Leaf Disease Detection

B_TECH-IT Semester - VIII

Prepared at

ISO 9001:2008
ISO 27001:2013
CMMI LEVEL-5

Bhaskaracharya National Institute for Space Applications & Geo-informatics
Ministry of Electronics and Information Technology, Govt. of India.

Gandhinagar

Prepared By

Shrey Parekh.	Shivam Rank.	Nir Patel.	Urvish Soni
ID No. 17IT059	ID No. 17IT094	ID No. 17IT073	ID No. 17IT112

Guided By:

Prof. Priyanka Patel

Department of IT

CSPIT, Changa, Gujarat.

External Guide:

Vishal Patel

Project Scientist

BISAG- N, Gandhinagar.

SUBMITTED TO

CHARUSAT
CHAROTAR UNIVERSITY OF SCIENCE AND TECHNOLOGY

SMT. KUNDANBEN DINSHA PATEL DEPARTMENT OF INFORMATION TECHNOLOGY

Chandubhai S. Patel Institute of Technology

At: Changa, Dist: Anand – 388421

April 2021

Bhaskaracharya National Institute for Space Applications and Geo-informatics

MeitY, Government of India

Phone: 079 - 23213081 Fax: 079 - 23213091

E-mail: info@bisag.gujarat.gov.in, website: https://bisag-n.in/

ISO 9001:2008
ISO 27001:2013
CMMI LEVEL-5

CERTIFICATE

This is to certify that the project report compiled by ***Mr. Shrey A Parekh, Mr. Shivam D Rank, Mr. Nir P Patel and Mr. Urvish S Soni*** *students of 8th Semester* ***B_Tech-IT from Chandubhai S. Patel Institute of Technology, CHARUSAT, Changa*** *have completed their final Semester internship project satisfactorily. To the best of our knowledge this is an original and bonafide work done by them. They have worked on Web-based application for "****Leaf Disease Detection****", starting from January 18th, 2021 to April 30th, 2021.*

During their tenure at this Institute, they were found to be sincere and meticulous in their work. We appreciate their enthusiasm & dedication towards the work assigned to them.

We wish them every success.

Vishal Patel

Project Scientist,

BISAG- N, Gandhinagar

T. P. Singh

Director General,

BISAG- N, Gandhinagar

CHAROTAR UNIVERSITY OF SCIENCE AND TECHNOLOGY

CERTIFICATE

This is to certify that the report entitled "**Leaf Disease Detection**" is a bona fide work carried out by **Shrey A Parekh(17IT059), Shivam D Rank(17IT094), Nir P Patel(17IT073) and Urvish S Soni(17IT112)** under the guidance and supervision of **Prof. Priyanka Patel and Vishal Patel** for the subject **Software Project Major (IT447)** of 8th Semester of Bachelor of Technology in **Information Technology** at Faculty of Technology & Engineering – CHARUSAT, Gujarat.

To the best of my knowledge and belief, this work embodies the work of candidate himself, has duly been completed, and fulfills the requirement of the ordinance relating to the B.Tech. Degree of the University and is up to the standard in respect of content, presentation and language for being referred to the examiner.

Under supervision of,

Prof. Priyanka Patel
Smt. Kundanben Dinsha Patel Department of Information Technology
CSPIT, Changa, Gujarat.

Vishal Patel
Project Scientist
BISAG.

Dr. Parth Shah
Head & Associate Professor
Smt. Kundanben Dinsha Patel Department of Information
CSPIT, Changa, Gujarat.

About BISAG- N

ABOUT THE INSTITUTE

Modern day planning for inclusive development and growth calls for transparent, efficient, effective, responsive and low cost decision making systems involving multi-disciplinary information such that it not only encourages people's participation, ensuring equitable development but also takes into account the sustainability of natural resources. The applications of space technology and Geo–informatics have contributed significantly towards the socio-economic development. Taking cognizance of the need of geo-spatial information for developmental planning and management of resources, the department of Science and Technology, Government of Gujarat established "Bhaskaracharya National Institute for Space Applications and Geo-informatics - N" (BISAG- N). BISAG- N is an ISO 9001:2008, ISO 27001:2005 and CMMI: 5 certified institute. BISAG- N which was initially set up to carryout space technology applications, has evolved into a centre of excellence, where research and innovations are combined with the requirements of users and thus acts as a value added service provider, a technology developer and as a facilitator for providing direct benefits of space technologies to the grass root level functions/functionaries.

BISAG- N's Enduring Growth

Since its foundation, the Institute has experienced extensive growth in the sphere of Space technology and Geo-informatics. The objective with which BISAG- N was established is manifested in the extent of services it renders to almost all departments of the State. Year after year the institute has been endeavouring to increase its outreach to disseminate the use of geo-informatics up to grassroots level. In this span of nine years, BISAG- N has assumed multi-dimensional roles and achieved several milestones to become an integral part of the development process of the Gujarat State.

BISAG- N Journey

2003-04

Gujarat SATCOM Network

2007-08

Centre for Geo-informatics Applications

2010-11

Academy of Geo-informatics for Sustainable Development

2012-13

A full-fledged Campus

Activities

Satellite Communication..

for promotion and facilitation of the use of broadcast and teleconferencing networks for distant interactive training, education and extension.

Remote Sensing..

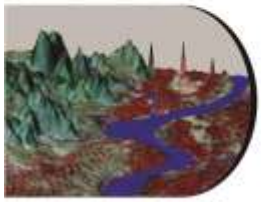
for Inventory, Mapping, Developmental planning and Monitoring of natural & man-made resources.

Geographic Information System..

for conceptualization, creation and organization of multi purpose common digital database for sectoral/integrated decision support systems.

Global Navigation Satellite System..

for Location based Services, Geo-referencing, Engineering Applications and Research.



Photogrammetry..

for Creation of Digital Elevation Model, Terrain Characteristic, Resource planning.

Cartography..

for thematic mapping, value added maps.

Software Development..

for wider usage of Geo-spatial applications, Decision Support Systems (desktop as well as web based), ERP solutions.

Education, Research and Training..

for providing Education, Research, Training & Technology Transfer to large number of students, end users & collaborators.

Applications of Geospatial Technology for Good Governance: Institutionalization

Through the geospatial technology, the actual situation on the ground can be accessed. The real life data collected through the technology forms the strong foundation for development of effective social welfare programs benefiting directly the grass root level people. The geospatial data collected by the space borne sensors along with powerful software support through Geographic Information System (GIS), the vital spatio-temporal maps, tables, and various statistics are being generated which feed into Decision Support System (DSS).

