Breast Cancer Detection and Classification

Submitted in partial fulfillment of the requirements of the degree of

Bachelor of Engineering

By

Anne Pinto	**53**
Minita Joshee	**23**
Amaan Nizam	**46**
Abhiram Pillai	**52**

Supervisor: **Dr. Satishkumar Chavan**

Department of Electronics and Telecommunication

Don Bosco Institute of Technology

University of Mumbai

2020 - 2021

Don Bosco Institute of Technology

(Affiliated to the University of Mumbai)
Premier Automobiles Road, Kurla, Mumbai - 400070)

CERTIFICATE

This is to certify that the project entitled **BREAST CANCER DETECTION AND CLASSIFICATION** is a bonafide work of

Anne Pinto	**53**
Minita Joshee	**23**
Amaan Nizam	**46**
Abhiram Pillai	**52**

submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **"Undergraduate"** in **"Bachelor of Engineering"**. This approval does not necessarily endorse or accept every statement made, opinion expressed or conclusion drawn as recorded in the report. It only signifies the acceptance of the report for the purpose for which it has been submitted.

Dr. Satishkumar Chavan
Project guide

Dr. Ashwini Kotrashetti
HOD,EXTC Department

Dr. Prasanna Nambiar
Principal

Don Bosco Institute of Technology

(Affiliated to the University of Mumbai)
Premier Automobiles Road, Kurla, Mumbai - 400070)

Project Report Approval for B.E.

This project report entitled **BREAST CANCER DETECTION AND CLASSIFICATION** by **Anne Pinto, Minita Joshee, Amaan Nizam and Abhiram Pillai** is approved for the degree of Bachelor of Engineering in Electronics and Telecommunication.

Examiners

1.

2.

Date: 27/12/2020
Place: **Kurla, Mumbai**

DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea / data / fact / source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Anne Pinto **Minita Joshee**

Amaan Nizam **Abhiram Pillai**

Date: 27/12/2020

ACKNOWLEDGEMENT

A project is a teamwork which involves the contribution of many people. We would like to thank everyone who have contributed by taking interest in our work and motivating us all the way through. Sincere thanks to our parents, teachers and our guide Dr. Satishkumar Chavan who made this project a possibility.

Date: 27/12/2020

ABSTRACT

One of the most prevalent cancers among women is breast cancer. Accurate diagnosis of breast cancer at an early stage can reduce the mortality associated with this disease. Diagnosis can be done with screening mammography. The main challenge of screening mammography is that it is associated with high risk of false positives and false negatives. In this work, we propose an automated system for detection, segmentation of tumours and classifying them into Benign and Malignant using mammograms in Cranial-Caudal and Medial-lateral oblique (CC and MLO) views. This will be done using deep learning models to improve the accuracy of screening mammography. Such an automated system will further help radiologists with the predictive accuracy of screening mammography. The comparative analysis of different classification models, present in this work, shows the highest accuracy models based on simple detection. We will also show the segmentation of tumors or infected tissues from the mammogram. The dataset used in this work, namely MIAS dataset, is popular and free to use as well and can for further experimenting and segmentation, we intend to way on DDSM and TMH datasets.

Contents

List of Figures

List of Tables

List of Abbreviations

ANN Artificial Neural Network

CCN Convolutional Neural Network

CC Cranial Caudal

DDSM Digital Database for Screening Mammography

ILSVRC ImageNet Large Scale Visual Recognition Challenge

MIAS Mammographic Image Analysis Society

MLO Mediolateral Oblique

MRI Magnetic Resonant Imaging

TMH Tata Memorial Hospital

VGG Visual Geometry Group

Chapter 1

Introduction

Breast cancer has the second highest mortality rate in women next to lung cancer [1]. According to the global cancer statics, the number of new cases in 2018 was estimated to be 18,078,957 and deaths 9,555,027 (52.85%) globally. The cases of breast cancer amounts to 2,088,849 (11.55%) and the deaths is estimated to be 626,679 (6.56%). 60% of the deaths occur in low income developing countries like Ethiopia [2] [3] [4] [5]. If the cancer is detected early, it increases expectancy of survival rate/mortality of patient. Existence of breast cancer can be indicated through many abnormalities such as masses, micro-calcifications and areas of symmetry and distortion. Among these, masses are the most representative and common lesion type. . However, masses can be easily hidden by overlapping breast tissues, making it difficult to detect them. Masses can be of two types, namely, undetected and misindentified. An undetected mass is a false negative, which delays a patient diagnosis until the next screening. A misidentified mass is a false positive, which leads to additional tests including re-screening and biopsy, causing unnecessary anxiety and pain to patients [6]. Many mammographic density ratings, ranging from manual categorical (e.g. BI-RADS) to automatic continuous scores, have been suggested. Radiologists characterized the mammographic appearance in the early years by a series of intuitive yet poorly defined patterns of breast tissue that have been shown to contribute to the risk of breast cancer. As obtained by Cumulus-like thresholding, the present gold standard is semi-automated continuous ratings.The strength threshold is set by the radiologist in Cumulus to distinguish radio dense (white appearing) from fatty (dark appearing) tissue. The computer then calculates the percentage of dense to total breast area known as the mammographic density percentage (PMD). User-assisted thresholding, however is subjective and time consuming, and is therefore not appropriate for high epidemiological thresholds. Moreover, doctors experience tiredness and fatigue going through 50-60 mammograms every day. A trend has been towards completely automating PMD scoring, but most of these methods depend on handcrafted features with many parameters that needs to be monitored. Hence, there arises the need for a more robust, fast, accurate, and efficient noninvasive cancer detection system. an automated

system is required for achieving error-free detection of breast cancer using mammogram.

1.1 Breast Anatomy

As we know human beings are mammals and every mammal has a mammary gland, breasts are supposed to be the mammary glands in humans. They are specialized organs on the anterior of the chest and are found to be more developed in females than in males because primary function in females is lactation. They produce milk for the infant.

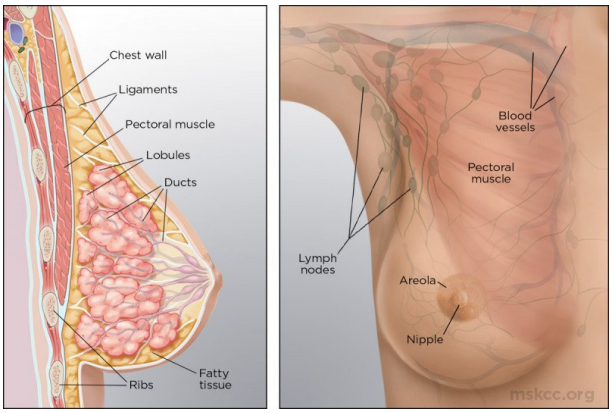


Figure 1.1: Structural anatomy of the breast

As shown in Figure 1.1, breasts are mainly made up of different kinds of tissues, glands, ducts and lymph nodes. Fatty, Connective (Glandular) or Fibrous and Glandular tissues consist of lobules and ducts. Fibrous tissues include Ligaments and scar tissues. Fatty tissues fills gap between glandular and fibrous tissue. Milk glands (lobules) that produce and supply milk. They are the sack like structures. Gland that produces mil there are 15-20 glands present which are known as lobes each of which has smaller lobules and they are arranged in clusters. Special ducts that transfer milk from the lobules to the nipple. These are thin tubes that carry milk from the lobes to the nipple and are located in the middle of the areola. Breast cancers can form in lobes and ducts. Fat fills space between lobes and ducts. The nipple is the centre of the areola. The nipple is surrounded by lobes in a radial manner.Areola (pink/brown pigmented region surrounding the nipple) is the area surrounding the nipple. Lymph nodes and Lymph are small bean shaped organs that help fight infection. A colourless fluid called lymph is produced by them which contain WBCs (lymphocytes).

