.43**

From the Table2 and Figure4, it can be seen clearly that the average accuracy for all types of Benign cancer category of images under 40X magnification factor for Convolution Neural Network based model is 87.91 and for Machine Learning based model is 78.42. The highest accuracy is 88.53 for AlexNet among all types of CNN models. On the other hand the highest accuracy is 76.73 for KNN among all types of Machine Learning Models.

Information shown in the Table3. and Figure.5, the average accuracy for all types of Benign cancer category of images under 100X magnification factor for Convolution Neural Network based model is 87.80 and for Machine Learning based model is 78.52. The highest accuracy is 94.7 for AlexNet among all types of CNN models. On the other hand the highest accuracy is 85.7 for PCA among all types of Machine Learning Models.

Figure 4. Classification on the basis on basis of Accuracy for 40X Benign Cancer Image Set

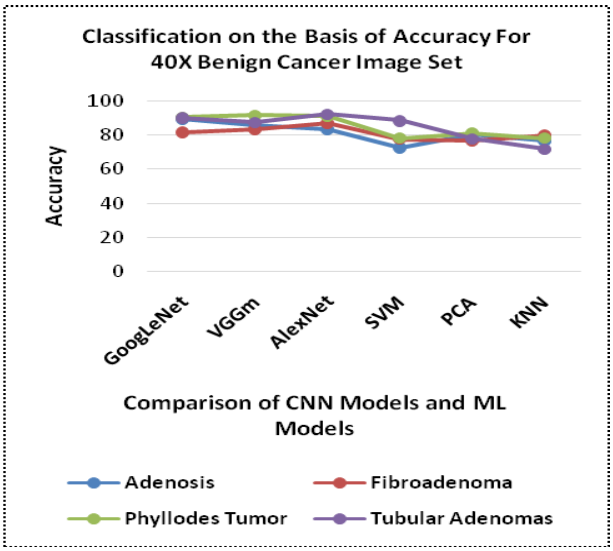


Figure 5. Classification on the basis on basis of Accuracy for 100X Benign Cancer Image Set

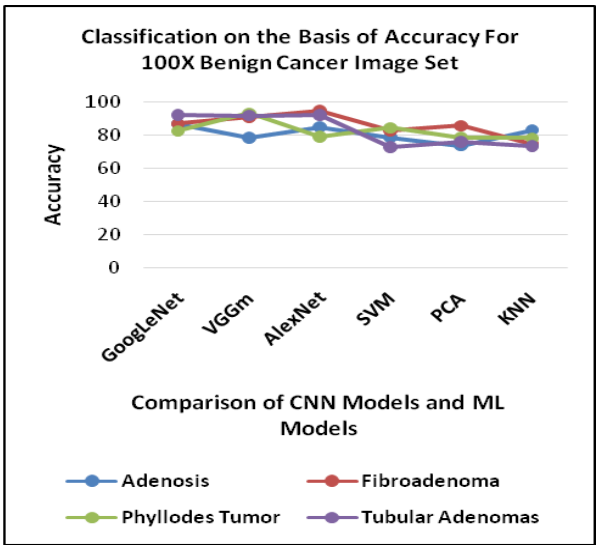


Table4. Classification on the Basis of Accuracy for 200X Benign Cancer Image Set

Test For Benign	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Adenosis	**93.2**	**93.3**	**87.5**	**74.6**	**75.8**	**76.8**
Fibroadenoma	**77.9**	**84.7**	**88.1**	**79.2**	**78.6**	**76.1**
Phyllodes Tumor	**84.2**	**91.2**	**87.5**	**77**	**74.9**	**75.7**
Tubular Adenomas	**91.4**	**81.8**	**88.2**	**81.2**	**79.6**	**78.7**
Average	**86.68**	**87.75**	**87.83**	**78**	**77.23**	**76.83**

Table5. Classification on the Basis of Accuracy for 400X Benign Cancer Image Set

Test For Benign	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Adenosis	**87.2**	**93.3**	**91**	**83.8**	**73.9**	**81.4**

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Fibroadenoma	**90.5**	**80.7**	**87.2**	**78.8**	**73.7**	**75.8**
Phyllodes Tumor	**81.8**	**81.1**	**82.1**	**77.1**	**78.5**	**72.7**
Tubular Adenomas	**87.2**	**89.2**	**86.4**	**72.8**	**70.7**	**74.1**
Average	**86.68**	**86.08**	**86.68**	**78.13**	**74.20**	**76.00**

The analysis of Table4. and Figure.6 clears the picture that the average accuracy for all types of Benign cancer category of images under 200X magnification factor for Convolution Neural Network based model is 87.42 and for Machine Learning based model is 77.35. The highest accuracy is 91.4 for GoogLeNet among all types of CNN models. On the other hand the highest accuracy is 81.2 for SVM among all types of Machine Learning Models.

