Mini Project Report On

CANCER DETECTION USING IMAGE PROCESSING AND TRANSFER LEARNING

Submitted in partial fulfilment of the requirements for the award of the

Bachelor of Technology

In

Computer Science and Engineering

By

Ashley Edgar Dcunha	**20241A05O6**
Chelmela Sai Harshith	**20241A05P2**
Rachapally Pavan Kumar	**20241A05S1**
Sasidhara Kashyap Chaturvedula	**20241A05S5**
Vellore Anand Kumar Sahil	**20241A05T5**

Under the Esteemed guidance of

Ms. K. Anusha

Asst.Professor

Department of Computer Science and Engineering

GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY

(Approved by AICTE, Autonomous under JNTUH, Hyderabad)

Bachupally, Kukatpally, Hyderabad-500090

Mini Project Report On

CANCER DETECTION USING IMAGE PROCESSING AND TRANSFER LEARNING

Submitted in partial fulfilment of the requirements for the award of the

Bachelor of Technology

In

Computer Science and Engineering

By

Ashley Edgar Dcunha	**20241A05O6**
Chelmela Sai Harshith	**20241A05P2**
Rachapally Pavan Kumar	**20241A05S1**
Sasidhara Kashyap Chaturvedula	**20241A05S5**
Vellore Anand Kumar Sahil	**20241A05T5**

Under the Esteemed guidance of

Ms. K. Anusha

Asst.Professor

Department of Computer Science and Engineering

GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY

(Approved by AICTE, Autonomous under JNTUH, Hyderabad)

Bachupally, Kukatpally, Hyderabad-500090

GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY
(Autonomous)

Hyderabad-500090

CERTIFICATE

This is to certify that the mini project entitled "**Cancer Detection Using Image Processing and Transfer Learning**" is submitted by **Ashley Edgar Dcunha(20241A05O6), Sai Harshith Chelmela(20241A05P2), Rachapally Pavan Kumar(20241A05S1), Sasidhara Kashyap Chaturvedula(20241A05S5), Vellore Anand Kumar Sahil(20241A05T5)** in partial fulfillment of the award of degree in BACHELOR OF TECHNOLOGY in Computer Science and Engineering during the Academic year 2022-2023.

Internal Guide	**Head of the Department**
Ms. K.Anusha	**Dr. K. Madhavi**
Assistant Professor	**Professor-HOD**

External Examiner

ACKNOWLEDGEMENT

Many people helped us, directly and indirectly, to complete our project successfully. We would like to take this opportunity to thank one and all. First, we would like to express our deep gratitude towards our internal guide **Ms.KAnusha,** Department of Computer Science and Engineering, for his/her support in the completion of our dissertation. We wish to express our sincere thanks to **Dr. K. Madhavi,** Head of the Department**,** and our principal **Dr. J. PRAVEEN,** for providing the facilities to complete the dissertation. We would like to thank all our faculty and friends for their help and constructive criticism during the project period. Finally, we are very much indebted to our parents for their moral support and encouragement to achieve goals.

Ashley Edgar Dcunha(20241A05O6)

Sai Harshith Chelmela(20241A05P2)

Rachapally Pavan Kumar(20241A05S1)

Sashidhar Kashyap Chaturvedula(20241A05S5)

Vellore Anand Kumar Sahil(20241A05T5)

DECLARATION

We hereby declare that the mini project titled "Cancer Detection Using Image Processing and Transfer Learning" is the work done during the period from 12th Jan 2023 to 3rd June 2023 and is submitted in the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering from Gokaraju Rangaraju Institute of Engineering and Technology (Autonomous under Jawaharlal Nehru Technology University, Hyderabad). The results embodied in this project have not been submitted to any other University or Institution for the award of any degree or diploma.

Ashley Edgar Dcunha(20241A05O6)
Sai Harshith Chelmela(20241A05P2)
Rachapally Pavan Kumar(20241A05S1)
Sashidhar Kashyap Chaturvedula(20241A05S5)
Vellore Anand Kumar Sahil(20241A05T5)

	Table of Contents	
Chapter	**TITLE**	**Page No**

List of Figures

List of Tables

ABSTRACT

The objective of this project is to develop a cancer detection system that utilizes image augmentation techniques, specifically the ImageDataGenerator, and MobileNetV2, a transfer learning approach. The system focuses on identifying the presence of nevus, melanoma, and basal cell carcinoma. By applying image augmentation techniques, we aim to enhance the dataset and improve the model's ability to generalize to different variations of skin lesions. MobileNetV2, a pre-trained deep learning model, is used as a feature extractor to leverage its learned representations for accurate classification. Transfer learning enables us to fine-tune the MobileNetV2 model using a dataset specifically curated for nevus, melanoma, and basal cell carcinoma detection. Performance evaluation metrics, including accuracy, precision, recall, and F1 score, are employed to assess the effectiveness of the proposed approach. The developed cancer detection system can be deployed on mobile devices, allowing for convenient and accessible diagnosis. The combination of image augmentation and transfer learning with MobileNetV2 enables real-time or near-real-time inference on resource-constrained devices. This system holds great potential in assisting medical professionals with early cancer detection, leading to improved patient outcomes and reduced healthcare costs. Further research can explore the extension of this approach to detect additional types of skin cancer and expand the system's capabilities.

CHAPTER 1
Introduction

Our project is focused on the development of a system for the detection of cancer. We used image processing and transfer learning to analyze images of skin cancer and identify the types of cancer.

Early detection plays a critical role in improving the prognosis and treatment outcomes for individuals diagnosed with cancer. By detecting cancer at an early stage, healthcare professionals can initiate timely interventions, resulting in better patient care and increased chances of successful treatment.

Our system incorporates image processing techniques and machine learning algorithms to enhance and analyze the images of skin lesions. These techniques enable us to extract valuable information and patterns from the images, which are then used to determine the likelihood of cancer presence.

To achieve accurate results, our system utilizes pre-trained models and transfer learning. This approach enables us to leverage existing knowledge and adapt it to the specific task of cancer detection. By building upon the foundations of established models, we can optimize our system's performance even with limited training data.

In conclusion, our project endeavours to contribute to the field of cancer detection by developing an advanced system that utilizes image processing and machine learning techniques. By providing a reliable and efficient method for early cancer detection, we aim to improve patient outcomes and support healthcare professionals in making informed decisions.

1.1 Existing

The base paper focuses on developing a cancer detection system using image processing and machine learning techniques. The methodology involved two main steps: image processing and machine learning.

In the image processing step, the biopsy images were enhanced to improve their quality and segmented into different regions of interest. Important features such as shape and texture were extracted from these regions to understand the characteristics of cancer cells.

The machine learning step utilized the Naive Bayes classifier to train the system on a dataset of annotated images. The system learned to distinguish between cancerous and noncancerous cells based on the extracted features. Testing was performed to evaluate the system's performance in classifying new images as positive or negative.

The overall objective of this work was to create an automated system that accurately detects and classifies cancer cells in biopsy images. By combining image processing and machine learning, the system aimed to provide efficient and reliable cancer detection.

1.1.1 Limitations in Existing System

1. Sensitivity to image quality: The accuracy of the system heavily relies on the quality of the biopsy images, and lower-quality images may impact its performance.
2. Unable to identify specific cancer types: While the system can determine if cancer is present or not in the images, it may not have the capability to classify the specific type of cancer.

