A

PROJECT REPORT ON

Medicinal Plant Identification using Android Application based on Leaf Image

SUBMITTED TO

SHIVAJI UNIVERSITY, KOLHAPUR

IN THE PARTIAL FULFILLMENT OF REQUIREMENT FOR THE AWARD OF DEGREE BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

Submitted by

MISS. PRAGATI MAHESH NANWANI	**19UCS087**
MISS. ADITI AVINASH MAGDUM	**20UCS303**
MISS. RUTUJA CHANDRAKANT PATIL	**20UCS305**
MISS. AYESHA MUBARAK ATTAR	**20UCS314**
MISS. SNEHAL SANJAY BARAD	**20UCS315**

Under The Guidance Of

PROF. T. I. BAGBAN

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

D.K.T.E. Society's

TEXTILE AND ENGINEERING INSTITUTE, ICHALKARANJI

2022-2023

D.K.T.E.SOCIETY'S
TEXTILE AND ENGINEERING INSTITUTE, ICHALKARANJI
(AN AUTONOUMOUS INSTITUTE)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that, project work entitled

"Medicinal Plant Identification using Android Application based on Leaf Image"

is a bonafide record of project work carried out in this college by

MISS. PRAGATI MAHESH NANWANI	**19UCS087**
MISS. ADITI AVINASH MAGDUM	**20UCS303**
MISS. RUTUJA CHANDRAKANT PATIL	**20UCS305**
MISS. AYESHA MUBARAK ATTAR	**20UCS314**
MISS. SNEHAL SANJAY BARAD	**20UCS315**

is in the partial fulfillment of award of degree Bachelor of technology in Computer Science & Engineering prescribed by Shivaji University, Kolhapur for the academic year 2022-2023.

PROF. T. I. Bagban
(Project Guide)

PROF. (DR.) D. V. KODAVADE
(HOD CSE DEPT.)

PROF. (DR.) MRS. L. S. ADMUTHE
(DIRECTOR)

EXAMINER: _____

DECLARATION

We hereby declare that, the project work report entitled "Medicinal Plant Identification Using Android Application Based on Leaf Image" which is being submitted to D.K.T.E. Society's Textile and Engineering Institute, Ichalkaranji affiliated to Shivaji University, Kolhapur is in partial fulfillment of degree B.Tech. (CSE). It is a bonafide report of the work carried out by us. The material contained in this report has not been submitted to any university or institution for the award of any degree. Further, we declare that we have not violated any of the provisions under Copyright and Piracy / Cyber / IPR Act amended from time to time.

NAME OF STUDENT	**PRN**	**SIGNATURE**
MISS. PRAGATI MAHESH NANWANI	19UCS087	
MISS. ADITI AVINASH MAGDUM	20UCS303	
MISS. RUTUJA CHANDRAKANT PATIL	20UCS305	
MISS. AYESHA MUBARAK ATTAR	20UCS314	
MISS. SNEHAL SANJAY BARAD	20UCS315	

ACKNOWLEDGEMENT

With great pleasure we wish to express our deep sense of gratitude to Mr. T. I. Bagban for his valuable guidance, support and encouragement in completion of this project report.

Also, we would like to take opportunity to thank our head of department Dr. D. V. Kodavade for his co-operation in preparing this project report.

We feel gratified to record our cordial thanks to other staff members of Computer Science and Engineering Department for their support, help and assistance which they extended as and when required.

Thank you,

NAME OF STUDENT	**PRN**
MISS. PRAGATI MAHESH NANWANI	19UCS087
MISS. ADITI AVINASH MAGDUM	20UCS303
MISS. RUTUJA CHANDRAKANT PATIL	20UCS305
MISS. AYESHA MUBARAK ATTAR	20UCS314
MISS. SNEHAL SANJAY BARAD	20UCS315

ABSTRACT

In Ayurveda medicine field accurate identification of medicinal plants is of great importance. Plants are identified by human experts according to their aroma and visual features. Manual identification of medicinal plants process is a very time consuming and requires the help of experts for plant identification. Incorrect identification of medicinal plants may lead to unfavorable results. To defeat this problem, automatic identification and classification of medicinal plants species is necessary for greater satisfaction to mankind. Plant identification and classification can be automated using visual morphological features such as the texture, color and shape of the leaves.

Here we develop a system which can identify the plant species based on the input leaf sample. MediLeaf presents how medicinal plants were identified with high accuracy by applying deep learning capabilities. For this, a dataset is used of scanned images of leaves of commonly used Ayurveda medicinal plants. Both the front and back sides of leaves were captured. The leaves are classified based on the unique characteristics combination. The plant species are classified by the features on each extracted leaf using CNN classifier with the accuracy of around 76 %.

MediLeaf provides this work as a modern smartphone application for leaf picture recognition tool for medicinal plants. The application operates on the Android operating system. MediLeaf has two key functions, i.e. recognition of medicinal plants and related information of that medicinal leaf. In this study, 30 species of medicinal plants were used and there are 40-50 digital pictures of their species.

Index

Chapter 1

Introduction

1.1 Problem Definition

Problem Statement

To design an automatic system to recognize and identify the medicinal plant leaf using CNN.

1.2 Aim and objectives

Aim:

To design an automatic system to recognize the medicinal plant leaf using CNN.

Objectives:

Non-Functional Requirements:

- Our primary goal is to design an automatic system to recognize the medicinal plant leaf using CNN.
- To develop a user-friendly interface.

Functional Requirements:

- To capture the image of leaves.
- To extract features from images.
- To train the model with a sample dataset of leaves.
- To build a model that identifies medicinal plants and displays basic information about identified plants such as local name, scientific name, origin of that plant, Main feature of that plant and medicinal use.
- To test the model with the test image.
- System will display the list of diseases cured by the identified plant.

1.3 Scope and limitations

Scope:

Our automated system will identify a few medicinal plant leaves. After identification of a medicinal plant, the system will display basic information about the plant such as local name, scientific name, origin of that plant, Main feature of that plant and medicinal use.

Our automated system will display the list of diseases that can be cured by an identified plant.

Limitations:

The Medicinal Plant Identification Using Android Application Based on Leaf Image has certain limitations, including:

- Only a few medicinal plants are identified.
- Our system identifies the medicinal plant only from characteristics of leaves of the plant.
- Our system's proposed methods are not suitable for tiny leaves or plants without a proper leaf.

1.4 Project Timeline

We started the project by gathering the related documents to the project at the end of July 2022. Gathering the requirements and all the analysis tasks was done by mid of August 2022. After that System design was started in the month of September 2022 and completed by the start of November along with the UML diagrams and Synopsis with a rough idea of the project.

In November 2022 we started making the detailed SRS documents along with deciding the methodology for the project which was completed by mid-December 2022. By the start of January 2023, we started coding by dividing it into 2 modules and completed the 1st module by the end of January 2023. Another module was completed by the end of March 2023. We started testing the project alongside designing the GUI which was completed in the first week of April 2023.

