Detection And Analysis of Breast Cancer Using Convolution Neural Network for Mammogram Imaging System

A DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF DEGREE

OF

MASTER OF TECHNOLOGY

IN

SOFTWARE ENGINEERING

SUBMITTED BY

Rajat Singh Arora
(2K20/SWE/18)

Under the supervision of

Dr. Abhilasha Sharma
(Assistant Professor)

DEPARTMENT OF SOFTWARE ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY, 2022

CANDIDATE'S DECLARATION

I, Rajat Singh Arora, Roll No. 2K20/SWE/18 student of MTech (Software Engineering), hereby declare that the project Dissertation titled " Detection And Analysis Of Breast Cancer Using Convolutional Neural Networks For Mammogram Imaging System" submitted by me to the Department of Software Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology in Software Engineering, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any degree, diploma associateship, fellowship or other similar title or recognition.

Rajat

Place: Delhi

Rajat Singh Arora

Date: 30/05/2022

CERTIFICATE

I hereby certify that the Project Dissertation titled "Detection and Analysis of Breast Cancer Using Convolutional Neural Networks for Mammogram Imaging System" submitted by Rajat Singh Arora, 2K20/SWE/18 in Department of Software Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology in Software Engineering, is a record of the project work carried out by him under my supervision. To the best of my knowledge this work has not been submitted in part or full for any degree or diploma to this university or elsewhere.

Place: Delhi

Date: May 30, 2022

Abhilasha Sharma

Dr. Abhilasha Sharma

SUPERVISOR

Assistant Professor

Delhi Technological University,

Shahbad Daulatpur,

Main Bawana Road,

Delhi 110042.

ACKNOWLEDGEMENT

The success of this project requires the assistance and input of numerous people and the organization. I am grateful to everyone who helped in shaping the result of the project.

I express my sincere thanks to **Dr. Abhilasha Sharma**, my project guide, for providing me with the opportunity to undertake this project under her guidance. Her constantsupport and encouragement have made me realize that it is the process of learning which weighs more than the end result. I am highly indebted to the panel faculties during all the progress evaluations for their guidance, constant supervision and for motivating me to complete my work. She helped me throughout with new ideas,provided information necessary and pushed me to complete the work.

I also thank all my fellow students and my family for their continued support.

Rajat

Rajat Singh Arora

ABSTRACT

Breast cancer is one of the most common life-threatening diseases in women, as well as a leading cause of cancer death. Mammography is one of the most effective diagnostic methods for detecting breast cancer. However, breast cancer researchers have advocated for the use of intelligent-based techniques by medical experts and radiologists over the last decade. Deep learning, convolutional neural network (CNN), is providing effective performance in accurately classifying mammograms, which can assist imaging specialists. The CNN model should be trained with a larger number of labelled mammograms to achieve an accurate classification of mammograms. However, it is not always possible to obtain additional labels for mammograms. The primary goal of this experiment is to use CNN with dense layers to perform highly accurate mammogram classification. In our research, we created classification CNN-based models with a single dense layer. In this case, the dense layers serve as the foundation for the CNN model's accurate mammogram classification. This work aims to improve the performance of CNN with more dense layers, including multi-view preprocessed mammograms. We have used CNN with multiple activation functions (like Linear, Sigmoid, SoftMax, Relu etc.) and their combinations to improve the performance of the model and have achieved the maximum accuracy of 85% of the model which uses the activation function in combination of Relu and Sigmoid.

CONTENTS

LIST OF FIGURE(S)

LIST OF ABBREVIATION(S)

ML: Machine Learning

CNN: Convolution Neural Network

DCNN: Deep Convolution Neural Network

MIAS: Mammographic Image Analysis Society

AI: Artificial Intelligence

DNA: Deoxyribonucleic Acid

ROI: Region of Interest

IDC: Invasive Ductal Carcinoma

ILC: Invasive Lobular Carcinoma

Px: Pixels

CBIS-DDSM: Curated Breast Imaging Subset of Digital Database for Screening Mammography

CHAPTER 1

INTRODUCTION

Breast cancer rates are even worse than other known cancers. In India, they are increasing patients are diagnosed with breast cancer. Regardless of age, religion or gender, breast cancer iseverywhere. The most common cancer in Indian women. It has been noted that breast cancer affects one in 22 women in India. Of the two who develop it, one dies.

Breast cancer in India is different from in the West. Here, it affects young women and more than half of them send them to advance stages. By 2030, it is estimated that breast cancer will lead tothe deaths of many women in India than any other cancer [1]. Forethought is the key to overcoming this disease. Right now, the best Early detection of mammography is always a proven method.

It was a fall seen in both serious breast cancer and death in patients undergoing surgery mammography, for early detection and treatment. However, a common test the systems provide aradiologist who examines several tests, which raises the possibility of incorrect diagnosis.

Also, mammograms are very difficult to read and translate as they are. Mammograms are X-raytests of low power to assist in early detection cancer. Digital mammography provides electronicimages of breasts. The machine is equipped find very little of the lumps and help detect the presence of cancer cells.

With this research, we aim to develop a reliable tool for diagnosing any tumor in the breast usingmammogram images. Different Image Processing Methods are used as a framework for this thesis. This chapter begins with an introduction to breast cancer and its problems and challenges experienced during the diagnosis of breast cancer. We then explain the motive for the study, the objectives, contributions, and draft framework for this thesis is presented.

1.1 MOTIVATION

The causes of cancer in India are like other regions of the world i.e., biological, chemical, and other natural causes of uncontrolled, irregular growth of cells, known as carcinogens. Under unusual circumstances, carcinogens interact with the DNA of normal cells initiate a series of complex processes that are made up of many triggering mechanisms proliferation of uncontrolledcells or tumors [2]. Both external and internal variables may result cancer. External variables includenatural variables such as food, tobacco, radiation, or other infectious diseases while internal mutations involve hormones and immune systems. A major deviation has been observed and recorded because of lifestyle choices, styles and diet habits. It has been also observed that Breast cancer can also start in the milk ducts this type of cancer is known as IDC or even start in milk producing glands this type of cancer is called as ILC [3]. There are many other factors to consider like family history, age, changes in the gene, exposure to the radiation [4].

It is very important to diagnose a dangerous disease to create the most effective treatment. Chest Cancer is first diagnosed with individual symptoms or more serious symptoms such as a rare lump near breast tissue or abnormal discharge from the breasts. After a personal examination, further testing is taken to confirm the possibility of cancer cells. Mammography and clinical chest tests performed by a qualified physician using pain verification, although other strategies are available, but these are the most effective level of acquisition. But it still does not provide a 100% success rate. These types of diagnoses are called lies the good news is that many of these symptoms can lead to many diagnostic procedures and additional testshave pressure on patients. So, we need better strategies with better results and more effective ones detection of cancer cells.

The early-stage detection of the cancer can help in reducing the death rates because of this disease; hence the regular screen plays a vital role in it the most important and effective way to detect breast cancer at the early stage is the Mammographic image system [5,6].It can detect the different anomalies in breast even before any sign of symptom beforehand[7].

1.2 OBJECTIVE

Determining breast cancer from mammograms with the help of preexisting systems only covers demonstration of schemes for obtaining a further framework. The objectives of this thesis include:

- To remove all types of sounds from mammography images, usually external one'sfeatures.
- To preview the image for further analysis of image enhancement techniques.
- To Separate the breast implant candidate in other regions of the gray scale valuesuch as different types of sound, dense tissue and counting as well.
- To identify quantitative features such as position and location of the plant.
- To improve low image quality that is expected to contain unusual contentregions.
- To give specialists radiation, a "second eye," one that can be both consistent andconsistent.

CHAPTER 2

LITERATURE SURVEY

2.1 OVERVIEW

A significant amount of work has been done in the field of detecting the breast cancer using mammographic images. In our country, one out of every eight women will be diagnosed with breast cancer at some point in their lives. Breast cancer is the most frequent no cutaneous cancer in women; experts project that 276,480 women will be diagnosed in 2020, with 42,170 of them dying [8]. Breast cancer is the most common non-cutaneous cancer in the world. Every year, over two million women are diagnosed with cancer.

2.2 MAX AND AVERAGE POOLING

Training with large number of features becomes computationally expensive as many dimensions must be taken into consideration for feature extraction purpose. To tackle this problem, researchers pool the output of convolutional layers using maximum and average pooling. The maximum pixel value of a 2D window is used in max-pooling, while the average of all pixel values is used in average-pooling. By using the dark orange tensors, pooling finds the most important feature and summarize that group based on that feature and reduce the dimensions. Researchers frequently add these procedures in their hidden layers to limit the number of trainable parameters and reduce computation time minimal mathematical calculations [9,10]. The leaky ReLU improves on the standard ReLU by adding a modest hyperparameter, and mammography researchers have used it in their CNN hidden layers [11].

2.3 UNSUPERVISED METHOD FOR DETECTING THE CANCER

CNNs can produce significant results even when datasets are completely unlabeled and unsupervised. Unsupervised learning tasks include dimensionality reduction and segmentation [12]. By decreasing the reconstruction error, autoencoders optimize low-dimensional coding. For deep feature extraction and segmentation, stacked autoencoders are frequently used in end-to-end machine learning systems [13]. CNN layers encrypt feature map to z in deep convolutional autoencoders. the decoder expands the image bringing original features of the image back. Because well-annotated datasets are scarce in breast cancer CAD, unsupervised semantic segmentation is particularly powerful.