1.2 Proposed System

We propose a cancer detection method that uses image processing and MobileNetV2, a transfer learning technique. Our system aims to detect nevus, melanoma, and basal cell carcinoma.
We begin by normalizing the Dermoscopic images depicting skin lesions. Then, we employ MobileNetV2, a pre-trained deep learning model, to extract meaningful features from the images.
The performance of our approach is evaluated using accuracy, precision, recall, and F1 score metrics on a separate test dataset.
Our system can run on web pages, enabling easy and accessible cancer diagnosis.
By leveraging transfer learning with MobileNetV2, we achieve real-time inference on resource-constrained devices. This system has the potential to assist in early cancer detection, leading to better patient outcomes and reduced healthcare costs.
Future work can explore expanding the system to detect other types of skin cancer.

1.2.1 Advantages of Existing System

1. Accurate detection: The use of advanced image processing techniques and transfer learning with MobileNetV2 enhances the accuracy of cancer detection, reducing the chances of false positives and false negatives.
2. Efficient utilization of resources: Transfer learning with MobileNetV2 allows the system to leverage pre-trained models, which saves computational resources and reduces the need for extensive training on large datasets.
3. Mobile compatibility: The system is designed to run on mobile devices, making it convenient and accessible for medical professionals. It enables on-the-go cancer detection, eliminating the need for specialized equipment or dedicated workstations.
4. Time-efficient: The efficient MobileNetV2 architecture enables faster cancer detection, ensuring timely diagnosis and potentially improving patient outcomes.

CHAPTER 2
Literature Survey

1. Title: "Deep learning-based detection of malignant melanoma using convolutional neural networks"

Authors: Esteva, Andre, et al.

Published: Nature, 2017

Summary: This study presents a deep learning approach for the detection of malignant melanoma, a deadly form of skin cancer. The authors utilize a convolutional neural network (CNN) to analyze dermoscopic images and distinguish between benign and malignant lesions. The proposed model achieves comparable accuracy to dermatologists, demonstrating the potential of AI in melanoma detection.

2. Title: "Automated Melanoma Recognition in dermoscopy images via very deep residual networks"

Authors: Han, Seung Shin, et al.

Published: IEEE Transactions on Medical Imaging, 2017

Summary: The authors propose a deep learning model based on very deep residual networks (ResNet) for automated melanoma recognition in dermoscopy images. The ResNet architecture enables the extraction of complex features from images, improving the accuracy of melanoma detection. Experimental results show promising performance and highlight the effectiveness of deep learning in this domain.

3. Title: "Skin cancer recognition using ensemble deep learning with weighted average fusion"

Authors: Wadhawan, Tarun et al.

Published: Journal of Medical Systems, 2018

Summary: In this study, the authors propose an ensemble deep-learning framework for skin cancer recognition. They combine multiple deep-learning models using a weighted average fusion technique to enhance classification accuracy. Experimental results demonstrate the effectiveness of the ensemble approach and its potential for accurate detection of skin cancer.

4. Title: "Automated cervical cancer screening using deep learning in colposcopy images"

Authors: Hu, Lifang, et al.

Published: Journal of Biomedical Optics, 2018

Summary: This paper presents an automated approach for cervical cancer screening using deep learning applied to colposcopy images. The authors develop a CNN model to analyze the images and detect abnormal cervical cells indicative of cervical cancer. The results show high sensitivity and specificity, suggesting the potential of deep learning in improving the efficiency and accuracy of cervical cancer screening programs.

5. Title: "Glioma classification using transfer learning from deep convolutional neural networks"

Authors: Havaei, Mohammad, et al.

Published: Medical Image Analysis, 2017

Summary: This research investigates the classification of brain gliomas, a type of brain tumor, using transfer learning from deep convolutional neural networks. The authors fine-tune pre-trained CNN models on a dataset of MRI images to differentiate between different grades of gliomas. The study demonstrates the effectiveness of transfer learning in achieving accurate glioma classification, offering a promising approach for assisting in brain tumor diagnosis.

CHAPTER 3
Software Requirement Specifications

3.1 Introduction

3.1.1 Purpose of the Requirement document

The purpose of the requirement document for the cancer detection system using image processing and transfer learning with MobileNetV2 is to clearly explain why we need this system and what it should do. The document helps everyone involved understand the goals and requirements of the project.

Clear Communication: We want to make sure that everyone understands the purpose, scope, and expected outcomes of the project. The document helps avoid confusion and ensures that all stakeholders are on the same page.

Define Requirements: We need to identify and document what the system should be able to do. This includes the specific features, functions, and performance criteria that the system must meet.

Guide Development: The document serves as a guide for the development team. It provides them with the necessary information and instructions on how to build the system, including technical specifications and integration requirements.

Support Decision Making: Stakeholders need to make informed decisions throughout the project. The requirement document helps them by providing a detailed analysis of the system's requirements, constraints, risks, and dependencies, enabling them to make the right choices.

Validation and Verification: We want to ensure that the system we develop meets the specified requirements and achieves its intended goals. The requirement document acts as a reference for validation and verification activities, helping us confirm that the system is working as expected.

3.1.2 Scope of the Product

The scope of the project is to develop a cancer detection system using image processing and transfer learning with MobileNetV2. The system will accept input images of cancer cells and employ image processing techniques to preprocess the images. This may include resizing, and normalization to optimize the images for subsequent analysis. Transfer learning with the MobileNetV2 model will be applied to extract relevant features from the preprocessed images. The original classification layer of MobileNetV2 will be replaced with a new classification layer specific to cancer detection. The system will provide predictions on whether the cells in the images are Basal Cell Carcinoma, Melanoma, or Nevus. It will be developed to achieve high accuracy, and considerations will be given to portability, performance, usability, and data security to ensure a robust and reliable cancer detection solution.

3.1.3 Definitions, acronyms, and Abbreviations

Machine Learning (ML): A branch of artificial intelligence that enables computer systems to learn and make predictions or decisions without being explicitly programmed.

Software Requirements Specification (SRS): A document that describes the requirements and specifications of a software system, including its functionalities, interfaces, and constraints.

Acronyms and abbreviations

AI: Artificial Intelligence
ML: Machine Learning
SRS: Software Requirements Specification
UI: User Interface.

3.1.4 Reference

1) https://ieeexplore.ieee.org/document/8471295
2) https://www.manning.com/books/deep-learning-with-python
3) https://scikit-image.org/docs/stable/api/api.html#module-skimage.transform
4) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5568765/

3.1.5 Overview

We are working on a special system that helps doctors detect three types of cancer cells(Basal Cell Carcinoma, Nevus, and Melanoma) in microscopic images. We use image processing and transfer learning with a model called MobileNetV2. These techniques help us understand important features in the images that can indicate cancer. Instead of looking at the whole picture, we focus on specific parts that might contain cancer cells. This helps us make more accurate predictions. For each cell, we can tell if it is cancerous or not. We want our system to be very accurate in detecting cancer cells, and work quickly. By combining these techniques, we hope to improve cancer detection and assist doctors in providing better care to patients.

