CancerCare: A Reliable and Secured Self-supervising and Interactive System Using Deep Learning

by

Pushpakali Khisa
19101241
Arisha Hossain
19101061
Anika Raihana
19101095
Alisha Khan
19101411
Tamanna Sultana Tonu
19101155

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
January 2023

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Pushpakali Khisa

Pushpakali Khisa
19101241

Arisha Hossain
191010611

Anika Raihana

Anika Raihana
19101095

Alisha Khan
19101411

Tamanna Sultana Tonu
19101155

Approval

The thesis/project titled "CancerCare: A Reliable and Secured Self-supervising and Interactive System Using Deep Learning" submitted by

1. Pushpakali Khisa (19101241)

2. Arisha Hossain (19101061)

3. Anika Raihana (19101095)

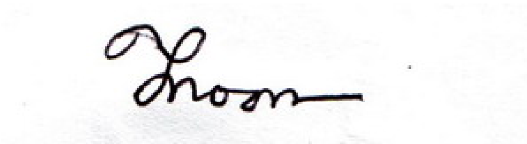
4. Alisha Khan (19101411)

5. Tamanna Sultana Tonu (19101155)

Of Fall, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January, 2023.

Examining Committee:

Supervisor:
(Member)



Jannatun Noor
Senior Lecturer
Department of CSE
BRAC University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam
Professor
Department
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Cancer is the ultimate global health issue in the 21st century, as its burden is increasing day by day. In the year 2020 [36], 18.1 million cancer cases were estimated, where 9.3 million were men and 8.8 million were women. Among these, many of the cases are detected at a very crucial stage due to a lack of advanced technologies to detect early symptoms, misinformation and ignorance. In recent years, the innovation of many healthcare systems are capable of contributing to raising awareness and providing significant assistance to both oncologists to detect cancer disease and patients, which are progressively surging in popularity in the medical sector. However, [15] manual detection of cancer cells from the histopathological image is a very tiring, time-consuming process for histopathologists and many human errors can occur. Therefore, many computer-based detection processes have been invented, giving better results than the manual detection process. Although several architectures have been introduced, it becomes a question of which architecture gives us the best result for detecting cancer cells. In this proposed framework, we have analyzed five deep Convolutional Neural Network architectures such as VGG16, MobileNetV3, InceptionV3, Xception, and DenseNet121, which have been trained and tested on the lung cancer and colon cancer datasets, present the performance comparison between them and found out the best image recognition and classification architecture which have given us the utmost accuracy for detecting any type of cancerous histopathological cell. Moreover, we have also designed a prototype for a user-friendly, self-supervising and reliable platform ("CancerCare" mobile application) with some key features after conducting a survey in Bangladesh to make it easier for oncologists as well as patients to deal with this fatal disease. Besides, it also perpetuates smoother communication between patients and oncologists on a regular basis via live chat and video consultation. At present times, misuse of data in mHealth applications is one of the most noted risks. Therefore, we have established authentication using firebase.

Keywords: Histopathological cell image, Deep Convolutional Neural Network, VGG16, VGG19, ResNet50, ResNet152, Xception, DenseNet12, Survey, Image recognition, Oncologists

Acknowledgement

Firstly, all praise to the Almighty Allah for whom our thesis has been completed without any major interruption.

Secondly, to our Supervisor Miss Jannatun Noor, for her kind support and advice in our work. She helped us whenever we needed help.

Thirdly, the two oncologists, Dr. Md. Golam Zel Asmaul Husna (Resident, Radiation Oncology, NICRH) and Dr. Saiful Alam, (Resident, Radiation Oncology, NICRH), who helped us throughout the whole research with their constant support. Also, A.B.M. Alim Al Islam, BUET, who gave us his valuable feedback and reviews on our work really helped us to learn a lot.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer, we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

ϵ	Epsilon

Chapter 1

Introduction

1.1 Cancer and It's Extremeness

Over the years, disease can be scrutinized as a matter of distress for the general people as well as a subject of interest for medical researchers. Unraveling the intricate riddles behind diseases and medicine has pushed us closer to the groundbreaking medical discovery. Scientific evolution such as the invention of modern medical technological tools, and the usage of sophisticated devices such as laptops and smartphones have made communication easier in recent years. As a result, the mortality rate of people suffering from crucial diseases like cancer has decreased significantly. However, people are still experiencing lifetime suffering, even death as complete recovery from cancer has not been possible yet. Despite this, suffering can be lessened if we can take crucial steps for the patients.

In the current world, where diseases have been spreading widely, the detection of cancer cells quickly in the human body has become a challenge to oncologists. Therefore, many image classification models have been proposed to detect cancer cells from histopathological image cells but they have not fully been implemented in the detection of cancer cells and as one model's performance varies according to the variation of cancer types, it has become a question which CNN based image classification model will perform best and give us best accuracy to detect the histopathological images of all types of cancer cells. Deathly illness like cancer is a very critical and complex disease itself where it can add extra burden when it becomes very time-consuming to detect the cancer cells in early stages and as patients face a lot of issues while communicating with their respective doctors such as sometimes patients can not understand the handwriting of their doctors and later when they cannot recall the suggestions of the doctors, it creates confusion among them. During emergencies, patients often use WhatsApp, imo, Email, Viber and many other social media platforms to consult with their doctors where the vital information shared by the doctors cannot be stored as well as sometimes doctors cannot keep track of their provided information. Patients over 60 years, which is considered the older age defined by WHO, will increase by almost two-thirds to reach 15.8 million by 2031 in a survey of the UK and it is a bigger challenge to care for the aging population rather than the younger one [7]. Therefore, they need proper sequential consultation during their diagnoses and immediate suggestions in the time of any emergencies. One of the most convenient approaches is a self-supervising mobile application.

Nowadays, mobile applications persist at the core of simulating development. In case, if anyone has a dilemma in this digital era, they look for a smartphone app to alleviate their inconvenience. Keeping pace with the rapidly evolving world, healthcare workers are continuously upgrading their scheme of rendering service to those who are in dire need of urgent treatment and medical assistance. Especially, young and adults who are diagnosed with cancer require continuous supervision, additional guidance and mental support. So far, hundreds of cancer-focused apps have been initiated. For example, "Bone and soft tissue tumors case studies (BoSTT)", "Personalized sarcoma care", "ABCanDroid" and many more have the potential to deliver receptive treatment in real-time by monitoring diverse symptoms and the physiological manifestations of any ailment. According to a study related to smartphone apps on cancer, the majority of these apps lacked the ability to upload modifications, reliability, track health complications due to the treatment, share treatment updates with family members or relatives, admin privileges, data integrity and so on [18]. Focusing on these ambiguities, we have stepped forward to design a mobile application which will integrate best-performed cancer-detecting architecture and could assist both clinicians and patients in establishing a communication bridge that will elevate expectations within patients. With a focus on reliability and sustaining the aspiration of cancer patients, the contents and instructions will be approved and advised by registered oncologists.

1.2 Problem Statement

According to recent studies done by WHO, it has been seen that nearly one out of every six patients diagnosed with cancer dies [28]. The United States of America is expected to see 1,918,030 new cases of cancer and 609,360 deaths related to cancer in 2022, with around 350 fatalities each day from which lung cancer is the main root of death [33]. Though there have been many innovations in the medical sector in the past decade to tackle this fatal disease, there still remains a lot of misinformation about cancer and most importantly communication gap between patients and doctors as well as the time-consuming cancer biomedical detection process such as the biopsy technique. It can mislead the patients about their diagnosis and self-supervising, which can subside the improvement of their health condition. While doing our research on this issue, we came up with a novel idea to design a mobile health (mHealth) application in such a way where we will implement the best-performing CNN image classification model where doctors or histopathologists will be able to detect any type of histopathological image of cancer cells. Our "CancerCare" application can also guide patients in self-monitoring and to ease their daily struggles. During 2013 and 2014, the worldwide usage of mobile phones surged from 406 million to 1.82 billion devices, while online activity rose by 81 percent due to which mobile application software in the medical field has exploded [14].

As cancer is a fatal disease, it requires intensive care and frequent monitoring. However, patients often face difficulty during their emergency because of traffic and transportation problems which can result in missing appointments. Later it gets tough to reschedule further appointments. Moreover, in some instances, patients forget to bring their necessary medical documents, or in some cases, they are confused about what documents they should bring which creates difficulty for the doctors in giving proper consultation and treatment. On the other hand, the reliability of the information in dealing with this condition can be a major issue because there is a lot of misinformation about it on the internet, which can lead to patients being misguided. For example, if patients experience some basic effects of cancer, such as nose-bleeding, vomiting, diarrhea, generalized weakness, difficulty in respiration, loss of appetite, insomnia and so on, they sometimes rely on the internet solution as a preliminary step or sometimes if they want to consult with the doctors about their problems through phone call or video call instantly, it becomes difficult for them to reach them. In order to tackle this communication gap, our research will provide smoother communication between patients and oncologists on a regular basis via video conversation, audio consultation and audio recordings of a live consultation. Moreover, patients can also easily book appointments through the app, where they do not have to go to hospitals or clinics.

In a survey, [10] carried out by UNESCO, the literacy rate in Bangladesh increased gradually from 59% to 74.7% during the year 2011 to 2019. However, many people still do not have proper knowledge about how to operate a smartphone. Keeping this in mind, our User Interface(UI) will be designed in such a way that it will be user-friendly and easily accessible by people (e.g. Whatsapp, Imo, Viber). For instance, a voice recognition feature, face recognition and fingerprint scanning will be added for people who cannot type so they can use their voice to login. Besides, our application will support both English and Bangla language so that people who do not know English can access the application. Furthermore, our application will be supported by iOS and Android operating systems. Therefore, people can access it from any device easily.

At present times, data misuse has become the greatest perceived risk of mobile health applications. In Germany, a web-based survey proclaimed about the misuse of data which was conducted among 206 orthopaedic and trauma surgeons. It identified the risk of data exploitation in the mHealth applications which is one of the most significant anticipated risks [19]. Therefore, many patients avoid using this type of software due to privacy concerns. To solve this issue, a private cloud will be installed for securing the confidentiality of patient information and data. Additionally, there are cases where accounts are being hacked and the data gets leaked. In order to solve this, we are using a two-factor authentication feature for securing the accounts of the users. While researching on this topic, we identified some of the above-mentioned problems that were causing stress as well as affecting the mental health of the patients. Since they are already undergoing a difficult chapter in their lives, we thought of minimizing some of their grief with the help of our proposed application 'CancerCare'.

1.3 Motivation

1.3.1 Motivation for Researching on CNN Models

As cancer is a very extreme kind of disease and people are dying every year because of this disease, it is very crucial to detect cancer as early as possible. For detecting the cancer cell, a biopsy test is conducted where the microscopic image of the cell or tissue, called the histopathological image, is tested by the histopathologist as well as human error. It requires a lot of time for the histopathologist to test the image manually [15]. Therefore, many CNN-based image classification models have been developed where many techniques have been used such as the ResNet101 model and have been complemented with a transfer learning model [30] for better performance. However, most of the studies focus on one type of cancer cell. For example, [30] a cancer cell detection-related paper, published on 22 January 2022 focuses on only breast cancer detection. Another journal, [15] published in 2019, they also focus on breast cancer detection. However, in Bangladesh, lung cancer and other types of cancer patients are increasing rapidly which we have found from our survey. Therefore, we have conducted our research on lung cancer and colon cancer and analyzed the better-performed convolutional neural network-based architecture among the five architectures after a solid analysis of their performance matrices. After that, our future plan is to implement the best-performed model in our "CancerCare" mobile application.

1.3.2 Motivation for Researching on Cancer Patients and Oncologists

From the day a person finds out they have cancer, their life changes completely. Each of their days passes with stress and agony living with this cruel illness. Moreover, it also drastically impacts the lives of their close ones as there is no complete cure or upright path that one will follow to get rid of the suffering. In recent times, a lot of technical improvements have been seen in the medical sector, from the collaboration between researchers and medical practitioners. All these motivated our study to deliver something that will slightly ease the difficulties patients have to go through regularly. Firstly, to get a better understanding, we consulted two oncologists. After exploring and researching different aspects, we have decided to curate and facilitate a survey where we planned to locate the types of problems both patients and oncologists face during the diagnosis. Moreover, we have also planned to analyze the interest and convenience of patients in using technology. Following this, we have interpreted the collected data and observed that most participants face various health-related and non-health-related issues during their diagnosis phase. Because of this, they are willing to use mHealth applications if it is designed in a way that will lessen their concerns. Based on the participants' feedback, we have modeled a mobile application prototype that includes suggested features from both patients' and oncologists' ends.

1.4 Research Objectives and Goals

- One of our research objectives is to find out the best performed deep neural network architectures for classifying and detecting histopathological images of cancer cells after the analysis of different image classification models, compile and fit them on different such as lung cancer and colon cancer histopathological image datasets.

- Another intent of this research is to identify the types of problems a cancer patient might face during the diagnosis. The problems are not only limited to health but also other factors, such as being unable to find a doctor during emergency situations, difficulty commuting to the hospital due to heavy traffic, transportation issues and financial crises. On the other hand, the study also aims to pin down problems from the oncologists' end.

- Additionally, our work also focuses on finding out whether there is any existing solution to tackle the communication gap, specifically the use of different communication platforms, for example, Whatsapp, Email, Viber etc., in Bangladesh. Moreover, our survey study aims to find how many respondents are comfortable using a smartphone and how experienced they are in doing so. Finally, this research will also determine whether patients and oncologists are willing to use a mHealth app in Bangladesh.

- Our main goal of this research is to highlight the problems faced by mainly cancer patients, such as communication gap, commuting to the hospital, difficulty in getting an appointment promptly, lack of proper knowledge about the illness they are suffering from and so on. By keeping these problems in mind, we would draw attention to the fact that technology can play a crucial role in solving these issues in the form of a mobile application.

Chapter 2

Background

2.1 Survey

A survey is a means of gathering information from a fragment of the population of a targetted group. The group of individuals are termed as 'sample'. A survey can be conducted in numerous ways depending on its data collection method. For instance, there is the remote-based survey, in-person surveys, and even in the form of telephones.

2.2 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a type of Deep Learning Algorithm and Network Architecture primarily used for data processing and image recognition. Its structure holds node layers mainly consisting of one input, output layer and one or many hidden layers in-between. Moreover, all nodes are interconnected with a relative weight and threshold. It is highly used in the field of oncology in recent times.

2.3 Softmax

Softmax is the activation function for deciding the neuron to be activated or not. In any classification problems, softmax is used to return the probability scores by converting the raw outputs. The equation is given below:

$$f(a)_i = \frac{e_i^a}{\sum_{j=1}^{n} e^{a_j}}$$

Here, a represents the raw output vector and j represents the number of neurons.

2.4 Fine Tuning

Fine-tuning, a more developed transfer learning model, uses the process of back-propagation to improve pre-trained knowledge of data of a CNN. It can mainly be used to modify the pre-trained CNN through tuning or tweaking and turning to a completely different dataset, which can be used to perform a similar function.

According to a study [13], fine-tuning can be applied to different medical imaging detection. It increases the accuracy of the model.

2.5 Confusion Matrix

In machine learning problems confusion matrices are used to measure the correct and incorrect values for solving classification problems. From the confusion matrix, we get to know 'TP' means true positive, 'TN' means true negative, 'FN' means false negative and 'FP' means false negative. TP values are positive values which have been classified accurately. TN means the values which are negative values but have been classified accurately. FN means the number of actually positive data classified as negative data and FP means the number of actual negative data classified as positive. From the confusion matrices, we can find the performance metrics of an architecture which are classification accuracy, precision, recall, and F1 score.
The formula for the accuracy is:
Accuracy=

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

The Formula of the precision for finding the true positive value is:
Precision=

$$\frac{TP}{TP + FP}$$

The formula of the recall for finding the actual positive value is:
Recall=

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

The formula of the recall for finding the actual positive value is:
F1 Score=2$\times \frac{Precision \times Recall}{Precesion + Recall}$

2.6 SPSS Software

SPSS is basically a windows based program that can be utilized for data entry, analyzing those data, and constructing graphs and tables. Large databases and files can be processed by SPSS. Besides, it can also undergo all of the analysis discussed in any literature in addition to many other operations. However, different options to adjust existing parameters, perform statistical studies, and compute descriptive and inferential statistics can be done using this software [9]. Also, it allows showing data that is categorized by a particular variable by choosing specific instances for further research. Many forms of graphs, such as line graphs, box plots, bar charts, histograms etc. can be produced.

2.7 Prototype

An original design of a proposed system is called a Prototype. Usually, in terms of mobile application development, prototypes are a working model designed to help software developers create the app swiftly. On the other hand, a hardware prototype is nothing but an analysis of whether or not the proposed system meets the required specifications.

Chapter 3

Related Works

3.0.1 Deep CNN Architecture in Medical Image Classification

Cancer is a very vast and critical disease in the 21st century that nearly causes the death of 10 million people. Approximately 30% of cancer cases [28] are identified in low and lower-middle countries such as Yemen, Afghanistan, Bangladesh, Uganda, and many more, which are caused by the infectious diseases Human Papillomavirus (HVP) and hepatitis [37]. According to a cross-sectional survey of Ghanians, it was observed that most of the people diagnosed with cancer basically wanted to have three pieces of information about the disease [26]. Among them, 90.6% of the population wanted to know the name and type of cancer, 92% wanted to know if the cancer is curable or treatable, and 88.3% wanted to know about all the possible treatments [26]. However, the most challenging part of cancer treatment is building better and smooth communication between doctors and patients. To solve this issue, a lot of mobile applications have been developed, and surveys have been conducted for nearly a decade. Some of them are related to raising awareness and delivering information about cancer, and auto-recording consultations between doctors and patients.

According to a research conducted by IARC (International Agency for Research and Cancer), more than 13,000 women in Bangladesh are diagnosed with breast cancer and it is causing the death of more than 7000 women per year [23]. The reason behind this massive rate of death is not detecting cancer on an early basis. Therefore, an android application, "ABCanDroid" has been proposed with the knowledge of Convolutional Neural Network(CNN) which is integrated with transfer learning to detect the early breast cancer which gained 99.58% accuracy [30]. It is a prompt and reliable cloud-distributed android application for the diagnosis of breast cancer where doctors upload a histological picture and get the result easily with the existing methods such as biopsy and biomedical imaging techniques. These methods require radiation which can be harmful to the human body and also requires time and well-experienced oncologists [30]. The transfer learning model has been trained with 15,616 images and 3904 images are used for testing purposes for this proposed work [30]. However, for their future work, they have proposed to use blockchain for the security of the data and it can also be made suitable for low-power IoT devices [30].

According to a research in Kazakhstan (Abduokhapov, 2022) [27] the most common type of cancer is Breast Cancer with cases over 2.2 million reported in 2020. To overcome this matter a group of classified malignant and benign histopathological pictures were collected and pre-trained with SGD optimizer and 12 epochs of 64 batch size under vital deep learning models like Efficientnetv2, Mobilenet-v2, Resnet V2- 50, VGG16 and many more. Among these the most accuracy was shown by Efficientnetv2 (94.5%), Resnetv2-50 (94.075%), and VGG16 (93.78%).

Another study in India (Subramanian et al., 2022) [34] presented a rare yet consequential type of skin cancer known as Melanoma where datasets from different sources were used to train Visual Geometry Group (VGG16) model, a deep learning model along with Convolutional Neural Network to improve the treatment of skin lesions. Moreover, they implemented confusion matrix to represent the performance of the classification process done by the classifier model where the accuracy of the classification process was significantly increased from 81% to 85% when transfer learning technique was used in Deep Convolutional Neural Network. The researchers also explained how the VGG16 model is formulated with Convolution and max pool layers and sigmoid function to get better results like an accuracy of 85.45%, a sensitivity of 90%, and a specificity of 81.66%. However, to avoid overfitting which can be a significant issue the researchers want to focus on using a broad dataset along with fine-tuning the hyper-parameters of the classifier model for future works.

