An Application of Deep Learning and Transfer Learning for Breast Cancer Detection Using Mammography Images

By

Nadia Sultana - 011161229

Md.Amir Hamza Faisal - 011161159

Sharika Nargis - 011161125

Tahsin Tabia - 011162139

Shirmine Naher - 011142122

Submitted in partial fulfilment of the requirements
of the degree of Bachelor of Science in Computer Science and Engineering

July 7, 2021

Department of Computer Science and Engineering
United International University

Abstract

Breast cancer is the most common cancer and second most common cause of death. Statically, about 1 in 8 U.S women is identified with breast cancer last year. For the detection of breast cancer, detailed analysis of medical images obtained with breast cancer is highly important. There have been many deep learning methods which are used to diagnose the medical images but particularly we are using Convolutional Neural Network (CNN) to classify different variants of mammograms. To evaluate the visual images and analyze image classification, CNN has already gained the attention of researchers. To extract the image patches, we propose a technique to train the CNN and the sequence of these patches for classification. We propose the newly generated transfer learning method- where a model is created for a project in first phase, then again it will be reused in the second phase as starting point. With single and multiple layer of CNN and ANN, two other methods- Auto-Encoder and VGG16 are introduced to validate and compare results among the methods with different datasets.

Acknowledgements

This project was supported by United International University and partially supported by Bardem General Hospital and National Cancer Research Institute and Hospital. We thank Dr. Swakkhar Shatabda for supervising us. We like to thank Dr. Dewan Md. Farid for sharing his pearls of wisdom with us during the course of this research. We are also grateful to our parents and faculties for supporting us always.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, the some informations and statistics are shared about breast cancer. The main motivations are described alongside with the objectives of this project. As it is one of the most caused cancer in the world, so we decided to do something that can be trusted by the people and we tried to develop our application with a purpose to gain more effective results.

1.1 Project Overview

Breast cancer is a significant threat to women's life and health. According to mortality rate, breast cancer is ranked second out of all female diseases [1]. In 2019, an estimated 1,762,450 new cases of invasive breast cancer are expected to be diagnosed in women in United States and It is estimated that 606,880 deaths from breast cancer will occur this year [2]. Breast cancer can be defined as the abnormal growing of tissue in the breast or a lump in the breast or an armpit [3]. The mortality rate of breast cancer can be reduced if they are detected early.

The mammographic images are widely using in early screening of breast cancer because of its relatively low expense [4]. This is the first step to estimate the risk of having and developing breast cancer [5]. In recent times, the accuracy of detecting cancer from mammograms has been showing impressive performance. Sometimes in actual diagnosis, radiologist can't identify cancer because of the complex structure of breast where computer-aided diagnosis (CAD) for breast cancer can identify more perfectly [4]. The classical CAD of breast cancer consist of three ways: (a) First of all, processed mammograms images and locate the region of tumor (b) then, extract feature from the shape, texture and density of tumor (c) then it diagnoses benign and malignant using all features [4].

Although, the classical CAD accuracy percentage is good but its accuracy still has to be more improve [6]. In recent years, deep learning methods, such as convolutional neural network (CNN) have shown exceptional performances for classification of medical images such as mammograms and it have been successfully applied with a great improvement on accuracies in many applications, such as image recognition, speech recognition, and natural language processing. CNN consist of three layer: (a)convolution layer (b) pooling layer (c) activation layer which help to extract hierarchical features from image data without the manual selection, which is also called objective features [7]. According to medical view deep learning is used to detect cancer cells automatically. From the beginning, train a deep convolution network is difficult because it needs very large amount of data for training. Recent studies showed that CNNs are highly useful to predict high accuracy on different types of medical images classification.

1.2 Motivation

Breast cancer is one of the main causes of death for women internationally [6]. Early detection of breast cancer would be a noble work for the women out there in the society. It is estimated that 42,260 deaths (41,760 women and 500 men) from breast cancer will occur this year . And it is expected that more than 8% of women will develop breast cancer during their lifetime [6]. So, this came up to our mind if we could make an application that can give faster and accurate breast cancer detection result. If any woman can have the result earlier of her breast cancer result, she would take the major steps to prevent this as early as possible. Mammography is the most efficient strategy for an early diagnosis of the oddity which could spot a tumor [7]. The most frequently used and efficient technique for breast cancer detection is digital mammography (DM) [1]. Contemplative studies show that, in current breast cancer screenings, 10%–25% of the tumors are missed by the radiologists [7]. The causes of these false-negative screening exami-nations are not plesant [7]. These mammograms are used as the dataset to train this application. Previously, mammograms of different datasets were used to predict the result but we propose to use a new collection of breast cancer mammograms as datasets. This would give us a more reliable and factual result to detect the problem.

1.3 Objectives

The main objective of this project is to use prediction algorithms to gain better and faster performance. Convolutional neural network (CNN) is used widely to classify the medical images as we are also using transfer learning to predict which is new deep learning method interested by the researchers. In recent times, many works have done using CNN but our aim is to predict the specific detection outcomes where false positive results would be less than any other predictive algorithms.

We are using a complete new collection of real mammograms based dataset that are collected from hospitals of breast cancer predicted women. This new collection of dataset would be much helpful to assist the model for achieving better performance.

Our aim is to make such an application where non computer science students or other general people can use it. We have seen that computer science based students can recognize these problems as they are familiar about the computer based terms and processes related to these things. This would be a great achievement in the bioinformatics field as anyone who has the idea about the problem can operate this application.

1.4 Methodology

Our cancer detection system consists of five steps: breast image pre-processing, mass detection, feature extraction, training data generation, classifier training. Currently breast cancer recognition follow two most common ways for designing image recognition system [8]. The approaches we are following is deep learning trend, where convolutional neural network (CNN) mainly we are using for breast cancer detection and we also using transfer learning process in our project. Mammographic Image Analysis Society (MIAS) and Digital Dataset for Screening Mammography (DDSM) are two datasets used for development of breast cancer classification system. (a) DDSM consists of 55,890 images where 14% is positive and 89% is negative and every images size is 299×299 and all images are pre-processed . (b) MIAS is consists of 322 images where every image size is 128×128 .