Figure 6. Classification on the basis on basis of Accuracy for 200X Benign Cancer Image Set

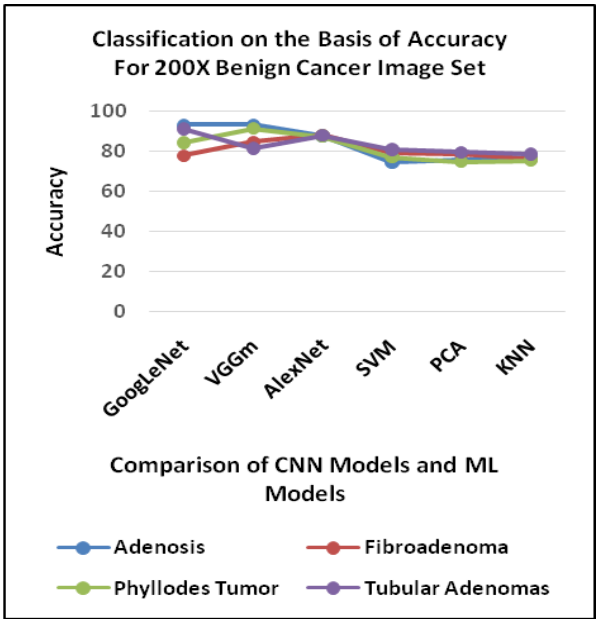


Table5. and Figure.7 presents average accuracy for all types of Benign cancer category of images under 400X magnification factor for Convolution Neural Network based model is 86.48 and for Machine Learning based model is 76.11. The highest accuracy is 90.5 for GoogLeNet among all types of CNN models. On the other hand the highest accuracy is 83.8 for SVM among all types of Machine Learning Models.

Figure 7. Classification on the basis on basis of Accuracy for 400X Benign Cancer Image Set

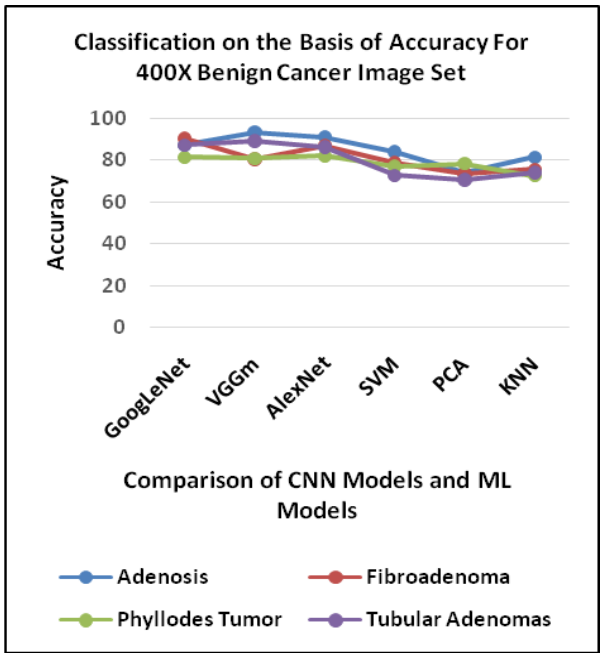


Table6. Classification on the Basis of Accuracy for 40X Malignant Cancer Image Set

Test For Malignant	**Convolution Neural Network**			**Machine Learning Model**		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Ductal Carcinoma	**90.6**	**93.5**	**89.1**	**80.7**	**83.2**	**79.7**
Lobular Carcinoma	**81.2**	**82.7**	**94.8**	**73.5**	**77.8**	**78.4**
Mucinous Carcinoma	**83.2**	**87.4**	**82.4**	**73.3**	**77.1**	**77.3**
Papillary Carcinoma	**91.7**	**86.3**	**85.1**	**81.6**	**74.1**	**71.2**