The research study based in Turkey (Abdullah, 2022) [31] explained that the death rate of breast cancer can be reduced if it is diagnosed at an earlier stage. It can be detected by analyzing the abnormality in human breast tissues displayed in mammogram images. Now to differentiate between the Malignant and Benign images VGG16 and ResNet50 are used which are considered two of the major Convolutional Neural Network models. Now talking about the models it is seen that VGG16 was created in 2014 by Simonyan & Zisserman which comprises 13 convolutional layers with 5 max-pooling layers, 2 dense layers and one output layer whereas ResNet50 which is used for training large datasets is designed and originated by He et al. which comprises 50 deep layers like pooling, residual, dense etc. Both models also take square RGB images of size 224. Therefore after the experiments were completed the results proved that VGG16 shows better accuracy of 80% than ResNet50 which shows 60% accuracy. Lastly, by involving the TL concept and preserving the original structures of the VGG16 and ResNet50 models the researchers have also developed a clinical diagnostic tool.

In 2022, A cancer study in Turkey (Toprak et al., 2022) [35] stated that in the USA it was predicted 31% which is one-third of cancer patients will be diagnosed with breast cancer. Similarly, in Turkey, the rate of breast cancer prognosis is anticipated to increase by 0.5% every year since 2000. Since then radiologists utilized a lot of CNN-based computer-aided tools to ease the whole procedure of the diagnosis. Moving forward with implementing new theories a lot of research was successfully presented using the U-Net and VGG16 approach. The researchers proposed a methodology where images were processed followed by mass segmentation on U-Net + VGG16 architecture. Then thresholding is applied with an experimental approach. Finally, the values are compared and more accurate cancerous lesion segmentation was achieved.

A paper based on India (Deshmukh et al., 2021) [21] stated Breast cancer to be the second most common cancer among women globally. On the other hand, there have been a lot of advancements recently in the E-healthcare sector which tells us how people are willing to have remote consultations with doctors. Correlating all these, this paper proposes an android mobile health application which will be used to diagnose breast cancer and also help patients to enhance the quality of their treatment. Another key purpose of creating the app is to help radiologists distinguish between benign and malignant tumors. Now coming to design the researchers used java, various ML prediction techniques and algorithms to classify the tumor and cancerous cells accurately. Moreover, they also discussed an existing mobile-based health application called Consult-AI which helps to predict diseases by using Support Vector Machines (SVMs). But however, there was a disadvantage in the algorithm that it is mainly effective in cases where the number of dimensions is more than the samples. So in this paper, the XGBoost classifier was implemented instead of SVMs in Wisconsin Diagnostic Breast Cancer(WDBC) dataset obtained from the UCI depository and it showed significant accuracy of 98.24%, credited to its fast computation speed and efficient performance. In 2017 [11], another convolutional neural network model has been proposed, named "Xception" meaning "Extreme Inception" trained on ImageNet, which gave 0.945 accuracies where VGG16 model gave 0.901, ResNet152 gave 0.933 and Inception V3 gave 0.941 accuracies in their analysis.

For the last few years, a lot of mhealth applications were built for analyzing skin cancer using different deep learning models. Likewise, this study in the Netherlands (Sangers et al., 2022) [32] also executed a study introducing a CE-marked mHealth app in Europe, Australia and New Zealand for both iOS or Android devices in 2020 for skin cancer patients using CNN. They compared the results with histopathological diagnosis by evaluating the sensitivity and specificity of the app to detect premalignant and malignant skin lesions. However, the iOS device showed greater sensitivity(91.0%) than the android one (83.0%). Moreover, they also performed a cross-sectional study among patients aged greater than or equal to 18 years to estimate diagnostic accuracy at Albert Schweitzer Hospital of the Netherlands.

In another study [19] about android applications, two technical possibilities are discussed for the detection of tumors which can later be spread as cancer. One of the techniques explained the usage of a phone's camera to identify a tumor with augmented reality. Another one is about developing a musculoskeletal tumor for post-surgery patients. While doing the research, they have identified two more applications used by health professions "Bone and soft tissue tumors case studies (BoSTT)" and "Personalized Sarcoma Care". These applications were designed to help patients in decision-making and determining the survival rate for dealing with sarcoma cancer.

3.0.2 Contribution of mHealth Applications in Cancer Patient Care

As reported in a survey done by the IARC, the prevalence of Colorectal cancer is about 4.2% in males and about 3.3% in females globally [17]. The paper [5] introduces a smart application which helps in the early screening of high-risk groups of potential Colorectal cancer patients. The application mainly uses data mining using a decision tree algorithm. In this paper, the application mainly consists of 3 models grounded on a decision tree algorithm, the Simple model (75.1% accuracy), the Intermediate model (85.8% accuracy) and the Best Model (88.0% accuracy). Patients input their data into the app, based on the 3 models, and the app would suggest whether or not they have a risk of Colorectal cancer. Moreover, if a patient was at high risk, the app would also refer them to necessary links of medical services to prevent complications. The study also shows that their proposed model gives a more accurate result than FOBT (Fecal Occult Blood Test), with only an accuracy of 45.5%, which is the traditional way of screening Colorectal cancer. Although the study showed quite successful results in early detection screening, the only limitation was that the study was diagnostic-based rather than population-based.

Recent app evaluation on Prostate cancer [25], shows that during cancer therapy, patient-reported outcomes were obtained using an app called the Interaction app. It gives a statistical analysis of self-care and literacy level on health. A significant number of prostate cancer patients had minimal health-related education throughout the radiotherapy procedure, which improved with the app's help. While assessing the health literacy, they performed the analysis using two methodologies, among which one is the Swedish Functional Health Literacy Scale (FHL), and the other is the Communicative and Critical Health Literacy Scale (CCHL) [25], whereas the statistics of the intervention and control groups were compared using Chi-square analysis. According to one of the studies on the app's benefits, intervention groups reported lower stress symptoms than emotional states. Additionally, the patients who used an app to report and manage symptoms successfully improved their literacy and skills related to healthcare [25].

Amid the top five cancers in the world, cervical cancer had more than 342,000 deaths in 2020. Among these, 90% of the death cases reportedly turned out to be from economically developing and underdeveloped nations [29]. Even in Bangladesh, about 50 million women are prone to enduring cervical cancer [22]. Despite the fact that cervical cancer can be treated if discovered at the primary stage and screening helps to decrease the mortality rate, follow-up screening has to be done in order to avoid the recurrence of cervical cancer. A study [20] was done to observe how mobile health interventions, such as text messages, emails, phone calls, etc., improve the screening rate of cervical cancer, and the results show that text messages and phone calls are a viable way of reminding patients. According to their review, screening rates increased by up to 17% due to mHealth interventions. Although there was a concern about whether or not the patient received those reminders due to inoperative phone numbers, poor network connection, as well as being technologically challenged. The study [20] concluded that these mHealth follow-up screening reminders might be favorable for a success rate if they are done every 3-year interval rather than as a one-time intervention.

According to a recent study [16], it was seen that a lot of patients want to auto-record their live consultations with doctors. They want to do this so that they can listen to it later for a better understanding of the medical procedure and also share it with other members of the family for making further decisions. But one of the main drawbacks of doing this is failing to provide security for medical information, as it can get leaked easily and hamper the privacy of clinicians. By keeping all these in mind, they decided to develop a hospital-authorized iOS application named "SecondEars", which will allow patients to audio-record the consultations while maintaining all the possible measures to secure medical data and information. They started the developing process by conducting six co-design workshops within three months where the research team involved stakeholders, app designers, patients, consumers, oncologists, and staff from the medical IT department to brainstorm and discuss different mythologies about how they could build a simple, secure, user-friendly, useful app. In the workshops, they used different kinds of techniques like journey mapping and the MoSCoW method to deduce important features of the application. After that, a wireframing of the app was done by the app designers, which led to a final prototype after various feedback review sessions from the research team, co-design team, consumers and other attendees of the workshops. Finally, after an acceptance test, "SecondEars" was introduced on the iOS App Store in Test Flight mode. However, to cut down further costs, the research team made two decisions. Firstly, instead of automating the app's connection to the medical records, a cloud-based model 9 named Amazon Web Servers was used to securely save medical records until the app had been officially operated by the clinics. Secondly, before subsidizing an Android version, the app was initially created on iOS as it is more expensive to develop for its complex structure. To ensure further medical and legal reasons, a sample of the audio documentation will be stored in each individual's medical record by default. To conclude, it has been said that no similar apps have been developed before, so the app should include other advanced interface options, such as translation and text-to-voice options, in future versions [16].

Over the last decade, technology has advanced a lot, changing our lives with the help of various mobile applications. Even in the healthcare sector, a vast number of innovative mobile health applications have been made, which are termed as mHealth apps, limited not only to cancer care or cancer detection but also to curing diabetes, mental health, fitness tracker etc. However, a question remains whether or not patients are in favor of using these applications. A research [12] was conducted in Germany to study patients' acceptance of various cancer care applications, and the results showed that around 48.5% of the participants were in favor of sending their medical data/history through an app. Although 43.5% of the participants were not willing to share their medical data due to privacy concerns, not having smartphones as well as lack of technical skills.

One such mHealth app, named "Healthier Together," was designed to target the minority population, mainly the non-Hispanic Black recently diagnosed cancer patients. The main objective of the paper [24] was to assess the users' acceptability and their involvement in the app. Amid numerous features in the app, the user can set as well as modify a weekly goal among exercise, diet, and weight tracking along with the discontinuation of smoking. Most importantly, users can get reminders and can keep track of their goals. The study was conducted for two months, and the outcome showed that the targeted group found the app to be reliable, with a System Usability Scale (SUS) of 87. Furthermore, on average, the participants successfully completed their goals in 5.1 out of the 8-week survey. However, the failure rate was 6.7%. At the end of the research, an exit survey was conducted in the form of Likert scale statements focusing mainly on the smooth use of the app and its information delivery.

In the treatment of many health-related problems, diseases and their behaviors, smartphone treatments have played a potential role in this sector. The outcomes of a comprehensive inspection of health care therapies which are delivered via mobile phones with necessary features such as texting, messaging and so on, have improved by 61% [1]. The trend in healthcare therapies has been enriched due to the greater presence of patients in medical appointments, better treatment, and enhanced diagnosis. Significant improvements in the clinical sector, such as enhancements in blood sugar management, self-efficiency, and asthma symptoms, were possible as mHealth applications played a crucial part in it.

Facilities provided by mHealth applications, such as sending messages via text, cameras for scanning documents etc., are not well acknowledged by the patients. The paper [1] also reviewed several apps related to cancer which are accessible to the common people, among which apps on Breast cancer and general cancer are prominent. A large percentage of mobile applications (64.1%, 189/295) did not imply their affiliation. There were 594 researchers analyzed in the medical literature, but none of them explored a cancer-focused mobile phone application. According to [1], numerous cancer-focused applications have contributed to improving behavior change, tracking a range of biological traits and symptoms of cancer, and delivering therapies, all while being simple and inexpensive. However, there is insufficient evidence to support their effectiveness or safety. In the future, the evidence base should be strengthened and integrated into a public whitelist.

There was a study [6] done particularly on breast cancer in which it was shown how patients and medical professionals learn and understand various health information related to the disease. Moreover, it also explains the interest of patients in using the existing health information-sharing tools with the assistance of oncologists. To begin with, the researchers noted down 23 different health factors and evaluated them by carrying out patient interviews and surveys correlating them with the methods breast cancer patients and doctors use to share health-related information. Following that they also planned on how to improve the ways of sharing information by picking out four design implications. They mainly focused on the usage of Personal Health Tracking Tools (PHRs) that provide lots of benefits to the patient such as providing them with a sole place where they can collect and know about their medical history. On the other hand, this tool fails to provide ownership of data to the respective users. However, the patient surveys and interviews provide an overview of some of the most important health factors such as biometric factors(height, weight), health factors(nausea, diarrhea etc) and emotional factors(stress, anxiety etc) which will ultimately help to reduce the challenges and promote the tools in creating an impact.

A research addressed how a person has to adapt to many changes once they get diagnosed with cancer [4]. They have to go through a chain of medical networks coordinating with doctors and healthcare providers to get through proper treatment. This takes a huge toll on their lives starting from mental strain to financial crisis and many more. Now to lessen the burden of them many healthcare and government organizations have merged to establish independent cancer navigation programs that will serve one to one medical, financial and other support to patients from the beginning to the completion of their treatment. Therefore, the navigation programs are designed focusing on five vital themes like duration, resource, knowledge, management and development. For each of them, a rigorous study was operated to develop strategies which will minimize the challenges and create optimum results and scope for the future.

A paper [3] described the ways in which cancer research data is mostly collected following a certain structure of guidelines. For instance, most data is gathered from a group of people in the form of a survey form which includes certain sets of questions particularly related to the topic. The questions are generally kept simple for the participants to answer them easily and quickly. Moreover, similar types of questions are distributed among distinct populations regardless of gender as the results portray quite similar information which ultimately saves the time from creating different datasets. Another definite nature of data is that it can never be duplicated as they are collected from a specific population at a definite time and location. Another important aspect to keep in mind during any kind of medical data research is to take permission and follow the proper guideline from a senior researcher before beginning the actual proceeding.

A research [2] carried out in Eastern England at a local hospital portrayed the understanding of patients regarding the confidentiality of their privately shared health-related data to benefit research in the respective medical field. The whole concept of maintaining a strong ground in protecting patients' private data has been one of the top prioritized criteria in the healthcare world. Therefore, by observing all these regulations this study planned to further encourage humans towards the complete process of data security by giving them an effective perspective of the existing design of various technological solutions which is used for consent management and handling sensitive data.

Chapter 4

Methodology

4.1 Research Plan

Nowadays, cancer patients are increasing day by day and patients have to go through various obstacles during their diagnosis period. Even people are facing death because of cancer. Considering this current situation of the patients, we were inspired to develop an idea with the help of modern technology which can help cancer patients to get proper treatment as well as doctors to give them consultation easily and effectively. We planned to propose a prototype of a mobile application "CancerCare" which will assist patients and doctors during diagnosis time as well as it will help to detect the cancer cell accurately without any human error.
Our whole research plan is given in figure 4.1

4.1.1 Working on Architectures

Firstly, we have collected histopathological images of lung and colon cancer datasets from a reliable source. Then, we pre-processed our image data which included image resizing to 224x224x3, augmentation and normalization. After that, we have splitted into training with 80% data and testing with 20% data. We have trained and tested both datasets on the five architectures which are VGG16. MobileNetV3, DenseNet121, Xception and InceptionV3 and find the accuracies of all the architectures. We have handled any overfitting that has occurred during the training of these five architectures on the two datasets. Finally, we have analyzed and compared all of the architectures' average accuracy to find out the best-fitted and performed architecture and add the detection of cancer histopathological cell feature in our proposed "CancerCare" mobile application prototype.

4.1.2 Meeting With the Oncologist

We conducted a meeting with two oncologists from Bangladesh; Dr. Md. Golam Zel Asmaul Husna (Resident, Radiation Oncology, NICRH) and Dr. Saiful Alam, (Resident, Radiation Oncology, NICRH) to share our ideas with them. They appreciated our idea and confirmed us that if there is a mobile application which will assist patients to contact doctors during emergencies instantly, help remote and rural patients to get a consultation with any oncologists in Bangladesh via video chat, live chat, get appointments and so on, then it will absolutely assist both patients

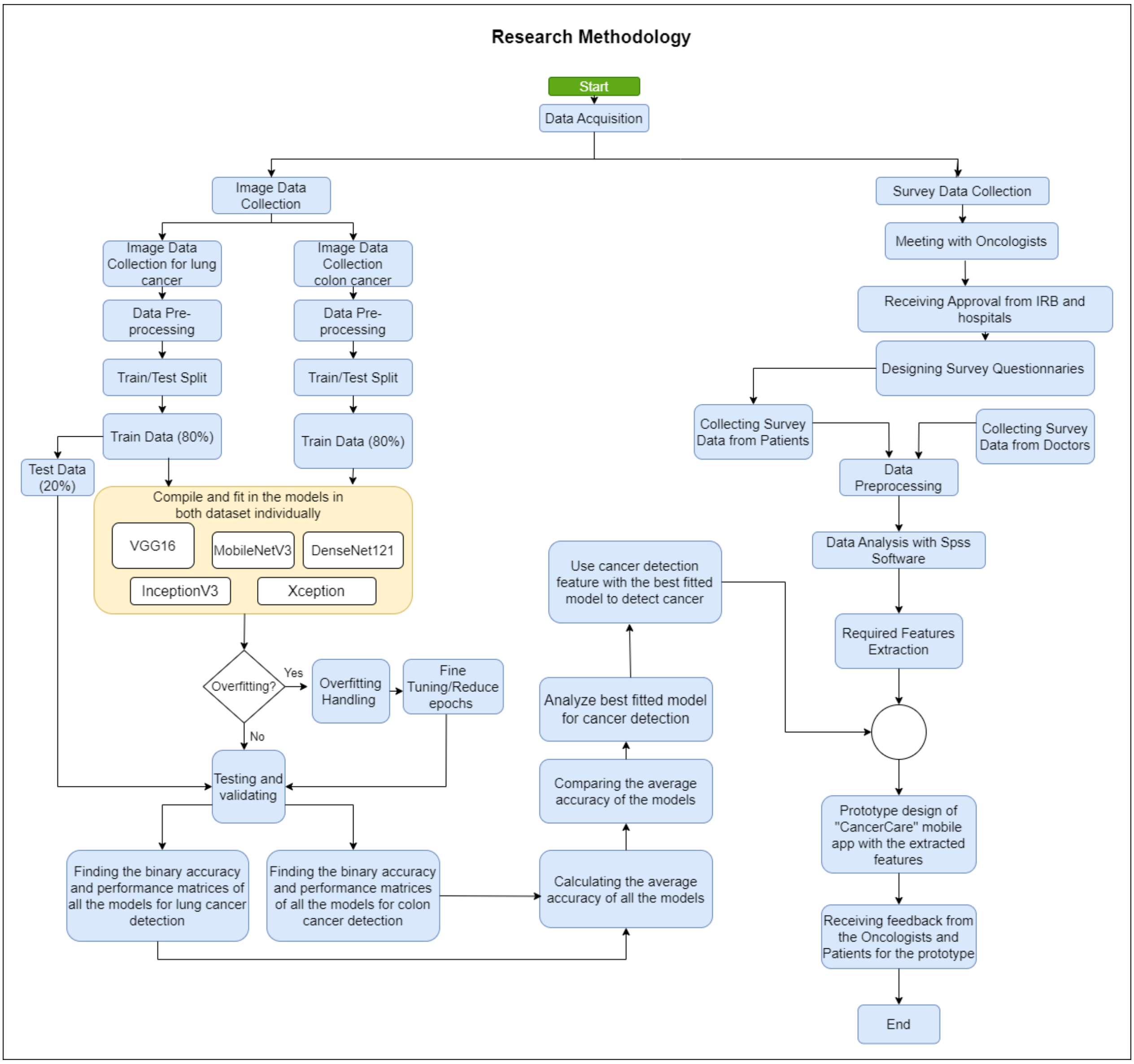


Figure 4.1: Work plan flowchart

and oncologists. They also said that, *"We want a system where only patients and oncologists will be able to access"*
After hearing their thoughts, we processed them according to our plan.

