TULSI HERB'S INFECTION CLASSIFICATION AND PREDICTION USING ARTIFICIAL INTELLIGENCE

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Electronics and Communication Engineering

By

Manjot Kaur

Registration Number: 42000143

Supervised By

Dr Someet Singh

Associate Professor

Electronics and Communication Engineering

Lovely Professional University

Co-Supervised by

Dr Anita Gehlot

Professor

Uttaranchal Institute of Technology

Uttaranchal University

LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2025

DECLARATION

I, hereby declared that the presented work in the thesis entitled **"Tulsi Herb's Infection Classification And Prediction Using Artificial Intelligence"** in fulfilment of degree of **Doctor of Philosophy (Ph. D.)** is outcome of research work carried out by me under the supervision of Dr Someet Singh, working as Associate Professor, in the Department of Electronics and Communication Engineering of Lovely Professional University, Punjab, India and Dr. Anita Gehlot, Professor, in Uttaranchal University, Dehradun. In keeping with general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of other investigator. This work has not been submitted in part or full to any other University or Institute for the award of any degree.

Manjot

Name of the scholar: Manjot Kaur

Registration No.: 42000143

Department/school: Department of Electronics and Communication Engineering/ School of Electronics Communication and Electrical Engineering

Lovely Professional University,

Punjab, India

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled **"Tulsi Herb's Infection Classification And Prediction Using Artificial Intelligence"** submitted in fulfillment of the requirement for the award of degree of **Doctor of Philosophy (Ph.D.)** in the Department of Electronics and Communication Engineering/ School of Electronics Communication and Electrical Engineering, is a research work carried out by **Manjot Kaur, 42000143**, is bonafide record of his/her original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.

Name of supervisor: Dr Someet Singh

Designation: Associate Professor

Department/school: School of Electronics Communication Engineering and Electrical Engineering

University: Lovely Professional University

Name of Co-Supervisor: Dr Anita Gehlot

Designation: Professor

Department/school: Uttaranchal University of Technology

University: Uttaranchal University

ABSTRACT

This thesis investigates the development and implementation of a disease prediction model for Tulsi flowers (Ocimum sanctum). The broad-spectrum purpose of this study is to meet the pressing need for early detection of disorders in Tulsi farming to reduce production losses and ensure long-term crop health management. With modern technologies such as computer vision and machine learning, this research focuses on automating disease prediction using leaf images as primary diagnostic information. Different image processing techniques, feature extraction methods and predictive approaches were investigated with a view to improve the accuracy and efficiency of disease prognosis. Through a systematic approach, the team came up with a robust prediction model that could diagnose some common diseases affecting Tulsi plants based on visual symptoms observed from images taken on leaves. Designed models are expected to assist better precision agriculture practices as well as plant health monitoring strategies which offer bright prospects for researchers, practitioners, stakeholders involved in Basil cultivation or agricultural production. The aim of this research is to develop an automated system for detecting tulsi diseases using leaf images. This paper carefully investigates various image processing steps, which include acquisition, pre-processing, segmentation and feature extraction to prepare visual data for analysis. Also, the application and performance of several machine learning classifiers such as Support Vector Machines (SVMs), Random Forests and Convolutional Neural Networks (CNNs) are investigated in terms of accurately identifying and categorizing Tulsi diseases. Image processing is a crucial aspect of automatic plant disease detection and classification. In this regard, several major image- processing techniques have been employed for tulsi leaf analysis that prepare the visual data for machine learning classification in subsequent stages.

Improved performance and robustness are often realized when ensemble methods that combine multiple classifiers are used. Different classifiers predictions can be combined using simple majority voting or weighted voting schemes. Stacking occurs when a meta-classifier is trained on output from base classifiers to potentially learn how to rectify systematic errors. Boosting algorithms such as Adaboost or Gradient boosting work towards sequentially improving model performance by assigning more weight to misclassified samples. Disease prediction models need to be reliable and effective, hence they undergo rigorous evaluation processes to improve accuracy before being implemented. To evaluate the model's performance and generalizability on different data subsets, cross-validation techniques like k-fold cross-validation are used. Different performance metrics are used to measure classifier's efficiency such as accuracy (which determines how many predicted diseases are true), precision, recall, F1-score (measure of the models' predictions for each disease class), confusion matrix (a detailed classification of correct and wrong classifications) and ROC-AUC (used for assessing how well the model can distinguish between classes). Consequently, a hyper-parameter configuration that is optimal for every classifier must be determined through methods such as grid search optimization, random search optimization or Bayesian optimization. Best performing models are selected and combined into robust ensemble classifiers. The development of a model to predict tulsi flora infection is an important breakthrough in agricultural production and offers practical solutions for enhancing crop safety and yield. The use of machine learning combined with

computer vision technologies effectively enables the analysis of diseases using visual data thus allowing timely interventions and customized treatment options. In this regard, collaboration among disciplines such as agriculture, technology and IT drive innovative progress in precision farming techniques as well as sustainable methods of agriculture. Other future research areas include further refining and optimizing the predictive model, integrating it with other real time monitoring systems, and making it relevant to broader agricultural contexts. These changes aim to provide farmers with unique tools for identifying and managing diseases, thereby strengthening the resilience and sustainability of the agricultural systems. Such progress will help in creating more environment-friendly and efficient techniques for crop health care management through continuous innovation as well as improvement of these technologies, thus benefiting all stakeholders in rural areas.

ACKNOWLEDGEMENT

I would like to offer my heartfelt gratitude to my supervisor, Dr. Someet Singh and my co-supervisor, Dr. Anita Gehlot for their exceptional guidance, unwavering support, and insightful feedback throughout the entire process of completing this thesis. Their combined expertise and encouragement have been invaluable in shaping this research and enhancing its quality. I am also grateful to academic committee members and for their critical feedback and ideas that helped to improve this study.

I would like to extend my sincere thanks to the Research and Development Center (RDC) and the Center for Research and Development Planning (CRDP) for their outstanding efforts in organizing smooth and enriching seminars throughout the course of my PhD studies. These seminars provided invaluable opportunities for intellectual exchange, learning, and networking within our academic community. I am grateful for the dedication and professionalism of the RDC and CRDP teams in coordinating these events, which have greatly contributed to my academic and professional development. Your commitment to fostering a stimulating academic environment is deeply appreciated and has been instrumental in shaping my research journey.

Thank you, family and friends, for your unwavering love, encouragement, and faith in my skills. Finally, I want to thank lovely Professional University for providing the resources and facilities that helped us perform this study. Each of you has played an important role in this scholarly endeavour, and I am grateful for your support.

Name of the scholar: Manjot Kaur

Registration No.: 42000143

Department/school: Department of Electronics and Communication Engineering/ School of Electronics Communication and Electrical Engineering

Lovely Professional University,

Punjab, India

29 May 2025

PUBLICATIONS and CONFERENCES

Journal indexing (Scopus/UGC/Web of Science)	Journal Name	Title of the Paper	Volume, Issue Number & page number	ISSN Number, Impact Factor/SJ R
Scopus	International Journal of Intelligent Systems and Applications in Engineering	Enhancing Precision in Tulsi Leaf Infection Classification: A Stacking Classifier Ensemble Strategy	Vol. 12 No. 20s (2024) 2147-6799	IF: 1.74, SJR: 0.234
Scopus/Web of Science	Journal of Electrical Systems	Unveiling the Potential: Machine Learning and Image Processing for Early Disease Detection in Tulsi Herb	Vol. 20 No. 5s (2024) 1112-5209	IF-0.70,SJR-0.218
UGC/Scopus till 2022	European Chemical Bulletin	Machine Learning Classification Of Infection In Ocimum Tenuiflorum Using Predictive Modelling	2063-5346	IF: 3.71, SJR:0.247
Scopus Conference	1st International Conference on Recent Advances in Computing Sciences (RACS-2022) 4th to 5th Nov 2022	Optimized prediction model using support vector machine for classification of infection in Tulsi leaf	ISBN: 9781032521954 (pbk), ISBN: 9781003405573 (ebk),DOI: 10.1201/9781003405573	-
Scopus Conference	5th International Conference on Intelligent Circuits and Systems (ICICS-2023)" October 12-13th, 2023, Taylor and Francis Books India Pvt Ltd	From Pixels to Prognosis: Machine Learning for Infection Segmentation and Disease Prediction in Tulsi leaf	ISBN: 9781003521716 (ebk)	-
Copyright	Diary No-40570/2024-CO/L Request No. - 173756 Receipt No. - 156802	Enhancing Infection Diagnosis with Tulsi: AI-driven Stacking Classification and Prediction Model"	-	-

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

S.No.	ABBREVIATION	DESCRIPTION
1	AA	Average Accuracy
2	AI	Artificial Intelligence
3	ALEXNET	Alex Convolutional Neural Network
4	ANN	Artificial Neural Network
5	APS-DCCNN	Attention-based Parallel Structure-Deep Convolutional Neural Network
6	AUC	Area Under curve
7	BB	Bounding Box
8	BOF_DP	Bag-of-Features with Dynamic Pooling
9	BOF_SC	Bag-of-Features with Spatial Context
10	CNN	Convolutional Neural Network
11	CV	Computer Vision
12	DICOM	Digital Imaging and Communications in Medicine
13	DLNN	Deep Learning Neural Network
14	DT	Decision Tree
15	DT	Decision Tree
16	FFT	Fast Fourier Transform
17	GBM	Gradient Boosting Machines
18	HOG	Histogram of Gradients
19	HSV	Hue Saturation value
20	IF	Impact Factor
21	kNN	k-Nearest Neighbors
22	LB	Local Binary
23	LBP	Local Binary Patterns
24	LDA	Linear Discriminant Analysis
25	MCA	Mean Classification Accuracy
26	ML	Machine Learning
27	MLC	Multi-Layer Classifier
28	MLP	Multi-Layer Perceptron
29	NB	Naïve Bayes
30	OMNCNN	Optical-Flow based Multi-Task Convolutional Neural Network
31	ORB	Oriented Fast and Rotated BRIEF
32	PCA	Principal Component Analysis
33	RBENN	Round based exact fit neural network
34	RF	Random Forest
35	RGB	Red Green Blue
36	RIA	Radioisotope Analyzer
37	ROC	Region of Curve
38	RPN	Region Proposal Network
39	RRDN	Residual Reconstruction Deep Network
40	SCIE	Science Citation Index Expanded
41	S-CNN	Spatial Convolutional Neural Network
42	SIFT	Scale-Invariant Feature Transform
43	SL	Structured Learning
44	SPSS	Statistical Package for the social sciences
45	SSD	Single shot multi-box detector

S.No.	ABBREVIATION	DESCRIPTION
46	SURF	Speeded-Up Robust Features
47	SVM	Support Vector machine
48	VGGNet	Visual Geometry Group Network
49	WOS	Web of Science

Chapter 1
Introduction

Tulsi, often known as "The Embodiment of Herbaceous Plant," is a notable annual sensitive plant. This tree is widely spread in the tropics and warm parts of India and gives out a nice aroma. It is widely grown throughout the country and serves as both a culinary garden staple and an interior plant with religious and medical properties. The essential oils derived from Tulsi are well known. India, which is still in the developmental stage, is significantly reliant on agriculture for growth and prosperity. However, agriculture has several obstacles, the most notable of which is considerable yield losses. Identifying plant leaf diseases using traditional methods remains a difficult task in the agriculture sector.

The main objective of this research was to determine the optimal number of image processing steps as well as develop a machine learning model for classifying diseased and healthy leaves with a reasonable degree of accuracy, precision, and recall, in the effort of aiming for optimal performance. Most of the time, image post-processing techniques are very useful in identification and in the classification of plant diseases with the latest technologies. Over the past 25 years, great advances have been made in improving criteria such as image analysis in the diagnosis, classification, and identification of Tulsi herb infections. These advances also include the application of computer vision and machine learning technologies which lead to more accurate values for computation, training, testing as well as evaluation. Image processing techniques combined with AI methods gives excellent quality and expected precision and outcomes.

With respect to computer solutions using AI, specific tasks were set and classifier models made for them, and so a deep learning algorithm of some sort had to be selected for getting high accuracy in different environments. As major parts of the plants, leaves possess many attributes like color, texture, and veins, which can differentiate infected leaves in many situations. Systems which have AI incorporated are important to improve systems. The use of low resolution images in this study is noteworthy. An AI based image processing program can set threshold levels for recognition and labeling irrespective of image quality.

That is why an efficient artificial intelligence (AI) system for recognizing the leaves of a Tulsi herb plant is in demand. In this approach, a plant image along with its diseased leaves is processed by a computer which classifies the disease using algorithms meant for identifying leaf diseases and is supposed to classify the disease into one of many pre-defined categories fully automatically. Such systems would indeed benefit farmers growing these crops since they accurately diagnose the disease and only necessitate purchasing pesticides that would help in treating the infection. In addition, the electrical system is not based on complex technologies and therefore requires little power, making it inexpensive and very eco-friendly. Cutting edge techniques for computer image processing of plant’s leaf diseases included the use of MATLAB, OpenCV, and others. Then let’s consider a suitable test sample images with different classifiers to obtain improved outcomes.

Tulsi leaves, blossoms, and extracts are used to make a variety of products with health and therapeutic advantages as shown in figure 1.1. Tulsi tea is famous due to its relaxing benefits and immune-boosting characteristics, whereas Tulsi essential oil is utilized in aromatherapy to encourage relaxation and mental focus. Tulsi capsules and supplements are popular due to their adaptogen properties, which help the body cope with stress and promote general vitality. Tulsi is used in skincare because of its antibacterial and anti-inflammatory characteristics, which improve skin health and treat diseases such as acne and inflammation. Furthermore, Tulsi-based hair care products nourish the scalp and stimulate healthy hair development. Whether eaten as herbal treatments, put into personal care items, or employed in culinary purposes [1].

Figure 1.1. Tulsi byproducts

1.1 Tulsi types and benefits

Tulsi, usually referred to as the sacred herb (the Ocimum tenuiflorum), is a highly valued herb in India, frequently associated with spiritual and therapeutic benefits. Tulsi comes in various types, each with its own distinct properties. Here are the most popular types as shown in figure 1.2:

Figure 1.2. Assembly of different types of tulsi leaves

1. **Rama Tulsi** (Ocimum sanctum) has become one of the most popular types of tulsi, also known as Holy Basil in English. It is revered in Ayurvedic medicine and has spiritual, medical, and therapeutic value in Indian culture. The following contains extensive information about Rama Tulsi. Rama Tulsi's leaves are bright green, smooth, and oval-shaped. The leaves are normally 2 to 4 cm long, with slightly serrated edges. The Krishna Tulsi's reddish stems stand out in a stark contrast to the green ones. These are minute flowers that are either white or purple in colour and are found in clusters. Rama Tulsi is a robust shrub that can grow to a height of approximately 1-2 feet along with a thick bushy coverage. Rama Tulsi is known for its medicinal value as well as the fact that it has been used in many Ayurvedic practices from a long time ago. This herb is considered to be an adaptogen, i.e., it helps the body to cope with stress and maintain equilibrium. It has got numerous health benefits which include the following:

   a. Respiratory Aid: Rama Tulsi is often regarded as an option for treating respiratory diseases like asthma; bronchitis; and the common cold. It has anti-inflammatory and anti-bacterial effects that helps clear the lungs.
   b. Potency Booster: It also has powerful antioxidants and antibacterial substances that work to enhance one's immunity and fight off sickness. Rama Tulsi helps with removing gas, acidity, and indigestion, thus making the digestive system more efficient. Stress, worry, and tension can be modulated with Rama Tulsi due to its well documented adaptogenic properties.
   c. Decreasing Inflammations: It also has anti-inflammatory properties and is widely used for easing join pain and swelling [2-3].