A. Convolutional Neural Network: Convolutional neural networks(CNN) have currently made huge success in the field of image recognition, object detection and segmentation of images [9]. In this paper, we have tried to improve the breast cancer detection performance based on CNN.

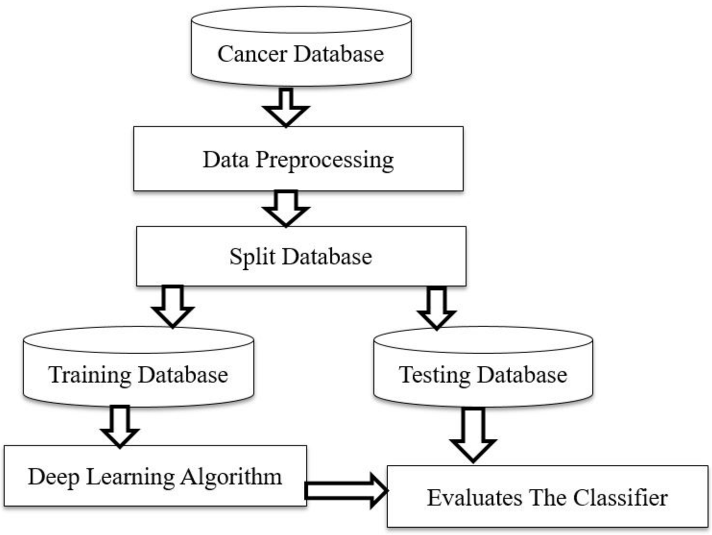

Figure 1.1: Methodology using Deep Learning Method

B. Transfer Learning: We are using the following steps to evaluate the classifier. Firstly, we bhave built deep convolutional neural network model to classify mammograms to benign and malignant variants. In addition to data augmentation we applied transfer learning technique to overcome insufficient data and training time. We have trained the CNN model with MIAS dataset and then we applied CBIS-DDSM dataset on the same model to check the results. This process is called transfer learning where a trained model with a set of dataset is applied with another dataset. Again, we have trained VGG16 (a method of CNN) model with MIAS dataset and applied with CBIS-DDSm dataset. By this process, we got good performance results which is shown in chapter-5.

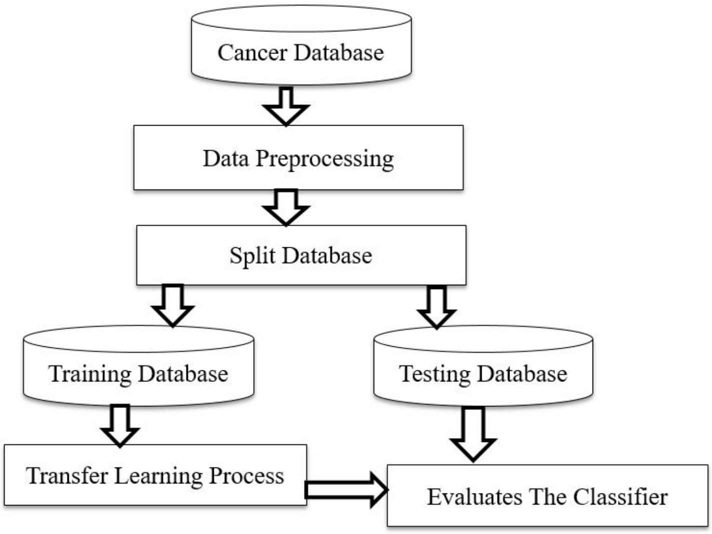

Figure 1.2: Methodology using Transfer Learning Method

1.5 Organization of the Report

The rest of the report are described following chapterwise. In the next chapter, we discussed about the background knowledge of our project and reviewed some other works related to this. In chapter-3, we shared all the specific details about how we designed our project. This includes the dataset we have used and all the other methods and models we have followed. The following chapter-4 is about the environment we used to develop our project included with the results gained from oue models. Chapter-5 is about the ethical standards we followed and in chapter-6, we described the summary of our project and the future plans related to it.

Chapter 2

Background

This chapter includes the background knowledge about our project. Many studies were done in the last few years to improve the detection procedure of breast cancer. Here, we have reviewed some of them and described how the studies were done with all the significant details. There were many good works done and we tried to gain some knowledge from those studies to make our project a better one.

2.1 Preliminaries

Different works suggested that histopathological images, 3D ultrasound images and mammography images are mostly used for classifying the variants. Most of the cases DDSM, MIAS and BCRP datasets are used to train and test. We decided to use MIAS dataset of mammography images.

Since 1990s, many applications of CNN can be identified based on medical images. An important feature of CNN which is "Transferability", placed in CNN pre-trained model. Earlier research presents that in the aspect of medical imaging, transfer learning can be divided between two different categories. Firstly, it needs to use a pre-trained model for feature extraction and to train a different classifier those features are needed. Secondly, a new specific layer is used instead of fully connected layers and the other layers of pre-trained model is used also.

