DEVELOPING A BREAST CANCER DISEASE DETECTION MODEL USING CNN APPROACH

A Thesis Presented

By

Mihret Gizaw Debebe

To

The Faculty of Informatics

Of

St. Mary's University

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In

Computer Science

June ,2022

ACCEPTANCE

Developing Breast Cancer Disease Detection Model Using CNN Approach

By

Mihret Gizaw Debebe

Accepted by the Faculty of Informatics, St. Mary's University, in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

Thesis Examination Committee:

____

Internal Examiner

____

External Examiner

____

Dean, Faculty of Informatics

June,2022

DECLARATION

I, the undersigned, declare that this thesis work is my original work, has not been presented for a degree in this or any other university, and all sources of materials used for the thesis work have been duly acknowledged.

Mihret Gizaw Debebe

Student

Signature

Addis Ababa

Ethiopia

This thesis has been submitted for examination with my approval as an advisor.

Million Meshesha (PhD)

Advisor

Million

Signature

Addis Ababa, Ethiopia

June, 2022

Acknowledgments

Foremost, I would like to express my deepest and sincere gratitude to my advisor Dr. Million Meshesha for the continuous support of my M.Sc. thesis, and for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me a lot in all parts of the thesis. Besides my advisor, I would like to thank every person who had contributed his/her part in any means to support me for the past two years. These individuals include Dr. Edome from Tikur Anbesa Hospital breast cancer section medical director, Dr. Adugaa, Dr. Ruth, Dr. Frehune, and Dr.Danel from Korea Hospital Radiologists and medical director as well as Ato Mesfine from Korea Hospital ICT Head. Finally, nothing is impossible without the support and motivation of my beloved family.

Thank You All.

Table of Contents

Abbreviations and Acronyms

APA----------------------	Application Portability Profile
API-----------------------	Application Programming Interface
BI-RADS----------------	Breast Imaging Reporting And Database System Score
BCE----------------------	Binary Cross-Entropy
CCE----------------------	Categorical Cross-Entropy
CPU----------------------	Central Processing Unit
CAD----------------------	Computer-Aided Diagnosis
CNN-----------------------	Convolutional Neural Network
CRNN---------------------	Convolutional Recurrent Neural Network
CNTK---------------------	Computational Network Toolkit
DCNN-------------------	Deep Convolutional Neural Network
DL-------------------------	Deep Learning
DNA-----------------------	Deoxyribonucleic Acid
FP---------------------------	False Positive
FN--------------------------	False Negative
FC--------------------------	Full Convolutional
GPU------------------------	Graphical Processing Unit
ICIAR----------------------	International Conference on Image Analysis & Recognition
IOS--------------------------	International Organization For Standardization
IEEE-------------------------	Institute Of Electrical And Electronics Engineer
MRI--------------------------	Magnetic Resonance Imaging
MSE-------------------------	Mean Squared Error

NN--------------------------- Neural Network

RAM------------------------ Random Access Memory

ROI-------------------------- Region Of Interest

RGB-------------------------- Red, Green, Blue

TP---------------------------- True Positive

TN---------------------------- True Negative

WHO------------------------- World Health Organization

List of figures

List of tables

Abstract

Cells that divide uncontrollably and spread into the surrounding tissues are what cause cancer. DNA changes are a cause of cancer. The majority of DNA alterations that lead to cancer occur in regions of DNA known as genes. One of the cancer diseases that is commonly recognized from a variety of angles as being quite diverse is breast cancer. It is among the main causes of death for females between the ages of 20 and 59 worldwide. According to the World Health Organization's (WHO) 2020 cancer country profile report, breast cancer has the highest age-standardized mortality rate of 22.9 per 100,000 people in Ethiopia, making it the most common cancer there. Early detection and care help patients receive adequate treatment and, as a result, reduce the risk of breast cancer morbidity. According to research, most experienced physicians can diagnose cancer with 79 percent accuracy, while machine learning techniques can achieve 91 percent accuracy. The main aim of this study is to develop a model that can assist a physician in detecting breast cancer and classifying it. Mammogram images were collected from the Korea hospital repository and used for developing a deep learning model. A pre-trained model such as VGG16, Inception, and SDDmobilenet are used as transfer learning for fine tuning. Also, there was a CNN model built from scratch with learning rate, batch size, and epoch and optimizer parameter optimization technique. The model built on InceptionV3 score the highest accuracy of 88% on training. The developed models have the capability of categorizing breast cancer. But the data is not sufficiently available for some classes. To solve the problem the researcher applied augmentation to overcome the problem of overfitting. Therefore collecting a large amount of data for all classes and developing a more reliable classification model is the future work of this thesis.

Key words: - deep-learning, breast cancer, convolutional neural network, detection, classification,multipleclassification

CHAPTER ONE

1. INTRODUCTION

1.1 Background

Cancer is a disease caused when cells divide uncontrollably and blowing out into surrounding tissues [1]. Cancer is caused by fluctuations in DNA. Most cancer-causing DNA fluctuations occur in sections of DNA called genes. And It is also a collective name for a large number of illnesses that manifest as malignant forms of abnormal cell proliferation in one or more body organs. The unexpected appearance of abnormal cells that grow beyond their usual boundaries, allowing them to infect nearby body components and spread to other organs, is one of the characteristics of cancer. In 2020(see figure 1.1), an estimated 19.3 million new cancer cases (18.1 million excluding non-melanoma skin cancer) were diagnosed worldwide, with almost 10.0 million cancer deaths (9.9 million excluding non-melanoma skin cancer) [2]. With an estimated 2.3 million new cases (11.7 percent), female breast cancer has exceeded lung cancer as the most frequently diagnosed cancer, followed by lung (11.4 percent), colorectal (10.0 percent), and prostate (7 percent) [3].

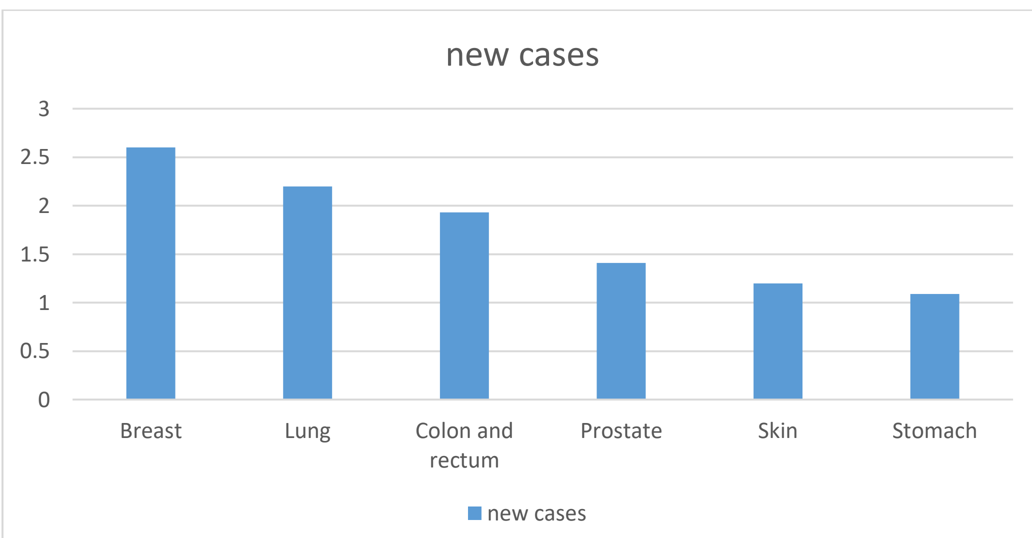

Figure 1.1 new cases of cancer in 2020 [4]

Breast cancer is one of the cancer diseases universally regarded as a highly heterogeneous disorder from a variety of perspectives [5]. Different forms of this neoplasm have different histopathological and biological characteristics, as well as different clinical outcomes and responses to systemic therapies. And also Breast cancer can be explained as a group of diseases that results from uncontrolled growth and changes in breast tissue typically resulting in a lump or mass. It is the second leading cause of cancer among females worldwide and a growing public health burden in low-resource settings. As of World Health Organization (WHO) 2020 cancer country profile report, breast cancer is leading cancer in Ethiopia with the highest age standardized mortality rate of 22.9 per 100,000 population [6]. As evidence shows an important determinant factor for breast cancer survival is the degree to which cancers are detected at early stages while advanced stage and large tumor size at diagnosis are associated with decreased survival and worse clinical outcomes.

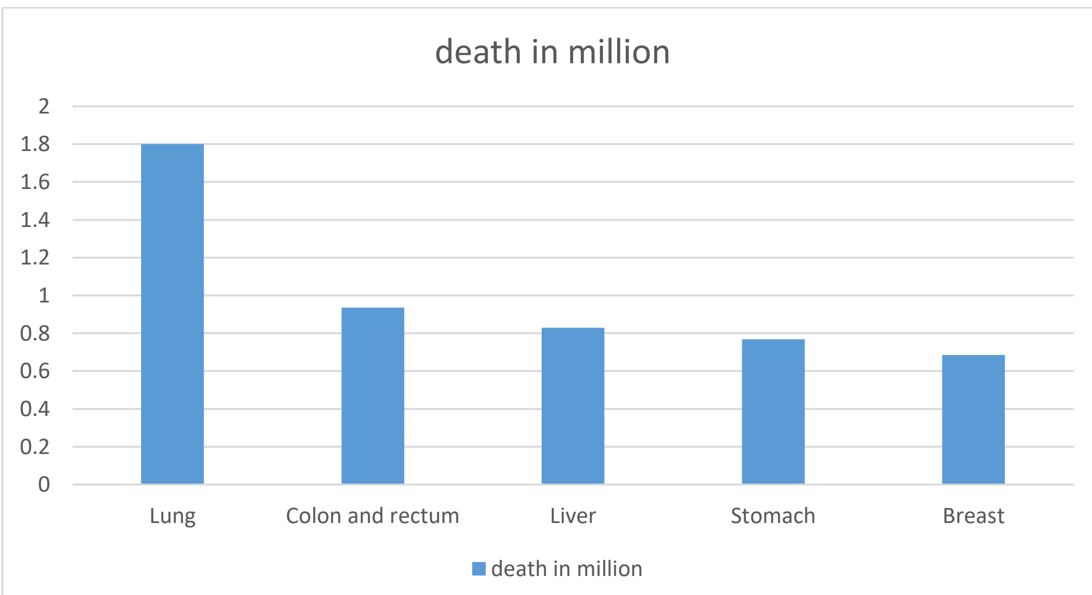

Figure 1.2 common cause of cancer in 2020 [4]

Deep learning is a relatively new discipline of data science research. It is a hybrid of artificial intelligence and machine learning techniques [7]. It is one of nature inspired techniques that is brain inspired. This has demonstrated higher flexibility and has resulted in the development of more neurons. Compared to machine learning, models are more accurate [8],.

Artificial intelligence (AI) is a technology that enables computers and systems to conduct operations or provide recommendations based on data that is extracted from digital images, videos, and other visual inputs [9]. If artificial intelligence allows computers to think, computer vision allows them to see, watch, and comprehend. Computer vision of deep learning has been applied in solving the problem of Image Classification, Image Classification with Localization, Object Detection, Object Segmentation, Image Style Transfer, Image Colorization, Image Reconstruction, Image Super-Resolution, Image Synthesis, and Other Problems [10]. This research emphasizes customizing the computer vision concept for detecting and classifying breast cancer disease using colposcopy images.

Deep learning models have become increasingly relevant in medical research in recent years [11]. It is particularly useful in the identification of cancer images because it outperforms competitors in tasks requiring massive and complex training datasets. A good disease classification should be scientifically valid, clinically useful, easy to apply, and widely reproducible. Unfortunately, amid all of the previous and recent attempts, there is still no ‘perfect' classification for breast cancer [12].

1.2 Motivation of the Study

One of the widely used breast imaging modalities is mammography. The imaging system does have some limitations, though. First, there is a lack of radiologists to analyze mammography images in hospitals across the nation. Second, it is difficult for today's radiologists to precisely read the mammography image. The Computer-Aided Diagnosis (CAD) system using Deep Learning (DL) plays a significant role in medical image analysis to solve the problem of mammography reading and interpretation. Convolutional Neural Networks (CNNs) are well-known and widely used deep learning (DL) models for the job of classifying images since they outperform other models in terms of accuracy and performance. It was created specifically for use with two-dimensional data, including photographs from a medical setting [12]. The CAD system has been used in a variety of deep learning-based single-label breast cancer screening studies. However, no additional multi-label classification for breast cancer screening was conducted. More specific information regarding a single picture, such as its density, discovery, and pathology, is provided via multi-label classification.to benefit from its advantage, this study was inspired to explore the earlier work to explore the multi-label classification. Many medical devices are used to diagnose a disease. Most devices need a high interference of physicians to interpret the image output from the device. So the model is needed for early detection to interpret the image from those devices.

1.3 Statement of the Problem

Breast cancer is one of the leading causes of death for women aged 20 to 59 around the world. In Ethiopia, the number of new cases of breast cancer is increasing, leading to high rates of morbidity and mortality [13]. Breast cancer is the most common cancer in women, accounting for one out of every three cases and one out of every five in the general population. There are many challenges to reducing the mortality rate of breast cancer in Ethiopia. In its nature breast cancer cannot be treated as a single clinic-pathological entity due to its high degree of heterogeneity [14]. Breast cancer is often discovered after symptoms occur, although often people with the disease have no signs or symptoms which leads women at high risk. To overcome such a challenge regular breast cancer screening is so important. There is a lack of physicians, Oncologists, and Pathologists in Ethiopia to diagnose, detect and classify breast cancer in less time and more accurately regularly. According to the Ministry of Health in Ethiopia[12], there are many hospitals, health centers, and health stations in Ethiopia but many of them have no physicians. It is reported that [11], overall, there is 1 physician for 57,876 people, but in southwest and west central Ethiopia, 1 physician serves between 200,000 and 300,000 people [15], [16].

Early detection and care help patients receive adequate treatment and, as a result, reduce the risk of breast cancer morbidity. According to research, most experienced physicians can diagnose cancer with 79 percent accuracy, while machine learning techniques can achieve 91 percent accuracy [17].This is why applying different type of Deep learning and machine learning technique can play a major role in detecting and classifying a breast cancer.

There are many online datasets that can be used in developing a deep learning model but they cannot have customized fully in Ethiopia or for the people who have a black skin. The machine learning model customized in Ethiopia showed that a people who have a black skin considered as a cancer infected. To overcome this problem,I prepare a dataset from medical images collected from hospitals in Ethiopia. So as To apply the most recent deep learning based breast cancer detection models. To solve such problem and support experts developing a model for detecting abreast cancer disease is needed.The rapid development of deep learning, a family of machine learning techniques, has spurred much interest in its application to medical imaging problems. This study aims to explore and develop a deep learning algorithm based breast cancer detection model. Using local image data set.

1.4 Research Questions

This study attempts to address the following research questions in the proposed solution.

RQ1. What is the current CNN approach to detecting breast cancer disease and classifying more than one disease in real time by low computational power devices?

RQ2. How to develop a CNN model for detection of breast cancer?

RQ3. What is the performance of the developed breast cancer detection in detecting disease.

1.5 Objective of the study

1.5.1 General Objective

The main objective of this study is to Developing Abreast cancer Disease Detection model using CNN Approach.

1.5.2 Specific Objectives

The following specific objectives are accomplished to achieve the general objective of the study.