Rama Tulsi holds a significant place in Hindu society prominence as it is regarded as a sacred plant that protects and is used as a decoration in most homes as well as temples. The following active principle are contained in Rama Tulsi leaves:

   a. There is a compound called Eugenol, which exerts analgesic and anti-inflammatory effects.
   b. Ursolic acid is a compound that possesses both tumoricidal and antibacterial properties. Linalool is a compound that relieves stress and promotes relaxation.
   c. Rosmarinus acid is a potent antioxidant effective against oxidative damage and inflammation.

It has many other uses:

   a. Rama Tulsi leaves are dried and made into herbal teas which enhance lungs and gut health.
   b. The leaves are also processed into essential oil which is popular in the treatment of diseases in the practice of aromatherapy and traditional medicine.
   c. Cooking: In some regions, Rama Tulsi is cooked with as it is a flavouring herb that adds a lovely fragrance to the food.

Ancient Ayurvedic literatures such as Charaka Samhita or Sushruta Samhita mention Rama Tulsi and its medicinal advantages. A good number of studies have substantiated its usefulness in healing. For example, a study that appeared in the Journal of Ayurveda and Integrative Medicine outlined tulsi's important function as an adaptogen, which increases one's capability of coping with stress, while other studies looked at its antibacterial and antioxidant properties.

2. **Krishna Tulsi (Ocimum tenuiflorum)** on the other hand best known with the name of Shyama Tulsi is a particular type of tulsi with dark purple leaves and dark purple stalks. Just like any other variety of tulsi, Krishna Tulsi is highly respected in Ayurvedic medicine and in Hinduism because of its great physical and spiritual qualities. Krishna Tulsi has leaves which are dark purple or purple green but are slightly shorter and less smooth than Rama Tulsi's leaves. These leaves tend to cut at the edges, displaying a serrated edge. Often, Krishna Tulsi has stems that are purple which differentiates from other types of tulsi where for instance Rama Tulsi has green stems. The flowers are small, violet coloured and mostly come in clusters. Krishna Tulsi grows as a bushy shrub and ranges approximately from 1-2 feet in height. Krishna Tulsi is one of the strongest types of tulsi especially in terms of therapeutic efficacy [4-6]. It contains PHYTOCHEMICAL which is said to have a positive impact on a variety of health benefits:

    a. Respiratory Health: Krishna Tulsi is often known to be effective in Asthma, Bronchitis, Sinusitis and other respiratory ailments. Its inflammatory and antibacterial features assist in clearing respiratory pathways and in unblocking congested airways.
    b. Its antibacterial properties allow it to be widely used in folk medicine specifically in the treatment of skin diseases, infections and acne.
    c. Health of the Digestive System: Krishna tulsi helps in the process of digestion, helps in healing stomach ulcers, and helps in reducing the amount of acid produced in the stomach.
    d. Anti-Oxidant and Anti-Cancer Agent: Like containing Rosmarinus acid, which is a known Cancer Chemo preventive, Krishna Tulsi contains Flavonoids, these help to combat oxidative stress caused by free radicals, which are sources of cancer development.
    e. Antibacterial properties: Some of the antibacterial essential oils in Krishna Tulsi are eugenol, linalool are effective against bacteria, viruses and fungi
    f. Diabetes Management: Some studies suggested that Krishna Tulsi is useful in diabetic people as it has the potential to lower blood sugar levels.

Krishna Tulsi is also considered to be devoted to lord Krishna which is the reason why it has that name. It is considered as holy and used during pujas and offerings especially in temples of Lord Krishna and Vishnu. Krishna Tulsi like other types of tulsi is sown around homes and temple in Hindus for blessings and protection. It is also believed to purify the area as well as provide peace and serenity in the hearts of the residents. Krishna Tulsi also has a lot of bioactive compounds, which are responsible for its medicinal properties.

a. Eugenol: An important oil with anti-bacterial, anti-inflammation properties with pain-relieving effects
b. Carvacrol: A chemical known for its anti-bacterial and fungal properties
c. Ursolic Acid: A chemical located in numerous herbal plants, having anti-inflammatory and anti-cancer properties.
d. Rosmarinic acid decreases the likelihood of cells sustaining oxidative stress as it is a strong antioxidant.
e. Linalool is stress and anxiety calming and relieving having a soothing effect on the nervous system.

Krishna Tulsi is known to be good for respiratory conditions, stress relief and aiding in digestion and is widely used in teas and infusions. It is also used to make essential oils that have anti-bacterial and therapeutic properties. Due to its antibiotic properties, it is one of the constituents of many formulations in the ayurvedic treatment of acne vulgaris, rash and many other skin diseases. Krishna Tulsi is taken in the form of herbs, capsules and tinctures as it is said to be an adaptogen which assists in stress and balance. Recent research has substantiated some of the traditional uses of Krishna Tulsi. One of the studies in the journal of Ethnopharmacology shows that tulsi has great potential as an antibacterial and anti-inflammatory agent particularly for skin infections and inflammation. Similar benefits were also concluded in a study by Phytotherapy Research which showed antidiabetic properties of tulsi extracts along with the ability to regulate blood glucose levels [7-10].

3. **Vana tulsi (Ocimum gratissimum)** The leaves of Vana Tulsi are quite a piece - they are finely textured, pretty big and oval, and dark green in color. The leaves vary in size from 3 to 6 cm in length and have a degree of uncleanness due to hairs being present on the leaf especially on the base. The stems are green and can measure up to 1 metre. It has tiny whitish-or purple-flowered heads clustered together. It has many medicinal properties:

    a. Respiratory Health: It is suggested that Vana Tulsi may be useful in the management of respiratory conditions such as asthma and bronchitis. It supports mucus ESP actors removers while also improving congestion remote.
    b. Digestive Aid: It is used in the management of other gastrointestinal disorders such as excess gas and dyspepsia.
    c. Antimicrobial: Composed of substances that are bactericidal, fungicidal, and virucidal, which are helpful in control of a range of infections.
    d. Traditional Medicine: Used often for herbal medicine, teas, and fortification.
    e. Aromatic Use: Vana Tulsi is known for the scent and has been largely used in the composition of perfume and incense.

Vana Tulsi is in most cases almost regarded as a scratch of wilderness and is said to be mostly found in forests and mountains.

4. **Kapoor Tulsi (Ocimum Kilimandscharicum)** These plants are also known as Kapoor Tulsi and have a slightly thicker hair texture and a smaller diameter in comparison to either Rama or Krishna Tulsi. The leaves are ovate in shape and have an acuminate tip. Kapoor Tulsi is herbaceous, with green stems and a bushy shape, growing up to 3 feet in height. These plants have minuscule flowers-likely of purple or white color-which sprout in lengthy varieties. It has medicinal properties:

   a. Plant help with breathing problems: This is a very prominent feature of Kapoor Tulsi and for this reason, it is employed in asthma, cough, and treating common colds.
   b. Plant of a relief: Contain substances that help underlie irritation and heal swelling in the respiratory one.
   c. Anti infection: Displays robust antifungal and antibacterial properties
   d. Analgesic: These have the capacity to relieve pain especially when taken in oil or extract form.
   e. Kapoor tulsi is commonly used for extracting essential oils which are widely used in aromatherapy because of the soothing and relaxing help it provides.
   f. Aromatherapy can be applied using massage oils, diffusers or even in the ordinary treatment combination.

Kapoor Tulsi is sometimes employed in religious practices and has meanings of cleansing and safeguarding.

5. **Amrita tulsi (Ocimum sanctum)** In contrast with other types of tulsi, Amrita Tulsi has broader dark green leaves possessing a stronger odour and has a taller stature ranging between 1 to 2 ft with a rough green stalk. Also, its white blossoms are considerably obscure as compared to those of Vana and Kapoor Tulsi, and not so much prominent. And notice, it has many medicinal purposes which Includes:

   a. Stress Relief: Amrita Tulsi is praised for its tranquilizing qualities, with the essential task of alleviating stress and improving the mindset.
   b. Immune Boosting: improves immunity and promotes good health.
   c. Digestive Aid: Also like other types of tulsi, it enhances digestion and may relieve some stomach discomfort.
   d. Antioxidant: Protects the body against oxidative stress by providing a range of antioxidants.
   e. Teas and other herbal and traditional formulations often incorporate Amrita Tulsi.
   f. Culinary Use: In different regions, its leaves are used to flavour and enhance food products because of their aromatic properties.

Amrita Tulsi is in fact spiritual in nature, which is why it is one associated with immortality and health and the word ‘Amrita’ means immortal in Sanskrit.

Table 1.1. Tulsi types comparison

Tulsi type	Appearance	Medicinal properties	Uses	Cultural significance
Vana Tulsi	Large green leaves, hairy texture	Respiratory aid, antimicrobial, stress relief	Herbal teas, traditional medicine	Found in wild, nature-associated
Kapoor Tulsi	Glossy green leaves, bushy plant	Respiratory benefits, analgesic, antimicrobial	Essential oils, aromatherapy	Used in rituals, purification
Amrita Tulsi	Bright green leaves, aromatic	Stress relief, immune boosting, antioxidant	Herbal remedies, culinary use	Associated with immortality
Rama Tulsi	Bright green, smooth, oval-shaped leaves with green stems.	Respiratory health (asthma, bronchitis, colds) Boosts immunity Aids digestion Stress relief	Teas and infusions. Essential oils	Associated with Lord Vishnu and Krishna. Planted for spiritual protection
Krishna Tulsi	Dark purple or purple-green leaves and purple stems.	Treats skin infections and acne. Relieves respiratory issues. Regulates blood sugar. Stress relief.	Herbal teas and skincare products Essential oils	Sacred to Lord Krishna Used in religious rituals and offerings

Shah et al. (2019) carried a study on adoption rate of Tulsi Farming from year 2016 to 2019 [11]. The research demonstrated that the usage of Tulsi is increasing every year but there is crop damage as well every year, which is degrading the harvesting yield per year.

Table 1.2. Classifier model evaluation by prevailing literature on tulsi

Year	No of Cultivators	Tulsi volume (in Tones)	Total value 100000-INR
2013	**200**	**87**	**6.10**
2014	**211**	**126**	**10.12**
2015	**234**	**140**	**11.23**
2016	**259**	**155**	**12.43**
2017	**400**	**241**	**19.28**
2018	**552**	**440**	**30.45**

Figure 1.3 and Table 1.1 depicts the consumption and wastage in INR for consecutive seven years, where the main area of focus is to minimize the crop damage using various prediction classifier models.

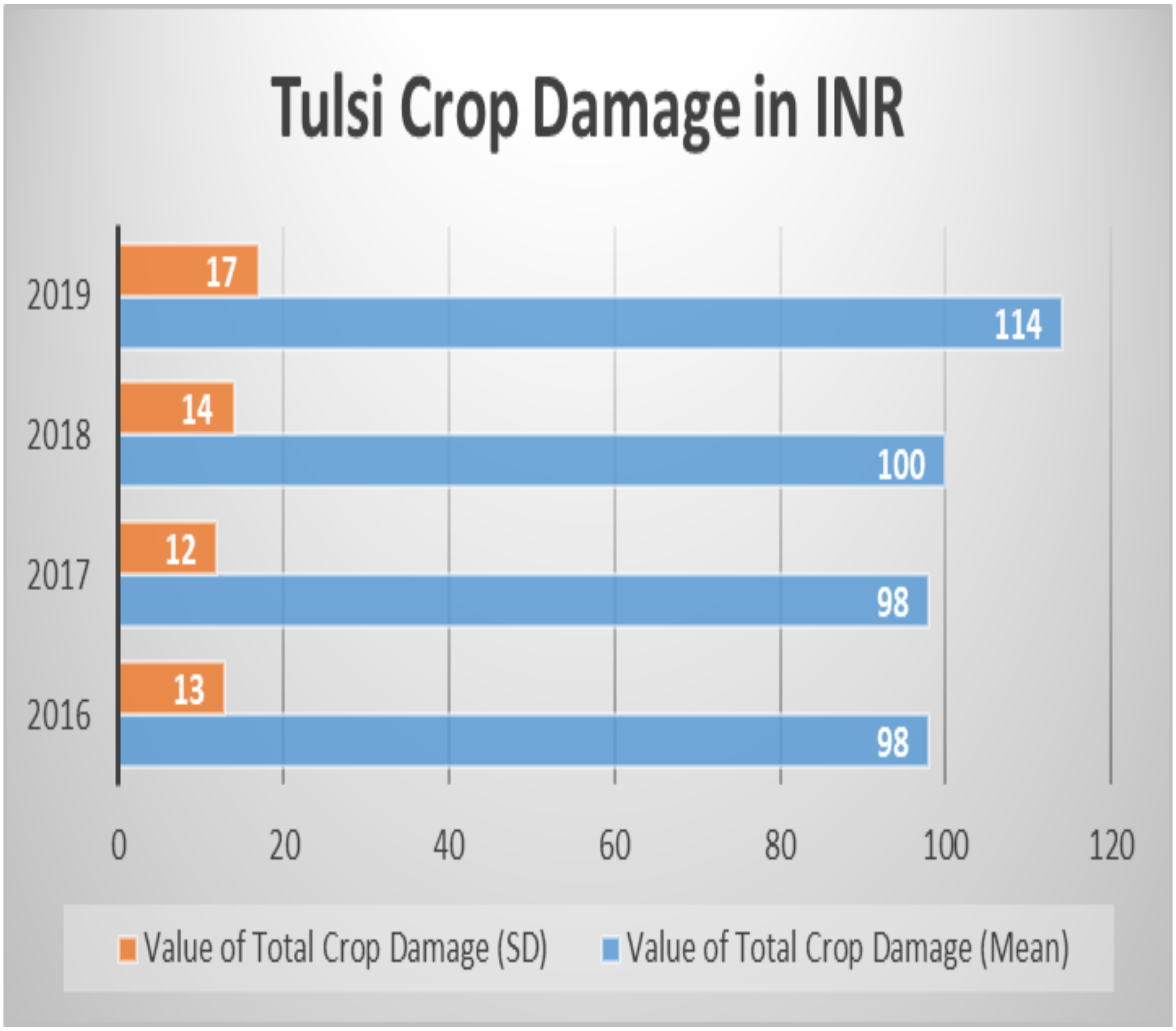


Figure 1.3. Tulsi crop damage from 2016-2019

Table 1.3. Mean and standard deviation crop damage analysis

Year	Mean	SD
2016	98	13
2017	98	12
2018	100	14
2019	114	17

1.2 Tulsi Infection Analysis

Tulsi leaves have the tendency to get affected by various types of diseases like fungal, bacterial and insects. Table 1.3 explains the causes and symptoms in different types of tulsi leaf areas.