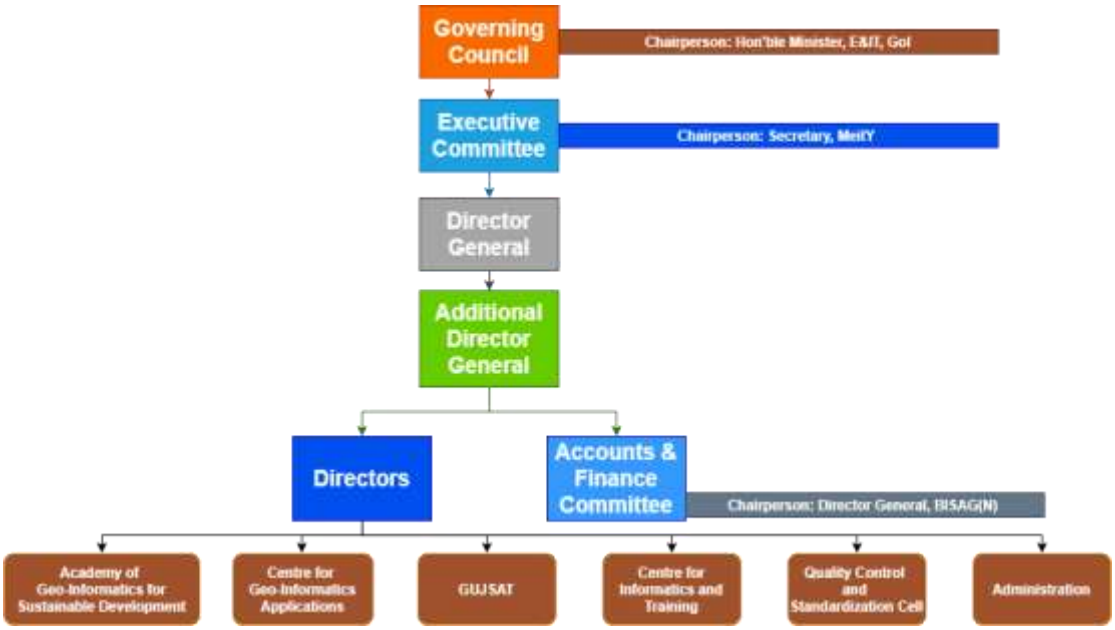
A multi-threaded approach is followed in the process of institutionalization of development of such applications. The 5 common threads which run through all the processes are: *Acceptability, Adaptability, Affordability, Availability and Assimilability*.
These are the "Watch Words" which any application developer has to meet. The "acceptability" addresses the issue that the application developed has met the wide acceptability among the users departments and the ultimate end beneficiary by way of providing all necessary data and statistics required. The "affordability" addresses the issue of the application product being cost effective. The "availability" aspect looks into aspect of easily accessible across any platform, anywhere and anytime. The applications should have inbuilt capability of easy adaptability to the changing spatio- and temporal resolutions of data, new aspects of requirements arising from time to time from users. The assimilability aspect ensures that the data from various sources / resolutions and technologies can be seamlessly integrated.

ACCEPTABILITY	▪ Problem definition by users • Proof of Concept development without financial liability on users ▪ Execution through collaboration under user's ownership
ADOPTABILITY	▪ Applications as per present systems & database ▪ Maximum Automation ▪ Minimum capacity building requirement at the user end
AFFORDABILITY :	▪ Multipurpose geo-spatial database, common, compatible, standardized (100s of layers) ▪ In house developed/open source software ▪ Full Utilization of available assets
AVAILABILITY:	▪ Departmental /Integrated DSS ▪ Desired Product delivery anytime, anywhere in the country
ASSIMILABILITY	▪ Integration of Various technologies like RS, GIS, GPS, Web MIS, Mobile etc.

Organizational Setup

The Institute is responsible for providing information and technical support to different Departments and Organizations. The Governing Body and the Empowered Executive Committee govern the functioning of BISAG- N. The Institute is registered under the Societies Registration Act 1860. Considering the scope and extent of activities of BISAG- N, its organizational structure has been charted out with defined functions.

Organizational Setup of BISAG- N



Governing Body

For smoother, easier and faster institutionalization of Remote Sensing and GIS technology, decision makers of the state were brought together to form the Governing Body. It is the supreme executive authority of the Institute. The Governing Body comprises of ex-officio members from various Government departments and Institutes.

- Hon'ble Minister of Electronics and Information Technology ..... Chairperson (Ex-Officio)
- Hon'ble Minister of State Electronics and Information Technology.....……Deputy Chairperson (Ex-Officio)
- Secretary of Government of India: Ministry of Electronics and Information Technology..... Executive Vice Chairperson (Ex-Officio)
- Chief Executive Officer, Niti Aayog .....……Member (Ex-Officio)
- Chairman, Indian Space Research Organization.....……Member (Ex-Officio)
- Secretary to Government of India: Department of Science and Technology .....……Member (Ex-Officio)
- Additional Secretary to Government of India: Ministry of Electronics and Technology.. ……Member (Ex-Officio)
- Chief Secretary to Government of Gujarat.....……Member (Ex-Officio)
- President & Chief Executive Officer, National e-Governance Division, Ministry of Electronics and Information Technology.....…… Member (Ex-Officio)
- Financial Advisor to Government of India: Ministry of Electronics and Information Technology…Member (Ex-Officio)
- Distinguished Professionals from the GIS field-Three (3) (To be nominated by the Chairperson)
- Director-General, Bhaskaracharya National Institute for Space Application and Geo-Informatics {BISAG(N)} ..... Member Secretary (Ex-Officio)

Centre for Geo-informatics Applications

Introduction

The objective of this technology group is to provide decision support to the sectoral stake holders through scientifically organized, comprehensive, multi-purpose, compatible and large scale (village level) geo-spatial databases and supporting analytical tools. These activities of this unit are executed by a well-trained team of multi-disciplinary scientists. The government has provided a modern infrastructure along with the state-of-the-art hardware and software. To study the land transformation and development over the years, a satellite digital data library of multiple sensors of last twenty years has been established and conventional data sets of departments have been co-registered with satellite data. The geo-spatial databases have been created using conventional maps, high resolution satellite 2D and 3D imagery and official datasets (attributes).The geo-spatial databases include terrain characteristics, natural and administrative systems, agriculture, water resources, city survey maps, village maps with survey numbers, water harvesting structures, water supply, irrigation, power, communications, ports, land utilization pattern, infrastructure, urbanization, environment data, forests, sanctuaries, mining areas, industries. They also include social infrastructure like the locations of schools, health centres, institutions, aganwadies, local government infrastructure etc. The geospatial database of nagar-palikas includes properties and amenities captured on city and town planning maps with 1000 GIS layers. Similar work for villages has been initiated as a pilot project.

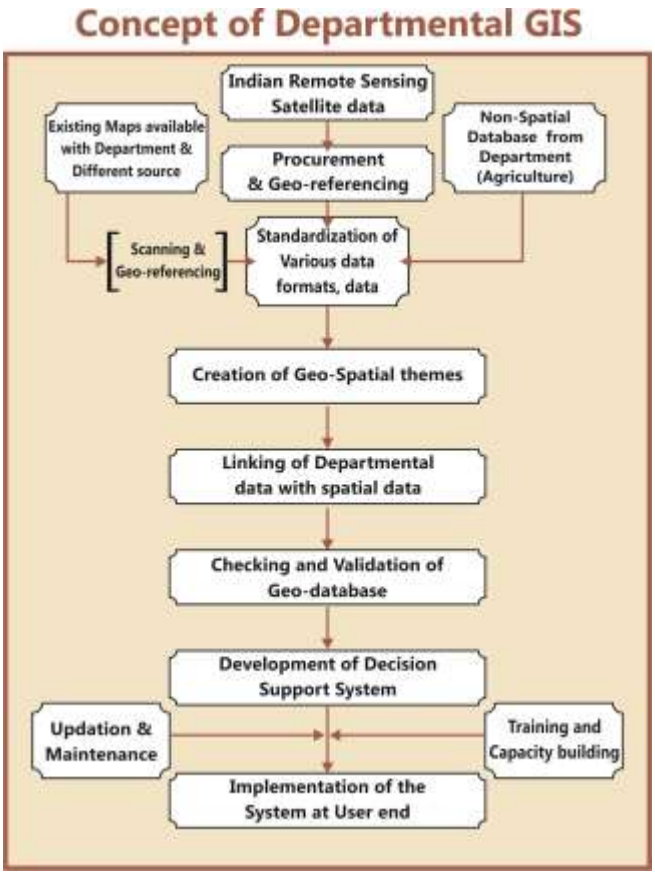
The applications of space technology and geo-informatics have been operational in almost all the development sectors of the state. Remote sensing and GIS applications have provided impetus to planning and developmental activities at grass root level as well as monitoring and management in various disciplines.