4.1.3 Receiving Approval From IRB and Hospitals

Surveying was one of the most challenging parts of our Research. Cancer is a very sensitive health issue which needs to be handled delicately. We needed some real-time data from the patients and oncologists about their every day facing problems to communicate with each other, to know if they agree with our idea of a mobile application, and what types of features they want to have in the application. Therefore, it requires several criteria to be fulfilled to collect data such as rules and conduct of communicating with patients, privacy concerns, their emotional, mental as well as psychological conditions and so on. As we had to conduct research involving interaction with individuals, the first and foremost task was to seek formal approval from the IRB as well as hospitals; "Bangladesh Cancer Society Hospital and Welfare Home" and "National Institute of Cancer Research and Hospital (NICRH)". We had to submit an IRB form about our motivation, intention, research work and questions. It took two months to get approval for collecting survey data from the patients and oncologists. After that, we started designing our survey questionnaires.

4.1.4 Designing Survey Questionnaires

As cancer disease is a very delicate issue, we had to give extra effort to make survey questions in such a way that we could collect unique and required information from the patients and doctors. After receiving some ideas about survey questions, we made an initial draft of survey questions and shared those with oncologist Dr. Md. Golam Zel Asmaul Husna (Resident, Radiation Oncology, NICRH) and Dr. Saiful Alam, (Resident, Radiation Oncology, NICRH) for their feedback. After getting the feedback, we moderated the questions and again received feedback from A. B. M. Alim Al Islam and made changes to our survey questions according to it. Afterwards, we started our survey data collection.

4.1.5 Conducting Survey With Patients

We went to the hospitals and collected in-person data from the patients. We talked about their communication problems, health emergencies, financial conditions and so on. We collected a total of 72 data from the patients as most of the patients were not very much interested, and some of the patients were not in the condition to talk.

4.1.6 Conducting Survey With Oncologists

For collecting data from oncologists, we give the survey data form to our collaborator oncologists. They helped us to distribute the survey questions to the other oncologists in Bangladesh and from them, we collected a total of 44 responses.

4.1.7 Initial UI Design of the Mobile Application Prototype

Our UI flow chart maps out the process of our proposed mobile application prototype. Here, the represented UI that we thought initially is given in figure 4.2.

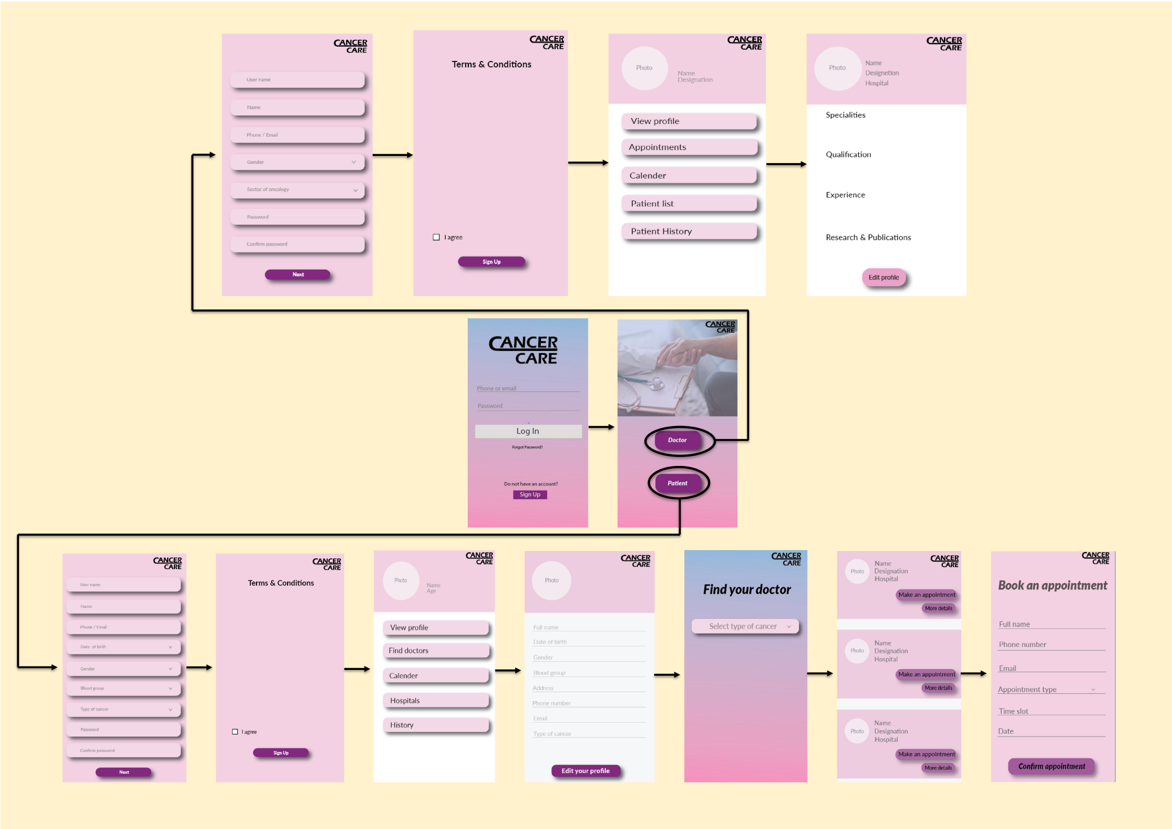


Figure 4.2: Flowchart of the UI

Chapter 5

Data Analysis

As we are researching the cancer histopathological images classification architectures for detecting cancer cells as well as cancer patients and oncologists, our datasets are divided into two types which are:

- Cancer Histopathological Image Datasets
  (Lung Cancer dataset and Colon Cancer dataset)

- Survey Datasets
  (Patients Dataset and Oncologists Dataset)

5.1 Cancer Histopathological Image Datasets Analysis

We have collected histopathological images of lung cancer cells and colon cancer cells from a reliable source: kaggle [38].
The lung cancer dataset has been classified into two classes which are Lung begin tissue which are normal cells, and lung adenocarcinoma cells which are infected cells. There are a total of 10,000 images in this dataset. We have split the data between the train with 80% data and the test with 20% data. Sample histopathological images are given below:

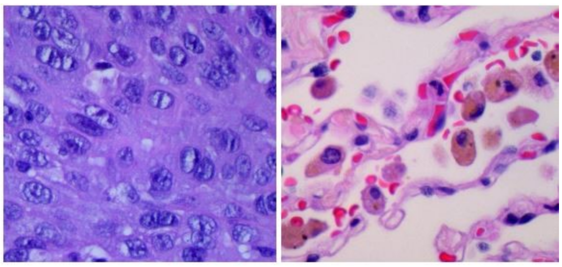


Figure 5.1: Sample histopathological image data of normal lung cell and cancerous cell

In figure 5.1, the left image is the beginning or normal lung cell histopathological image and the right top picture is the lung adenocarcinoma cell histopathological image.

The splitting of the dataset is illustrated in table 5.1.

Classes	Training Set	Testing Set	Total
Lung Beginning Tissue or Normal Cells	4,000	1,000	5,000
Lung Adenocarcinoma cells	4,000	1,000	5,000
Total	8,000	2,000	10,000

Table 5.1: Lung cancer histopathological image dataset classification in tabular format

The colon cancer dataset has been also classified into two classes which are colon begin tissue which are normal cells and colon adenocarcinoma cells which are infected cells. There are a total of 10,028 images in this dataset. We have split the data between the train with 80% data and the test with 20% data. Sample histopathological images are given below:

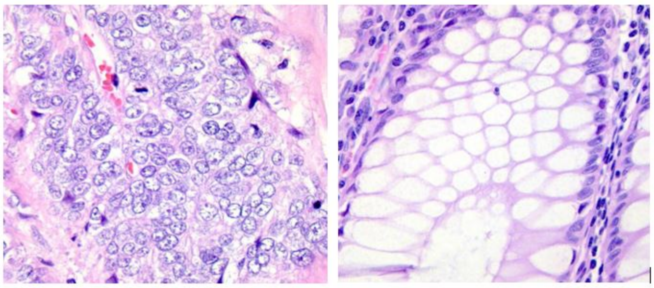


Figure 5.2: Sample histopathological image data of colon cancer cells

In figure 5.2, the left image is the beginning or normal colon cell histopathological image and the right top picture is the colon adenocarcinoma cell histopathological image.

The splitting of dataset is illustrated in table 5.2.

Classes	Training Set	Testing Set	Total
Colon Beginning Tissue or Normal Cells	4,000	1,000	5,000
Colon Adenocarcinoma cells	4,028	1,000	5,028
Total	8,028	2,000	10,028

Table 5.2: Our colon cancer histopathological image dataset classification in tabular format

5.2 Image Data Pre-Processing

Before analysis, we have pre-processed our data which are described below:

5.2.1 Image Resizing

We have analyzed VGG16, MobileNetV3, DenseNet121, Inception and Xception. Before training, testing and validating the models, we have resized the data into 224x224x3 RGB images using Tensorflow and Keras.

5.2.2 Image Augmentation

The image augmentation process has been used to improve the accuracy of the result. Therefore, ImageDataGenerator has been imported from TensorFlow and Keras. We have done random horizontal and vertical flipping as well as random rotation of each image with fill_value 0.2.

5.2.3 Normalization

We have normalized images by scaling from (0,255) to a range (-1,+1) as we have loaded pre-trained ImageNet weights in our base model.

5.3 Survey Data Analysis

Through surveys, a wide range of information was gathered which was analyzed using a variety of methods.

5.3.1 Sample Data From Survey

Through the process of data collection, we managed to collect 72 responses from patients' survey questionnaires and 44 from doctors' survey questionnaires. For simplicity and easier data processing, we have divided the survey data into two parts labelled - Patients' Data and Doctors' Data. We did our preliminary analysis based on the collected data.
Figure 5.3 and 5.4 represents a sample of how the collected data is stored in their respective excel file.

Field of Oncology	In which medical sector are you in?	Have you faced any problems while communicating with the patients?	Types of problem faced
Radiation	Government	Yes	• Patient could not come timely during the appointment time • Patient forgot to bring the necessary reports • I couldn't explain the treatment scheme
Radiation	Private	No	• None
Medical	Government	Yes	• Patient could not come timely during the appointment time • Patient forgot to bring the necessary reports
Other	Government	No	• None
Radiation	Private	Yes	• I could not come at hospital in time because of traffic jam or transportation issues

Figure 5.3: A sample of the doctors' survey data

At which age (years) were you diagnosed with cancer?	What type of cancer have you been diagnosed with?	Have you faced any problems while communicating with the doctors?	Types of problem faced
Above 64	Lung	No	• None
Under 18	Brain	No	• None
45 - 54	Kidney	Yes	• I could not timely come during the appointment time • I could not explain the problems properly
45 - 54	Cervical	Yes	• I could not timely come during the appointment time • I forgot to bring some necessary reports
18 - 24	Blood Cancer (Leukemia)	Yes	• I could not timely come during the appointment time

Figure 5.4: A sample of the patients' survey data

5.3.2 Data Filtering

The raw data that was collected through survey forms for doctors and patients were placed into two different .csv Microsoft Excel files and then renamed accordingly. The .csv files contained some redundant columns like Timestamp and names which were deleted.

5.3.3 Handling and Encoding Missing Values

The audiences adopted a variety of techniques to complete the survey questionnaires for which we used data filtering. A number of the submitted responses were written in Bangla some of which were transliterated and others were disregarded. In addition, a few unnecessary symbols, emoticons and punctuation marks, including full stops, colons, semicolons, and commas, were omitted. Responses that were too lengthy were also disregarded. Eventually, to get rid of the words that don't add significantly to a sentence, stop words were added. In addition, some of the important fields were unfilled. An alternative value has been used employing the imputation technique to address these missing variables. In order to fill in the missing values, the most frequent and mean values were placed. Since each of the responses contained categorical values, these were encoded using Label Encoder to numerical values.

5.3.4 Analyzing Survey Responses

Python was used to analyze the file after encoding and create graphs to display the results. To obtain a much more precise and high-quality graph, SPSS software was utilized. Numerous percentage bar graphs and pie charts were constructed using the survey data to highlight the relationships between various outcomes. These graphs that were produced, reflected some demographic characteristics, several information related to cancer, cancer patients as well as doctors.

Chapter 6

Researching on Deep CNN Architectures

We have researched five deep CNN image classification architectures, which are VGG16, MobileNetV3, DenseNet121, InceptionV3 and Xception from keras. We have trained, tested and validated two cancer histopathological image datasets such as the Lung cancer dataset and colon cancer dataset, on these five models to analyze and compare their performance parameters, such as binary accuracy, recall, precision and F1 score with each other. We have used fine-tuning to increase the model accuracy of the models. After that, the best-performing architecture has been presented.

6.1 VGG16

VGG16 is a Convolutional Neural Network based architecture which has 16 layers consisting of 13 convolutional layers and 3 fully connected layers. 224x224x3 RGB images have been passed through the convolutional layers, max pooling and fully connected layers with 3x3 filter size, 2x2 filter size for max pooling and stride 1. The architecture of VGG16 is shown in figure 6.1.

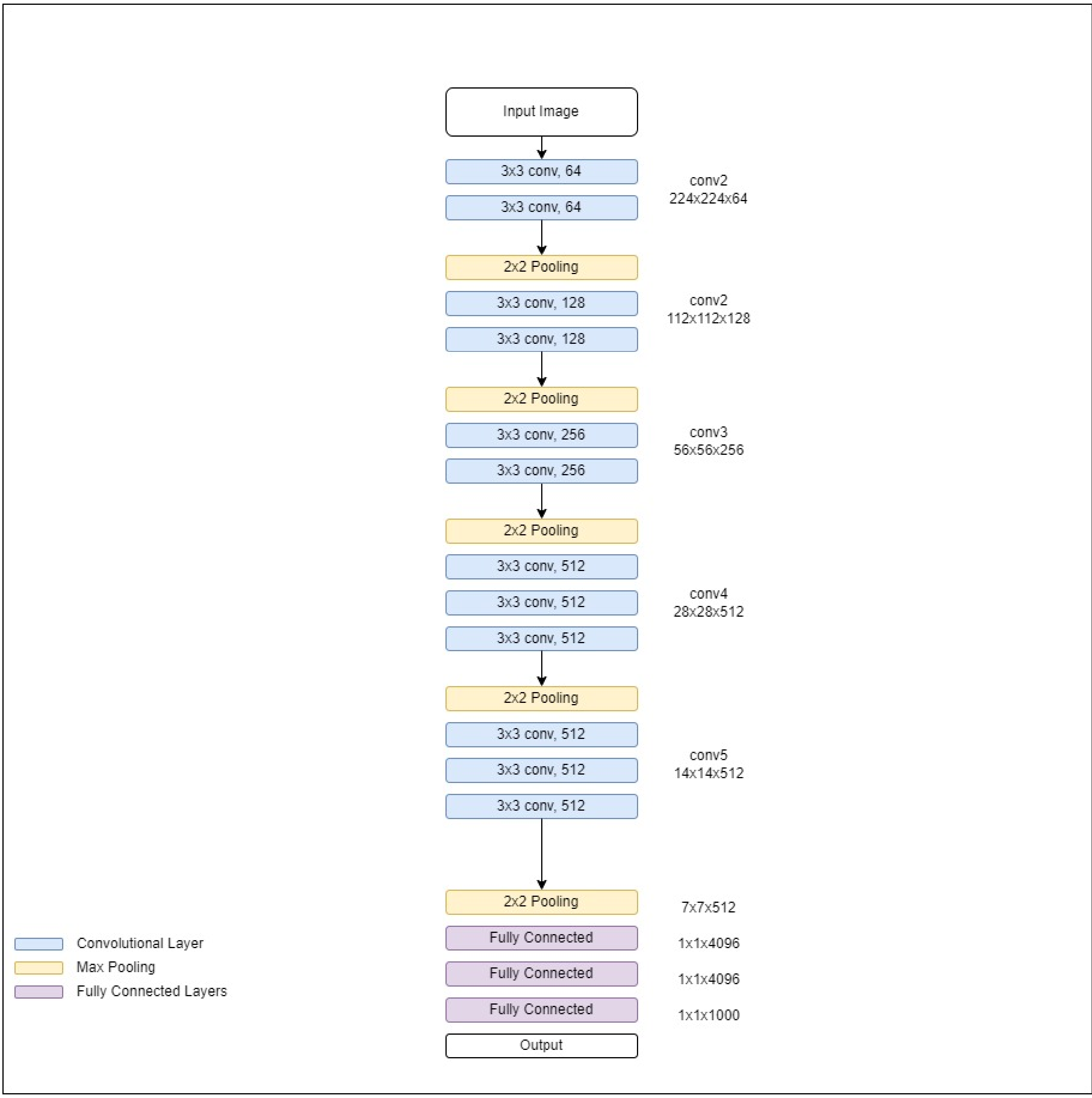


Figure 6.1: VGG16 Internal Architecture

Binary Accuracy, Loss, F1 Score of VGG16 are displayed in table 6.1.

For the Lung Cancer Dataset:

Binary Accuracy	Loss	F1 Score
98.44%	10.84%	98.91%

Table 6.1: Percentages of the performance parameters of VGG16 lung cancer histopathological image dataset

Model accuracy graph:

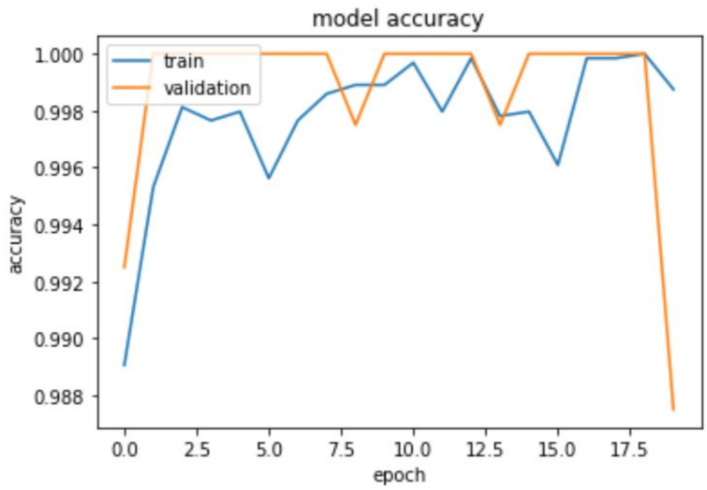


Figure 6.2: VGG16 accuracy graph of histopathological lung cancer cell dataset

Model Loss Graph:

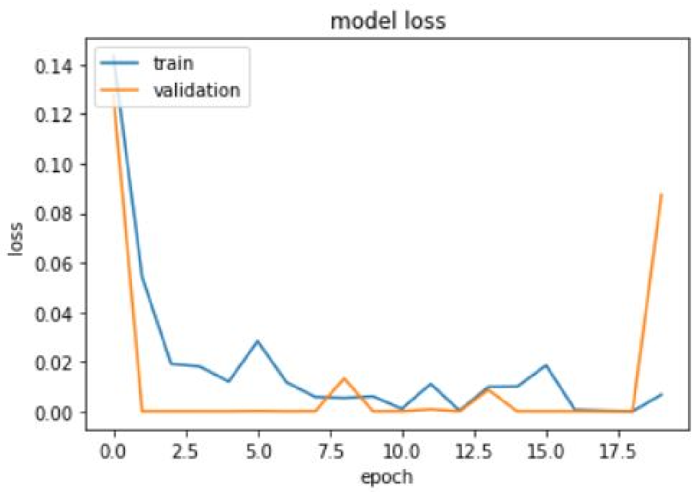


Figure 6.3: VGG16 loss graph of histopathological lung cancer cell dataset

Confusion Matrix:

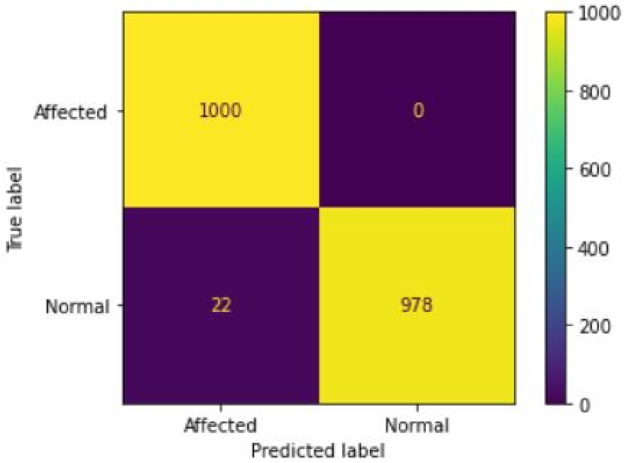


Figure 6.4: VGG16 confusion matrix of histopathological lung cancer cell dataset

For the colon Cancer Dataset:

Binary Accuracy	Loss	F1 Score
99.69%	1.2%	99.35%

Table 6.2: Percentages of the performance parameters of VGG16 colon cancer histopathological image dataset

Model accuracy graph:

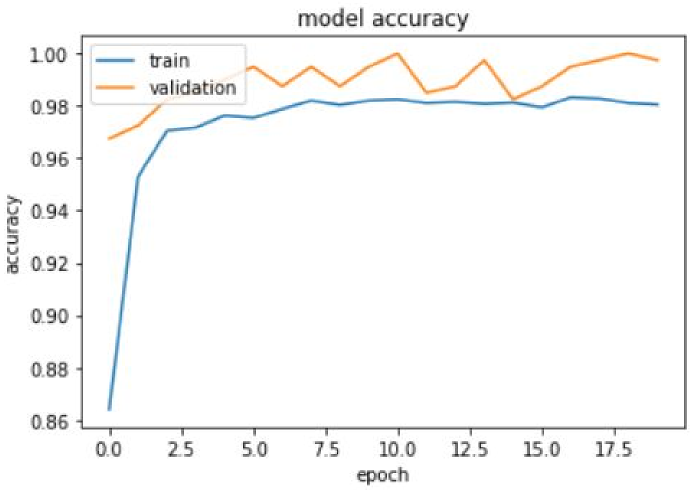


Figure 6.5: VGG16 accuracy graph of histopathological colon cancer cell dataset

Model Loss Graph:

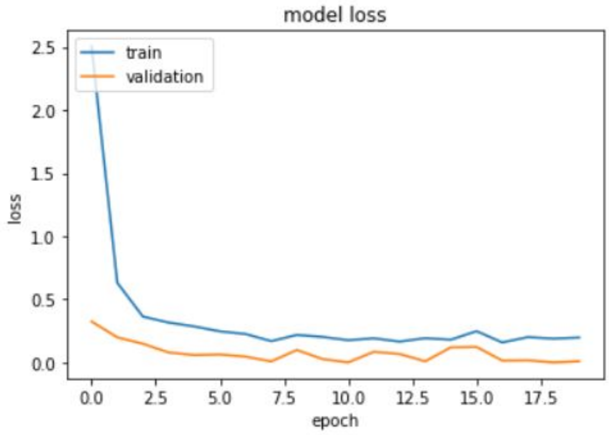


Figure 6.6: VGG16 loss graph of histopathological colon cancer cell dataset

Confusion Matrix:

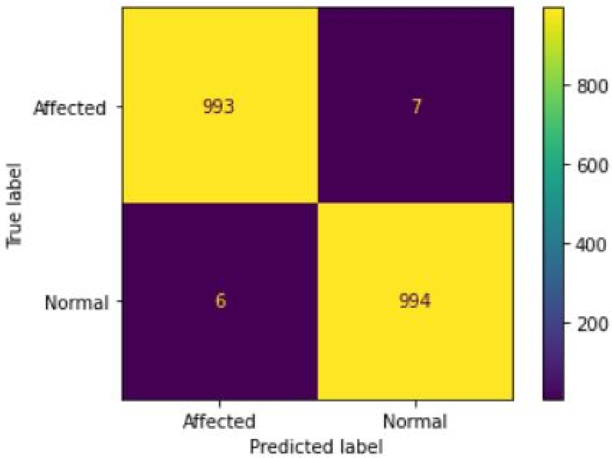


Figure 6.7: VGG16 confusion matrix of histopathological colon cancer cell dataset

6.2 MobileNetV3

MobileNetV3 has hard-swish, which introduces nonlinearity to improve the accuracy instead of ReLu. It has a squeeze and excite layer, depth-wise and point-wise convolutions. Firstly, the input image passes through the expansion convolution, and the output generated from it passes to the depth-wise convolution where the kernel size is 3x3. Point-wise convolutions are used to shrink the image back to its original size.
Binary Accuracy, Loss, F1 Score of MobileNetV3 are given below:

For the Lung Cancer Dataset:

Binary Accuracy	Loss	F1 Score
99.06%	1.16%	98.91%

Table 6.3: Percentages of the performance parameters of MobileNetV3 lung cancer histopathological image dataset

Model accuracy graph:

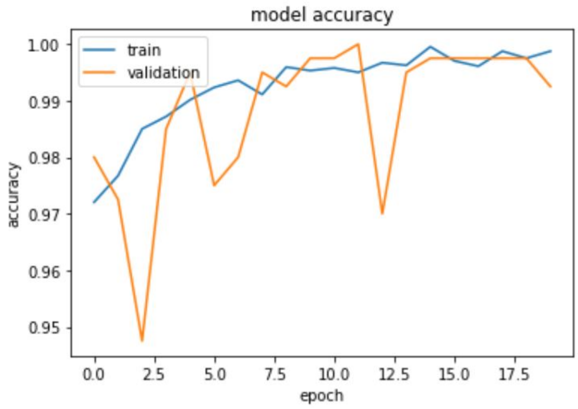


Figure 6.8: MobileNetV3 accuracy graph of histopathological lung cancer cell dataset

Model Loss Graph:

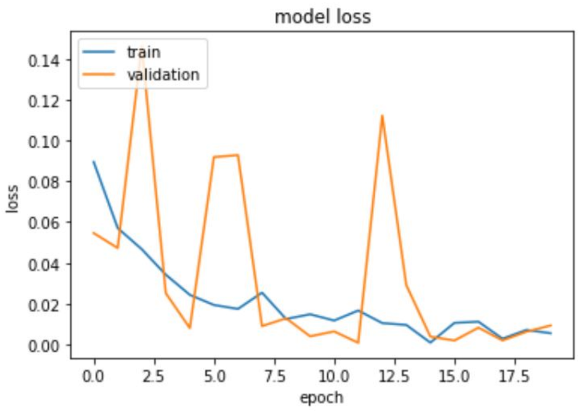


Figure 6.9: MobileNetV3 loss graph of histopathological lung cancer cell dataset

Confusion Matrix:

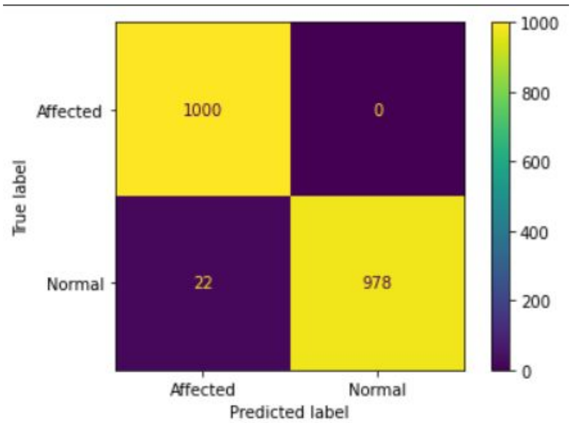


Figure 6.10: MobileNetV3 confusion matrix of histopathological lung cancer cell dataset

For the colon Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
96.25%	5.26%	97.84%

Table 6.4: Percentages of the performance parameters of MobileNetV3 colon cancer histopathological image dataset

Model accuracy graph:

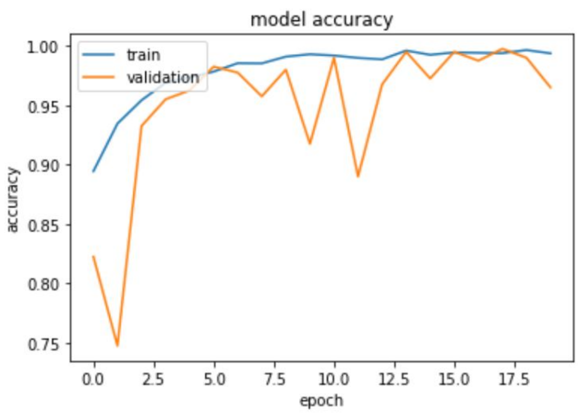


Figure 6.11: MobileNetV3 accuracy graph of histopathological colon cancer cell dataset

Model Loss Graph:

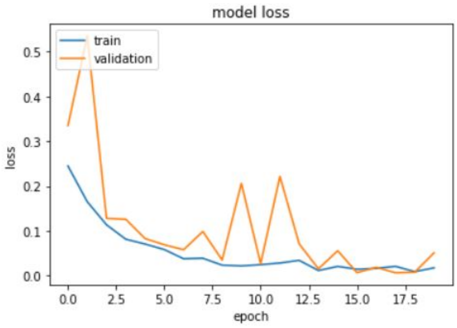


Figure 6.12: MobileNetV3 loss graph of histopathological colon cancer cell dataset

Confusion Matrix:

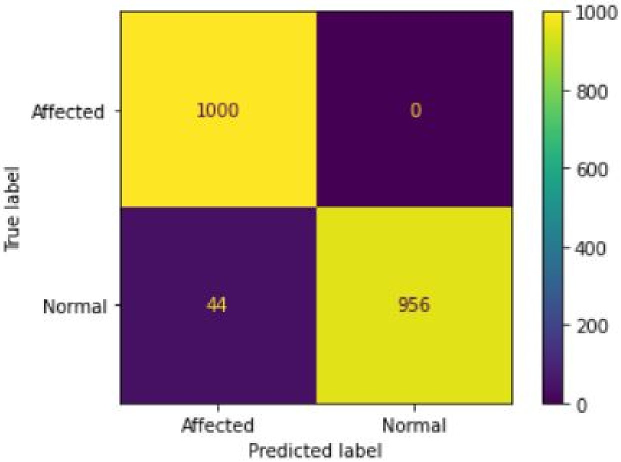


Figure 6.13: MobileNetV3 confusion matrix of histopathological colon cancer cell dataset

6.3 DenseNet121

DenseNet121 consists of 121 convolutional layers with 4 dense blocks and 3 transition layers. Input images go to the first convolutional layer with filter size 7x7 and stride 2, the pooling layer with filter size 3x3. Inside every dense block, each convolutional layer is connected to other convolutional layers directly. Therefore, one layer's generated feature map receives the other layer's feature map with its own generated feature maps and this feature size must be the same inside a feature block to be concatenated with each other. The architecture of DenseNet121 is given below:

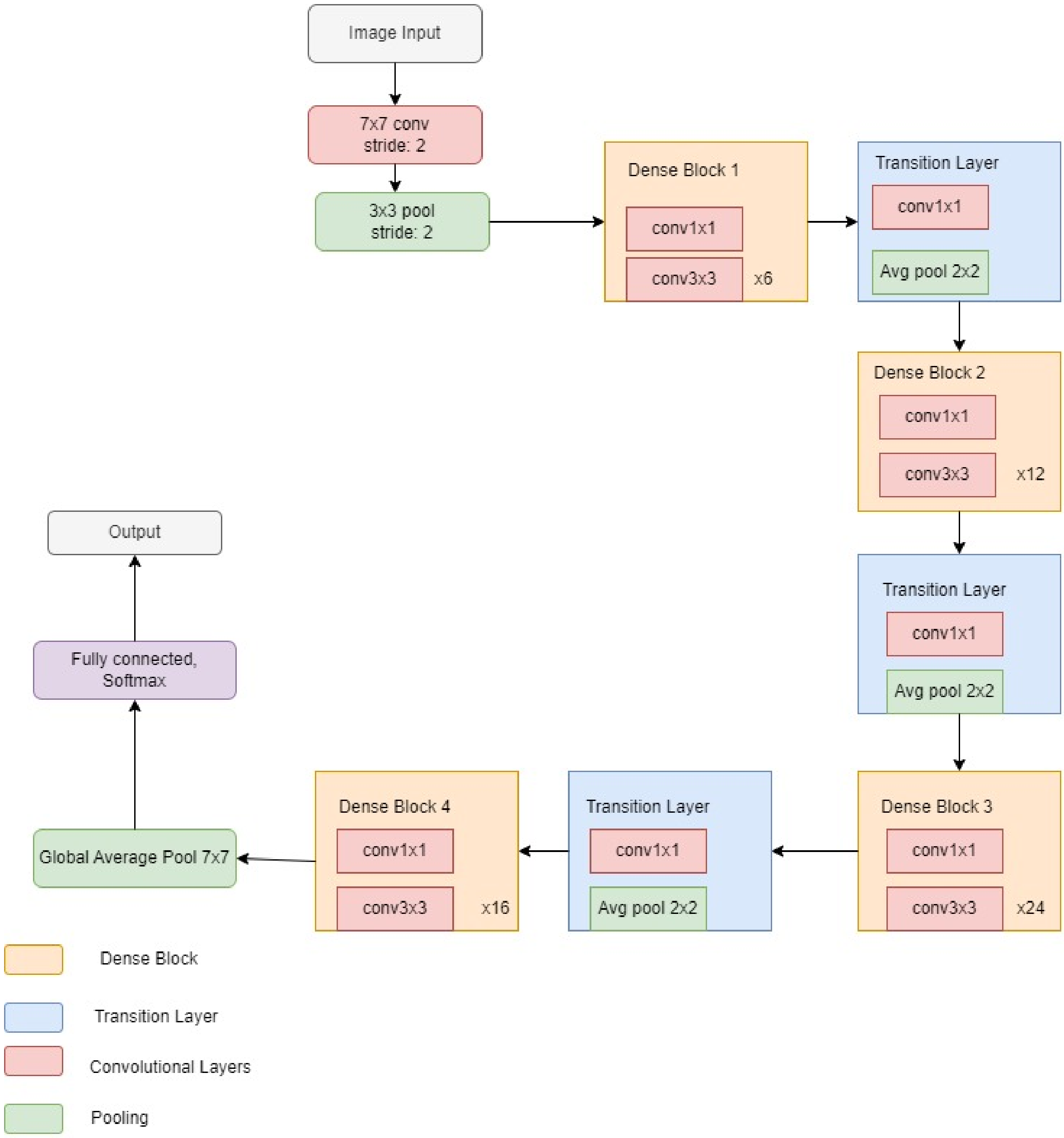


Figure 6.14: DenseNet121 architecture

Binary Accuracy, Loss, F1 Score of DenseNet121 are given below:
For the Lung Cancer Dataset:

Binary Accuracy	Loss	F1 Score
99.06%	7.84%	99.23%

Table 6.5: Percentages of the performance parameters of DenseNet121 lung cancer histopathological image dataset

Model accuracy graph:

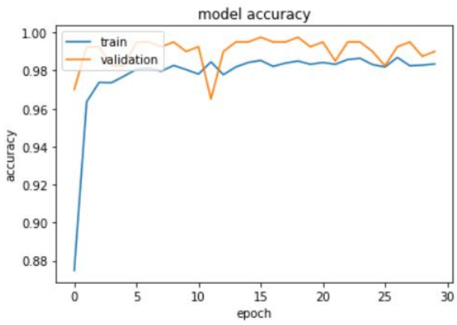


Figure 6.15: DenseNet121 accuracy graph of histopathological lung cancer cell dataset

Model Loss Graph:

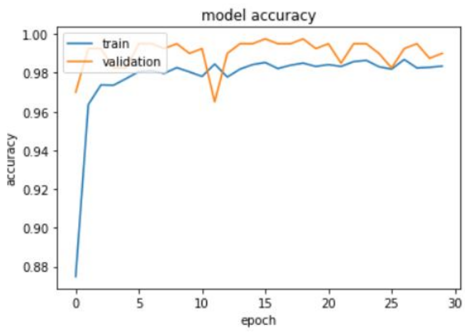


Figure 6.16: DenseNet121 loss graph of histopathological lung cancer cell dataset

Confusion Matrix:

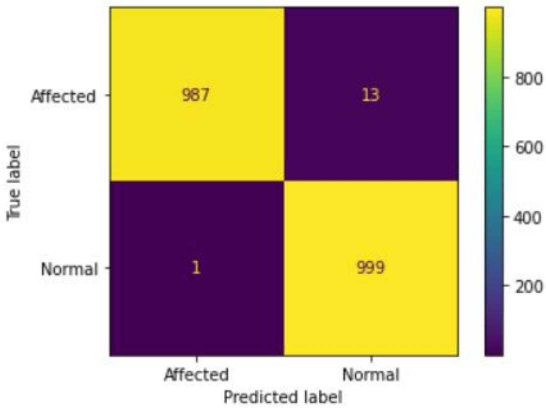


Figure 6.17: DenseNet121 confusion matrix of histopathological lung cancer cell

For the colon Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
83.75%	49.65%	85.32%

Table 6.6: Percentages of the performance parameters of DenseNet121 colon cancer histopathological image dataset

Model accuracy graph:

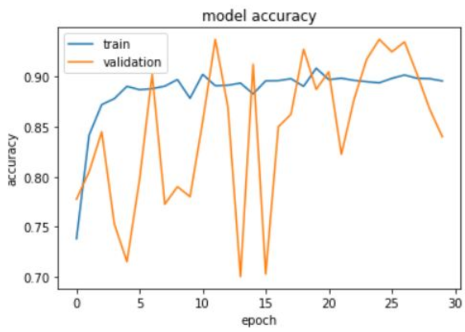


Figure 6.18: DenseNet121 accuracy graph of histopathological colon cancer cell dataset

Model Loss Graph:

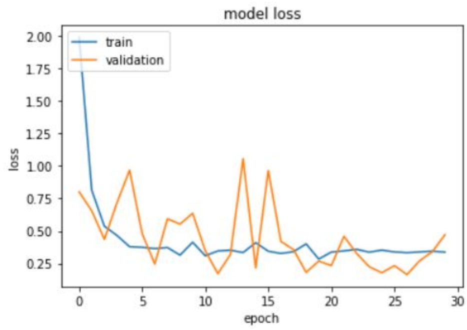


Figure 6.19: DenseNet121 loss graph of histopathological colon cancer cell dataset

Confusion Matrix:

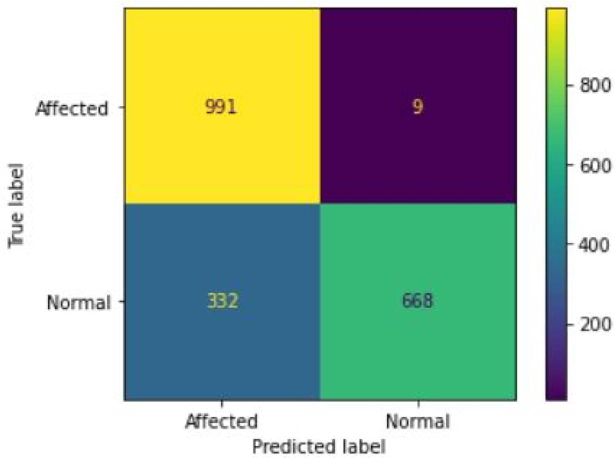


Figure 6.20: DenseNet121 confusion matrix of histopathological colon cancer cell dataset

6.4 InceptionV3

InceptionV3 architecture has 42 layers which use a Label Smoothing classifier as well as factorized 7x7 convolutions and it also has an auxiliary classifier.