- To conduct comprehensive literature review so as to identify methods, algorithms and approaches used in this study
- The data is labeled by their experts when collecting the mammogram images.
- To prepare training and test data set after collecting a breast mammogram.
- To identify suitable deep learning algorithms for supervised learning
- To develop an optimal model to detect and classify breast cancer.
- To test and evaluate the performance of the proposed model

1.6 Significance of the Study

Deep learning is one of the computer vision technologies that can solve the problem in health. Specifically supporting a physician or experts in Ethiopia. Early prediction and detection of a cancer is very important in boosting the diagnosis process of a cancer which can increase survival rate of life. The findings of this study are applied in the medical facility when screening mammography images of women to assess breast cancer risk. Due to the fact that women are the backbone of the family, finding and diagnosing breast cancer saves lives. When breast cancer is detected early on, therapy is simple and has a good probability of success. Therefore, it is advised that women begin screening at the age of 30 and once a year after that. The specific individual who benefited from this work includes the following:

- **Radiologists**: The study helps the radiologist to provide accurate on time and effective.
- **Woman**: Since the proposed model can confidently or effectively diagnose cancer status, it can be treated in a manageable form.
- **Families:** Given that women are the foundation of families, once cancer has been detected through the screening process, they can live comfortably without fear or properly adhere to treatment.

1.7 Scope and Limitation of the study

The study emphasized on developing deep learning based model for breast cancer disease detection. A model to be developing used in detecting and classifying a breast cancer which is malignant or not. A data set used for developing proposed model was prepared by collecting images collected from Korea hospital in Ethiopia starting from 2016 -2020. The images were collected from mammogram data. The classification of a cancer type depends on an image of breast mammogram. The image collected was resized to 417 *417 to effective training of model.

The research has limitation of covering the images of X-ray and Colposcopy. And also there are also a lack of category 2 and category 6 that causes unbalance data.

1.8 Structure of the Thesis

The remainder of the thesis is structured as follows. The next chapter deals with the literature review and summary of related works. The literature review explains a summarized report about breast cancer disease, computer vision, machine learning, and deep learning concepts in general. In the same chapter, under the summary of the related works, works done by different authors that are highly relevant and more related to breast and other cancer disease detection using a deep learning approach were included. Then, Chapter three discusses the research methodologies. Subsequently, Chapter four explains the experiment and its results. Additionally, Chapter five explains the results and discussions of the experiment. Finally, Chapter six summarizes the study using conclusion, contribution, and recommendations.

CHAPTER TWO

2. LITERATURE REVIEW AND RELATED WORKS

In this chapter literature review and related works are discussed which helps to explain what previous scholars have done, and also identify a research gap. Section 2.1 description of breast cancer is explained, Section 2.2 explain the types of breast cancer disease, Section 2.3 explain digital image processing and how it used for breast cancer detection, Section 2.4 describe breast disease detection using Convolutional Neural Network, and Section 2.5 lists related works done on the domain, and finally, a summary of the literature review was mentioned.

2.1. Breast Cancer Overview

A disease state that prevents cells from reacting to common stimuli is known as cancer. Tumors can be caused by abnormal cells that grow and spread uncontrollably [5]. Invasive tumors are those that have spread past the area where they first appeared and are now encroaching on nearby healthy tissues. Breast cancer is the most common type of cancer in women and is in the top 5 cancers that cause death, after lung, colon, liver, and stomach [5].

The exocrine glands, which include the breast, are an organ of the human body that can produce breast milk. Numerous tiny structures known as lobes and fat make up the majority of its structure. Numerous alveoli, or cavities, can be found in each lobe. My epithelial cells encircle these cavities and connect them with milk-releasing epithelial cells. My epithelial cells create ducts, which are tiny tubes where milk is extracted from the lobes. A network of pipes is connected. The larger pipe that results from the connection of these pipes ends at the nipple [18]. Breast mammograms are marked as density, mass, or calcification depending on the affected portion of the anatomy of the breast.

The Breast Imaging Reporting and Database System score is known as the BI-RADS score. It's a method of rating mammography results that radiologists employ. A mammogram is an X-ray imaging procedure used to assess the health of the breasts. It is the most effective method for assisting in the early detection of breast cancer. When medical professionals discover suspicious masses during a clinical breast exam, they might use it as a follow-up tool.

This test can aid in identifying anything odd even if it cannot medically diagnose breast cancer. It's important to remember that not all aberrant findings are cancerous [18].

2.1.1 Breast cancer category

The BI-RADS system is used by doctors to categorize abnormal findings. The range of classifications is 0 to 6. (see figure 2.1). Women who are 40 years of age and older frequently receive scores between 0 and 2, which denote normal results or indicate that abnormal results are benign, or noncancerous. Doctors and radiologists advise a follow-up appointment or a biopsy to determine the next step if you receive a score of 3 or above.

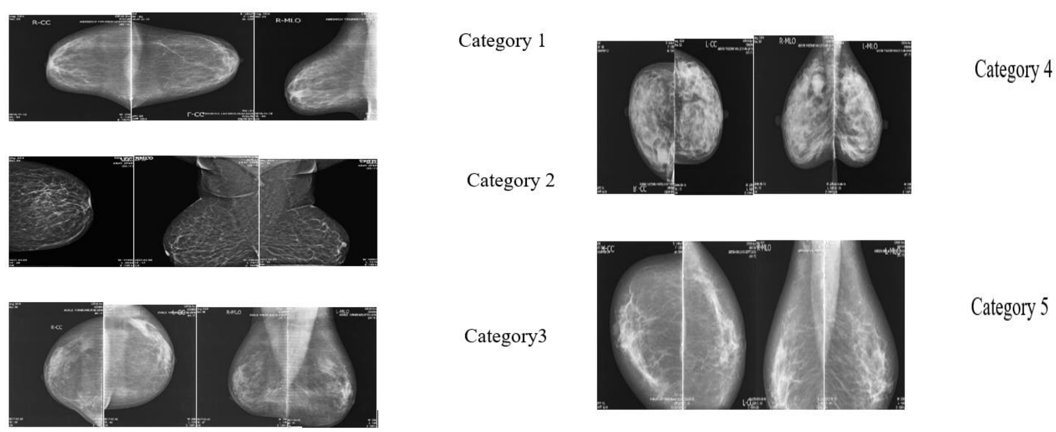

Figure 2.1 breast cancer category images

Category 0: - A test that has not been completed has a score of 0. It's possible that the mammography images were challenging to read or interpret. To see if there have been any changes, doctors may occasionally compare these new photos with earlier ones. A definitive evaluation cannot be made with a BI-RADS score of 0, as it necessitates additional tests and photos.

Category 1: - This result indicates with certainty that your mammography was negative. A score of 1 indicates that you have no cancer and equally dense breasts. Nevertheless, it's crucial to keep up with routine screenings.

Category 2: - Your mammography results are likewise normal if you receive a BI-RADS score of 2. Although there are no signs of cancer, the doctor might find any benign lumps or cysts that should be mentioned in your report. Regular visits are advised based on this score. Future discoveries will be compared to the note in your report.

Category 3: - A score of 3 denotes a 2 percent possibility of cancer despite the likelihood that your mammography results are likely normal. Doctors advise a follow-up visit within six months in this situation to confirm the findings are benign. Additionally, until your results improve and any abnormalities have normalized, you'll need to schedule routine appointments. Regular checkups aid in avoiding numerous, pointless biopsies. Additionally, they support cancer diagnoses that are made early.

Category 4: -A category 4 score denotes an anomaly or suspicious finding. There is a 20 to 35 percent likelihood of cancer in this situation. Your doctor must conduct a biopsy to examine a small tissue sample to make sure. Under the doctor's level of suspicion, this score is divided into three more groups.:

There are three levels of suspicion for cancer or malignancy: 4A, which indicates a low likelihood, 4B, which indicates a moderate likelihood, and 4C, which indicates a high likelihood.

Category 5: - A score of 5 denotes a strong suspicion of cancer. There is a minimum 95% likelihood of breast cancer in this situation. A biopsy is strongly advised to confirm the findings and establish the course of treatment.

Category 6:- Only after having undergone a biopsy and learning that you have breast cancer may you receive a score of 6. Using the accompanying photographs as a comparison, this category demonstrates how the cancer is reacting to important medical interventions like chemotherapy, surgery, or radiation.

2.2. Introducing image processing

Image processing is defined as manipulating images in various ways, in order to reduce distortion or noise, enhance the image and extract information from the image. Hence, image processing is used for improving the visual appearance of images to a human viewer and preparing images for measurement of the features and structures present.

Image processing techniques includes image acquisition, image enhancement, and image segmentation [19].

Levels of Digital Image Processing: To extract information, analyze a digital image, more importantly, to take full advantage of digital image processing, the processing is broken down into three levels namely: low, mid, and high-level processes [20].

Low-Level Processing: It entails simple procedures like sharpening, contrast improvement, image scaling, and image preprocessing to get rid of and minimize noise from images. The ultimate goal of this primitive operation is to improve the nature of the image to give better information. Both the input and the output of the low-level processing is an image [19], [20].

Mid-level Processing: includes tasks such as image segmentation, description of image, object recognition, etc. In these level, the input is a processed image and the final output is a feature or attribute (e.g., edges, contours, regions, etc.) which extracted from an image [20].

High-level processing: involves complex image processing tasks "making sense" from a group of the recognized objects. This process is normally associated with computer vision. In these levels of processing the input is attributes and the output is understanding of the digital image by extracting information from it [17], [20].

2.3. Steps followed in image processing

The majority of works involve the following key phases for the successful construction of an application using digital image processing. Each essential stage has a number of sub-steps and approaches. The stages shown in Figure 2.3 are thought of as the major ones; the actual work is carried out throughout the numerous substages that are described next. Each phase's drop-down steps link to its succeeding phases in the same way that an input step does.

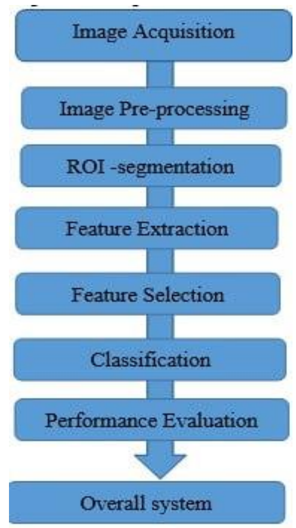

Figure 2.2 steps in image processing

2.3.1 Image Acquisition

The first phase in every image processing project is image acquisition, during which an original input image is obtained from the original source (where an image can be found) [21]. There are numerous techniques to implement image acquisition. This stage is frequently only thought of as taking a photograph of an actual environmental situation, but it can also involve gathering an image file that already exists from any electronic source using any technique of collection. When viewed from the corner of the targeted image processing program, the acquired image is completely unprocessed. The hardware that was used to create the initial image is what produced it, and having a consistent baseline from which to work can be quite crucial in some fields. The hardware equipment itself can range from a small desktop scanner to a huge optical telescope.

For the majority of applications, such as face recognition, object or animal identification, plant disease detection, raw images can be captured from environmental scenes using a digital camera, cell phone camera, or web camera. Depending on the application or type of process, raw images are acquired from various sources. For medical diagnosis purposes, such as tumor detection and cancer detection, images may be obtained through medical imaging equipment including X-Ray machines and MRI machines. For managing geographic information and related applications, images can also be obtained from satellites [22]. The research's data came from magnetic resonance imaging.

2.3.2 Image Preprocessing

The practice of enriching image data before computational processing is known as image preprocessing, and it is the second crucial stage. The preprocessing of an image may involve filtering, formatting, removing artifacts, and shade correction. The goal of filtering is to remove low-frequency background noise, normalize the intensity of the individual particle images, remove reflections, and mask portions of images to improve the stability of the image characteristics. Formatting deals with storage representation and setting the attributes of the image [23].

Although optical and analog image processing is also feasible, digital image processing is the most common type. The following details a few of the jobs involved in image pre-processing. Images can be shrunk to a smaller, lighter version of the original image for streamlined computation; this process is known as **image resizing** [24]. However, this is only advised when the scaling procedure has no impact on the final image's meaning or appearance relative to the original.

Image conversion One aspect of preprocessing is image conversion, which might involve transforming a picture between several formats or color spaces to facilitate computation [25]. For instance, as photos with fewer color ranges are simpler to analyze than those with more, such as RGB or other sorts, colorful photographs are often converted into grayscale or black and white images. But image conversion is not always necessary; it may be disregarded in circumstances when the outcome may change the meaning of a picture. For instance, compared to grayscale photographs, color images offer more detailed information. For the diagnosis of breast cancer illnesses, keeping the color image unconverted or reverting it back to color is crucial.

Image enhancement Another component of picture pre-processing called "image augmentation" uses various methods to give an image a deeper, more meaningful meaning. The most popular method for improving images is contrast arrangement. The sub-step of picture restoration deals with how to enhance the visual appeal of an image [26]. Picture restoration is objective as opposed to image enhancement, which is subjective. This is thus because methods for picture restoration frequently involve probabilistic or mathematical models of image degradation. Contrarily, image enhancement is dependent on subjective human choices for what makes a "good" enhancement outcome.

2.3.3 Image Segmentation

Image segmentation, also known as Region of Interest (ROI) Identification, is a term designated for any method that divides or groups an image into various useful sections or parts in the context of digital image processing. These areas often correlate to the portion of the image that humans can easily distinguish and perceive as discrete objects. The image's segmentation method is based on a variety of data. This could be a fragment of an image, color information, or boundary information. The dependability of feature extraction and the precision of detection is directly influenced by the segmentation result's precision.

Numerous image segmentation approaches are now in use in the fields of image processing and computer vision, each with its benefits and drawbacks depending on the issue at hand, or the image that needs to be segmented [26]. The two fundamental approaches to segmentation, region-based and edge-based approaches, can be used to find all of these techniques. Each technique can be used to achieve the necessary segmentation on various photos. These strategies can all be divided into three groups [27]. Next, a few of the more popular methods are discussed.

- Structural Segmentation Techniques: These are image segmentation methods that rely on the understanding of the structure of the needed region to be segmented within the image [20], [28].

- Stochastic Segmentation Techniques: These image segmentation methods use discrete pixel values from the image rather than the structural information of a region to segment the image [28].

- Hybrid Techniques: These image segmentation methods combine the concepts of the two techniques mentioned above, using both discrete pixels and structure information [29].

The threes holding method, edge-based detection techniques, region-based techniques, clustering-based techniques, watershed-based techniques, and artificial neural network-based approaches are some of the common picture segmentation techniques. Regarding the approach taken for segmentation and all of these strategies differ from one another.

2.4. Breast Image Analysis

Utilizing the volume of data and exploring and displaying the information effectively for particular medical tasks requires the medical image. As a result of the development of DL algorithms, this field is receiving a lot of attention and has improved accuracy. In radiology, the well-known medical image analysis is used to spot anomaly kinds and severity levels in X-ray images and determine whether they are normal or affected. Although it has a lot of false positives, CAD systems were developed and implemented in the healthcare system in the 1980s [30]. The sole goal of medical image processing, in contrast to conventional image processing, which has one main goal (such as improving image aesthetics or creative art), is to improve the readability of the generated content. This could entail improving the image itself to improve the perception of specific features and automatically extract information [30]. Making an accurate differential diagnosis using the medical images of each patient is one of the radiologists' most critical responsibilities. This can include determining the type of malignancy to evaluating the presence or absence of disease [31]. For doctors, examining the medical image is a difficult and time-consuming task. The CAD system has been created to evaluate the medical image since 1980[30] to lessen the challenge and free up the physician's time for other tasks. However, the CAD system that was in place at the time produced more false positives than doctors, which prolonged assessment times and necessitated unnecessary biopsies [31]. The use of medical images to enhance image capture, disease detection, therapy, and prediction has evolved and improved as a result of advancements in computer vision. It may create 3D and 4D data that facilitates human comprehension using texture, shape, contour, previous examples, and contextual information from a picture sequence. Diagnostic and therapeutic outcomes were considerably improved by the automatic processing of 3D medical images. New research questions in the areas of computer vision, graphics, and robotics are generated by this automation. To learn a well-executed hypothesis, a big number of diagnosed samples must be located.