The GIS based Applications Development

The GIS software is a powerful tool to handle, manipulate and integrate both the spatial and non-spatial data. The GIS system operates on the powerful backend data base and Sequential Query Language (SQL) to inquiry the data bases. It has the capability to handle large volume of data and process to yield values of parameters which can be input to very important government activity as Decision Support System (DSS). Its mapping capabilities help the users and specialists in generating single and multi-theme wise maps.

The GIS based applications development has been institutionalized in BISAG- N. This process can be listed as (Refer Figure for Details)

- Making the users aware of the GIS capabilities through introductory training programme and by exposing to already developed projects as success stories.
- Helping the users in defining the GIS based projects.
- Digitizing the data available with the users and encouraging them to collect any additional data as may be required.
- Generating the appropriate data bases with the full involvement of the users following the data bases standards

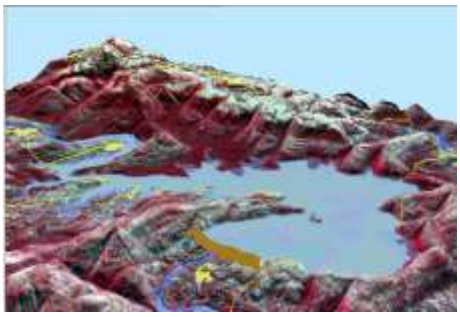


Remote Sensing and GIS Sectoral Applications:

Geo-informatics based Irrigation Management and Monitoring System

- The Geo-spatial information system for Irrigation water Management and Monitoring system for command areas in Sardar Sarovar Narmada Nigam Limited (SSNL) has been developed. Satellite image-based Irrigation monitoring system has been developed in GIS. From the multi-spectral Satellite images of every month, the irrigated areas were extracted.
- The irrigated area were overlaid on the geo-referenced cadastral maps and the statistics of area irrigated has been estimated.
- The user friendly Customized Decision Support System (DSS) has been developed.

Preparation of DPR of Par–Tapi-Narmada Link using Geo-informatics for National Water development Agency (NWDA)



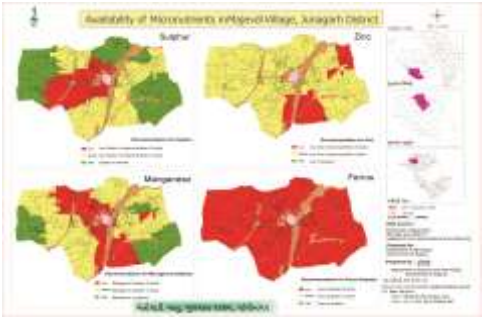
- The main objective of Par–Tapi-Narmada Link project is to divert surplus water available in west flowing rivers of south Gujarat and Maharashtra for utilization in the drought prone Saurashtra and Kachcha. On the request from NDWA, preparation of various maps for proposed DPR work was undertaken by the BISAG- N. Land use and submergence maps of proposed dams along with its statistics have been prepared by the BISAG- N. The detailed work consisted of generation of Digital Elevation Model (DEM), contour generation, Land use mapping, forest area generation of submergence extent at different levels etc.

Agriculture

District and Village-level Crop Inventory

- Remote Sensing (RS) based Village-level Crop Acreage Estimation at was taken up in two villages of Anand and Mehsana districts of Gujarat state. The major objective of this study was to attempt village-level crop inventory during two crop seasons of Kharif (monsoon season) and Rabi (winter season) using single-date Indian Remote Sensing (IRS) LISS-III and LISS-IV digital data of maximum vegetative growth stage of major crops during each season.
- District-level crop acreage estimation during three cropping seasons namely Kharif, Rabi and Zaid (summer) seasons was also carried out in all the 26-districts of Gujarat State. Summer crop acreage estimation Gujarat State was carried out during 2012.

Spatial Variability Mapping of Soil Micro-Nutrients



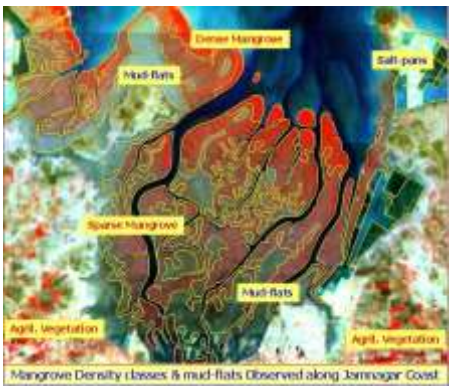
- The spatial variability of soil micro-nutrients like Fe, Mn, Zn and Cu in various villages of different districts, Gujarat state was mapped using geo-informatics technology. The major objectives of this study were i) to quantify the variability of Mn, Fe, Cu and Zn concentration in soil; ii) to map the pattern of micro-nutrient variability in cadastral maps, iii) suggest proper application of micro-nutrients based on status of deficiency for proper crop management and iv) preparation of village-level atlases showing spatial variability of micro-nutrients.

Geo-spatial Information System for Coastal Districts of Gujarat

- The project on development of Village-level Geo-spatial Information System for Shrimp Farms in Coastal Districts of Gujarat, was taken with major objective of development of Village-level Geo-spatial Information System for Shrimp/Scampi areas using Remote Sensing (RS) and GIS. This project was sponsored by the Marine Products Export Development Authority (MPEDA), Ministry of Commerce & Industry, Government of India for scientific management of Scampi farms in the coastal districts which can help fishermen to better their livelihood and increase the economic condition on sustainable basis. The customized query shell was developed using the open source software for sharing the information amongst the officers from MPEDA and potential users. This has helped the farmers to plan their processing and marketing operations so as to achieve better remunerations.

Environment and Forest

Mapping and Monitoring of Mangroves in the Coastal Districts of Gujarat State



- Gujarat Ecology Commission, with technical inputs from the Bhaskaracharya National Institute for Space Applications and Geo-informatics - N (BISAG- N) made an attempt to publish Mangrove Atlas of the Gujarat state. Mangrove atlas for 13-coastal districts with 35-coastal talukas in Gujarat, have been prepared using Indian Remote sensing satellite images. The comparison of mangrove area estimates carried out by BISAG- N and Forest Survey of India (FSI) indicates a net increase in the area under mangrove cover. The present assessment by BISAG- N, has recorded 996.3 sq. km under mangrove cover, showing a steep rise to the tune of 88.03 sq. km. In addition to the existing Mangrove cover, the present assessment also gives the availability of potential area of 1153 sq. km, where mangrove regeneration program can be taken up.

Academy of Geo-informatics for Sustainable Development

Introduction

- Considering the requirement of high end research and development in the areas having relevance of geo-informatics technology for sustainable development, a separate infrastructure has been established. In collaboration with different institutes in the state as well as in the country, R&D activities are being carried out in the areas of climate change, environment, disaster management, natural resources management, infrastructure development, resources planning, coastal hazard and coastal zone management studies, etc. under the guidance of eminent scientists.