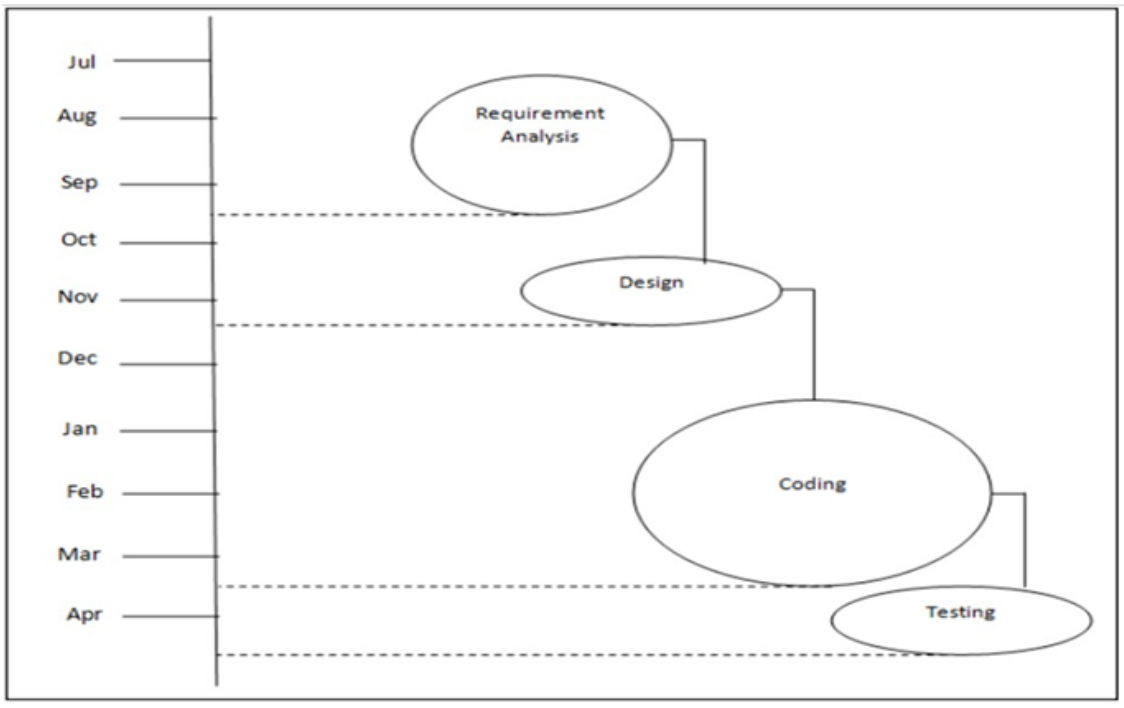


Figure 1.1: Project Timeline

1.5 Project Management Plan

Task	Period	Start Time	End Time	Priority
Domain Selection	7 days	15-07-2022	22-07-2022	High
Analysis of various problems suitable in the selected domain	15 days	23-07-2022	07-08-2022	High
Domain finalization	10 days	08-08-2022	18-08-2022	Medium
Problems detected	7 days	19-08-2022	26-08-2022	High
Research on various problem statements detected	10 days	27-08-2022	05-09-2022	Medium
Problem statement Finalize	14 days	06-09-2022	19-09-2022	Medium
Study of research papers	20 days	20-09-2022	09-10-2022	High
Synopsis Documentation	15 days	10-10-2022	24-10-2022	High
Requirement Collection	10 days	25-10-2022	03-11-2022	High
Module identification	10 days	04-11-2022	13-11-2022	Medium
SRS documentation Presentation	14 days	14-11-2022	28-11-2022	Medium
Collecting datasets	10 days	28-11-2022	08-12-2022	High
Study various algorithms	10 days	08-12-2022	18-12-2022	High
Coding implementation 30%	25 days	19-12-2022	14-01-2023	High
Coding Implementation 70%	30 days	15-01-2022	15-02-2023	High
Testing and accuracy improvements	20 days	16-02-2023	06-03-2023	High
Code updating	10 days	07-03-2023	17-03-2023	High
Coding 90%	20 days	18-03-2023	08-04-2023	High
Testing and Implementation 100%	8 days	09-04-2023	13-04-2023	High

Table 1.1: Project Management Plan

1.6 Project Cost

Sr. No.	Required Hardware / Software	Cost Rs.
1	Computer system with i7 11th generation or above	50000
2	8 GB or above RAM	5000
3	Python IDE to run machine learning modules.	0

Table 1.2: Hardware and Software Requirements

Estimated cost by considering other factors will be approximately – RS 65,300/-

COCOMO Model

In this Project, the Cost Estimation based on COCOMO (Constructive Cost Model) is calculated as below:

Line of code: To develop the system 398 lines of codes are required.

KLOC: KLOC is the estimated size of the software product indicates in Kilo Lines of Code.

Effort: The effort is only a function of the number of lines of code and some constants evaluated according to the different software systems.

Time: The amount of time required for the completion of the job, which is, of course, proportional to the effort put in. It is measured in the units of time such as weeks months.

Persons Required: Persons required is nothing but effort divide by time.

The coefficient ab, bb, cb and db are given below:

Software Project	Ab	bb	Cb	Db
Organic	2.4	1.05	2.5	0.38
Semi-Detached	3.0	1.12	2.5	0.35
Embedded	3.6	1.20	2.5	0.32

- KLOC: = LOC / 1000 = 405/1000 = 0.405

- Effort: = a * (KLOC) ^{b} = 2.4 * (0.405) $^{1.05}$ = 2.4 * 0.387 = 0.9288
- Time: = c *(Efforts) ^{d} = 2.5 * (0.9288) $^{0.405}$ = 2.5 * 0.97 = 2.425
- Persons Required: = Efforts / Time = 0.9288/2.425 = 0.38

Chapter 2

Background Study and Literature Overview

2.1 Literature Overview

Paper 1:

Title: Medicinal Plant Identification Using Android Application Based on Leaf Image.

Summary: MediLeaf-a mobile medicinal plant identification application centered on leaf picture has been created. The program has two key purposes, i.e. the recognition of a medicinal plant and the search for medicinal plants. MedPlant is a machine-aided medical plant identification device using visual processing technology, computer vision and smart information systems. In order to classify the image, they used local binary designs to extract the texture of the leaf and probabilistic neural network.[1]

Limitations: The recognition of medicinal plants is based only on the texture of the leaves. The exact illustration of medicinal plants based on leaf texture is 56.33% . It also allows people, organizations and populations to build unused and untapped expertise to utilize medicinal plant value. As a consequence, the power, money and economic prosperity would rise.

Paper 2:

Title: Real Time Identification of Medicinal Plant using Machine Learning Techniques.

Summary: The article proposed a robust technique using CNN for the identification of rare medicinal plants. Test accuracy of 90% was obtained by using the TensorFlow on the dataset that we created. This immunity was obtained from the extraction of the relevant features of the leaf image. The samples were formed using neural networks for the recognition of statistical models. The accuracy increases with the number of training which is meant by epoch. When the number of epoch increases, it can get high accuracy. [2]

Limitations: In this research, we determined the problem of identifying the species of medicinal plant by the examination of leaf images achieved directly from their habitat and disregarding of lighting conditions.

Paper 3:

Title: Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers

Summary: The trained model for the neural network is tested over the test set to get their individual classification accuracy. The regularization parameters change to achieve the best model features from the dataset. Accuracy of training set using this model is acquired up to 100%. This accuracy depends on the number of epochs that commits. While tested on 500 images each of 10 sparse medicinal plants, the test set accuracies were between 95% - 99%.[3]

Limitations: This research was only applied on 10 rare species of medicinal plant

2.2 Critical appraisal of existing research and projects

Paper 1:

Title: Medicinal Plant Identification Using Android Application Based on Leaf Image. Leaf pictures were evaluated using feature equations as textures. The features include grey textures, GTSDM, and Local Binary Pattern (LBP) operators. Three elements of the herb namely morphology, form and texture were used to classify medicinal plants. The recognition of medicinal plants was based only on the texture of the leaves.

They have used the Probabilistic Neural Network to identify medicinal plants. The features were combined using Product Decision Rule (PDR). They have also used Visual processing technology, computer vision and smart information systems. In order to classify the image, they have used local binary designs to extract the texture of the leaf and probabilistic neural network.

Paper 2:

Title: Real Time Identification of Medicinal Plant using Machine Learning Techniques. System could identify the plant species based on the input leaf sample. ExG-ExR was used to obtain more information from the images. The reason there was, it fixes a built-in zero threshold and hence there was no require to use otsu or any threshold value selected by the user. The ExG-ExR index identifies a binary plant region of interest. Using Logistic Regression classifier the plant species are classified by the color and texture features of each extracted leaf. The Dataset used for this work contains 5 classes of medicinal plant where 20 images in each class. 70% of the images is used as train data and 30% of the images is used as test data. The leaf in the input image is segmented using ExG-ExR method. The texture and color features can be extracted from the segmented leaf. The Weka is used to classify the medicinal plant species.

Paper 3:

Title: Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers This paper implements how rare medicinal plants were identified with high accuracy by applying concepts of image processing and machine learning proficiency. For this research, a database was created from collecting scanned images of flowers and leaves of sparse medicinal plants. Both the front side and back side of leaves and flowers were captured. The leaves are classified based on the unique characteristics combination. Identification accuracy rates up to 98% have been achieved when tested over 10 plants.

2.3 Investigation of current project and related work

This investigation focuses on the development and implementation of an Android application for medicinal plant identification based on leaf images. The project aims to provide a user-friendly tool that can assist individuals in identifying medicinal plants by simply capturing images of their leaves using a mobile device. The investigation involves various stages, including dataset collection, preprocessing, feature extraction, machine learning model training, and app development. The ultimate goal is to create an accurate and reliable application that can positively impact individuals' access to medicinal plants for healthcare and conservation purposes.

We started this project with the aim of identifying the leaf and displaying related information of that leaf.

In the previous system the recognition of medicinal plants is based only on the texture of the leaves. So we decided to create a system that recognizes medicinal plants based on Shape, vein, Texture, Color of the leaf.

In the previous system, the accuracy of medicinal plant identification based on leaf texture was 56.33%. In our current project the accuracy of medicinal plant identification is 76%.

"Leafsnap" by Columbia University, University of Maryland, and Smithsonian Institution: Leafsnap is an electronic field guide developed for iOS and Android devices. It enables users to identify tree species using photographs of their leaves. The system uses computer vision techniques, such as shape and vein pattern analysis, to match leaf images to a database of known species.

"Smart Leaf" by Smart Leaf Inc.: Smart Leaf is an Android application that utilizes machine learning algorithms for plant identification based on leaf images. It employs a deep learning model trained on a large dataset of leaf images to classify and identify plants accurately. The app provides users with detailed information about the recognized plant species.

These are just a few examples of existing Android applications and research projects related to medical plant identification based on leaf images. The field is rapidly evolving, and there may be newer works and advancements since my knowledge cutoff.

Chapter 3

Requirement Analysis

3.1 Requirement Gathering

- **User -**

  1. As a user, I want to capture leaf image of medicinal plants through Camera.
  2. As a user, I want to upload leaf image of medicinal plants from my gallery.
  3. As a user, I want to get an error message if an image is not uploaded successfully for some reason.
  4. As a user, I want to view the name of the correct medicinal plant of uploaded image.
  5. As a user, I want to view the additional information about the medicinal plant of uploaded image as local name, scientific name, origin of that plant, Main feature of that plant and medicinal use.
  6. As a user, I want to view the list of diseases cured by an identified plant.

- **User Stories:**
  User-

  1. Users should capture images through the camera.
  2. Users should upload images from the gallery.
  3. Users are able to view the additional information about the medicinal plant of the uploaded image as local name, scientific name, origin of that plant, Main feature of that plant and medicinal use.
  4. Users are able to view the list of diseases cured by an identified plant.