1.2 Types of Breast Cancer and Detection Techniques

The most common medical conditions found in breasts is cancer. As depicted in Figure 1.2, breast cancer can be of two types i.e. Malignant (Cancerous) and Benign (Non-Cancerous) conditions. Cancerous medical conditions can include formation of lump, discharge from nipple pitting, change in shape etc. Benign conditions include Fibrosis, Simple cyst formation. Cancers can also be differentiated as Ductal or Lobular. Further the cancers can be classified into In situ or Invasive Carcinoma and Proliferative or Non-proliferative.

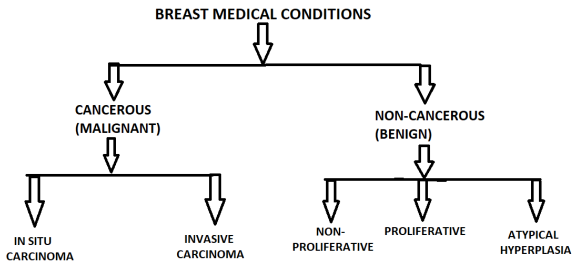


Figure 1.2: Classification of Breast Cancer

There are two main types of breast cancer namely: Ductal and Lobular. Figure 1.3 represents the two types of cancers and where exactly does the cancer form. Ductal Carcinoma is a common type of breast cancer that starts in cells

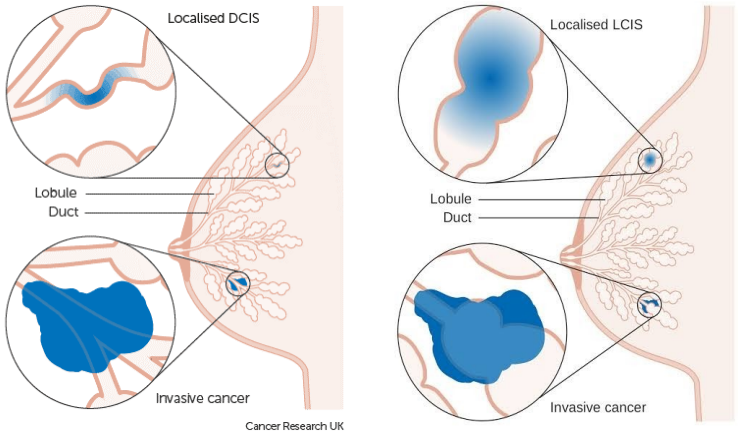


Figure 1.3: Ductal and Lobular Carcinoma In-situ

that line the mil ducts, which carry milk to the nipple. It is divided into 2 types: Ductal Carcinoma In-situ also known as intraductal carcinoma accounts for one of every five new breast cancer diagnoses. It's an uncontrolled growth of cells within the breast. It is non-invasive meaning it hasn't grown into the breast tissue outside of the ducts. The phrase in-situ means "it's in the original place". It's a stage 0 breast cancer. Invasive Ductal Carcinoma accounts for about 80% of all invasive breast cancers in women and 90% in men. It begins in the cells of the milk duct, then it grows into the duct walls and into the surrounding breast tissue. It can spread to the other parts of the body as well.
Lobular Carcinoma is another type of breast cancer that is commonly observed in women. It is also divided in 2 types:

Lobular Carcinoma In-situ isn't cancer, but it is thought to be a sign that breast cancer may develop. With this condition, there are abnormal cells in multiple lobules of the breasts, but these cells rarely spread to other parts of breasts. Even though this condition doesn't spread, it's important to keep an eye on it. Between 20% to 40% of women with this condition will develop a separate invasive breast cancer that will grow outside its original location within the next years. Invasive Lobular Carcinoma is the second most common form of invasive breast cancer. It begins in one of the breast lobules, then spreads to other parts of the breast. It is most likely to be found in both breasts. It can spread to other areas of the body.
Breast cancer can be symptomatic or asymptomatic. Symptoms for the most common breast cancer include a breast lump or tissue thickening, breast pain, swelling in most part of breast, changes in appearance of the skin near the breast. There are cases where no symptoms can be presented but the only way to detect is through a mammogram or an ultrasound.
To locate the region of abnormality, breasts are divided into four quadrants as shown in figure 1.4. Vertical quadrants are divided as inner and outer portions whereas horizontal quadrants are divided as upper and lower portions. Physicians sometimes describe the location of a breast tumour. There are various

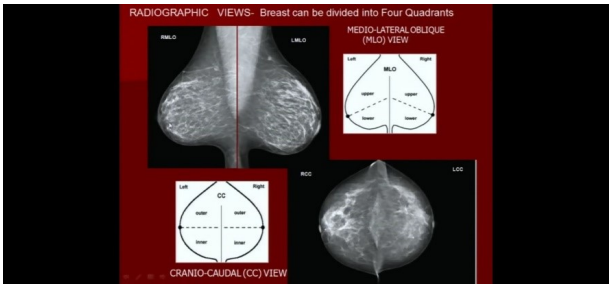


Figure 1.4: Quadrants of the breast

screening techniques available for the detection of breast cancer. The most common being mammography. To detect early stage breast cancer, the individual patient's risk profile can serve as a good guideline for the entire screening pro-

cess. During mammography, the key factor is the position of the breast. There are two most common views which are CC and MLO as depicted in Figure 1.5 and Figure 1.6.