2.4 RESNET

He et al. [14] pioneered the use of residual networks (ResNets). They contain leftover blocks linked by shortcut connections, allowing inputs to transfer information to subsequent layers while avoiding weighted activations. The skip connection enables researchers to avoid costly gradient calculations for these layers during backpropagation, resulting in significant efficiency gains. These leftover blocks are made up of frequently stacked convolutional layers. The result of the forwarding layer, f(x), are combined with the connection that are skipped of the values, x, to produce a result, h. (x). When the weight optimization model problem is framed according to this manner, the equation results a method for getting the value of f(x) from x and h(x) => f(x) = h(x) – x

2.5 U-NET

The U- net's building blocks include both CNN layers and skip connections. The Ronneberger et al. [14] model design is unique in that it includes both a contracting and expanding path. The contracting links keeps on generating the smaller feature map till it reaches the bottleneck. Up-sampling on the other hand keeps on expanding the value of feature map till it reaches the output map.

CHAPTER 3

THEORETICAL CONCEPTS

3.1 ANATOMY OF BREAST

Breast cancer is caused by the breakdown of damaged cells. It is therefore very dangerous whichis the result of inactive cell division. To understand breast cancer, of course it is necessary to have a basic knowledge of breasts and their structure. Each breast has 15 to 20 parts called lobesaround the nipple area speakers on a wheel like a building or light system. These lobes consist ofsmall sections known as lobules. At the end of these lobes, there are small “bulbs” for milk production. These bulbs are connected by small tubes called ducts, which carry milk to the nipples. Spaces between lobes and oil-filled pipes.

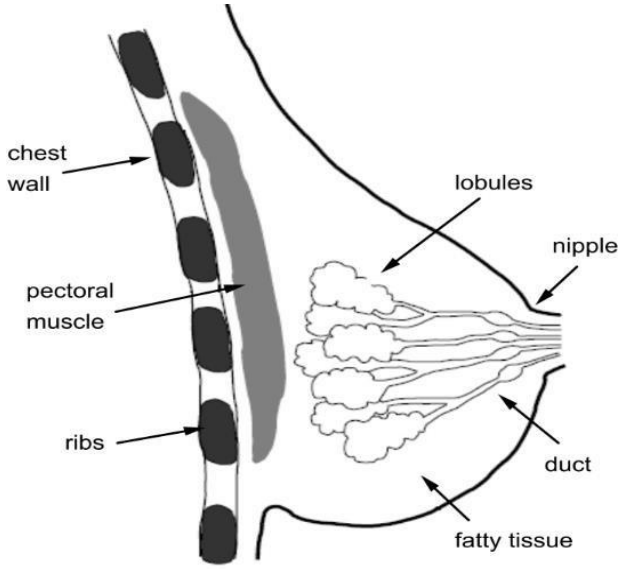


Figure 3.1 Anatomy of Breast [15]

3.1.1 Mammary Glands

Breast glands are altered sweat glands and cover 15 to 20 a series of secret lobules and pipes.

3.1.2 Lactiferous Duct

The Lactiferous duct is a structure like a single tube or tube that draws multiple alveoli

3.1.3 Connective Tissue Stroma

The supporting structure around the mammary glands is called the Connective Tissue Stroma. It mainly contains oily and fibrous components. Fibrous stroma later develops suspensory lines connecting the breasts with the pectoral fascia and separating the secretion breast lobes.

3.1.4 Pectoral Fascia

A sheet of connective tissue that connects the breasts to the pectoralis major muscles called the Pectoral Fascia. It is the attachment point of the suspensor ligaments

3.2 BREAST CONDITIONS

There are many breast conditions that can affect the breasts. A few are listed below:

3.2.1 Breast Cancer

Cancer occurs in the breast when the killer cells or cells have a specific defect they multiply andspread throughout the body. Both men and women can they are prone to breast cancer, althoughit is more common in women. Chest cancer can be detected by leaking a bloody nipple, changesin the skin or lump detection.

3.2.2 A Simple Breast Implant

This is a non-cancerous or malignant tumor that has a sac-like structure and is usually normal filled with liquid. It is most seen in women in their 30s or 40s. Mildness can it causesbreast cysts and fluid can leak out of the cyst.

3.2.3 Phyllodes Plant

This is a rare, fast-growing plant that can be dangerous or dangerous, most common in women intheir 40s. The tumor looks like a fibroadenoma on ultrasound and is usually large in size.

3.2.4 Fibrocystic Asthma

A cancer-free breast lump that changes size throughout the menstrual cycle and it causesdiscomfort.

3.3 Types of Cancer

Cancer is majorly divided into two types i.e., Benign and Malignant cancer.

3.3.1 Benign

These are the tumors that are not life threatening to the patient. They are non-cancerous. A benign tumor can easily be removed with less chances of it growing back. This is because cells in benign tumors do not spread.

3.3.2 Malignant

In contrast to benign tumors malignant are cancerous. The cells of a malignant tumor are abnormal and divide rapidly**.** Malignant tumor cells can migrate to other parts of the body andform new tumors. These cells act aggressively**,** attacking surrounding tissues**.**

3.4 CANCER STAGES

Cancers are classified according to stage**.** Tumor stage is determined by how tumor cells look under a microscope. Staging helps the body determine the type of cancerthe body has and prescribes treatment for patients**.** It determines how advanced the cancer **is** and spreads throughout the body. Thereare many stage systems. The most common system is TNM. "T" represents the size of the tumor, "N" represents the number of affected lymph nodes**,** and"M" represents the Metastasis, that is, the extent to which cancer has spread through the

circulatory or lymphatic system to other organs in the body. A lower stage, better, indicates a less advanced cancer that has probably not spread to the body and mayindicate better outcomes with treatment.

Cancer Stage

Stage 4 = Cancer in an organ other than where it started.
Stage 3 = Larger cancer also in the lymph nodes.
Stage 2 = Larger cancer that may or may not have spread to the lymph nodes.
Stage 1 = Small cancer found only in the organ where it started.
Stage 0 = Precancerous

3.5 BREAST CANCER

Breast cancer is distinguished by how the cells look under a microscope. Most of the types of breast cancer are carcinomas. Carcinoma is a type of cancer. In this type of cancer, there are cells lined up in organs and tissues such as the breast. There is no cure for cancer and no effective way to prevent it. The statistics of breast cancer are dire compared to other known cancers. In India, more and more patients are being diagnosed with breast cancer. Regardless of age, religion or gender, breast cancer is found everywhere. It is the most common cancer in Indian women. It has been observed that 1 in Indian women will develop breast cancer. One of the two people who developed it dies. Breast cancer in India is different from that in the West. Here it hurts young women, with morethan half of them submissive to themselves in later stages. By 2030, it is estimated that breast cancer will cause more deaths among women inIndia than any other type of cancer. Early detection is the most successful tactic for dealing with this epidemic. Mammography is currently the best method for early detection and remains the proven method. In patients undergoing routine mammography, prior detection and treatment resulted in a reduction in serious breast cancer and death cases. However, the routine screening program provides radiologists with many tests, which increases the chanceof a misdiagnosis. Also, mammograms are very difficult to read and interpret as-is. Mammography is a low-energy X-ray examination that helps detect breast cancer early.

3.6 BREAST CANCER LESIONS

Bilateral asymmetry, microcalcifications (MC), structural distortions, and massesare several breast cancers lesions.

Microcalcifications

Microcalcifications are small calcium deposits that are difficult to detect on mammography dueto their small size and low contrast. The size of these calcium deposits varies from 0.33 mm to 0.7 x mm. Because of their small size, isolated microcalcifications are not as easily detected as clustered microcalcifications. The accumulation of microcalcifications contained 3 or more microcalcifications containing microcalcifications in an area of 1 cm. The detection of microcalcifications is very important forearly detection of cancer. It is important to be able to differentiate between malignant and benign microcalcifications as they can cause cancer.

Mass

Mass is like normal breast parenchyma and is more difficult to detect on mammography than microcalcification [16]. As shown in the figure below, neoplasms may have various shapes, suchas round, oval, lobular or irregular. Characteristics of the texture and shape of the mass.

Benign tumors are round and have a smooth, distinct texture. Malignant tumors, on theother hand, are irregular in shape with blurred borders.

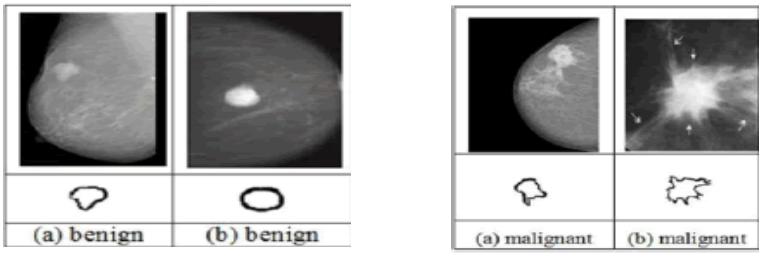


Figure 3.2 Benign and Malignant Tumors [17]

3.7 METHODS TO DETECT BREAST CANCER

There are several ways to detect the cancer some of them are listed below

3.7.1 X -ray

As there is no cure currently for cancer, early detection of breast cancer tumors is very important for the besttreatment. Early detection is the most winning tactic in the fight against this epidemic. Mammography is currently the best method for early detection and remains the proven method . Early detection and treatment resulted in a reduction in fatal breast cancer and mortality in patients who underwent routine mammography. However, the routine screening program provides radiologists with many tests, which increases the chanceof a misdiagnosis. Also, mammograms are very difficult to read and interpret as-is. Mammography is a low-energy X-ray examination of the breast that helps detect cancer early. A diagnostic mammogram is taken and read by a radiologist to screen for breast disease in women with symptoms such as nipple discharge or swelling. The mammogram contains more than images of that area. A diagnostic mammogram may show:

1. Abnormalities should not be cause for concern and patients may have regular annualmammograms. There is no need for additional testing.
2. Abnormal tissue or lesions are more likely to be benign or noncancerous. In these cases,patients have mammograms every 4-6 months.
3. Whether the tissue is malignant or suspected of being cancerous, and whether a biopsy isneeded to determine whether it is cancerous

3.7.2 Ultrasound

Ultrasonography, or commonly known as Ultrasound, is an imaging technique which emits highfrequency sound waves. These sound waves bounce off tissues and internal

organs, and their echo produces an image called a sonogram. These sound waves cannot be heard by human ears.Ultrasound is equipped with a transducer which is a handheld instrument and to obtain a sonogram, a special gel is applied on the interested area. Then with the help of a transducer, sonogram is obtained by pressing it over the gel against the skin. The emits sound waves that reflect off body tissue and produce echoes that are picked up by the machine and produce a flat, white image on a computer screen [18]. This image is then analyzed by aradiologist. This is a no-radiation and painless test.