3.2 Requirement Specification

No.	Requirement	Essential/ Desirable	Description of Requirement	Remarks
RS1	The system must be able to capture leaf Image.	Essential	The leaf image captured by the user through camera must be stored by the system.	Phone camera should be in good resolution.
RS2	The system must be able to upload leaf image.	Essential	The leaf image uploaded by the user from gallery must be stored by the system.	Only leaf images are allowed
RS3	The system should be able to generate error message.	Essential	The system generates error message if an image is not uploaded successfully for some reason.	Error must be generated
RS4	The system must be able to identify correct medicinal plant	Essential	After uploading the leaf image system will displays name of the medicinal plant.	Result should be displayed
RS5	The system must be able to display additional Information and medical uses of identified Medicinal plant	Essential	Displays additional information about identified medicinal plant such as local name, scientific name, origin, Main feature of that medicinal plant.	Result should be displayed
RS6	The system must be able to display list of diseases cured by identified plant	Essential	Displays list of diseases cured by that identified plant	Result should be displayed

3.3 Use Case Diagram

Figure 3.1: Use Case Diagram

Chapter 4

System Design

4.1 Architectural Design

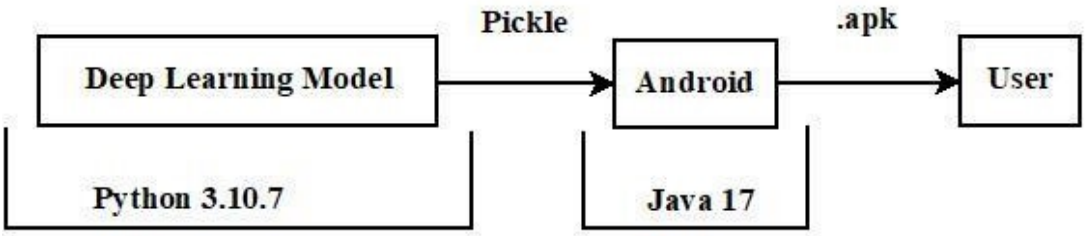
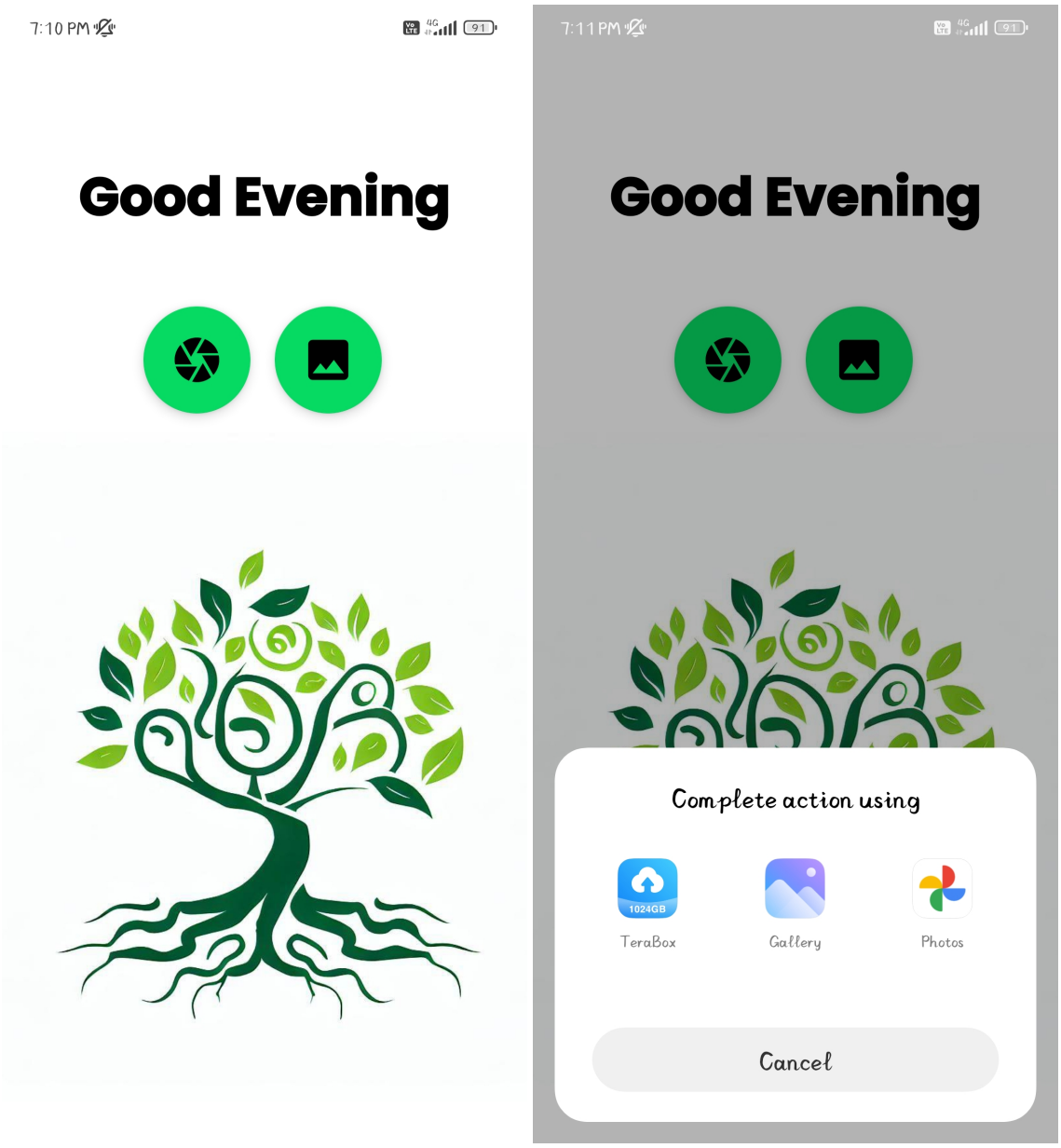


Figure 4.1: Architectural Design

4.2 User Interface Design



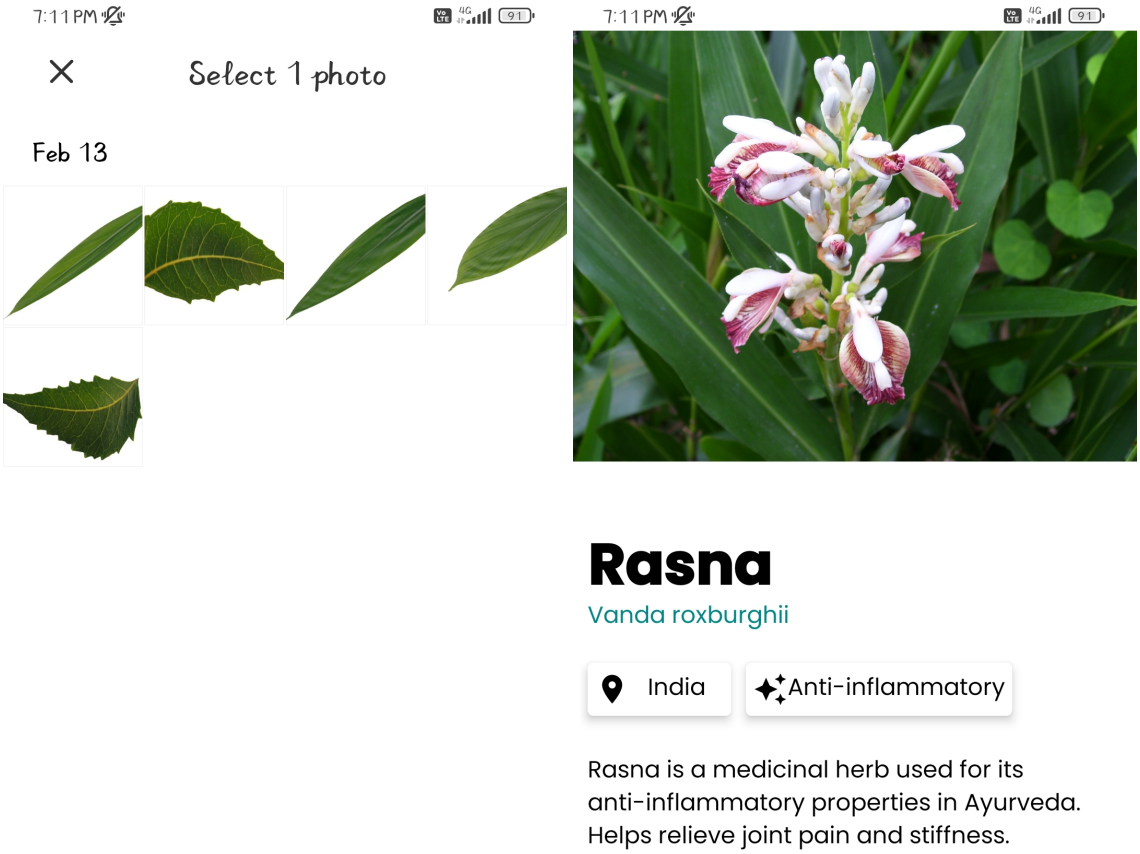


Figure 4.2: User Interface Design

4.3 Algorithmic description of each module:

Our project aims at identifying the leaf and giving related information about that leaf. It is composed of various modules. The modules present in our project are as follows:

1. Capture or upload image
2. Submit image
3. Identification of medicinal plant
4. Display information of identified medicinal plant
5. Showing medicinal uses of identified plant

1. **Capture or upload image:** Users will capture an image of a leaf through a camera or upload it from a gallery.
   Algorithm:
   (a) Open app.
   (b) Click on the camera icon or upload from the gallery icon.
   (c) If image upload successful
   (d) then image acquisition is successful
   (e) else
   (f) error: cannot get image at this moment

2. **Submit image:** Captured Image is given as an input to a deep learning model.
   Algorithm:
   (a) if leaf image acquisition done
   (b) then image preprocessing done
   (c) else
   (d) incorrect results

3. **Identification of medicinal plant:** After image processing, the deep learning model has to identify the medicinal plant which is in the given image.
   Algorithm:

(a) if image preprocessing done

(b) then system will identify accurate medicinal plant

(c) else

(d) system will identify medicinal plant incorrectly

4. **Display information of identified medicinal plant:** After identification of the medicinal plant our system will display its Local Name, Scientific Name etc.
   Algorithm:

   (a) if image identification done

   (b) then display Local Name, Scientific Name of the medicinal plant.