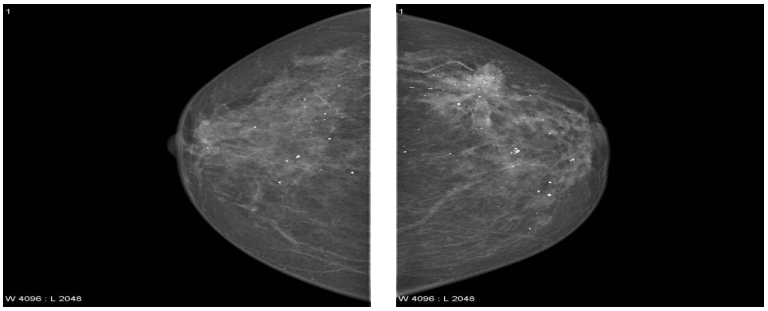


Figure 1.5: CC view in a mammogram

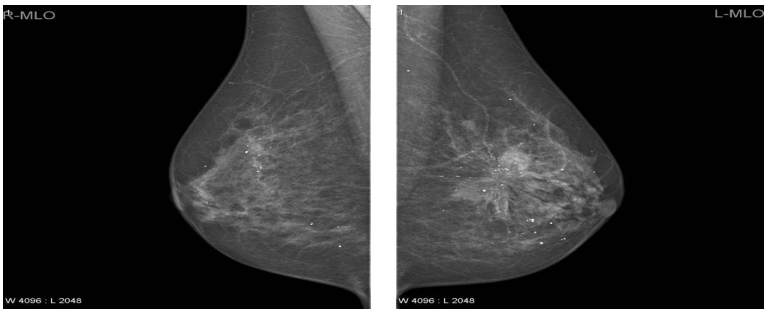


Figure 1.6: MLO view in a mammogram

Though the most common, mammography has its limitations which include low sensitivity in dense breasts. For this reason, ultrasound or MRI can be valuable adjunctive in clinical examinations. A combination of mammography, ultrasound, and MRI has shown to give very good results in detecting ductal carcinoma in situ and invasive cancer. Their combined use increases the likelihood of detecting breast cancer, and reduces the rates of false detections. However, further improvements of these medical imaging techniques are needed in order to overcome current limitations and to increase their effectiveness in detecting breast cancer.

Chapter 2

Literature Survey

Mammography can be used as a noninvasive method for screening purposes and assisting the radiologists in prognosis and treatment. Bray et al. [3] have evaluated a status report of the various types of cancer that are diagnosed in both sexes combined and mainly among females, breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death. In both sexes, lung cancer is the most commonly diagnosed cancer (11.6% of the total cases) and the leading cause of cancer death(18.4% of the total cancer deaths), closely followed by female breast cancer (11.6%), prostate cancer (7.1%), and colorectal cancer (6.1%) for incidence and colorectal cancer (9.2%), stomach cancer (8.2%), and liver cancer (8.2%) for mortality. Islam et al. [7] have explored the different breast cancer imaging techniques such as mammography, magnetic resonance imaging (MRI), and ultrasound for breast cancer diagnosis and compared their effectiveness, advantages, and disadvantages for detecting early-stage breast cancer. Singh et al. [8] says that error-free diagnosis of breast cancer at an early stage can reduce the mortality associated with the disease. Thermographic image and usage of artificial neural networks have improved the accuracy of thermography in early diagnosis of breast abnormality. They evaluated the role of image thermography in early breast cancer detection. [9] introduced a multistage system for detecting MCs in mammograms. They used a back-propagation neural network to find candidate calcification regions first, cleaned network output to remove thin elongated structures and used a measure of local density for final classification. [10] applied a two-level algorithm for the detection of MCs using diverse-AdaboostSVM. Six features (four wavelet plus two gray level features) were computed for neural network to detect candidate MC pixels. As a result, 25 features from candidate MCs were extracted and further reduced with geometric linear discriminant analysis (GLDA). The classifier was built with diverse Adaboost SVM.

A mass in mammogram is defined as a space-occupying lesion seen in more than one projection [11]. The general procedure for detecting mass is first to detect suspicious regions, then extract shape and texture features, and finally detect mass regions through classification or removing false positive regions [12].

[13] used texture features to distinguish mass and non-mass regions. [14] used an adaptive density-weighted contrast enhancement filter to obtain potential masses and used Laplacian Gaussian for edge detection. Morphological features were extracted for classifying normal and mass ROIs. [15] first segmented the boundary of ROI using an edge-based approach and then computed geometric and shape features. Neural networks were trained to distinguish true mass from normal regions.

Since the 1990's CNN's applications can be traced in the field of medical diagnostics just when calcification where detected in digital mammogram. Due to their outstanding performance CNN's have attracted significant attention for feature extraction. There was a breakthrough in 2008 for image classification. Krizhevsky et al. [16] trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. Simonyan et al. [17] investigated the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting.Increasing depth using an architecture with very small (3x3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16-19 weight layers. Selvanthi et al. [1] deployed the deep learning techniques such as convolutional neural network,sparse autoencoder, and stacked sparse autoencoder. The performance of these techniques is analyzed and compared with the existing methods. From the analysis, it is observed that the stacked sparse autoencoder performs better compared to other methods.

Chapter 3

Methodology

The overall plan followed throughout the work is depicted in 3.1.

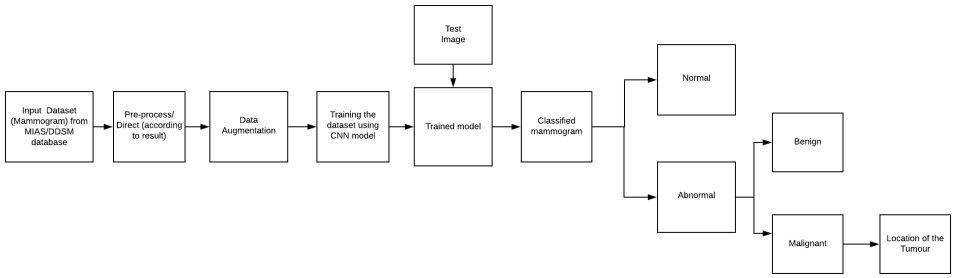


Figure 3.1: Block Diagram

To start with the project, we worked on a simple image classification algorithm using CNN because CNNs can be thought of automatic feature extractors from the image. It effectively uses adjacent pixel information to down-sample the image first by convolution and then uses prediction layer at the end. After studying the four classification models, we were left to decide whether to go for supervised or unsupervised learning model [18].
In a supervised learning model, the algorithm learns on a labeled dataset, providing an answer key that the algorithm can use to evaluate its accuracy on training data, where as an unsupervised model, in contrast, provides unlabeled data that the algorithm tries to make sense of by extracting features and patterns on its own.
There are two main areas where supervised learning is useful: classification problems and regression problems.
Classification problems ask the algorithm to predict a discrete value whether

the image belongs to one of the many mentioned classes i.e identifying the input data as the member of a particular class or group. In a training dataset of animal images, that would mean each photo was pre-labeled as one defined class or the other. The algorithm is then evaluated by how accurately it can classify new images of other similar classes of data.

3.1 Dataset

The input dataset was divided into two categories i.e training dataset and test dataset. Training dataset includes the images which were trained with four models while test dataset was the one which provided unbiased evaluation of the model on the training dataset.
After successfully implementing the above models and analysing the efficiency of each, the top 2 models were chosen to be implemented on the MIAS dataset.

MIAS Dataset: The MIAS database contains the original 322 images (161 pairs) at 50 micron resolution in "Portable Gray Map" (PGM) format and associated truth data which depict CC and MLO views of the scanned images [19]. It is a supervised dataset which consists of the class, radius, severity and X and Y coordinates of each scan. 3.2 shows a Cancerous mammogram in MIAS dataset.