Using Ultrasound:

1. Ultrasound in conjunction with mammography is useful when scanning high-riskwomen with dense breasts.
2. In practice, ultrasound is used to guide radiologists for the invasive treatment of a variety of cancers, including prostate and liver cancer.
3. Useful for evaluating lesions that are difficult to detect on mammography.
4. Also used as part of other medical imaging technologies.

3.7.3 Magnetic Resonance Imaging

MRI, short for Magnetic Resonance Imaging, has also proven to be very helpful in detecting breast cancer in high-risk women. MRI scans use magnets and radio waves stronger than X-rays. The contrast agent, gadolinium, is injected intravenously hours before a breast MRI . It helps to show more details. The use of MRI in conjunction with mammography, high-risk women was found to be beneficial for close follow- up and better evaluation. It can be used to better study suspicious areas on a mammogram. It is also used to determine the true size of a tumor in the breast

3.7.4 Positron Emission Mammography Detection

Early detection is the most winning tactic in the fight against this epidemic. Mammography iscurrently the best method for early detection and remains the proven method. Studies have shown that early detection and treatment reduces both fatal breast cancer and mortality rates in women who undergo routine mammography. However, the routine screening program provides radiologists with many tests, which increases the chanceof a misdiagnosis. Also, mammograms are very difficult to read and interpret as-is. Mammography is a low-energy X-ray examination of the breast that helps detectcancer early . A diagnostic mammogram is taken and read by a radiologist to check for breast disease in women with symptoms such as nipple discharge orswelling. The mammogram contains more than images of that area.

Diagnostic mammography may show

1. Abnormalities should not be cause for concern and the patient may have an annual standardmammogram. There is no need for additional testing.
2. Abnormal tissue or lesions are more likely to be benign or not cancerous.In these cases, people have mammograms every 4 to 6 months.
3. Whether the tissue is suspected of being malignant or cancerous and a biopsyis required to determine whether it is cancerous

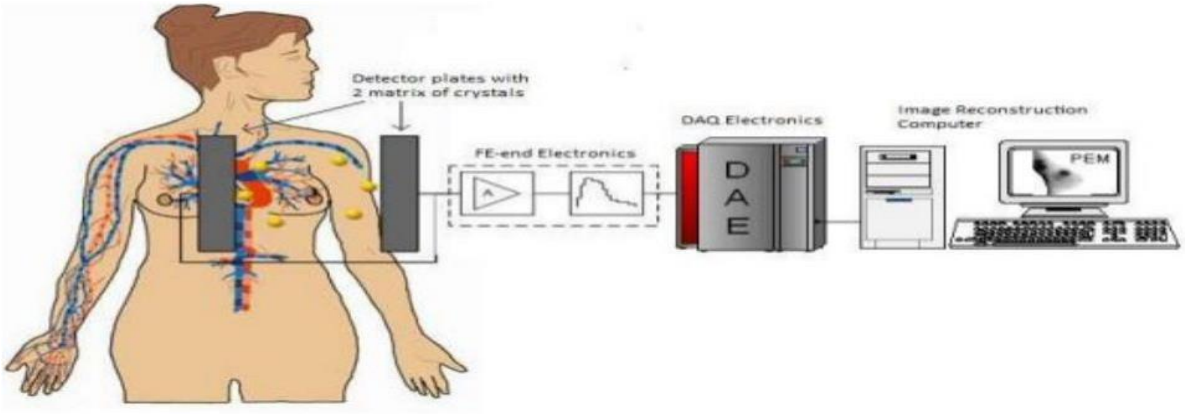


Figure 3.3 PEM Architecture [19]

3.8 ADVANTAGES AND DISADVANTAGES OF MAMMOGRAPHY

For all types of cancer, the most important step is to find the cancer before symptoms appear. Especially in breast cancer, mammography is the most used method to detect symptoms before they appear. Once symptoms appear, this indicates that the cancer has already spread, and the prognosis is poor. This study helps radiologists find missing breast lesions and prevent advanced cancer from developing. Here are some advantages and disadvantages of the mammography images

3.8.1 Advantages of Mammography

1. Mammography is the most effective strategy for detecting clinically occult illness, and it is theonly image-based tool for breast cancer screening that is approved.
2. Breast cancer can be dramatically reduced by mammography. Breast cancer mortality in a well-organized screening program over the population this detection method that reduces mortality the most.
3. Mammography reduces the risk of going through chemotherapy

3.8.2 Disadvantages of Mammography

1. Abnormalities might get masked by complex structures.
2. False positive and false negative due to misinterpretation.
3. Interana observer variability is high.
4. Overdiagnosis and overtreatment due to mistranslation.
5. Fabric overlaps, causing unnecessary trouble.
6. With dense breasts, cancerous growths may be hidden behind the normal structure of thebreast.

3.8.3 Mammographic Image Cancer Detection at Early Stage

Early detection is the most winning tactic in the fight against this epidemic.

Mammography iscurrently the best method for early detection and remains the proven method. In patients undergoing routine mammography, prior detection and treatment resulted in a reduction in serious breast cancer and death cases. However, the routine screening program provides radiologists with many tests, which increases the chance of a misdiagnosis. Also, mammograms are very difficult to read and interpret as-is. Mammography is a low-energy X-ray examination of the breast that helps detect cancer early. Digital mammography provides an electronic image of the breast. The machine detects small bumps and helps detect thepresence of cancer cells. In general, a typical digital mammography system involves two stages.

1. Mammography Image Preprocessing

a. Noise Removal

b. Image Enhancement

2. Mammogram Image Post-Processing

a. edge detection

b. division

c. bilateral asymmetry

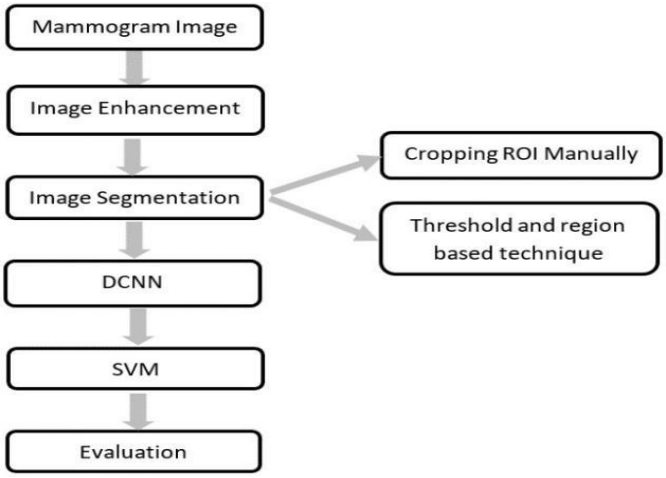


Figure 3.4 Steps of Image Processing of Mammogram Image[20]

CHAPTER 4

PROPOSED WORK

4.1 DATASETS DECISION

From a clinical aspect, the dataset is notable because it includes normal and abnormal patients,resulting in 3 groups (normal, benign & malignant). The MAIS Data set is selected for the experiment purpose of the thesis. The data contains the column for the type of cancer present where B represents the Benign and M represents Malignant cancerous tissues, column is provided in the dataset which includes the region of interest (ROI) in terms of x and y coordinates. There is total 322 images of 1024 *1024 Px. the images have been centered in the matrix already

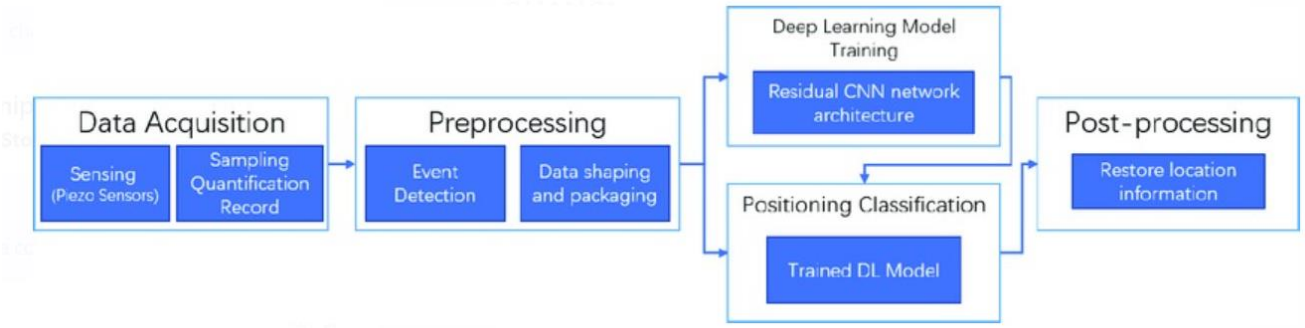
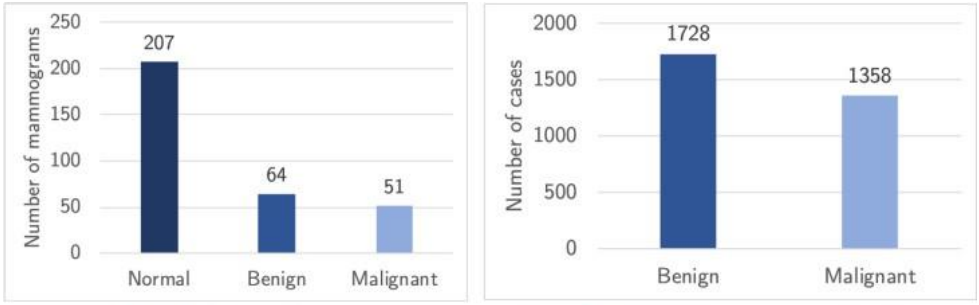


Figure 4.1 Data Preprocessing and Model Training [21]

4.1 PRE-PROCESSING OF THE DATA

4.1.1 Dataset Balance

Since it is a task of the classification, it is important for seeing the distribution of the datasetacross the classes to see if some classes are much more common than the others (dataset bias).