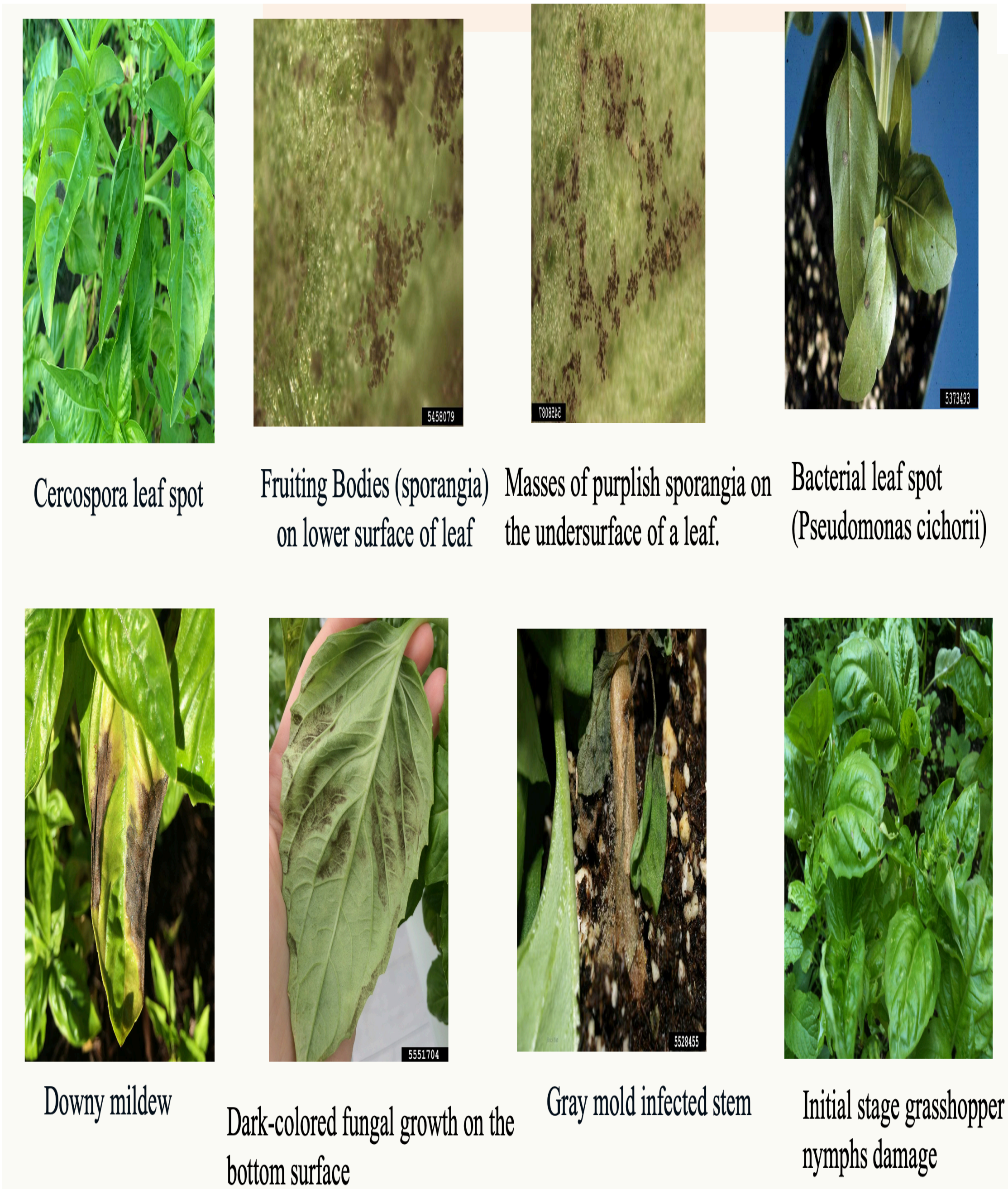


Figure 1.4. Tulsi leaf infection types

Table 1.4. Types of infections on tulsi

Infection	Cause	Symptoms
Root rot	Fungal	Common seed germination and seedling growth concerns include seed failure, collapse of germinated seedlings, brown, shriveled patches appear near the stem's base, as do brown, water-soaked roots. These signs frequently signal numerous plant diseases or environmental stresses that impair seedling health. It is critical to quickly identify and address the underlying reasons to avoid further harm and ensure healthy plant establishment.
Flea beetle	Insects	The presence of microscopic holes or pits in leaves produces a characteristic "shot hole" effect on foliage. This damage is most obvious on young plants and seedlings, which are more vulnerable. Plants that have been infected may develop slower, and severe damage can eventually lead to plant death. Identifying and correcting the source of this leaf damage as soon as possible is critical for preventing future damage and maintaining plant health.
Fusarium wilt	Fungal	The yellowing and dying of leaves, small brown differences on lower leaf surface or the overall death of the sapling indicates that there is serious issue related to its health. For plant's efficiency biggest task is managing the issue as a shortage of time can further deteriorate the state of plant.
Pests	Insects	Aphid can cause mass death of growth in shoots and plants while also causing yellowing and putting spots all over the leaf. There is also a by-product of aphids which is called as honeydew a syrupy substance which is likely to contribute in the spread of sooty fungal growth in the plants. So it is important to have adequate measures ready against aphids otherwise the condition of plants is likely to get worse.
Leaf spot	Bacterial	In case of bacterial infection, in order to contain it leaves or stems of the plant would exhibit presence of black or dark patches which would be in a circular or irregular shapes.
Gray Mold	Fungal	The existence of thick fuzzy mycelial coating of brown to grey coloration on the surface of stems and leaves as well as falling plant litter indicates a fungal disease. The brownish fuzzy growth on leaves and sundried lumps of dead plants may cause these plants to wither and fall off the plant, and deep infections on the stem can kill the plant if treatment is delayed. This kind of fungal condition should be recognized and dealt with as

		quickly as possible in order to stop the further contamination and safeguarding the health of the plant.
Slugs & snails	Mollusc	Regularly formed holes in leaves and stems may indicate pest damage. Flowers and fruits may also be harmed. Severe infestations can cause shredded leaves. Pest infestations must be addressed as soon as possible to reduce damage and maintain plant health.
Root knot nematode.	Nematode	The presence of galls on roots, which can vary in size but are normally less than 3.3 cm (1 in), might reduce plant Vigor. Affected plants may yellow and wilt, especially in warmer weather. To prevent further loss in plant health, the root gall issue must be identified and addressed as soon as possible.
Downy Mildew	Fungal	Yellowing of the leaves, which usually begins around the centre vein and spreads outward, may be grey, fuzzy, or downy growth on the lower leaf surface may occur after this. In addition, the plant may develop brown to black angular necrotic patches. Identifying and treating these signs as soon as possible is critical for preventing further harm and maintaining plant health.
Grasshoppers	Insects	Nymphs and adult bugs feed on the leaves, buds as well as delicate stems. In the first stage, the nymphs circle holes on the leaves to make great holes. The bug maturing starts eating out all the leaves on a particular plant causing more severe loss on that plant. Their early detection and especially treatment are very important to avoid loss of health and growth of the plants significantly.
Cutworms, owlet moths, loopers and underwings	Insects	In the early stages, the larvae of these insects feed on the plant's terminal clusters. As they mature, later-stage larvae skeletonize the leaves, leaving only the veins. Furthermore, they are known to shear seedling stems near the base, resulting in considerable losses in plant establishment and growth. Detecting and treating these pests as soon as possible is critical for minimizing damage and ensuring healthy plant growth.
Cercospora leaf	Fungal	The spot on the leaf is a fungus infection that forms round to irregular dark areas on brightly coloured leaves. This fungal infection commonly affects plant leaves and can be caused by a variety of fungal infections. After the infection spreads, the black patches that we see are actually lesions that are left after the fungus has attacked the tissue of a leaf and caused damage to it.

To be able to address the leaf spot infection problem, it is most essential to establish the particular fungal pathogen underlying the infection since the control of the stated pathogen may require different strategies. Cultural practices such as suitable spacing plants, circulation of air, and exclusion of watering from above the plants also help to reduce the chances for the occurrence of any fungal disease. Apart from that, in certain conditions, any medical breach such as treatment with fungicides may be needed to protect the plant and control outbreak. Regular Observation and Prevention are of major importance when it comes to dealing with plants, leaf spots in particular.

Two of the major issues challenging crop management that AI and machine learning will be able to help address include the optimization of resources and meeting the specific problems. This technology may help farmers map and forecast properly the yields and plan better accordingly in order to meet the supply needs of the market and avoid meeting the resources in a wasted fashion. Furthermore, AI derived solutions can assist farmers in making better decisions as to when to harvest the crops as well as when to sell them based on current market prices, thereby increasing the profitability of the business.

Moreover, the algorithms built into machine learning make it easier to find and cut off the dangerous weeds without human resources and chemical solutions which not only strengthens the plants themselves but also makes the farming techniques eco-friendly. Image based AI systems can identify the diseases in the plants in a swift manner and such targeted treatments are delivered in time thus preventing the severe spread of the disease. In addition, it is worth noting that the correct recognition of weed species with the assistance of artificial intelligence improves the weed management strategies used, leading to targeted and efficient control measures based on species specific characteristics. Overall, the use of ML and AI technologies in agriculture allows for more efficient management of cultivated crops, resource optimization and the introduction of ecologically clean farming methods without compromising the security and profitability of farmers' food supply.

1.3 Image Classification Process

The image classification method was divided into six segments, highlighted here:

i. Pre-processing Image
ii. Segmentation of Image
iii. Extracting features
iv. Selecting features
v. Machine Learning Classifiers
vi. Classification Output

AI encompasses reasoning, learning, real-time identification of problems, and even adaptive and corrective measures. In Machine Learning, a model is trained to differentiate and classify data based on certain parameters. It consists of formulating data, selecting a suitable algorithm, implementing it, and evaluating the results, taking into consideration many significant parameters. Deep learning forms an important part of the machine learning area while machine learning assists the Artificial intelligence process. As machine learning is a subset of AI, it's important to note that not all AI qualifies as machine learning. Aspects of model performance, accuracy, interpretability and computational resources are to be considered in detail when determining which AI model is best suited to a problem statement. Seeds, regarding traditional agricultural methods, are regarded and evaluated by agriculture professionals' visual assessment of plant diseases. However, due to innovations in technology, machine learning models can assess images of plants' leaves to analyse and check for infestations even in automate systems.

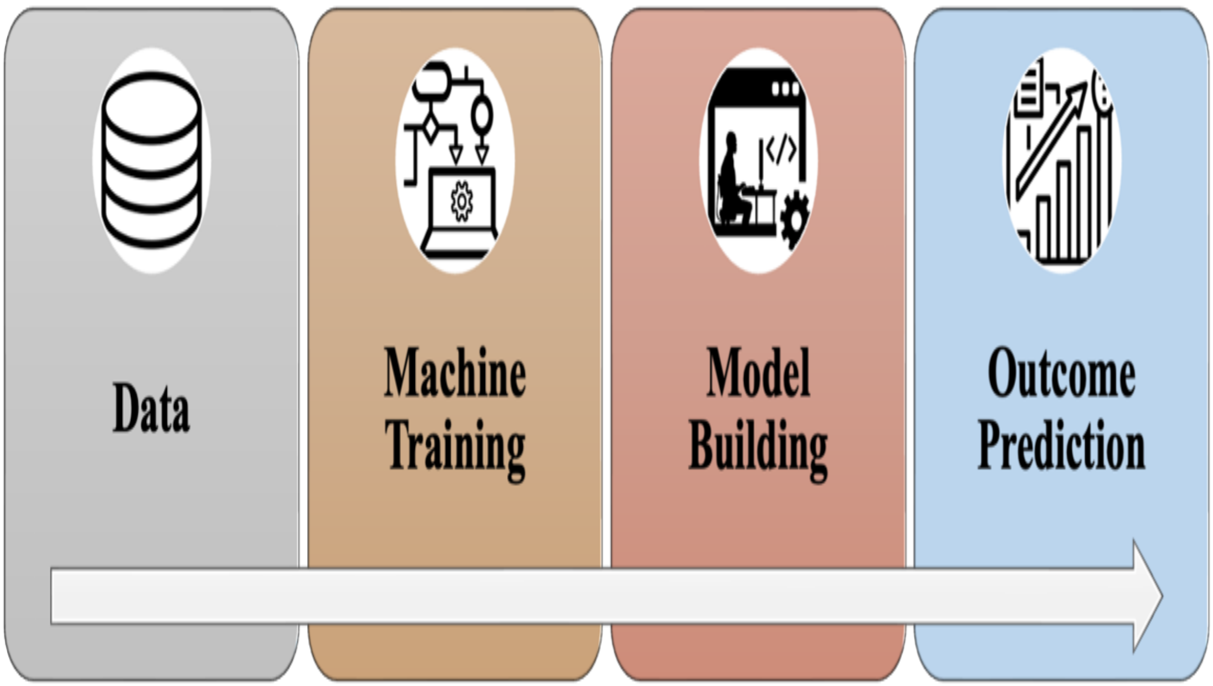


Figure 1.5. Intelligent model for predicting leaf infection

An Artificial Intelligent Model was developed with the purpose of determining diagnosis and classifying infected areas of a disease on plants leave in order to use appropriate strategies which would enhance the healthy growth of the plant and the productivity of the plant overall. So, decisions can easily be automated and understood by AI, which uses mathematical equations to demonstrate the cognitive processes involved in decision-making. This application undoubtedly represents a new wave in agriculture and it concerns the issues of precision agriculture, forecasting, mechanization, increased quantity and quality of crops grown. In addition, it allows for estimating yields and tracing maps, more rational methods of harvesting a crop, meeting the requirements for healthier foods without the excess production of such foods, grading possibilities, and accurately diagnosing and finding out how to make the plant disease suffer and recover from it.

To begin with, the first task of data collection was essential, meaning that it involves sourcing different datasets that contain plant characteristics, external factors, and disease manifestations. This might incorporate images of the plants, sensor information (temperature, humidity etc.), and genotypes. Preprocessing tasks consist not only of cleaning the statistics and reconstructing missing entries but possibly including datasets for causes of class imbalance particularly if some diseases are sparse or are seasonal. Feature extraction is of paramount importance as it allows to detect specific features from images or sensors that are relevant to the detection of a disease. The choice of the classifier to be used for plant disease identification was affected by the structure of the dataset and the level of explainability required. CNN's are good for images of plants and seek to facilitate accurate disease diagnosis through use of transfer learning from models trained on large datasets.

On the other hand, using decision trees or even SVMs might enable one to be more interpretive on which plant characteristics are most useful in predicting disease incidences. In regards to the methodology, model training and validation involves dataset partitioning, which includes the training, validation and test splits, hyperparameter optimization, and performance evaluation such as accuracy, or Recall among others. Lastly, integration of these models into agricultural processes warrants for yes scalability, ability to perform advanced analytics at the edge (e.g. edge AI), and integration with already established farm management systems for timely and efficient disease management solutions.

In essence, the work involved in preparing a model which predicts plant disease classes accurately was essential for the health of the population of crops, hence increasing the yield. Approaches based on machine learning are powerful in processing complex data and accurately diagnosing plant diseases based on several observable features. One of interesting strategies in classifiers construction was stacking, where multiple basic classifiers were used to achieve the specific targeted goal. This strategy incorporates a number of different learning algorithms within a stacking architecture, and as a result, makes effective the complementary features of a number of models to enhance the accuracy and reliability of diagnosis of plant diseases.

Employing stacking classifiers in predicting plant disease can be seen as the new frontier of machine learning in agriculture. This technique not only supports the collecting of plant images, environmental conditions, and previous disease incidences, but also facilitates the constructing of a more complex model that can handle many input variables working in conjunction. As this backdrop has been set, we will attempt to describe the basic concepts and processes related to the stacking classifiers for classifying plant diseases, and their enormous promise to transform disease control practices and assist sustainable agriculture. Models based on ML algorithms have the capability to identify possible plant diseases and pest attacks in advance by means of picture identification and data processing. As a result, farmers are able to take timely measures and thus minimize the losses as well as the use of wide-spectrum chemicals. Planting using ML also solves problems of distinguishing between a deficiency of nutrients and diseases, thus rendering more accurate treatment methods. Historical, climatological, edaphic, and farming data can be put together in order to develop algorithms that would predict yields. This information aids farmers in determining the correct time to plant, the quantity of resources to use, and how to market their crops, thereby improving production and profits. Machine Learning is basically self-directed and can improve or optimize each aspect of the agricultural supply chain by demand prediction, improving logistical issues and reducing waste. Such predictive analysis can properly time the harvesting and transportation of fruits to ensure that only minimum time elapses between harvesting and delivery to the market reducing postharvest losses.