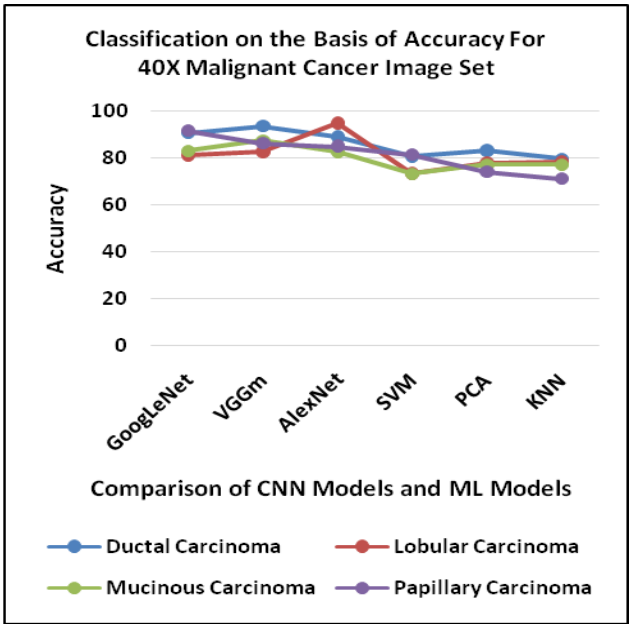
Average	**86.68**	**87.48**	**87.85**	**77.28**	**78.05**	**76.65**

Table7. Classification on the Basis of Accuracy for 100X Malignant Cancer Image Set

Test For Malignant	**Convolution Neural Network**			**Machine Learning Model**		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Ductal Carcinoma	**85.9**	**87.6**	**92.3**	**76.8**	**73.1**	**75.8**
Lobular Carcinoma	**91.5**	**86.7**	**86.8**	**77.8**	**80.1**	**76.8**
Mucinous Carcinoma	**85.2**	**87.1**	**86.3**	**75.8**	**78.1**	**74.3**
Papillary Carcinoma	**86.1**	**92.6**	**80.2**	**78.7**	**74.4**	**72.3**
Average	**87.18**	**88.50**	**86.40**	**77.28**	**76.43**	**74.80**

From the Table 6. and Figure.8, data shows that the average accuracy for all types of Malignant cancer category of images under 40X magnification factor for Convolution Neural Network based model is 87.33 and for Machine Learning based model is 77.33. The highest accuracy is 94.8 for AlexNet among all types of CNN models. On the other hand the highest accuracy is 83.2 for PCA among all types of Machine Learning Models.

Figure 8. Classification on the basis on basis of Accuracy for 40X Benign Cancer Image Set



In the Table 7. and Figure.9, it can be seen clearly that the average accuracy for all types of Malignant cancer category of images under 100X magnification factor for Convolution Neural Network based model is 87.36 and for Machine Learning based model is 76.17. The highest accuracy is 92.6 for VGGm among all types of CNN models. On the other hand the highest accuracy is 80.1 for PCA among all types of Machine Learning Models.

Figure 9. Classification on the basis on basis of Accuracy for 100X Benign Cancer Image Set

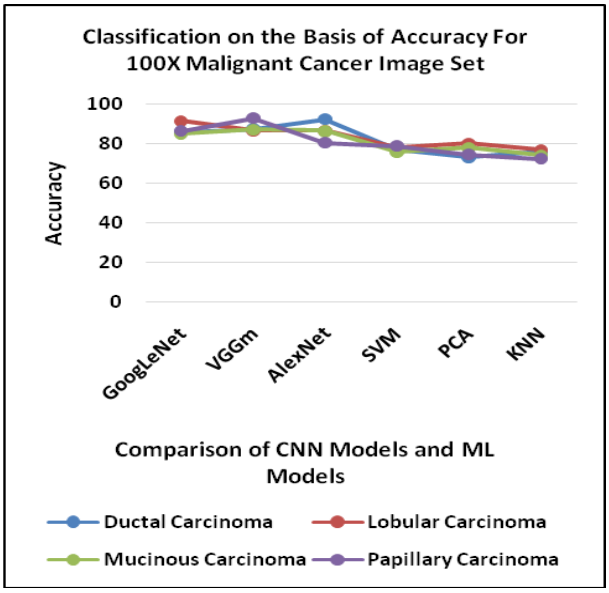


Table8. Classification on the Basis of Accuracy for 200X Malignant Cancer Image Set