(a) mini-MIAS class distribution. (b) CBIS-DDSM class distribution.

Figure 4.2 MIAS and the CBIS-DDSM Datasets[19]

The MIAS dataset is substantially imbalanced due to the non-uniform distribution, which should be considered to minimize training errors. Convolution neural model is incorrect. Possibleapproaches for redressing this imbalance include:

- under sample the dataset by removing all photos.
- data augmentation can be used to oversample the dataset.
- include class weights to give underrepresented groups more weight.

Under sampling the dataset is inefficient since it reduces the number of possible outcomes. Samples from which the model could learn by dropping samples most of the population Under sampling the datasets would be a bad idea because they are already so small. strategy, as itmay cause useful features to be overlooked. As a result, oversampling It has been demonstrated that producing fresh artificial images that are like the original data is a viable option. enhance precision A less expensive solution in terms of processing power is also available. To add class weights, you don't have to touch the dataset. The result can be a loss. weighted average,with less frequent classes receiving greater weight.

4.1.2 Dataset Split

Splitting of the dataset is done to prevent the biases results of the prediction. The split is done to get the training and the testing sample, training sample is the larger chunk of the split which contain the unbiased amount of the sample in this thesis we will be taking the 70% of theMIAS Dataset as the training split and the rest is for testing the model ,i.e., to check if the modelthat we have made is the predicting accordingly on the new ,untested sample

4.1.3 Loading of Data

With only 322 pictures, the MIAS dataset is quite small (339 Mb before preprocessing, 202 Mbafter preprocessing). As a result, it can be loaded directly into RAM without the use of any dataloading optimization algorithms. The CBIS-DDSM is quite large data set of 163GB with 10,239 images the for the loading of the data in the model it must be done in batches. For this thesis we are using MIAS dataset.

4.1.4 Normalization of Data

Images are scaled to a suitable size after import to avoid having uneven input sizes of MIAS [22]. When looking at the photo pixel intensities in the files, it's evident that they result in numbers with values ranging from 0 - 255. However, the weight of the neural network is taken to be small, large input values may cause the training process to be disrupted and slowed, resulting inlower accuracy.

4.1.5 Encoding of Labels

The labels for each mammogram must be encoded into a numerical format because they are in categorical string format. Sparse nature of the labels (As there are three groups (normal, benign &malignant) makes use the one hot encoding , It is used for the MIAS dataset, where a single digit have the value of 1 while others have the value of 0 to identify the distinct labels. Hot singlefrom the labels as shown in Table.

Categorical format	One-hot encoding
Normal	1 0 0
Benign	0 1 0
Malignant	0 0 1

Categorical format	Binary encoding
Benign	0
Malignant	1

4.1.6 Training The Model

The photos after preprocessing are now ready to be fed in the CNN model for the training where the images with the values are compared and the model keeps on iterating until it reaches to thepoint where the result starts to get almost equal to the previous one this this basically training ofthe model.

4.1.7 CNN MODEL

As Illustrated in the figure below the images form the MIAS dataset are fed to the first layer ofthe model now since the image be large in the size it is important for us the pooling , pooling is basically the process of grouping the pixels of the picture and make the image size reduce by taking max value (pixel), average ,mean etc. from the group of the that area which generally results in the reducing the size/dimension of the image ,now this reduced image size is fed to next layer and same pooling process is repeated and finally the image is then fed to the predicting layer where we can use Relu ,SoftMax etc. to increase the result of the prediction

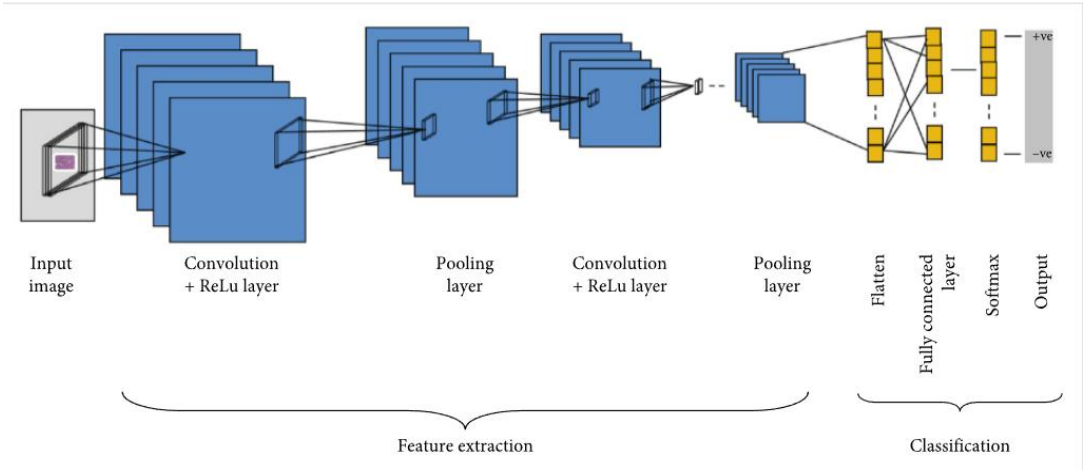


Figure 4.3 Layout of CNN Model [23]

4.1.8 Data Fitting

The figure show the ideal figure for the function used SoftMax, sigmoid and as a matter of fact relu which is sown in the later part of the thesis , the data fitting means that that according to the layer that we used as the last layer in out model the prediction on MIAS image should follow thelines of the ideal figure of the model till that not happens or until we react to the near the model we will keep on training the model , but we have to be careful that our model do not completely replicate the ideal model as it will result in the overfitting as overfitting results in the reduction of the efficiency of the model as its starts taking the noises into the consideration too

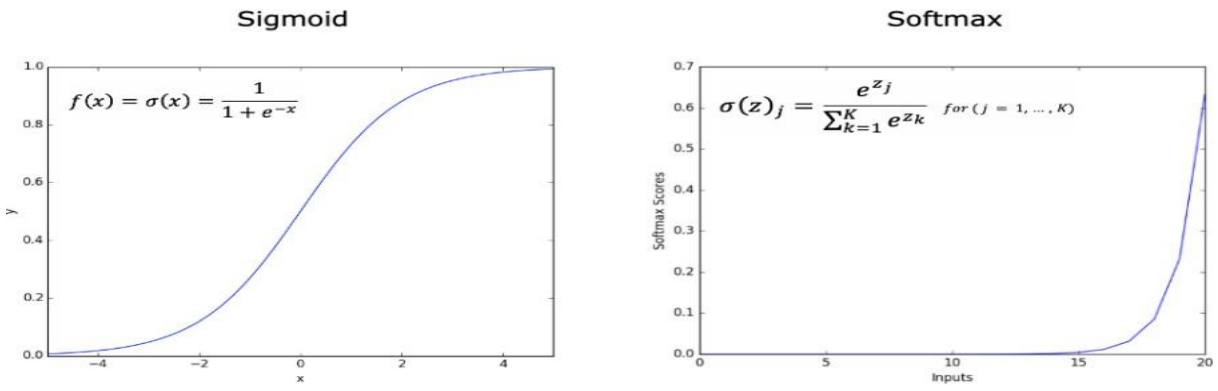


Figure 4.4 Activation Function Ideal Graph[24]

Loss Function it may be used for both binary and multiclass problems, one of the most important used loss functions is cross entropy. Because probability is calculated using the Crossentropy is the perfect loss function because it largely relies on sigmoid and SoftMax activation functions.

Optimizer, because of the model's depth, it's critical to keep the number of hyperparameters under control to a minimum. Traditional optimizers for e.g., SGD, that are slow and require morefine tuning, usually generalize better than adaptive learning rate algorithms. Adaptive moment estimation (Adam) is the most generic adaptive optimizer, which is also being used in on modelfor optimization

4.2 RESULT VISUALIZATION

The metrics described below are defined using the following terminology:

- True Positives (positive cases accurately predicted as positive).
- True Negatives (negative cases correctly forecasted as negative).
- False Positives (negative cases wrongly projected as positive).
- False Negatives (negative cases incorrectly predicted as negative); (positive caseincorrectly predicted as negative).

4.3 ACCURACY

When analyzing the classifiers' performance, the imbalanced class distributions must be considered. Scores. Indeed, adopting a metric for evaluation Overall accuracy, for example, would be inaccurate because it would not reflect how well the system worked. The data suited the classifier. Detecting FPs and FNs is also important in breast cancer

detection. It is critical to prevent misinterpreting malignant tumors as benign and vice versa. It may cause injury to the patient and eventually death Creating a simple classifier that always classifies photos as "normal"is one example. despite never looking for anything, obtains 64.28 percent accuracy on the MIAS dataset anomalies. Following is the formula used for the accuracy calculation

$$Accuracy = \frac{TP + TN}{P + N} \quad \text{.....(4.1)}$$

On the MIAS dataset, for example, generating a simple classifier that always labels photos as"normal" achieves 64.28 percent accuracy despite never looking for anomalies. As a result, avariety of extra measures should be employed to assess how well the model learns from mammography data and generalizes to cases that aren't visible.