In this dataset [10], for feature extraction many methods with classifiers are used like descriptive convolutional neural network, wavelet transforms, support vector machine, fusion of cosine transform etc. For feature extraction, different comparisons have been done. One is by comparing some medical images and the other is comparing different classifiers. For classification of features, SVM classifier is used among normal, benign and malignant variants with dense SIFT [10]. In this work, IRMA dataset is used and it had an accuracy of 81.83% for discrete wavelet transform and 83.74% for curvelet transform [10]. C- mean clustering method has been also used in some works. It proposed that with the help of genetic algorithm, a better result has come to detect and extract the affected region [11]. 3D breast ultrasound images were also used to detect breast cancer [12]. These 3D images are used to create differences in fatty and non-fatty tissues [12]. A work was done to detect breast cancer using breast thermograms [13]. Using K-means clustering, they categorized the hot region and analyzed the color segmentation of breast thermograms [13]. Another work has been done for classification using back propagation neural network (BPNN) classifier [14]. It had a good classification rate of 88.9% where BPNN worked as a microcalcification classifier [14]. A work proposed the extraction of dense region of a tumour based on the ratios of ultrasound images [15]. An edge-based algorithm is used to the contour search which is verified by operating a threshold operator [7]. In terms of receiver operating characteristic (ROC) curve, it gained about 89.2% classification rate on the test set [7]. Using an alternate dense area identification method a research has been done on 32 different medical image features [16]. A set of logistic regression models were used to examine the method [16].

2.2 Summary

This research introduces a possible solution to classify benign from malignant breast cancer in mammograms by an approach which contains a CNN and transfer learning based methods. This approach applies the concept of neural network which is very useful when there is a lot of data. We are trying to implement an application where these methods can be very useful to detect the benign and malignant variants of breast cancer. At first we tried to make this approach with a small amount of data. But then, we continued our program to to run the classifier with a large amount of data. The evaluation methods used to evaluate the performances of the CNN and transfer learning based approaches are confusion matrix, sensitivity, specificity and classification accuracy. Thus we accepted the challenge for implementing such an application where women can be helped by detection of the breast cancer as fast possible.

Study	Input	Purpose	Data set	Classifier	Results
Qiong et al.	Mammograms	Breast mass detection and mass diagnosis	400 cases of female mammograms	CNN, USELM	(i) 80.75% accuracy (ii) 80.48% sensitivity (iii) 82.43% accuracy in benign (iv) 78.97% accuracy in malignant
M. M. Mehdy et al.	Mammograms	Breast tumor detection	Mammography images set	ANN	Better accuracy, sensitivity and also positive predictive value
Neeraj et al.	Mammograms	Increasing the accuracy rate	(i) DDSM-BCRP (ii) INbreast	R-CNN	(i)True positive rate of 0.86 at 1.2 false positive per image on INbreast (ii) True positive rate of 0.75 at 4.8 false positive per image on DDSMBCRP
S.A. Taghanaki et al.	Mammograms	Achieves superior classification results	949 mammograms	Nondominated sorting genetic algorithm	98.45% accuracy
Fabio et al.	Histopathological Images	Improve the classification performance	WDBC	CNN	(i) Accuracy of 99.28% (ii) sensitivity of 98.65% (iii) Specificity of 99.57%

Table 2.1: Summary of Literature Review

Chapter 3

Project Design

This chapter is about all the specific details and informations of the models we designed to demonstrate our project. The datasets we have used, informations of the methods we applied with how they perform with the algorithms -these things are described in this chapter. In the last section, the performance measurement formulas are shown.

3.1 Brief Overview of the System

Our study includes several methods of CNN and ANN to predict the results. We applied different layers in the CNN and ANN to evaluate the performances and then compare the results with different parameters. And a computer-aided system is introduced to predict a single x-ray image's result where the system would tell with positive/negative outcome.

3.2 Description of the Datasets

We have used two datasets to train and predict results- MIAS and CBIS-DDSM.

Dataset	Size	Positive number of images	Negative number of images	Color of image
MIAS	128×128	161	161	gray
CBIS-DDSM	299×299	7825	49742	gray

Table 3.1: Datasets information

3.2.1 MIAS

MIAS consists of 322 images where every image size is 128×128 and all mammograms are in PGM and gray format. And all mammograms are pre-processed.All mammograms are divided into normal and cancer part.

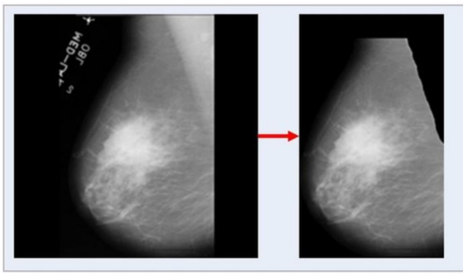

Figure 3.1: MIAS image

3.2.2 CBIS-DDSM

Curated Breast Imaging Subset of DDSM (CBIS-DDSM) is an improved version of DDSM (Digital Database for Screening Mammography). The dataset consists of 10239 mammography images and categorized in normal, benign and malignant classes. The images are converted to DICOM format. For our research, we have used 1400 mammogram images from the dataset [17].

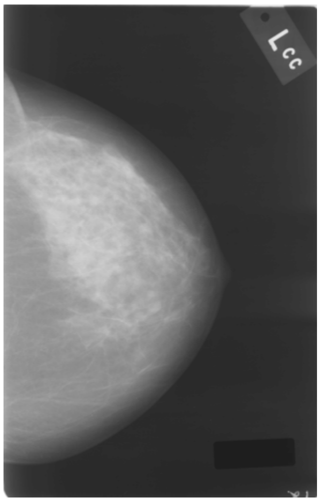

Figure 3.2: An image sample from CBIS-ddsm dataset.

3.3 Methods and Materials

In this section, we present the algorithms and models used and developed in our project.