2.5. Deep learning

A subset of machine learning techniques called deep learning seeks to extract numerous properties from distributed representations. The distribution of data with many levels of abstraction can be learned by computational models that stack multiple processing layers. Additionally, it is making significant strides in providing solutions to issues that the artificial intelligence community faces [32].This field is revolutionary because it’s computationally feasible and achieved good results in areas like image recognition, voice recognition, and other tasks that require a complex operation which has quite large data. Each output layer in deep learning is used as an input for the next layer. The learning process of deep learning or machine learning can be unsupervised, supervised, and semi-supervised based on the nature of a problem. The deep learning area of study began a few years ago and successfully achieved compared to some other areas of study, but the field was not explored further to solve our community's problem[9].

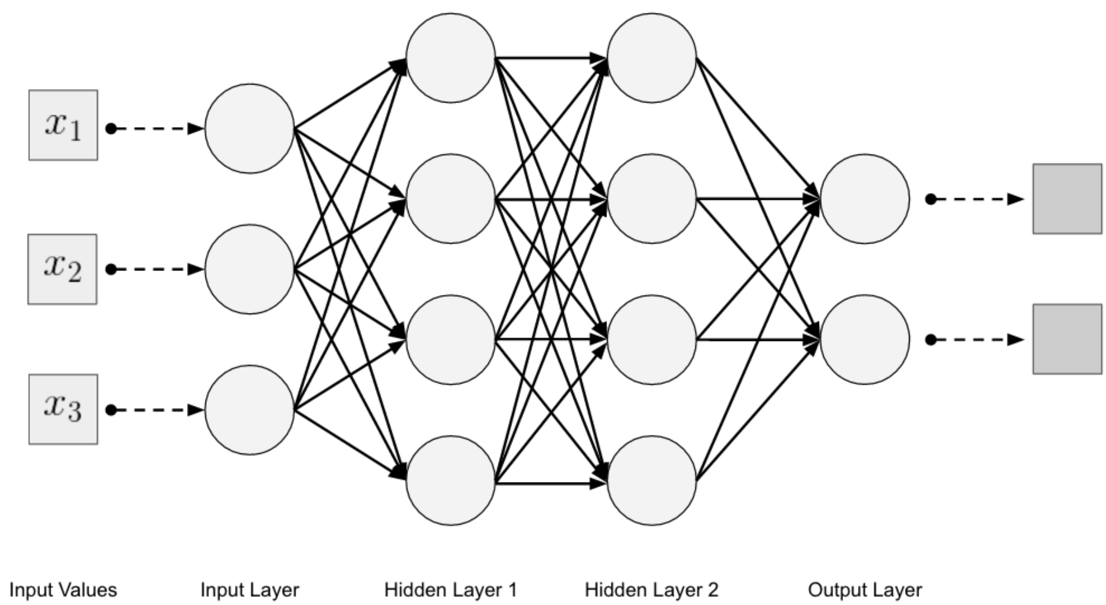

Figure 2.3 Deep Learning architecture

The two halves of the CNN architecture are feature extraction and classification. Feature extraction refers to the process of using a convolution tool, pooling, or stride to separate and identify different aspects of an image, such as patterns, lines, and edges, for analysis.

Contrarily, classification uses a fully connected NN to process the results of the feature extraction and forecast the image's class [33].

Christopher Thomas mentioned that deep CNN has shown an excellent performance in computer vision and machine learning problem [34]. The major critical reason to use deep learning's CNN is largely due to the use of multiple feature extraction stages that can automatically learn a representation of data form to a given input dat. Factors like, availability of a huge amount of data, and improvement of hardware technologies have improved the research of CNN. Besides, in recent years several interesting CNN architectures was reported, and that architecture used different loss and activation functions, parameter optimization, architectural innovation, etc. But the major improvement of CNN came from architectural innovation. Deep learning is a representation learning method. Learning representation model algorithms make an optimization to find out the most convenient way to represent data. One of the characteristics of deep learning is not dividing the feature extraction and classification as a separate task, because the model about to build automatically learn the features during the phase of training a model [32].

2.5.1. ComparingCNN vs. Classical Classification methods

CNN is different from other classical classification methods. Khan mentioned that most random studies used handcrafted feature extraction methods like texture analysis, followed by random forest and support vector machine [8]. The difference between the above methods and CNN are the following.

- CNN does not require an expert-based feature extraction.
- The CNN architecture does not require the segmentation of features by human experts.
- Needs a lot of data because of its millions of learnable parameters to estimate
- More computationally expensive requires hardware like GPU (Graphical Processing Unit) to effectively trains a model.

A common neural network used to handle issues with image classification, object detection, face recognition, image recognition, and other issues is the convolutional neural network (ConvNets). There are numerous applications for the convolutional neural network, some of which include speech recognition, natural language processing, image processing, image restoration, and bioinformatics. CNN can study things at a high level and can readily separate a feature with multiple layers since it requires less pre-processing to detect and classify objects [35].CNN has four layers: convolutional layer, pooling layer, activation function, and fully connected layer.

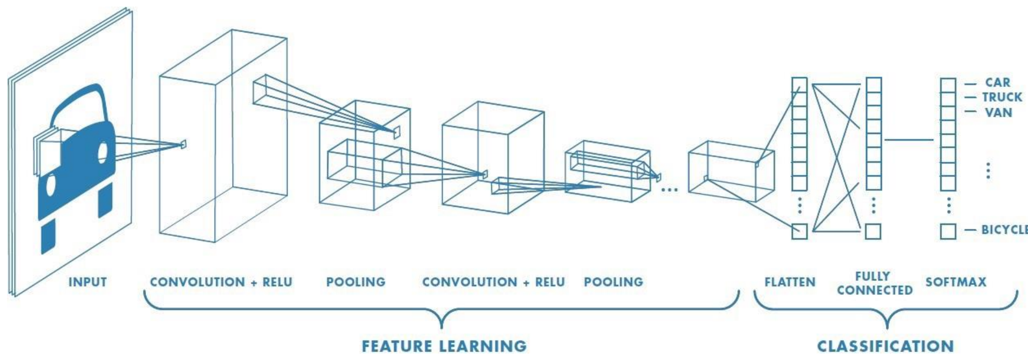

Figure 2.4 Example of CNN Layers

Convolutional Layer the CNN algorithm took its name from the convolution layer. In this layer series of mathematical operations i.e. matrix operation performed to extract a feature map from the input image [36]. The mathematical operation is depicted in Figure 2.5.1. First, the filter is shifted step by step starting from the upper left corner of the image. At each step, the values in the image are multiplied by the values of the filter (kernel) and the result is summed. A new matrix with a smaller size is created from the input image.

Pooling Layer: -the convolutional layer operation is followed by the second layer called the pooling layer. The pooling layer provides a typical downsampling operation which reduces the in-plane dimensionality of feature maps to introduce a translation invariance to a small shift and distortion and decrease the number of subsequent learnable parameters [37]. After an operation is performed in the convolution layer then a feature map is generated as an output. The output of the Conv layer's feature map is reduced in the pooling layer. Different filter sizes are used in the pooling layer, but usually, a 2x2 size filter is used, and also several functions are used to list some max pooling, average pooling, and sum pooling are some of the functions used.

Activation Function: this is one of the parameters of CNN architecture, in deep learning, many complex problems like image classification, object recognition, object detection, and others use a non-linear activation function with the neural network [38]. The task of the function is to decide whether a function is to be fired or not by calculating the weighted sum and adding bias. There are many types of activation functions and selecting the right one to enhance the learning process of the model. The activation function in a neural network solves a different complex problem like image detection.

Fully Connected Layer (Dense layer): this layer is an essential component of CNN that recognizes and classifies an image in a computer vision successfully. Once the input image is fed to the CNN algorithm, and after it passes the convolution and pooling layer the image breakdown into a feature and is analyzed independently. FC layer flattens the output of the pooling layer to classify an image [39].

Overfitting A statistical model is said to be overfitted when a model trains with data and starts to learn noisy and inaccurate data entries from the dataset. Overfitting is a critical challenge in deep learning and machine learning because the overfitted model is not capable of generalizing new data during testing the model [40]. Therefore, a test set is critical for the performance measurement of a deep learning model. There are several methods proposed to minimize overfitting and those are listed in Table 2.1.

Table 2.1 Mitigation method of overfitting

How to mitigate overfitting	
1	More training data
2	Data Augmentation
3	Regularization(weight decay, dropout)
4	Batch Normalization
5	Reduce architecture Complexity

2.6. Transfer Learning

Transfer learning, in which a neural network is taught for a certain task with data and the knowledge acquired would be transferred to another task, is one of the most potent concepts in deep learning. Training a deep learning model based on the problem complexity and dataset size it might take hours or weeks. And one of the methods that help to overcome this problem is to reuse the model's weight from a pre-trained model that was developed for standard computer vision benchmark datasets, such as Image Net image recognition tasks. Top-performing models can be downloaded and used based on the problem. In deep learning, transfer learning is a method whereby a neural network model is trained on a problem similar to the problem at hand. And then over one or more layers from the training model going to be used for a new model. Some of the advantages of using transfer learning are: Transfer learning uses pre-trained models like VGG16, InceptionV3, ResNet, and others [39]. These models were trained with millions of images with thousands of classes and with a high computation machine. Therefore, the problem of having little data is solved by transfer learning, which requires less computation power and training time. The next section explains some of the pre-trained models studied in this research, which are VGG (VGG16, VGG19), and InceptionV3.

Alex Net

Alex Net is a deep convolutional neural network model trained on 1.2 million images under 1000 classes from the ILSVRC dataset. It was the winner of the competition in 2012. The architecture has about 650,000 neurons and 60 million parameters, and the components of the model are arranged as 5 convolution layers, 2 normalization layers, 3 max-pooling layers, 3 fully connected layers, and Softmax in the output. Dropout regularization was applied to decrease the overfitting problem and in each convolution layer, the Relu activation function was used [41].

VGG19

VGG19 is another type of pretrained model of a deep convolutional neural network developed by K. Simonyan and A. Zisserman. The model is trained for the ILSVR competition and has more than 15 million tagged high-resolution images. To build the 20 models the dataset partitioned into 1.3 million images for the training set, 50,000 images for the validation set, and 100,000 images for the testing set were used. The largest VGGnet model has 144 million parameters from 16 convolution layers with a kernel size of 3x3, five max pooling with a size of 2x2, 3 fully connected layers, Softmax activation function in the output layer, output regularization in the fully connected layers, and Relu activation function in the convolution lay [42].

GoogleNet

GoogleNet Is a pre-trained model that won the 2014 ILSVRC competition also known as InceptionV3, The objective of the GoogleNet architecture is to reduce computational cost. The design of GoogleNet has increased the width and depth of the network while decreasing the computational cost [43].

2.7. Related works

Farhadi et al. [41] proposed an accurate and efficient deep transfer learning method to handle the imbalanced data problem which is prominent in breast cancer data. It describes that the underlying theory of most machine learning classification algorithms assumes that the class distributions in the data are balanced. However, most real-world application data are imbalanced which causes a series of performance bottlenecks for machine learning algorithms. The study focused on structured data which available publicly to generate the pre-trained model. The central hypothesis of the study was deep transfer learning-enabled solutions based on structured breast cancer datasets can improve the early detection and classifications of malignant breast cancer. Finally, based on the experimental results the researcher’s concluded that the proposed deep transfer learning on structured data can be used as an efficient method to handle imbalanced class problems in clinical research [44].

Kausani et al. [42]stated that an early and accurate diagnosis of breast cancer may significantly increase the survival rate of patients. The authors aimed to develop a fully automatic, deep-learning-based model. To achieve the aim of the study, image process technology was applied to stained histological breast cancer tissue images that were pre-processed using the stain normalization technique in the first phase. Next to pre-processing, data augmentation procedures are performed to minimize the problem of the limited size of datasets and lastly, high-level features are extracted from pre-processed images by using the proposed DCNN model. to be an input of standard multi-layer perceptron classifier. Inception v3, Inception ResNetv2, Xception, and Two VGGNet models are DCNN architectures employed as feature extractors. The experiment has carried out by 400 Breast Cancer Histology (BACH) images obtained from ICIAR 2018 Grand challenge. Finally, the proposed architecture achieves 92.5% accuracy which is better than other models [45].

A study conducted by Kim, Park, and Hong [43]stated that early detection and accurate diagnosis of breast cancer are crucial for reducing the associated death rate. This article proposed an edge extraction algorithm and a modified convolutional recurrent neural network (CRNN) model to accurately assess breast cancer based on medical imaging. The proposed Elfa-CRNN model analyses the line segments of the mass which is called line feature analysis. 250 ultrasonic breast image datasets from the Mendeley database were used. The Elfa proposed algorithm, canny and Sobel was the algorithm that delivered the highest accuracy of 98%, and the proposed Elfa –CRNN model achieve the highest accuracy of 99.75% when compared to CRNN, AlexNet, and VGG models [46].

Lu, Loh, and Huang [44] discuss that the incidence rate of breast cancer continued to rise and also stated that the screening strategy of the breast has been based on classic X-ray Imaging. Median filter, contrast-limited adaptive histogram equalization, and data augmentation were used for preprocessing 9000 mammograms. A convolutional neural network was used as a classifier by trained data. The study concluded that the accuracy of the model with preprocessed images significantly outperformed the model without preprocessed images [47].

According to Anupama, Sowmya, and Soman [45], One of the most serious types of cancer that can affect women is breast cancer. This study used histology images to categorize various forms of breast cancer. Breast cancer image from BACH 2018 grand challenge dataset was used as an input for training and testing the capsule network model. The research also recommends that the performance of the convolutional architectures can be improved by data preprocessing and transfer learning (parameter tuning) [48].

A summary of the related works done for breast cancer detection, classification, and diagnosis is presented below in table 2.2

Table2.2: Summary of Related Work

Author Name& year	Problem studied	Type & size of the dataset	Techniques & algorithms	Performance result
Farhadi et al. [45](2019)	Detection and Classification of breast cancer	357 benign and 212 malignant cases image from WDBC	DeepLearning CNN(Image Net architecture))	Best performance more than other machine learning techniques with 90%
Kassani et al [46](2019)	Diagnosis	400 BACH Image from ICIAR	Stain normalization Deep learning(DCNN model)	Accuracy- 92.5
Kim, Park, and Hong [47],2020	Classification and detection	250 ultrasonic breast image	Edge extraction algorithm, canny, and Sobel	Canny and sobel-98% and eLFA-CRNN-99.75% of accuracy when compared to CRNN, AlexNet, and VGG
Lu, Loh, and Huang [48],2019	Detection and classification	9000 mammograms image	CNN- for train	The accuracy of pre-processed model > non-pre-processed model
Anupama, Sowmya, and Soman [49],2019	Classification	Histology images from BACH 2018	Deep Learning(CNN) VGG16, Capsule Network InceptionV3	Capsule network-87% Inception ResNet-86% CNN-77.8% Deep CNN-87.2%

CHAPTER THREE

3. METHODOLOGY

3.1 Overview

This chapter focuses on explaining where the image data was collected, and how the data was prepared and analyzed for developing a classification model using the selected modeling techniques. Finally, evaluation metrics are presented that are used for evaluating the performance of the developed model

3.2 Research design

This study follows an experimental research method, to achieve the objective of this thesis. A study that follows a scientific research strategy is known as experimental research. The goal of experimental research is to find a link between two variables: the dependent and independent variables. A correlation between a specific property of an entity and the variable being researched is either supported or rejected when an experimental research study is completed. so in this thesis, the research is conducted to determine the availability of disease or not on mammogram images.