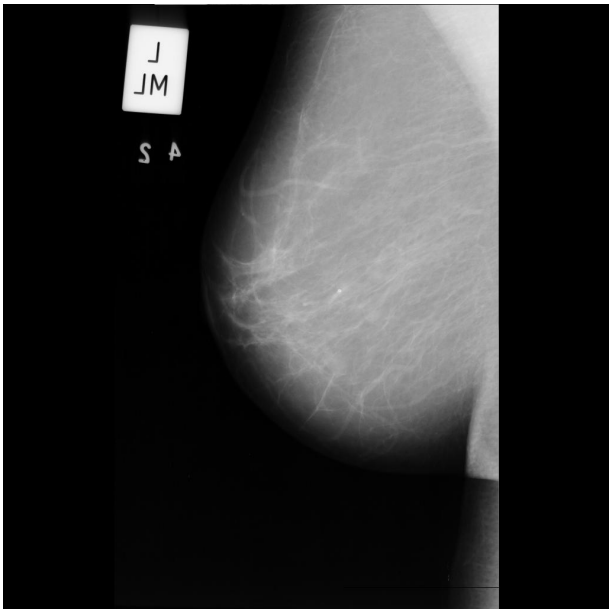


Figure 3.2: Sample image of cancerous mammogram used in MIAS dataset

DDSM Dataset: Another database used is DDSM which is a database of digitized film-screen mammograms with associated ground truth and other information. The purpose of this resource is to provide a large set of mammograms in a digital format that may be used by researchers to evaluate and compare the performance of computer-aided detection (CAD) algorithms. The database was completed in the fall of 1999. It contains 2620, four view, mammography screening exams [20]. This dataset contains images that have been pre-processed and converted to 299x299 images by extracting the ROIs. The data is stored as tfrecords files for TensorFlow. The dataset contains 55,890 training examples, of which 14% are positive and the remaining 86% negative, divided into 5 tfrecords files. Following the entire plan, the algorithm will detect normal or abnormal scan and from the abnormal scans, it will further detect a benign or malignant tumor.

3.2 Pre-processing

Based on the results obtained, pre-processing of dataset can be done to reduce the noise and convert the raw data into understandable format. It helps to remove unwanted information from the image and allows users to have more valuable datasets. Existing datasets are always incomplete and data cannot be sent to a training model directly. If done so, it always would cause errors. Thus, data pre-processing plays a major role in any classification algorithm.

3.3 Data Augmentation

The problem with small datasets is that models trained with them do not generalize well data from the validation and test set. Hence, these models suffer from the problem of overfitting. Data augmentation is another way we can reduce overfitting on models, where we increase the amount of training data using information only in our training data [21]. Data augmentation refers to increasing the amount of data available by changing various parameters of the image. The amount of data can be increased by rotating the image, changing the dimensions, flipping the image. It is an effective way to increase the efficiency of the model and reduce the losses while training.

3.4 Classification Techniques in Deep Learning

3.4.1 Convolutional Neural Networks

In the past few years, Deep Learning has become an area of interest to the researchers. CNN overcomes the limitations of traditional machine learning approaches. It is a deep learning approach that is widely used for solving complex problems. Machine learning promises to reduce the efforts by making the machines to learn themselves through past experiences [22] using three approaches

of learning namely, learning under supervision, without supervision and semi-supervised learning [23]. Deep belief networks, recurrent neural networks, convolution neural networks etc. are different deep learning architectures. Convolution Neural Network, often called ConvNet, has deep feed-forward architecture and has astonishing ability to generalize in a better way as compared to networks with fully connected layers [24].

[25] describes CNN as the concept of hierarchical feature detectors in biologi-

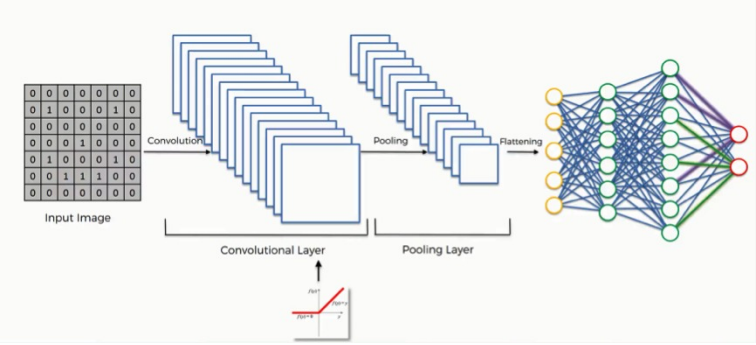


Figure 3.3: Convolutional Neural Network

cally inspired man. There are many considerable reasons why CNN is considered above other classical models. First, the key interest for applying CNN lies in the idea of using concept of weight sharing, due to which the number of parameters that needs training is substantially reduced, resulting in improved generalization. Due to lesser parameters, CNN can be trained smoothly and does not suffer overfitting. Secondly, the classification stage is incorporated with feature extraction stage, both uses learning process. Thirdly, it is much more difficult to implement large networks using general models of ANN than implementing in CNN.

3.4.2 AlexNet

AlexNet is a convolutional neural network, designed by Alex Krizhevsky in 2012. This network was entered in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second best entry [16]. This is a large, deep convolutional neural network consisting of eight layers as shown in Figure 3.4. The key feature of this model is that it can be trained using multiple GPUs for improved speed in training as well as usage of bigger datasets.

ImageNet is a dataset consisting of 15 million labeled high-resolution images, over 22,000 categories. As it consists of variable-resolution images, the authors down-sampled the images to a fixed resolution of 256x256. The primary concern on this dataset is to prevent overfitting, which was observed using ReLU nonlinearity which also influenced the performance of the training model by faster

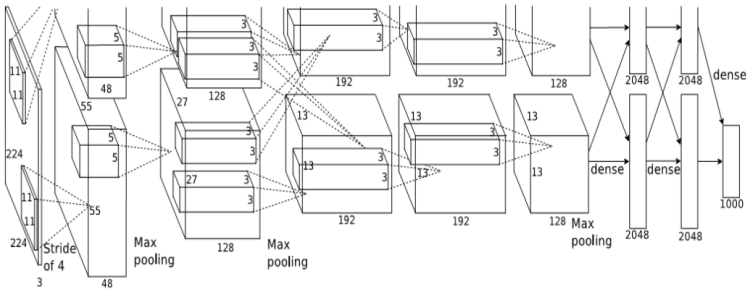


Figure 3.4: AlexNet Architecture

learning. Overlapping pooling reduces computation and controls overfitting.
AlexNet consists of 8 deep layers in total which includes 5 convolutional layers and 3 fully connected (dense) layers as shown in figure 10. Layer one is a convolutional layer with input image size of 224x224x3. Number of filters used in this layer are 96 with filter size of 11x11x3 and stride 4. The output dimension of layer one is 55x55x96 which is split across two GPUs, so 55x55x48 for each GPU. Layer two is a max pooling layer followed by a convolution with input image of size 55x55x96. The max pooling layer results to 27x27x96. Number of filters used are 256 with filter size of 5x5x48. The output dimension changes to 27x27x256 which is again split across two GPUs, so 27x27x128 each GPU. Layer three, four and five follow on similar lines as shown in figure 11. Layer six is fully connected layer with an input of 13x13x128 which is transformed into a vector. The output then results to 1x2048. Layer seven and eight follow on similar lines as shown in figure 11. Lastly the output is passed through softmax activation in order to normalize classification vector.