   (c) else

   (d) system will display medicinal plant information incorrectly

5. **Showing medicinal uses of identified plant:** System will display the list of diseases cured by the identified medicinal plant to the user.
   Algorithm:

   (a) if image identification done

   (b) then display list of diseases cured by leaf

   (c) else

   (d) system will unable to display list of diseases cured by leaf

4.4 System Modeling

4.4.1 Dataflow Diagram

DFD Level 0:

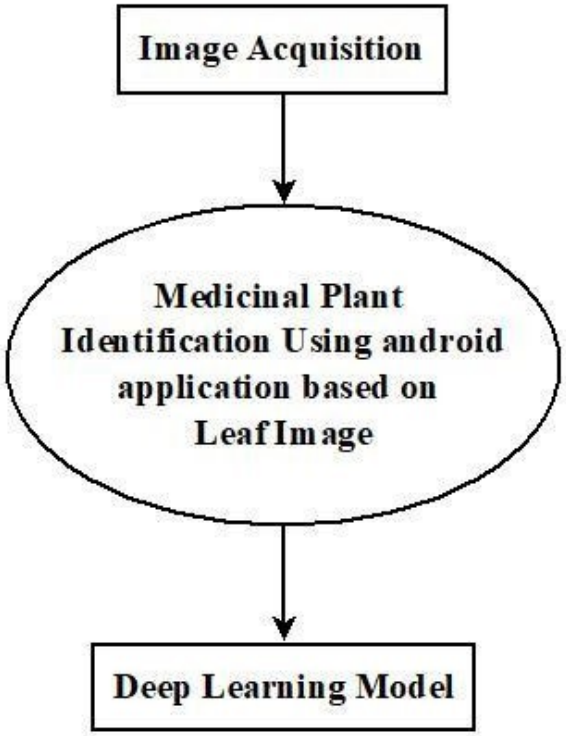


Figure 4.3: DFD Level 0

DFD Level 1:

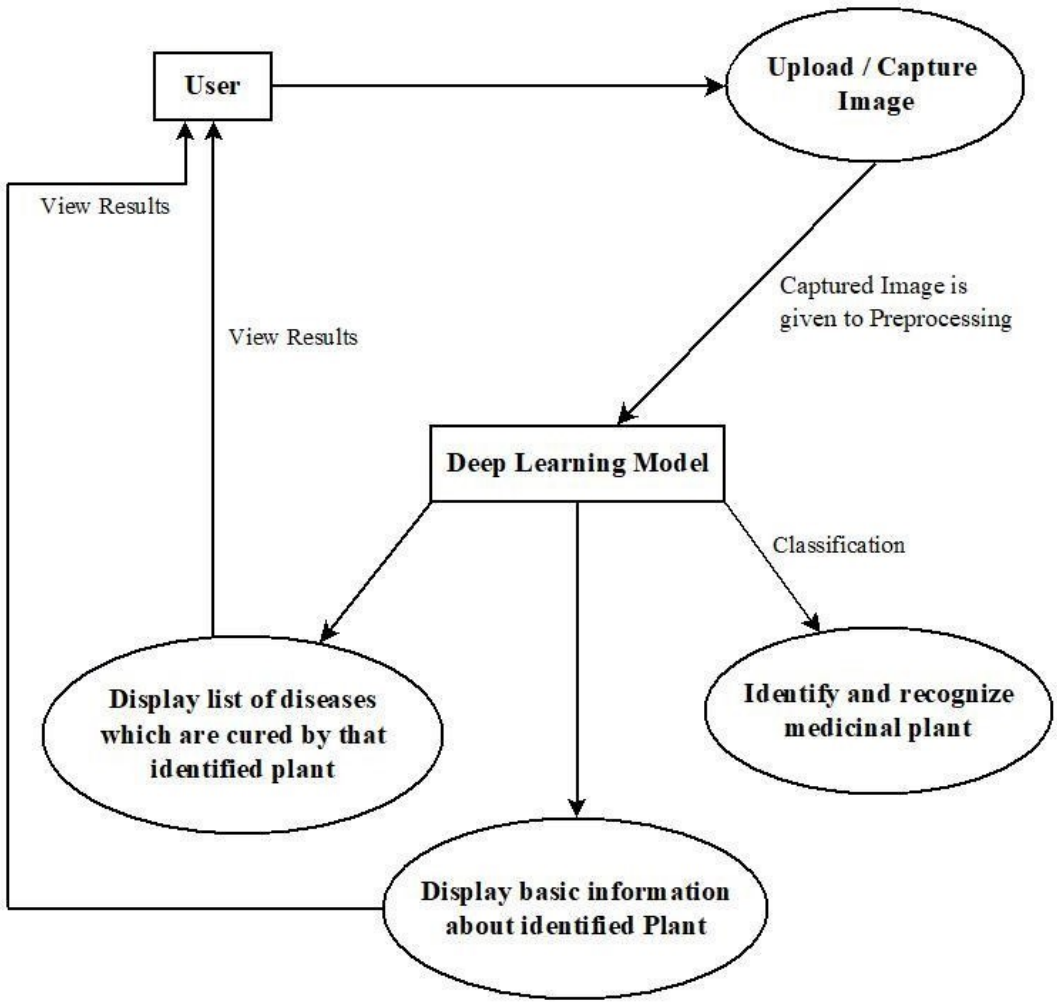


Figure 4.4: DFD Level 1

4.4.2 Sequence Diagram

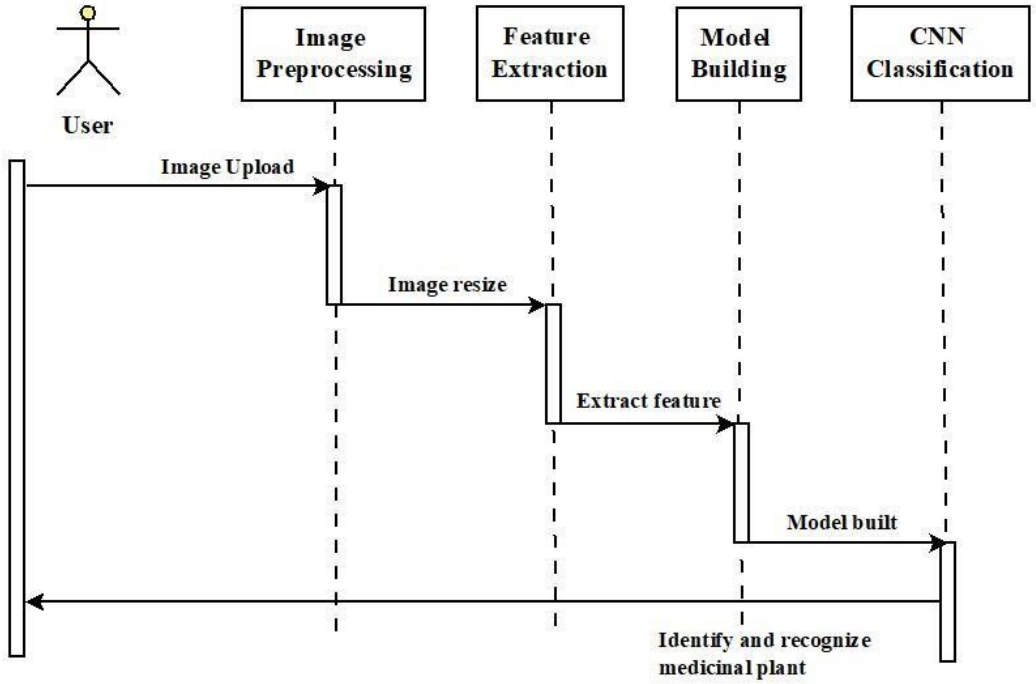


Figure 4.5: Sequence Diagram

4.4.3 Activity Diagram

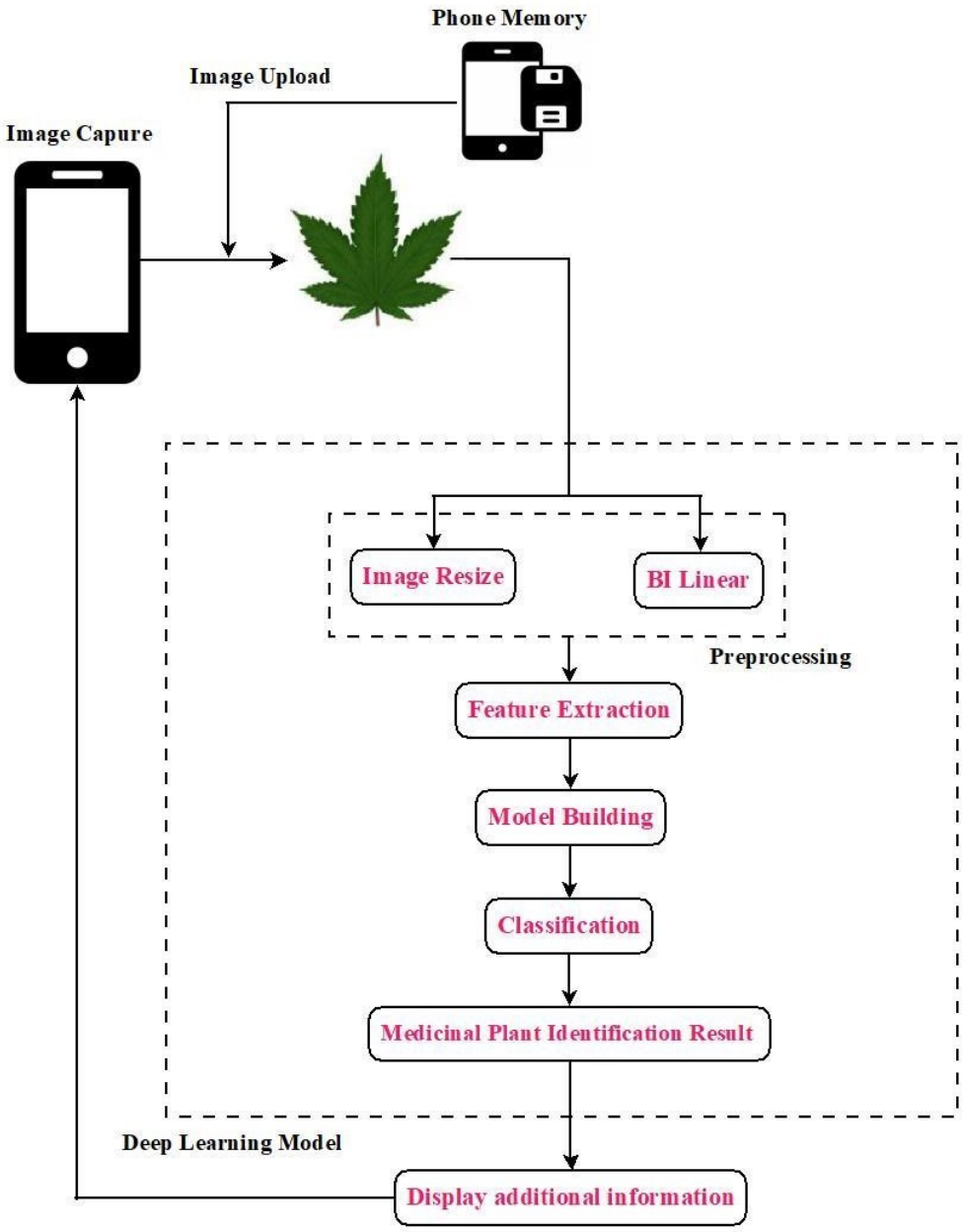


Figure 4.6: Activity Diagram

4.4.4 Component Diagram

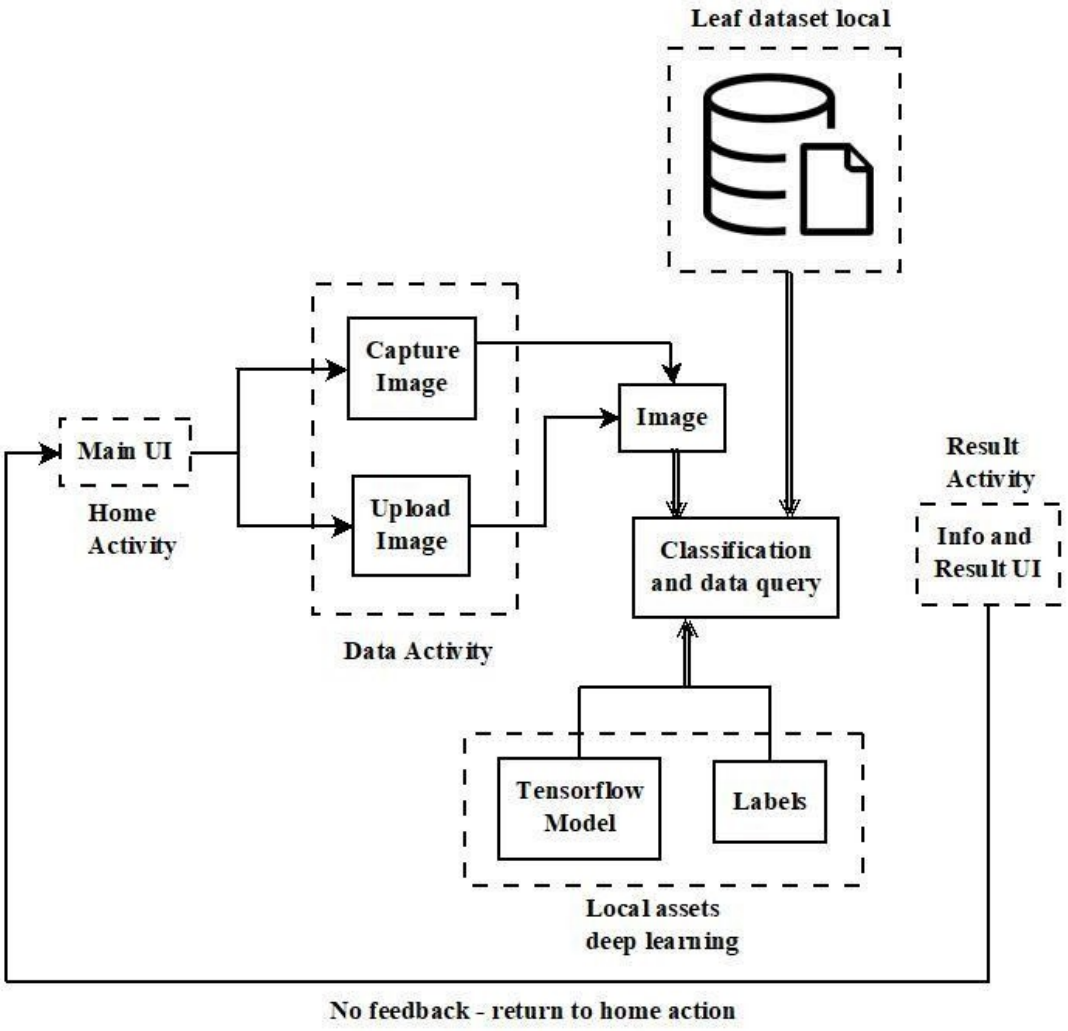


Figure 4.7: Component Diagram

4.4.5 Deployment Diagram

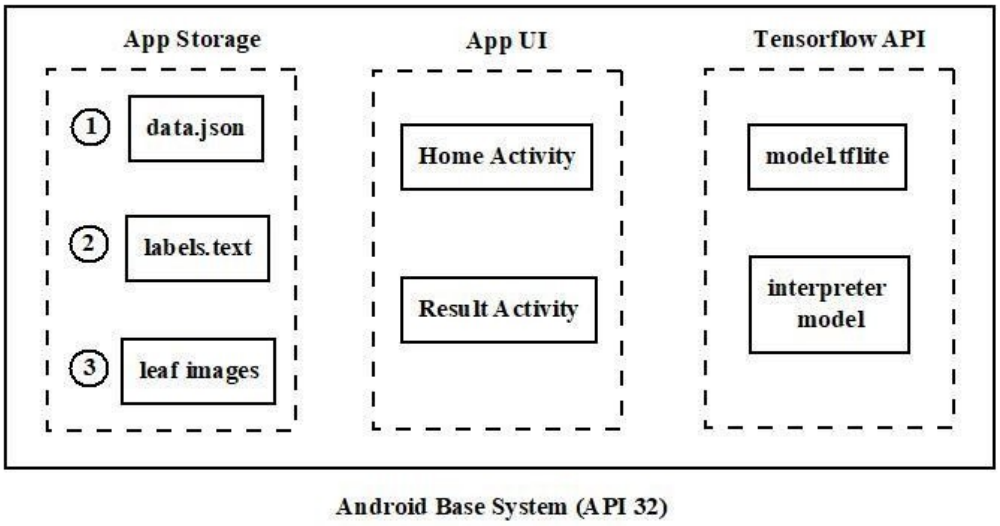


Figure 4.8: Deployment Diagram

Chapter 5

Implementation

5.1 Environmental Settings

- Windows Operating System: 8, 10 or above
- Processor: Intel core i5 or above.
- 64-bit, quad-core, 2.5 GHz minimum per core
- Ram: 4 GB or more.
- Hard disk: 10 GB of available space or more. 6. Python: 3.10.7
- Android: Android 11
- Java 17
- JSON
- **Python:** Python is an interpreted , high-level , general purpose programming language created by Guido Van Rossum and first released in 1991. It provides code readability and helps the programmers to write transparent, logical code for small and large scale projects.
- **PIP:** It is the package management system used to install and manage software packages written in python.
- **TensorFlow:** TensorFlow is an open-source library developed for deep learning applications such as neural networks.
- **Keras:** Keras is an open-source neural network library written in Python. It is capable of running on top of TensorFlow. It is designed for developing and evaluating deep learning models.
- **Android:** Android is an open source and Linux-based Operating System for mobile devices such as smartphones and tablet computers. Android is used to develop interactive mobile applications.
- **Java:** Java is a programming language and a platform. Java is a high level, robust, object-oriented and secure programming language. We used java as a programming language for android.
- **JSON:** JSON is a lightweight format for storing and transporting data. JSON data is written as name and value pairs. We stored additional information of 30 plants in JSON format.