3.2 General Description

Our project aims to create a system that can detect different types of cancer. We will make the system smarter by adding variations to the images, like flipping or changing their size. This helps the system learn better and become more reliable.
We will use machine learning, which is a way for computers to learn from examples, to teach the system how to recognize cancer. By showing it many images of cancer and non-cancer cases, the system will learn to identify the signs of cancer in new images. We will test the system to make sure it works well. We will use different measurements to see how accurate it is in detecting cancer. This will help us make any necessary improvements to the system.
By creating this system, we hope to make it easier for doctors to detect cancer early on. Early detection can lead to better treatment and improved outcomes for patients.

3.2.1 Product Perspective

The cancer detection system using image processing and transfer learning with MobileNetV2 is a standalone software application. It operates independently, without relying on other systems or components. The system analyses skin images to detect cancer cells and provides predictions. From a user's perspective, the system will have a user-friendly interface, allowing healthcare professionals to easily interact with it. Users can input images and receive results indicating the presence of cancer cells. It will handle images securely and ensure data protection during transmission and storage.

3.2.2 Product Functions

Image Upload: Users can upload medical images of cells suspected of having cancer into the system.

Pre-processing: The system applies image processing techniques to enhance the quality of the uploaded images, ensuring optimal conditions for analysis.

Feature Extraction: Utilizing transfer learning with MobileNetV2, the system extracts relevant features from the pre-processed images to capture important characteristics indicating cancerous or non-cancerous cells.

Classification: The system employs a customized classification layer to classify cells as cancerous or non-cancerous based on the extracted features.

User Interface: The system provides a user-friendly interface, allowing healthcare professionals to easily interact with the system, upload images, and view results

3.2.3 User Characteristics

The primary users of the system include Medical Professionals, Patients, and Doctors. The users must have basic Knowledge of computer literacy and can be able to understand the characteristics of cancerous cells.

3.2.4 General Constraints

These constraints can affect how we develop and implement our system. One constraint is the availability and quality of data. We need to make sure we have access to reliable data related to our project's objectives. Another constraint is the computational resources we have. We need to consider the limitations of our hardware and software in processing and analyzing large amounts of data. Time is also a constraint for us. If there are disruptions in data collection it can impact the performance of our system. We should ensure that our project is scalable and adaptable to handle future changes. We need to be prepared to integrate new technologies, accommodate evolving requirements, and work well with existing systems or frameworks.

3.2.5 Assumptions and Dependencies

Assumptions:
We assume that we will have access to a sufficient amount of reliable and relevant data that is needed for training and evaluating our models. This data will play a vital role in ensuring the effectiveness and accuracy of our system. We assume that the collected data will undergo necessary preprocessing.

Dependencies:
Our project relies on evaluation metrics and procedures to assess the performance and effectiveness of our models. These metrics and procedures will allow us to measure the performance of our system and its intended purpose. They will help us identify any areas that require improvement or optimization.

3.3 Specific Requirement

3.3.1 Software Requirement

- Operating System : Windows 8+
- Server-side Script : Python 3.7+
- IDE : PyCharm
- Libraries used : pandas, numpy, cv2, os, random, PIL, TensorFlow,keras, keras.preprocessing, sklearn.metrics

3.3.2 Hardware Requirement

- Processor : Intel i3+
- RAM : 6GB+
- Hard Disk : 64GB
- Monitor : Any

3.3.3 Functional Requirements

- **Image analysis:** The system must be able to analyze medical images and detect cancer.

- **Transfer learning**: The system must be able to use transfer learning to improve accuracy and performance.

- **Image Augmenting:** The system must be able to Augment medical images and increase the dataset virtually.

- **Data storage:** The system must have a secure database for storing medical images and results.

- **Reporting:** The system must be able to generate reports with visual representations of the results and diagnostic information.

3.3.4 Non-Functional Requirements

- **Performance:** The system must deliver fast and accurate results, and the performance should be continuously monitored and improved.

- **Scalability:** The system must be scalable to handle an increasing volume of medical images as the project grows and evolves.

- **Security:** The system must adhere to strict data privacy regulations and secure medical images to protect patient confidentiality.

- **User experience:** The system must have an intuitive and user-friendly interface for medical professionals to access and analyze images.

- **Reliability:** The system must be reliable and consistently deliver accurate results.

3.4 Feasibility Study

Our project, which aims to develop a cancer detection system using image processing and transfer learning, undergoes a feasibility study to assess its viability and potential success. The study focuses on several key aspects:

3.4.1 Technical Feasibility

Firstly, image processing techniques, such as resizing and normalization, are well-established and widely implemented, ensuring the efficient preprocessing of input images. Secondly, MobileNetV2, a popular and efficient deep learning model, has been successfully used in various computer vision tasks, making it a suitable choice for feature extraction in cancer cell classification. Transfer learning allows leveraging the pre-trained MobileNetV2 model, saving time and computational resources during training. Additionally, libraries and frameworks such as TensorFlow and Keras provide robust support for implementing image processing and deep learning algorithms. GPU acceleration can be utilized to speed up image processing and classification tasks, enabling real-time or near-real-time performance. Compatibility with different operating systems and platforms ensures portability and accessibility. However, the success and accuracy of the system are dependent on the availability of high-quality and diverse training data, as well as continuous evaluation and refinement of the model.

3.4.2 Economic Feasibility

The implementation of such a system can yield significant economic benefits. Early and accurate detection of cancer can lead to timely treatment, reducing healthcare costs associated with advanced stages of the disease. By automating the detection process, the system can enhance the efficiency of healthcare providers, allowing them to analyze a larger volume of images in less time. This, in turn, can improve patient throughput and reduce waiting times. The use of transfer learning with the MobileNetV2 model enables leveraging existing pre-trained models, minimizing the need for extensive computational resources and reducing development and training time, thus lowering overall costs. Additionally, the system's portability and compatibility with various platforms make it accessible to a wider range of healthcare facilities. However, there might be initial costs associated with acquiring suitable hardware and software infrastructure, as well as ongoing costs for maintenance, updates, and the acquisition of diverse and high-quality training data. Proper cost-benefit analysis and assessment of potential savings should be conducted to ensure the long-term economical feasibility of the system.

3.4.3 Operational Feasibility

The availability of established image processing techniques and pre-trained deep learning models, such as MobileNetV2, ensures that the system can be implemented effectively. These techniques have been extensively studied and widely used in various applications, indicating their reliability and operational readiness. Additionally, the use of transfer learning allows for leveraging the expertise and knowledge captured by the pre-trained model, enabling faster development and reducing the need for extensive training on large datasets. The system can be designed with a user-friendly interface, making it easy for healthcare professionals to interact with and interpret the results. Moreover, the system's ability to process images in real-time or near-real-time ensures timely and efficient analysis, enhancing operational efficiency in healthcare settings. However, proper integration with existing healthcare systems and workflows should be considered, and appropriate training and support should be provided to ensure the successful adoption and operation of the system within the intended environment.

CHAPTER 4
Design

4.1 Project Description

In our project, we focus on developing a system for cancer detection using Image processing and Transfer Learning. We aim to improve the accuracy and reliability of cancer diagnosis by leveraging advanced technologies.