3.4.3 VGG-16

VGG-16 was proposed by Karen Simonyan and Andrew Zisserman of the Visual Geometry group lab of Oxford University in 2014. This particular network was entered into the 2014 ImageNet challenge [17]. It had 16 weight layers as shown in Figure 3.5. The design of this is very similar to Alexnet except that they have made this a little systematic. They have also stuck to only one filter size, 3, which is the smallest filter size throughout the layers. There are around 138 million parameters.

Input- VGG takes in a 224x224 pixel RGB image. For the ImageNet competition, the authors cropped out the center 224x224 patch in each image to keep the input image size consistent.
Convolutional layers- The convolutional layers in VGG use a very small receptive field ie. 3x3, smallest possible size that still captures left/right and up/down. The convolution stride is fixed to 1 pixel so that the spatial resolu-

Figure 3.5: VGG-16 Architecture

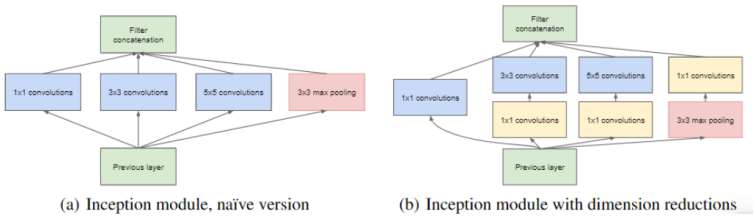
tion is preserved after convolution.
Fully connected layers- VGG has three fully connected layers. 4096 channels each and the last one has 1000 channels softmax layer.
Hidden layers- All of VGG's hidden layers use RelU (a huge innovation from Alexnet that cut training time.) The input to the network is an image of dimension (224x224x3).

The first two layers have 64 channels of 3x3 kernel with padding 1 and stride 1 followed by a maxpool layer of stride 2, size 2x2. The output dimension results in 112x112x65. The third and fourth convolutional layers have 128 channels of 3x3 kernel, followed by a maxpool layer same as before. The output dimension changes to 56x56x128. The fifth, sixth and seventh layers have 256 channels of 3x3 kernel, followed by a maxpool layer. The output dimension then changes to 28x28x256. The eighth, ninth and tenth convolutional layers have 512 channels of 3x3 kernel, followed by a maxpool layer. The output dimension results in 14x14x512. The eleventh, twelfth and thirteenth layers also have 512 channels of 3x3 kernel, followed by a maxpool layer. The output dimension then changes to 7x7x512.
After the stack of convolution and max pooling layers, we flatten the output to make it a feature vector as shown in figure 12. After this there are three fully connected layers, which takes in input from the last feature vector and outputs a 1x1x4096 vector. The output of the three fully connected layers is passed to the softmax layer in order to normalize classification vectors. All the hidden layers use RelU as it's activation function. RelU is more computationally efficient because it results in faster learning and it also decreases the likelihood of vanishing gradients.

3.4.4 GoogleNet

GoogleNet is a deep convolutional network architecture proposed by Google (with collaboration of many universities) in 2014. It is the winner of ILSVRC 2014. It has provided a significant decrease in error rate as compared to previous winners AlexNet (Winner of ILSVRC 2012) and ZF-Net (Winner of ILSVRC 2013) and significantly less error rate than VGG (2014 runner up). It has 22 layers of deep architecture [26]. The architecture was designed to keep compu-

tational efficiency in mind. The idea behind that the architecture can be run on individual devices even with low computational resources. The architecture also contains two auxiliary classifier layers connected to the output of Inception (4a) and Inception (4d) layers.



(a) Inception module, naïve version (b) Inception module with dimension reductions

Figure 3.6: Inception module of GoogleNet Architecture

GoogleNet uses many different methods such as 1x1 convolution and global average pooling that enables it to create a deeper architecture. The inception architecture uses 1×1 convolution in its architecture. These convolutions used to decrease the number of parameters (weights and biases) of the architecture. 1x1 convolution is used to reduce the number of parameters of the architecture further which increases the depth.Global average pooling is another method used in the architecture of GoogleNet. It is used at the end of the network. This layer takes a feature map of 7x7 and averages it to 1x1. This also decreases the number of trainable parameters to 0 and increases accuracy of 0.6%. GoogleNet has an inception module which is different from previous architectures such as AlexNet, ZF-Net. In this architecture, there is a fixed convolution size for each layer as shown in Figure 3.6. In the Inception module 1×1, 3×3, 5×5 convolution and 3×3 max pooling performed in a parallel way at the input and the output of these are stacked together to generate the final output. The idea behind that convolution filters of different sizes will handle objects at multiple scales better.

Chapter 4

Experimental Results

The Dogs vs. Cats dataset is a standard computer vision dataset that involves classifying photos as either containing a dog or cat. The collection consists of images of dogs and cats given from a much larger dataset of 3 million professionally annotated photos as a subset of photos. 25,000 labelled images were given by the Kaggle competition: 12,500 dogs and the same number of cats.
Classification models ask the algorithm to predict a discrete value whether the image is a cat or dog.
Model summary of various networks for classification of cats and dogs are as shown in Figure 4.1, Figure 4.2, Figure 4.3 and Figure 4.4.

Comparision of all these four models can be found in Table 4.1

Table 4.1: Comparison of various networks for classification of cats and dogs

Name of Model	Number of Layers	Accuracy (%)	Validation Accuracy (%)	Loss(%)	Validation Loss(%)	Parameters (Million)
Alex Net	8	49.93	49.69	71.29	69.49	62
VGG-16	16	84.75	90.45	34.18	23.2	138
Efficient Net	17	94.55	96.90	49.17	70.64	5.3
Google Net	22	91.45	96.2	31.36	13.2	22.2