The ML supported decision support systems provide farmers recommendations by integrating real time and historical data. Therefore, these technologies provide means for effective and efficient decisions on the determinants of crop yield such as planting and soil health management and pest control. In its essence, machine learning in agriculture lies in increased efficiency, greater availability of resources to carry out operations and greater returns on investment while some of the repercussions include resolution of climate change, resource scarcity and even food security concerns. The most evident outcome of ML are expansion of possibilities for development of essentially different approaches to farming and its management.

Educational papers clearly show that agricultural diseases are widespread and that they depend on climatic and environmental conditions. But not much effort has been put to specifically find areas of infection on the leaves of Tulsi, which is an important herb in the herbal and ayurvedic therapies. Accurate identification of the disease along with its minimal impact was the foremost requisite to guarantee optimum quality and production of the yield and prevent further outbreaks. For the enhancement of the infection, we must design a computerized automatic diagnosis system to detect the problem and major sicknesses through images with the use of machine learning concepts.

We can however develop accurate and reliable learning algorithms for diagnosing plant diseases by using publicly available data sets. Technologies have integrated systems which help to reduce errors, assist in monitoring over time, reduce manual activities, foster digital assistance, enhance quick operations, and enhance rapid decision making. Therefore, it is expected that artificial intelligence will greatly assist in changing the face of agriculture by ensuring that every stage of plant growth is monitored and thus more food will be produced.

1.4 Research Objectives

The study aims to fill the gaps in the existing approaches of Automatic illness detection model, empowering it towards a disease detection solution that is rapid, accurate, efficient and cost effective by focusing on disease detection in its nascent stage. The aims of this particular system include:

i. Develop a variety of Tulsi herb disease damage and classification levels.
ii. Analyse the disease type present on the Tulsi leaves with the help of image processing techniques.

iii. Devise a strong training model for predicting disease from pictures of the leaf.
iv. Carry a wide range of testing in order to assess the performance and effectiveness of the prediction model.

These aims serve to each other in forming an improved disease detection system that facilitates an early detection and control of fungus diseases in Tulsi plants. The techniques proposed include image processing and machine learning to enhance the accuracy and efficiency of disease diagnosis including better agricultural practices and plant health management.

This automatic illness detection model for Tulsi herb plants could turn out to be extremely useful and beneficial on the grounds that it caters to the particular challenge of effective diagnosis and management in agriculture. Further, the purpose of the prototype was to get insights about the patterns and symptoms once focusing on the detection and classification of various species specific diseases of Tulsi. The system is based on novel digital image processing techniques to classify virgin leaf photographic images by their mineral disease characteristics such as texture, colour, shape, and venation pattern. This makes possible the training of a specialized machine learning system that utilizes the input photos of leaves to forecast the possibility of developing certain diseases. This would allow for prompt actions and measures to be taken aimed at reducing losses of crops while ensuring the plants are sound and healthy.

The goals of the suggested method were practical since they focus on in-depth examination and model evaluation in order to confirm a prototype's viability and reliability. The last objective was to check the accuracy, repeatability, and compliance of the model in identification and classification of diseases, which will be achieved through extensive trial runs on numerous specimen sources. Finally, the use of image processing and algorithms for learning for agricultural purposes has the potential to alter disease control practices, allowing producers to make more educated decisions and adopt focused treatment tactics to keep Tulsi plants healthy and productive. This work advances agricultural practices by utilising technological breakthroughs to address important challenges in plant health management and disease control, paving the way for a robust and environmentally friendly agriculture system.

This work focused on the detection and classification of infections in the Tulsi (Holy Basil) leaves using image processing and machine learning techniques. It elaborates on the workflow which begins with understanding the plant and culminates to designing an effective model for classification utilizing visuals of the plant. In chapter one, the reader is introduced to the Tulsi plant and its cultural, medicinal, and environmental value, setting the premise for the rest of the book. The chapter discusses the different types of Tulsi and some of the common infections that are seen in its leaves. The chapter also discusses the rationale behind the use of image classification techniques, goals, and overarching aims of the research. In chapter two, the related works are presented and previous studies on plant disease detection and Tulsi research is provided. The chapter discusses the importance of the problem and early detection in agriculture while analysing the existing literature critically. All these elements are important for framing the justification for the approach and innovations in the current study. In chapter three the research methodology is explained in detail. It encompasses the entire process of image acquisition and classification. The chapter discusses the methods applied for image pre-processing, feature extraction and selection, and training of the various classification models. This chapter explains how the research was conducted in a systematic approach.

The results achieved through the execution of the offered models are included in chapter 4. It evaluates the effectiveness of various techniques such as Support Vector Machines and stacking classifiers for infection detection on Tulsi leaves. Other discussions provided in the chapter include the models' accuracy and efficiency as well as how comprehensively the proposed solution met the aims of the research. In Chapter 5 final remarks of the work are presented and these are the major findings as well as the observations from the study. The results are practical and offer useful value to the domains of agriculture and computer vision. This section also suggests other unaddressed issues that need further refinement and investigation. The report has been well-organized with references and appendices that contain publications, conference participations, and other professional activities that accompanied the research.

Chapter 2
Literature Review

Ocimum Tenuiflorum, commonly known as Holy Basil or Tulsi, is believed to have originated in Southeast Asia. This herb has been widely used in Indian Medicinal practices for thousands of years. all parts of the herb have been claimed to be beneficial to the physical, moral, and spiritual aspects of a person. There are a lot of studies that have been done regarding the entire Tulsi plant and are in favour of supporting its therapeutic applications due to the medical benefits that Tulsi has. The range of ailments that Tulsi can help initiate recovery from includes respiratory, digestive, cardiac and dermatological disorders. Tulsi has many other benefits including maintaining good heart health, anti-ageing effects, reducing stress on kidneys, relieving headaches, preventing acne, reducing temperatures, improving oral and ocular health, curing respiratory tract diseases, and absorption of Vitamin K. As such Tulsi has an apparent role against a variety of diseases ranging from the mild ones to chronic conditions. Experimental investigations have shown that Tulsi herb has strong cytoprotective and immune-modulating characteristics as well as potent anti-cancer benefits. According to research from the Journal of Ayurveda & Integrative Medicine, Tulsi has anti-anxiety and anti-depression matrices which work similarly to standard anti-depressants. Also, the secondary products of the Tulsi plant are commercially significant, as they yield aroma substances present in the essential oil of its leaves and flowers.

The large scale of harvesting of tulsi herbage during the blooming season both nationally as well as internationally translates into roughly 10,000 kilograms of fresh herbage per hectare annually. Such herb which is directly exported produces the salient oil which ranges between 0.1 to 0.23 percent per Kilogram of fresh tulsi which is estimated to yield 10 to 20 Kilograms per hectare in essential oil .With irrigation increases the Tulsi yield production with oil content reaching an average of 40kg per hectare which translates to 20 tons in yield over the period of irrigation .Tulsi is available in oil, powder and capsule form. There are many uses for Tulsi in the world market such as personal care products, medicine, food and drinks.

2.1 Background and Importance of Tulsi

Tulsi, or Ocimum sanctum as it is scientifically called, is an herbaceous plant rather popular among Ayurveda practitioners and is often called as the 'herb for all reasons'. This is mainly due to the fact that it has a vast range of medicinal and beneficial properties. This sacred herb has been used for thousands of years in traditional medicine systems such as Ayurveda and is famous for its potency to heal numerous health issues and wellbeing in general. Mostly Tulsi has been known to aid in treating disorders or diseases that are on the more modern side such as lesions, stress, respiratory issues and most commonly, immune issues. Due to the strong ability of Tulsi to work as an adaptogen, it helps in coping with a wide range of issues, improving overall holistic health, and ensuring perfect synchrony between the mind, body and soul. Known to be a very strong adaptogen, Tulsi helps in coping with various physical stimuli, emotions and even environmental stress, alongside boosting immunity. Due to Tulsi's detoxifying traits, the herb allows the patient to get rid of various toxins and impurities getting rid of unwanted strain on the liver and allowing the entire body to detox. Tulsi also boosts the effectiveness of the immune system and throws the body into homeostasis balancing out the system and counteracting a wide array of stresses.

Due to its properties that combat germs that cause disease and infection, Tulsi Is useful as an adjunct in the prevention and management of many infections like skin and respiratory diseases. Ayurveda is said to call Tulsi "liquid yoga", which is said to calm the mind and promote relaxation thereby reducing anxiety, to achieve clear thinking and a stable emotion. Tulsi has a sacred significance in Indian culture and is usually found in Indian households and temples which symbolize health, safety, and purity.

Tulsi embodies the nurturing and curative aspects that nature has endowed upon humans, which appeals towards a more natural and multi-faceted approach towards one's health and wellness. All in all, the whole host of benefits and therapeutic uses offered by Tulsi underline the potential of this herb in enhancing disease prevention and holistic healing. Tulsi is understood to provide a range of benefits from increasing immunity levels and detoxifying the organism to relieving stress and enhancing mental acuity, and is still revered as a multi-functional plant that helps solve issues of the present day while keeping in mind age-old knowledge and the bounteous aspects of nature.

Different types of Basil or Tulsi are available across the world; they are classified into two categories: a) Holy basil and b) Mediterranean basil. Holy Basil, which is part of category one, possesses remarkable medicinal properties and finds application in Ayurveda. Mediterranean basil or simply basil is the common variety mostly grown in Asian, European, American and African continents and is utilized in cooking recipes. Table 1.1 provides an overview of the significance and diverse uses of the two types of the Tulsi plant that are most commonly used globally. The similar table featured below presents different other research works undertaken towards the prediction and classification of infection in the plants.

Tulsi, also known by its botanical name Tulsi (Ocimum sanctum), is a widely cultivated plant due to its significance in Indian Ayurvedic medicine. Other forms of Basil, also known as Mediterranean, on the other hand, have culinary uses adding taste and aroma to thèy typical encompass different sub species. It is common to use Holy Basil for various herbal formulations as it possesses limitless healing properties. Mediterranean and Holy Basil both have great value in the economy as they serve different purposes or form part of various cuisines and are also regarded as important in herbal therapy due to their stress alleviating effects, assist lung functioning and also support the immune system. Additionally, the two emerge from different continents whereby Mediterranean emerged from Asia, America, and Africa whereas Holy Basil has its origin from the Mediterranean regions. It is apparent that Holy Basil has a significant economic plantation demand globally which spans to Great Britain, the US, and Canada; countries which have included Tulsi in their herbal medicine regimes for Its overwhelming therapeutic qualities.

In the previous literature, developed background knowledge is established including automated diagnosis of basil plant diseases as few of them have concentrated on the prediction and classification of plant infection and although these techniques employ advanced algorithms such as algorithms based on images and algorithms based on deep learning. It is anticipated that improved algorithms capable of efficiently and accurately addressing the issues will be developed through this research, in which many authors are engaged in analysing leaf images using different parameters such as leaf texture, colour, and shape. This approach is multi-disciplinary involving agricultural engineering, computer science and data science with the potential of the future massive growth of the basil plant industry across the globe.

2.2 Tulsi leaf disease classification analysis

The authors Patel, R., et al., have reported on "Automated Detection and Classification of Diseases in Tulsi Leaves using Deep Learning Tools", Tulsi (Ocimum sanctum) leaves are a common drug that has a variety of uses in herbal medicine and other areas. This research proposed a deep learning approach for the automatic detection and classification of cloves (Ocimum sanctum) cysts, or leaf tunnels. The research seeks to apply deep learning techniques built on convolutional neural networks (CNNs) for image feature extraction and subsequent classification to increase the diseases detection rate. According to the authors S. Khan, et al, "Machine Learning Approaches for Disease Detection in Basil (Ocimum basilicum) Leaves: A Comparative Study." This comparative research benchmarked the capabilities of several machine learning methods towards illness detection in the sister species of Tulsi, Basil leaves. This research also tried to look into the efficacy of Naive Bayes, Decision Trees, Support Vector Machines, and Random Forest classifiers in the identification of prevalent leaf fungal diseases. Jain, P., et al. (2020). Deep learning-based Multiclass Disease classification of Ocimum sanctum (Tulsi) Leaves. This study aims to examine the applicability of deep learning models in multiclass disease classification in Tulsi leaves. The study looks at transfer learning using CNNs targeted at feature extraction and disease recognition, and it gives good results for disease detection. S. Mishra, et al., "Automated Identification of Tulsi Leaf Diseases using Texture Analysis and Machine Learning Techniques", This work discusses a texture based approach in conjunction with machine learning algorithms for the automatic identification of Tulsi leaf disease. The study seeks to texture feature extraction and classification through Support Vector Machines (SVM) of healthy and diseased leaves [12-14].

Singh, P et al. (2012): Disease management in Tulsi (Ocimum sanctum L.) – A Holy medicinal plant. this reference is on certain strategies used in the management of disease focusing on Tulsi (Ocimum sanctum) as a medicinal plant with some religious importance. I suppose the article may also include issues of plant disease such as the most common ailments of tulsi, how to prevent its occurrence, and how to treat it accordingly in order to sustain the plant and the medicinal properties that it offers [15]. Padalia, R.C., et al., in their publication, Notice that the paper 'Chemical Composition and Antibacterial Activity of Essential Oil Extracted From Various Sections of Ocimum Sanctum L.' that was published looks at the chemical composition of basic oils that were extracted from different varieties of Tulsi (Ocimum sanctum) herbs available in India, more so the chemical antibacterial aspects of these oils [16]. Singh, S.K., et al., Disease management in Tulsi (Ocimum sanctum L.) The disease management aspects of Tulsi were likely highlighted in this article. The article is probably a summary and explains types of common diseases affecting Tulsi, their symptoms, and integrated control measures recommended [17]. Pandey, A. K., et al, Treatise on chemistry and bioactivities from aromatic compounds from various Ocimum species [18]. This in all these review the subject on aromatherapy using essential oils obtained from various species of Ocimum including Tulsi hadbeen elaborately documented. Singh, S., Singh, A., et al., Several essential oils exhibited antifungal action against fungal infections of Ocimum sanctum L. and their phytochemical analysis. This article likely explored the antifungal properties of essential oils derived since Tulsi (Ocimum sanctum) that investigates their effectiveness against fungal pathogens that affect the plant [19].

Bharti, P., et al., Detection and management techniques for Tulsi leaf spot disease (Ocimum sanctum L.). This publication probably focuses on the detection methods and management strategies for leaf spot diseases specifically affecting Tulsi plants [20].
Tandon, S., et al., Leaf spot disease in Tulsi (Ocimum sanctum L.) and its treatment. This article likely discusses the symptoms, causal agents, and management options for leaf spot disease in Tulsi. Verma, R.S., et al., The chemical makeup of Ocimum sanctum L. oil of essential and its efficacy against storage fungi [21]. This publication may look into the chemical substance of Tulsi essential oil and it's likely the use as a fungicidal agent against storage fungi [22]. Sah, A.N., et al., Fungi related with Tulsi leaf spot disease (Ocimum sanctum L.). This article likely identifies and discusses fungal pathogens associated with leaf spot disease in Tulsi plants [23-24].

The table 2.1 below summarizes extensive literature from 2004 to 2024, emphasizing salient research contributions applicable to plant disease detection—namely for Tulsi and other medicinal leaves. It encompasses author names, publication sources, year of publication, status of being indexed (Scopus, SCI, etc.), and primary findings or conclusions. The table is a starting reference point in the literature review, aiding in:

i. Following the development of methodologies (from classical ML to ensemble models)
ii. Describe research gaps (such as sparse attention to Tulsi)
iii. Justify the novelty and methodological decisions of the current study.