- Various innovative methodologies/models developed in this academy through the research process have helped in development of various applications. There are plans to enhance R&D activities manifold during coming years.
- This unit also provides training to more than 600 students every year in the field of Geo-informatics to the students from various backgrounds like water resources, urban planning, computer Engineering, IT, Agriculture in the areas of Remote sensing, GIS and their applications.

- This Academy has been established as a separate infrastructure for advanced research and development through following schools:
  - School of Geo-informatics
  - School of Climate & Environment
  - School of Integrated Coastal Zone Management

- School of Sustainable Development Studies
- School of Natural Resources and Bio-diversity
- School of Information Management of Disasters
- School of Communication and Society

During XIIth Five year Plan advance applied research through above schools shall be the main thrust area. Already M. Tech and Ph.D. students of other Universities/ Institutes are doing research in this academy in applied sciences under various collaborative programmes.

M. Tech. Students' Research Programme

The academy started M. Tech. students' research programme in a systematic way. It admitted 11 students from various colleges and universities in Gujarat, Rajasthan and Madhya Pradesh for period of 10 months from August 2011 to May 2012. All the students were paid stipend of Rs. 6000 per month during the tenure. The research covered the following areas:

- Cloud computing techniques
- Mobile communication
- Design of embedded systems
- Aquifer modelling
- Agricultural and Soils Remote Sensing
- Digital Image processing Techniques (Data Fusion and Image Classification).

The research resulted in various dissertations and publications in national and international journals.

• Now nine students, one from IIT, Kharagpur, three from GTU, one from M. S University, Vadodara and four from GU, are undergoing their Ph. D programme. Out of nine, two thesis have been submitted. Two students are from abroad. One each from Vietnam and Yemen. Since then (after approval of research programme from the Governing Body), 200+ papers have been published by the Academy.

CANDIDATE'S DECLARATION

We declare that 8th semester internship project report entitled "**Leaf Disease Detection**" is our own work conducted under the supervision of the external guide **Vishal Patel** from BISAG-N (Bhaskaracharya National Institute for Space Applications & Geo-informatics). We further declare that to the best of our knowledge the report for this project does not contain any part of the work which has been submitted previously for such project either in this or any other institutions without proper citation.

Shrey Parekh

Student ID: 17IT059

Shivam Rank

Student ID: 17IT094

Nir Patel

Student ID: 17IT073

Urvish Soni

Student ID: 17IT112

Submitted To:

CHARUSAT UNIVERSITY,
SMT. KUNDANBEN DINSHA PATEL DEPARTMENT OF INFORMATION TECHNOOGY,
Chandubhai S. Patel Institute of Technology (CSPIT), Changa, Gujarat.

ACKNOWLEDGMENT

We are grateful to **T.P. Singh,** Director (BISAG-N) for giving us this opportunity to work the guidance of renowned people of the field of MIS Based Portal also providing us with the required resources in the company.

We would like to express our endless thanks to our external guide **Mr. Vishal Patel**, And Admin Staff **Mr. Sidhdharth Patel** at Bhaskaracharya National Institute of Space Application and Geo-informatics for their sincere and dedicated guidance throughout the project development.

Also, our hearty gratitude to our Head of Department, **Dr. Parth Shah** and our internal guide **Prof. Priyanka Patel** for giving us encouragement and technical support on the project.

Shrey Parekh
Student ID: 17IT059

Shivam Rank
Student ID: 17IT094

Nir Patel
Student ID: 17IT073

Urvish Soni
Student ID: 17IT112

ABSTRACT

India is a country where agriculture is of great importance and is one of the main sources of income for a large number of people. Agriculture has a great contribution in the overall GDP and economy of the country. Crop failure is a problem that is affecting the overall production, it is a problem to be addressed. We have developed a Web-based project "Leaf Disease Detection" a Deep Learning Model deployed on Flask Server. The Model Classifies 14 different plants and 32 diseases that any person can check using web-based application, can upload a real-time image and get information whether the leaf of the plant is diseased or not. This application can help the farmers detect the disease of the leaf and can take precautionary measures before it spreads through the entire plant. This application is built on 4 main model

1. Base Image classification Model
2. CNN Architecture
3. VGG16 Model
4. RESNET34 Model

The result of CNN architecture is fed to one of the two options available to us VGG16 or RESNET34, this are the transfer learning techniques that are used to get the desired output. As the number of classes are more we have used this transfer learning techniques.

CONTENTS

List of Figures

List of Tables

1. INTRODUCTION

1.1 PROJECT OVERVIEW

“Leaf Disease Detection” project is plant disease recognition whose base model is leaf image classification. We have used “Deep Convolutional Network” for the development of the project. The data base used for the project consist of 14 different kind of plants and it can detect 38 different types of plant diseases.

Plants:

- Tomato
- Apple
- Blueberry
- Grape
- Peach
- Corn
- Soyabean
- Potato
- Orange
- Strawberry
- Pepper
- Raspberry
- Squash
- Cherry

Diseases:

- 'Apple___Apple_scab', 'Apple___Black_rot', 'Apple___Cedar_apple_rust', 'Apple___healthy'
- 'Blueberry___healthy'
- 'Cherry_(including_sour)__Powdery_mildew', 'Cherry(including_sour)___healthy'
- 'Corn_(maize)__Cercospora_leaf_spot_Gray_leaf_spot', 'Corn(maize)_Common_rust', 'Corn(maize)__Northern_Leaf_Blight','Corn(maize)___healthy',
- 'Grape___Black_rot', 'Grape___Esca_(Black_Measles)', 'Grape___Leaf_blight_(Isariopsis_Leaf_Spot)', 'Grape___healthy'
- 'Orange___Haunglongbing_(Citrus_greening)'
- 'Peach___Bacterial_spot', 'Peach___healthy'
- 'Pepper,_bell___Bacterial_spot', 'Pepper,_bell___healthy',
- 'Potato___Early_blight', 'Potato___Late_blight', 'Potato___healthy'
- 'Raspberry___healthy'
- 'Soybean___healthy'

- 'Squash___Powdery_mildew'
- 'Strawberry___Leaf_scorch', 'Strawberry___healthy',
- 'Tomato___Bacterial_spot','Tomato___Early_blight', 'Tomato___Late_blight', 'Tomato___Leaf_Mold','Tomato___Septoria_leaf_spot', 'Tomato___Spider_mites Two-spotted_spider_mite', 'Tomato___Target_Spot', 'Tomato___Tomato_Yellow_Leaf_Curl_Virus', 'Tomato___Tomato_mosaic_virus', 'Tomato___healthy'

The dataset contains:

- 70295 Training Images.
- 17572 Testing Images.

1.2 OBJECTIVE

"INDIA" where agriculture is the primary source of livelihood for about 58% of India's population, Crop failure is one of the biggest problems being faced by the farmers now a days. There are various reasons why crops fail:

- No proper tools available for farming.
- New plant diseases that the farmer doesn't know of.
- No appropriate use of fertilizers.
- Climate change
- Etc….

Crop failure because of pests and disease amount to Rs.50,000 crore According to a study by the Associated Chambers of Commerce and Industry of India. Which is a huge amount to bear and is increasing every year. To prevent this, we need to minimize crop failure and one of the effective ways of doing this is timely detecting the diseases. This will solve two of our major problems that is crop failure will decrease and we will be able to stop further spread of the disease. By doing this we will be able to achieve food security. There are traditional ways of detection of disease but that have to be done manually and needs to be done by experts, it is time consuming and costly. To find a better time effective and more feasible solution we have developed a model for 14 different plants along with 38 different diseases to detect.