Test For Malignant	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Ductal Carcinoma	83.2	82.8	86.2	73.2	78.1	70.8
Lobular Carcinoma	83.8	83.5	84.7	74.5	76.1	75.1
Mucinous Carcinoma	82.3	82.1	94.3	76.7	71.6	73.8
Papillary Carcinoma	84.6	82.3	85.1	72.5	77.2	72.5
Average	83.48	82.68	87.58	74.23	75.75	73.05

Table9. Classification on the Basis of Accuracy for 400X Malignant Cancer Image Set

Test For Malignant	Convolution Neural Network			Machine Learning Model		
Type of Method	**GoogLeNet**	**VGG**	**AlexNet**	**SVM**	**PCA**	**KNN**
Ductal Carcinoma	85.2	86.3	86.2	76.1	75.6	76.3
Lobular Carcinoma	82.7	84.4	80.5	74.8	78.8	72.9
Mucinous Carcinoma	87.5	79.5	85.5	73.8	79.2	79.5
Papillary Carcinoma	78.9	83.2	84.2	75	75.9	76.7
Average	83.58	83.35	84.10	74.93	77.38	76.35

From the Table 8. and Figure.10, it can be seen clearly that the average accuracy for all types of Malignant cancer category of images under 200X magnification factor for Convolution Neural Network based model is 84.58 and for Machine Learning based model is 74.34. The highest accuracy is 94.3 for AlexNet among all types of CNN models. On the other hand the highest accuracy is 78.1 for PCA among all types of Machine Learning Models.

From the Table 11. and Figure.8, it can be seen clearly that the average accuracy for all types of Malignant cancer category of images under 400X magnification factor for Convolution Neural Network based model is 83.68 and for Machine Learning based model is 76.22. The highest accuracy is 86.3 for VGGm among all types of CNN models. On the other hand the highest accuracy is 76.22 for PCA among all types of Machine Learning Models.

Figure 10. Classification on the basis on basis of Accuracy for 40X Benign Cancer Image Set

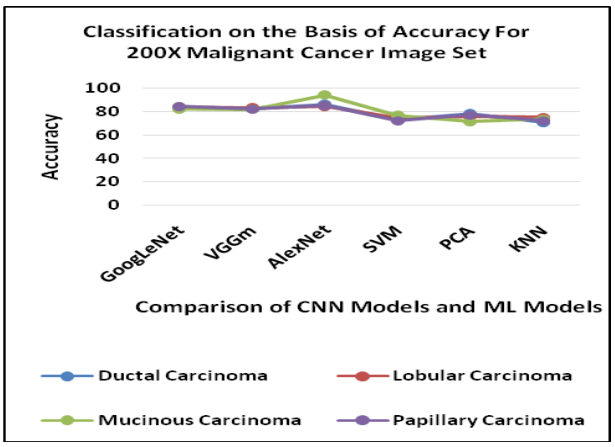
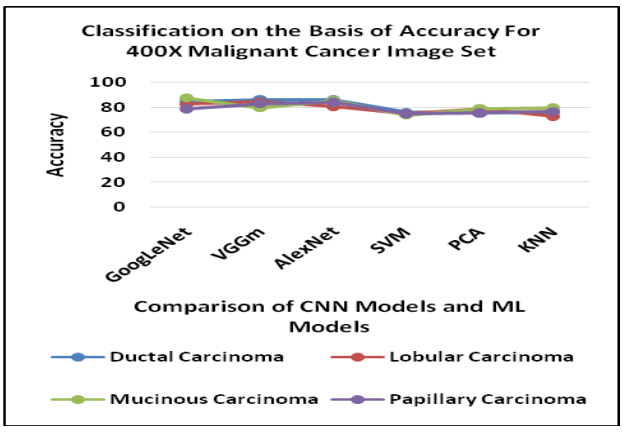


Figure 11. Classification on the basis on basis of Accuracy for 40X Benign Cancer Image Set