Table 2.1. Classifier assessment with the literature on tulsi

Author Name	Detail of the journal/ Book / Book chapter/ website link	Year of Publication	Indexing of journal (Scopus/ SCI index etc.)	Main findings or conclusion relevant to proposed research work
Prakash et al. [25]	Indian Journal of Physiology and Pharmacology- Scientific Scholar	2004	SCOPUS, Web of Science IF0.302	Tulsi is suggested to possess many useful actions related to health issues.
Anami et al. [26]	International Journal of Computer Applications- Foundation of Computer Science	2010	SCOPUS EBSCO IF-0.835	SVM and the round basis exact fit neural network (RBENN) were utilized for training and testing image samples belonging to three classes. The overall classification accuracy of chroma and edge texture parameters reaches 74% and 80%, respectively. Combining colour and texture features improves accuracy to 90%.
Pattnayak et al. [27]	Pharmacogn Rev	2010	Index Copernicus, Indian Science Abstracts	Endorse the use of such plants in treating human and animal diseases, as well as draw attention to the ethnobotanical perspective as an alternative source of bioactive compounds.
Valliammal et al. [28]	Computer Science & Engineering: An International Journal- ISROSET	2011	EBSCO IF-3.802	PIS Image Classification Hence, it is important to explain leaf veins in order to be able to separate living plants. All these aspects are of particular interest in this undertaking i will describe as the main goal of this work is to allow users in the field to take photographs of a newly discovered plant, insert it into a laptop based system and have the system itself classify the species and display pictures of closest specimens within a few seconds.

Tewari et al. [29]	International Journal of Ayurvedic Medicine	2012	WOS IF-1.56	The nutritious and medical features of the tulsi herb in its natural state were the result of an effective blend of various relevant phytochemicals; thus, the comprehensive impact of Tulsi cannot be adequately recreated using single substances or extracts.
Jabal et al. [30]	Journal of Computer Science- Science Publications	2013	SCOPUS IF-0.169	Examine the types of leaf attributes that must be extracted, the external factors that must be addressed before commencing the extraction process, and the various extracts and method of classification that can be used for crop classification and identification.
Kumar et al. [31]	International Journal of Agronomy and Plant Production- Hindawi	2013	WOS IF-0.564	Economic Benefits of Tulsi
Larese et al. [32]	Pattern Recognition- Elsevier	2013	SCI, WOS, SCOPUS IF-7.74	Hit/Miss Transform Adaptive Thresholding SVM, PDA, Random Forests, PDA- 98%
Nilamudeen et al. [33]	Progressive Horticulture- Indian Society of Horticultural Research and Development	2013	Horticultural Abstracts (CABI), Indian Science Abstracts and Indian Citation Index 3.25 NAAS	Cultivating tulsi in farming systems with susceptible crops

Annathi et al. [34]	Journal of NanoScience & NanoTechnology- World Scientific	2014	SCIE, WOS IF-1.134	Recognition Rate using SPSS software Employing grayscale, canny edge detector, morphological and neural network algorithms individually, a powerful single leaf image extraction computer package was developed.
Cohen et al. [35]	Journal of Ayurveda and Integrative Medicine- Medknow Publiactions	2014	SCOPUS, WOS IF-0.72	Tulsi-Herb for all reasons- medicinal uses
Dubey et al. [36]	Journal of Intelligent Systems- Walter de Gruyter GmbH	2015	SCOPUS, WOS IF-2.050	Fruit and vegetable disease classification-SVM & kNN comparison
Gomathi et al. [37]	Research Journal of Pharmaceutical, Biological and Chemical Science	2015	SCOPUS IF-0.221	Validation performance and the efficiency obtained by both the neural networks were compared and the modified LM algorithm have better recognition accuracy
Vijaysree et al. [38]	Middle-East Journal of Scientific Research	2015	SCOPUS, ROAD IF-0.688	GLCM features like energy, uniformity of variance a correlation and comparison have just been evaluated for tulsi leaves. Eleven attributes were gathered, and the leaves were chosen. The distinction outcome has been calculated using the information found characteristics and exhibited.
Wang et al. [39]	Neural Comput & Applic- Springer	2015	SCOPUS IF-1.492	The entropy sequence produced by PCNN is used as a main feature for classification, with some form and texture features serving as assistance features.
Yun et al. [40]	Int J Agric & Biol Eng	2015	SCOPUS, WOS, SCIE IF-1.731	The experimental results on three Cucumber diseased leaf image databases, i.e. mildew, downy, anthracnose and blight, and

				anthracnose, showed that crop diseases can be effectively acknowledged by the combined use of leaf image processing technological advances, disease meteorological information, and the accurate identification rate exceeded than 90%, indicating that the PNNs that were classifier developed on the infection characteristic coefficients.
Sladojevic et al. [41]	Computational Intelligence and Neuroscience- Hindawi	2016	SCOPUS, WOS, SCIE IF-3.633	Caffe is a multiplying model developed by the Berkley Vision and Learning Centre, and was used to train the very deep CNN developed. The precision values of the tests on the established model ranged from 91% to 98%, with the average over independent class tests standing at 96.3%.
Tsaftaris et al. [42]	Trends in Plant Science- Elsevier	2016	IF-11.911	Although machine learning can employ for characterization, the segmentation procedure must be performed to design suitable attributes.
Zhou et al. [43]	Applied Sciences- MDPI	2019	SCIE, WOS IF-2.679	Colour, Edge Detection Based Algorithm Random Forest, Mean Intersection Over Union 0.81 hue saturation value hue saturation value.
Muneer et al. [44]	IEEE Access	2020	SCIE, SCOPUS, WOS IF-3.367	The SVM result exhibits a recognition accuracy of 74.63% but the DLNN result exhibits a recognition accuracy of 93% both for the emulated model and the mobile application created.
Ganantra et al. [45]	International Journal on Emerging Technologies- Research Trend	2020	SCOPUS IF-3.1	Random Forest, SVM, kNN, A73.88% FScore-71.98% Recall-72.90% Precision-72.88%

Gautam et al. [46]	Journal of Ayurveda and Integrative Medicine- Medknow Publications	2020	SCOPUS, WOS IF-1.505	Ayurveda & Ayush Kwath importance in terms of Tulsi leaf
Ji et al. [47]	Information Processing in Agriculture- Elsevier BV	2020	SCOPUS IF-6.41	Of the models analysed, the United Model exhibited the best average precision of 99.05%, recall of 98.88% and F1-score of 98.96%. VGGNet achieved the lowest scores for all the metrics: precision of 94.16%, F1-score of 94.40%, and recall of 93.32%.
Malvankar et al. [48]	International Journal for Research in Applied Science & Engineering Technology	2020	ISI, Index Copernicus IF-1.451	The Naive Bayes Classifier has an accuracy of 65%, while K Nearest Neighbour has an estimated accuracy of 80%.
Mishra et al. [49]	International Journal of Pharmacy and Biological Sciences (IJPBS)- Jp Research Publications	2020	Index Copernicus IF0.888	The conservation of Tulsi plant species to make them constant in our lives and to identify the active bio-compounds
Monzano et al. [50]	International Journal of Environmental Research and Public Health-MDPI	2020	SCOPUS, WOS IF-3.390	The study identifies gaps in agricultural research on medicinal plant domestication, production, and genetic/biotechnological development.
Ngugi et al. [52]	Information Processing in Agriculture- Elsevier	2020	SCOPUS IF-6.41	SVM, kNN, NB, MLC, DT, PCA, RF ANN 98.5%-AA
Saleem et al. [53]	Plants-MDPI	2020	Scopus, SCIE (WOS) IF-3.935	The Xception architecture optimised with the Adam optimizer achieved the greatest accuracy in validation and The score for F1 of 99.81% and 0.9978, correspondingly.

Sharma et al. [54]	Information Processing in Agriculture- Elsevier BV	2020	SCOPUS IF-6.41	S-CNN obtained a prediction accuracy of 98.6%, versus 42.3% as previously estimated for F-CNN.
Sun et al. [55]	IEEE Access	2020	SCIE, SCOPUS, WOS IF-3.367	SSD-Single Shot multi-box detector Algorithm, Caffe framework RPN Network 91.83% - Highest mean average precision
Wang et al. [56]	Information Processing in Agriculture- Elsevier BV	2020	SCOPUS IF-6.41	The research' findings confirm that when the ideal mist parameters were used during different development phases, the average optimal index throughout all three growth stages was 42.29%, which was up to 62.24% higher than what was attained without differentiating development periods.
Yang et al. [57]	IEEE Access	2020	SCIE, SCOPUS, WOS IF-3.367	APS-DCCNN Average Accuracy-96.37%
Zeng et al. [58]	IEEE Access	2020	SCIE, SCOPUS, WOS IF-3.367	Deep Convolutional Generative Adversial Networks Average Accuracy-92.6%
Zhang et al. [59]	Applied Sciences- MDPI	2020	SCIE, WOS IF-2.679	To obtain shape features of an object, a method called BOF_SC is adopted that works on a bag of contours segment. We further use BOF_DP and LDA to extract features describing the text which reduces the dimensionality of the feature space. Finally, both characteristics were classified using a support vector machine,

				which is linear in nature, and the above approach outperforms existing algorithms on standard leaf datasets.
Zhang et al. [60]	Network Pharmacology- International Academy of Ecology and Environmental Science	2020	EBSCO	Created a sophisticated MATLAB-based application (imageProcAnal Version 1.0) for processing images and analysis
Ji et al. [61]	Information Processing in Agriculture- Elsevier BV	2021	SCOPUS IF-6.41	In particular, this method identifies the maize tassels accurately in the field images, and is more consistent with time with methods not previously reported. The precision, recall, and F1 measure scores were 86.30%, 91.44%, and 88.36%, respectively.
Naem et al. [62]	Agronomy-MDPI	2021	Scopus, SCIE (WOS), AGRICOLA, AGRIS IF-3.34	MLP, LB, B, RF, SL 99.10% Tulsi-AA
Tavakoli et al. [63]	Computers and Electronics in Agriculture- Elsevier	2021	SCIE, SCOPUS IF-5.565	CNN 95.86% -Level 1 91.37%- Level 2 86.87%-Level 3 MCA Classification Accuracy) (Mean
Zhou et al. [64]	IEEE Access	2021	SCIE, SCOPUS, WOS IF-3.367	RRDN Deep CNN AA-95%

Vishnoi et al. [65]	Journal of Plant Diseases and Protection- Springer International Publishing AG	2021	SCI, IF-1.652	The use of computer vision and soft computing techniques for automated plant disease detection, focusing on leaf image analysis and highlighting the merits, demerits, and effectiveness of modern feature extraction methods across various crop categories.
Vishnoi et al. [66]	Multimedia Tools and Applications	2021	IF-2.517	wide range of image qualities, including colour, texture, and shape, for diverse illnesses
Madhavan et al. [67]	Computers, Materials & Continua	2021	SCI, IF-3.772	SVM MATLAB 98.39% accuracy
Ashwinkumar et al. [68]	Materials Today: Proceedings, Science Direct	2022	SCOPUS, IF-1.8	The OMNCNN model precision was 0.985 while accuracy was 0.987 recall score 0.9892 F score 0.985 and kappa score 0.985.
Bhujel eta al. [69]	Agriculture, MDPI	2022	Scopus, SCIE, IF-3.049	CNN, average accuracy 99.69%
Ravi et al. [70]	Expert Systems, John Wiley & Sons Ltd	2022	Scopus, IF2.587	CNN, Cassava, logistic regression
Subramanian et al. [71]	Neural Computing and Applications, Springer	2022	Scopus, IF5.606	Deep learning model accuracy 93%
Singh et al. [72]	Multimedia Tools and Applications	2022	Scopus, IF2.757	CNN 99.16%
Khan et al. [73]	Applied Sciences, MDPI	2022	Scopus, SCIE, IF-2.679	Deep Entropy 98.4% accuracy
Kaur et al. [74]	Sensors, MDPI	2022	Scopus, SCIE, IF-3.576	CNN 98.7% accuracy
Sathiya et al. [104]	International Journal of intelligent Systems and Applications in Engineering	2023	Scopus, IF 1.03	CNN- Inception V3 model 77.55% accuracy

Kaur et al. [110]	Multimedia Tools and Applications	2023	Scopus, WOS(SCI), IF-3.0	DeepLabV3+ 95% accuracy
Alom et al. [111]	Journal of Agriculture and Food Research	2023	WOS. IF-5.85	ResNet50 96.7% accuracy
Meenakshi et al. [112]	Wireless Personal Communications	2023	Scopus	Logistic Regression 98.56% accuracy
Mahum et al. [113]	Human and Ecological Risk Assessment: An International Journal	2023	Scopus, SCIE, IF-3.56	DenseNet-201 CNN 97.8% accuracy
Yashwant et al. [115]	IEEE	2023	Scopus, WOS	SVM 95.2% accuracy
Muthunayagam et al. [123]	Journal of Phytopathology	2024	SCIE, IF-2.1	Weighted Ensemble 96.3% accuracy
Sharma et al. [124]	Advances in Signal Processing and Communication Engineering	2024	Scopus	CNN 92.7% accuracy
Kumar et al. [125]	Complex & Intelligent Systems	2024	Scopus, SCIE, IF-3.6	CNN 99.24% accuracy
Kumar et al. [126]	ICMET 2023	2024	Scopus	CNN
Lakshmanrao et al. [127]	IEEE	2024	Scopus	Transfer Learning, Hybrid classifiers
Deshmukh et al. [128]	Indian Journal of Science and Technology	2024	Scopus	MLP, CNN 99.01%, 98.3%

2.3 Research gap identification

Identifying research gaps in the proposed AI-driven image processing model for diagnosing plant illnesses in Tulsi herb plants can provide useful insights for further refining and improving the approach. Here were the major research gaps that can be identified:

i. Dataset Representation and Diversity: Availability and diversity of the datasets required for image analysis and machine learning methods were one of the primary research needs. The quality, quantity as well as the representation of the data which specifies the amount that is used for training these predictive models were very crucial. Studies should focus on developing composite data sets that consist of healthy and diseased Tulsi plant leaves collected from different environmental and disease severity conditions.
ii. Generalization and Adaptability: Another important issue is how the verified model generalizes and whether it is adaptable to other geographical regions, portions, and types of Tulsi plants. Particularly the research backed the phylogenetic model aims to how well each model performs in comparison to the rest when exposed to unfamiliar environmental conditions and need changes at a certain point to remain efficient.
iii. Instant Performance and Resource Usage: It's necessary to determine how the model handles the resources entrusted to it, while at the same time assessing how well it performs, such as Before being deployed in low-resource environments such as rural farming areas, the system must pass such testing. There is the need for further work in developing accurate low-power models suitable for mobile devices concentrating on the compute and memory requirements optimization.
iv. Enhanced explainable artificial intelligence: Who could benefit from a better comprehension of which criteria drove to defining a certain diagnosis based on the images of the plant leaves includes farmers and other stakeholders. Greater efforts should be devoted to finding highly visual or textual formats which will represent or underlie the algorithm's reasoning processes and decision making processes.
v. Artificial Intelligence model and its diagnosis: The model and its diagnosis were key to the inclusion with the practice. It is also important to examine how farmers would integrate this technology considering usability, user interface, and modes of interaction.
vi. Control over time and treatment management: a question that should be addressed is that of treatment management whereby control over the time factor is of utmost importance. The study should move to a diagonal level focusing on not only the primary disease diagnosis but also predicting the future spread of diseases, their evolution and possible treatment depending on the state of the plant at a given time.
vii. Validation and Field Testing: Finally, the resulting model has to be thoroughly validated, and applied in field experiments in real conditions. It is necessary to ascertain the framework's validity, strength and adaptability in real world farming practices with the farmers and agricultural scientist thorough participatory feedback and validation exercises.