3.3.1 ANN

Artificial Neural Network (ANN) is a multi-layer fully-connected neural network that includes an input layer, multiple hidden layers and an output layer. The below figure shows that every node in a layer is connected with all the other nodes in the next layer. The network can be made deeper by adding more hidden layers. We used only dense layer as

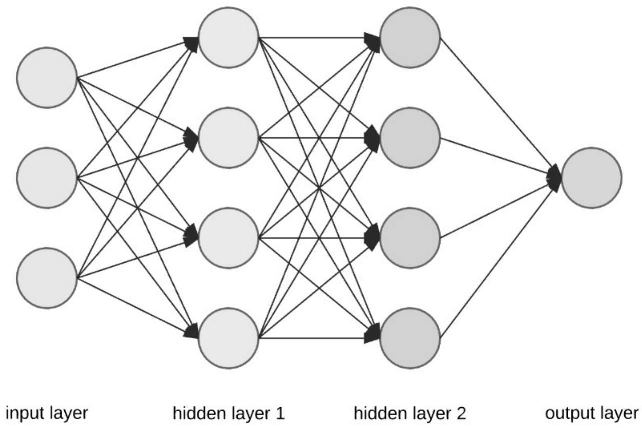

Figure 3.3: Basic ANN model

input and output. Here 256 input neurons are used and for input layer we used ReLU as activation function. For output layer only 2 neurons are used as the output should be normal or cancer and for output layer we used softmax as activation function. To evaluate the softmax function, the (n-1)th layer's output values are increased by some weights. An activation function passed the values and added up to a vector that contains a discrete value for every class. Softmax function derived every vector component in the interval of 0 and 1 which followed a probability distribution added up to 1.

The softmax function $\sigma : \mathbb{R}^K \to \mathbb{R}^K$ can be defined by the equation

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \text{ for } i = 1, \ldots, K \text{ and } \mathbf{z} = (z_1, \ldots, z_K) \in \mathbb{R}^K$$

3.3.2 CNN

A convolutional neural network (CNN) is a specific type of neural network that can take images as input, specifies the effect of characteristics in the images and be able to differenciate among them. Specially, high volumed and multi-challenged images(3D images) are processed by CNN that terminates the input values by some hidden layers to output classes. The neurons in a layer of CNN is connected by its previous layer. The most common layers in CNNs are convolutional layers, pooling layers, RElu, pooling layers alongside with input and output layers [18].

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix

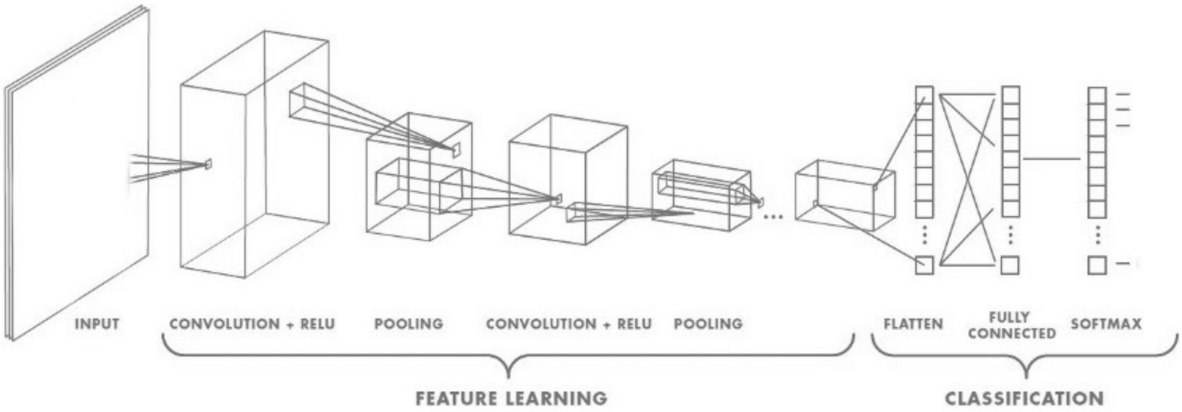

Figure 3.4: Basic CNN model

and a filter or kernel. If we consider a 5 x 5 whose image pixel values are 0, 1 and filter matrix 3 x 3 as shown in below. Then the convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is called "Feature Map".

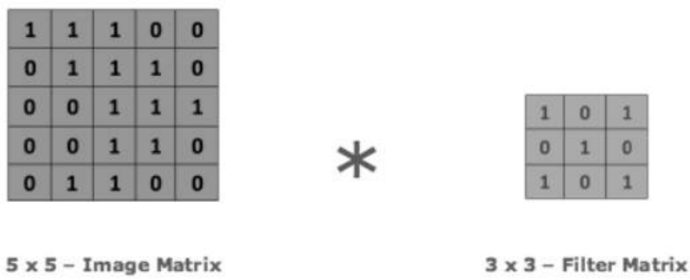

Figure 3.5: Convolutional layer in CNN

Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or downsampling which reduces the dimensionality of each map but retains important information. Spatial pooling can be of different types: Max pooling, average pooling, sum pooling.

Max pooling takes the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling. If we have a 4x4 matrix representing our initial input, then we have a 2x2 filter that we'll run over our input. We'll have a stride of 2 (meaning the (dx, dy) for stepping over our input will be (2, 2)) and won't overlap regions.For each of the regions represented by the filter, we would take the max of that region and create a new, output matrix where each element is the max of a region in the original input.

We used multiple layers of CNN. Four 2D convolutional layers and four 2D pooling layers used for extracting feature. 3*3 matrix was used in every convolutional layer and

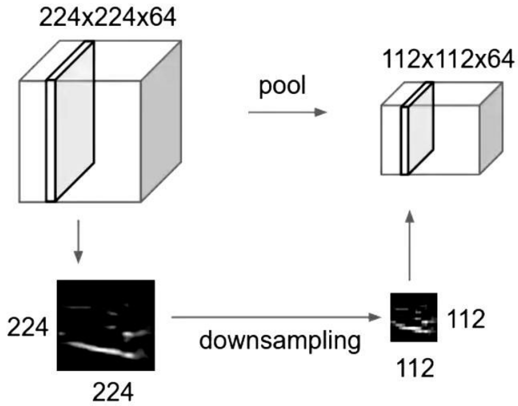

Figure 3.6: Max pooling layer

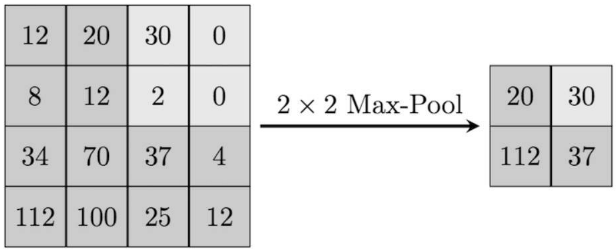

Figure 3.7: Real life max-pooling

2*2 matrix was used in every pooling layer. Again, 3 hidden layers are used and for the activation function Rectified Linear Unit (RelU) was applied. We used two fully connected layers where one activation function is ReLU and another one is softmax.