To do this, the process flow that is shown in Figure 3.1 is used, which comprises three primary steps. In the first stage, the problem's domain is identified to better comprehend it through a review of various types of literature. Then, the thesis's objectives general and specific are established. Data preparation is covered in phase 1 of the thesis's design, while phase 2 is dedicated to creating a research prototype. The Korea Hospital in Ethiopia provided the data for the creation of the data, which was subsequently labeled by medical professionals and divided into training, validation, and testing subsets. The model is designed after the data are prepared. The third step focuses on putting the thesis into practice. During this phase, the designed model is put into use using the proper tools and techniques. With the right data, the designed model is trained and tested. The model's performance is assessed while it is being trained. The best model is found during evaluation, and then it is evaluated using test data. The model is then contrasted with other previously trained models.

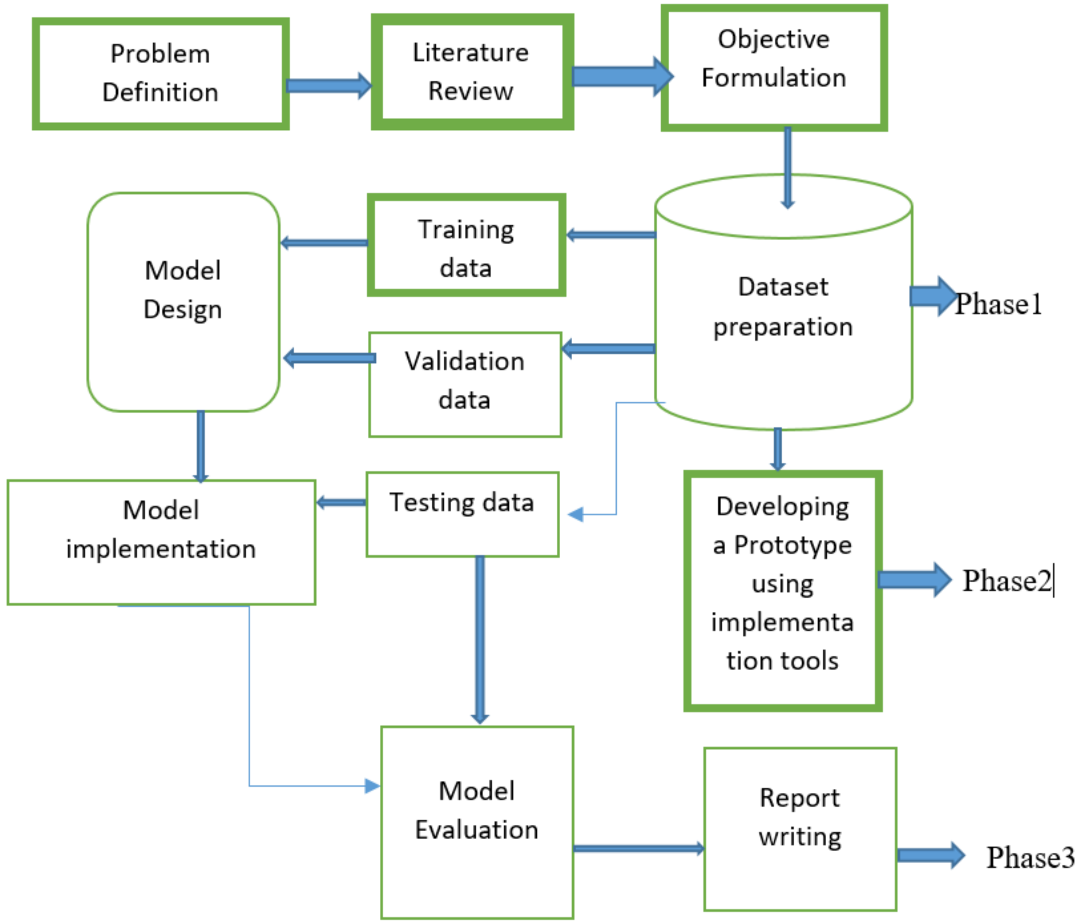

Figure 3.1 Rsearch Flow Model

3.3. Data Preparation

Obtaining the data necessary to train the neural network model is crucial if we wish to use neural networks or deep learning algorithms in our research. Breast cancer mammography imaging data are the primary input for the model in this thesis. However, there isn't a publicly accessible database with thousands of breast cancer photos of Ethiopian women that we could obtain and use to train the algorithm. So, we collected mammogram images of breast cancer from hospitals in Ethiopia with the help of Physicians and medical experts. All diseased and healthy images are collected from a Korean hospital.

3.3.1. Image Data Preprocessing and analysis

Data pre-processing is the process of preparing an image and making it suitable for a deep learning model. The MRI image from health sectors was processed before training a model for removing noise, getting high-quality image resolution, and enhancing the computational performance of a model. Before being fed into the neural network or deep learning algorithm, the raw data is also transformed. However, CNN (as mentioned in Section 2.5) eliminates the necessity for explicit preprocessing on the dataset by learning the features from the image's raw pixels. However, the created dataset's photos have various sizes. Because the model is trained on a conventional PC with constrained hardware resources, such as processor and memory, size normalization is done in the dataset to acquire similar size of all images for the CNN algorithm and to shorten the computational time of training. Finally, the dataset's photos are all downsized to 417 417 pixels.

3.3.2. Data Partitioning

First off, the dataset is split into two sections: training and test. The model is trained using the training split, and the model is tested using the test split, which is not used during model training. The validation split is used to evaluate the performance of the model created during training and to adjust model parameters in order to choose the model that performs the best overall. 2493 unique pictures were taken at a hospital in Korea. According to the literature [49], the training split should range from 60% to 90% of the total dataset, with the remaining 20% being used for testing. In this thesis, experiments are carried out with a 9:1 ratio.The training dataset is used to train the model by iteratively providing into the network using forward propagation and learnable parameters are updated into the weights of the model via back propagation. Due to the fact that updating of weights during training would not be biased in favor of one of the categories, using an equal number of images in each class for validation and training helps to prevent the issue of over fitting.

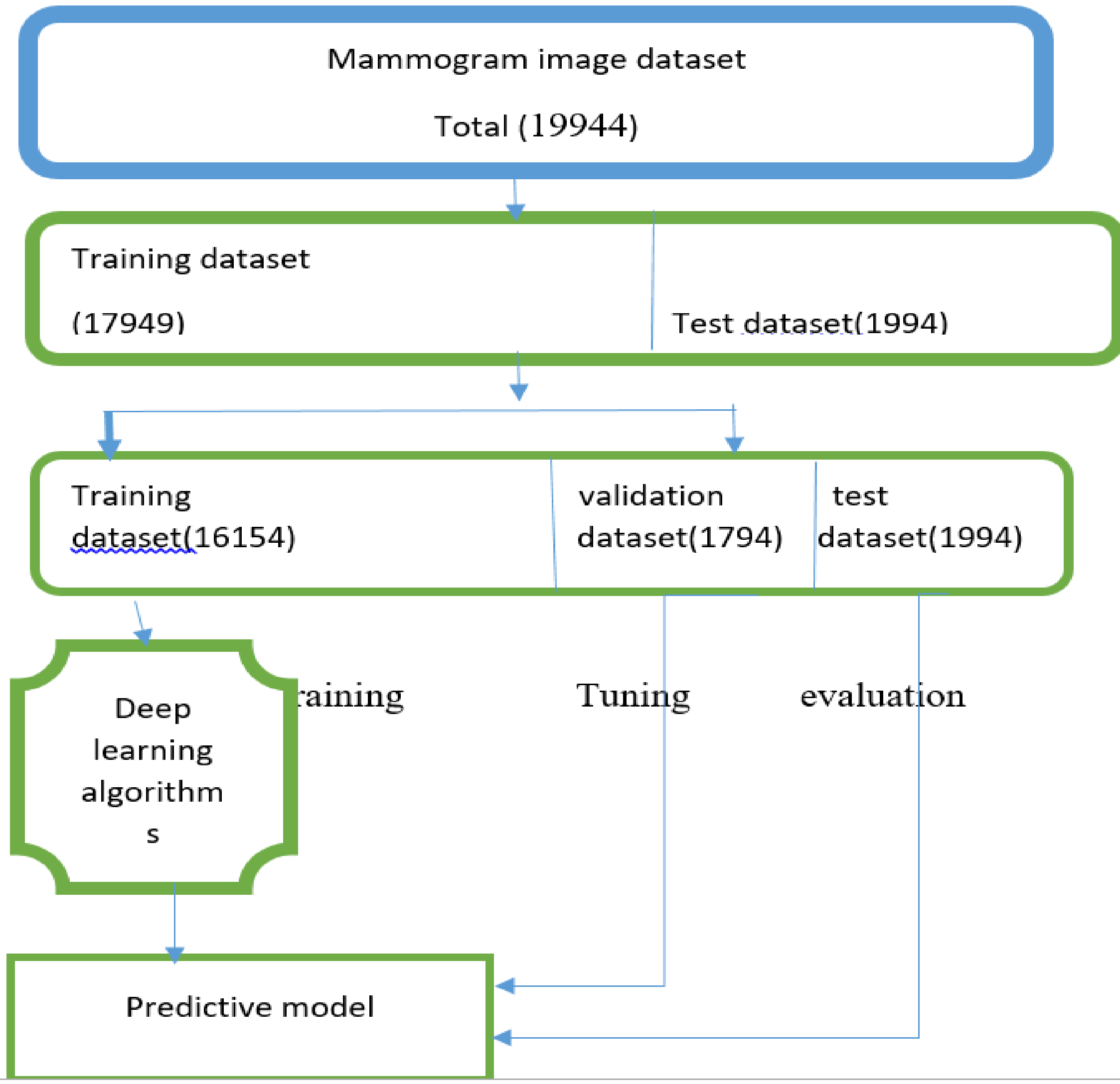

Figure 3.2 Data Partitioning

3.3.3. Data Augmentation

Research datasets for image categorization are typically very big (especially for applications involving deep learning). Data augmentation is the process of producing extra data from the current training sample in order to increase the amount of training data points in a dataset [50]. Even though the data collection is enormous, it is crucial to increase the number of data points. It assists the network in extracting more intricate characteristics from the data and mitigates the overfitting issue [51]. In order to obtain more photos for our data collection, several data augmentation techniques have been applied to the original images in this thesis. Before feeding the data to build the model, we can undertake data augmentation, or we can do it while training. In this thesis, data augmentation is carried out utilizing Keras libraries during network training. The original image serves as the basis for each image that was delivered into the network during training.

3.4. Implementation Technique and Tools

For this research, many types of development tools are used to design and implement the proposed model. Here under a description with justification of the implementation tools is given.

3. 4.1 Software Tools

In order to choose the best software tool for breast cancer detection from mammography pictures using the CNN algorithm, a review of the software products that are currently on the market with their libraries is done. We discovered during the inquiry that some tools, like Python, are broad and can be used with both machine learning and deep learning algorithms, while others, like SVM, are particular and can only be used with one of them.

We took into account the following factors before choosing the tools since they enable us to find the right software tools and the accompanying libraries. The primary factor is the computer language used to carry out the algorithm. The second is to choose tools that come with adequate learning resources, such as free video lessons and previous expertise. The third requirement is that the tools be utilized in computers with restricted resources (like CPU only). Python has been used as the programming language, together with the Tensor Flow and Keras libraries, in the Anaconda environment to develop the CNN algorithm. These tools meet every need for consideration, and they work in the well-known programming language Python.

TensorFlow The most well-known and quickest deep learning library at the moment is TensorFlow, a free and open-source program created by Google [52]. It can be utilized in the cloud as a service, on mobile devices like iOS and Android, and on any desktop computer running Windows, macOS, or Linux. TensorFlow's design is effective for data preprocessing, model construction, model training, and model estimation. Tensors (n-dimensional arrays) are used in all TensorFlow computations to represent various types of data. For the graphical depiction of the series of computations during training, TensorFlow additionally uses a graph framework. It has two CPU and GPU distributions.

Keras TensorFlow, Theano, or Microsoft Cognitive Toolkit are supported by the high-level neural network API Keras, which is written in Python (CNTK). The pretrainedCNN models VGG16 and Inception that we utilize in the experiment are included, and it is very easy to develop a model, is user-friendly, and is easily expandable with Python. It supports both CNN, RNN, and the combination of the two and enables quick and simple prototyping [53].

CUDA: - is a parallel computing platform and programming model for general-purpose computing that makes using a GPU straightforward and elegant.

CUDNN: -NVIDIA CUDA Deep Neural Network (cuDNN) is a GPU-accelerated library of primitives for deep neural networks. It provides highly tuned implementations of routines arising frequently in DNN applications.

Mendeley:is Desktop software that is used as a PDF reader and also helps to easily cite documents with desired formats like IEEE, APA, and others. Mendeley automatically loads all information that is used for citation, unlike Microsoft Word which takes time and effort to fill all the information, because of this; it's a great tool for researchers. It has a plugin that can be synchronized with Microsoft Word.

Hardware Tools: Hardware tools that are used for the implementation of this research are hard disk: for the storage of the datasets, GPU: for increasing the computation for an efficient training purpose, CPU: for testing the models, and RAM: for effectively train the model with CPU and GPU cooperatively. Figure 3.3 shows specification of the laptop used for experimentation.

Figure 1 Laptop Specification

- Computer type: - Dell core i5 7th Gen
- Operating System: - Windows 10 Home
- Processor:-Intelcore(7200)I5-7200CPU@2.50GHZ2.70GHZ
- Installed RAM: -8.00 GB
- Storage Disks:-1 TB

3.5. Evaluation methods

To determine how useful and well-performing a model is, it is important to examine it. Evaluation measures like accuracy, precision, recall, and F1-score are used to determine which instance belongs to which class in computational problems like classification and detection. The classification metrics were used to calculate those measures. The classification metrics provide information on how well a model performs for each class. All of the metrics listed above, including TP (True Positive), TN (True Negative), FP (False Positive), and FN, were derived based on the confusion matrix value (False Negative). In Table 3.1, the confusion matrix is displayed.

Table 1 Table 3.1 Confusion matrix

Predicted value	Actual value		
		Negative	Positive
	Negative	TN	FN
	Positive	FP	TP

Where:

TP: Actual class of the instance is positive and predicted as positive.

TN: Actual class of the instanceis negative, and predicted as negative.

FP: Actual class of the instanceis negative, but predicted as positive.

FN: Actual class of the instanceis positive, but predicted as negative

Based on the summarized information on confusion matrix, the following evaluation metrics are computed as follows.

Accuracy is defined as the overall correct classification of instances into their belongingness so it answers the question about how often the model predicts the class's correctly i.e. healthy mammogram and infected mammogram images. It is computed using the following formula [53]

$\boldsymbol{Accuracy}$= …………… (3.1)

Precision gives insight into how often a positive value prediction is correct. Example: Predicting image as a disease infected, how often the prediction precisely predicts. It is calculated using the below formula.