5.2 Detailed Description of Methods

- **Upload Image:** Upload leaf Image of medicinal plant which has to be processed further for identification.
- **Data Collection:** There are several open-source datasets available for medicinal plants leaf Images. We have downloaded a dataset from Mendeley Data website which consists of nearly 1500 images. The dataset comprises of thirty species of healthy medicinal herbs such as Santalum album (Sandalwood), Muntingia calabura (Jamaica cherry), Plectranthus amboinicus / Coleus amboinicus (Indian Mint, Mexican mint), Brassica juncea (Oriental mustard), and many more. Each species consists of 60 to 100 high-quality images. The folders are named as per the species botanical/scientific name. Training Images - 75% Testing Images - 25%
- **Classification using CNN (Image classification) :** Classification is the best approach for identification of images like any kind of medical imaging. All classification algorithms are based on the prediction of an image, where one or more features and that each of these features belongs to one of several classes. An automatic and reliable classification method Convolutional Neural Network (CNN) will be used since it is robust in structure which helps in identifying every minute details. A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other. The preprocessing required in a ConvNet is much lower s compared to other classification algorithms. For this step we need to import Keras and other packages that we're going to use in building the CNN. Import the following packages:

    1. Sequential is used to initialize the neural network.
    2. Convolution2D is used to make the convolutional network that deals with the images.
    3. MaxPooling2D layer is used to add the pooling layers.
    4. Activation layer defines how the weighted sum of the input is transformed into an output from a node or nodes in a layer of the network.
    5. Dropout layer is used to prevent a model from overfitting.
    6. Flatten is the function that converts the pooled feature map to a single column that is passed to the fully connected layer.
    7. Dense adds the fully connected layer to the neural network.

5.3 Implementation details

- **User Details:** The user is asked to capture or upload an input image as a leaf image of a medicinal plant. This Android application is implemented using Java and JSON. JSON is used as an intermediate in android application to transfer data between android. Deep Learning is used to build model for correct indentation of medicinal plants using python language.

- **Backend Implementation:** CNN is used for image classification and its role is to reduce the images into a form whichis easier to process, without losing features which are critical for getting a good prediction. Layers used are as follows –

    1. **Sequential:** To initialize the neural network, we create an object of the Sequential class.

       model = Sequential ()

    2. **Convolutional:** To add the convolution layer, we call the add function with the classifier object and pass in Convolution2D with parameters. The first argument feature detectors which is the number of feature detectors that we want to create. The second and third parameters are dimensions of the feature detector matrix. The next parameter is input shape which is the shape of the input image. The images will be converted into this shape during preprocessing. If the image is black and white it will be converted into a 2D array and if the image is coloured it will be converted into a 3D array The final parameter is the activation function. Classifying images is a nonlinear problem. So, we use the rectifier function to ensure that we don't have negative pixel values during computation. That's how we achieve non-linearity.

       model.add(Conv2D(filters=32, kernel_size=(3,3), activation='relu',
       input_shape=(150,150,3)))

    3. **Pooling:** The Pooling layer is responsible for reducing the spatial size of the convolved feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.
       There are two types of Pooling: Max Pooling and Average Pooling. Max

Pooling returns the maximum value from the portion of the image covered by the Kernel. On the other hand, Average Pooling returns the average of all the values from the portion ofthe image covered by the Kernel. Generally, we use max pooling. In this step we reduce the size of the feature map. Generally, we create a pool size of 2x2 for max pooling. This enables us to reduce the size of the feature map while not losing important image information.

model.add (MaxPooling2D (pool_size= (2,2)))

4. **Dropout:** Dropout layer is used to prevent a model from overfitting.

   model.add (Dropout (0.4))

5. **Flatenning:** In this step, all the pooled feature maps are taken and put into a single vector for inputting it to the next layer. The Flatten function flattens all the feature maps into a single long column.

   model.add (Flatten ())

6. **Fully Connected:** The next step is to use the vector we obtained above as the input for the neural network by using the Dense function in Keras. The first parameter is output which is the number of nodes in the hidden layer. You can determine the most appropriate number through experimentation. The higher the number of dimensions the more computing resources you will need to fit the mode.

   model.add (Dense (output = 64))

   The next layer we have to add is the output layer. In this case, we'll use the softmax activation function since we expect a multiple classes outcome. If we expected more than two outcomes, we would use the SoftMax function. Num_of_classes here are 30.

   model.add(Dense(num_of_classes, activation='softmax'))

Chapter 6

Integration and Testing

6.1 Description of the Integration Modules

No.	Module	Input	State	Output	Result
1	Capture Image	Capture leaf Image through camera	Flag =1	Captured image successfully	Ok
2	Upload Image	Select leaf Image from gallery	Flag =1	Image selected successfully	Ok
3	Identification of Medicinal Plant	Captured/Uploaded image	Flag =1	Identification of accurate medicinal plant successfully	Ok
4	Display information of identified Medicinal plant	Identified medicinal plant	Flag =1	Local name, scientific name, origin, Main feature of medicinal plant displayed to the user	Ok
5	Showing medicinal uses of medicinal plant	Identified medicinal plant	Flag =1	Appropriate list of diseases Displayed to the user	Ok

6.2 Testing

No.	Module	Description	Expected Output	Actual Output	Result
1	Capture or Upload Image	User have to capture an image of a leaf through a camera or upload it from a gallery	Image Acquisition Successfully	Image Acquisition successfully	Pass
2	Submit Image	Captured Image is given as an input to a deep learning model	Image should submit	Appropriate captured image is taken as input for further model processing	Pass
3	Identification of Medicinal Plant	After image processing, the deep learning model has to identify the medicinal plant which is in the given image	Our System will identify accurate medicinal plant	Our System identified accurate medicinal plant	Pass
4	Display information of identified Medicinal plant	After identification of the medicinal plant our system will display its Local Name, Scientific Name etc.	System will display additional information about corresponding medicinal plant	Local name, scientific name, origin, Main feature of medicinal plant displayed to the user	Pass
5	Showing medicinal uses of medicinal plant	System will display the list of diseases cured by the identified medicinal plant to the user	Appropriate list of diseases will be displayed to the user	System displayed list of diseases cured by the identified medicinal plant to the user	Pass

Chapter 7

Performance Analysis

Performance Analysis

The performance analysis of the project is based on the classification and prediction model. If the model predicts the output with high accuracy, then the system will successfully classify the leaf images and will tell the name of the correct medicinal plant with its additional information. To train the prediction model we use neural network (Convolutional Neural Network). Using this Convolutional Neural Network we get accurate classification of leaf images. For the leaf image classification task, we have used images of 30 species of medicinal herbs and each species consists of 60 to 100 high quality images. This task is done through CNN algorithm. We have a dataset of 1500 leaf Images used to correctly identify the medicinal plant. Overall system gives accuracy between 70% to 80%. Predicted output will be displayed on the android application.