We used multiple layers of CNN. Four 2D convolutional layers and four 2D pooling layers used for extracting feature. 3*3 matrix was used in every convolutional layer and 2*2 matrix was used in every pooling layer. Again, 3 hidden layers are used and for the activation function Rectified Linear Unit (RelU) was applied. We used two fully connected layers where one activation function is ReLU and another one is softmax.

3.3.3 VGG16

VGG16 is a convolution neural net (CNN) architecture which was used to win ILSVR (Imagenet) competition in 2014. It is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG16 is that instead of having a large number of hyper-parameter they focused on having convolution layers of 3x3 filter with a stride 1 and always used same padding and maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC(fully connected layers) followed by a softmax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million (approx) parameters.

3.4 Performance Measurement

The most common evaluation methods in the field of medical imaging process were used for evaluating the performance of the proposed hybrid approach on the classification of benign from malignant cases. To specify the classifier performance of CNN and ANN models, widely used metrics are- AUC and ACC, SEN and SPE. Moreover accuracy, specificity and sensitivity are calculated based on confusion matrix using TP (true positive value), TN(true negative value), FP(false positive value) and FN(false negative value). The results using the models which are shown in table-2 are calculated according to these formulas:

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

Chapter 4

Implementation and Results

A brief description of results with analysis is discussed here in this chapter. To conduct this work, many tools were needed. There is a list of the tools that we have been using to execute our project. Performance measurement results are briefly discussed with all the methods included with graphs in the following sections. Also a comparison of the results between two datasets are shown with different parameters are delivered with informations in the tables in the sections.

4.1 Environment Setup

Ideally, this is a research project and so far we have been using different tools to implement our application. We have been using different tools such as-

1. Google Python 3 notebook for writing the codes.
2. www.overleaf.com for writing the thesis paper.
3. Microsoft PowerPoint Presentation to make presentation based on our project.
4. https://asana.com to manage our tasks and project management.

4.2 Implementation

In this study, we are following five steps for breast cancer detection: Image pre-processing, training data generate, feature extraction, model training and cancer detection. In Breast image pre-processing, denoising processes will be applied on the original mammogram images. Then for every pre-processed images the model would be trained using their extracting feature. After that, perfectly trained classifier would be able to classify the cancer and normal mammogram.

There are many approaches exist for denoising images but in this case we selected adaptive mean filter algorithm to reduce noise from images. The main concept is to calculate mean, variance and special correlation values for detecting, in which part has noise in an image. If the noise is detected, then the pixel value would be replaced with mean value.

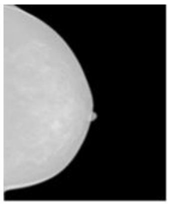

Figure 4.1: Normal mammographic image

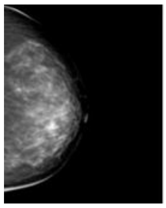

Figure 4.2: Denoising

For extracting features, CNN is being used to extract from mammography images. Here table(tablenummber of CNN) presents 6 layer CNN architecture where 3 convolutional layer, 2 max-pooling layer and 1 fully connected layer.The input size is 128*128 dimension.

4.3 Results

4.3.1 Classification with Multiple layers of CNN:

The ROC curve for the final performance using for this system is shown on Fig.5. From the curve, we can see true positive rate which is sensitivity and false positive rate using this model are respectively 0.6 and 0.3. Considering this model, the AUC rate was 0.57. This model achieved a good percentage of training accuracy and validation accuracy which show in Fig.4. Using this model, we got almost 94.8% accuracy and the sensitivity and specificity results are 66.67% and 33%. Fig.4 presenting the results of training accuracy,validation accuracy and training loss,validation loss.

4.3.2 Single hidden layer of ANN:

Using this model, we get training accuracy about 60.25%, sensitivity and specificity are about 66% and 33%. Fig-6 is representing the graph of training, validation accuracy, training and validation loss.

4.3.3 Training Accuracy and Training loss:

Training accuracy means measurement of performance after training a model by an algorithm. How perfectly the model is predicting is measured by comparing with true data. And training loss is opposite of training accuracy, that means how bad the model predicts

Algorithm	Epochs	Training Acc	Training Loss	Validation Acc	Validation Loss	Sn	Spc	AUC	MCC	F1
Multiple layer of CNN	30	94.8%	13%	47.37%	78%	66.67%	33%	0.57	0.73	0.46
Single hidden layer of ANN	30	60.25%	66.08%	71.43%	64.10%	66.67%	33%	0.66	0.73	0.46
Auto Encoder	40	51.47%	60%	71.43%	62.45%	75%	33%	0.625	0.25	0.6
encoding and sequential model combination	60	60.25%	22.26%	55.55%	24.85%	66.67%	47.72%	0.5	0.0	0.64

Table 4.1: A summary of CNN and ANN based method using MIAS dataset

to identify accurate data class. If the model can't be able to give accurate prediction loss with increase and if it can give perfect prediction loss will 0. Here, table-2 shows that using multi layer CNN, it gives an accuracy rate of 94.8% and training loss 13% and using single layer ANN, it gives training accuracy and loss about 62.5% and 66.08%.