$\boldsymbol{Precision}$= ………………….... (3.2)

Recall: Also known as sensitivity it describes how sensitive the classifier is while detecting positive instances. It is calculated as follows [53].

$\boldsymbol{Recall}$= ………………………. (3.3)

F1-score Is the harmonic mean of the precision and recall, and the lowest value of F1-score is 0 which means one of the metrics has a value of 0. It indicates perfect precision or recall. It is calculated as follows [53].

$\boldsymbol{F1 - Score}$= …… ………. (3.4)

CHAPTER FOUR

4. METHODS AND APPROACHES

4.1. Overview

The design of the suggested model and its experimental setup are the main topics of this chapter. The creation of the suggested model, its descriptions, the features extracted, the classification is done, and the use of other pretrained models using the so-called transfer learning method are all briefly covered.

4.2. The proposed architecture

The system design starts with the preparation of training and validation data, and the data is labeled by three expertswhen collecting the mammogram's images, then the raw data passes through data preparation. In this stage the images prepared using data normalization, cropping, and labeling into its categorical diseases. The preprocessed images have 417*417*3 dimensions and clearly show the area of interest. Each image's important features are retrieved from the CNN staked layers, and classification is then done based on the extracted feature for model creation.During training, the model features of the breast were extracted, and based on the feature extraction the classifier detected from the mammogram patterns of benign and malignant breast. The model's performance was measured using the validation dataset while the model is training. In short, after being the model is trained, the validation set is used to assess how well the model is performing. The highest performing model was built from scratch and several pretrained models were experimented with, and using model evaluation the top performing model was selected. Finally, under the testing phase, the top-performing model was tested with unseen and unlabeled data. The model finally gives class prediction which is the probability of the image belongs to one of the given class during the training (in our case the classes are benign and malignant).

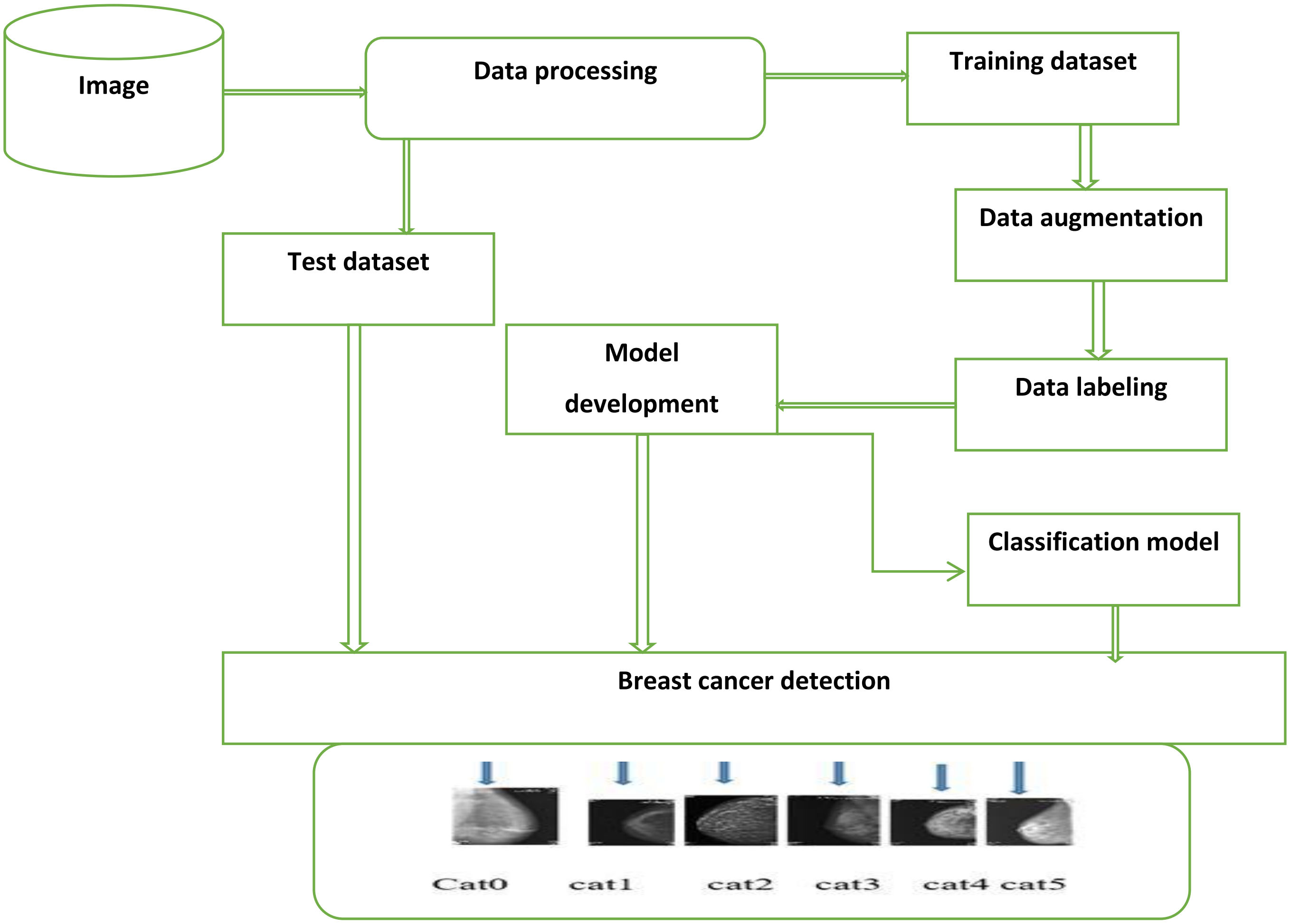

Figure 4.1the proposed Architecture for Breast Cancer Detection

The proposed system architecture of breast cancer disease detection is depicted above in Figure 4.1. Each process of the architecture is explained above.

4.3. Image data preparation

4.3.1. Data Preprocessing

The collected images are resized to 417*417 by using python resize script source code. The resized images are labeled or categorized to the respective categories of cancer stages by using LabelImage tools. When the image is annotated the file is in the form of.xml file. Then the .xml file had changed to .csv file through .xml_to_.csv and then the object detector read the ft. Record. So both train.csv and test.csv are converted to train.tfrecord and test.tfrecord for fed into model.

4.3.2. Data Augmentation

The data annotated are uploaded to rob flow workspace and some augmentation techniques were applied automatically. In this workspace there is an option for augmenting the annotated data. Rotating 90 degree rotate, blur and flip are applied on original image.

CNN has millions of parameters, to build a robust model massive amount of data is required. Insufficient data would cause a problem called overfitting which is explained in Section 2.3. As mentioned in Section 3.4.2, augmentation is one of the techniques that help to increase the number of image data which would resolve the model over fitting problem[55]. Data Splitting

The data was partitioned by train test split method to overcome the problem of over fitting. Train dataset was used for developing a model while the test dataset was used for evaluating developed models.in this thesis the dataset are split into 90% to 10% training testing ratio. The validation dataset was needed to reduce overfitting.so 90% of training dataset was split into training and testing dataset in training 80% training and 10% validation. Finally 80% Training, 10% Validation and 10% for testing dataset were used.

4.3.3. Training Components of the Proposed Model

There are no built-in deep learning models which solve all machines learning problem. Many machine learning problems build a model as per the computational problem. Having a small dataset and less computational resource affect proposing the CNN model. The dataset has two classes i.e. the healthy breast and infected breast. The data partitioned to 80 percent for training, 20 percent for testing,the scratch model has three convolution layers followed by three pooling layers, two fully connected layers, and the output layer used a sigmoid activation function. In each convolution (Conv2d, Conv2d_1, and Conv2d_2) and the first dense layer used a non-linear activation function called ReLu is used.

To solve our multiple image classification problems the model was built from the scratch using different hyper-parameters mentioned below. The value of the hyper-parameters experimented in different cases.

Input layer: the input layer of our CNN model accepts RGB images of size 417 × 417 × 3 with five different classes (category1, category2,category3,category4 and category5). This layer only passes the input to the first convolution layer without any computation. Therefore, there are no learnable features and the number of parameters in this layer is 0.

Convolutional layer: The scratch model has 3 convolution layers i.e. (Con2d, Conv2d_1, and Conv2d_2. The input image with dimensions of (None, 417, 417, 3) passes from the input layer. Then Con2d takes the output of the input layer as input layer i.e. (None, 417, 417, 3), and generates output (None, 417, 417, 16).

```
model =tf.keras.models.sequential([
       tf.keras.layer.conv2D(16,(3,3),activation='relu', input_shape=(417,417,3))
       tf.keras.layers.maxpooling2D(2,2),
       tf.keras.layers.conv2D(32,(3,3),activation='relu'),
       tf.keras.layers.maxpooling2D(2,2),
       tf.keras.layers.conv2D(64,(3,3),activation='relu'),
       tf.keras.layers.maxpooling2D(2,2),
       tf.keras.layers.Flatten(),
       tf.keras.layers.Dense(512,activation='relu'),
       tf.keras.layers.Dense(1,activation='sigmoid')
])
model.compile(optimizer=RMSprop(lr=0.001, loss='binary-crossentropy', metrics=['acc']))
```

```
history= model.fit(training_set,
                   epochs=30,
                   verbose=1,
                   validation_data=
                   validation_set)
```

Figure 4.2 CNN model

As Figure 4.2 shows the Conv2d layer has filter=16, kernel size=3x3, stride, and dilation rate=1 as a default. The output shape is calculated as:

(Image Height - KernelHeight+Stride,ImageWidth-KernelWidth+Stride, Filter).

Given all the parameter of the formula, the output shape becomes, (417-3/2+1, 417-3/2+1, 16) this becomes (415, 415and 16) and the number of a parameter in this layer calculated as:

*(KernelHeight*KernelWidth*InputChanal*OutputChanal+OutputChanal if bias used)*

Using the formula, the number of parameters calculated as (3*3*3*16+16) equals 448 parameters. The max-pooling layer takes input from the Conv2d layer's output i.e. (None, 415, 415, 16).

Pooling layer: The first, second, and fifth convolutional layers of the proposed model are followed by three max-pooling layers. With a filter of size 3*3 and stride 1, the first max-pooling layer reduces the output of the first convolutional layer. The output of the second convolutional layer serves as the input for the second max pooling layer, which pools data using stride 1's 2 *2 filters. The filter in the third max-pooling layer is size 2*2 with stride 2. This layer just down samples data along the spatial axis of the input volume and lacks any learnable features.

Fully Connected (FC) layer

The output layer is one of three completely connected layers in the suggested model. 64 neurons each make up the first two fully linked layers, whereas there is just one neuron in the model's output layer's last layer. After turning the 3D volume of input into a vector value, the fifth Conv layer output is accepted by the first FC layer (Flattening). This layer calculates the class score and the number of neurons that were predefined for the layer throughout the model's development. It is the same as regular NN, and as its name suggests, every neuron in this layer is linked to every number in the layer behind it.

Output layer: The model's final output layer, or third FC layer, contains one neuron with a softmax activation function. Since the model is built to categorize six different categories, it uses multiple categorization.

4.4 Hyper-parameter Tuning

Hyper-parameters are settings that are independent of the deep learning algorithm and whose values are predetermined before training. There is no set formula for selecting the ideal hyper-parameters for a particular issue. As a result, many experiments are run to select the hyper-parameters. The hyper-parameters selected for the model are discussed in the paragraphs that follow.

Optimization algorithms: To train a neural network an optimization algorithm selection plays a major role because it allows a model to learn faster and achieve better performance. The backpropagation of the error procedure is used to update the weights. There are many different types of optimization algorithms, among them Adam and RMSprop, which were tested on learning rates of 0.0001 and 0.00002 to reduce the error rate. In deep learning research, gradient descent is by far the most well-liked and frequently applied optimization strategy. At the same time, Keras and other gradient descent optimization methods are implemented in every modern deep learning package (used in this thesis). It adjusts parameters and updates the model's weight to minimize the loss function. The Adaptive Moment Estimation (Adam) optimizer is used to enhance the gradient descent. Adam calculates the adaptive learning rate for each parameter and scales the learning rate using squared gradients and the gradient's moving average.

Learning rate: The amount of weight updated during training is called the learning rate. It's a configurable hyperparameter used in training a neural network and the value often ranges between 0 and 1.0. It's the most important hyperparameter because it controls how fast a model adapts to a given problem. A small learning rate requires many training epochs and changes made to the weight are updated and having a large learning rate results in a rapid change to the training epoch. One of the challenges of building a neural network model is the selection of a learning rate that is not too large or too small. After several, experimentation the learning rate takes is 0.0001 and 0.00002.

Loss function: The sort of problem we are attempting to solve and the activation functions utilized in the output layer (final fully connected layer) of the model are closely related to the loss function used (whether regression or classification). In the suggested model, the final fully connected layer uses the sigmoid activation function. We are dealing with a classification problem, specifically a binary classification challenge. The loss function for our model is Binary Cross-Entropy (BCE) loss. Even though there are other loss functions, such as Categorical Cross-Entropy (CCE) and Mean Squared Error (MSE), which assess the gap between the desired and actual output, they work best for models that output probabilities. BCE loss and CCE loss were both used in the experiment.

Activation Function: Two distinct activation functions are utilized in the experiments: SoftMax in the suggested model and ReLuused in the hidden layer. The Soft max activation function, which is the optimum option for a multiple classification problem, is utilized at the output layer.

Epoch: Describe the number of times training data is shown to the neural network. During experimentation as a baseline, the number of epoch takes was 100 and by interpreting the accuracy validation graph for a proposed model epoch of 30 were used and for other pre-trained models, the number of epoch used was 65.

Batch size: Batch size is the number of subsamples given to the network after which parameter update happens. The default value of batch size is 32, 64, 128, and so on. After multiple experiments batch sizes achieved the best score is 32.

CHAPTER FIVE

5. EXPERIMENTATION AND DISCUSSION OF RESULTS

Overview

This chapter presents the implementation of Breast Cancer disease detection using a CNN algorithm. The detail of the experiments is explained with multiple scenarios, and also the hyperparameter used for experimentation is explained.

5.1. Experimental Setup

The experimental setup of the study includes four scenarios to efficiently detect and classify breast cancer disease.

The first scenario builds a model from scratch (using CNN algorithm with default parameter and breast cancer dataset prepared in this study.)

The other experiments are done based on transfer learning the available models. Accordingly, the second experiment uses the InceptionV3 model, while the third scenario is attempt is to use VGG16 model and finally the fourthexperiment is SSDMobilenetV2 model.

5.2. CNN features

The input 417*417*3 means the image with RGB are inserted into the model with filter 32 and filter size 3*3with stride 2 and the image fed into convolutional layer with 415*415*16 output is max pooled and the 208 *208 *16 are inserted into the second layer (206*206*16) are maxpooled with second max pooling layer then 103 *103*32 output fed into the last convolutional layer 101*101*64 output are max pooled and 50*50*64 are flattened to display the output of 6 class.Softmax activation function used of full connected layer to select and display one of the classes.

5.3. Augmentation parameters

The augmentation parameter used in the experiment is explained in Figure 5.1. As a result, using data augmentation the number of images increased from 2893 of all class images of the training set to 19944 images which help the model to learn more features

```
rescale=1./255,
rotation_range=40,
width_shift_range=0.2,
height_shift_range=0.2,
shear_range=0.2,
zoom_range=0.2,
horizontal_flip=true,
fill_mode='nearest'
```

Figure 5.1 Augmentation Parameter during training

Table 5.3 shows hyper-parameter setting for pre-trained models.