The main objective behind this project is that a person can detect whether the leaf of the plant is affected by any kind of diseases mentioned for any of the mentioned plants. This project in particular builds a model and train the model in a way that we can choose either of the two-model used in the project. The user on the other hand can choose an image on their own, upload it and can know whether the leaf has any kind of disease or not and if any then which disease it has from the ones described.

1.3 SCOPE (SCOPE – WHAT IT CAN DO AND CAN'T DO)

We have used various CNN architecture, combination of multiplayer CNN with linear layers and two different model options:

1. VGG16
2. Resnet34

Transfer learning methods are used to classify images into multiple classes as we have 32 different diseases, we have used transfer learning methods so as to feed the results generated by the CNN architecture to any of the two transfer learning techniques.

The model is going to be deployed on flask server, where the user can upload any image, he/she wants to during run-time and can find whether the plant is diseased or not and can get all kind of information if the plant is diseased. This will help the farmers to predict any kind of mentioned disease so the farmer can take precautionary steps to stop further spreading of the disease.

The Model however is not capable of all the plants as the dataset is limited to 14 different plants and 32 different kind of diseases. The restriction is the model can only detect disease if the disease is available in the dataset.

1.4 TOOLS & TECHNOLOGY USED

1.4.1 GOOGLE COLAB

- We have used google colab for execution of the python code.

Fig. 1.1 Google Colab

1.4.2 FLASK

- Flask platform is used to deploy the built model on it, so as to provide the user with a run-time application to find whether their leaf has any kind of disease or not.

Fig. 1.2 Flask

2. PROJECT MANAGEMENT

2.1PROJECT PLANNING

2.1.1 PROJECT DEVELOPMENT APPROACH AND JUSTIFICATION (PROCESS MODEL USED)

We have been provided with a task which we need to complete, there's no specific feedback system required in our project and we know all the requirements, Design, Development, Testing, Deployment and Maintenance of the project. So we decided to use the most appropriate model for our that is the "Waterfall Model".

Waterfall model:

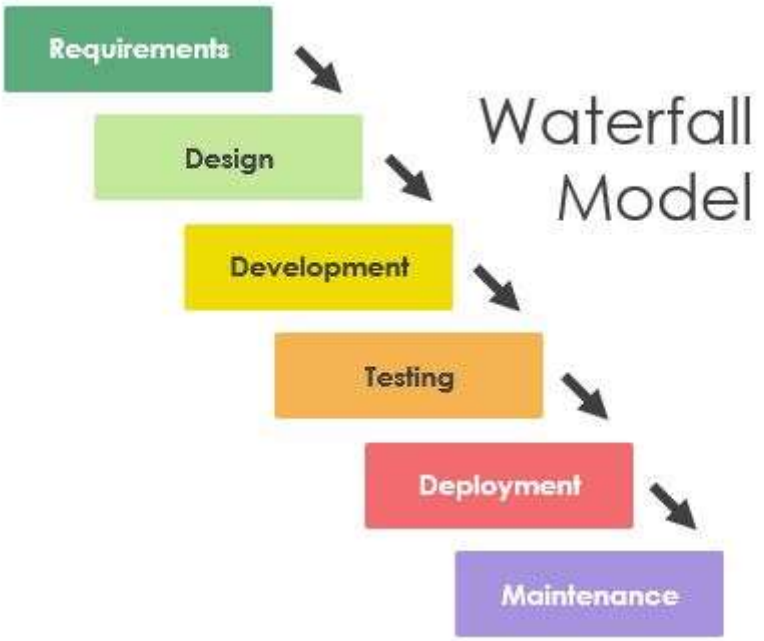


Fig. 2.1 Waterfall Model

Requirements:

We knew all the requirements of the projects as this is not a client assigned project, so the requirements were all known from the beginning.

Design:

No such design specifics are required for the project, it is a web-based application where the entire project is being deployed.

Development:

Development of the project in divided into parts as all the four contributors of the project are going to work on 4 most important models that are the base of the project. All the other work was divided equally amongst all.

Testing:

Testing will be done using the images that are taken from outside the dataset. All the other testing will be done before deploying the model to the web.

Deployment:

After all the testing is done, the project is deployed on the flask platform for it to run on the web.

Maintenance:

There is no requirement for maintenance of the project built. We think it is the most suitable model for our project.

2.2PROJECT WORK SCHEDULING

We have divided the project into different parts and different roles have been assigned to various group members. There are four major division of the project as the basic four model required for the project that is:

- Base Image Classification Model
- CNN Model
- VGG16 Model
- Resnet34 Model

Thus, the four models were divided amongst 4 members of the group and all the discussions of the project were done via Google Meet. We have used ASANA project management tool for the scheduling of the project throughout the period of the project. Here is the Screenshot of the same:

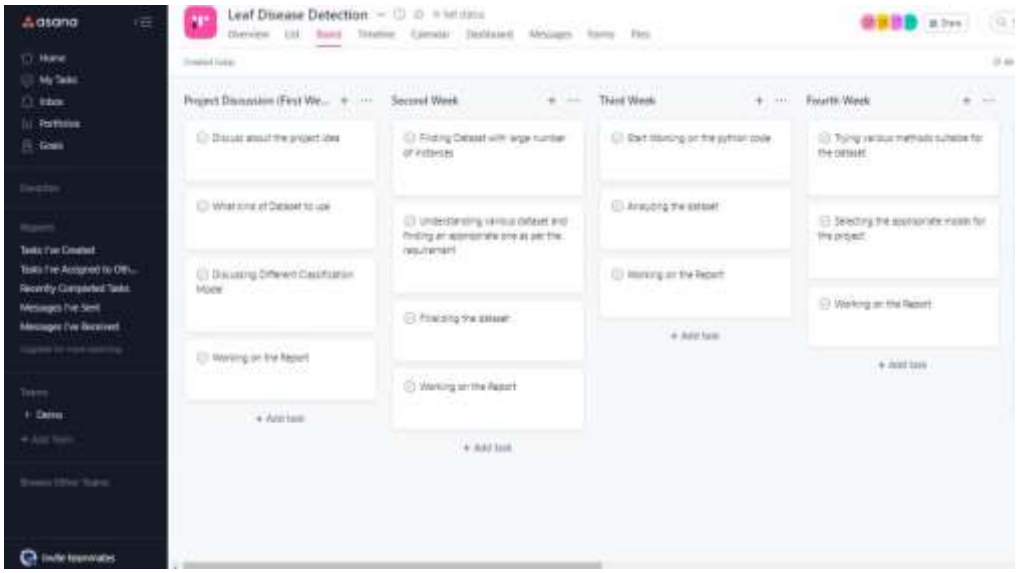


Fig. 2.2 Asana Tool Chart

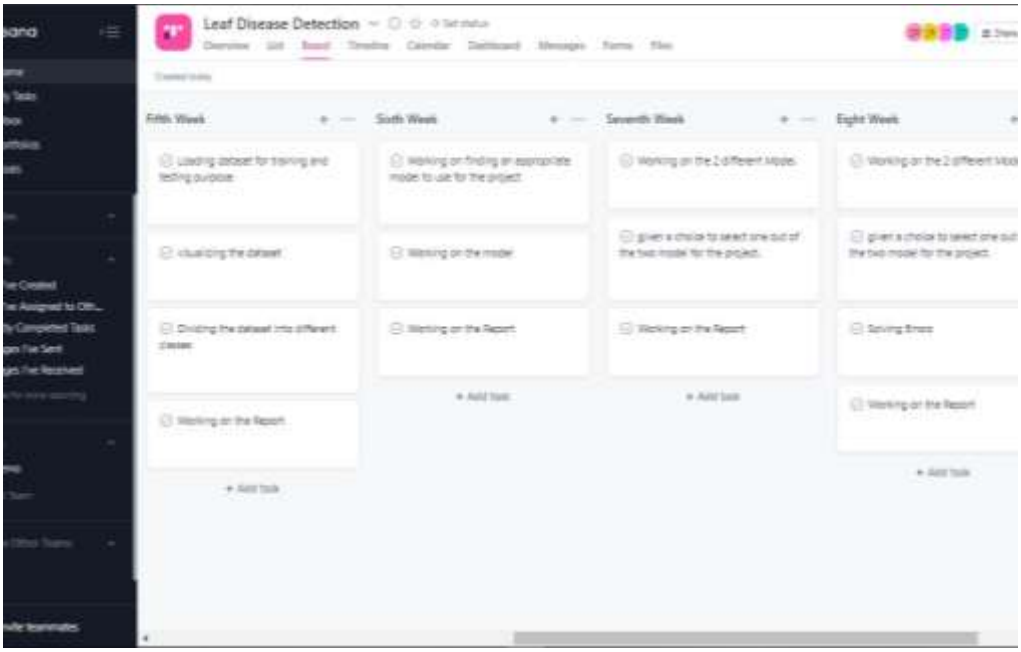


Fig. 2.3 Asana Tool Chart

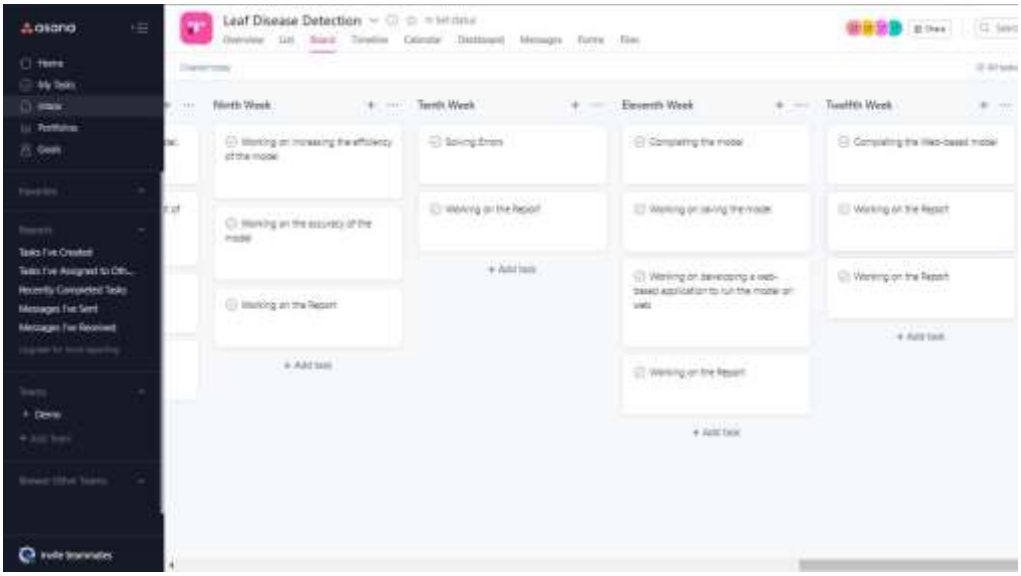


Fig. 2.4 Asana Tool Chart

3. SYSTEM REQUIREMENT STUDY

3.1 USER CHARACTERISTICS

- **End Users**
  - End users in the farmers who need to find the disease in leaf of different plants and if the plant has disease, then what kind of disease does it have.
  - The User need to upload his/her image on the web-based application, by clicking on the check button he/she can get the results.
  - The results would be in this way:
    - It will show you the information of the disease if it can detect the disease, if not then it will not be able to classify the disease.
- **Administrator**
  - Administrator of the project will handle the web-based model in the flask platform.
  - There will be update where the application can get better after the administrator adds new features like image segmentation, more generalised model of the project etc….

3.1.1 EXECUTION DIAGRAM

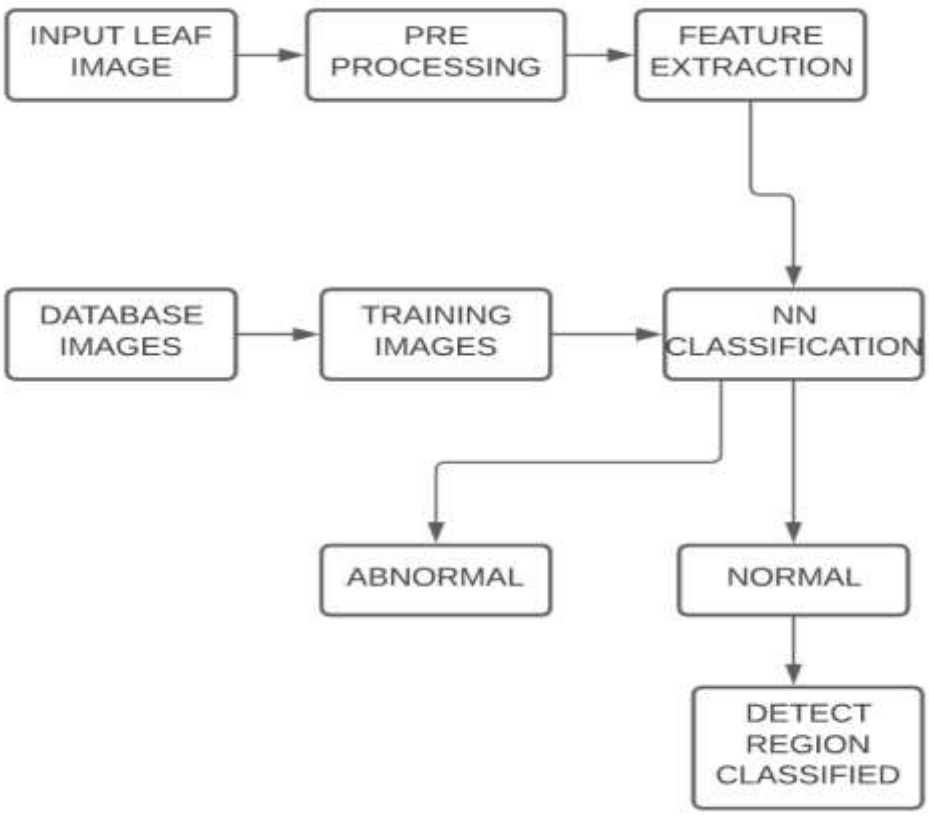


Fig. 3.1 Use case diagram

3.2 HARDWARE AND SOFTWARE REQUIREMENTS

3.2.1 HARDWARE SPECIFICATIONS

- For implementing the model any system with support of Python can be used.
- As the data we have used is not enormously large the processing power required is quite less, any computer with following standards can run this model
  - 4 GB ram
  - 256 MB Dis space
  - Any PC Having browser with internet connection

3.2.2 SOFTWARE SPECIFICATIONS

- Python v (with all supporting libraries)
- Flask v

3.3 ASSUMPTIONS AND CONSTRAINTS

- For this project development we had a dataset where there were plant diseases for some of the specific plants only and there, we predefined diseases it can detect. So, her we assume that the data available is enough to train the model and provide us with the kind of results we need.
- We assume that the user will be testing his/her plant from the predefined categories of plants used for training of the model. The real-time images can only be classified if they fall under any of the plants used in the model as the data.
- The constraint of the project is the image cannot further be segmented enough to get the precise location of the affected area.