We start by working with a dataset of skin lesion images, which are known to be indicative of various types of cancer. To enhance the dataset and make our model more reliable, we employ image augmentation techniques. This involves applying different transformations to the images, such as rotation, scaling, and flipping. By doing so, we create a larger and more diverse set of images for training our model.

Next, we utilize Transfer learning algorithms, specifically the MobileNetV2 architecture, for training and classification tasks. The model is trained to learn patterns and features associated with different types of skin cancer, including nevus, melanoma, and basal cell carcinoma. The goal is to accurately classify new images as either nevus, melanoma, or basal cell carcinoma based on the patterns learned during training.

4.1.1 System UseCase Diagrams

4.1.1.1 Data Collection and Preprocessing Module

We look for a suitable data set from different resources. After acquiring the dataset we download it to the local drive. Later we apply various Image pre-processing techniques such as rescaling, randomizing, Normalizing, and augmenting.

Rescaling: In this, we change all the images to a particular size.

Normalizing: we apply constant scaling to normalize the data

Augmentation: Here we add images that are randomly rotated, flip or zoom in/out to increase the dataset size virtually

Further, we save the pre-processed data.

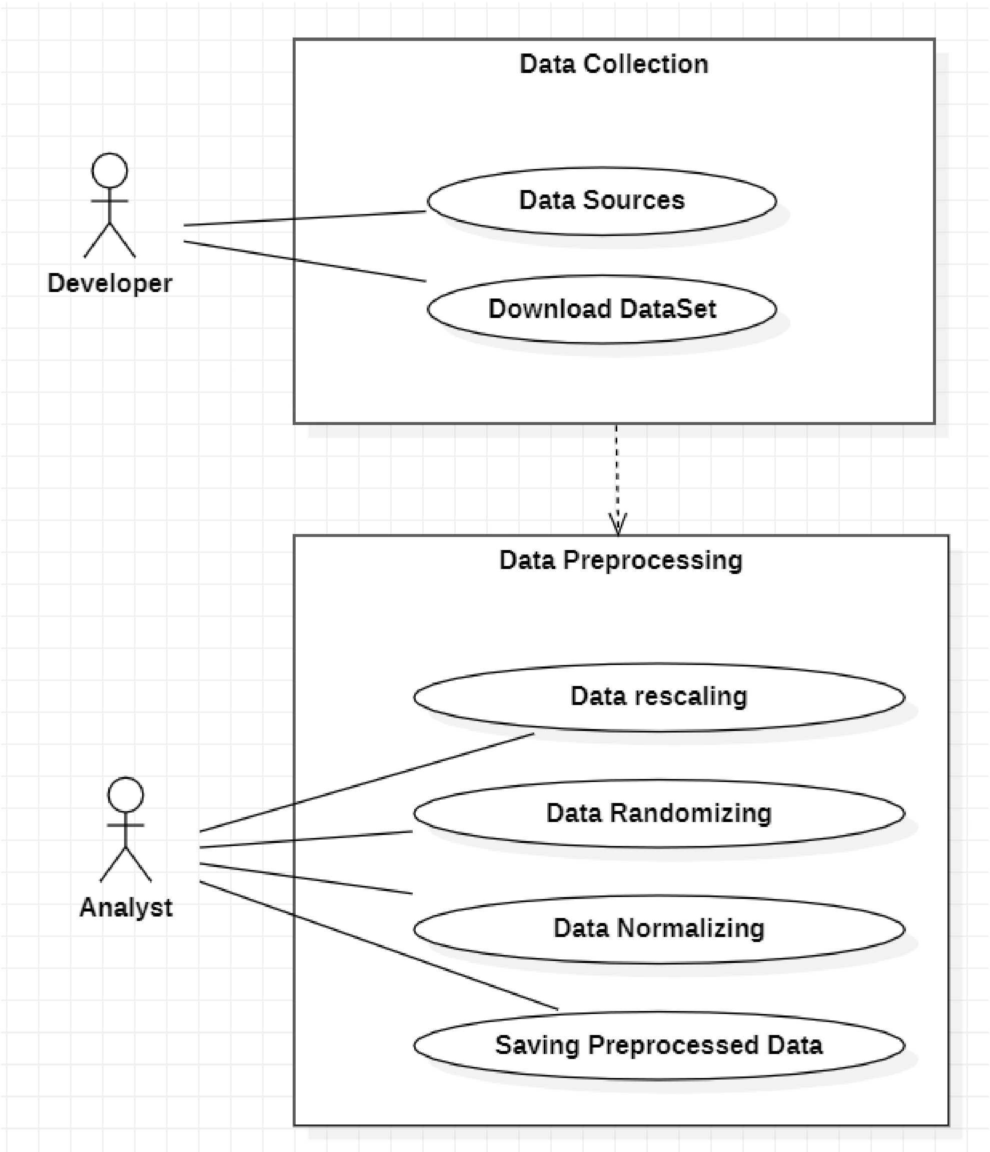


Fig.1 Data Collection and Preprocessing Usecase Diagram

4.1.1.2 Model Development

Select the type of model. Add the dense and drop layers to your model giving certain values. Later fit the model on the dataset to obtain a trained model which can predict the type of cancer. Save the model with the .h5 extension. Use accuracy, precision, f1-score, and support to evaluate the matrix.