Parameter	Epoch	Batch Size	Activation Function	Loss Function	Optimization Algorithm	Learning rate
Value	90	32	ReLu	BCE	Adam	0.001
	30	64	Softmax	CCE		0.002
	65	128	Sigmoid			

5.4. Experimental Result

5.4.1 Breast cancer disease classification using a model built from scratch

Before using the pre-trained model the experiment began by building a CNN model from scratch and compared it with other pre-trained models. It has good training and validation accuracy, but the performance of the model with unseen data is less when compared to the pre-trained models like VGG16, InceptionV3, and SSDmobilenet. Although, building a model from the scratch is not recommended using the relatively small dataset, because the model is learn few features of images and testing with unseen data might not yield a good result. As a result, the model built from scratch has shown good performance on the training and validation set, but testing with unseen data has a low performance. Generally, the models were experimented with using different scenarios.

Scenario 1: Checking with an image dimension of 300x300x3 and 224x224x3.

Scenario 2: Checking with a learning rate of 0.00002 and 0.0001.

Scenario 3: Changing the last layer activation function, with ReLu and Softmax.

Scenario 4: Changing the optimization algorithm with RMSprop and Adam with the mentioned learning rate.

All of the four scenarios have experimented and the best result achieved was the scratch model with an image dimension of 300x300x3, a learning rate of 0.001, and last layer activation of ReLu and RMSprop optimization algorithm. The result is shown in terms accuracy, recall and precision in table 5.2 below.

Table 5.4.Summary of experimental results

	Accuracy	Recall	Precision
Scenario 1(224*224)	86	81	82
Scenario 2(0.001)	87	83	84
Scenario 3(Softmax)	88.2	86	86.5
Scenario 4(Adam)	88.6	86.8	87

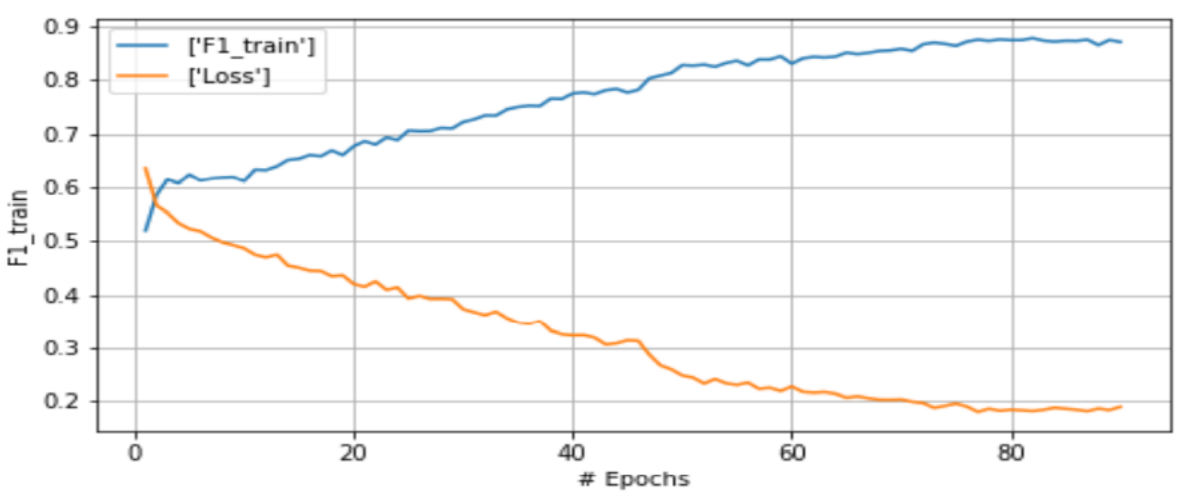

Figure 5.2: graph of Scratch model

Table 5.5 train accuracy vs. train loss of scratch model

Metrics	Mean Accuracy			Mean Loss		
	Training	Validation	Testing	Training	Validation	Testing
Value	88%	89%	86%	5.01%	5.6%	14%

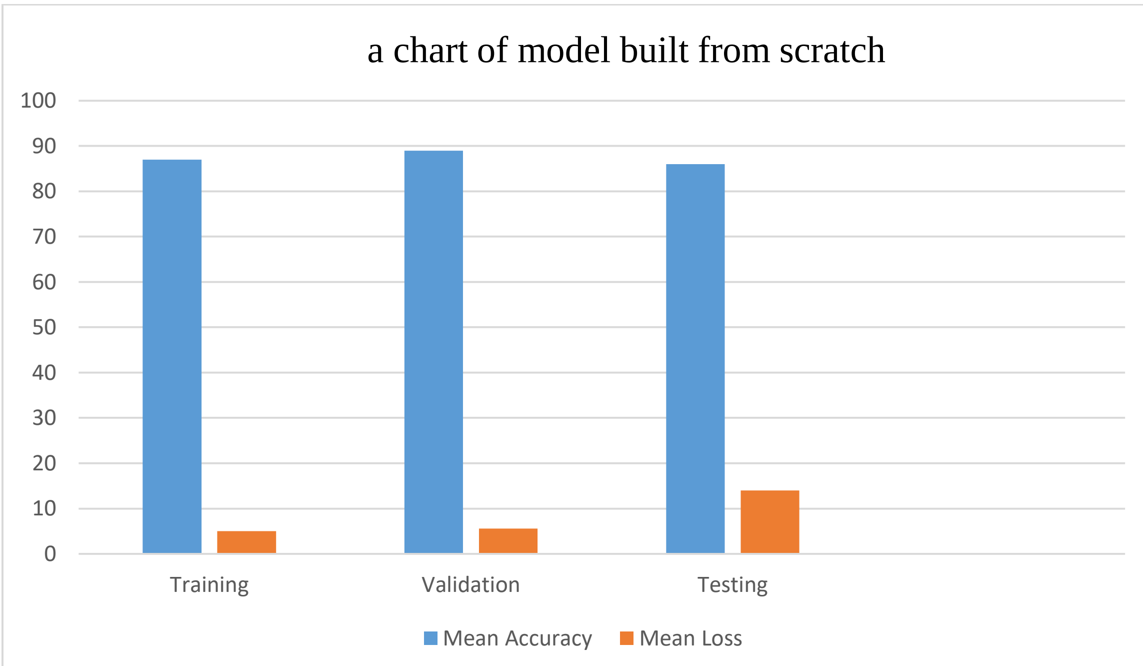

Figure 5.3. Chart Model for Scratch Model

5.4.2 Breast cancer disease classification using Pre-trained model

The Three pre-trained CNN models: VGGInceptionV3 and SSDmobilenetthat are widely used and with pre-trained architectures in Image Net are used and fine-tuned. Those selected pre-trained model are selected for their performance of breast cancer detection with low computation power device such as mobile and tablet. The InceptionV3model is employed due to its complex characteristics, the VGG model is picked for its simplicity, and SSDmobilenet is chosen due to its cheap processing requirements[56]. To determine the classification accuracy of these models in our dataset, tests are run on both a reasonably basic model and a sophisticated one. The same dataset and hyper-parameter configuration are used for all of the studies.

Experiment Result of VGG16

All implementation information and hyper-parameters from the baseline works are used in the VGG16 experiment. To initialize the kernel using Adam optimizer, the first 15 epochs were trained using Image Net weight. Another 90 epochs were trained using weight that was SGD optimizer-trained on the first 15 epochs. The architecture has been completely fine-tuned. The learning rates used are 0.001, 0.002, and 0.2 for dropout rates.

Table 5.6 VGG16 Experiment Setup

	Learning rate	Accuracy	Recall	Precision
Image Net weight with Adam optimizer	0.001	78	76	75
	0.002	77	75.2	74.6
15 epochs with SGD optimizer	0.001	74	72.4	72
	0.002	73	71	79.6

Figure 5.4 displays a loss of VGG16 of 7% and training accuracy of 77.8% for customized F1_score. Table 5.5 provides the experiment results of VGG16 on the dataset for the detection of breast cancer as measured by evaluation metrics.

Table 5.7: VGG16 Model Accuracy and Loss Result

Metrics	Mean Accuracy			Mean Loss		
	Training	Validation	Testing	Training	Validation	Testing
Value	77.8%	78.4%	76%	7.0%	7.70%	22.5%

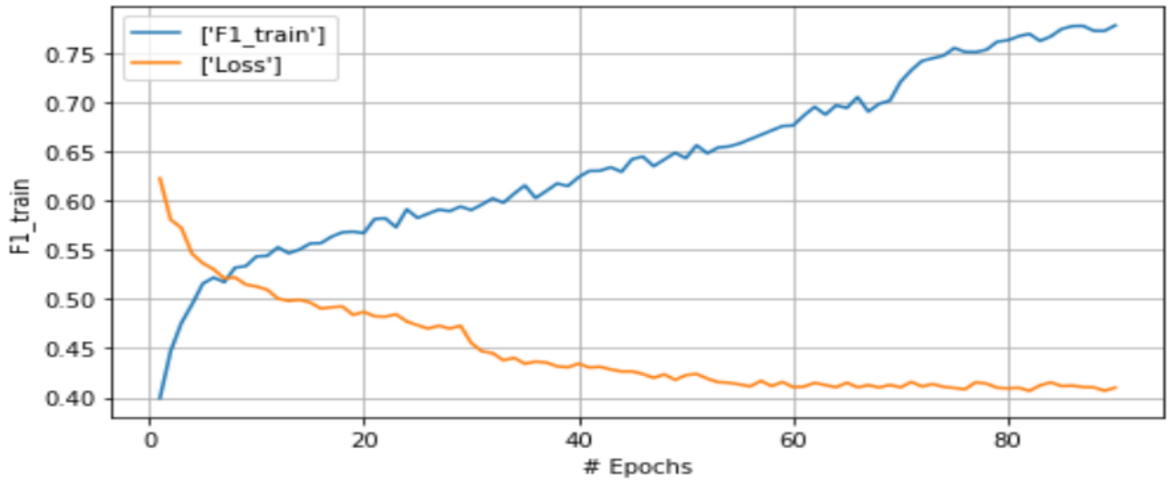

Figure 5.4 Training accuracy and loss of vgg16

The graph shows that as an epoch are increases the learning ability of the model increase and also the loss of the model decrease. And also Adam optimizers change the performance of the model.

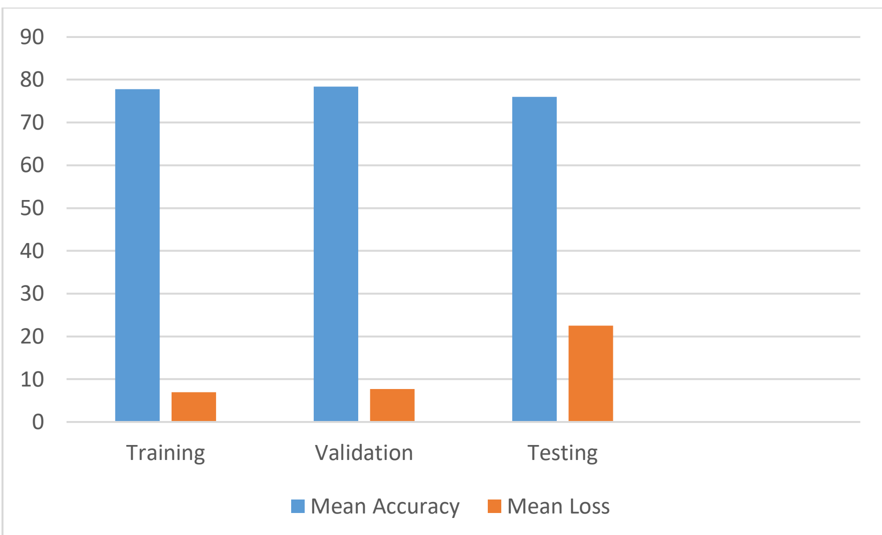

Figure 5.5 Graph VGG16 Mean Accuracy and mean Loss

Experiment Result of InceptionV3

The hyper-parameter setup directly influences how well the model performs when the neural network is being trained. With a 10^{-4} learning rate, 0.2 dropouts, and training over 90 epochs, InceptionV3 is implemented. Table 5.6 displays the InceptionV3 experiment's classification of breast cancer results.

Table 5.8 InceptionV3 Experimental Result

	Learning rate	Accuracy	Recall	Precision
Image Net weight with Adam optimizer	0.001	88.4	87	86.4
	0.002	87.8	85.6	85.8
15 epochs with SGD optimizer	0.001	86	83.6	85.4
	0.002	85	84	86

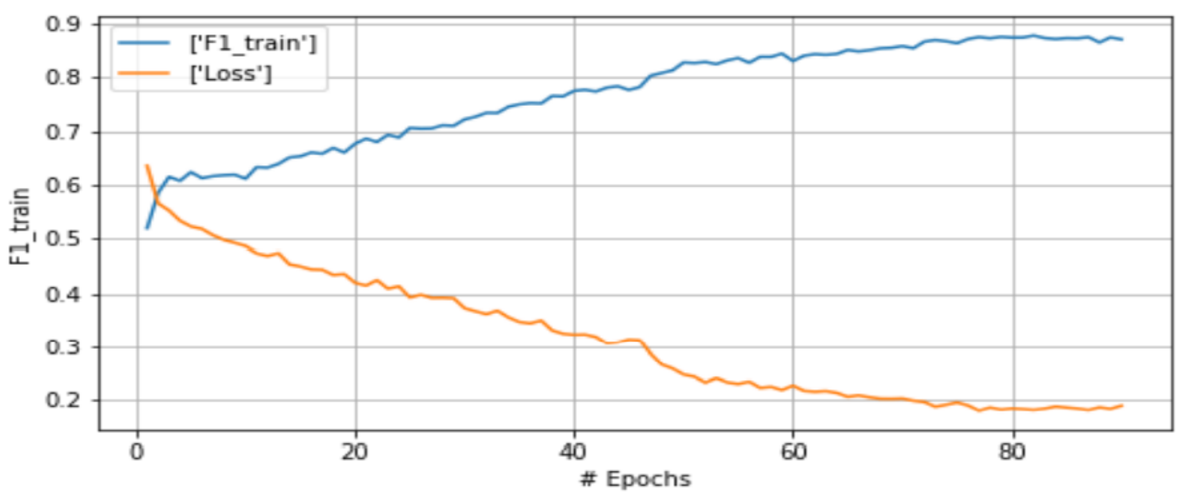

Figure 5.6 InceptionV3 Training and Loss Graph

Metrics	Mean Accuracy			Mean Loss		
	Training	Validation	Testing	Training	Validation	Testing
Value	88.4%	88.7%	85%	6.72 %	5.5%	9%

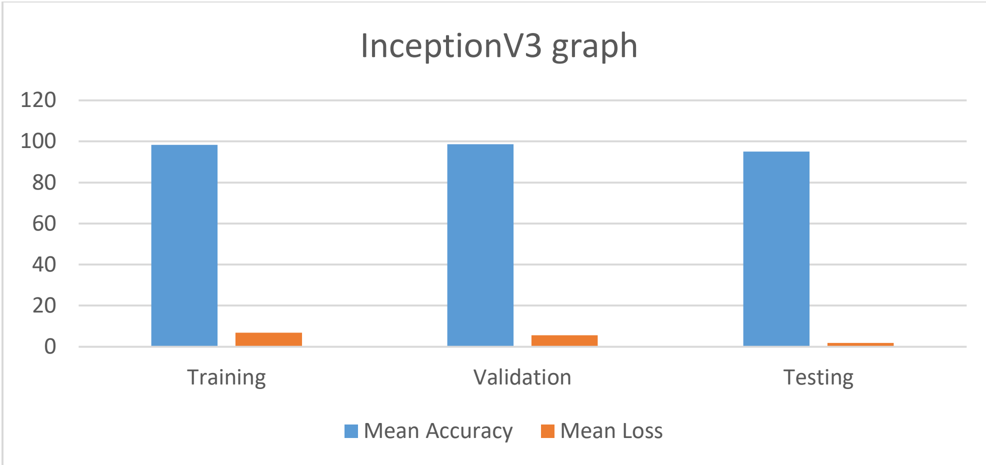

Figure 5.7 InceptionV3 graph mean accuracy and mean loss

The InceptionV3 perform well on training and testing. The figure 5.7 show that the learning ability of the model is increased with increasing epoch and Adam optimizer perform better than SGD optimizer on optimizing developed model.