Binary Accuracy, Loss, F1 Score of InceptionV3 are given below:

For the Lung Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
99.37%	1.09%	99.65%

Table 6.7: Percentages of the performance parameters of InceptionV3 lung cancer histopathological image dataset

Model accuracy graph:

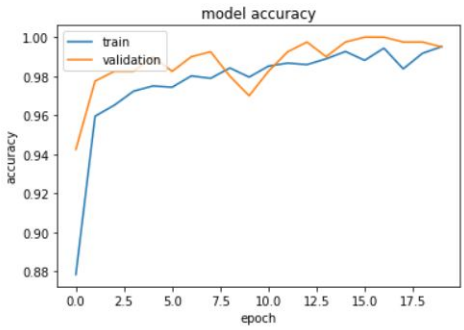


Figure 6.21: InceptionV3 accuracy graph of histopathological lung cancer cell dataset

Model Loss Graph:

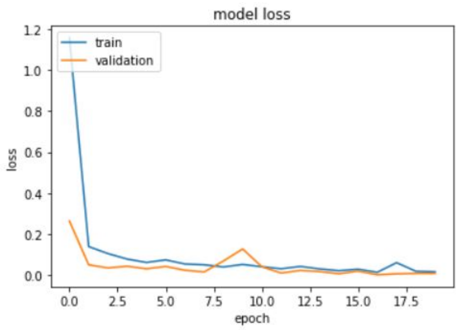


Figure 6.22: InceptionV3 loss graph of histopathological lung cancer cell dataset

Confusion Matrix:

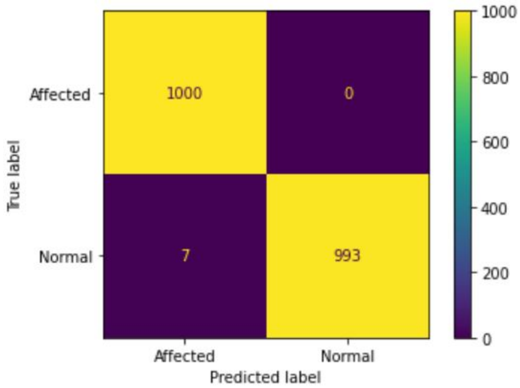


Figure 6.23: InceptionV3 confusion matrix of histopathological lung cancer cell

For the colon Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
83.75%	49.65%	85.32%

Table 6.8: Percentages of the performance parameters of InceptionV3 colon cancer histopathological image dataset

Model accuracy graph:

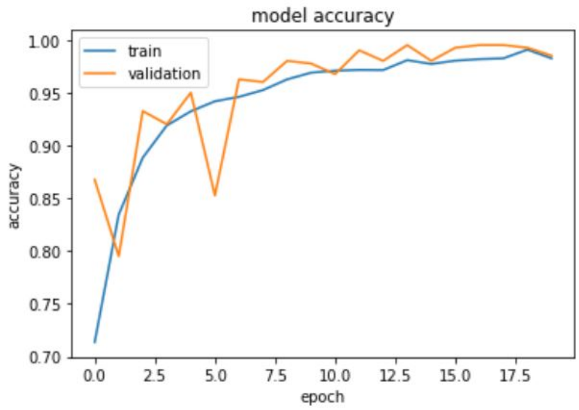


Figure 6.24: InceptionV3 accuracy graph of histopathological colon cancer cell dataset

Model Loss Graph:

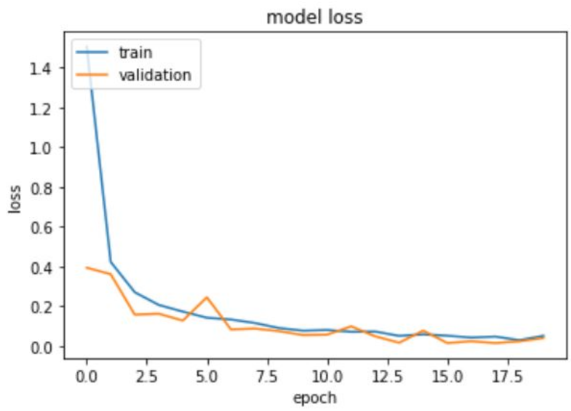


Figure 6.25: InceptionV3 loss graph of histopathological colon cancer cell dataset

Confusion Matrix:

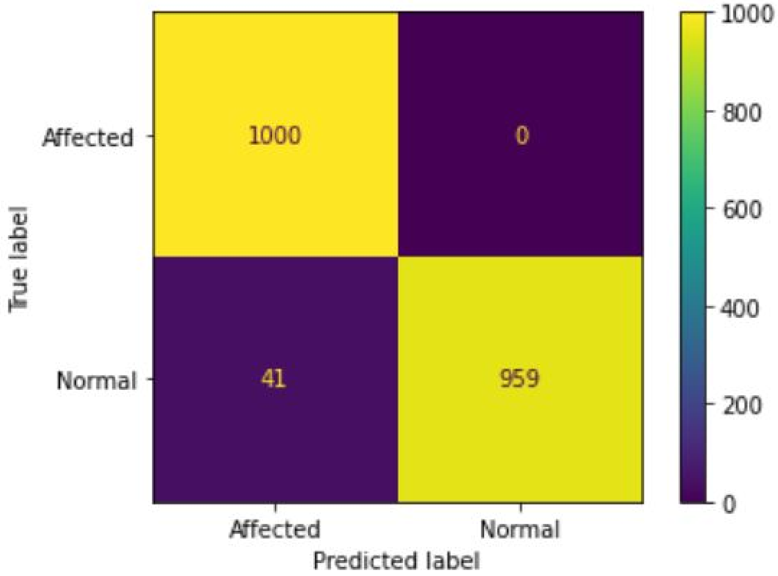


Figure 6.26: InceptionV3 confusion matrix of histopathological colon cancer cell dataset

6.5 Xception

The xception model has 36 convolutional layers with 14 modules. All the layers, without first and last modules, are surrounded by the linear residual connections. In the Xception architecture, data goes through three different flows; i) Entry Flow, ii) Middle Flow and iii) Exit Flow.
Binary Accuracy, Loss, and F1 Score of Xception are given below:

For the Lung Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
99.69%	1.74%	99.64%

Table 6.9: Percentages of the performance parameters of Xception lung cancer histopathological image dataset

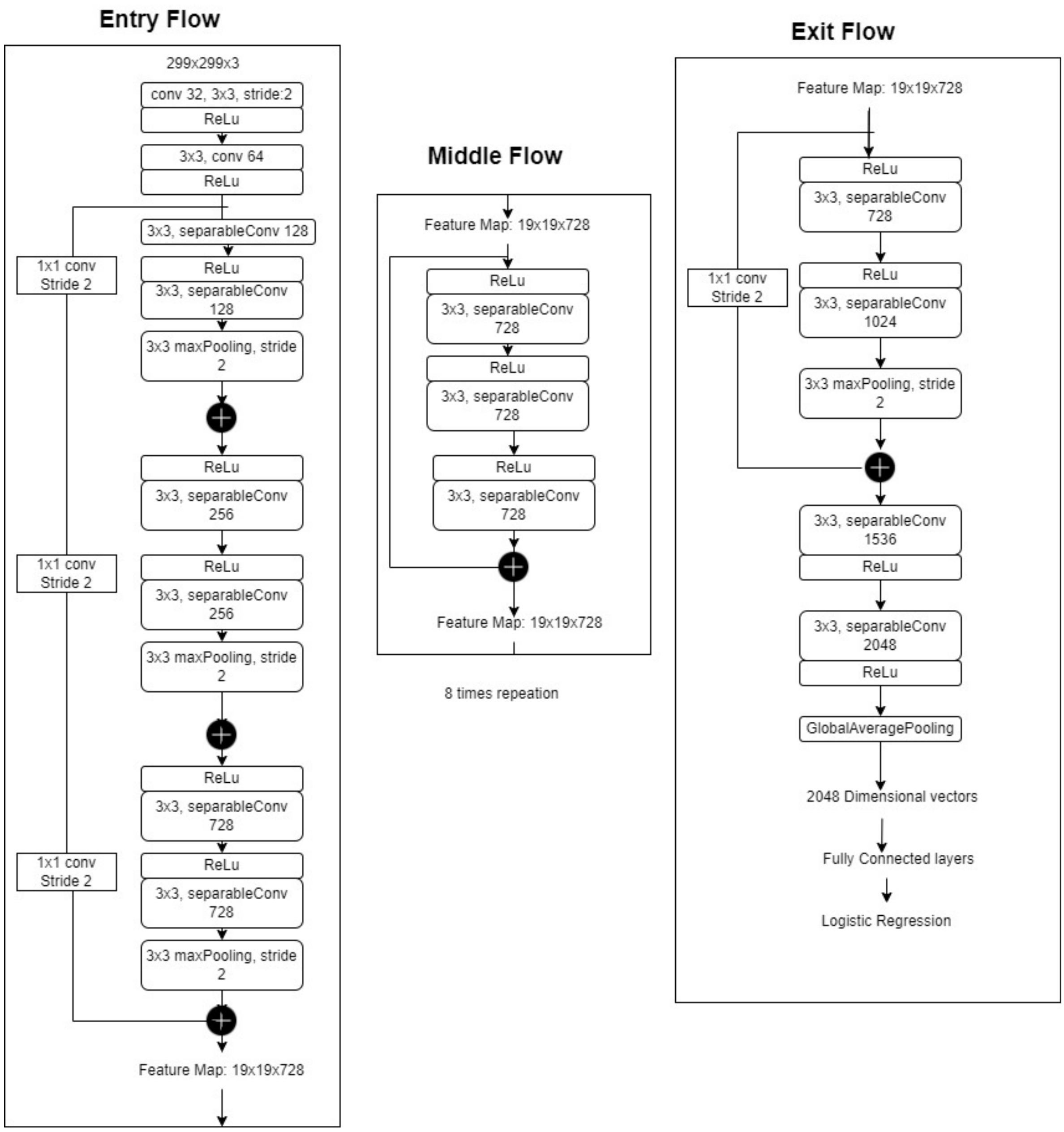


Figure 6.27: Xception architecture internal view

Model accuracy graph:

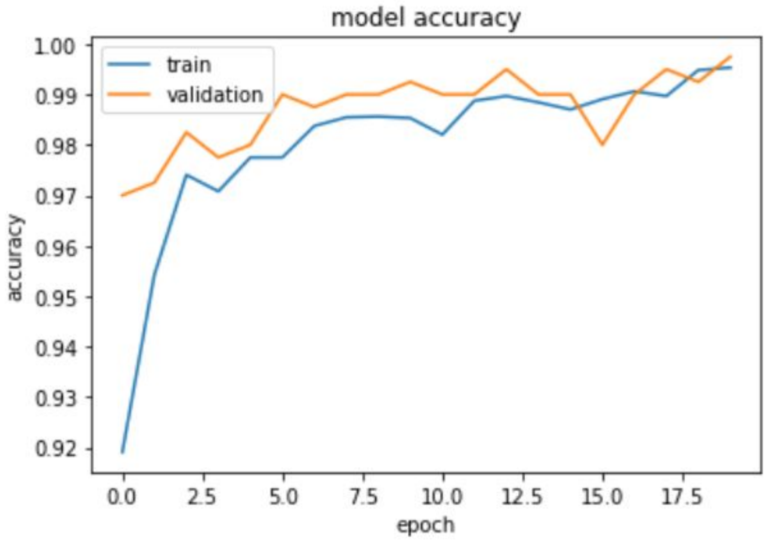


Figure 6.28: Xception accuracy graph of histopathological lung cancer cell dataset

Model Loss Graph:

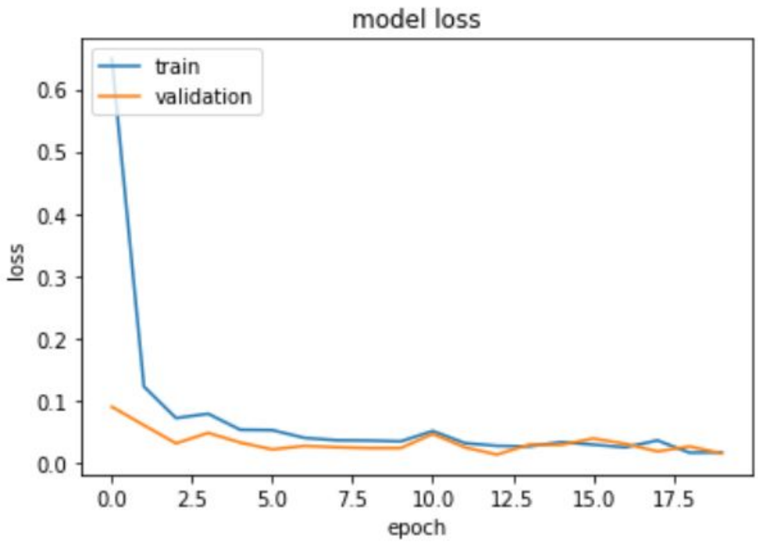


Figure 6.29: Xception loss graph of histopathological lung cancer cell dataset

Confusion Matrix:

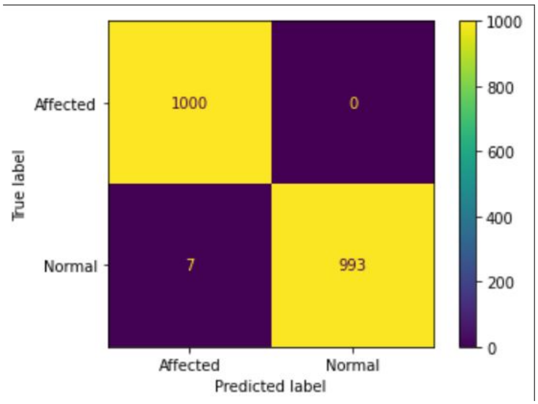


Figure 6.30: Xception confusion matrix of histopathological lung cancer cell dataset

For the colon Cancer Dataset:

Binary Accuracy	**Loss**	**F1 Score**
95.54%	7.25%	96.24%

Table 6.10: Percentages of the performance parameters of Xception colon cancer histopathological image dataset

Model accuracy graph:

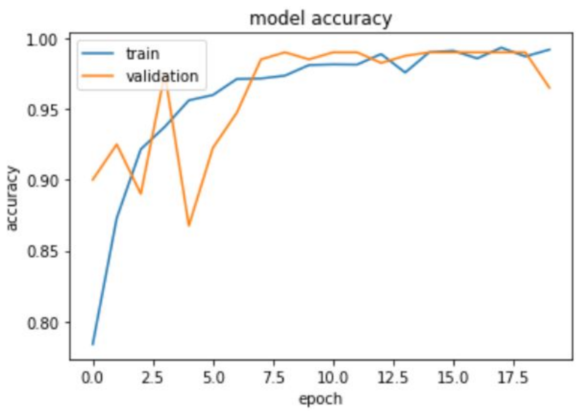


Figure 6.31: Xception accuracy graph of histopathological colon cancer cell dataset

Model Loss Graph:

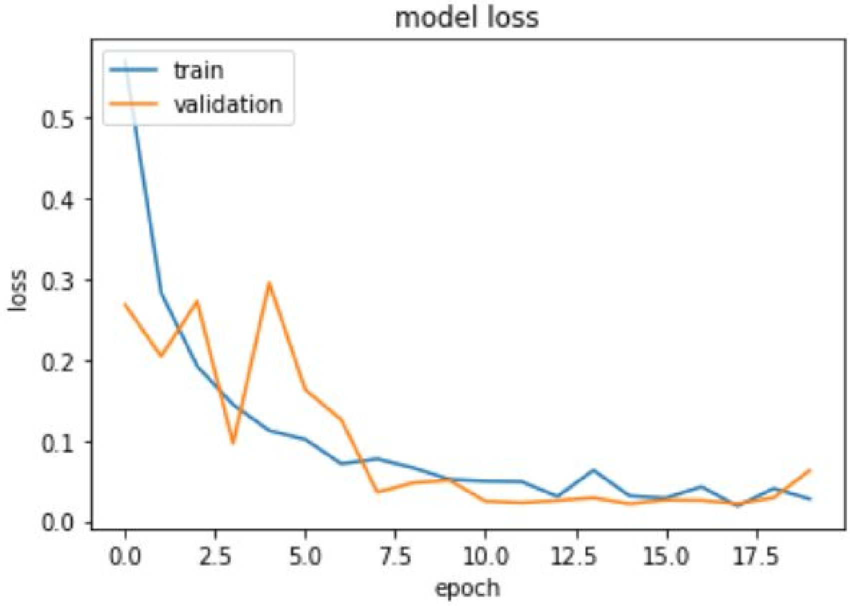


Figure 6.32: Xception loss graph of histopathological colon cancer cell dataset

Confusion Matrix:

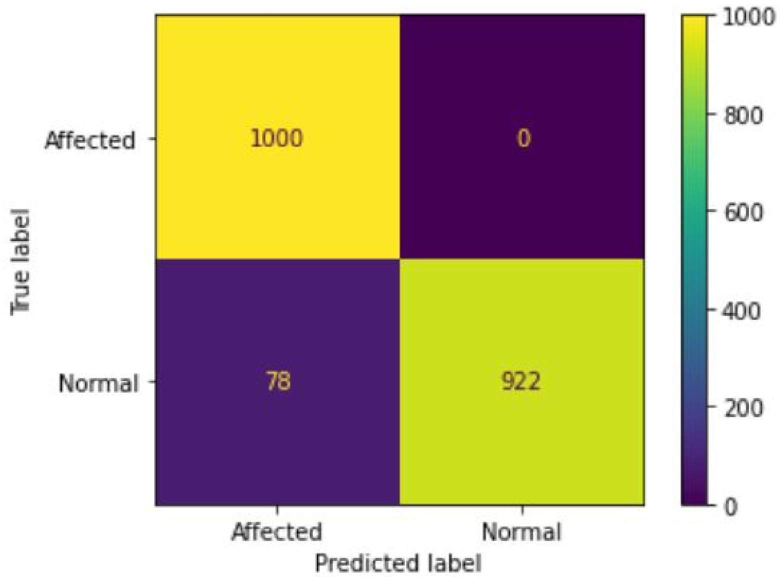


Figure 6.33: Xception confusion matrix of histopathological colon cancer cell dataset

Chapter 7

Prototype Development

7.1 Use Case Diagram

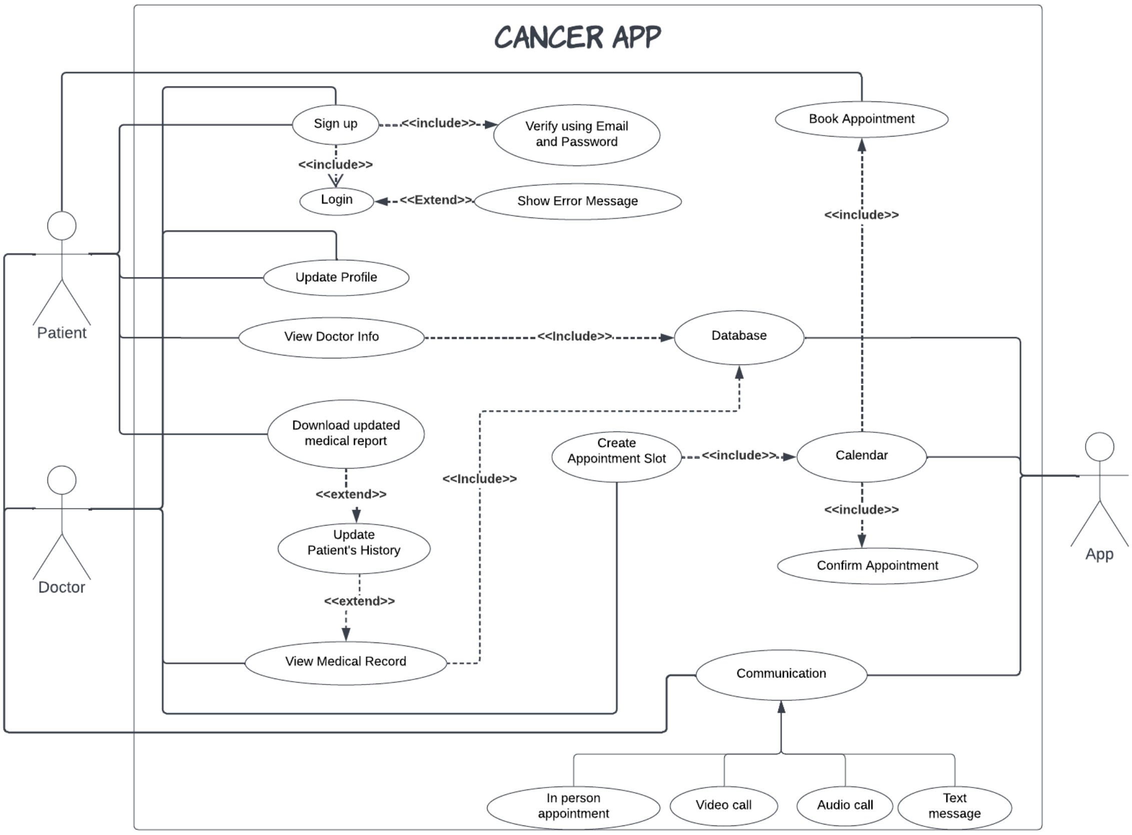


Figure 7.1: Use case diagram

According to figure 7.1 'patients' and 'doctors' are the primary actors who give input and perform the actions of roles in the mobile application and 'System' is the secondary actor who performs the administrative functions. Moreover, each use case has a unique name which describes the functions of the app. Each actor is connected with use cases via a straight solid line classified as an association. Besides that, 'communication' has a generalization relation between four use cases (in-person appointment, video call, audio call and text message). Lastly, two types of

Features Suggested to implement in the App	**Percentage (%)**
Patients would be able to chat with the doctors	30.0
Phone call recording and audio-recording system	30.0
Patients will be able to communicate with the doctors through video conversation	42.5
Patients would be able to send their reports to the doctors	75.0
Doctors will have access to patients medical history	80.0
Doctors can prescribe medicine and give feedback to their patients	55.0
An option to immediately contact at any emergency situation	60.0
The system will provide basic information through bot (do's/don'ts) for cancer patients	55.0

Table 7.1: Suggested features by the doctors for implementation in the app

dependencies are shown in the figure, which are $<<$ *include* $>>$ and $<<$ *extend* $>>$, where include indicates a compulsory relationship between the use cases and extend indicates an optional relationship.