4.3.4 Validation Accuracy and Validation loss:

Validation accuracy and validation loss is like training accuracy and training loss which explained in sec A but the difference is its accuracy and loss are based on validation set.Validation set is like not part of testing set, it can be also a part of training set. Validation set is actually used for tuning the parameters of a model. If validation accuracy percentage comes larger than training accuracy, it means the model is overfitting. Table-2 shows that using CNN validation accuracy and loss percentage comes about 47.37% and 78% and on the other side, using ANN the accuracy has come to 71.43% and loss about 64.10%.

4.3.5 Comparison among Results:

Table-3 shows that a training accuracy of 97.22% is gained from multiple layer of CNN using CBIS-DDSM dataset. Using same model of CNN, the training accuracy gets better

Algorithm	Epochs	Training Acc	Training Loss	Validation Acc	Validation Loss	Sn	Spc	AUC	MCC	F1
Multiple layer of CNN	20	97.22%	8.56%	81.63%	97.55%	83.33%	75%	0.53	0.073	0.78
Single hidden layer of ANN	20	55.25%	65.98%	47.13%	71.55%	6.52%	60%	0.508	0.035	0.11
Auto Encoder	10	91.58%	23.91%	69.00%	1.03%	50.70%	51%	0.464	0.023	0.50
Encoding and sequential model combination	10	39.67%	77.31%	47.30%	73.33%	45.08%	10%	0.5	0.34	0.48
VGG16	5	72.70%	50.34%	70.59%	55.34%	45.93%	47%	0.4584	0.083	0.45

Table 4.2: A summary of CNN, ANN, auto encoder, encoding and sequential model combination and VGG16 method using CBIS-DDSM dataset

than the result gained from MIAS dataset. The results of other models are shown on the table including Auto-Encoder and VGG16 (A new model of CNN) using CBIS-DDSM dataset. About 91.58% training accuracy and 69% validation accuracy came using Auto-Encoder and 72.7% training accurcay came using VGG16. If we check the results of table-2 and table-3, the difference of results using two different dataset is clearly understandable. The training loss of multile layer of CNN is only 8.56% whereas it was 13% using the MIAS dataset. Using single hidden layer of ANN, the training loss is quite similar for the both datasets. But the validation loss for the both models are higher using CBIS-DDSM dataset.

4.4 Discussion

For the last few years, many computer-aided diagnosis (CAD) system have been developed to determine breast cancer using mammogram images.This methods or techniques applied to help radiologist to analyse mammogram images which cant be remarkable using eye.Researchers are working to develop all techniques which can be use in mammograms to improve the percentage of accuracy.Many neural networks, like ANN,CNN proved as a good classifier for classification of masses from mammography images.As for classifying breast cancer allover use mammography images that's why neural network is the best option as a classifier.In image processing most of the time CNN and ANN used and it act like very good classifier.In this kind of models have too many layers that can easily filter all feature from a image and easily can make difference from other images. In our project we are working with neural network algorithms.As we are using two dataset, MIAs and CBIS-DDSM dataset.For each dataset we got different type of result for every algorithms.Some algorithm work good for MIAS dataset and some work good for CBIS-DDSM dataset.CNN work good for CBIS-DDSM dataset where training accuracy got 97.22% and validation accuracy come out 81.63% which is better than MIAS dataset.In our project we also use vgg16 which is a pretrained keras model and using this model we also got very good percentage of training and validation accuracy both.But from the table 4.1 and table 4.2 it can be say that CBIS-DDSM dataset work most better than MIAS dataset.

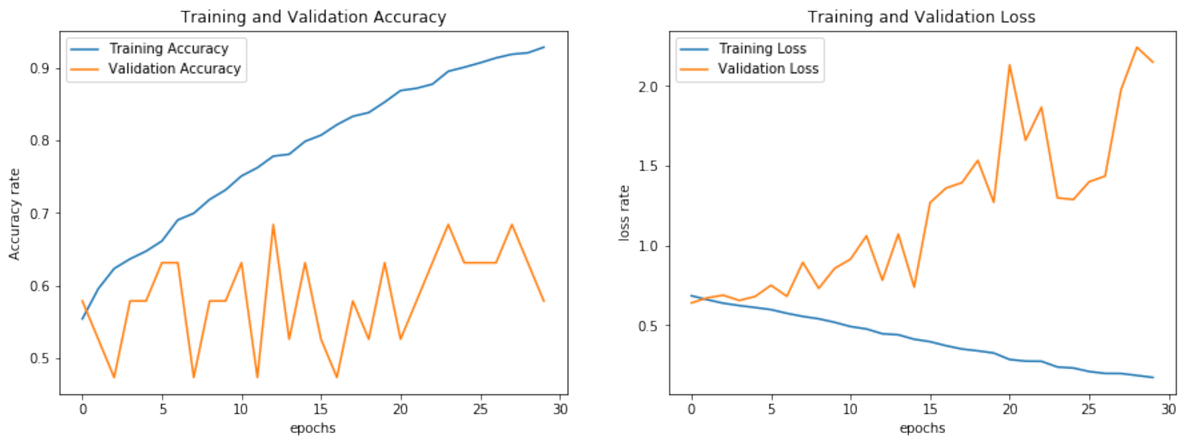

Figure 4.3: Training accuracy and Training loss using CNN(MIAS dataset)