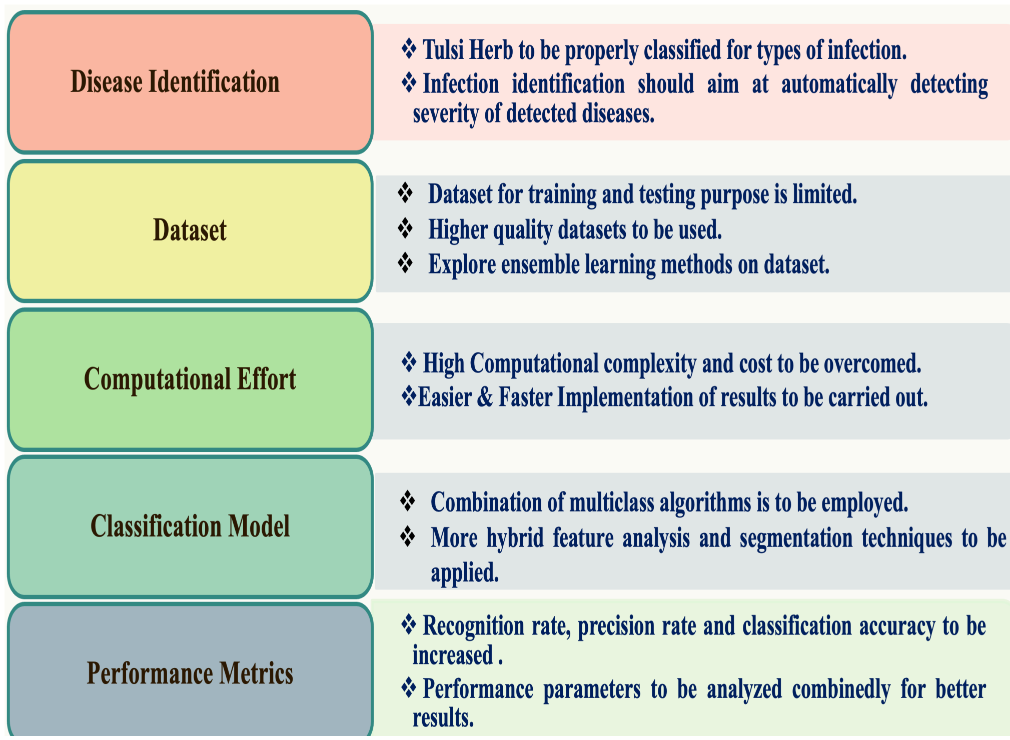


Figure 2.1. Research Gap overview

This chapter examined earlier works pertaining to plant disease detection using machine learning and deep learning approaches, concentrating on medicinal plants. Special emphasis was placed on Tulsi (*Ocimum sanctum*) because it is underrepresented in literature. Earlier works mostly relied on CNN-based architectures along with transfer learning models like VGG16, ResNet50, and InceptionV3 to classify plant leaf diseases in crops such as tomatoes, potatoes, and basil. Even though these models achieved considerable accuracy, they were mostly focused on non-medicinal plants. Research works focusing on medicinal plants did not possess sufficient scalability and generalizability. Furthermore, most works were based on a single classifier and did not attempt ensemble learning. Studies focusing on the classification of Tulsi leaves were limited and many available were poorly supported by datasets and lacked validation in real-world scenarios.

The current study tried to fill the gap by designing an ensemble stacking classifier model that coupled multiple base learners with the same task to improve prediction accuracy. This strengthened classification performance, reduced overfitting, and enhanced the reliability of disease detection on Tulsi leaves in changing environmental conditions. The rationale for the use of ensemble models was to address the shortcomings of the approaches that relied solely on a single model.

Chapter 3

Methodology

3.1 Research Methodology

The research technique for developing an illness prediction model in plants using a stacking classifier model begins with the crucial phase of data collection and preprocessing. This entails accumulating a diverse set of plant characteristics, environmental variables, and disease classifications. Data cleaning processes were used to correct missing values, outliers, and inconsistencies in datasets, ensuring their integrity and accuracy. Feature engineering is critical in identifying useful features from raw data or applying domain-specific information to improve the prediction capabilities of the model. To provide unbiased assessment of models and confirmation, the image dataset is formed into three sets: training images, verification of images, and testing of images.

Table 3.1. Tools/ Instruments used

Objective	Software	Source Links	Analysis undertaken
To identify and categorize types of infections in Tulsi Herb.	CV Studio Open CV Python	Open Source https://vision.skills.network/ Python Microsoft visual studio code Kaggle	Identification & Categorization and create a csv file and dataset collection in Kaggle
To apply image processing for classification of infection in leaves.	CV Studio Open CV Python	Open Source https://vision.skills.network/ Python Microsoft visual studio code Kaggle	To find and implement the novel methods for image pre-processing
To develop the training model for prediction of infection in leaves.	CV Studio Open CV Python	Open Source https://vision.skills.network/ Python Microsoft visual studio code Kaggle	Upload, Annotate, Training dataset
Testing of the designed prediction model.	CV Studio Open CV Python	Open Source https://vision.skills.network/ Python Microsoft visual studio code Kaggle	Test different models and find the best accuracy model

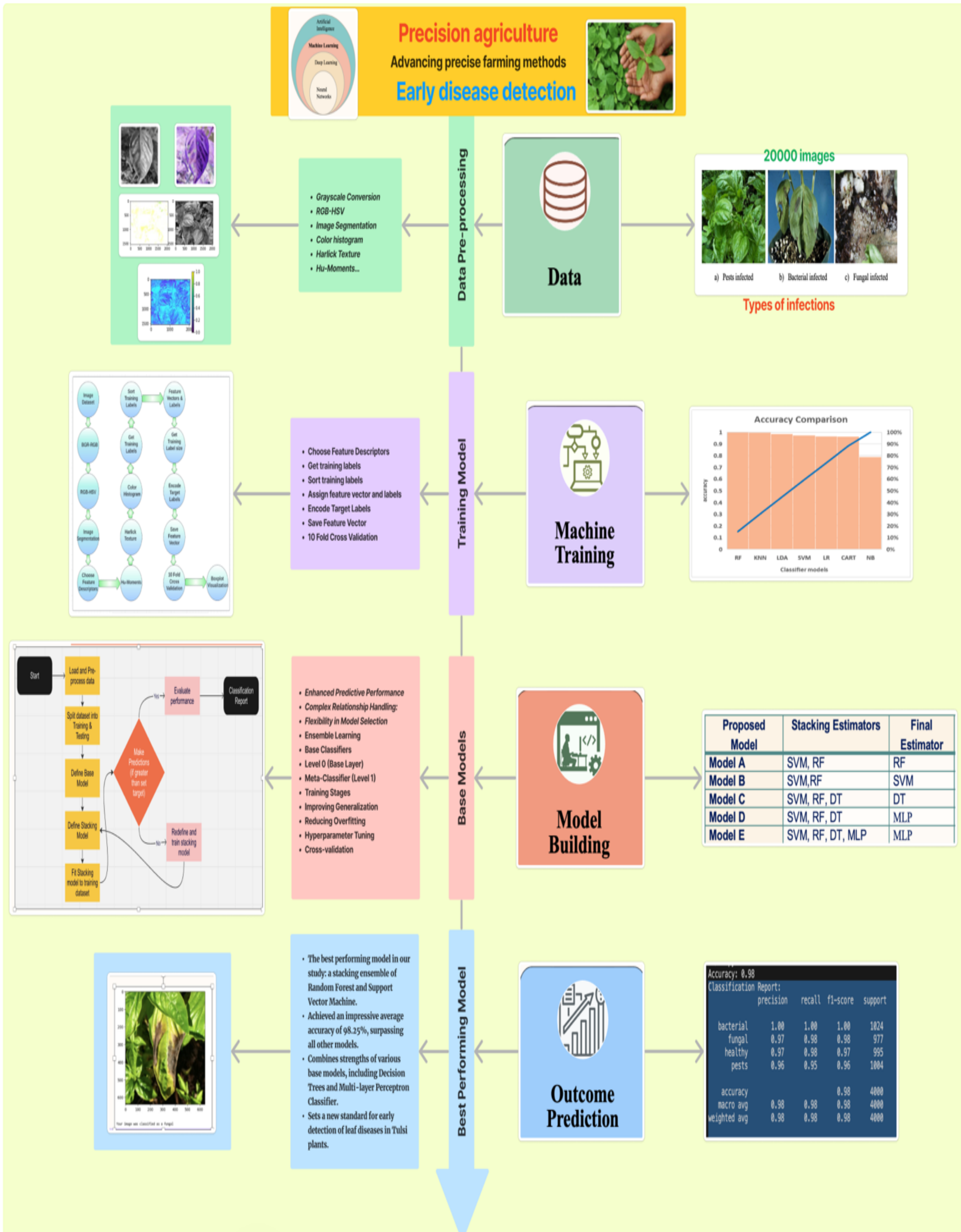


Figure 3.1. Methodology for disease detection

3.1.1 Dataset Creation

The dataset used in this study, named "Tulsi Leaf Train and Test Dataset", was developed and shared on Kaggle in 2023. It was made to fill the gap of publicly available, labelled image datasets centred around Ocimum sanctum (Tulsi) leaf disease classification. The dataset consists of a total of 20,000 images, divided equally into four different classes:

i. Bacterial infections – 4,000 images
ii. Fungal infections – 4,000 images
iii. Pest damage – 4,000 images
iv. Healthy leaves – 4,000 images

Original base images were gathered and curated based on visual references from the PlantVillage platform (https://plantvillage.psu.edu/topics/basil/infos), a renowned source for plant and agricultural disease datasets. The dataset creation process included meticulous manual selection and classification of leaf images to simulate real-world visual symptoms of every condition. To improve model performance and generalizability, the dataset was subject to extensive image augmentation, including: Rotation, Flipping, Scaling, Brightness and contrast variation, Zooming and cropping. These enhancement methods contributed to boosting intra-class variability and mimicking real-world environmental conditions, thereby providing a stronger training dataset for machine learning algorithms. The resultant dataset was divided into training and test subsets with stratified sampling to preserve class balance. It is publicly uploaded and shared on Kaggle:

i. Title: Tulsi Leaf Train and Test Dataset
ii. Author: Manjot Kaur
iii. URL: https://www.kaggle.com/dsv/6493843
iv. DOI: 10.34740/KAGGLE/DSV/6493843

This proprietary dataset is a useful tool for the educational community and forms the basis of training and testing of the ensemble classification models introduced in this study.

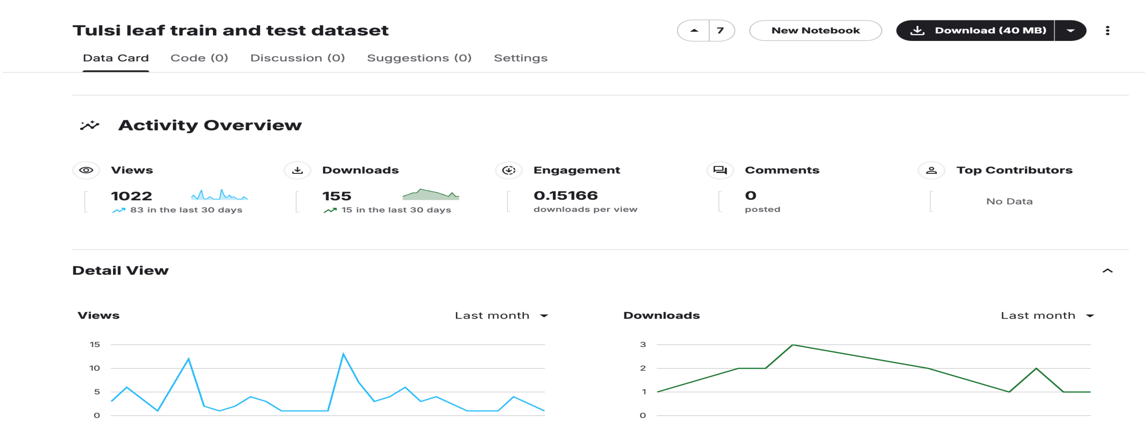


Figure 3.2. Kaggle dataset download statistics

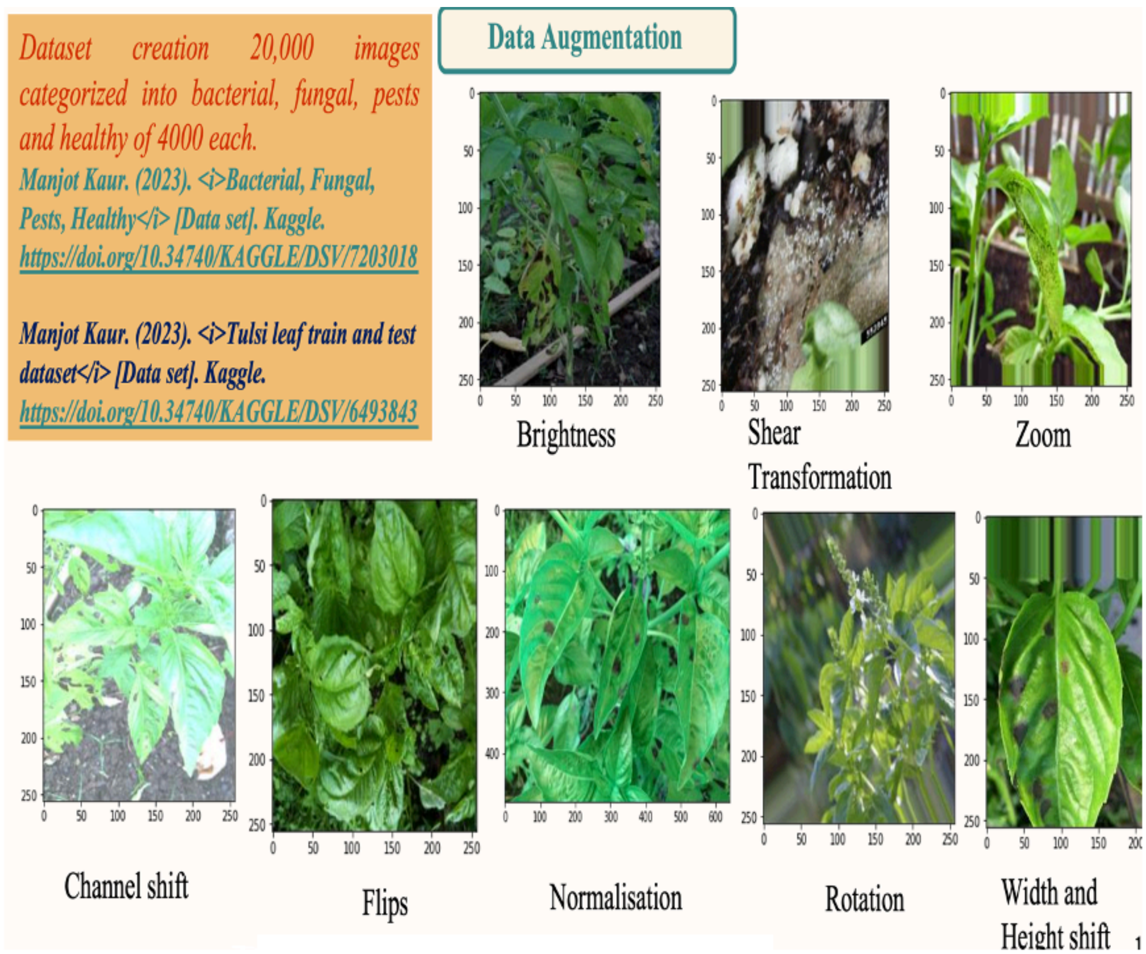


Figure 3.3. Data augmentation sample images

3.2 Pre-processing Techniques

Image preprocessing techniques and strategies were imperative in enhancing the quality of images employed to classify diseases, more so, in the medical field. These strategies assist in machine processing/deep learning algorithms boosting accuracy and efficiency of the images [65-67]. Below were some common images preprocessing strategies that aid in the classification of the diseases:

a. Image Data Acquisition: DICOM: Medical images are mostly sent to the DICOM format which may contain additional information in the formalities' of medicine and medical imaging. It is essential that these formats are captured accurately and converted into a usable format like JPEG or PNG prior to use.