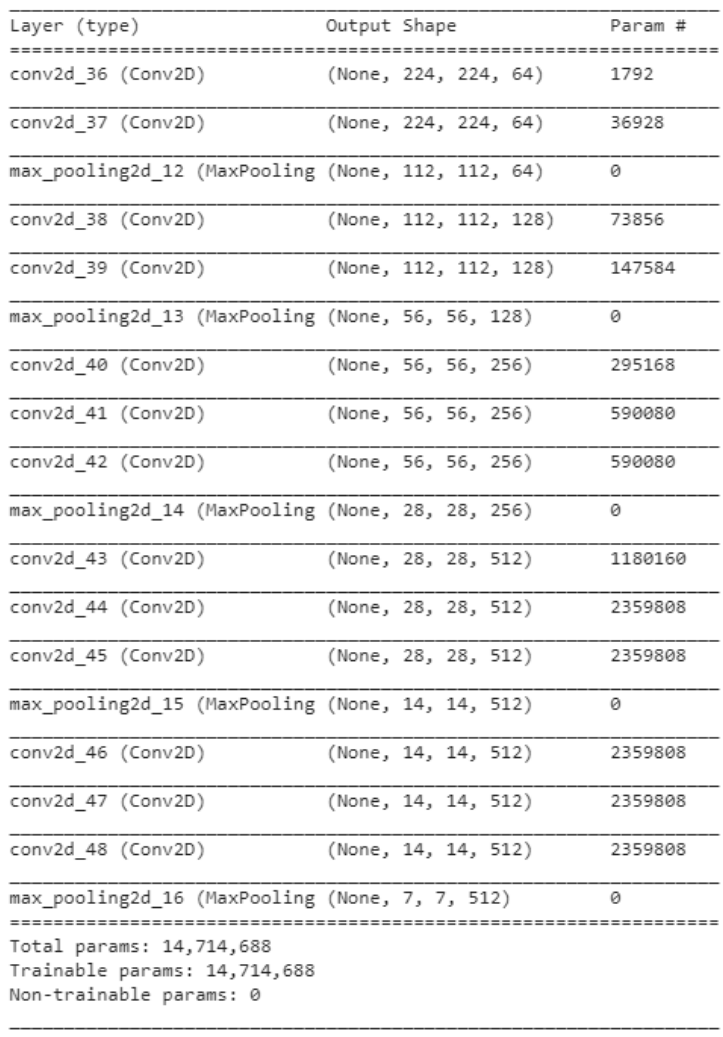
In all the above models we saw how when we add more layers we get more accurate. In short what we are doing is scaling up the models. So the goal of Efficient net is to systematically scale up depth, width and resolution of the CNN's. Process of scaling up CNN's is not well understood; many ways of doing this, mostly arbitrarily chosen. So by randomly increasing layers etc.there 3 types of scaling width scaling,depth and resolution scaling. depth scaling accounts to adding more feature maps.depth scaling adding more layers and making your neural

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 14, 14, 96)	34944
activation (Activation)	(None, 14, 14, 96)	0
max_pooling2d (MaxPooling2D)	(None, 7, 7, 96)	0
conv2d_1 (Conv2D)	(None, 1, 1, 256)	614656
max_pooling2d_1 (MaxPooling2	(None, 1, 1, 256)	0
conv2d_2 (Conv2D)	(None, 1, 1, 384)	885120
conv2d_3 (Conv2D)	(None, 1, 1, 384)	1327488
conv2d_4 (Conv2D)	(None, 1, 1, 256)	884992
max_pooling2d_2 (MaxPooling2	(None, 1, 1, 256)	0
dense (Dense)	(None, 1, 1, 4096)	1052672
activation_1 (Activation)	(None, 1, 1, 4096)	0
dropout (Dropout)	(None, 1, 1, 4096)	0
dense_1 (Dense)	(None, 1, 1, 4096)	16781312
dropout_1 (Dropout)	(None, 1, 1, 4096)	0
dense_2 (Dense)	(None, 1, 1, 1000)	4097000
activation_2 (Activation)	(None, 1, 1, 1000)	0

Total params: 25,678,184
Trainable params: 25,678,184
Non-trainable params: 0

Figure 4.1: Model Summary of AlexNet

network deep.or resolution scaling where the input image resolution is increased . And compound scaling refers to a combination of increasing the width adding more layers and increasing the resolution. Saturation happens if only one type of scaling is done. But if we overall increase the parameters we will get more accurate results.The earlier models like ResNet follow the conventional approach of scaling the dimensions arbitrarily and by adding up more and more layers. However, the paper proposes that if we scale the dimensions by a fixed amount at the same time and do so uniformly, we achieve much better performance. The scaling coefficients can be in fact decided by the user.Though this scaling technique can be used for any CNN-based model, the authors started off with their own baseline model called EfficientNetB0. MBConv stands for mobile inverted bottleneck Convolution(similar to MobileNetv2). They also propose the Compound Scaling formula with the following scaling coefficients:Depth = 1.20, Width = 1.10, Resolution = 1.15.



Layer (type)	Output Shape	Param #
conv2d_36 (Conv2D)	(None, 224, 224, 64)	1792
conv2d_37 (Conv2D)	(None, 224, 224, 64)	36928
max_pooling2d_12 (MaxPooling	(None, 112, 112, 64)	0
conv2d_38 (Conv2D)	(None, 112, 112, 128)	73856
conv2d_39 (Conv2D)	(None, 112, 112, 128)	147584
max_pooling2d_13 (MaxPooling	(None, 56, 56, 128)	0
conv2d_40 (Conv2D)	(None, 56, 56, 256)	295168
conv2d_41 (Conv2D)	(None, 56, 56, 256)	590080
conv2d_42 (Conv2D)	(None, 56, 56, 256)	590080
max_pooling2d_14 (MaxPooling	(None, 28, 28, 256)	0
conv2d_43 (Conv2D)	(None, 28, 28, 512)	1180160
conv2d_44 (Conv2D)	(None, 28, 28, 512)	2359808
conv2d_45 (Conv2D)	(None, 28, 28, 512)	2359808
max_pooling2d_15 (MaxPooling	(None, 14, 14, 512)	0
conv2d_46 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_47 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_48 (Conv2D)	(None, 14, 14, 512)	2359808
max_pooling2d_16 (MaxPooling	(None, 7, 7, 512)	0

Total params: 14,714,688
Trainable params: 14,714,688
Non-trainable params: 0

Figure 4.2: Model Summary of VGG-16

While implementing the models, we found out that AlexNet was very slow as compared to the other models with low efficiency. VGG16 improved the efficiency but the training time was too long. Efficient Net further increased the efficiency but the losses were more. The best efficiency with low number of losses was provided by GoogleNet.