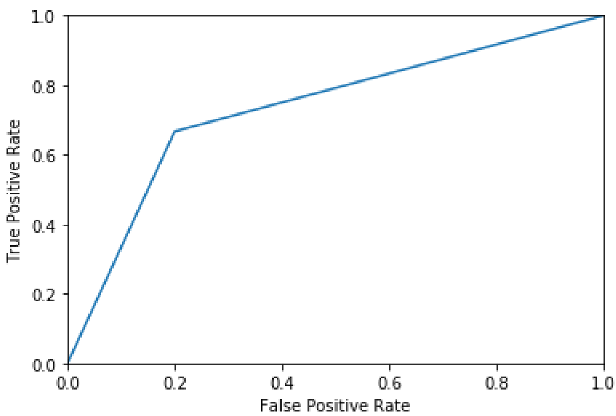

Figure 4.4: roc_curve using CNN

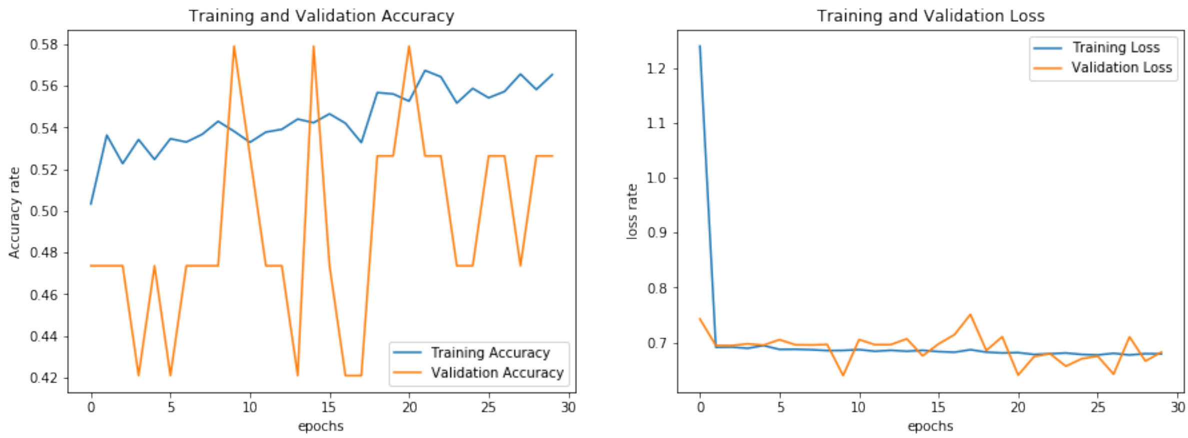

Figure 4.5: Training accuracy and training loss using ANN(MIAS dataset)

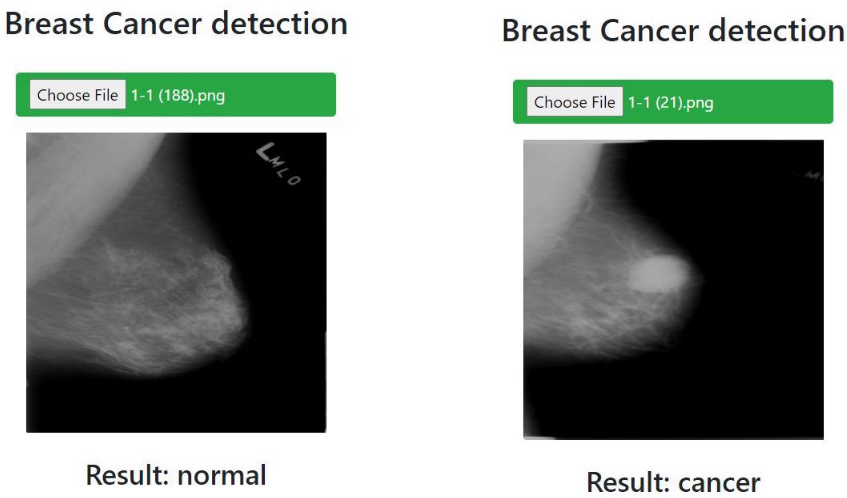

Figure 4.6: Image Classifier Web Application.

4.5 Web Application

We have also implemented a simple web application where a user can upload the image of any mamography data and the the web application can provide prediction based on the saved models. Screenshots of the model is given in Figure 4.6.

Chapter 5

Standards and Design Constraints

Always, a research work needs to be done with complete guidance and standards. There have been many standards and guidelines to be followed but particularly we have gone along with the ethical standards of IJAET journal. This chapter is all about the standards we have followed and the constraints we have faced to demonstrate our project.

5.1 Compliance with the Standards

In this section, we describe different standards followed in this project.

5.1.1 Ethical Standard

We have followed the IJAET journal for the ethical standards of our thesis work:

- Avoiding Plagiarism.
- Sources are cited but still plagiarized.
- To decide when to give credit.
- To make sure remaining safe when working.
- When to paraphrase and summarize.
- When to quote directly and indirectly.
- To decide if something is 'Common knowledge'.

5.1.2 Software Standard

We have created an application that can detect the benign and malignant variants of breast cancer. And we have followed ISO/IEC 9126 standards to complete this task which can be addressed as-

Functionality: We are trying to create an application where all the features will function differently with their perspectives.

Reliability: An application is reliable to the users when it can fill up the motive. We can rely on this software application where women can easily detect breast cancer and take necessary steps following that result.

Usability: This application is being made up very user friendly as non-technical people can also use that.

Efficiency: We are working through the relationship between the level of performance of the application and the amount of resources used, under stated conditions.

Maintainability: 'Analyzing' is also an attribute that relies on the effort needed to make specified modifications of our application. Maintaining the functions and the requirements is very important to continue a running application.

Portability: The application we are trying to make could be used from different environments. This application can be easily adopted by the people in medical environment and those who have a least knowledge of breast cancer

5.2 Design Constraints

Only mention the design constraints that are related to your project. This list is not complete.

5.2.1 Economic Constraint

Histopathological images, 3D ultrasound images are more costly than mammography screening test. So, it will be less-costly for mammography screening test and to use the mammography images as dataset.

5.2.2 Environmental Constraint

For collecting data we found some ethical constraints from the sources. The hospitals didn't allow to share their cancer databese for ethial issues.