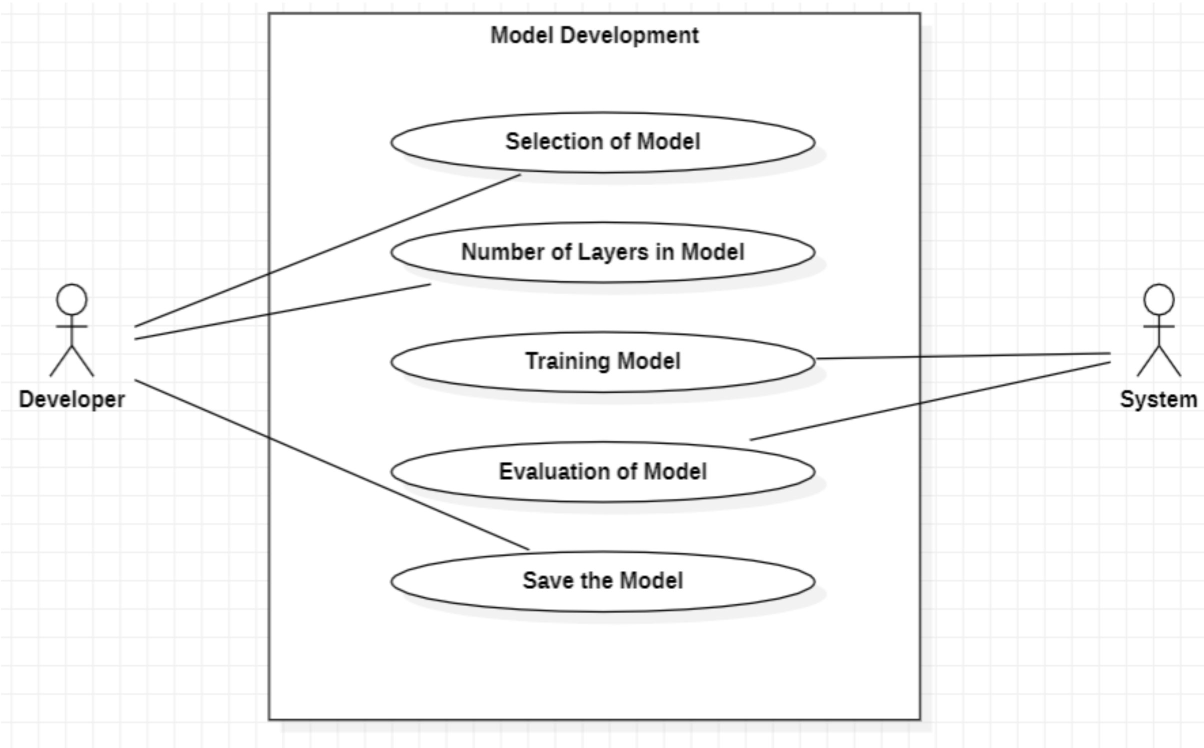


Fig. 2 Model Development UseCase Diagram

4.1.1.3 Web Interface

An image is given to the web page as an input. It uses Python to preprocess such as rescaling and normalizing the image. It loads the model which was trained on the dataset and predicts the result. The result is then displayed on the web interface.

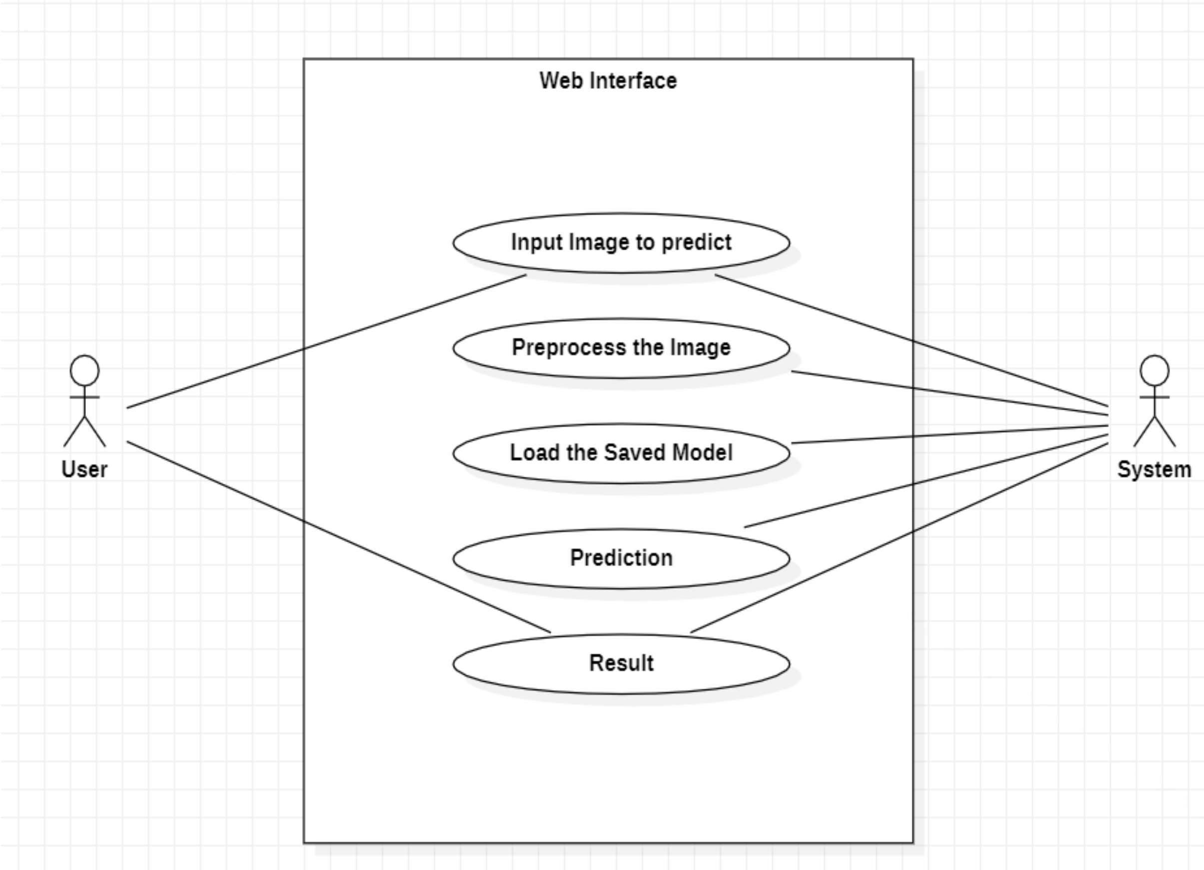


Fig. 3 Web Interface Usecase Diagram

4.1.2 Class Diagrams

The different classes in our project are Cancer Detection, import modules, load dataset, explore data, image processing, transfer learning, evaluation, model.h5, and Interface.
In import modules, we import all the required libraries.
In load dataset, we load the dataset into the program.
In image processing, we apply a few image processing techniques, such as resizing, normalizing, and image augmentation.
Transfer learning class consists of loading MobileNetV2 pre-trained model.
Few evaluations are performed in the evaluation class such as precision, recall, and f1-score.
Model.h5 consists of load model operation.
Coming to the interface we have a load model, input image, and display prediction.

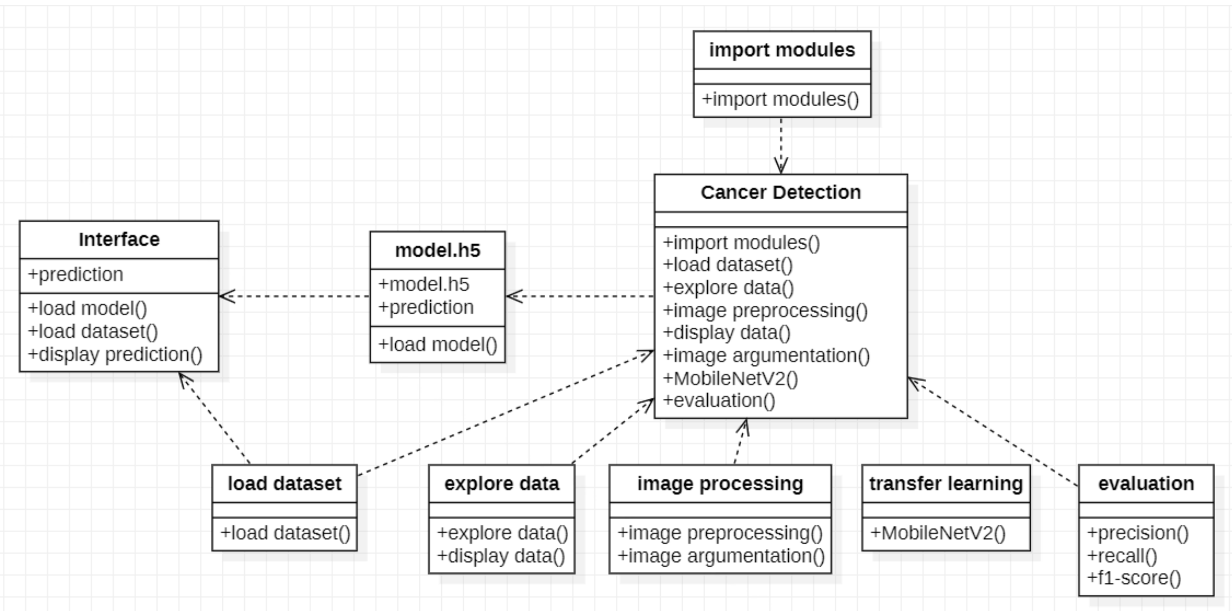


Fig. 4 Class Diagram

4.1.3 Data Flow Diagram (DFD)

The user provides the image to upload so that prediction can be made. After the image is uploaded, it is sent to the trained model for prediction. Predictions are then displayed to the user.