7.2 Features

Based on the survey question, "What type of features would you want the app to have?", asked the oncologists, we have analyzed our findings and have presented them in a table. Table 7.1 represents the percentage of doctors' desired features.

Other than the features listed in the table above, we have also planned to implement a few more features such as:

- Both patients and doctors will get a notification before their appointments
- Patients will be able to find the list of hospitals and also search for oncologists registered in the app.
- Doctors will be able to view their list of patients and check each individual's medical history.
- Patients can book an appointment with their preferred doctors directly from the app.
- Doctors can also check their daily list of appointments through the app.

7.3 Proposed Prototype

- Sign up - The figure below illustrates the sign-up process from both Doctor's and the Patient's end:

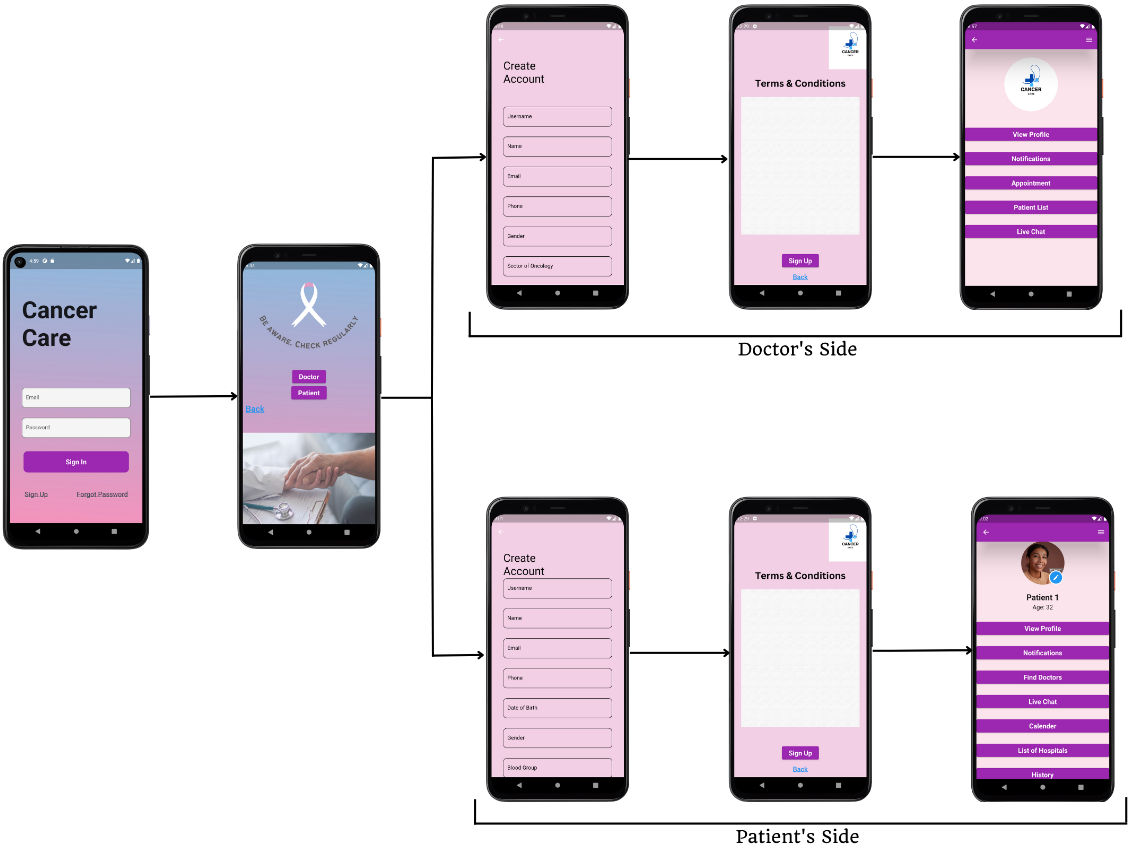


Figure 7.2: Flowchart of the signup process of the prototype

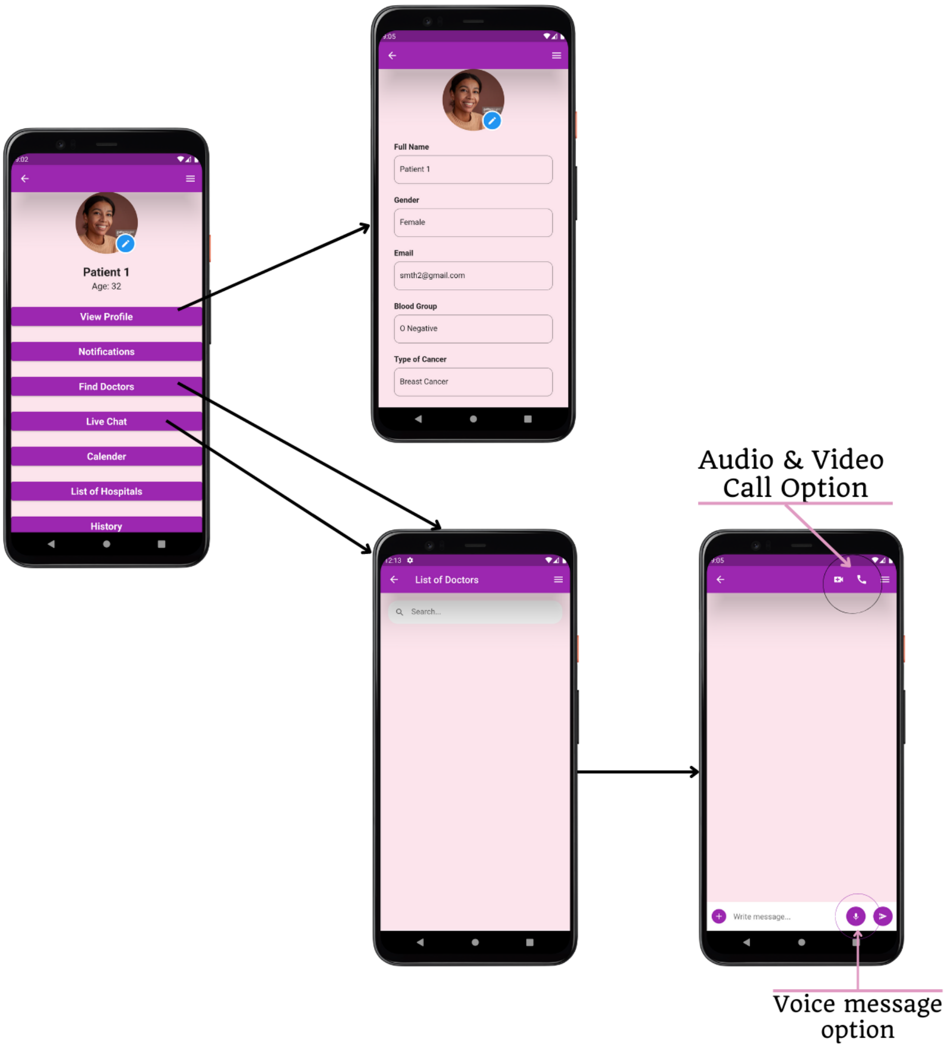


Figure 7.3: Flowchart of features from patient's homepage

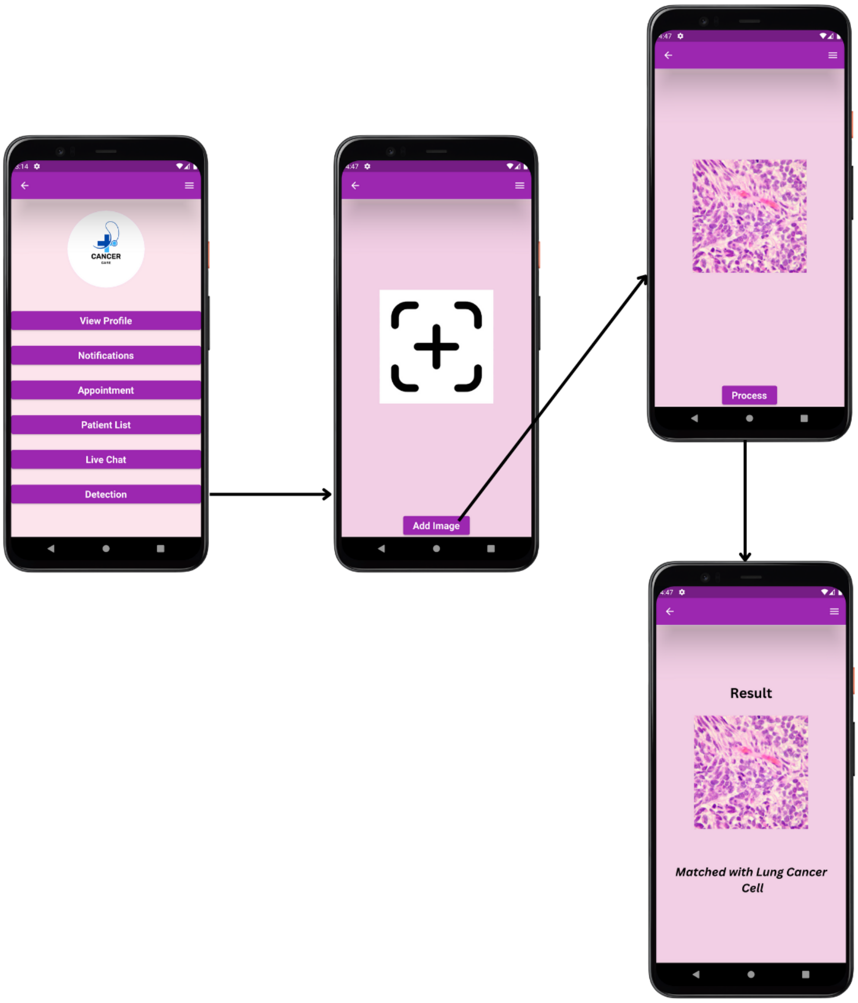


Figure 7.4: Flowchart of histopathological cancer cell detection features

Chapter 8

Result and Discussion

8.1 Comparison Among the Architectures

In the kaggle [39], comparisons have been shown between different architectures trained and tested on only lung cancer histopathological cell images. The architectures they have used to compare are EfficentNetB7, InceptionV3, ResNet50, ResNet50 drop SimpleCNN and VGG16. Therefore, we have used not only VGG16, and InceptionV3 models but also DenseNet121, Xception and MobileNetV3. Besides, we have also used histopathological images of the colon cancer datasets to compare the accuracy of these five architectures. We have trained all the architectures individually on each of the datasets and present the comparison of the average accuracy.

For measuring the performance of the architectures (VGG16, MobileNetV3, DenseNet121, InceptionV3, Xception), we have analyzed some common matrices of the performance such as accuracy, precision, recall, F1 score for both the datasets and compared the average model accuracy. Among these five architectures, VGG16 performance is better with 99.1% average model accuracy than other architectures which are MobileNetV3 (97.8%), DenseNet121 (90.6%), InceptionV3 (98.8%) and Xception (97.9%) shown in the table 8.1.

The average accuracy of VGG16 gives us 1.3% more accuracy than MobileNetV3, 8.5% more than DenseNet121, 0.3% and 1.2% more accuracy than InceptionV3 and Xception architectures.

Architectures	Lung cancer detection				Colon cancer detection				Average model accuracy
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	
VGG16	98.9%	100%	97.85%	98.91 %	99.35%	99.3%	99.4 %	99.35 %	99.1%
MobileNetV3	97.8%	100%	97.85%	98.91 %	97.8%	100%	95.78 %	97.84 %	97.8%
DenseNet121	98.3%	98.67%	99.8%	99.23 %	82.95%	99.1%	74.9 %	85.32 %	90.6%
InceptionV3	99.65%	100%	99.3%	99.65 %	97.95%	100%	96.06 %	97.96 %	98.8%
Xception	99.65%	100%	99.3%	99.64 %	96.1%	100%	92.76 %	96.24 %	97.9%

Figure 8.1: Comparison of performance matrices of the architectures

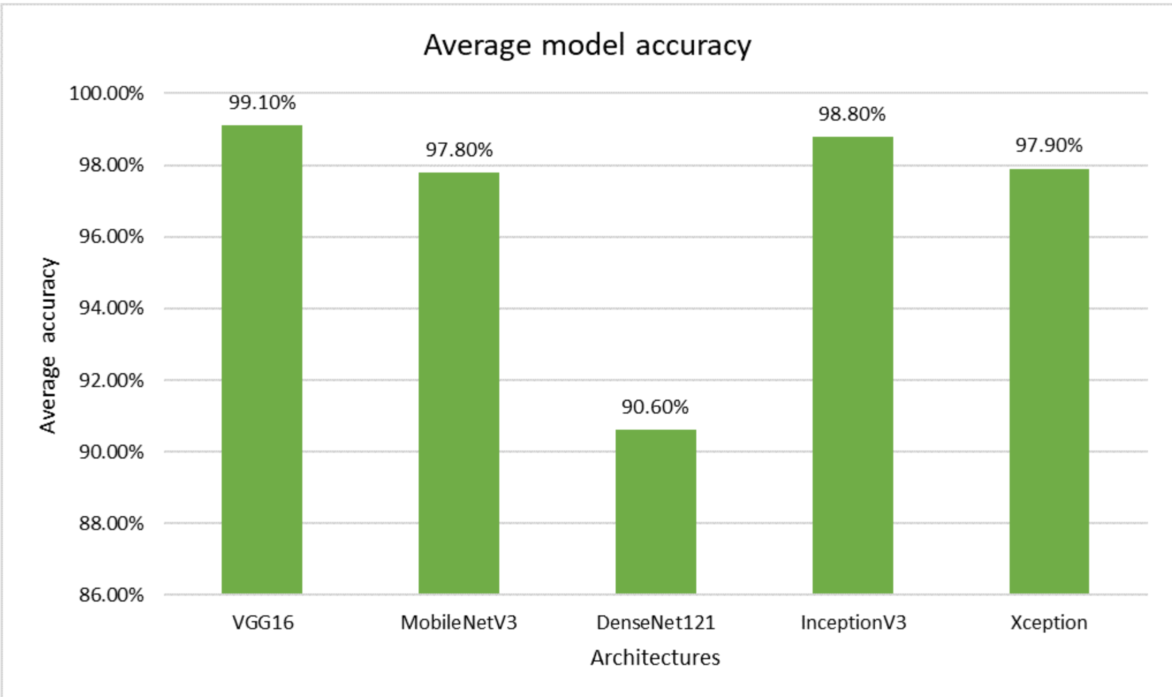


Figure 8.2: Demonstration of the average accuracy of the models

8.1.1 Overfitting Problem

During training and finding performance matrices, we have done fine-tuning of the pre-trained architectures to increase the accuracy rate. However, VGG16 and DenseNet121 models become overfitted by fitting exactly to the training set and give us 100% binary accuracy. Therefore, these two architectures were unable to give us accurate results on the new data.
The binary accuracy and loss that we have found for the overfitting of the models are given below:

Datasets	VGG16		DenseNet121	
	Binary Accuracy	Loss	Binary Accuracy	Loss
Histopathological Lung cancer cell image dataset	1.0000	5.7814e-10	1.0000	4.4454e-06
Histopathological colon cancer cell image dataset	1.0000	2.0718e-04	1.0000	3.4167e-04

Figure 8.3: Accuracy and loss of the overfitted architectures

Accuracy and loss graphs for overfitting are given below:

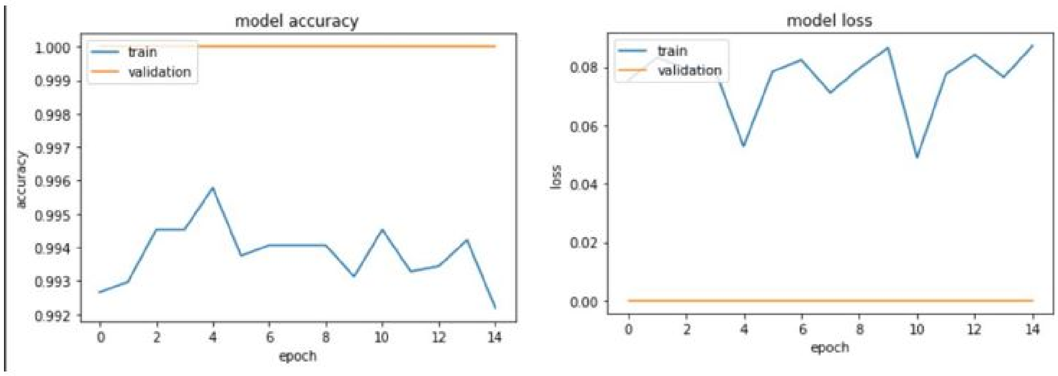


Figure 8.4: VGG16 accuracy and loss graph for overfitting

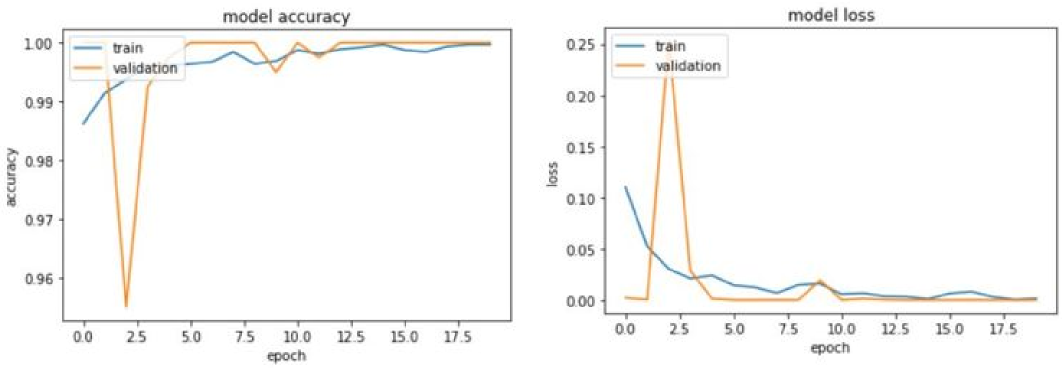


Figure 8.5: DenseNet121 accuracy and loss graph for overfitting

8.1.2 Solutions for Overfitting Problem

To solve the overfitting problem on VGG16 and DenseNet121, we have not performed fine-tuning and reduced the number of epochs from 30 to 20 for VGG16. After that, we have found the average accuracy of 99.1% of the VGG16 model and 90.6% of the DenseNet121 model.