5.2.3 Social Constraint

Many women in our society don't go to hospitals or clinics for asking help about breast related diseases to avoid embarassing situations. But they don't know that these diseases would affect them badly in future. So, we need to create awareness for this.

5.2.4 Manufacturability and Cost Analysis

An application will be made in a low cost such that people can easily use.

Chapter 6

Conclusion

So far what we have done in this project and how we have come to a conclusion to this work -a brief summary is described on the following section in this chapter. It took almost 8 months to conduct this work which is also a research topic. Also, the future plans of this thesis project is also mentioned in the following section.

6.1 Summary

In this paper, we have presented a CAD system for detection and classification of benign and malignant variants in digital mammograms. We have investigated and analyzed many feature extraction techniques, neural-network settings and transfer learning method. We have modified some traditional features and found that a combination of our modified features is a much better to distinguish a benign pattern from one that is malignant. This paper introduces some classification models of CNN and ANN which are used to classify the benign and malignant variants of breast cancer. A new network architecture VGG16 was followed in this work and the result came with a better percentage of accuracy. We have used Auto-Encoder to evaluate which is an approach of transfer learning. Implementing all these methods on MIAS and CBIS-DDSM datasets, the results are compared where multiple layer of CNN gave the most percentage of accuracy in two datasets.

Sensitivity, specificity, AUC and MCC - are the methods which were used to evaluate the results. We have been trying to create a dataset of our local breast cancer patients and the process is ongoing. In this study, we have trained a few models to classify and evaluate which can be a good opportunity to follow with in the future credentials.

6.2 Future Work

1. Improving the Classification Rate: There will always be a need to continue researching until a method is developed that classifies with 100% accuracy. Obviously, it is arguable whether this will ever eventuate. However, since the motivation to save human lives is inspiring researchers to develop accurate and efficient methods of detection and diagnosis,

research will and should continue well into the future.

2. Developing the CNN algorithm: We will try to develop our predicted algorithm for delivering a better performance than this.

References

[1] A. Jamal and F Bray. Global cancer statistics. *CA,CancerJ.Clinicians*, vol. 61(2):pp. 134, Mar. 2011.

[2] R.L. Seigal and K.D Miller. Cancer statistics. *CA,CancerJ.Clinicians*, vol. 69(1):pp. 7–34, Jan. 2019.

[3] M. M. Mehdy, P. Y. Ng, E. F. Shair, and N. I. Md Saleh. Artificial neural networks in image processing for early detection of breast cancer. *Computational and Mathematical Methods in Medicine*, Vol. 2017(Article ID 2610628):pp. 15, April. 3, 2017.

[4] ZHiquing Wang, Huaxia Wang, Hanyu Jiang, and Yudong Yao. Breast cancer detection using extreme learning machine based on feature fusion with cnn deep features. *IEEE Access*, vol. 7(1):pp. 105146–105158, Jan. 16, 2019.

[5] Neeraj Dhungel, Gustavo Carneiro, and Andrew P. Bradley. Automated mass detection in mammograms using cascaded deep learning and random forests. *Computational and Mathematical Methods in Medicine*, page pp. 8, jan. 2015.

[6] Neslihan Bayramoglu, juho kannala, and janne Heikkila. Deep learning for magnification independent breast cancer histopathology image classification. pages pp. 1–6, 2017.

[7] D. Cascio, F. Fauci, R. Magro, G. Raso, R. Bellotti, et al. Mammogram segmentation by contour searching and mass lesions classification with neural network. *IEEE Transactions on Nuclear Science*, vol. 53(5):pp. 2827–2833, Jan. 2006.

[8] Fabio A. Spanhol, Paulo R. Cavalin, Luiz S. Oliveira, and Caroline Petitjean. Deep features for breast cancer histopathological image classification. vol. 13:pp. 1–6, 2016.

[9] Liu Kui, Kang Guixia, Zhang Ningbo, and Hou Beibei. Breast cancer classification based on fully-connected layer first convolutional neural networks. *Digital Object Identifier,IEEE Access*, vol. 6:pp. 2169–3536, Feb 5, 2018.

[10] M Mohsin Jadoon, Qianni Zhang, Ihsan Ul Haq, Sharjeel Butt, and Adeel Jadoon. Three-class mammogram classification based on descriptive cnn features. *BioMed research international*, 2017:pp. 1–6, 2017.

[11] Shivali Sharma, Mansi Kharbanda, and Gautam Kaushal. Brain tumor and breast cancer detection using medical images. *International Journal of Engi-neering Technology Science and Research*, 2, 2015.

[12] Peng Gu, Won-Mean Lee, Marilyn A Roubidoux, Jie Yuan, Xueding Wang, and Paul L Carson. Automated 3d ultrasound image segmentation to aid breast cancer image interpretation. *Ultrasonics*, 65:51–58, 2016.

[13] Priya Hankare, Kesha Shah, Deepthy Nair, and Divya Nair. Breast cancer detection using thermography. *Int. Res. J. Eng. Technol*, 4(3):2395–2356, 2016.

[14] B. Verma and J. Zakos. A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. *IEEE transactions on information technology in biomedicine*, vol. 5(1):pp. 46–54, Jan. 2001.

[15] Woo Kyung Moon, I-Ling Chen, Jung Min Chang, Sung Ui Shin, Chung-Ming Lo, and Ruey-Feng Chang. The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound. *Ultrasonics*, 76:70–77, 2017.

[16] Fredrik Strand, Keith Humphreys, Abbas Cheddad, Sven Törnberg, Edward Azavedo, John Shepherd, Per Hall, and Kamila Czene. Novel mammographic image features differentiate between interval and screen-detected breast cancer: a case-case study. *Breast Cancer Research*, 18(1):100, 2016.

[17] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L Rubin. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific data*, 4:170177, 2017.

[18] Avijit Dasgupta and Sonam Singh. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 248–251. IEEE, 2017.