b. Image Resizing: Standardization involves resizing images to a consistent size for neural networks and algorithms. The most common sizes are 224x224, 256x256, or 512x512 pixels, depending on the model.
c. Normalization: Pixel Value Scaling: Normalizing pixel values (e.g., scaling them from 0 to 1) accelerates the convergence of optimization techniques and improves model performance. Mean Subtraction: Subtracting the dataset's mean pixel value from each pixel helps to centre the data and improve model training.
d. Security and Scope: Simply put, security and scope involves identifying who can access the server and what external authorities can the server communicate with. Data Scrambling: FPGAs have gained popularity for both fast machine busying and data scrambling functions:

    i. Flipping: The process of flipping images horizontally or vertically.
    ii. Rotation is the process of rotating images at different angles (for example, 90° and 180°).
    iii. Scaling involves randomly zooming in or out.
    iv. Translation is the process of shifting images along the x or y axes.
    v. Color jittering is the slight alteration of brightness, contrast, saturation, and hue.
e. Image Filtering and Smoothing Gaussian Blur: This approach decreases noise and detail in medical images, improving overall structure.
f. Median Filtering: This proves to be useful for the elimination of salt and pepper noise while ensuring edges are maintained which is fundamentally important for the perfect determination of the illness.
g. Edge Detection Techniques: Use of edge detection techniques (like Sobel, Canny) enables important features to be highlighted in medical images facilitating classifiers identification of those features.
h. Thresholding
    i. Binary Thresholding: transforms greyscale pictures to their binary components that may help in separating areas of interest for instance healthy and sick tissues.
    ii. Adaptive Thresholding: This technique uses a threshold which changes along the image so that it is applicable in different lighting conditions.
i. Morphological Operations

    i. Erosion and dilation: These are useful procedures that help in shaping the structures that are visible in an image and this improves feature extraction and classification.
    ii. Opening and Closing: These procedures can be employed to suppress very small noises or to fill small voids in areas of interest.
j. Histogram Equalization

i. Image Contrast improvement: This modifies the contrast of an image by changing the intensity levels of its pixels. It is most useful in cases of images that are overly bright or too dark in order to enhance visibility of their features.
ii. CLAHE is a method in which small parts of images are equalized while preserving contrast in those areas making them distinct.

k. Colour space transformation is an aspect that can reduce the amount of information while retaining key elements for instance when turning a colour image into greyscale which can also be handy in cases where colour isn't necessarily a dominant feature. Colour normalisation: Integrating colours in regards to images can be vital when dealing with images from different machines.
l. Image segmentation one of the most basic approaches is region based image segmentation that simply – removes background in order to preserve certain areas of interests such as tumours, or lesions. Another technique called semantic segmentation works by partitioning a given image by categorizing each pixel with the use of convolutional neural networks such as U-Net so that it is easier to see certain aspects of the body.
m. Feature identification using definition extract key features, in order to extract high resolution features from images apply methods such as scale invariant feature transform sift, histogram of orientated gradients hog, or models such as vgg16 or ResNet.

Classifying diseases through the use of medicinal images classification is a crucial part of the process, the image that is fed into the epoch as well as the model greatly determine the classification accuracy. Through utilizing such methods one is able to greatly improve the quality of data being fed into the ML models which in turn improves the accuracy of the classification and the diagnosis. The type of images being scrutinized, as well as the elements of the particular disease in question, are determining factors in determining the correct pre-processing strategies to deploy. Image properties are crucial in the above-mentioned tasks, if we consider the lost images and attempts to detect and classify plant diseases for instance. Such features include a variety of visual features extracted from plant images in order to differentiate healthy plants from those that are diseased or infested with pests. Such features are important as they provide information on the surface characteristics of leaves or plant organs .Texture features capture space changes of an arrangement of pixels in a given area. Such attributes are important as they reveal the arrangement of leaves or any other plant organ .Colour attributes can tell about the distribution and quantity of colours in a picture, such abnormalities may signal specific aspects of disease such patterns such as colour changes (discoloration) and lesions . Shape features are the outlines of the plants structures – leaf forms of various sizes and contours for instance which are also changed when diseases attack plants.

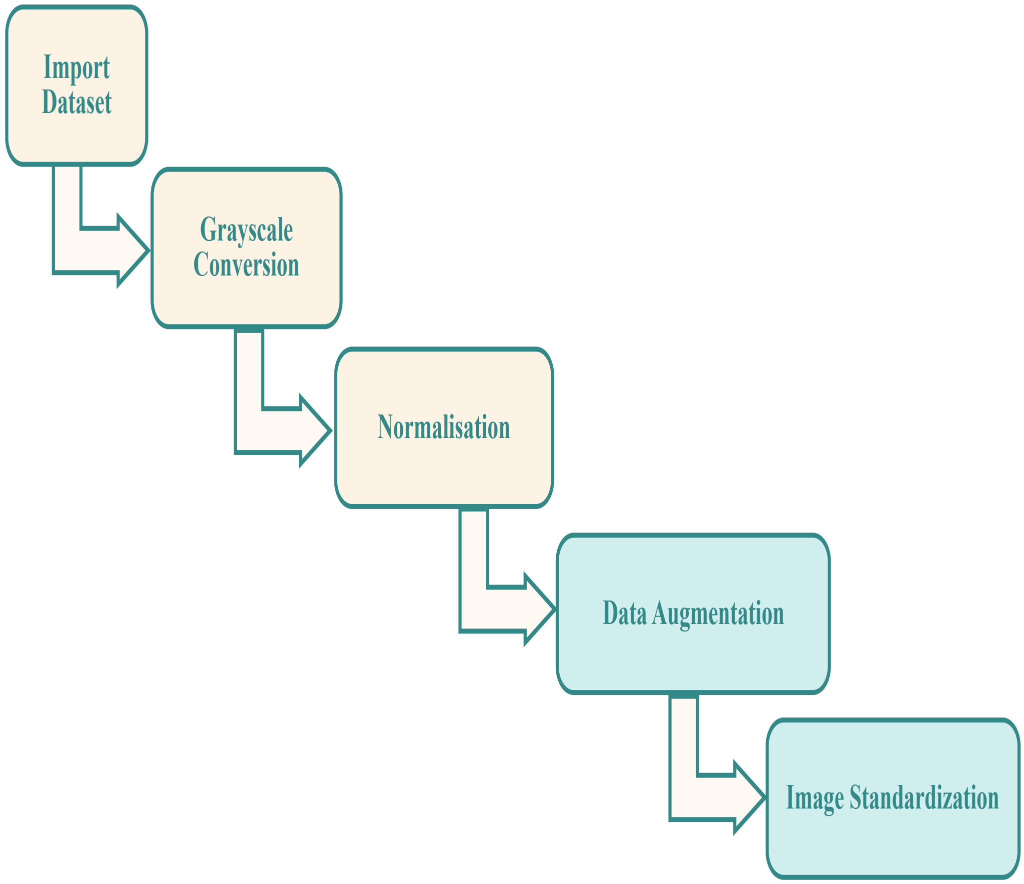


Figure 3.4. Methodology for image pre-processing

Table 3.2. Image Pre-Processing Techniques

Pre-Processing Technique	Advantages
Gabor Filter	Uniqueness, Specific to period & Scale, Fast FFT, Quantification of Stationary Signals
Adaptive Median Filter	Level non repulsive noise, Retain threshold information in high density impulse noises
Morphological Operations	Revealing of lesions of various dimensions and shapes
Mean Filter/Average Filter	Reduce the variance and easy to carry out
Image Normalization	Consistent among different pages in dataset, prevent printing problems
Histogram Equalization	Simple & Enhance Image Contrast
Weighted Median filter	Salt & Pepper noise Removal
Weiner Filter	MSE Minimization, Degradation Function & Noise Handling

Table 3.3. Image segmentation techniques

Algorithms	Technique
Pixel Based	Region Homogeneity Criterion Easy to Define
Edge Based	Computationally Fast, no prior Image content is required
Region Based	Easier to classify & implement
Deformable Model Based	Generate surface from Images
Texture Based	Complex problems using training data
ANN Based	Parallel nature of Neural net
Fuzzy Theory Based	Feature Based & Spatial Information
Genetic Algorithm Based	Contrast Enhancement, Solve Complex Optimization Problems

Visual inputs obtained from photographs can be utilized to train machine learning models along with image processing algorithms to diagnose, detect and categorize plant diseases. This approach of combining image features semantics with computational methods provides a strong basis for the establishment of reliable and effective disease diagnosis systems in the agricultural sphere, which paves the way for preventive and control measures to be undertaken in a timely manner and thus improve the overall agricultural output [66-68].

Table 3.4. Image Preprocessing methods

Methods	LBP,ORB	Spectra (SIFT,SURF)	Basic Space FFT Code Books	Blob Metrics
Illumination Adjustments	√	√	√	√
Blur & Focus Improvements	√	√	√	√
Filtering & Vibration Reduction	√	√	√	√
Threshold segment				√
Edge adjustments		√		√

Morphology				√
Segmentation Process				√
Region Filtering and Processing		√	√	√
Point Computation		√		√
Mathematics and Statistics Computing		√		√
Color Alignment Adjustment		√	√	√

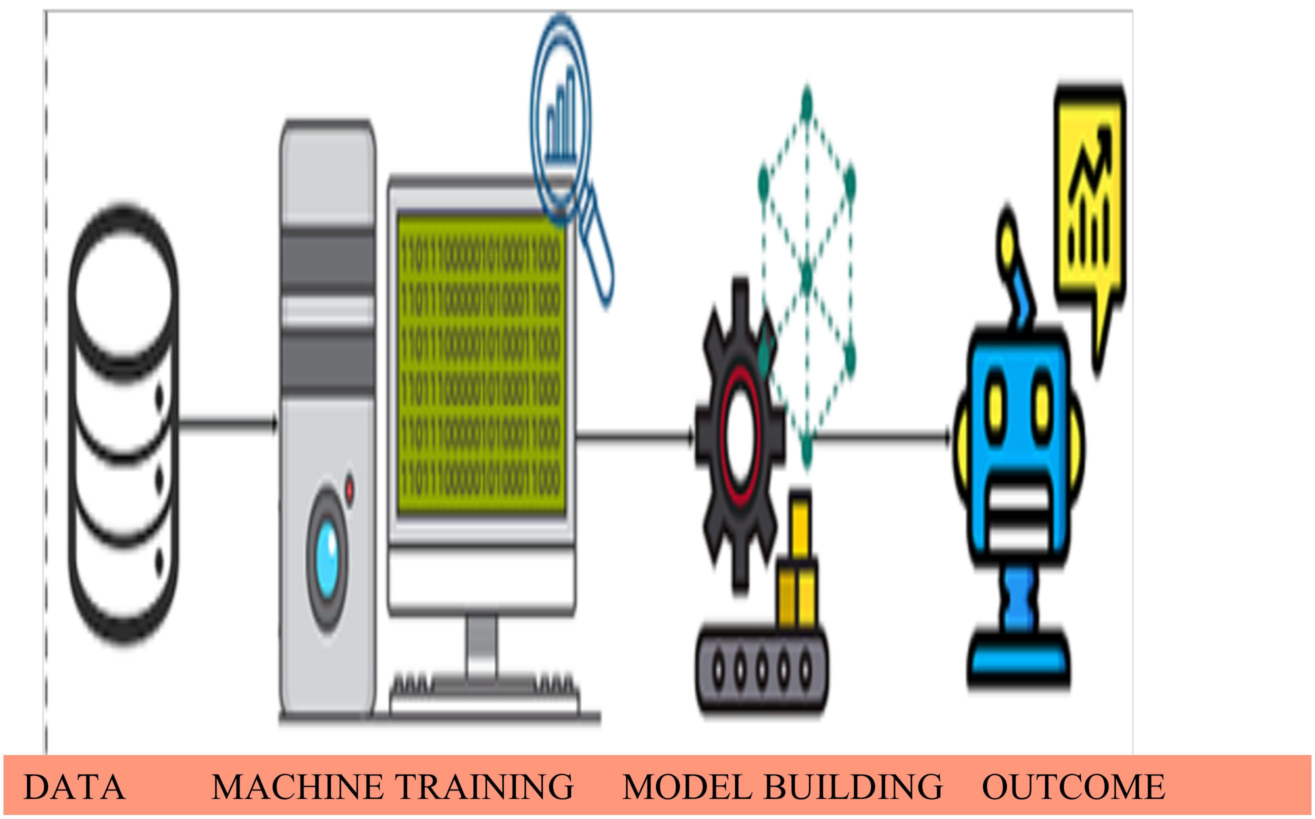


Figure 3.5. Prediction process flow

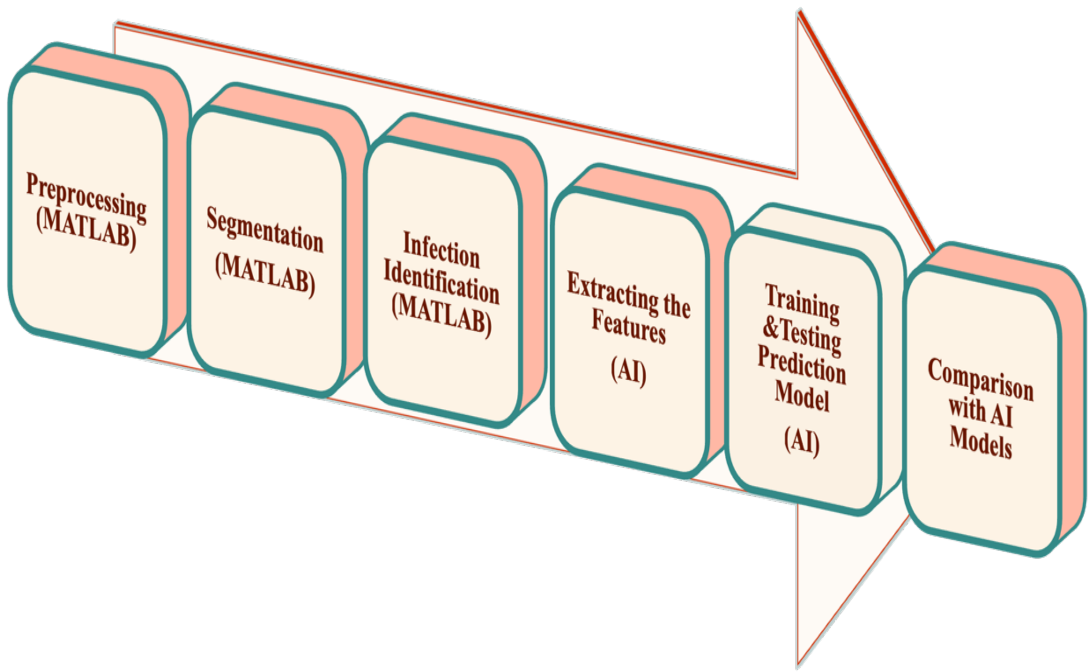


Figure 3.6. Proposed model process flow

3.3 Feature Extraction:

Signal processing and data analysis in agrophysics among other disciplines has improved with the use of feature extraction methods. These were techniques developed and employed to extract the pertinent features from raw data which makes learning algorithms more effective in discovering and analysing patterns. As in other agricultural uses such as plant diseases detection, agricultural uses of such key feature extraction approaches include historical significance extends back to times . First time domain analysis is concerned with the examination of signals over time: the amplitude, duration and frequency of events.
This approach allows the analysis of temporal variation as well as dynamics in signals related to the plant or the environmental parameters. Frequency domain analysis, on the other hand, means transforming the data onto the frequency domain by means of Fourier techniques. This method reveals the dispersion of signal energy at several frequencies. This information could be helpful in detecting some distinct features or abnormalities related to the disease of the plants.