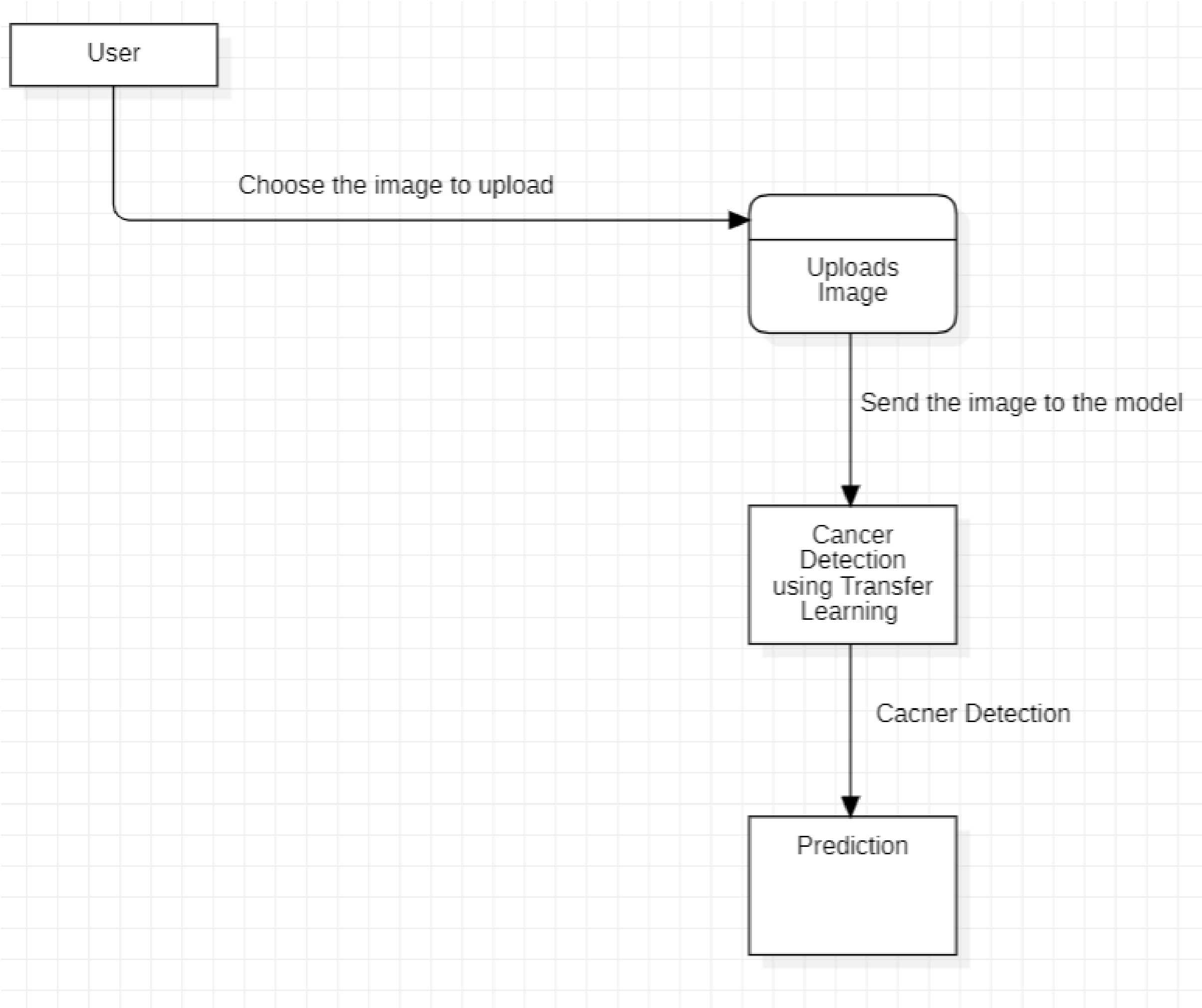


Fig. 5 Data Flow Diagram

CHAPTER 5
Implementation

5.1 Source Code

Data Preparation - Loading Images, Labels and Resizing

In [ ]:
```
data=[]
labels=[]
Basal=os.listdir("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Basal_cell_carcinoma/train/")
for a in Basal:
    try:
        image=cv2.imread("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Basal_cell_carcinoma/train/"+a)
        image_from_array = Image.fromarray(image, 'RGB')
        size_image = image_from_array.resize((224, 224))
        data.append(np.array(size_image))
        labels.append(0)
    except AttributeError:
        print("")

Melanoma=os.listdir("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Melanoma/train//")
for b in Melanoma:
    try:
        image=cv2.imread("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Melanoma/train/"+b)
        image_from_array = Image.fromarray(image, 'RGB')
        size_image = image_from_array.resize((224, 224))
        data.append(np.array(size_image))
        labels.append(1)
    except AttributeError:
        print("")
Nevus=os.listdir("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Nevus/train//")
for c in Nevus:
    try:
        image=cv2.imread("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/Nevus/train/"+c)
        image_from_array = Image.fromarray(image, 'RGB')
        size_image = image_from_array.resize((224, 224))
        data.append(np.array(size_image))
        labels.append(2)
    except AttributeError:
        print("")
```

In [8]:
```
#converting features and labels in array
feats=np.array(data)
labels=np.array(labels)

# saving features and labels for later re-use
np.save("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/feats_train",feats)
np.save("C:/Users/pavan/Downloads/ISIC_2019_Training_Input/labels_train",labels)
```

Fig. 6.1 Source Code Data Preparation and labelling

Train Test Split

In [15]:
```
# splitting cells images into 80:20 ratio i.e., 80% for training and 20% for testing purpose
(x_train,x_test)=feats[(int)(0.2*len_data):],feats[:(int)(0.2*len_data)]

(y_train,y_test)=labels[(int)(0.2*len_data):],labels[:(int)(0.2*len_data)]
```

Image Data Normalization

In [16]:
```
x_train = x_train.astype('float32')/255 # As we are working on image data we are normalizing data by dividing 255.
x_test = x_test.astype('float32')/255
train_len=len(x_train)
test_len=len(x_test)
```

In [17]:
```
y_train=to_categorical(y_train,3)
y_test=to_categorical(y_test,3)
```

Fig. 6.2 Source Code Train Test Split and Normalization

Image Augmentation

In [16]:
```
trainAug  = ImageDataGenerator(
featurewise_center=False,  # set input mean to 0 over the dataset
        samplewise_center=False,  # set each sample mean to 0
        featurewise_std_normalization=False,  # divide inputs by std of the dataset
        samplewise_std_normalization=False,  # divide each input by its std
        zca_whitening=False,  # apply ZCA whitening
        rotation_range=10,  # randomly rotate images in the range (degrees, 0 to 180)
        zoom_range = 0.1, # Randomly zoom image
        width_shift_range=0.1,  # randomly shift images horizontally (fraction of total width)
        height_shift_range=0.1,  # randomly shift images vertically (fraction of total height)
        horizontal_flip=False,  # randomly flip images
        vertical_flip=False)
```