4.3.1 Accuracy and Recall

Accuracy is the measure for avoiding the negative cases as the positive by counting the correct positive prediction

Precision corresponds to the number of correct positive predictions

$$Precision = \frac{TP}{TP + FP} \quad \text{.....(4.2)}$$

The Recall is the number of positive cases that were accurately predicted, indicating how wellthe model can locate all positive cases.

$$Recall = \frac{TP}{TP + FN} \text{..... (4.3)}$$

4.3.2 F1 Score

The combination of precision and recall is the f1 score to have the F1 score to be high theprecision and recall should be high as well

$$F_1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{TP}{TP + \frac{FN+FP}{2}} \text{..... (4.4)}$$

4.3.3 Confusion Matrix

It is a visual metric that displays total predictions produced for every class, with every row represents labels and each column representing a forecast. This is important for determining where the true class is encountered most frequently and where the predicted class is misclassified. To emphasize the misclassification and compare predictions with different classifiers, the confusion matrix is adjusted to display percentages rather than count

CHAPTER 5

EXPERIMENTAL SETUP

5.1 OVERVIEW

This section will describe the planned approach for doing histopathological image detection forcancer, which is implemented on a deep convolutional neural network model. In which we designed a deep convolutional neural network as from scratch.

5.2 DATASET USED

Images and labels / annotations for mammography scans make up the data. The 'Preview' kernel demonstrates how to appropriately parse the Info.txt and PGM files. There are a few other thingsto consider:

The films are organized in pairs, with each pair representing a single patient's left (even filenamenumbers) and right (odd filename numbers) mammograms. All the photos are 1024 pixels by 1024 pixels in size. The images in the matrix have been centered. When calcifications are present, the radii and center positions apply to clusters rather thanindividual calcifications. The bottom-left corner is the origin of the coordinate system [24,25]. Calcifications are sometimes dispersed throughout the image rather than concentrated at a singlelocation. In these circumstances, the radii and center placements are important are unsuitable andhave been removed.

5.3 DEEP LEARNING

Deep learning is subset of ML that contains artificial neural network, that are algorithms inspiredby biological and function of the human brain.

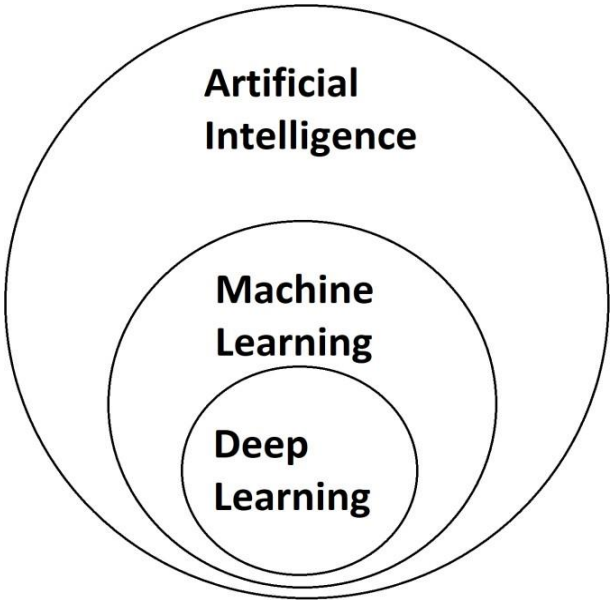


Figure 5.1 Machine Learning, Deep Learning and Artificial Intelligence[26]

Deep neural networks analyze data with a predetermined structured way to reach similar conclusions as humans [27]. Deep learning achieves this by employing a algorithm with multiple layers known as neural networks

We may use neural networks to accomplish a range of tasks, such as, classification, regression, and clustering. We can use neural networks to categorize or classify unlabeled data based on similarities between the samples. We may train the network on a labeled dataset to categorize thedata in this dataset into distinct categories in the classification stage.

5.3.1 CNN

Convolution neural network is a one type deep learning model. This neural network is feedforward neural network where signal flow from output of one neuron to input of next layerneuron means there is no feedback.

A CNN is a technique that takes an image is processed and assigns weights to the features in the image so that they can be discriminated [28,329]. Compared to other classification techniques, CNNs don't require that much preprocessing. The filters can be learned by the convolutions themselves.The architecture of CNN is the structure of neurons as a source of inspiration in the human brain.

CNN uses multiplication of an image matrix with a for extracting features and pre-determined characteristics from it. We use a channel to filter the image and get only the predominant important features. The images are matrices of pixel values, and the filters are commonly 3x3 or5x5. The filter is moved across the image with a specified stride, and the values are multiplied and added to provide a matrix output that is easier to understand.

$$F(x) = f_n\,(f_{n-1}(\dots f_i(x)) \quad \dots\dots\dots(5.1)$$

Where ‘n’ is number of hidden layer and ‘fi’ is function used in corresponding layer.In a CNN model there are basically 5 layers.

- Convolution layer
- Activation function layer
- Pooling layer
- Fully connected dense layer
- Predication layer

5.3.2 Convolution Layer

In convolution layer it uses filters size n*n and apply convolution operation all over the image using these filters and extract the tiny features of the image which we called featuremap [30]. For example: In nodule detection in lung cancer, we use to detect edges, shapes, abnormal cells etc.

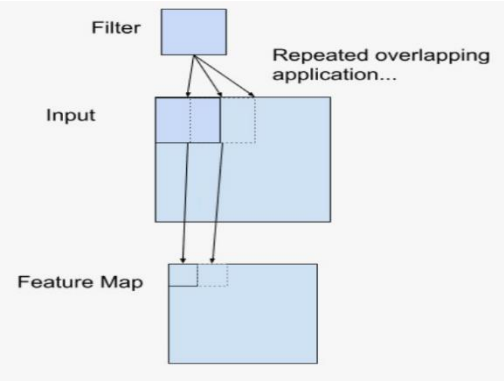


Figure 5.2 Applying Filter in Input Image[31]

5.3.3 Activation Function Layer

In this layer we use any nonlinear activation function to bring non-linearity in our model which speeds up training and faster to compute. It is used to learn and understand complex pattern in our data, and this is done to avoid the values from summing up to zero. Mostly we use RELU (rectified linear units) [32] as an activation function which can be expressed as $\boldsymbol{f(x) = \max(0, x)}$ There can be other activation function too like sigmoid, tan h and exponential linear unitsetc.

5.3.3.1 Linear Activation Function

The equation for a linear function is y= mx, which is the same as the equation for astraight line. Equation: f(x)=x Range: -∞ to ∞ ……...(6)

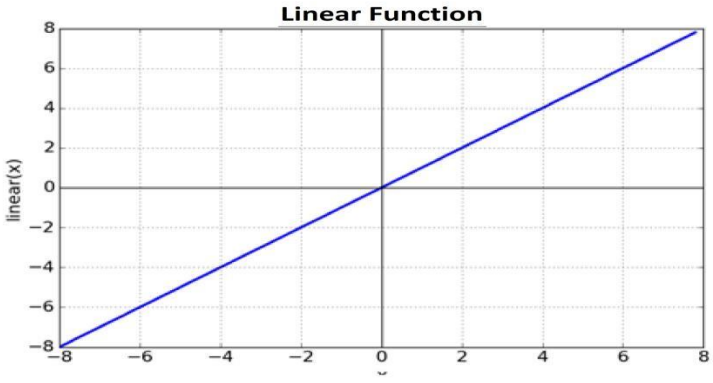


Figure 5.3 Linear Activation function[33]

5.3.3.2 SIGMOID FUNCTION

It's a function this is graphed as an S shaped it is not as rigid as the linear function whichdivides the data point form on single line it although has the value between 0 and 1 but itprovides a flexibility and extends accordingly to the data point to an extent [34]

Equation: $f(x)=\frac{1}{(1+e^{-x})}$

Range: 0 to 1 ……. (5.2)

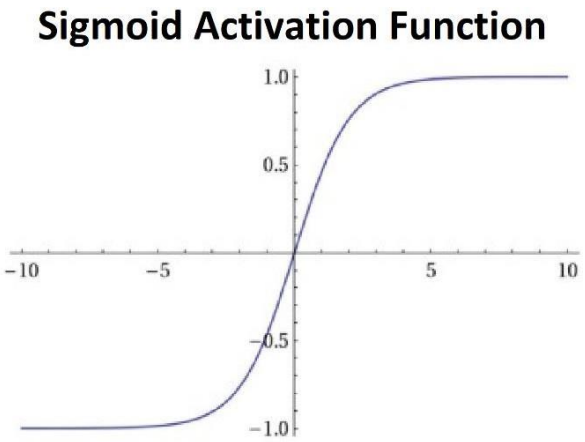


Figure 5.4 Sigmoid Activation Function[35]

5.3.3.3 ReLU Activation Function

Rectified linear unit is a term that refers to a linear unit that has been recalibrated. It's most often used activation technique in hidden layers of a deep neural networks. The ReLU function is non-linear, that implies we may easily back propagate errors and have several layers of neurons triggered by this[34]. Because it includes fewer mathematical calculations, ReLU becomes less computationally costly than tanh and sigmoid.

Only these few neurons are active at a time, making the deep neural network sparse and fast forprocessing.