Experimental result of SSDmobilenet

The hyper-parameter configuration for the SSDmobilenet is identical to that in table 5.1. 7 × 7 x 1024-pixel feature map is extracted. The computational performance of the model during the NN's training period is decreased by the little feature that was extracted. The experimental findings are shown in table 5.7.

Table 5.9 Experimental result of SSDMobilenetV2

	Learning rate	Accuracy	Recall	Precision
Image Net weight with Adam optimizer	0.001	88.8	87	87.2
	0.002	86	86.6	86.4
15 epochs with SGD optimizer	0.001	88.4	88.3	88.2
	0.002	87.2	87.6	87.8

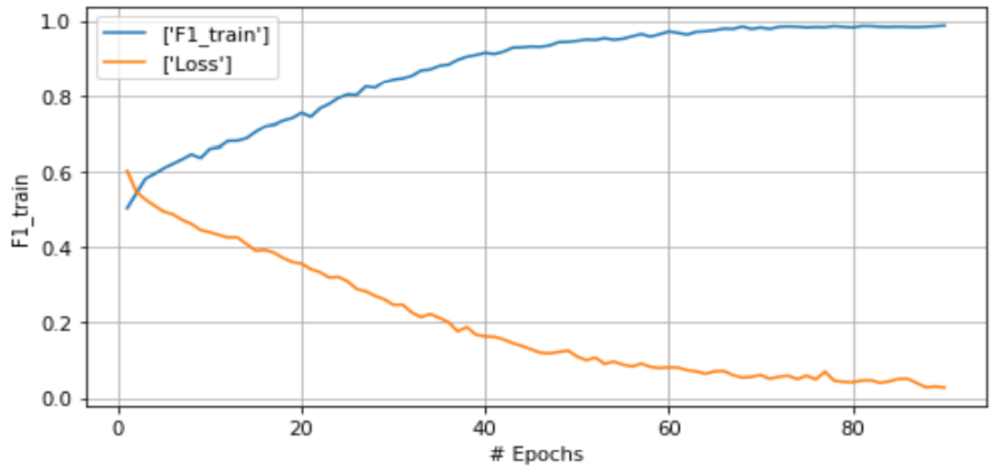

Figure 5.8 Training and Loss of SSDMobilenetV2

The above graph shows that as epoch increases the performance of the model increases. But after epoch 80 the graph becomes saturated.

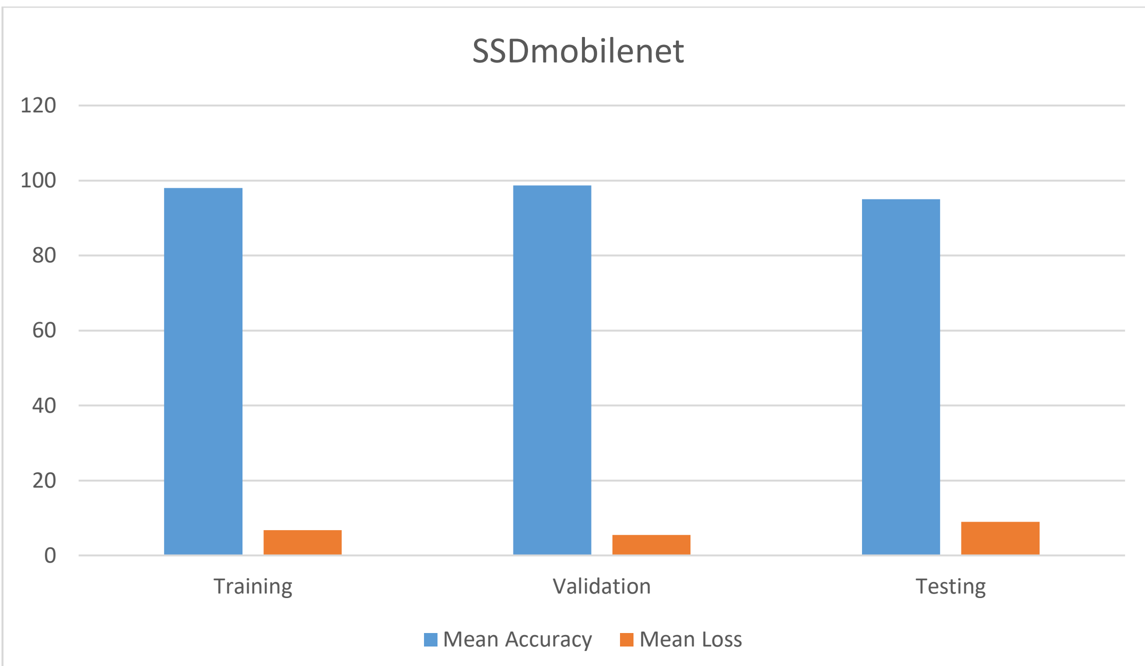

Figure 5.9. SSDmobilenet Mean Accuracy Vs. Mean Loss

Discussion

Comparison of performance result

Table 5.10 Comparison result of Models

Model Name	Accuracy	Loss
CNN model	87	14
VGG16	79	22
InceptionV3	89	9
SSDMobilenetV2	88	11

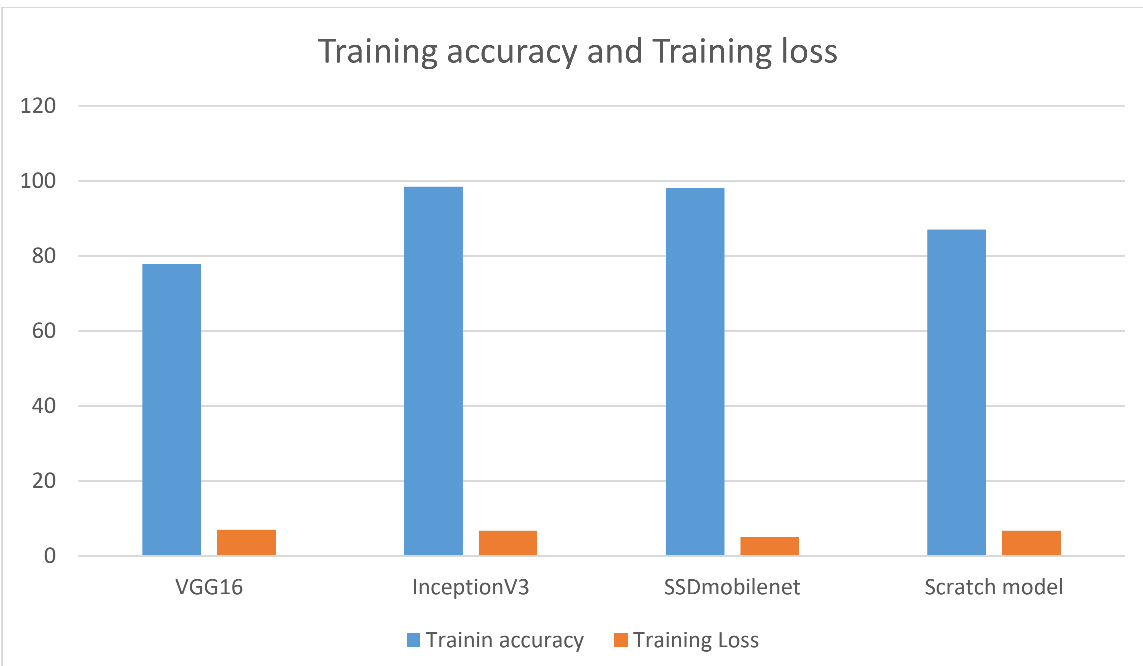

Figure 5.10. Model Comparison

The ultimate goal of the study was to build breast cancer disease classification model and to achieve the goal different literature was reviewed, experiments were conducted. The algorithm selected to detect and classify the breast cancer disease is CNN. The algorithm was experimented with in-depth by building a model from the scratch and also transfer learning using a pertained model was used. The drawback of building a model from the scratch is it requires a large data and the model performance depends on the amount of data. The training and validation accuracy of the model built from the scratch was good, but testing with unseen data the performance is poor. As a result, the pre trained model resolved the problem of the model built from the scratch. The pre trained model (VGG16, SSDmobilenet, and InceptionV3) experimented with the different hyper parameter settings.

High accuracy is displayed by InceptionV3 on the practice set. However, it has the lowest value based on the testing set and evaluation metrics. Overfitting is reduced using the regularization procedure, and it has been empirically proven to be an issue with the short dataset that was employed. This is because overfitting happened because to the high parameters that were calculated on a short number of data. On the other hand, the least accurate training model is VGG16. However, its values are higher than those of SDD mobile net in terms of evaluation metrics. The least expensive in terms of computation and training time is a model developed from scratch. However, it is less accurate than InceptionV3.

The mean % training accuracy for VGG16, InceptionV3, and the suggested CNN model, as shown in the following plots, is 86.6, 81.7, and 88.49, respectively. These demonstrate that the models produce good results on the training dataset. For the VGG16, InceptionV3, and suggested models, the mean percentage of validation accuracy is 86.8, 81.5, and 88.48, respectively. The suggested model almost has the same mean training accuracy and mean validation accuracy when we compute the difference between mean training accuracy and mean validation accuracy for each of the three experiments. These demonstrate that the models are not overfitted, and we may conclude that the suggested model has a high degree of generalizability.

When it comes to the mean training loss a metric that assesses the discrepancy between the value predicted and the actual value we got 7, 14.3, and 4.3 for the three experiments, VGG16, InceptionV3, and the proposed model, respectively. When we compute the difference between mean training loss and mean validation loss, the mean validation loss is 7.7, 14.3, and 5.7, which is very similar to the mean training loss.

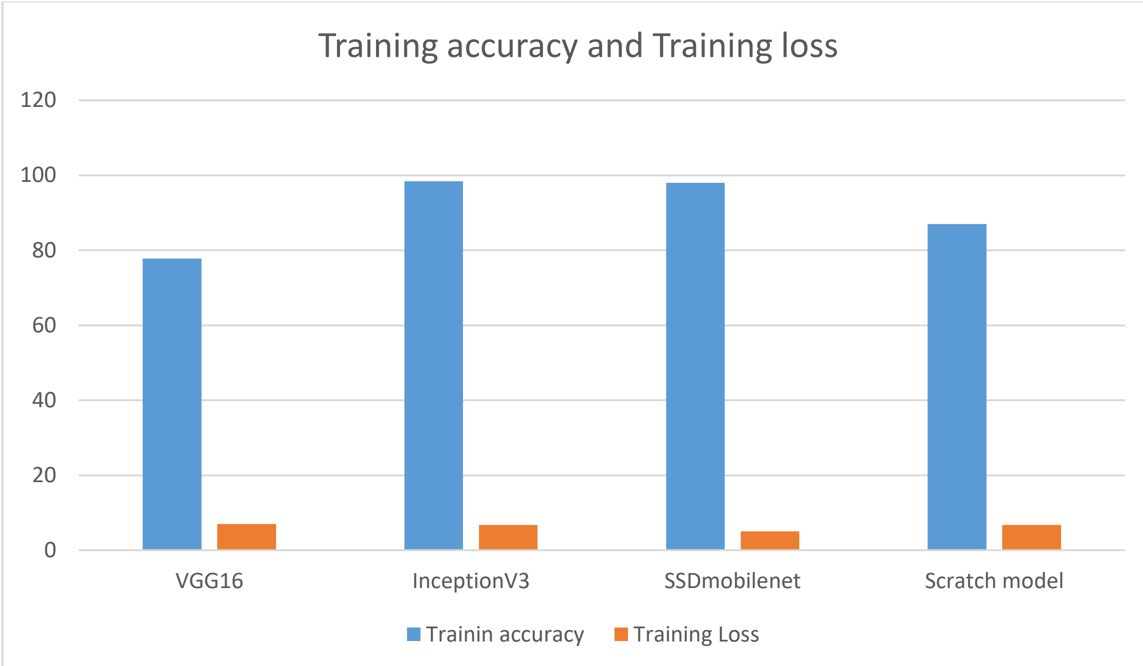

Figure 5.11.Mean accuracy of the four experiments

We have achieved a promising outcome by testing the models with unknown data. VGG16 test accuracy is 77%, InceptionV3 test accuracy is 88%, and SSDmobilenet test accuracy is 88% for the proposed model. The test results demonstrate that the suggested model performs more accurately than the pre-trained VGG16 and InceptionV3 models in classifying the provided image as either healthy or unhealthy.

The test loss for all experiments is depicted in Figure 5.9 above, and the suggested model's value is less than that of the two previously trained models. Test losses for VGG16 are 22.1, InceptionV3 are 27.2, and the proposed model has a test loss of 6.23, all of which are respectable values. The suggested model is therefore performing well on both the training dataset and the testing dataset.

The images in the dataset that we used to train the model are easily classified with human eyes, which is one of the primary reasons why the suggested model performs better. The second main reason is that our proposed model uses smaller sized filters in the convolution layer of the network. A very tiny chance of losing an important feature is created by using smaller-sized convolution, which aids in identifying extremely small features that are utilized to differentiate between the input image and the output image.

A large number of images (in the millions), tens of millions of parameters, and high-performance computing equipment with faster GPUs are used to train the majority of deep learning algorithms, particularly computer vision for image categorization issues. But with smaller networks, fewer parameters, less hardware use, and less data, we can train and achieve better outcomes. If the photographs in the dataset were taken in stable environmental conditions, such as stable object distance from the camera, good lighting, and perfect focus, more accurate findings were achieved. The model's accuracy is also improved by preprocessing the images to get rid of noise and undesirable characteristics.

CHAPTER SIX

6. Conclusion And Recommendation

Cancer is a disease caused by uncontrolled cell replication and breast cancer is one type that originated in the breast. Breast cancer cells usually form tumors, which can usually be seen on X-rays. It occurs mostly in women, and sometimes men also develop breast cancer. The origin of breast cancer is the tube that carries milk to the nipple. Sometimes it begins in the glands that make breast milk. So in Ethiopia women are highly susceptible to this cancer at a high rate. So if this cancer is detected and classified early the model has its contribution.

Single label classification with one specific category was done previously. Some use mass finding while others use calcification. Density is also used by other studies to screen for breast cancer. But mammogram gives more labels or information at the same time. Multiple classification handle this problem by giving multiple predictive reports at the same time.

In chapter one, in Section 1.4 research questions were formulated. Based on the research questions the study answered all the questions formulated. The answers to the questions are explained as follows.

Does the first research question ask for methods that could be used for preparing a dataset used for breast cancer disease detection?