Fig. 6.3 Source code Image Augmentation

Model Building

In [17]:
```
conv_base = MobileNetV2(
    include_top=False,
    input_shape=(224, 224, 3),
    weights='imagenet')

for layer in conv_base.layers:
    layer.trainable = True
```

```
Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/mobilenet_v2/mobilenet_v2_weights_tf_dim_ordering_tf_kernels_1.0_224_no_top.h5
9406464/9406464 [==============================] - 3s 0us/step
```

In [18]:
```
x = conv_base.output
x = layers.GlobalAveragePooling2D()(x)
x = layers.Dense(256, activation='relu')(x)
x = layers.Dropout(0.2)(x)
x = layers.Dense(64, activation='relu')(x)
x = layers.Dropout(0.1)(x)
predictions = layers.Dense(3, activation='softmax')(x)
model = Model(conv_base.input, predictions)
```

In [18]:
```
callbacks = [ModelCheckpoint('.mdl_wts.hdf5', monitor='val_loss',mode='min',verbose=1, save_best_only=True),
             ReduceLROnPlateau(monitor='val_loss', factor=0.3, patience=2, verbose=1, mode='min', min_lr=0.00000000001)]
```

In [20]:
```
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
BS = 64
print("[INFO] training head...")
H = model.fit(
    trainAug.flow(x_train,y_train, batch_size=BS),
    steps_per_epoch=train_len // BS,
    validation_data=(x_test, y_test),
    validation_steps=test_len // BS,
    epochs=30,callbacks=callbacks)
```