IX. CONCLUSION

This research work proposed to enhance image analysis to segment objects and detect other region of interest in histopathological breast cancer images. We have used feature based classification (like colour texture feature, nuclei detection etc.) models. The image analysis completed by two different approaches based on convolutional neural networks and machine learning techniques. There are many rule based machine learning and deep learning methods for digital histopathology image analysis. It can also be observed that deep learning perform much better in comparison to machine learning. In the field of medical image processing CNN has been used successfully to predict can like diseases. This method with a comprehensive analysis may be useful to minimize the subjective issue of knowledge or experience for medical technicians. We provide a complete picture for all types of breast cancer carcinomas in benign and malignant category with different type of images constructed on different magnification factors. We believe that this this experiment and study may useful to breast cancer prognosis and diagnosis of breast cancer.

REFERENCES

1. Komura, Daisuke, and Shumpei Ishikawa. "Machine learning methods for histopathological image analysis." Computational and structural biotechnology journal 16 (2018): 34-42.
2. Cruz-Roa, Angel, Hannah Gilmore, Ajay Basavanhally, Michael Feldman, ShridarGanesan, Natalie NC Shih, John Tomaszewski, Fabio A. González, and AnantMadabhushi. "Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent." Scientific reports 7 (2017): 46450.
3. Liu, Yun, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, TimoKohlberger, Aleksey Boyko, SubhashiniVenugopalan et al. "Detecting cancer metastases on gigapixel pathology images." arXiv preprint arXiv:1703.02442 (2017).
4. Spanhol, Fabio Alexandre, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. "Breast cancer histopathological image classification using convolutional neural networks." In 2016 International Joint Conference on Neural Networks (IJCNN), pp. 2560-2567. IEEE, 2016.
5. Araújo, Teresa, GuilhermeAresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, AntónioPolónia, and AurélioCampilho. "Classification of breast cancer histology images using convolutional neural networks." PloS one 12, no. 6 (2017): e0177544.
6. Cireşan, Dan C., Alessandro Giusti, Luca M. Gambardella, and Jürgen Schmidhuber. "Mitosis detection in breast cancer histology images with deep neural networks." In International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 411-418. Springer, Berlin, Heidelberg, 2013.
7. Albarqouni, Shadi, Christoph Baur, Felix Achilles, VasileiosBelagiannis, Stefanie Demirci, and Nassir Navab. "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images." IEEE transactions on medical imaging 35, no. 5 (2016): 1313-1321.
8. Cruz-Roa, Angel, Ajay Basavanhally, Fabio González, Hannah Gilmore, Michael Feldman, ShridarGanesan, Natalie Shih, John Tomaszewski, and AnantMadabhushi. "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks." In Medical Imaging 2014: Digital Pathology, vol. 9041, p. 904103. International Society for Optics and Photonics, 2014.
9. Wang, Haibo, Angel Cruz Roa, Ajay N. Basavanhally, Hannah L. Gilmore, Natalie Shih, Mike Feldman, John Tomaszewski, Fabio Gonzalez, and AnantMadabhushi. "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features." Journal of Medical Imaging 1, no. 3 (2014): 034003.
10. Xu, Jun, Xiaofei Luo, Guanhao Wang, Hannah Gilmore, and AnantMadabhushi. "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images." Neurocomputing 191 (2016): 214-223.
11. Malon, Christopher D., and Eric Cosatto. "Classification of mitotic figures with convolutional neural networks and seeded blob features." Journal of pathology informatics 4 (2013).
12. Madabhushi, Anant, and George Lee. "Image analysis and machine learning in digital pathology: challenges and opportunities." (2016): 170-175.
13. Rakhlin, Alexander, Alexey Shvets, Vladimir Iglovikov, and Alexandr A. Kalinin. "Deep convolutional neural networks for breast cancer histology image analysis." In International Conference Image Analysis and Recognition, pp. 737-744. Springer, Cham, 2018.