For problems like image classification, image recognition, and image detection CNN is a current state-of-art algorithm [8],[9],[10],[11]. The algorithm has shown promising results in many competitions. Generally, there are two ways to build a CNN model. i. Building a model from the scratch: In this method, the components like Convolution, Pooling, Activation Function, and Fully Connected layer are designed carefully. The design also contains the hyper-parameter configuration. The major problem of the method is the amount of data because having a small amount of data limits the neural network to acquire enough knowledge. Because of this, training a model from the scratch with limited data would result in low results in the testing set. ii. Transfer learning using a pre-trained model: Pretrained models are also used in different scenarios.

The second research question is concerned with, the deep learning algorithms to apply for developing an optimal model for the detection of breast cancer?

Since (a convolutional neural network) is the state-of-the-art deep learning algorithm to develop a breast cancer detection model. in this study we used it in which promising results were obtained.

The third research question is concerned with evaluating the performance of the proposed breast cancer disease detection model?

The performance of the models was assessed using performance metrics (Precision, Recall, F1-score, and Accuracy). The classes of the data have an equal number of images. For these types of data, the classification metrics mentioned above have equal value. Of the three pre-trained models the top performed model is InceptionV3. It scored 88.4% accuracy on unseen data.

Generally, InceptionV3 is investigated and identified as the best feature extractor by answering the first research question to be answered by this work. Among the conducted experiments SSDmobilenet model gave the best next to InceptionV3 performance results. This guarantees that InceptionV3 is the best feature extractor for multiple classifications. The results indicate that the proposed model achieves a substantial improvement over existing methods in terms of all conducted evaluation metrics.

6.1 Future Works

Finding the best solution for a problem on hand and giving direction for the future is the nature of research. Optimizing the feature extractor by cutting in different layers with compatible fully connected NN further is recommended since this increase the accuracy and decreases the parameters to be calculated. Besides, using more mammogram datasets and other modalities of breast datasets is extending this work in the future. Applying segmentation before CNN can enhance the performance of the model. Because in this thesis onlyRectangle box was only used for identifying the region of interest.

In addition, also standard dataset is needed for experimentation with a large amount to get a more reliable and consistent model. Finally, this research has to be extended to future to other health cancer problems.

References

[1] Z. Zheng, H. Zhang, X. Li, S. Liu, and Y. Teng, "ResNet-Based Model for Cancer Detection," *2021 IEEE Int. Conf. Consum. Electron. Comput. Eng. ICCECE 2021*, no. Iccece, p. pp. 325–328, 2021.

[2] M. Li, "Research on the detection method of breast cancer deep convolutional neural network based on computer aid," *Proceedings of IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers,*, pp. 536-540, 2021.

[3] H. Peng, Y. Jiang, X. Li, H. Luo, and S. Yin,, "A novel redundant information elimination aided classification approach for cervical cancer diagnosis,," *Proc. 2019 11th CAA Symp. Fault Detect. Supervision, Saf. Tech. Process. SAFEPROCESS 2019,* p. 631–636, 2019.

[4] Sung, Hyuna, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians,* Vols. 71, no. 3, p. 209–249, 2021.

[5] M. Slimi, B. Jmai, P. Mendes, and A. Gharsallah, "Breast cancer detection based on CPW antenna," *IEEE Access,* vol. 4, no. 8, pp. 3-6, 2019.

[6] B. T. Deressa, N. Cihoric, E. V. Badra, A. Tsikkinis, "Breast cancer care in northern Ethiopia-cross-sectional analysis,," *BMC Cancer,* vol. 19, no. 8, p. 1–7, 2019.

[7] U. Kamath, J. Liu, and J. Whitaker, Deep Learning for NLP and Speech Recognition, vol. 2, mumbai: IEEE, 2019.

[8] S. Khan, H. Rahmani, S. Afaq, A. Shah, and M. Bennamoun, "A Guide to Convolutional Neural Networks for Computer Vision," *IEEE access,* vol. 1, no. 8, pp. 29-224, 2020.

[9] Getachew, Sefonias, "Perceived barriers to early diagnosis of breast Cancer in south and southwestern Ethiopia: a qualitative study," *BMC Women's Health,* vol. 20, no.

1, pp. 1-8, 2020.

[10] Patterson, Josh and Gibson, Adam, Deep learning: A practitioner's approach, O'Reilly Media, Inc., 2017.

[11] T. G. Debelee, S. R. Kebede, F. Schwenker, and Z. M. Shewarega,, "Deep Learning in Selected Cancers’ Image Analysis—A Survey,," *Journal of Imaging,* vol. 6, no. 11, p. 121, 2020.

[12] I. Allaouzi and M. Ben Ahmed, "A Novel Approach for Multi-Label Chest X-Ray Classification of Common Thorax Diseases,," *IEEE Access,* vol. 7, p. 64279–64288, 2019.

[13] Deressa, Biniyam Tefera, "Breast cancer care in northern Ethiopia-cross-sectional analysis," *BMC Cancer,* vol. 19, no. 1, pp. 1-7, 2019.

[14] Akram Farhadi, David Chen, Rozalina McCoy, Christopher Scott, John A. Miller, Celine M. Vachon, Che Ngufor, "Breast Cancer Classification using Deep Transfer Learning on Structured Healthcare Data," in *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2019.

[15] Biniyam Tefera Deressa, Nikola Cihoric, Daniel Rauch, "Breast cancer care in northern Ethiopia cross-sectional analysis," *open accesss,* pp. 1-6, 2019.

[16] T. Assefa, D. Haile Mariam, W. Mekonnen, M. Derbew, and W. Enbiale, "Physician distribution and attrition in the public health sector of Ethiopia," *Risk Management and Healthcare Policy,* vol. 9, no. 2, p. 285–295, 2016.

[17] S. Kaya and M. Yağanoğlu, "An Example of Performance Comparison of Supervised Machine Learning Algorithms Before and After PCA and LDA Application: Breast Cancer Detection," Istanbul, Turkey, 2020.

[18] CHANG-MIN KIM , ROY C. PARK , ELLEN J. HONG, "Breast Mass Classification Using eLFA Algorithm Based on CRNN Deep Learning Model," vol.

8, pp. 197312-197323, 2019.

[19] K. Varshini, R. K. Sethuramamoorthy, V. Kumar, S. A. Shree, and S. Deivarani, "Breast cancer prediction using machine learning techniques,," *Int. J. Adv. Sci. Technol*, vol. 29 , no. 6 spatial issue, p. 2026–2032, 2020.

[20] Y. Cheng and B. Li, "Image segmentation technology and its application in digital image processing," *Proc. IEEE Asia-Pacific Conf. Image Process. Electron. Comput.*, vol. 3, no. 5, p. 1174–1177, 2021.

[21] M. Lyra, S. Lyra, B. Kostakis, S. Drosos, C. Georgosopoulos, and K. Skouroliakou, "Digital Mammography texture analysis by computer assisted image processing," *IST 2008 - IEEE Work. Imaging Syst. Tech. Proc*, vol. 1, no. 3, p. 73–76, 2008.

[22] L. M. G. Fonseca, L. M. Namikawa, and E. F. Castejon, "Digital image processing in remote sensing," *Tutorials of SIBGRAPI 2009 - 22nd Brazilian Symposium on Computer Graphics and Image Processing*, vol. 4, no. C, p. 59–71, 2009.

[23] R. Mékle, A. F. Laine, and E. X. Wu, "Combined MR data acquisition of multicontrast images using variable acquisition parameters and K-space data sharing," *IEEE Trans. Med. Imaging,*, Vols. 22,, no. 7, p. 806–823, 2003.

[24] Z. Kuixing and J. Mei, "Study on digital acquisition method of infantile finger veinlet image," *ICME Int. Conf. Complex Med. Eng.*, vol. 5, no. 6, pp. 421–424,, 2013.

[25] S. I. Workshop and M. Experience, "EFFECTIVE PREPROCESSING STAGE IN THE FOURIER TRANSFORM DOMAIN FOR IMAGE QUALITY ASSESSMENT," *Workshop, Sixth International*, vol. 5, no. 7, pp. 189-194, 2014.

[26] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," vol. 10, no. 1, pp. 1-22, 2021.

[27] Khatak, Anil, "A review on various filtering techniques in image processing," *Natl. J. Multidiscip. Res. Dev*, vol. 3, no. 1, pp. 1350-1355, 2018.

[28] ROSLIDAR , AULIA RAHMAN , BISWAJEET PRADHAN , KHAIRUL MUNADI, "A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection," *IEEE Access,* vol. 4, no. 6, pp. 1-20, 2020.

[29] Zhou, S. Kevin, "A Review of Deep Learning in Medical Imaging: Imaging Traits," *Proceedings of the IEEE,* vol. 109, no. 5, pp. 820-838, 2021.

[30] J. Latif, C. Xiao, A. Imran, and S. Tu, "Medical imaging using machine learning and deep learning algorithms: A review," *2019 2nd Int. Conf. Comput. Math. Eng. Technol. iCoMET 2019,* vol. 17, no. 3, p. 1–5, 2019.

[31] N. Singh and S. Singh, "Object classification to analyze medical imaging data using deep learning," *Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems, ICIIECS 2017,* Vols. 2018-Janua, pp. 1-4, 2018.

[32] M. Pandiya, S. Dassani, and P. Mangalraj, "Analysis of Deep Learning Architectures for Object Detection - A Critical Review," *Proceedings of 2020 IEEE-HYDCON International Conference on Engineering in the 4th Industrial Revolution, HYDCON 2020,* vol. 2, no. 7, 2020.

[33] B. P. Amiruddin and R. E. A. Kadir, "CNN Architectures Performance Evaluation for Image Classification of Mosquito in Indonesia," *Proceedings - 2020 International Seminar on Intelligent Technology and Its Application: Humanification of Reliable Intelligent Systems, ISITIA 2020,* vol. 2, no. 3, pp. 223-227, 2020.

[34] Ahmed, Eman, "A survey on Deep Learning Advances on Different 3D Data Representations," vol. 1, no. 1, pp. 1-35, 2018.

[35] Y. Seo and K. S. Shin, "Image classification of fine-grained fashion image based on style using pre-trained convolutional neural network," *2018 IEEE 3rd International Conference on Big Data Analysis, ICBDA 2018,* vol. 4, no. 8, pp. 387-390, 2018.

[36] Shaily, Tumun, "Bacterial Image Classification Using Convolutional Neural Networks," *2020 IEEE 17th India Council International Conference, INDICON 2020,* vol. 3, no. 9, 2020.

[37] Hakim, Luqman and Sari, Zamah and Handhajani, Handhajani, "Klasifikasi Citra Pigmen Kanker Kulit Menggunakan Convolutional Neural Network," *Jurnal RESTI,* vol. 5, no. 2, pp. 379--385, 2021.

[38] E. Chen, X. Wu, C. Wang, and Y. Du,, "Application of improved convolutional neural network in image classification," *Proceedings - 2019 International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2019,* pp. 109-113, 2019.

[39] K. Singh, A. Seth, H. S. Sandhu, and K. Samdani, "A comprehensive review of convolutional neural network based image enhancement techniques,," *2019 IEEE International Conference on System, Computation, Automation and Networking, ICSCAN 2019,* vol. 1, no. 1, pp. 1-6, 2019.

[40] T. A. Korzhebin and A. D. Egorov, "Comparison of Combinations of Data Augmentation Methods and Transfer Learning Strategies in Image Classification Used in Convolution Deep Neural Networks," *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021,* vol. 3, no. 5, pp. 479-482, 2021.

[41] A. Titoriya and S. Sachdeva, "Breast Cancer Histopathology Image Classification using AlexNet," *2019 4th Int. Conf. Inf. Syst. Comput. Networks,* vol. 3, no. 9, p. 708–712, 2019.

[42] R. Watanabe, N. Soga, and H. Noji, "Novel nano-device to measure voltage-driven membrane transporter activity," *IEEE Trans. Nanotechnol,* vol. 15, no. 1, p. 70–73,

2016.

[43] S. M. Badawy, A. E. N. A. Mohamed, A. A. Hefnawy, H. E. Zidan, M. T. Gadallah, and G. M. El-Banby, "Classification of Breast Ultrasound Images Based on Convolutional Neural Networks - A Comparative Study," *2021 Int. Telecommun. Conf. ITC-Egypt 2021 - Proc,* vol. 4, no. 1, pp. 7-12, 2021.

[44] Farhadi, Akram and Chen, David and McCoy, Rozalina and Scott, Christopher and Miller, John A and Vachon, Celine M and Ngufor, Che, "Breast cancer classification using deep transfer learning on structured healthcare data,," *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA),* pp. 277--286, 2019.

[45] Kassani, Sara Hosseinzadeh and Kassani, Peyman Hosseinzadeh and Wesolowski, Michal J and Schneider, Kevin A and Deters, Ralph, "Breast Cancer Diagnosis with Transfer Learning and Global Pooling," *IEEE,* pp. 519--524, 2019.

[46] C. M. Kim, R. C. Park, and E. J. Hong, "Breast mass classification using eLFA algorithm based on CRNN deep learning model," *IEEE Access*, vol. 8, no. 1, pp. 197312–197323,, 2020.

[47] H. C. Lu, E. W. Loh, and S. C. Huang, "The Classification of Mammogram Using Convolutional Neural Network with Specific Image Preprocessing for Breast Cancer Detection," *2019 2nd International Conference on Artificial Intelligence and Big Data, ICAIBD 2019*, vol. 1, no. 4, pp. 9-12, 2019.

[48] M. A. Anupama, V. Sowmya, and K. P. Soman, "Breast cancer classification using capsule network with preprocessed histology images,," *Proc. 2019 IEEE Int. Conf. Commun. Signal Process. ICCSP 2019,* vol. 1, no. 2, p. 143–147, 2019.

[49] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics Med. Unlocked,* vol. 20, no. 12, p. 100412, 2020.

[50] A. Mikołajczyk and M. Grochowski,, "2019 International Interdisciplinary PhD

Workshop," *IIPhDW 2019,2019 Int. Interdiscip. PhD Work. IIPhDW 2019,* vol. 4, no. 3, p. 117–122, 2019.

[51] F. Ertam, "Data classification with deep learning using tensorflow," *2nd Int. Conf. Comput. Sci. Eng. UBMK 2017*, p. 755–758, 2017.

[52] K. Jakhar and N. Hooda, "Big data deep learning framework using keras: A case study of pneumonia prediction," *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, vol. 6, no. 3, p. 1–5, 2018.

[53] H. Dalianis, "Evaluation Metrics and Evaluation," *Clin. Text Min*, vol. 1967, no. 2, pp. 45–53,, 2018.

[54] N. Arya and S. Saha,, "Multi-modal classification for human breast cancer prognosis prediction: Proposal of deep-learning based stacked ensemble model," *IEEE*, 2020.

[55] Sara Hosseinzadeh Kassani , Peyman Hosseinzadeh Kassani, Michal J. Wesolowski, "Breast Cancer Diagnosis with Transfer Learning and global pooling," in *ICTC 2019*, 2019.

[56] Hao-Chun Lu , El-Wui Loh ,Shih-Chen Huang, "The Classification of Mammogram Using Convolutional Neural Network with Specific Image Preprocessing for Breast Cancer Detection," in *IEEE*, 2019.

[57] Anupama M A , Sowmya V , Soman K P, "Breast Cancer Classification using Capsule Network with Preprocessed Histology Images," in *IEEE*, India, 2019.

[58] K. Nguyen, C. Fookes, and S. Sridharan, "IMPROVING DEEP CONVOLUTIONAL NEURAL NETWORKS WITH UNSUPERVISED FEATURE LEARNING Kien Nguyen , Clinton Fookes," *Int. Conf. Image Process*, vol. 4, no. 3, p. 2–6, 2015.