4. SYSTEM ANALYSIS

4.1 STUDY OF EXISTING SOLUTION

Previously many people have worked on the same area and there have been many different ways to detect leaf disease detection. For Example,

- Using partial classification for just classifying the affected part, real time monitoring where it continuously monitors the crop and alarms if it finds disease.
- Using neural networks for rubber tree leaves specifically, where it analyzes directly the RGB values of the pixels of a low resolution (15x15 pixels) of the leaves.
- Using thresholding where it discriminates between maize plants affected by fall armyworm from healthy ones using images.
- Using Dual-segmented regression analysis where it detects early deficiency of calcium in lettuce.

However, there have been performed for some specific plants only and for some specific in of diseases. Where in this project we try to cover more than 1 plant and there have been more than 1 disease for each plant, so it has variety of plants with variety of diseases.

4.2 LIMITATION OF EXISTING SOLUTION

The model that was used in the previously trained model by other people had some specific single plant and specific disease can only be detected. Where as in our model the way we have trained the model is more diversified and has more diseases for a single plant. As of now the accuracy of the model is good enough for a real time user to use the model.

Also, as we have not used image localization the model is not able to find the exact position of the leaf affected. There is no system where it suggests the possible solution. The dataset needs to have a greater number of images. The model can be more generalized if we have data for all the plants and all of their diseases.

4.3 REQUIREMENT OF PROPOSED SYSTEM

4.3.1 Functional Requirements

- The first and the most important requirement is the accuracy of the project, it shows how well the model is trained and how well can it predict the output.
- The developed model can be deployed on the web so should be efficient on space and time requirements.

4.3.2 Non-Functional Requirements

- There were no such non-functional requirements as we didn't have any client to work with.
- We designed a basic web-based model.

4.4 SYSTEM WORK FLOW

The flow of the built model is designed in a way that the project can be easily understood by any person looking at the model even for the first time. The project is divided into small chunks of code, main parts of the code are as follows:

1. Preprocessing on the dataset
2. Building a validation set
3. Loading data (Batches)
4. CNN architecture
5. Building Model
6. Training the Model
7. Evaluating the Model
8. Saving the Model
9. Deployment of the Model on Flask

A. Looking at the dataset

- The dataset is taken from Kaggle " https://www.kaggle.com/vipoooool/new-plant-diseases-dataset". The dataset consists of various images of diseased plants and not-diseased plants.

- The images are divided into batch size of 64 images in a single batch.

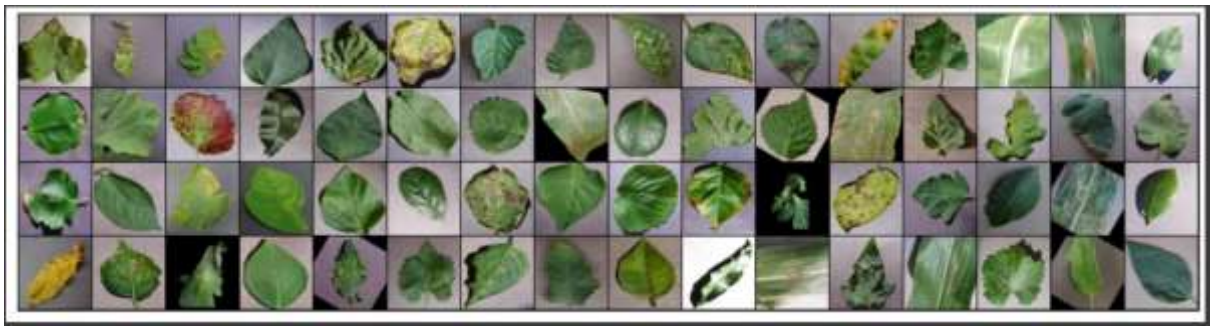


Fig. 4.1 One Batch of images

B. Pre-Processing Data

- The data needs to be transformed into 128x128 Pixels as it takes a lot of time to process with images having 256x256 pixel size.

```
[ ]  transform = transforms.Compose(
       [transforms.Resize(size = 128),
        transforms.ToTensor()])
```

Fig. 4.2 Image Transform

- Increasing the size of dataset by changing other properties of the dataset like inverting the image.
- This will provide us with more instances and the model can learn better.

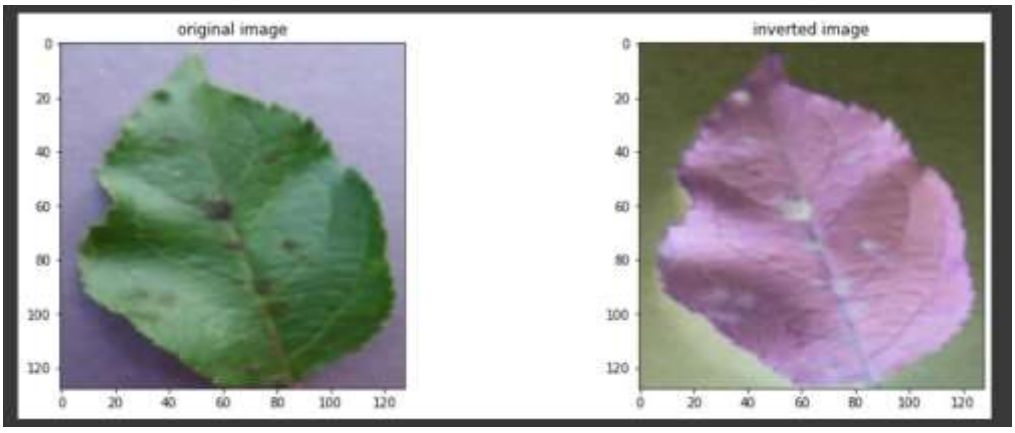


Fig. 4.3 Inverted Image

C. Building the Model

- There are 4 pillars of this project
  1. Base Image Classification Model
  2. CNN Model
  3. VGG16 Model
  4. Resnet Model

1. Base Image Classification Model

```
def accuracy(outputs, labels):
 _, preds = torch.max(outputs, dim=1)
 return torch.tensor(torch.sum(preds == labels).item() / len(preds))


class ImageClassificationBase(nn.Module):

 def training_step(self,batch):
  images,labels = batch
  out = self(images)
  loss = F.cross_entropy(out,labels)
  return loss

 def validation_step(self,batch):
  images,labels = batch
  out = self(images)
  loss = F.cross_entropy(out,labels)
  acc = accuracy(out,labels)
  return {'val_loss':loss,'val_acc':acc}

 def validation_epoch_end(self,outputs):
  batch_loss = [out['val_loss'] for out in outputs]
  epoch_loss = torch.stack(batch_loss).mean()
  batch_acc = [out['val_acc'] for out in outputs]
  epoch_acc = torch.stack(batch_acc).mean()
  return {'val_loss':epoch_loss.item(),'val_acc':epoch_acc.item()}
```

Fig. 4.4 Code

- This is the classification model for image classification.