14. Sirinukunwattana, Korsuk, Shan e Ahmed Raza, Yee-Wah Tsang, David RJ Snead, Ian A. Cree, and Nasir M. Rajpoot. "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images." IEEE Trans. Med. Imaging 35, no. 5 (2016): 1196-1206.
15. Huynh, Benjamin Q., Hui Li, and Maryellen L. Giger. "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks." Journal of Medical Imaging 3, no. 3 (2016): 034501.
16. Bejnordi, BabakEhteshami, MitkoVeta, Paul Johannes Van Diest, Bram Van Ginneken, Nico Karssemeijer, Geert Litjens, Jeroen AWM Van Der Laak et al. "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer." Jama 318, no. 22 (2017): 2199-2210.
17. Litjens, Geert, Clara I. Sánchez, NadyaTimofeeva, MeykeHermsen, Iris Nagtegaal, Iringo Kovacs, Christina Hulsbergen-Van De Kaa, Peter Bult, Bram Van Ginneken, and Jeroen Van Der Laak. "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis." Scientific reports 6 (2016): 26286.
18. Arevalo, John, Fabio A. González, Raúl Ramos-Pollán, Jose L. Oliveira, and Miguel Angel Guevara Lopez. "Representation learning for mammography mass lesion classification with convolutional neural networks." Computer methods and programs in biomedicine 127 (2016): 248-257.
19. Kowal, Marek, PawełFilipczuk, Andrzej Obuchowicz, JózefKorbicz, and Roman Monczak. "Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images." Computers in biology and medicine 43, no. 10 (2013): 1563-1572.
20. Wang, Dayong, Aditya Khosla, RishabGargeya, HumayunIrshad, and Andrew H. Beck. "Deep learning for identifying metastatic breast cancer." arXiv preprint arXiv:1606.05718(2016).
21. Wei, Benzheng, Zhongyi Han, Xueying He, and Yilong Yin. "Deep learning model based breast cancer histopathological image classification." In 2017 IEEE 2nd international conference on cloud computing and big data analysis (ICCCBDA), pp. 348-353. IEEE, 2017.
22. Chen, Hao, Qi Dou, Xi Wang, Jing Qin, and Pheng Ann Heng. "Mitosis detection in breast cancer histology images via deep cascaded networks." In Thirtieth AAAI Conference on Artificial Intelligence. 2016.
23. Arevalo, John, Fabio A. González, Raúl Ramos-Pollán, Jose L. Oliveira, and Miguel Angel Guevara Lopez. "Convolutional neural networks for mammography mass lesion classification." In 2015 37th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp. 797-800. IEEE, 2015
24. Veta, Mitko, Paul J. Van Diest, Stefan M. Willems, Haibo Wang, AnantMadabhushi, Angel Cruz-Roa, Fabio Gonzalez et al. "Assessment of algorithms for mitosis detection in breast cancer histopathology images." Medical image analysis 20, no. 1 (2015): 237-248.
25. Jiao, Zhicheng, XinboGao, Ying Wang, and Jie Li. "A deep feature based framework for breast masses classification." Neurocomputing 197 (2016): 221-231.
26. Carneiro, Gustavo, Jacinto Nascimento, and Andrew P. Bradley. "Unregistered multiview mammogram analysis with pre-trained deep learning models." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 652-660. Springer, Cham, 2015.
27. Bar, Yaniv, IditDiamant, Lior Wolf, Sivan Lieberman, Eli Konen, and Hayit Greenspan. "Chest pathology detection using deep learning with non-medical training." In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 294-297. IEEE, 2015.
28. Zheng, Yushan, Zhiguo Jiang, FengyingXie, Haopeng Zhang, Yibing Ma, Huaqiang Shi, and Yu Zhao. "Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification." Pattern Recognition 71 (2017): 14-25.
29. Spanhol, Fabio A., Luiz S. Oliveira, Paulo R. Cavalin, Caroline Petitjean, and Laurent Heutte. "Deep features for breast cancer histopathological image classification." In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1868-1873. IEEE, 2017.
30. Hatipoglu, Nuh, and GokhanBilgin. "Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships." Medical & biological engineering & computing 55, no. 10 (2017): 1829-1848.
31. Wahab, Noorul, Asifullah Khan, and Yeon Soo Lee. "Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection." Computers in biology and medicine 85 (2017): 86-97.
32. Li, Wenqi, SiyamalanManivannan, Shazia Akbar, Jianguo Zhang, Emanuele Trucco, and Stephen J. McKenna. "Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks." In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1405-1408. IEEE, 2016.