Equation: f(x)=max (0, x) Range: 0 to ∞ ……. (5.3)

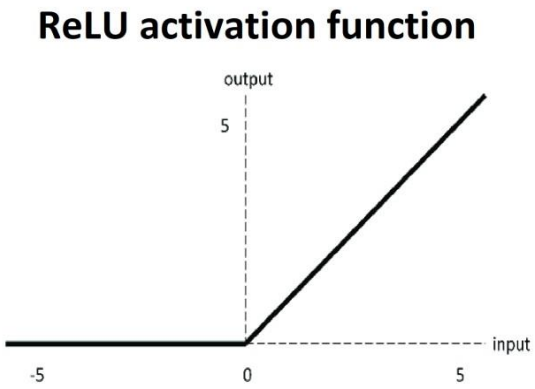


Figure 5.5 ReLU activation function[36]

5.3.4 Pooling Layer

This layer used to reduce the size or reduce the dimension of image (feature representation). By experiments it was found that Max pooling is used mostly. In Max pooling we use a window of size m*m and take maximum pixel value among all the pixel values in window of feature map and slides to stride of 'k' and by doing this it covers the whole feature map [37,38,39]. There are different types of pooling like max pooling, min pooling, average pooling etc.

5.3.4.1 Max Pooling

A pooling operation that determines the maximum value of sections of a feature map anduses it to construct a down - sampled (pooled) feature map is known as Max Pooling. After a convolutional layer, it's typically used. The feature space dimensions are decreased by using pooling layers. As a result, the set of parameters for learn and the complexity of processing in the network are both reduced.

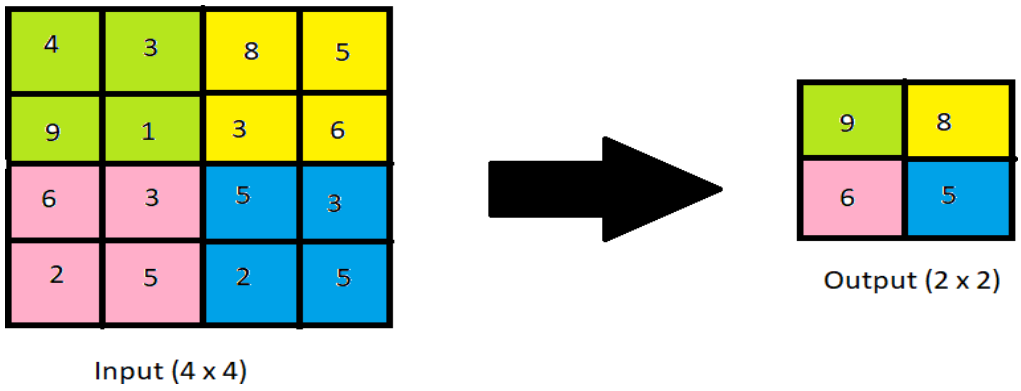


Figure 5.6 Example of Max Pooling of The Input Image[40]

5.3.4.2 Average Pooling

The average of the items present in the region of the feature space represented by the filter gets calculated applying average pooling. As a result, while max pooling returns themost significant aspect in a feature map patch, average pooling provides the average of all features present in that patch.

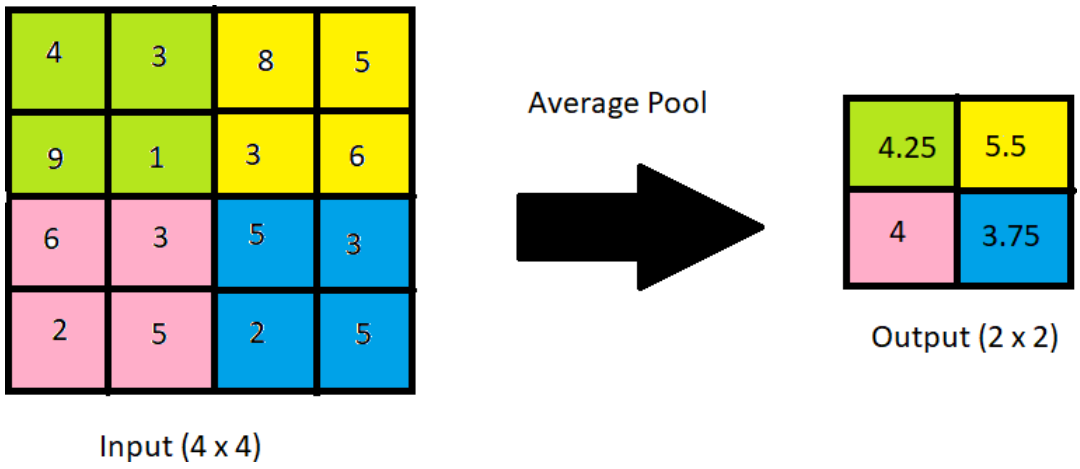


Figure 5.7 Example of Average Pooling of The Input Image[41]

5.3.5 Fully Connected Dense Layer

This is the final layer where the actual classification happens. Here we take our filtered and shrieked images and put them into a single list which is called vector In Fully connected dense layer each neuron is connected to every other neuron of next layer.

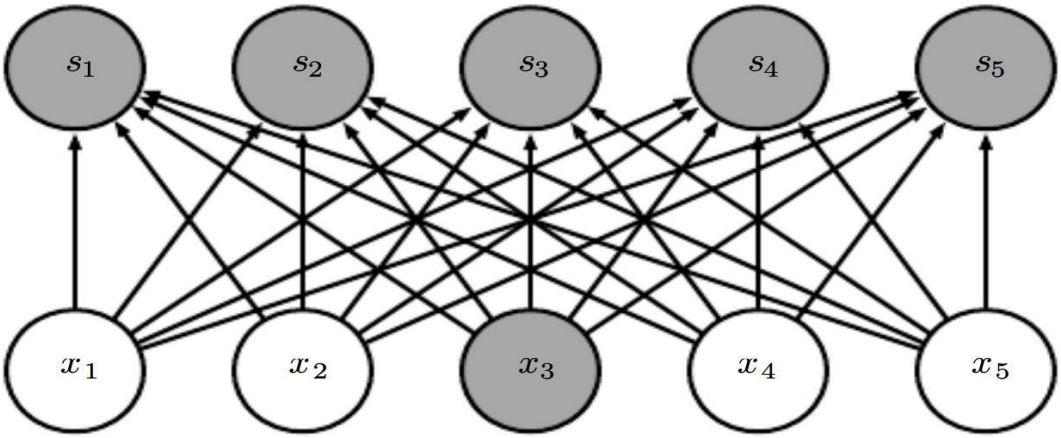


Figure 5.8 Fully Connected Dense Layer[42]

5.3.6 Dropout Layer

When a perceptron's/neuron in a neural network is turned off with a probability P during training, the term "dropout" is used. Assume a probability of P = 0.25, which means that throughout training, 25% of the neurons will be dropped. As a result, a quarter of the neurons within neural network would not be examined, and the neural network will become easier. To avoid overfitting, we add a dropout layer in the CNN. Simple terms, throughout the training process, a certain number of neurons are discarded from the deep neural network inthe dropout layer.

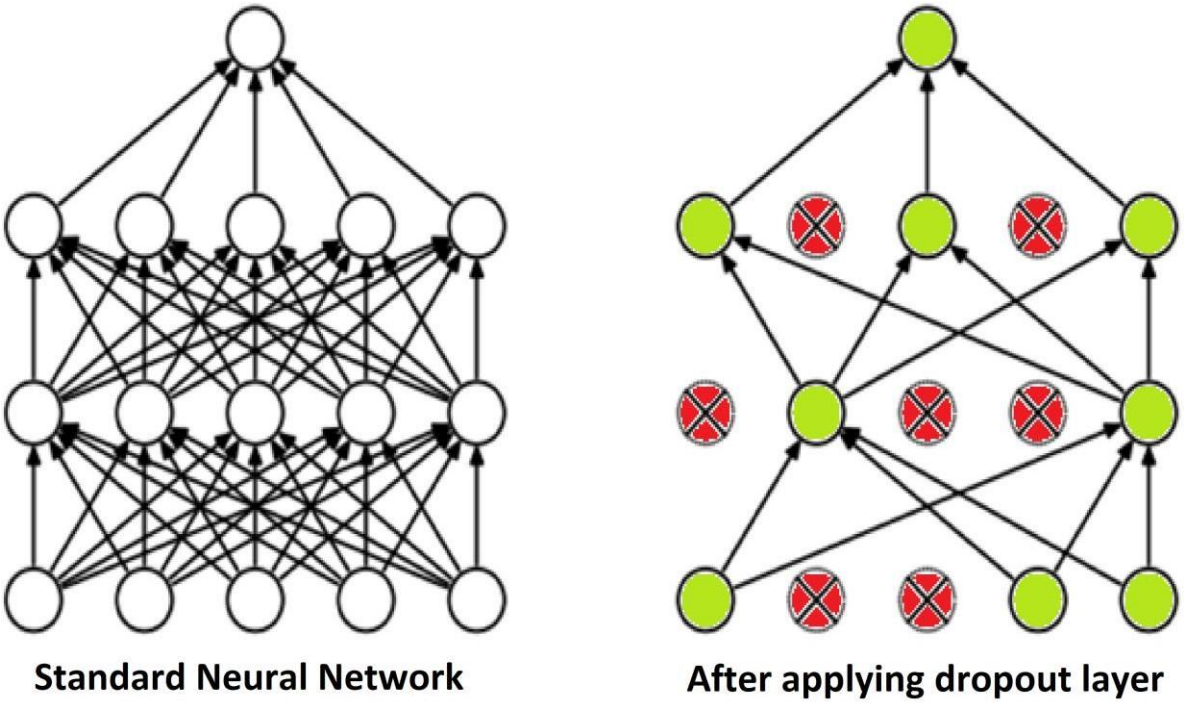


Figure 5.9 Standard and dropout CNN layer[43]

5.4 DATA PREPROCESSING

The data operations directory contains functions related to data preprocessing, such as extractingpicture paths and labels, processing images, encoding labels, loading data into memory, and performing transformations for data augmentation. The dataset processing scripts directory contains one-time scripts for parsing the picture paths and labels for each dataset. With only 322 pictures, the MIAS dataset is quite small (339 Mb before preprocessing, 202 Mb after preprocessing) [44]. As a result, it can be loaded directly into RAM without the use of any dataloading optimization algorithms. The CBISDDSM dataset, on the other hand, is substantially larger, with 10,239 photos covering 160 GB of storage space [45]. Hence CBISDDM dataset must be passed in the batches to the CNN model for processing. Here in this is thesis we have used the MIAS dataset for the experiment The image data is loaded from the csv file the classification of the images is done accordingly and the data having the missing values

are completely removed from the data by dropping that row completely

5.4.1 Data Processing

The MIAS images are first imported into python imaging library format using keras load image function. then they are resized to the target of 224*224 pixels in grey scale to reduce the dimensionality of the images before converting them to 1D array format. Region of interest is identified on the grey scaled image and then the images are then ready to be fed to the first layerof the CNN model

5.4.2 Data Set Segmentation

Function for Segmenting Training Tests Using a shuffle of 60/20/20 percent tier splits,ScikitLearn partitions the dataset into training, validation, and test sets.