2. Building a CNN Model
   - A CNN Model is used to reduce the size of image from 128x128 to 4x4.
   - Then Flattening it into a linear layer.

```
class Plant_Disease_Model(ImageClassificationBase):

 def __init__(self):
  super().__init__()
  self.network = nn.Sequential(
     nn.Conv2d(3,32,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.Conv2d(32,64,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.MaxPool2d(2,2), #output : 64*64*64

     nn.Conv2d(64,64,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.Conv2d(64,128,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.MaxPool2d(2,2), #output : 128*32*32

     nn.Conv2d(128,128,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.Conv2d(128,256,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.MaxPool2d(2,2), #output : 256*16*16

     nn.Conv2d(256,256,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.Conv2d(256,512,kernel_size=3,stride=1,padding=1),
     nn.ReLU(),
     nn.MaxPool2d(2,2), #output : 512*8*8
```

Fig. 4.5 Code

3. Building the VGG16 Model
    - This is a transfer learning method of deep learning.
    - This VGG16 Model is used to feed the last layer of the CNN architecture to it and get the desired output.

```
class Plant_Disease_Model1(ImageClassificationBase):

 def __init__(self):
  super().__init__()
  self.network = models.vgg16(pretrained=True)
  num_ftrs = self.network.classifier[-1].in_features
  self.network.classifier[-1] = nn.Linear(num_ftrs, 38)

 def forward(self,xb):
  out = self.network(xb)
  return out
```

Fig. 4.6 Code

4. Building a Resnet34 Model
    - This is a transfer learning method of deep learning.

- This Resnet34 Model is used to feed the last layer of the CNN architecture to it and get the desired output.

```
class Plant_Disease_Model2(ImageClassificationBase):

  def __init__(self):
    super().__init__()
    self.network = models.resnet34(pretrained=True)
    num_ftrs = self.network.fc.in_features
    self.network.fc = nn.Linear(num_ftrs, 38)

  def forward(self,xb):
    out = self.network(xb)
    return out
```

Fig. 4.7 Code

D. Training and Evaluation

- Training the Model and evaluating the model.
- Finding the accuracy of the Model.
- Accuracy of the Model is 98.42%.

```
[ ]  evaluate(model,val_loader)

     {'val_acc': 0.03333333507180214, 'val_loss': 3.8491690158843994}

     history = fit(10, 0.001, model, train_loader, val_loader, opt_func = torch.optim.Adam)

     Epoch [0], train_loss: 0.3521, val_loss: 0.5414, val_acc: 0.8465
     Epoch [1], train_loss: 0.1276, val_loss: 0.5559, val_acc: 0.8561
     Epoch [2], train_loss: 0.0686, val_loss: 0.1504, val_acc: 0.9527
     Epoch [3], train_loss: 0.0790, val_loss: 0.0907, val_acc: 0.9714
     Epoch [4], train_loss: 0.0522, val_loss: 0.1122, val_acc: 0.9660
     Epoch [5], train_loss: 0.0602, val_loss: 0.0827, val_acc: 0.9758
     Epoch [6], train_loss: 0.0419, val_loss: 0.2095, val_acc: 0.9447
     Epoch [7], train_loss: 0.0474, val_loss: 0.0530, val_acc: 0.9830
     Epoch [8], train_loss: 0.0367, val_loss: 0.0924, val_acc: 0.9735
     Epoch [9], train_loss: 0.0344, val_loss: 0.1542, val_acc: 0.9585
```

Fig. 4.8 Code

E. Flask

- Flask API is used to deploy the model on the server as a web-based model.

5. SYSTEM IMPLEMENTATION AND TESTING

5.1 CODING STANDARDS

Coding standards used here are same as suggested my PEP-8 norms. Some of the specifications are as follows:

- Encoding: UTF-8
- Class separation: 2 Blank line
- Method definition separation: 1 Blank line
- Indentations: 4 spaces
- Maximum line length: 79 characters
- Function separation: 1 Blank line
- Standard naming convention:

Identifier	Convention
Module	Lowercase
Type variables	CapWords
Class	CapWords
Functions	lowercase

Table 5.1 Standard naming convention

6. FUTURE ENHANCEMENTS

- The Model is for 14 plants and 32 diseases, we can increase the number of plants in the dataset and the diseases in the dataset so as to make it a generalized model.
- We can use Image Localization to find the specific part that is affected by the disease.
- We can also add remedies for the disease detected.
- We can implement proper method and use some techniques to prevent it from getting affected by the diseases.

7. CONCLUSION

7.1SELF ANALYSIS OF PROJECT VIABILITIES

According to me the project is highly feasible, I think the model designed and the option available to select from either of the two model makes it more viable. The transfer learning methods used help classify the data into different classes more effectively and is able to provide with high accuracy. The accuracy VS loss graph shows the effectiveness of the model. It is clear from the past instances that the model previously designed were for one specific plant and a specific disease. But this model has 14 different plants and 32 different diseases, which gives it a upper hand.

Then comes the part of deploying the model on flask, which helps the farmers struggling to know what the disease actually is finding id the disease is from the above trained 32 diseases. The data is not yet generalized, I think if more data is available then we can convert the model into a generalized one and more and more people will be able to use the model

7.2PROBLEM ENCOUNTERED AND THEIR SOLUTIONS

We faced quite a few problems during the project, there was on particular problem that made problem. We didn't know how to deal with so many classes and classify such a large number of classes. Then we worked on it and found a solution for that, "Transfer Learning". While exploring the technique we went through two different models and decided to use both the models and also discovered on how to classify images if number of classes are more than 1.

7.3 SUMMARY OF PROJECT WORK

Internship experience at BISAG has been great and we got to learn a lot of new things. I studied various research papers and learnt a lot about various methods where how people solved their own problems. After going through various papers, projects and different areas, we decided to go with this particular project. We worked with Leaf data having more than 78,000 images.

I learned on how to work on pre-processing of the data and increasing the size of the data by changing its various properties. I even worked on working with CNN Model and various transfer learning techniques to classify images in more than 2 groups. Previously I have worked with CNN but I have classified the images into 2 different classes only. It was the first time I worked with the classification dataset which needs to be classified in more than 2 classes. We tried our level best for the project. We added two different transfer learning model and even provided an option to the developer to choose from the two methods. The proposed models were VGG16 and RESNET34. We understood the concept and worked on the project with full energy. Still the concepts of transfer learning cannot be learnt overnight so we are still working on it. We submitted our weekly report so we got to work on our formal writing skills. It was a great experience to learn under a company like BISAG.

REFRENCES

1. Arnal Barbedo, J.G. Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus* **2,** 660 (2013). https://doi.org/10.1186/2193-1801-2-660
2. https://www.pantechsolutions.net/image-processing-projects/leaf-disease-detection-using-image-processing
3. Kaggle dataset: https://www.kaggle.com/vipoooool/new-plant-diseases-dataset

Bhaskaracharya National Institute for Space Applications and Geo-informatics

MeitY, Government of India

Phone: 079 - 23213081 Fax: 079 - 23213091

E-mail: info@bisag.gujarat.gov.in, website: https://bisag-n.in/

ISO 9001:2008
ISO 27001:2013
CMMI LEVEL-5

Report Verification Procedure

Date: 23/04/2021

Project Name: Leaf Disease Detection

Student Name & ID: 1. Shrey Parekh – 17IT059

2. Shivam Rank – 17IT094

3. Nir Patel – 17IT073

4. Urvish Soni – 17IT112

	Soft Copy	Hard Copy
Report Format:	☐	☐
Project Index:	☐	☐

Sign by Training Coordinator **Sign by Project Guide**