Fig. 6.4 Source Code Model Building

5.2 Result Screenshots:

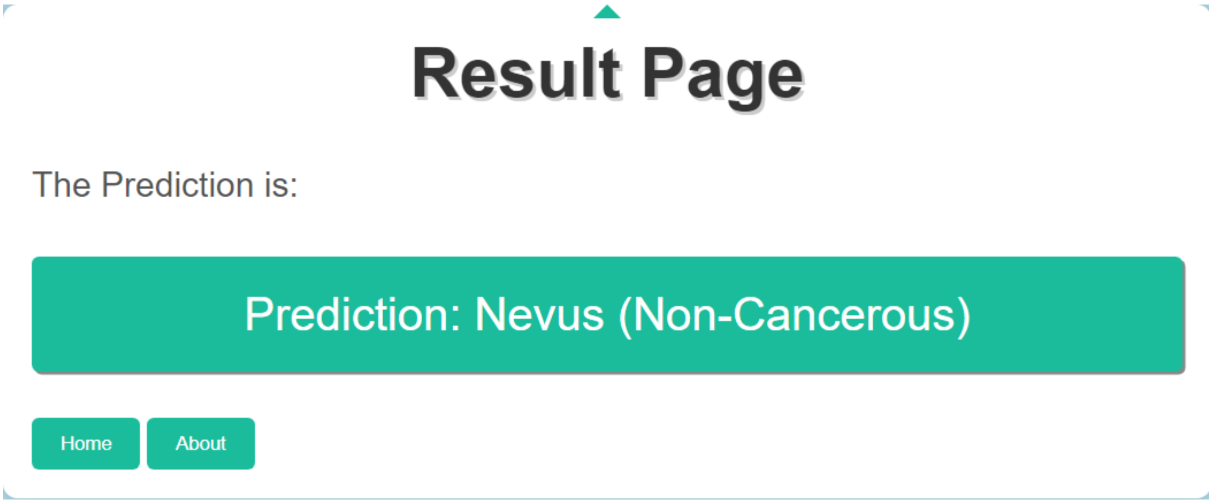


Fig. 7.1 Result Screenshot

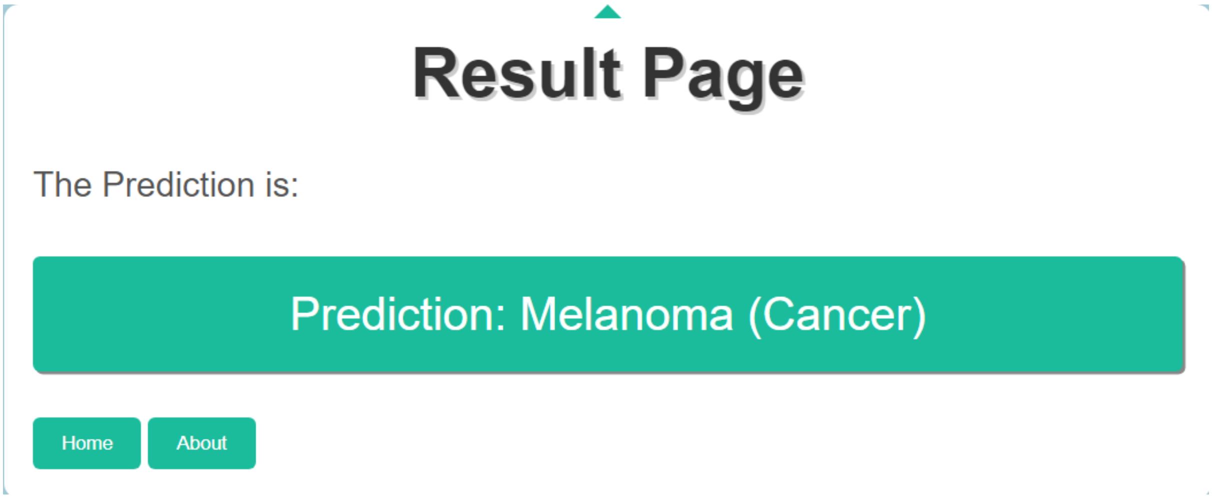


Fig. 7.2 Result Screenshot

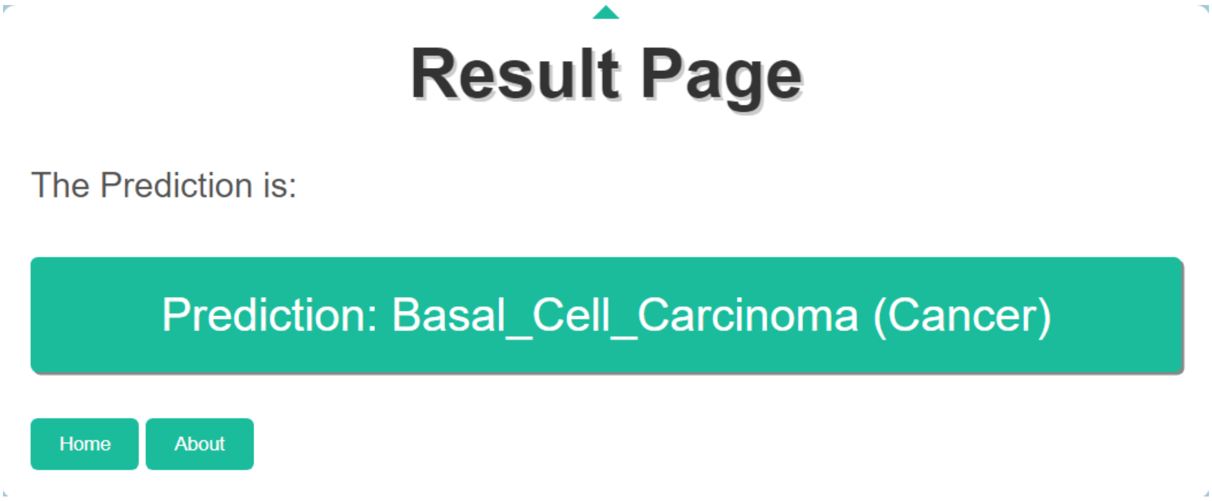
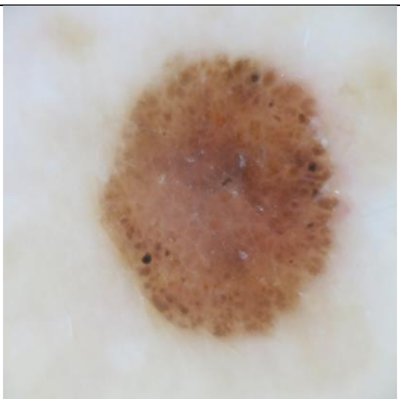
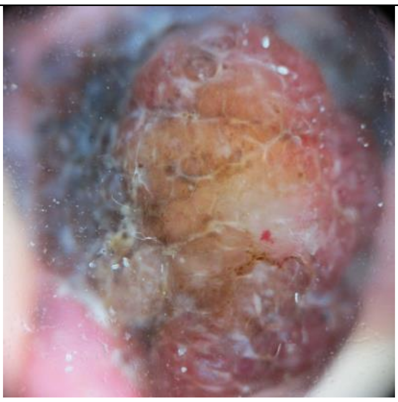
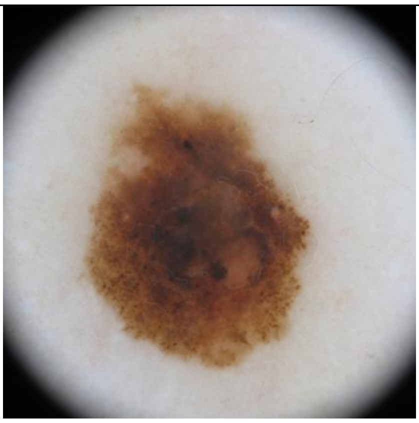
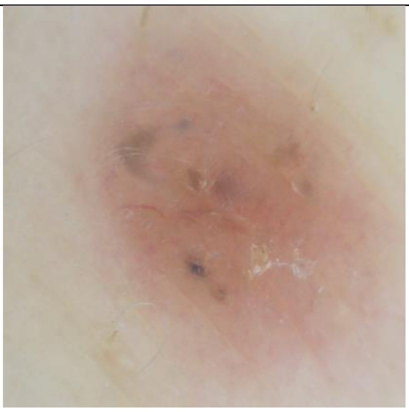
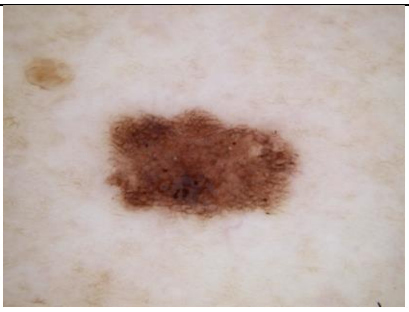


Fig. 7.3 Result Screenshot

CHAPTER 6
Testing

Test Case id	Test Case	Expected Result	Actual Result	Status Pass/Fail
1.		Basal Cell Carcinoma	Basal Cell Carcinoma	Pass
2.		Nevus	Nevus	Pass
3.		Melanoma	Melanoma	Pass
4.		Nevus	Nevus	Pass

5.		**Basal Cell Carcinoma**	**Basal Cell Carcinoma**	**Pass**
6.		**Melanoma**	**Melanoma**	**Pass**
7.		**Nevus**	**Nevus**	**Pass**
8.		**Melanoma**	**Melanoma**	**Pass**

9.		**Basal Cell Carcinoma**	**Basal Cell Carcinoma**	**Pass**
10.		**Melanoma**	**Nevus**	**Fail**

CHAPTER 7
Conclusion and Future Scope

In conclusion, our project on cancer detection using image augmentation (ImageDataGenerator) and MobileNetV2 (transfer learning) has successfully developed a system capable of detecting nevus, melanoma, and basal cell carcinoma. By enhancing and standardizing dermoscopic images, we prepared the data for analysis and leveraged the power of MobileNetV2, a pre-trained model, for our specific task. Our experiments demonstrated that the system accurately detects the targeted cancer types.

The system offers several advantages, including its ability to run on mobile devices, making it convenient and accessible for diagnosis. MobileNetV2 also enables fast detection on resource-constrained devices without compromising accuracy. This system has the potential to assist medical professionals in early cancer detection, leading to better outcomes for patients and reduced healthcare costs.

Moving forward, there are several avenues for further improvement and exploration. Firstly, incorporating more advanced feature extraction techniques could enhance the system's performance by capturing additional relevant information from the images. In addition, we are exploring the use of advanced learning models called convolutional neural networks to improve the accuracy and reliability of cancer detection. CNNs are powerful tools that can help the system better recognize and classify cancer-related features in skin lesion images.

Also, we know it's important to include many different types of cancer images in our dataset. By having a wide variety of cancer images, the system can get better at learning and detecting different types of cancer accurately.

CHAPTER 8
References

1. **Paper:** "Skin Lesion Classification Using Deep Convolutional Neural Networks" by Esteva et al. (2017)
   Link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5568765/

2. **Paper:** "MobileNetV2: Inverted Residuals and Linear Bottlenecks" by Sandler et al. (2018)
   Link: https://arxiv.org/abs/1801.04381

3. **Website:** TensorFlow documentation on MobileNetV2
   Link: https://www.tensorflow.org/api_docs/python/tf/keras/applications/MobileNetV2

4. **Book:** "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville (MIT Press, 2016)
   Website: http://www.deeplearningbook.org/

5. **Paper:** "A Review on Deep Learning Techniques for Image Classification and Segmentation" by LeCun et al. (2015)
   Link: https://arxiv.org/abs/1508.06576

6. **Paper:** "A Comprehensive Survey of Deep Learning Techniques for Skin Cancer Classification" by Haenssle et al. (2018)
   Link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6152942/

7. **Paper:** "Skin Lesion Analysis Towards Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging" by Codella et al. (2018)
   Link: https://arxiv.org/abs/1710.05006

8. **Website:** Scikit-image documentation on image augmentation techniques
   Link: https://scikit-image.org/docs/stable/api/api.html#module-skimage.transform

9. **Book:** "Deep Learning with Python" by François Chollet (Manning Publications, 2018)
   Website: https://www.manning.com/books/deep-learning-with-python

10. **Paper:** "Cancer Detection Using Image Processing Techniques: A Comprehensive Review" by Jyoti and Mittal (2018)
    Link: https://ieeexplore.ieee.org/document/8471295