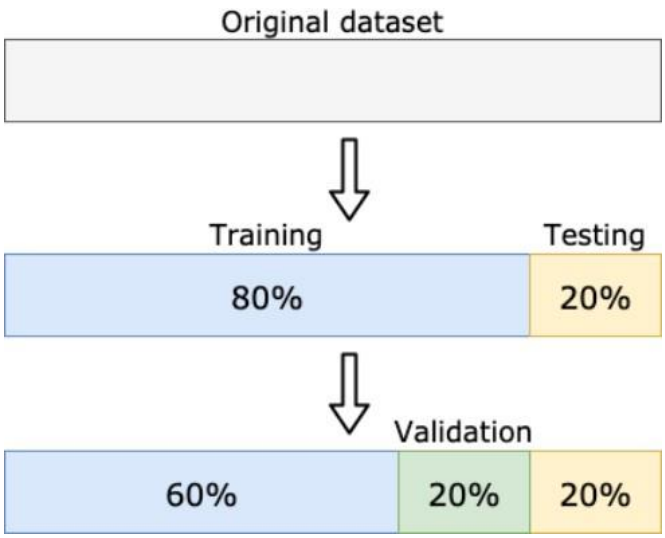


Figure 5.10 Using a 60/20/20 split, the original dataset was partitioned into training,validation, and testing sets.[46]

5.5 TRAINING THE MODEL

After preprocessing the data is now ready to be fed into the model for the training purpose.

5.5.1 Training Steps

After all the important steps which included importing the data and preprocessing comes the nextimportant step that is training the model. The preprocessed data is fed to the CNN outer layer the image are converted to the one-dimensional array and each feature is fed to different node of first layer the layer then do max pooling of the image set provided and reduce the dimensionality of the image and this process is continued further the , the result is then evaluatedand is checked with the output present in the training data. Then the back propagation starts to reevaluate the decision on the weights made if the output obtained by the model is wrong this process keeps on happening till the result starts getting repeated. Once this process over then we can say that model is trained and ready for the testing phase of the model.

CHAPTER 6

EXPERIMENTAL RESULT

6.1 MODEL EVALUATION

We use the following measurements based on confusion matrices conclusions for predictionevaluation.

6.1.1 Accuracy

The data that is correctly categorized divided by the whole dataset evaluated is how accuracy is calculated. It can also be calculated as a 1-error. The following equation can be used to calculatethe accuracy:

$$\text{Accuracy} = \frac{TruePositive(TP) + TrueNegative(TN)}{TruePositive(TP) + TrueNegative(TN) + FalsePositive(FP) + FalseNegative(FN)} \quad \text{……. (6.1)}$$

6.1.2 Precision

Simply expressed, precision refers to the percentage of actual positive results out of the totalpositive anticipated by the model. The sum of true positives (TP) throughout all classes is divided by the sum of true positives (TP) and false positives (FP) across all classes to get precision.

The following equation can be used to calculate precision values:

$$\text{Precision} = \frac{Sum\ x\ in\ X\ TruePositives\_x}{Sum\ x\ in\ X\ (TruePositives\_x + FalsePositives\_x)} \quad \text{……. (6.2)}$$

6.1.3 Recall

The rate at which the system can relearn information is referred to as recall. As a result,Recall calculates the number of true positive (TP) characteristics that our model detected andassigned a positive label to. The true positive (TP) sum of across all classes is divided by thetrue positive (TP) and false negatives (FN) summation across all classes to calculate recall.

The following equation can be used to calculate recall values:

$$\text{Recall} = \frac{Sum\ x\ in\ X\ TruePositives\_x}{Sum\ x\ in\ X\ (TruePositives\_x + FalseNegatives\_x)} \quad \ldots\ldots. (6.3)$$

6.2 MODEL PREDICTION

After cleaning, preprocessing, and analyzing the data, the first action we do is integrate it into a model that generates probabilistic results. The confusion matrix is a performance metric for machine learning classification. Precision, Recall, Accuracy, Specificity all are calculated using the confusion matrix. It is used to evaluate machine learning classification performance

It can be observed that the proposed model which uses Relu, and Sigmoid activation functionperformed reasonably well with 87% precision and 92% recall in the classification task. Nowmoving on to the accuracy and F1-scores of the proposed model, the proposed model achieved the highest score with 85% accuracy and 89% F1 score

```
              precision    recall  f1-score   support

           0       0.87      0.92      0.89     39736
           1       0.76      0.65      0.70     15769

    accuracy                           0.84     55505
   macro avg       0.81      0.78      0.80     55505
weighted avg       0.84      0.84      0.84     55505

[[36502  3234]
 [ 5563 10206]]
Accuracy: 0.8415097738942438
Specificity: 0.6472192275984526
Sensitivity: 0.9186128447755184
```

Figure 6.1 Result of the Proposed Model

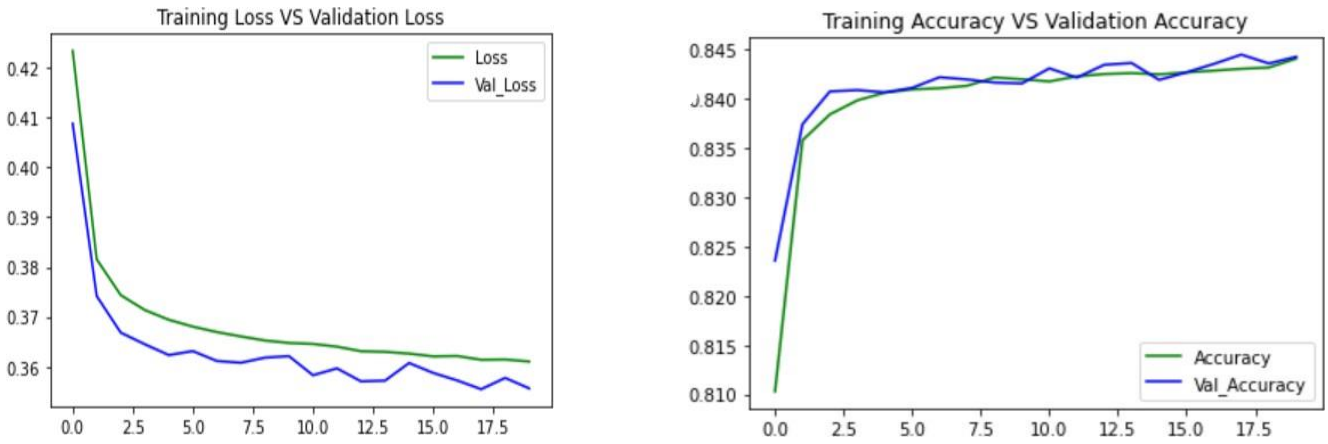


Figure 6.2 Accuracy Graph Vs Validation Graph for The Proposed Model

CHAPTER 7

CONCLUSION AND FUTURE WORK

Breast cancer is the most common disease in the women and also one of the major reasons for the cancer death one of the most efficient ways for the early detection of the cancer is through mammographic images. Through is experiment we tried to make a CNN model which can be used for the early detection of the cancer. The MAIS dataset that has been obtained from the Kaggle, is used for the experiment. The images are preprocessed by converting to the grayscale image for the reduction of the dimensionality of the images. The images are then fed to the CNN model where multiple activation functions are used to check the performance of the model, we have also used the various combinations of the activation function to increase the accuracy but keeping in mind to not overdo the various activation function as that might result into overfitting of the model and my give bad results in the testing phase. The model having Relu and Sigmoid as the activation function has resulted int the maximum accuracy of the 85% for the model.

As for the future work, work can be done to further improve the accuracy of the model, currently this model uses MIAS dataset which have relatively less sample size as compared to CBIS-DDSM, using this dataset in batches will surely improve the accuracy of the model, Overall, we can say that detection of the Breast cancer is a good way for predicting the disease at an early stage which can save many lives.

REFERENCES

[1] Sung H, Ferlay J, Siegel RL et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249. https://doi.org/10.3322/CAAC.21660

[2] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of alzheimer's disease with deep learning," IEEE, 2014.

[3] Chen Z, Yang J, Li S, Lv M, Shen Y, Wang B, Li P, Yi M, Zhao X, Zhang L, Wang L, Yang J (2017) Invasive lobular carcinoma of the breast: A special histological type compared with invasive ductal carcinoma. PLoS One 12(9):e0182397.