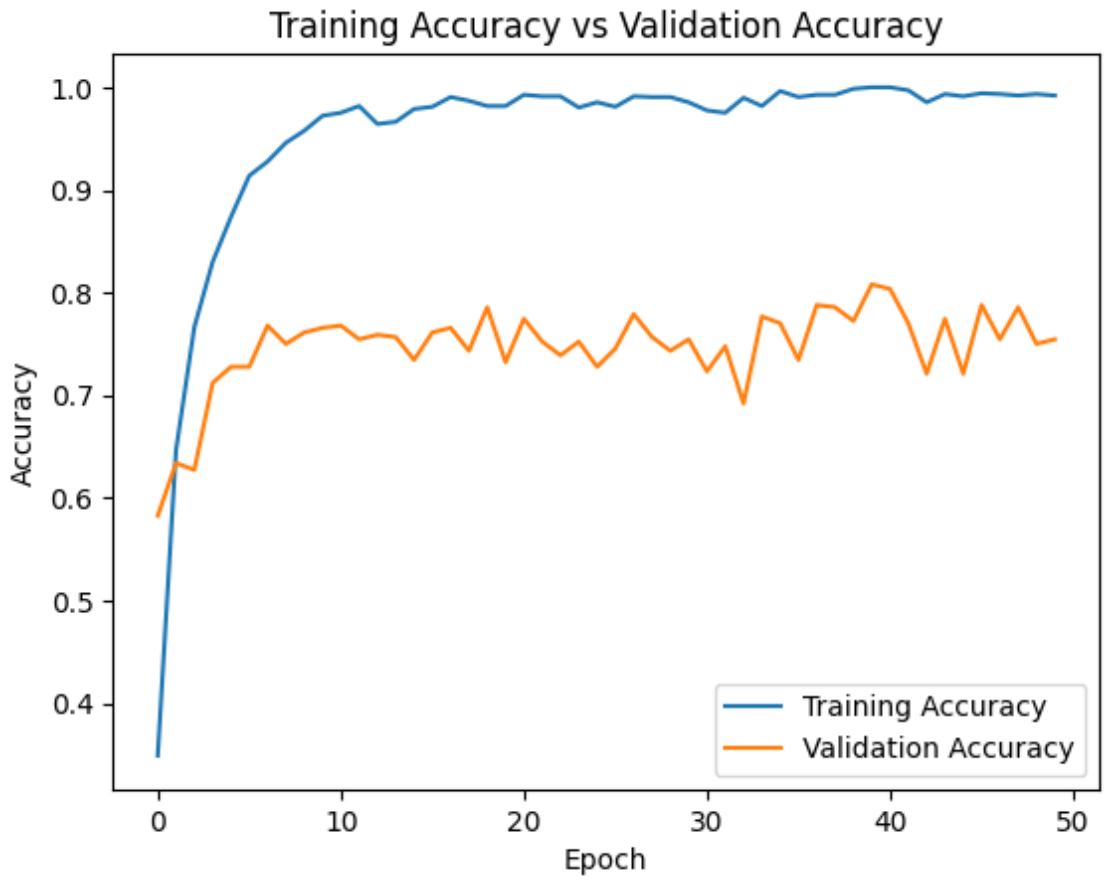


Figure 7.1: fig:Accuracy

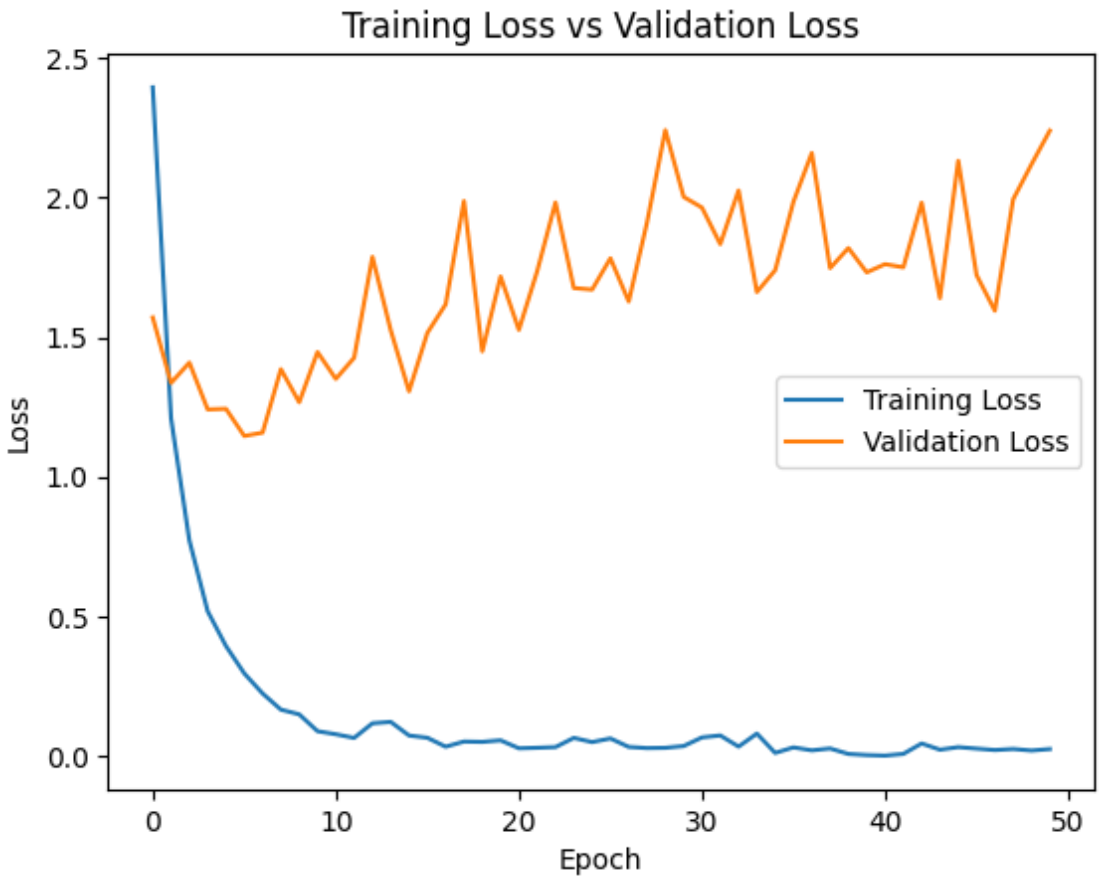


Figure 7.2: fig:Loss

According to the above figure, there is an optimal convergence between the training data and the validation data in the accuracy model. As well, we are correctly monitoring the increase in precision and loss between train and validation in epoch this reflects that with each epoch the model is learning more information. For the loss model, the accuracy and loss start to deviate steadily, this could be a sign of stopping training at an earlier time. The accuracy curve of training data with 90% to 98% and the accuracy curve of validation data with 70% to 80%.

Chapter 8

Future Scope

Future Scope

In future research in the area of medicinal plant identification, improved machine learning classifiers with some pre-processing and feature selection models will be used to solve the accuracy related issues and enhance the performance.

Our developed system is not worthy for tiny leaves or plants without a proper leaf. In future, efforts may be made to develop methods to identify these types of plants.

Chapter 9

Applications

Applications

Applications of our app are as below:

- To uncover new plant varieties and discover various plant species.
- To study plant taxonomy to improve knowledge of medicinal plants.
- To identify invasive species i.e. species that are not indigenous or native to a particular area.
- To identify species causing food poisoning to improve the health of human beings and create awareness among them.

Chapter 10

Installation Guide and User Manual

Installation Guide

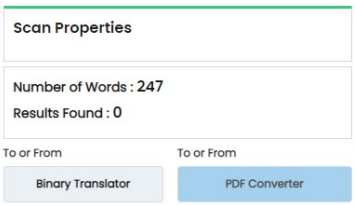
To manually set up Medileaf Application, please follow the steps below:

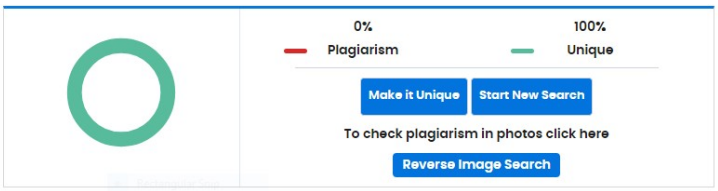
1. Download Medileaf Application on your phone.
2. User have to capture an image through a camera or upload an image from a gallery.
3. Results will be displayed as the name of the correctly identified medical plant.
4. View basic information about medicinal plant.
5. View list of diseases cured by identified plant.
6. Close the application.

Chapter 11

Plagiarism Report

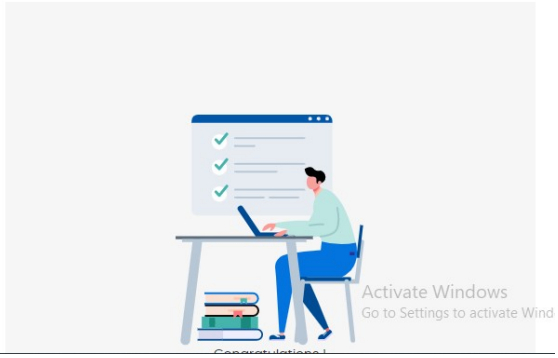
Plagiarism Report

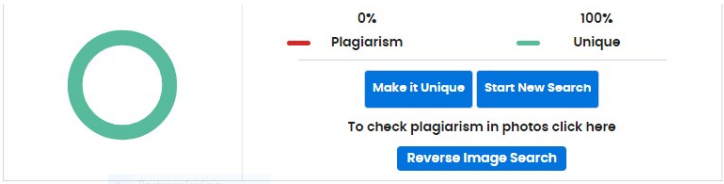




plants process is a very time consuming and requires the help of experts for plant identification. Incorrect identification of medicinal plants may lead to unfavorable results. To defeat this problem, automatic identification and classification of medicinal plants species is necessary for greater satisfaction to mankind. Plant identification and classification can be automated using visual morphological features such as the texture, color and shape of the leaves.

Here we develop a system which can identify the plant species based on the input leaf sample. MediLeaf presents how medicinal plants were identified with high accuracy by applying deep learning capabilities. For this, a dataset is used of scanned images of leaves of commonly used Ayurveda medicinal plants. Both the front and back sides of leaves were captured. The leaves are classified based on the unique characteristics combination. The plant species are classified by the features on each





a. Problem Definition:

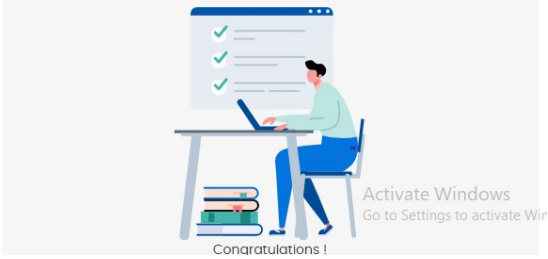
Problem Statement –

To design an automatic system to recognize and identify the medicinal plant leaf using CNN.

b. Aim and Objective of the Project:

Aim –
To design an automatic system to recognize the medicinal plant leaf using CNN.