8.2 Findings From Survey Studies

We conducted a survey which consisted of a total of 116 participants, patients and oncologists combined, which consists of 72 patients and 44 doctors. Now, while this data size may seem less, since our research method involves in-person data collection, according to a research [8], the mean sample size of in-person interviews is 18, whereas we conducted interviews with a total of 72 patients. Coming to doctors' data size, we have consulted with them and they have suggested,

"44 is more than enough data for doctors side since not every oncologist is willing to take part in the survey and not everyone is willing to share data"

On the basis of this, we analyzed the data of our survey and below we have discussed the important findings and why it is important.

8.2.1 Outcome From Patients Survey

Demographic Features of the Patients

We have evaluated which gender and age range people have been diagnosed with Cancer in the past few years. As a result, we found out that 43 responses were from females, and the remaining 29 were from male respondents. More than half of the patients diagnosed with cancer, which is 59.72%, were female. The age at which the individuals learned that they had cancer is illustrated as a bar chart in Figure 8.6. According to the survey, 19.44% of female patients and 6.94% of male patients are between the ages of 45 and 54, and 11.11% of male and 5.56% of female respondents are between the ages of 55 and 64. Most cancer diagnoses occur in middle-aged patients. Only 6.94% of female and 4.17% of male respondents were between the ages of 18 and 24 whereas 6.94% of males under 18 years had a cancer diagnosis.

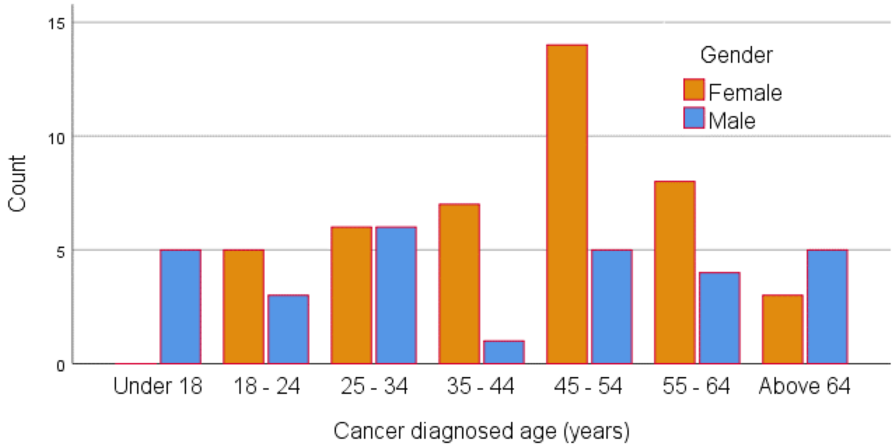


Figure 8.6: Percentage of male and female cancer patients respondents of different age range

However, the majority of the respondents to our questionnaire have decent educational backgrounds. A total of 23 respondents (31.94%) have earned a B.Sc. or BA, while 14 respondents (19.44%) have obtained an M.Sc. or MA. Given their better educational credentials, we can therefore speculate that the majority of responders have a basic knowledge and awareness of the ailment, cancer. However, 15.28% of respondents concluded their schooling before SSC, 6.94% finished their SSC, and 16.67% of respondents passed their HSC or A-level. Additionally, only an extremely small percentage of respondents, 2.78%, completed a diploma, and just 1.39%, a PhD. The remaining respondents, or 5.56%, have never attended a formal educational institution. This clearly states that a significant portion of responders may have only a rudimentary understanding of this fatal illness, its course of treatment, and its implications. The percentage of highest educational qualifications of the respondents is displayed in the pie chart in figure 8.7.

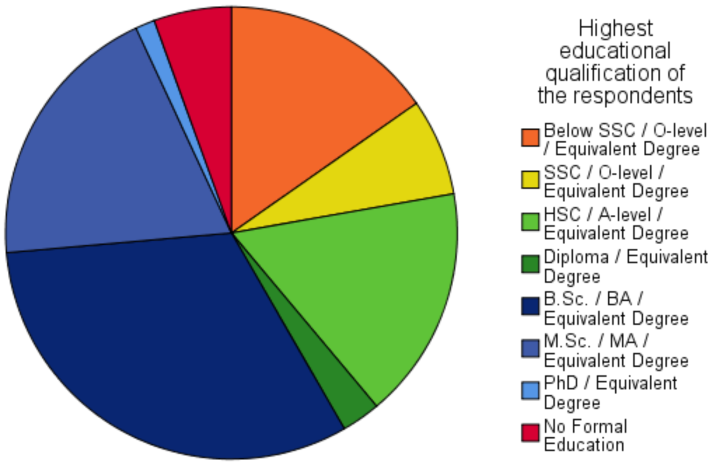


Figure 8.7: Percentage of Highest educational qualifications of the survey respondents

Besides, the study demonstrates that 47.22% of respondents came from middle-class families, 23.61% each is from lower middle-class and higher middle-class families, and the remaining 5.56% are mostly from lower-income families. The large percentage of responders, who come from middle-class households, are primarily students, housewives, public servants, private employees, retired personnel, and so on. Besides, students, teachers, private employees, retirees, business owners, and housewives make up the responders from lower middle-class and upper-middle-class families. Most unemployed people are from middle-class, lower-middle-class, or low-income homes. This demonstrates that the financial situation of the cancer respondents is not very satisfactory, with just an ordinary standard of living. The financial status of the respondents from various occupations is presented in the bar graph of figure 8.8.

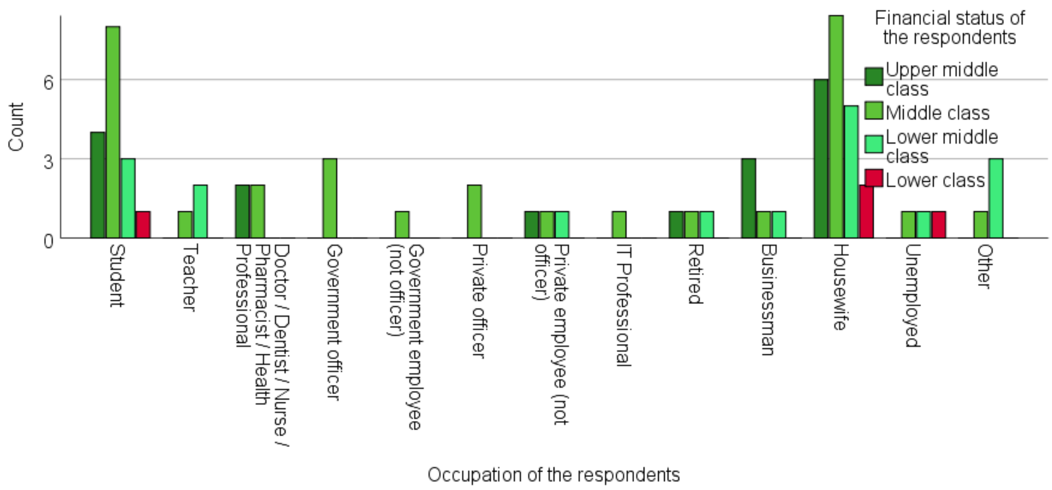


Figure 8.8: Percentage of respondents with different occupation and financial status

Since we conducted interviews with cancer patients at two Dhaka-area hospitals, the National Institute of Cancer Research and Hospital (NICRH) and Bangladesh Cancer Society Hospital and Wellfare Home, we obtained 40 (59.70%) responses from individuals who reside in the Dhaka Division. In addition, we received 11 (16.42%) responses from Chattogram respondents, 6 (8.96%) responses from Rangpur respondents, and 3 (4.48%) responses from Khulna respondents. We have received 2.99% of answers from the Barisal, Rajshahi, and Mymensingh Divisions, respectively. Unfortunately, we managed to get only 1 (1.39%) response from the Sylhet division. Figure 8.9's pie chart indicates the percentage of responders from various divisions.

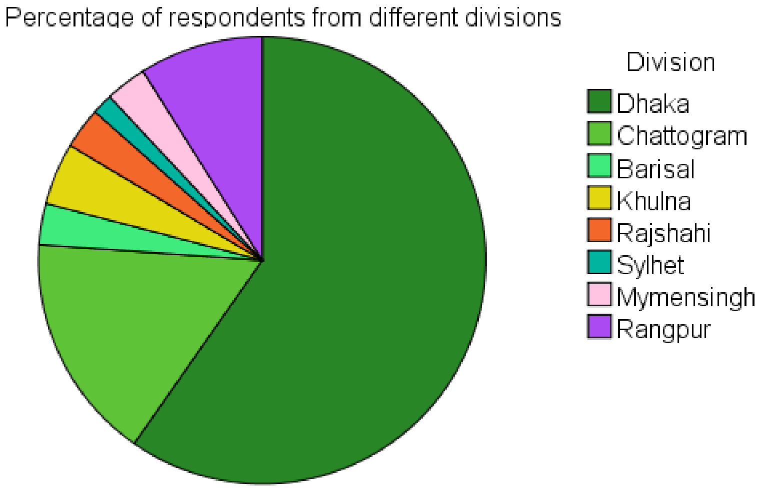


Figure 8.9: Percentage of respondents of different division

Types of Cancer Respondents Are Diagnosed With

The most frequent and prevalent cancer according to the data obtained from cancer institutes' patients as well as through survey responses is Blood cancer which has a frequency of 18.06%. Breast cancer and lung cancer, both of which account for 15.28% of cases, are the second most frequent cancers. The percentage for other malignancies, such as Cervical cancer is 8.33% whereas Head and Neck cancer, Lymphoma and Rectal cancer are 5.56% for each. Figure 8.10 displays the various cancers that respondents have indeed been diagnosed with.

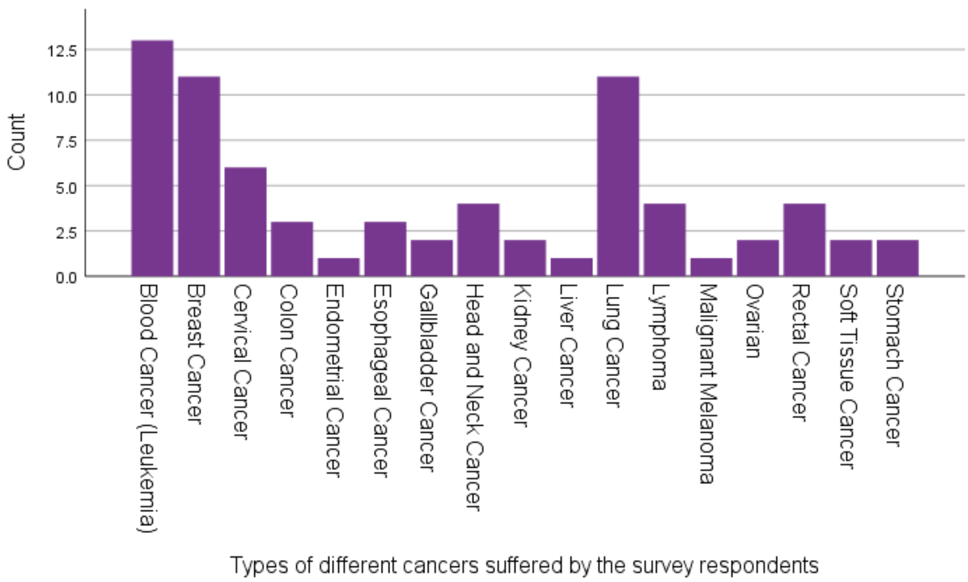


Figure 8.10: Types of different cancers the survey respondents are diagnosed with

Restricted Foods During Diagnosis and Basic Health-Related Issues

As cancer is a deadly disease patients have to follow proper diet charts strictly and exclude some foods from their dietary habits. They often face several common

List of food items restricted by doctors	Basic health issues patients experience
Fatty and oily food	Loss of appetite
Excessive sugar in presence of Diabetes	Vomiting
Salty foods in presence of Hypertension	Diarrhea
Grape	Weight Loss
Meat	Weakness
Pear	Coughing and Sneezing
Orange	Bleeding from any organ such as nose, anal etc.
Brinjal	Pain – may be localized or diffused
Fish	Insomnia
	Difficulty in respiration

Table 8.1: List of some foods that should be excluded from dietary habits as instructed by the doctors and basic health-related issues experienced by the Cancer patients

health-related complications for different types of cancer as well. We have listed some food items that were restricted by the doctors and some basic health-related issues that they often face from our survey data in table 8.1.

Communication Reliability With Doctors

During the survey, we collected information from the patients about their communication reliability with oncologists where we have found that patients face some difficulties while communicating with doctors during their treatment period which is illustrated in figure 8.11.

Figure 8.11: Problems faced by the respondents while communicating with doctors

At the time of our inquiry about their communication problems, 14.29% of the respondents said they forgot to bring required reports, 21.43% of them were unable to show up for their appointment, and 64.29% failed to convey their difficulties to the doctors. In addition, 69.2% of respondents said it was difficult to schedule a doctor's appointment, 61.5% claimed that they faced transportation problems and 46.2% said they were stuck in traffic.

Sample Interview With Patients

During the interview with a child who is suffering from Brain cancer, said to us, *"I was infected with cancer when I was 9 years old. Now, I am 12 years old. I had 3 surgeries on my brain. I am unable to see. I am not so educated, but I know many prayers. I always pray to Allah that I can overcome my sickness soon."*.

We also talked with his parents. His father told us that he was a farmer. He sold all his lands for the treatment of his child. They have a daughter. But she rarely comes to visit them as they are not financially developed.

From the interview with the patients, a heartbreaking message we have got is that cancer is not only a deadly disease but also very costly for patients who are not financially well-established.

Respondents and Their Proficiency in Using Smartphones

The bar chart in figure 8.12 depicts that 9.84% of respondents aged 18 to 24 and 8.24% of respondents aged 25 to 34 have expertise in using smartphones, respectively. 9.84% respondents of the age range 45-54 years are experienced in using a smartphone. However, 6.84% of individuals aged 55 to 64 have only basic or intermediate knowledge of how to operate a smartphone proficiently. The participants who are below the age of 18 and above the age of 64 have never operated a smartphone. In this technologically evolved era, present youth, particularly those who are 18 years of age and older, have a broader understanding of modern technology and its applications. For this reason, they prefer adopting more sophisticated technological tools that are simple to operate and effective.

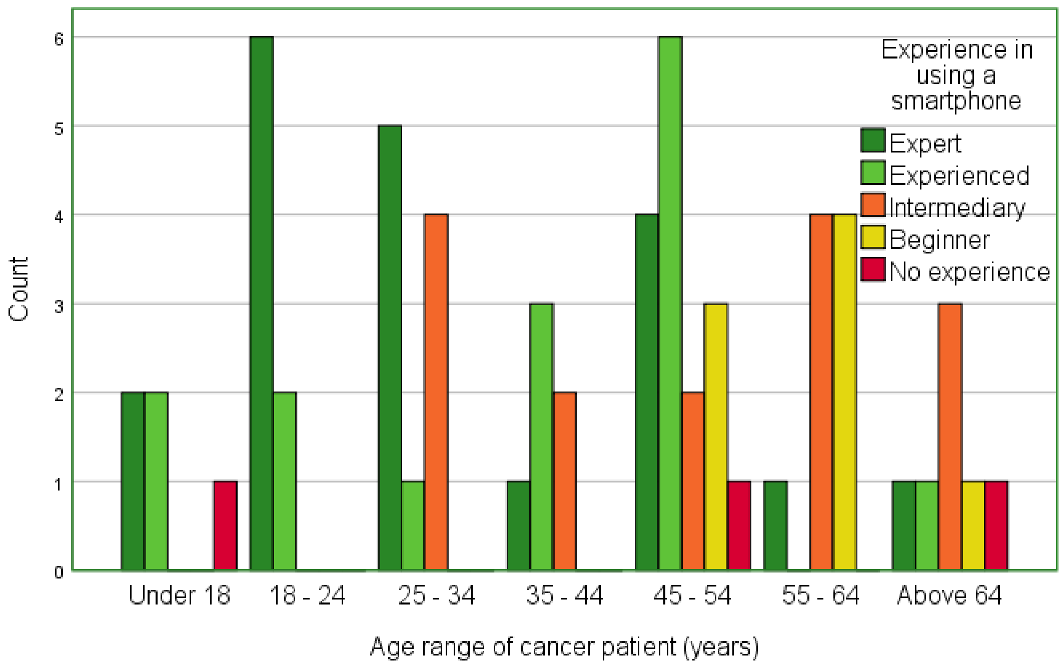


Figure 8.12: Proficiency of different age group respondents in using smartphone

Willingness of Patients to Use the App

Even though patients experience less trouble while interacting with doctors, they are indeed willing to rely on a System to ease communication. In fact, when asked if they would choose a system that could alleviate their hassle, 55 (76.39%) out of 72 participants answered "Definitely" and 11 (15.28%) respondents chose "Most Probably" as displayed in the pie chart in figure 8.13. This illustrates that the prototype we have put forth the attempt to design could be of enormous help to both the patient and the doctors in developing a more efficient interface to communicate.

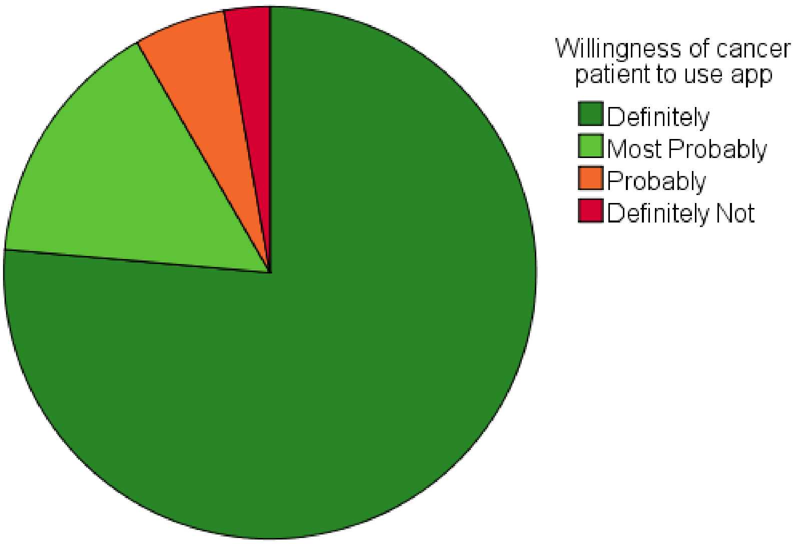


Figure 8.13: Willingness of patients to use app

8.2.2 Outcome From Doctor's Survey

Demographic Features of the Doctors

The doctor's demographic description from our findings is given in table 8.2.