[4] Ataollahi M, Sharifi J, Paknahad M, Paknahad A (2015) Breast cancer and associated factors: a review. J Med Life 8:6–11

[5] Misra S, Solomon NL, Moffat FL, Koniaris LG (2010) Screening criteria for breast cancer. Adv Surg 44:87–100. https://doi.org/10.1016/J.YASU.2010.05.008

[6] Welch HG, Prorok PC, O'Malley AJ, Kramer BS (2016) Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. N Engl J Med 375(15):1438–1447

[7] Baker JA, Rosen EL, Lo JY, Gimenez EI, Walsh R, Soo MS (2003) Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. AJR Am J Roentgenol 181(4):1083–1088

[8] National Cancer Institute, Cancer of the breast (female) - cancer stat facts, SEER (2020), 1–1.

[9] S. Guan, M. Loew, Breast cancer detection using synthetic mammograms from generative adversarialnetworks in convolutional neural networks, J. Med. Imag. 6 (3) (2019) 1–10, https://doi.org/10.1117/1.JMI.6.3.031411

[10] W. Lotter, G. Sorensen, D. Cox, A multi-scale cnn and curriculum learning strategy for mammogramclassification, in: In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2017, pp. 169–177.

[11] N. Dhungel, G. Carneiro, A.P. Bradley, A deep learning approach for the analysis of masses in mammograms with minimal user intervention, Med. Image Anal. 37 (2017) 114–128,

https://doi.org/10.1016/j.media.2017.01.009

[12] M.A. Al-Masni, et al., Simultaneous detection and classification of breast masses in digital mammograms via a deep learning yolo-based cad system, Comput. Methods Progr. Biomed. 157 (2018)85–94, https://doi.org/10.1016/j.

[13] A. Geron, Hands-on machine learning with scikit-learn, keras, and tensorflow: concepts, tools, and techniques to build intelligent systems, OReilly Media, Inc, 2019

[14] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in: In Proceedings of IcmlWorkshop on Unsupervised and Transfer Learning vol. 27, 2012, pp. 37–49, https://doi.org/10.5555/3045796.3045801.

[15] https://images.app.goo.gl/oQzTJ83jegdZDbZ99

[16] K.-L. Hua, C.-H. Hsu, S.C. Hidayati, W.-H. Cheng, Y.-J. Chen, Computer-aidedclassification of lung nodules on computed tomography images via deep learning technique, OncoTargets and therapy,8 (2015).

[17] Hamouda, S. K. M., El-Ezz, R. H. A., & Wahed, M. E. (-1, November 30). *Enhancement accuracy of breast tumor diagnosis in digital mammograms*. Journal of Biomedical Sciences. Retrieved May 25, 2022, from http://surl.li/cbbyl

[18] http://dspace.dtu.ac.in:8080/jspui/handle/repository/18100

[19] D. Gutman, N.C. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, A. Halpern, Skinlesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC), arXiv preprint arXiv:1605.01397, (2016).

[20] A.A.A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S.J. van Riel, M.M.W. Wille, M. Naqibullah, C.I.Sánchez, B. van Ginneken, Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks, IEEE transactions on medical imaging, 35 (2016) 1160-1169. https://doi.org/10.1109/tmi.2016.2536809\

[21] Sigmoid activation function graph - Retrieved from http://surl.li/cbcnq

[22] K.-L. Hua, C.-H. Hsu, S.C. Hidayati, W.-H. Cheng, Y.-J. Chen, Computer-aidedclassification of lung nodules on computed tomography images via deep learning technique, OncoTargets and therapy,8 (2015).

[23]CNN architecture model retrieved from http://surl.li/cbqgi

[24] N. Dhungel, G. Carneiro, A.P. Bradley, A deep learning approach for the analysis of masses in mammograms with minimal user intervention, Med. Image Anal. 37 (2017) 114–128, https://doi.org/10.1016/j.media.2017.01.009

[20] S. Hussein, R. Gillies, K. Cao, Q. Song, U. Bagci, TumorNet: Lung Nodule Characterization Using Multi-View Convolutional Neural Network with Gaussian Process, arXivpreprint arXiv:1703.00645, (2017).

[21] Q. Dou, H. Chen, L. Yu, J. Qin, P.-A. Heng, Multilevel Contextual 3-D CNNs for FalsePositive Reduction in Pulmonary Nodule Detection, IEEE Transactions on Biomedical Engineering, 64 (2017) 1558-1567.

[22] J. Suckling, J. Parker, and D. Dance, ‘‘The mammographic image analysis society digital mammogram database excerpta medica,’’ in Proc. Int. Congr. Ser., vol. 1069, 1994, pp. 375–378

[23] “Mammographic image analysis society digital mammogram database.” [Online]. Available: http://peipa.essex.ac.uk/info/mias.html

[24] A. Jalalian, S. Mashohor, R. Mahmud, B. Karasfi, M. I. B. Saripan, and A. R. B. Ramli, ‘‘Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection,’’ EXCLI J., vol. 16, p. 113, Feb. 2017

[25] S. Bridle, Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition, in Neurocomputing, Springer, 1990, pp. 227–236

[26] Artificial Intelligence Subset Model ,retrieved from https://levity.ai/blog/difference-machine-learning-deep-learning

[27] J. Deng, W. Dong, R. Socher, L-J. Li, K. Li, and L. Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009). IEEE, Miami, FL, USA, 248–255.

[28] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE transactions on pattern analysis and machine intelligence 35 (8) (2013) 1798–1828.

[29] C. D. Malon, E. Cosatto, Classification of mitotic figures with convolutional neural networks and seeded blob features, Journal of pathology informatics 4

[30] A. Cruz-Roa, A. Basavanhally, F. Gonz´alez, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, A. Madabhushi, Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks, in: Medical Imaging 2014: Digital Pathology, vol. 9041, International Society for Optics and Photonics, 904103, 2014

[31] V.K. Singh, et al., Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network, in: Expert Systems With Applications vol. 139, 2020, p. 112855, https://doi.org/ 10.1016/j.eswa.2019.112855

[32] J. Deng, W. Dong, R. Socher, L-J. Li, K. Li, and L. Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009). IEEE, Miami, FL, USA, 248–255.

[33] Linear activation function graph retrieved from http://surl.li/cbcnf

[32] V. Nair, G.E. Hinton, Rectified Linear Units Improve Restricted Boltzmann Machines, 2010

[33] J.S. Bridle, Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition, in Neurocomputing, Springer, 1990, pp. 227–236

[34] M.A. Al-Masni, et al., Simultaneous detection and classification of breast masses in digital mammograms via a deep learning yolo-based cad system, Comput. Methods Progr. Biomed. 157 (2018) 85–94, https://doi.org/10.1016/j. cmpb.2018.01.017.

[35] Sigmoid activation function graph retrieved from http://surl.li/cbcnq

[36] G. Liang, et al., Joint 2D-3D breast cancer classification, in: In 2019 Ieee International Conference on Bioinformatics and Biomedicine (Bibm), 2019, pp. 692–696, https://doi.org/10.1109/BIBM47256.2019.8983048.

[37] V.K. Singh, et al., Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network, in: Expert Systems With

Applications vol. 139, 2020, p. 112855, https://doi.org/ 10.1016/j.eswa.2019.112855.

[38] J. Suckling, et al., Mammographic Image Analysis Society (Mias) Database v1.21, 2015.

[39] M. Heath, et al., Current status of the digital database for screening mammography, Comput. Imag. Vision Digital Mammogr. (1998) 457–460

[40] Max Pooling retrieved from http://surl.li/cbcnx

[41] Average Pooling retrieved from http://surl.li/cbcoe

[42] Neural Network Model retrieved from http://surl.li/cbcoj

[42]Standard CNN model and dropout Layer model retrieved from http://surl.li/cbcon

[41] P. Elangovan, A. Hadjipanteli, A. Mackenzie, D.R. Dance, K.C. Young, K. Wells, OPTIMAM image simulation toolbox-recent developments and ongoing studies, in: International workshop on breast imaging, 2016, pp. 668–675, https://doi. org/10.1007/978-3-319-41546-8_83.

[42] M.D. Halling-Brown, et al., OPTIMAM mammography image database: a large scale resource of mammography images and clinical data, arXiv preprint arXiv, 2004, 04742, 2020

[43] Enakshi Jana, Dr.Ravi Subban, S. Saraswathi,Research on Skin Cancer Cell Detection using Image Processing, 2017 IEEE International Conference on Computational Intelligence and Computing Research. https://doi.org/10.1109/ICCIC.2017.8524554

[44] K.-L. Hua, C.-H. Hsu, S.C. Hidayati, W.-H. Cheng, Y.-J. Chen, Computer-aided classification of lung
nodules on computed tomography images via deep learning technique, OncoTargets and therapy, 8 (2015).

[45] Suckling, J. Parker, and D. Dance, ‘‘The mammographic image analysis society digital mammogram database excerpta medica,’’ in Proc. Int. Congr. Ser., vol. 1069, 1994, pp. 375–378

[46] S. Destounis, A. Santacroce, Age to begin and intervals for breast cancer screening: balancing benefits and harms, Am. J. Roentgenol. 210 (2) (2018) 279–284, https://doi.org/10.2214/AJR.17.18730.

[47] E. Sickles, W. Weber, H. Galvin, S. Ominsky, R. Sollitto, Baseline screening mammography: one vs two views per breast, Am. J. Roentgenol. 147 (6) (1986) 1149–1153, https://doi.org/10.2214/ajr.147.6.1149

[48] D. Houfani, S. Slatnia, O. Kazar, N. Zerhouni, A. Merizig, H. Saouli, Machine learning techniques for breast cancer diagnosis: literature review, in: Advanced

[49] A. White, et al., Cancer screening test use - United States, 2015, MMWR Morb. Mortal. Wkly. Rep 66 (8) (Mar. 2017) 201–206, https://doi.org/10.15585/ mmwr.mm6608a1