Objectives –



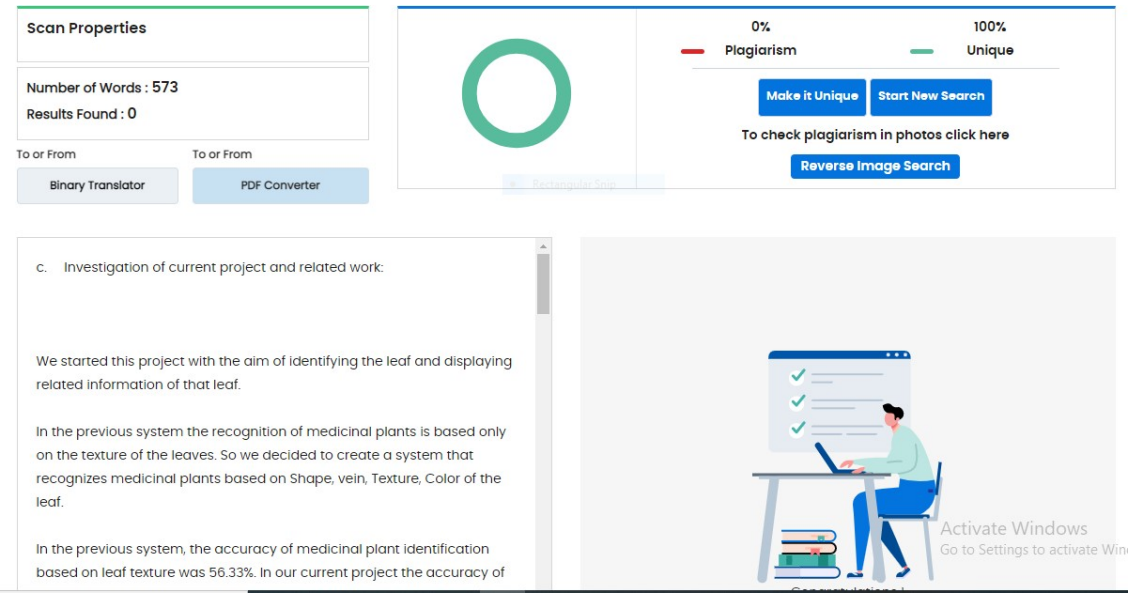


Figure 11.1: Plagiarism Report

Chapter 12

Ethics

Ethics

- **Creative Commons License:** Attribution-Share Alike 4.0 International (CC BY SA 4.0 This is human-readable summary of (and not a substitute for) the license.You are free to:

  1. Share: copy and redistribute the material any medium or format.
  2. Adapt: remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms.

- **Using the following terms:** Attribution you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any responsible manner, but not in any way that suggests the licensor endorses you or your use.

  1. Share Alike: you remix, transform, and build upon the material for any purpose, you must distribute your contributions under the same license as the original.
  2. No additional restrictions: You must apply legal terms or technological measures that legally restrict others from doing anything the license permits.
  3. Notices: You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an application exception or limitation. No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit hoe you use the material.

- **Ethical Practices for CSE Students:** As Computer Sc. & Engineering student, I believe it is unethical to:

  - Take credit for someone else's work
  - Hire someone to write an assignments
  - Purchase or submit a research or term paper from the internet to a class as one's own work
  - Cheat on a graded assignment
  - Cheat on an exam
  - Plagiarize other people's work without citing or referencing the work

- Add the name of a non-contributing person as an author in a project/research study
- Copy and paste material found on the Internet for an assignment without acknowledging the authors of the material
- Deliberately provide inaccurate references for a project or research study
- Knowingly permit student work done by one student to be submitted by another student
- Surf the internet for personal interest and non-class related purposes during classes
- Make a copy of software for personal or commercial use
- Make a copy of software for a friend
- Loan CDs of software to friends
- Download pirated software from the internet
- Distribute pirated software from the internet
- Buy software with a single user license and then install it on multiple computers
- Share a pirated copy of software
- Install a pirated copy of software

Chapter 13

References

References

1. Fadzilah Siraj. (2014). Flower image classification modeling using neural network. Available at: https://ieeexplore.ieee.org/document/7042605 [Accessed 8 Mar. 2019].

2. Huthaifa Almogdad. (2018).A Flower Recognition System Based On Image Processing And Neural Networks. Available at: http://www.ijstr.org/finalprint/nov2018/A-FlowerRecognition-System-BasedOn-Image-Processing-And-Neural-Networks.pdf [Accessed 8 Mar. 2019]

3. Thi Thanh Nhan Nguyen. (2016).Flower species identification using deep convolutional neural network Available at: https://www.researchgate.net/publication/308322586_Flower_species_identification_using_deep_convolutional_neural_networks?enrichId=rgreq69d385e2d892ea951dc14797d47a5fe4-XXX&enrichSource=Y292ZXJQYWdlOzMwODMyMjU4NjtBUzo0MDgyMzgyMTU0NTA2MjVAM//TQ3NDM0Mjk4Mzk//3Ng%3D%3D&el=1_x_3&_esc=publicationCoverPdf [Accessed 8 Mar. 2019].

4. P. Bhandarkar, H. Doshi, and M. Kadakia, "Leaf identification using morphology and structural decomposition - IEEE Conference Publication", Ieeexplore.ieee.org, 2014. [Online].Available at:https://ieeexplore.ieee.org/ielx7/6766412/6776//904/06776934.pdf?tp=&arnumber=6776934&isnumber=6776904 [Accessed: 07- Mar- 2019].

5. N. Deepak and R. Ahmed, "Neural networks based leaf identification using shape and structural decomposition - IEEE Conference Publication", Ieeexplore.ieee.org, 2016. [Online]. Available at: https://ieeexplore.ieee.org/ielx7/7948574/7955253/07955302.pdf?tp=&arnumber=7955302&isnumber=7955253&tag=1 [Accessed: 07- Mar- 2019].

6. V. T, "Leaf identification for the extraction of medicinal qualities using image processing algorithm - IEEE Conference Publication", Ieeexplore.ieee.org, 2017. [Online]. Available at: https://ieeexplore.ieee.org/ielx7/8316803/8321763/08321884.pdf?tp=&arnumber=8321884&isnumber=8321763 [Accessed: 07- Mar- 2019].

7. Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples. Available at:https://www.researchgate.net/publication/322059676_Iden//tification_of_ayurvedic_medicinal_plants_by_image_processing_of_leaf_samples[Accessed 8 Mar. 2019].

8. Introduction to Deep Learning Deep learning basics with Python, Tensorflow and Keras Available at: https://pythonprogramming.net/introduction-deep/ /learning-python-tensorflowkeras/[Accessed 9 Mar. 2019].

9. A Complete Guide to Types of Neural Network. Available at: https://www.digitalvidya.com/blog/types-of-neuralnet//works/ [Accessed 9 Mar. 2019].

10. Plant Identification Using Tensorflow. Available at: https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1271&context=cpesp[Accessed 9 Mar. 2019].

11. Convolution Neural Networks (CNN) : Step 1 Convolutional Operation. Available at: https://www.superdatascience.com/blogs/convolutional-neural-net/ /works-cnn-step-1-convolutionoperation [Accessed 9 Mar. 2019].

12. D. A. D. S. Jayalath, T. G. A. G. D. Amarawanshaline, D. P. Nawinna, P. V. D. Nadeeshan and H. P. Jayasuriya, "Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers," 2019 14th Conference on Industrial and Information Systems (ICIIS), 2019, pp. 125-129, doi: 10.1109/ICIIS4734 // 6.2019.9063275

13. P. M. Kumar, C. M. Surya and V. P. Gopi, "Identification of ayurvedic medicinal plants by image processing of leaf samples," 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2017, pp. 231-238, doi: 10.1109/ICRCICN.2017.8234512.

14. D. A. D. S. Jayalath, T. G. A. G. D. Amarawanshaline, D. P. Nawinna, P. V. D. Nadeeshan and H. P. Jayasuriya, "Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers," 2019 14th Conference on Industrial and Information Systems (ICIIS), 2019, pp. 125-129, doi: 10.1109/ICIIS4734 6.2019.9063275.

15. B. Pukhrambam and R. Rathna, "A Smart Study on Medicinal Plants Identification and Classification using Image Processing Techniques," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 956-962, doi: 10.1109/ICICV50876.2021.9388566.

16. D Venkataraman and Mangayarkarasi N, "Computer Vision Based Feature Extraction of Leaves for Identification of Medicinal Values of Plants", IEEE International Conference on Computational Intelligence and Computing Research, 978-1-5090-0612-0/16/$31.00 ©2016 IEEE.

17. http://www.iraj.in/journal/journal_file/journal_pdf/6-480-153448/
/609513-18.pdf

Chapter 14

Bibliography

Bibliography

1. **Research Papers:**

   (a) Sandya, Desta & Herdiyeni, Yeni. (2013). MedLeaf: Mobile Application For Medicinal Plant Identification Based on Leaf Image. International Journal on Advanced Science, Engineering and Information Technology. 3. 10.18517/ijaseit.3.2.287.

   (b) D. A. D. S. Jayalath, T. G. A. G. D. Amarawanshaline, D. P. Nawinna, P. V. D. Nadeeshan and H. P. Jayasuriya, "Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers," 2019 14th Conference on Industrial and Information Systems (ICIIS), 2019, pp. 125-129, doi: 10.1109/ICIIS47346.2019.9063275.

   (c) Sivaranjani, C. & Kalinathan, Lekshmi & Amutha, R. & Kathavarayan, Ruba Soundar & K J, Jegadish. (2019). Real-Time Identification of Medicinal Plants using Machine Learning Techniques. 1-4. 10.1109/ICCIDS.2019.8862126.

2. **Diagrams:**
   For drawing the diagrams, we have used - Dia Diagram Editor Software.

3. **Plagiarism Check:**
   For checking plagiarism, we have used - `https://smallseotools.com/plagiarism-checker/`