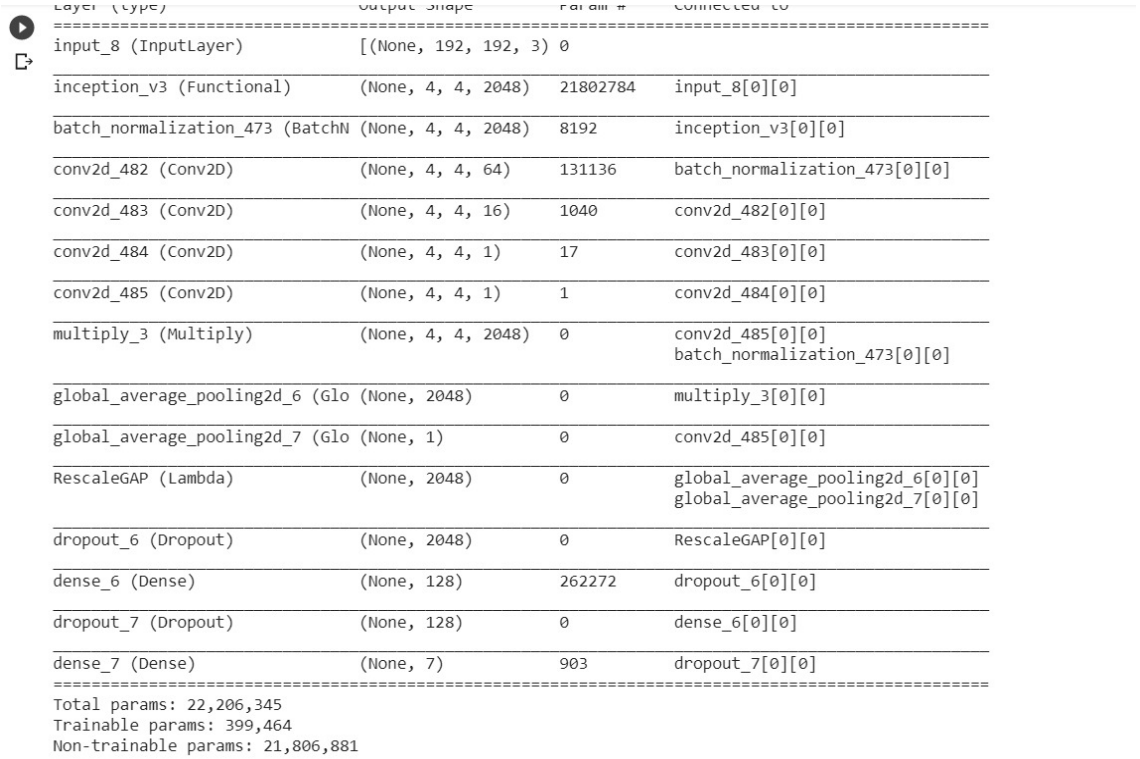
```
model.summary()
```

```
Model: "functional_1"
```

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 224, 224, 3)	0	
stem_conv (Conv2D)	(None, 112, 112, 32)	864	input_1[0][0]
stem_bn (BatchNormalization)	(None, 112, 112, 32)	128	stem_conv[0][0]
stem_activation (Activation)	(None, 112, 112, 32)	0	stem_bn[0][0]
block1a_dwconv (DepthwiseConv2D	(None, 112, 112, 32)	288	stem_activation[0][0]
block1a_bn (BatchNormalization)	(None, 112, 112, 32)	128	block1a_dwconv[0][0]
block1a_activation (Activation)	(None, 112, 112, 32)	0	block1a_bn[0][0]
block1a_se_squeeze (GlobalAvera	(None, 32)	0	block1a_activation[0][0]
block1a_se_reshape (Reshape)	(None, 1, 1, 32)	0	block1a_se_squeeze[0][0]
block1a_se_reduce (Conv2D)	(None, 1, 1, 8)	264	block1a_se_reshape[0][0]
block1a_se_expand (Conv2D)	(None, 1, 1, 32)	288	block1a_se_reduce[0][0]
block1a_se_excite (Multiply)	(None, 112, 112, 32)	0	block1a_activation[0][0] block1a_se_expand[0][0]
block1a_project_conv (Conv2D)	(None, 112, 112, 16)	512	block1a_se_excite[0][0]
block7a_se_expand (Conv2D)	(None, 1, 1, 1152)	56448	block7a_se_reduce[0][0]
block7a_se_excite (Multiply)	(None, 7, 7, 1152)	0	block7a_activation[0][0] block7a_se_expand[0][0]
block7a_project_conv (Conv2D)	(None, 7, 7, 320)	368640	block7a_se_excite[0][0]
block7a_project_bn (BatchNormal	(None, 7, 7, 320)	1280	block7a_project_conv[0][0]
top_conv (Conv2D)	(None, 7, 7, 1280)	409600	block7a_project_bn[0][0]
top_bn (BatchNormalization)	(None, 7, 7, 1280)	5120	top_conv[0][0]
top_activation (Activation)	(None, 7, 7, 1280)	0	top_bn[0][0]
flatten (Flatten)	(None, 62720)	0	top_activation[0][0]
dense (Dense)	(None, 1024)	64226304	flatten[0][0]
dropout (Dropout)	(None, 1024)	0	dense[0][0]
dense_1 (Dense)	(None, 1)	1025	dropout[0][0]

```
Total params: 68,276,893
Trainable params: 64,227,329
Non-trainable params: 4,049,564
```

Figure 4.3: Model Summary of Efficient Net



Layer (type)	Output Shape	Param #	Connected to
input_8 (InputLayer)	[(None, 192, 192, 3)	0	
inception_v3 (Functional)	(None, 4, 4, 2048)	21802784	input_8[0][0]
batch_normalization_473 (BatchN	(None, 4, 4, 2048)	8192	inception_v3[0][0]
conv2d_482 (Conv2D)	(None, 4, 4, 64)	131136	batch_normalization_473[0][0]
conv2d_483 (Conv2D)	(None, 4, 4, 16)	1040	conv2d_482[0][0]
conv2d_484 (Conv2D)	(None, 4, 4, 1)	17	conv2d_483[0][0]
conv2d_485 (Conv2D)	(None, 4, 4, 1)	1	conv2d_484[0][0]
multiply_3 (Multiply)	(None, 4, 4, 2048)	0	conv2d_485[0][0] batch_normalization_473[0][0]
global_average_pooling2d_6 (Glo	(None, 2048)	0	multiply_3[0][0]
global_average_pooling2d_7 (Glo	(None, 1)	0	conv2d_485[0][0]
RescaleGAP (Lambda)	(None, 2048)	0	global_average_pooling2d_6[0][0] global_average_pooling2d_7[0][0]
dropout_6 (Dropout)	(None, 2048)	0	RescaleGAP[0][0]
dense_6 (Dense)	(None, 128)	262272	dropout_6[0][0]
dropout_7 (Dropout)	(None, 128)	0	dense_6[0][0]
dense_7 (Dense)	(None, 7)	903	dropout_7[0][0]

Total params: 22,206,345
Trainable params: 399,464
Non-trainable params: 21,806,881

Figure 4.4: Model Summary of Google Net

Chapter 5

Plan of Work

There will always be tradeoffs when determining the right model. It is necessary to know that the MIAS dataset, even though containing 8000 images, is incomplete when constructing a deep learning model, that different types of images in dataset need different methods and resources, and that there will often be tradeoffs when deciding the correct model. Hence, choosing the most efficient model out of the four models, which also gives the least number of losses, is a big challenge in itself. Further, after determining, we would work on the MIAS dataset with the basic CNN model and also all the other models.
The roadmap to deep learning is riddled with trial and error. Challenges will emerge during this process particularly with image processing and deciding the correct model to preprocess the image. Six-eight months are usually required to provide us with improved efficient of any model that has been built.
We will be working on the DDSM dataset and finally a dataset provided by Tata Memorial. Additionally, we will also be trying to locate and segment the tumour, calculating the volume of cancerous tissue.
Our full plan of work is depicted in Figure 5.1.