Age (years)	Gender	Field of Oncology	Medical Sector	Professional Designation
25 - 34: 63.6% 35 - 44: 31.8%	Male: 54.5% Female: 43.2%	Radiation Oncologist: 79.5% Medical Oncologist: 18.2%	Private Sector: 29.5% Government Sector: 68.2%	Resident Physician: 31.8% Medical Officer/Assistant Registrar: 36.4%
45 - 54: 2.3%	Do not want to disclose: 2.3%	Surgical Oncologist: 2.30%	Defense:2.3%	Resident: 6.8%
55 - 64: 2.3%				Trainee: 6.9%
				Student: 6.9% Assistant Professor / Junior Consultant: 4.5% Registrar: 4.5% Trainee Officer: 2.3%

Table 8.2: Different demographic features of doctors

From the demographic features received through the survey, we have found out that most oncologists are aged between 25 to 34 years (63.6%) and 35 to 44 years (31.8%) who are mostly Radiation oncologists. The majority of respondents were oncologists who were male (54.5%), female (43.2%), and who also did not want to reveal their identity (2.3%). The oncologists who took part in the survey and interviews came from three various sectors: the public sector (68.2%), the private sector (29.5%), and the defense sector (2.3%). There were additionally physicians (31.8%), medical officers or assistant registrars (36.4%), and residents (6.8%). Only a small number of trainees, students, assistant professors or junior consultants, registrars, and trainee officers, however, also took part in the questionnaire.

Communication Reliability With Patients

Since the prime goal of this proposed application prototype is to improve the communication gap between patients and doctors, in our survey we asked the respondents (oncologists) if they ever faced any problems while communicating with the patients during their diagnosis and as Figure 8.14 illustrates that 39 (88.64%) of them selected ‘Yes’ as their response. When asked what kind of communication problems the oncologists faced, at least 35 (89.7%) respondents out of 44 selected “Patient forgot to bring the necessary reports” and 20 (51.3%) out of 44 respondents chose “Patient could not come timely during the appointment time”. One oncologist even mentioned that communication problems are not only from the patient’s end but sometimes the doctors cannot explain the treatment scheme properly.

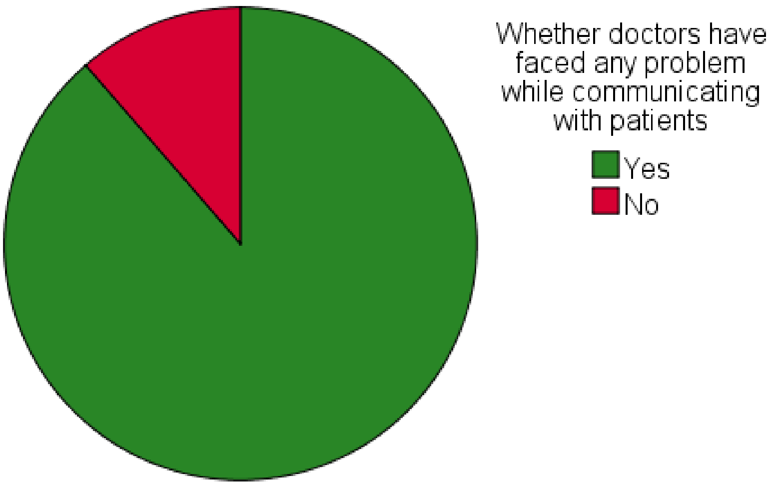


Figure 8.14: Percentage of doctors who faced difficulty while communicating with the patients

Sample Interview With Oncologists

When we talked with the Oncologist about the patients, one of the Oncologists said to us,
"Most of the patients are not well aware of their physical condition. For this, they could not explain the problems clearly. Communication problems are not only from the patient's end, but sometimes we cannot explain the treatment scheme properly.".

Willingness of Doctors to Use the App

When asked whether or not the doctors would use a system (if there is one) for easier communication with their respective patients, 40 out of 44 respondents opted for 'Yes', as shown in figure 8.15, implying that there is a scope for the system in the future.

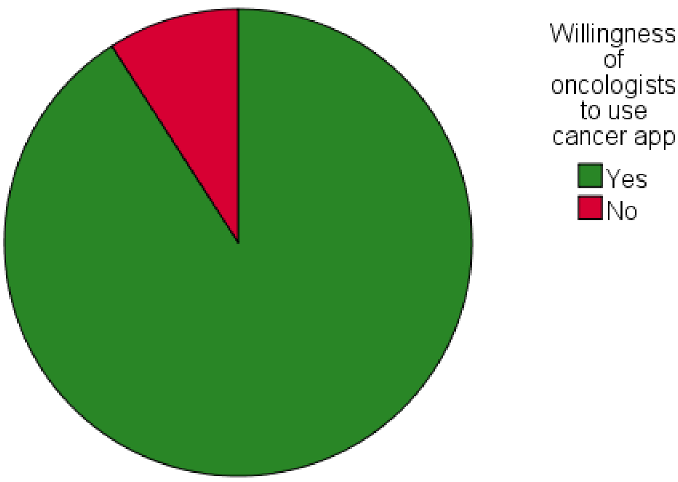


Figure 8.15: Willingness of patients to use app

8.3 Result and Discussion of the Prototype

We have designed the prototype to be as simple as possible to ensure that cancer patients can easily operate it without hassle when the application is complete and deployed since many people are technologically challenged. Moreover, we have tried to keep the most basic features needed for a cancer patient, such as audio calls, video calls, etc. Now, a question might arise what happens to individuals who have little to no knowledge of how to operate a smartphone? From our survey, we have analyzed that since participants from a younger age group are more experienced in smartphone use, they can help the cancer patient(if family) operate the mobile application easily.

But the main question remains whether or not our proposed app will be feasible in the long run if deployed. Our data from Table 8.3 shows that doctors and patients have previous experience with online consultation through other communication platforms, such as WhatsApp (Patient: 92.5%, Doctor: 91.5%), Messenger (Patient: 17.5%, Doctor: 35.3%), Viber (Patient: 30%, Doctor: 11.8%) etc. Hence, we can deduce that the participants are quite proficient with smartphones.

One of our survey questions included a question for the doctors on what kind of features they would like the application to have. Based on their suggestions, we have designed our prototype and added features such as sending reports as attachments, doctors having access to patient's medical history etc. After the initial prototype was completed, we presented it to a few patients and two doctors for some constructive feedback. Besides that, we have also discussed with the two doctors collaborating with us about detecting cancer cells through our app. Keeping these features in mind, we will be incorporating those features while developing the application in our future work.

Other communication platforms used	By Doctors (%)	By Patients (%)
WhatsApp	91.2	92.5
Messenger	35.3	17.5
Viber	11.8	30
Email	2.9	7.5
imo	20.6	0.00
Skype	0.00	2.5

Table 8.3: Other communication platforms used by doctors and patients

8.3.1 Feedback From Patients

- **Patient 1:**"It seems that it is quite easy to use the app, which will be good for individuals of all age groups. It would also be great if there were an option to call for basic queries."
- **Patient 2:**"The UI is simple and well-designed. The features shown seem quite useful."
- **Patient 3:** "There should be an option to rate the doctors; otherwise, it is quite good."

8.3.2 Feedback From Doctors

- **Doctor 1 (Dr. Md. Golam Zel Asmaul Husna):** "The prototype of your planned app is nicely designed; however, you should definitely include more details in the patient profile. For instance, how long they have been suffering from cancer, prescription history etc."
- **Doctor 2 (Dr X):** "Try to add a feature that will convert the text-to-voice and also try to implement a translation feature. Besides that, the design looks simple and easy to use."

8.3.3 Comparison With Other mHealth Applications

While numerous mHealth applications are available for people in Bangladesh, only a few are cancer-focused. For instance, there is an application called "Doctime" which mainly provides online healthcare services such as video consultations, healthcare packages, online medicine delivery etc. However, unlike our proposed prototype, Doctime does not have any slot booking system, bot feature, and cancer detection system. On the other hand, there is another app called "Maya", which focuses on woman's healthcare specially for issues like family planning and mental health. In comparison, our proposed prototype is solely focused on Cancer patients only.

Chapter 9

Conclusion and Future Works

Every year a huge number of people are being diagnosed with cancer all over the world. Therefore, we have worked with the oncologists and cancer patients in Bangladesh to know about their present conditions, financial status, food requirements and restrictions in order to provide the necessary support for the treatment and the problems faced while communicating with doctors. Alongside working with the oncologists and patients all over Dhaka, we have also analyzed different deep neural network architectures such as VGG16, MobileNetV3, InceptionV3, Xception and DenseNet121 and found out the model that gives the best accuracy to correctly identify the histopathological image of infected cancer cells. Moreover, we compiled and fitted the models in different cancer histopathological image datasets. Through our investigation and research, we presented the best-fit architecture which will correctly detect the histopathological images of not only one type of cancer but also every type of cancer. On the other hand, through our survey targeting cancer patients and oncologists, we also developed a potential prototype with all the necessary features to build our CancerCare mobile application which will help oncologists and patients deal with this vital disease.
Our future work involves mainly the following parts, which we are hoping to further work on:

- First, we will conduct our survey throughout Bangladesh and collect as much data as possible from oncologists and patients so that we can accurately find out the present demographic situations, health conditions of cancer patients, requirements and problems faced by both parties.

- Simultaneously, we will be working on completing our proposed mobile application, "CancerCare", and will also be implementing the necessary features based on the survey data. For instance, a bot feature to answer the basic queries of the patients. Besides that, we will also be integrating the cancer detection features in our to-be-built application with the help of google cloud API.

- After developing the application with its basic feature, we plan to make an initial release to a small group of participants for user testing. Based on the feedback of those respondents, we will keep improving our app and deploy it to more participants for further study.

- Finally, we are also planning to incorporate some additional features, such as a fingerprint scanner for easy login and a private cloud to secure and store all the patient data.

Bibliography

[1] J. L. Bender, R. Y. K. Yue, M. J. To, L. Deacken, and A. R. Jadad, "A lot of action, but not in the right direction: Systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer," *Journal of medical Internet research*, vol. 15, no. 12, e2661, 2013. DOI: 10.2196/jmir.2661. [Online]. Available: https://doi.org/10.2196/jmir.2661.

[2] A. A. O'Kane, H. M. Mentis, and E. Thereska, "Non-static nature of patient consent: Shifting privacy perspectives in health information sharing," in *Proceedings of the 2013 conference on Computer supported cooperative work*, 2013, pp. 553–562. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2441776.2441838.

[3] B. Rolland and C. P. Lee, "Beyond trust and reliability: Reusing data in collaborative cancer epidemiology research," in *Proceedings of the 2013 conference on Computer supported cooperative work*, 2013, pp. 435–444. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2441776.2441826.

[4] M. Jacobs, J. Clawson, and E. D. Mynatt, "Cancer navigation: Opportunities and challenges for facilitating the breast cancer journey," in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, 2014, pp. 1467–1478. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2531602.2531645.

[5] H.-C. Wu, C.-J. Chang, C.-C. Lin, M.-C. Tsai, C.-C. Chang, and M.-H. Tseng, "Developing screening services for colorectal cancer on android smartphones," *Telemedicine and e-Health*, vol. 20, no. 8, pp. 687–695, 2014. DOI: 10.1089/tmj.2013.0288. [Online]. Available: https://doi.org/10.1089/tmj.2013.0288.

[6] M. L. Jacobs, J. Clawson, and E. D. Mynatt, "Comparing health information sharing preferences of cancer patients, doctors, and navigators," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2015, pp. 808–818. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2675133.2675252.

[7] A. D. Torrance, S. L. Powell, and E. A. Griffiths, "Emergency surgery in the elderly: Challenges and solutions," *Open Access Emergency Medicine: OAEM*, vol. 7, p. 55, 2015. DOI: 10.2147/OAEM.S68324. [Online]. Available: https://doi.org/10.2147/OAEM.S68324.

[8] K. Caine, "Local standards for sample size at chi," in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 981–992. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2858036.2858498.

[9] M. Nagaiah and K. Ayyanar, “Software for data analysis in spss on over view,” Feb. 2016. [Online]. Available: https://www.researchgate.net/publication/348740813_Software_for_Data_Analysis_in_SPSS_On_over_view.

[10] Apr. 2017. [Online]. Available: http://uis.unesco.org/en/country/bd.

[11] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html.

[12] K. A. Kessel, M. M. Vogel, C. Kessel, *et al.*, “Mobile health in oncology: A patient survey about app-assisted cancer care,” *JMIR mHealth and uHealth*, vol. 5, no. 6, e7689, 2017. DOI: 10.2196/mhealth.7689. [Online]. Available: https://doi.org/10.2196/mhealth.7689.

[13] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, “An ensemble of fine-tuned convolutional neural networks for medical image classification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 31–40, 2017. DOI: 10.1109/JBHI.2016.2635663. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7769199/.

[14] G. Giunti, D. Giunta, E. Guisado-Fernandez, J. Bender, and L. Fernández-Luque, “A biopsy of breast cancer mobile applications: State of the practice review,” *International journal of medical informatics*, vol. 110, pp. 1–9, 2018. DOI: https://doi.org/10.1016/j.ijmedinf.2017.10.022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386505617304021.

[15] S. Dabeer, M. M. Khan, and S. Islam, “Cancer diagnosis in histopathological image: Cnn based approach,” *Informatics in Medicine Unlocked*, vol. 16, p. 100 231, 2019, ISSN: 2352-9148. DOI: https://doi.org/10.1016/j.imu.2019.100231. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352914819301133.

[16] R. Lipson-Smith, F. White, A. White, *et al.*, “Co-design of a consultation audio-recording mobile app for people with cancer: The secondears app,” *JMIR Form Res*, vol. 3, no. 1, e11111, Mar. 2019, ISSN: 2561-326X. DOI: 10.2196/11111. [Online]. Available: https://formative.jmir.org/2019/1/e11111/.

[17] 2020. [Online]. Available: https://gco.iarc.fr/today/data/factsheets/populations/50-bangladesh-fact-sheets.pdf.

[18] D. H. Charbonneau, S. Hightower, A. Katz, *et al.*, “Smartphone apps for cancer: A content analysis of the digital health marketplace,” *Digital Health*, vol. 6, p. 2 055 207 620 905 413, 2020. DOI: 10.1177/2055207620905413. [Online]. Available: http://dx.doi.org/10.1177/2055207620905413.

[19] J. Berger-Groch, M. Keitsch, A. Reiter, S. Weiss, K. Frosch, and M. Priemel, “The use of mobile applications for the diagnosis and treatment of tumors in orthopaedic oncology–a systematic review,” *Journal of medical systems*, vol. 45, no. 11, pp. 1–7, 2021. DOI: 10.1007/s10916-021-01774-z. [Online]. Available: https://doi.org/10.1007/s10916-021-01774-z.

[20] S. Bhochhibhoya, P. D. Dobbs, and S. B. Maness, “Interventions using mhealth strategies to improve screening rates of cervical cancer: A scoping review,” *Preventive Medicine*, vol. 143, p. 106 387, 2021. DOI: 10.1016/j.ypmed.2020.106387. [Online]. Available: https://doi.org/10.1016/j.ypmed.2020.106387.

[21] S. Deshmukh, N. Umredkar, E. Sharma, and R. Chalke, “Smart doctor android application for breast cancer risk prediction and diagnosis,” Apr. 2021. [Online]. Available: https://www.researchgate.net/publication/360877161_SMART_DOCTOR_ANDROID_APPLICATION_FOR_BREAST_CANCER_RISK_PREDICTION_AND_DIAGNOSIS.

[22] M. R. Hoque, E. Haque, and M. R. Karim, “Cervical cancer in low-income countries: A bangladeshi perspective,” *International Journal of Gynecology & Obstetrics*, vol. 153, no. 1, pp. 19–25, 2021. DOI: doi.org/10.1002/ijgo.13400. [Online]. Available: https://doi.org/10.1002/ijgo.13400.

[23] M. Kamruzzaman, *Central database, mass awareness to address breast cancer in bangladesh*, 2021. [Online]. Available: https://www.aa.com.tr/en/asia-pacific/central-database-mass-awareness-to-address-breast-cancer-in-bangladesh/2401288.

[24] D. Resnick, M. M. Schapira, J. M. Smith, *et al.*, “Promoting collaborative goal setting for cancer prevention among primary care patients through mhealth: Mixed methods evaluation of a new app,” *JMIR Formative Research*, vol. 5, no. 7, e22510, 2021. DOI: 10.2196/22510. [Online]. Available: https://doi.org/10.2196/22510.

[25] K. Sundberg, V. Lindström, L.-M. Petersson, and A. Langius-Eklöf, “Supporting health literacy using an interactive app for symptom management during radiotherapy for prostate cancer,” *Patient Education and Counseling*, vol. 104, no. 2, pp. 381–386, 2021. DOI: 10.1016/j.pec.2020.08.003. [Online]. Available: https://doi.org/10.1016/j.pec.2020.08.003.

[26] E. C. Wilson, E. M. Jalil, C. M. Jalil, *et al.*, “Results from a peer-based digital systems navigation intervention to increase hiv prevention and care behaviors of young trans women in rio de janeiro, brazil,” *Journal of Global Health Reports*, vol. 5, e2021077, 2021. [Online]. Available: https://www.joghr.org/article/28347.pdf.

[27] A. Abduokhapov, “Early detection of breast cancer using deep neural network classification models on histopathological images,” 2022. [Online]. Available: http://nur.nu.edu.kz/handle/123456789/6196.

[28] *Cancer*, Feb. 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer.

[29] *Cervical cancer*, Feb. 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer#:~:text=Key%5C%20facts,%5C%2Dincome%5C%20countries%5C%20(1)..

[30] D. Chowdhury, A. Das, A. Dey, *et al.*, “Abcandroid: A cloud integrated android app for noninvasive early breast cancer detection using transfer learning,” *Sensors*, vol. 22, no. 3, p. 832, 2022, ISSN: 1424-8220. DOI: 10.3390/s22030832. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/832.

[31] D. Ç. Ertuğrul and S. A. Abdullah, “A decision-making tool for early detection of breast cancer on mammographic images,” *Tehnički vjesnik*, vol. 29, no. 5, pp. 1528–1536, 2022, ISSN: 1330-3651. DOI: 10.17559/TV-20211221131838. [Online]. Available: https://doi.org/10.17559/TV-20211221131838.

[32] T. Sangers, S. Reeder, S. van der Vet, *et al.*, “Validation of a market-approved artificial intelligence mobile health app for skin cancer screening: A prospective multicenter diagnostic accuracy study,” *Dermatology*, pp. 1–8, 2022. DOI: 10.1159/000520474. [Online]. Available: https://www.karger.com/Article/PDF/520474.

[33] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. J. DVM, “Cancer statistics, 2022,” *CA: A Cancer Journal for Clinicians*, 2022. DOI: 10.3322/caac.21708. [Online]. Available: https://doi.org/10.3322/caac.21708.

[34] B. Subramanian, S. Muthusamy, K. Thangaraj, *et al.*, “A new method for detection and classification of melanoma skin cancer using deep learning based transfer learning architecture models,” 2022. DOI: 10.21203/rs.3.rs-1857063/v1. [Online]. Available: https://doi.org/10.21203/rs.3.rs-1857063/v1.

[35] A. G. Toprak, D. Ünay, T. Cerit, and K. Üğüdücü, “Cancerous lesion segmentation for early detection of breast cancer by using cnn,” 2022. [Online]. Available: https://www.researchgate.net/publication/362568994_Cancerous_Lesion_Segmentation_for_Early_Detection_of_Breast_Cancer_by_using_CNN.

[36] *Worldwide cancer data*, Apr. 2022. [Online]. Available: https://www.wcrf.org/cancer-trends/worldwide-cancer-data/.

[37] *Low income countries 2023*, 2023. [Online]. Available: https://worldpopulationreview.com/country-rankings/low-income-countries.

[38] *Lung and colon cancer histopathological images*. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images.

[39] *Lung cancer classification: Best models comparison*. [Online]. Available: https://www.kaggle.com/code/dreamtim1/lung-cancer-classification-best-models-comparison.