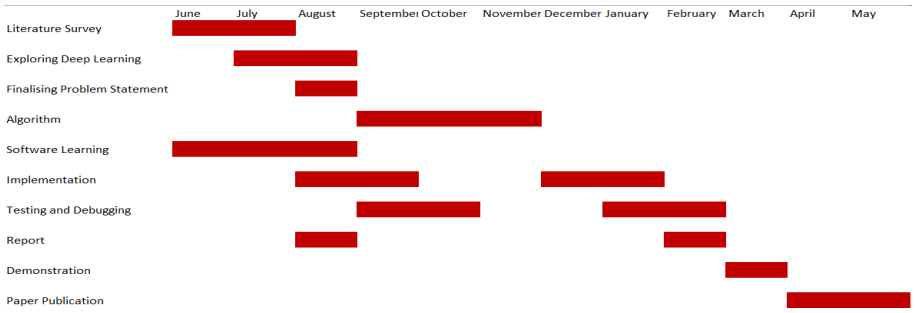


Figure 5.1: Plan of Work

Chapter 6

Conclusion

In the proposed work, the model we developed detects mass abnormality and classify both in to Benign and Malignant. We also compare the recent deep learning models in breast cancer detection and will segment the tumours in the mammograms. We worked on mammographic images available publicly such as MIAS dataset and will work on DDSM and TMH datasets in future. GoogleNet, despite being a complex CNN model achieves better results and produces much better efficiency for the mass detection as compared AlexNet, VGG16 and ResNet. This is because of high accuracy obtained due to more trainable parameters and faster network. The models that use balanced number for the classes the model shall detect, are more fair and achieve better overall accuracy than the unbalanced sized data models.
The proposed methodology for the work includes methods for image pre-processing, mass detection, data augmentation, and finally predicting the outcome of new data given to the model. These techniques have resulted in a much more normalized dataset and increased the training as well as testing images.
This helps with early diagnosis of breast cancer. The classification of breast cancer using the tool we have developed will assist the radiologists to prioritize the mammograms. It will help provide early treatment of patients which will be useful for planning treatments such as chemotherapy and radiotherapy. This will be used to detect the survival ratio based on the survival rates.

Bibliography

[1] D. Selvathi and A. A. Poornila, "Deep learning techniques for breast cancer detection using medical image analysis," in *Biologically rationalized computing techniques for image processing applications*, pp. 159–186, Springer, 2018.

[2] S. Hadush, Y. Girmay, A. Sinamo, and G. Hagos, "Breast cancer detection using convolutional neural networks," *arXiv preprint arXiv:2003.07911*, 2020.

[3] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[4] E. Hadgu, D. Seifu, W. Tigneh, Y. Bokretsion, A. Bekele, M. Abebe, T. Sollie, S. D. Merajver, C. Karlsson, and M. G. Karlsson, "Breast cancer in ethiopia: evidence for geographic difference in the distribution of molecular subtypes in africa," *BMC women's health*, vol. 18, no. 1, pp. 1–8, 2018.

[5] S. V. Sree, E. Y.-K. Ng, R. U. Acharya, and O. Faust, "Breast imaging: a survey," *World journal of clinical oncology*, vol. 2, no. 4, p. 171, 2011.

[6] G. Hamed, M. Marey, S. Amin, and M. Tolba, *Deep Learning in Breast Cancer Detection and Classification*, pp. 322–333. 03 2020.

[7] M. S. Islam, N. Kaabouch, and W. Hu, "A survey of medical imaging techniques used for breast cancer detection," *IEEE International Conference on Electro-Information Technology , EIT 2013*, pp. 1–5, 2013.

[8] D. Singh and A. K. Singh, "Role of image thermography in early breast cancer detection- past, present and future," *Computer Methods and Programs in Biomedicine*, vol. 183, p. 105074, 2020.

[9] Z. Yang, M. Dong, Y. Guo, X. Gao, K. Wang, B. Shi, and Y. Ma, "A new method of micro-calcifications detection in digitized mammograms based on improved simplified pcnn," *Neurocomput.*, vol. 218, p. 79–90, Dec. 2016.

[10] F. Harirchi, P. Radparvar, H. A. Moghaddam, F. Dehghan, and M. Giti, "Two-level algorithm for MCs detection in mammograms using diverse-adaboost-SVM," in *2010 20th International Conference on Pattern Recognition*, IEEE, aug 2010.

[11] *ACR BI-RADS®-Atlas der Mammadiagnostik*. Springer Berlin Heidelberg, 2016.

[12] Y. Li and H. Chen, "A survey of computer-aided detection of breast cancer with mammography," *Journal of Health & Medical Informatics*, vol. 7, no. 4, 2016.

[13] A. Petrosian, H.-P. Chan, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis," *Physics in Medicine and Biology*, vol. 39, pp. 2273–2288, dec 1994.

[14] N. Petrick, H.-P. Chan, B. Sahiner, and D. Wei, "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection," *IEEE Transactions on Medical Imaging*, vol. 15, no. 1, pp. 59–67, 1996.

[15] D. Cascio, F. Fauci, R. Magro, G. Raso, R. Bellotti, F. D. Carlo, S. Tangaro, G. D. Nunzio, M. Quarta, G. Forni, A. Lauria, M. Fantacci, A. Retico, G. Masala, P. Oliva, S. Bagnasco, S. Cheran, and E. Torres, "Mammogram segmentation by contour searching and mass lesions classification with neural network," *IEEE Transactions on Nuclear Science*, vol. 53, pp. 2827–2833, oct 2006.

[16] A. Krizhevsky and I. Sutskever, "H. geoffrey e.,"alex net,"," *Adv. Neural Inf. Process. Syst*, vol. 25, pp. 1–9, 2012.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[18] R. Sathya and A. Abraham, "Comparison of supervised and unsupervised learning algorithms for pattern classification," *International Journal of Advanced Research in Artificial Intelligence*, vol. 2, no. 2, pp. 34–38, 2013.

[19] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, *et al.*, "Mammographic image analysis society (mias) database v," vol. 1, no. 21, 2015.

[20] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, "The digital database for screening mammography, iwdm-2000," in *International Workshop on Digital Mammography. Medical Physics Publishing, Madison WI*, pp. 212–218, 2001.

[21] J. Wang, L. Perez, *et al.*, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Networks Vis. Recognit*, vol. 11, 2017.

[22] J. Carbonell, R. Michalski, and T. Mitchell, "An overview of machine learning in "machine learning," *An Artificial Intelligence Approach", Morgan Kaufman Publishers (Ed.)*, 1983.

[23] S. C. Dharmadhikari, M. Ingle, and P. Kulkarni, "Empirical studies on machine learning based text classification algorithms," *Advanced Computing*, vol. 2, no. 6, p. 161, 2011.

[24] C. Nebauer, "Evaluation of convolutional neural networks for visual recognition," *IEEE transactions on neural networks*, vol. 9, no. 4, pp. 685–696, 1998.

[25] J. Fieres, J. Schemmel, and K. Meier, "Training convolutional networks of threshold neurons suited for low-power hardware implementation," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 21–28, IEEE, 2006.

[26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.