Time-frequency domain techniques are hybrid techniques that cover both temporal and frequency domains allowing the signals to be analysed with respect to their time shifts while considering the frequency aspects. This category includes the use of techniques like wavelet transformations and spectrogram analysis which have been found to possess higher resolution and sensitivity on transient signals. Spatial analysis techniques focus on the assessment of the spatial patterns and correlations of multi-dimensional data sets. In agriculture, for example, it could be looking at the spatial distribution of features or symptoms of diseases, covering fields and or crops. With these techniques of feature extraction, it is possible for researchers and practitioners to manage complex agricultural data employing time domain, frequency domain, time-frequency domain and common spatial techniques. This makes it possible to develop advanced machine learning algorithms and decision support systems of precision agriculture and plant performance analysis. Thanks to these strategies it is possible to identify early signs of plant disease which may be expressed as minutes patterns and signals for quick interventions, narrowed treatment and improved management of crops [69].

Table 3.5. Feature extraction algorithms

Algorithms	**Technique**	**Advantages**
Principal Component Analysis	Simplify a dataset, identify patterns with similarities and difference, Powerful Tool for analysing data	Data Compression, Low Noise Sensitivity, Decreased Requirement for Capacity and Memory
Discrete Cosine Transform	Spatial to Frequency Domain, Correlation with neighboring pixels	De-correlation, Energy Compaction, Concentrating Energy into lower order coefficients
Linear Discriminant Analysis	Data reduction, dimensionality reduction, class scatter measure	No illumination problems
Independent Component Analysis	Second order and higher order dependencies, statistically independent variables	Reconstructs data better than PCA in n dimensional space

3.4 Feature Selection:

Feature selection in predictive modelling as well as data assessment analysis is an important process, said analysis improves due to determining the most useful attributes out of a large dataset. Let's discuss three important feature selection methods:

i. Principal Component Analysis (PCA) is a well-known procedure for dimensional reduction intended for a large dimensional dataset to transform it into a simpler form by retaining only the most relevant features. In our context PCA works by transforming these original attributes into a new set of orthogonal components, the principal components, whose number is equal to the number of the largest variances in data. By doing so and retaining only principal components, which met the criteria such as those that account for a large volume of variance, while a high proportion of insignificant and weak features are eliminated gives appropriate selection criteria. This is very useful for tasks such as image processing and pattern recognition since these types of tasks have a lot of related variables in the data sets.

ii. According to filter bank methods, a set of filters are required to be utilized over a range of orientations and scales which means that some information that may be useful is located within the signals or images. In processing images together with the use of spatial frequencies, Gabor and wavelet filters}, image filter banks are also said to be used to texture patterns. To put it differently, in the bank, every filter is focused on some particulars in the input, like edge, corner, texture, or color detail. Filter bank approaches first do the inverse, and design filters based on salient properties of the images being displayed – this step is known as feature extraction [70-74].

iii. Evolutionary Algorithms apply optimization strategies inspired by the processes of biological evolution through genetic algorithms. While performing feature selection this is a search strategy that aims at determining the optimal subset of features for a specific task, for example, a classification task, and then evaluates the performance of the particular subset to cost. Typically, EAs begin with a pool of feature set subsets and then they suit selection, crossover and mutation strategies throughout multiple generations to achieve enhanced results. Once the optimate part is completed, each individual undergoes a performance evaluation based on several targets. Such algorithms are

iv. also said to combine a variety of complex principles that enable them to perform the task of pruning a feature space optimally, allowing and finding subsets that enhance prediction accuracy at better costs and lower total overfitting.

Table 3.6. Feature selection algorithms

Method/Parameters	Filter	Wrapper	Embedded
Implementation	Numerical Measure	Optimization Process	Amalgamation of Filter & Wrapper
Computational area	Less	More	More
Excessive Dimensional Records	Suitable	Not Suitable	Not Suitable
Computational amount	Cheaper	Expensive	Expensive
Generalization	High	Less	Less
Involvedness	Low	High	High
Reckoning Time	Efficient	Slow	Slow
Cognitive Productivity	Efficient	Inefficient	Inefficient
Benefit	Useful as preprocessor	High Classification accuracy	Reduce Computational Time
Drawback	Does not have to contend with features that are redundant.	Increased runtime	Classifiers Dependencies

To summarize, EAs, Filter Bank Techniques and Principal Component Analysis (PCA) purposes the best feature selection techniques for the application of machine learning and image processing information. Thus, the methods assist in the determination of the specific and most relevant features from complex datasets which leads to more accurate and interpretable, and more robust forecasting models across many fields, including plant disease detection and classification.

3.5 Classifier Model:

In simplified terms, classifier models function by providing a label to given data items based on their corresponding attributes. They can be classified into several categories according to factors which include the algorithms used, the area of the data, and the type of output. In the following paragraphs, we will discuss the classification of classifier models in detail.

a. Supervised learning

With supervised learning, a model is trained using a specific dataset meaning each training input sample has a corresponding output label. The aim is to understand how the input and output data are related. Some of the examples are:

i. Logistic regression is a technique that uses a logistic function to link to binary dependent variables.
ii. Support Vector Machines (SVM): These are sophisticated classifiers that determine the surface which best divides them. Our goal assigned to each machine learning algorithm is proximal to the information.
iii. Trees are directed acyclic graph models which are per branches of features and outcomes.
iv. Random Forests: A combinatorial method uses many binary tree classifiers to enhance accuracy of estimates without causing overfitting.
v. Neural Networks: Algorithms aimed at uncovering relationships in data by mimicking human brain processes.

b. Unsupervised Learning

Unsupervised learning means that you train your algorithm on without provided correct answers. Your task is to analyze a set of data points and determine its structure. i. Linear models:

i. Clustering algorithms (e.g., K-Means and DBSCAN): While these are most employed for clustering, they can also be utilized for classification tasks based on recognized groups.
ii. Binary Classifiers: These classifiers produce one of two possible classifications. Examples are SVM and Logistic Regression.
iii. Naive Bayes, multiclass classifiers: These classifiers can classify an input into one of three or more classes. Multinomial Logistic Regression is an extension of logistic

regression to several classes. Random Forests and neural networks, including SoftMax output layers for multi-class classification.

iv. Multilabel Classifiers: These classifiers can assign numerous labels to the same instance. Binary Relevance: Each label is treated as a separate binary classification issue. Classifier Chains: Link binary classifiers to consider label interdependencies.

c. Based on Algorithmic Approach

i. Approaches where the relations between the input variables and output class is treated as linear. E.g: Logistic Regression, Linear SVM.

ii. Non-linear Models: These models can capture more complex relationships across features and classes. Examples of ensemble methods are neural network, decision tree, neural network, support vector machines kernel [75-78].

d. **Ensemble approaches**

It uses many models to improve classification performance for instance:

i. Bagging: Consider language modelling as an example of a computational language processing problem. A language model is an important component of most of the NLP tasks such as automatic speech recognition, information retrieval, machine translation, N-gram classification, etc. Image classifiers are design for image recognition tasks. Convolutional neural networks (CNNs) are often used for image classification.

ii. Boosting: Boosting trains models sequentially, with every new one focusing on the previous model's flaws. Interpretable models include AdaBoost, Gradient Boosting Machines (GBM), and XGBoost. These models are straightforward to understand and communicate. Examples include logistic regression, decision trees, and black-box models. These models are more sophisticated and difficult to comprehend. Examples consist of neural networks and support vector machines.

e. Classifiers based on data type (e.g. text)

Classifier approaches are highly varied owing to the vastness of the language. This is beneficial because it enables the selection of the model that will best solve a given case thereby increasing effectiveness and interpretability. There are certain characteristics associated with a classification problem that can assist in deciding the most suitable classifier while at the same time assessing the pluses and minuses of each classifier in relation to the entire dataset and the particular problem at hand.

The second phase of the entire process is to select the appropriate base classifiers to be used in the stacking ensemble. Some of the elementary models that are used to extract a specific information or make a specific kind of predictions from the data include decision trees models, regression and support vector machines models. For each basic classifier's results to be enhanced, some hyperparameter tuning is done. After ensuring that all the base classifiers have been established and fine-tuned, then it is time to specify the structure of the stacking model. Reconstruction of the layers and arrangement of the ensemble and also the strategy that will be employed to aggregate the predictions of base classifiers into a final prediction are included in this plan. The stacking prototypical is then trained by providing it the training data set together with the prediction of the base learner.

After constructing the stacking model, we proceeded to test and conduct a thorough analysis of the model's performance. A couple of cross validation approaches like k fold cross validation has been used to test the area under the ROC curve AUC ROC. Accuracy, recall, precision, and F1 score are calculated with respect to the model under consideration. Relative comparisons of the stacking model and individual base classifiers and other ensemble models, qualitative characteristics of success are set. Feature classification including feature importance classification is done in order to determine which features and which classifiers are important in making likelihood predictions by the model. Ultimately, the objectives focus on simplicity of the model in terms of its deployment, scalability, efficiency, and compatibility with the set-up of already established agricultural systems. In relation to all these procedures, permanent monitoring procedures are applied with regard to each model so as to maintain accuracy during the evolution processes of the model so as to aid in the prediction and control of diseases in agricultural environments [79-81].

Table 3.7. Classifier model parameters

Classifiers	**Decision Trees**	**Neural Networks**	**Naïve Bayes**	**kNN**	**SVM**	Rule Learners
Accuracy	**	***	*	**	****	**
Speed of Learning (Attributes & Instances)	***	*	****	****	*	**
Model Parameter Handling	***	*	****	***	*	***
Tolerance to Irrelevant Attributes	***	*	**	**	****	**
Acceptance to missing Values	***	*	****	*	**	**
Tolerance to Highly Interdependent Attributes	**	***	*	*	***	**

Acceptance to redundant Attributes	**	**	*	**	***	**
Transparency of Classification	****	*	****	**	*	****
Noise Tolerance	**	**	***	*	**	*
Speediness of Classification	****	****	****	*	****	****

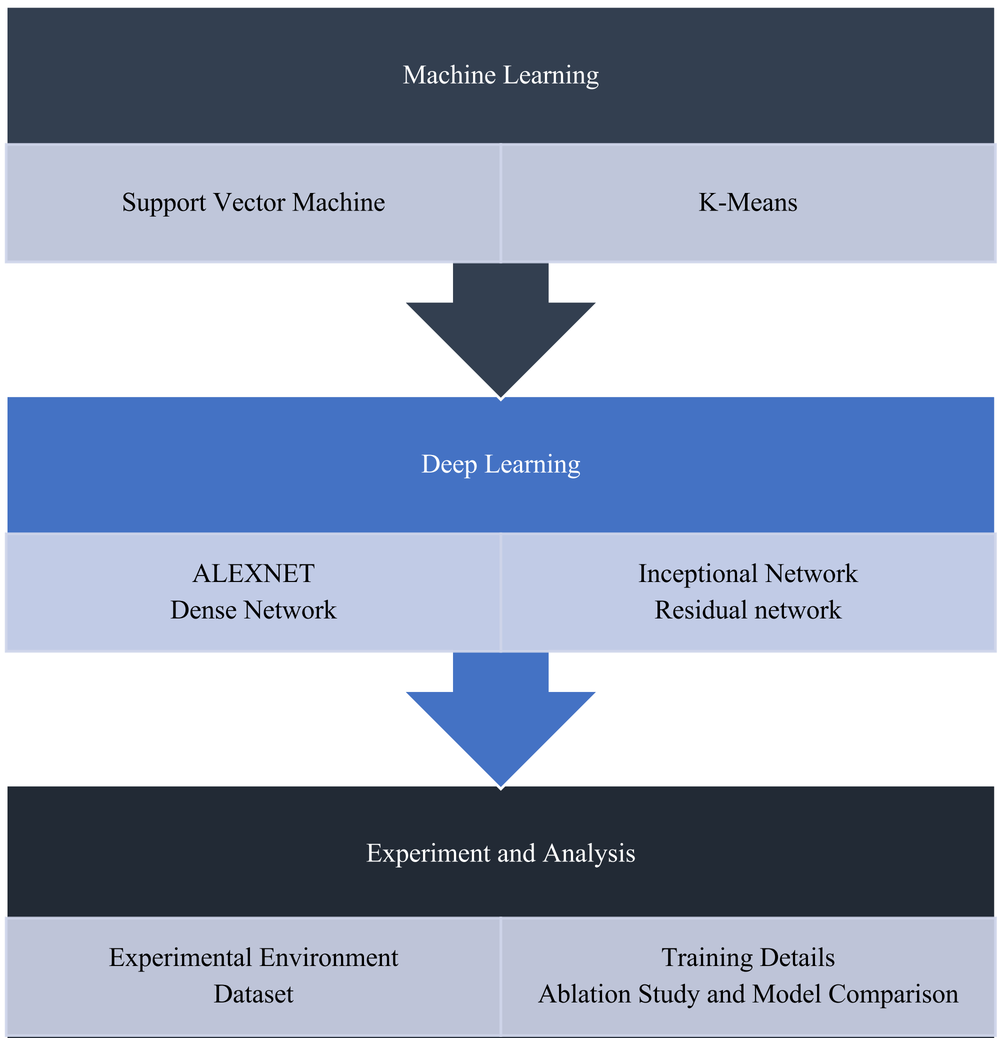


Figure 3.7. Techniques for the Prediction Model

3.6 Classification Algorithms:

Classification algorithms form a crucial part of machine learning systems as they categorize the incoming data into predefined levels or classes depending on the characteristics observed. Some of the most widely used classification methods include naive Bayes models, decision trees, random forest methods, support vector machine algorithms (SVM), and K-Nearest Neighbours. The naive Bayes classification describes a family of classifiers based on the use of Bayes theorem, under the restriction of the assumption of feature conditional independence. It considers specific features for determining each class and by the use of the class with the most features, like Naive Bayes and large datasets, gets a decision. Decision Trees utilize a cross-industry standard of ordered segments in a network. Each node represents a feature and makes a binary decision leading to the next node or leaves. Decision trees are intuitive and have the ability to process both continuous and categorical data. Random Forest is a ensemble learning technique that builds several decision trees and combines the outputs based on voting or averaging methods. It improves the accuracy and reduces the chances of overfitting due to the randomization effect in the process of tree building. Support Vector Machines (SVM) involves finding the maximum separating hyperplane by maximizing the margin which is the distance between the class of points involving the closest points of that class (support vectors). SVM is effective in cases with high dimensional space and enable non-linear boundary decision making by using kernel methods.

K-nearest neighbours (K-NN) is a simple and intuitive method that involves the use of the K closest data points to a particular data point in numerous parameters. K-NN is a non-parametric and instance based method, and therefore, it is suitable for small and medium data sets because it does not require a training stage. Each classifier possesses unique advantages and disadvantages, also the selection of the classifier is defined by the dataset, the difficulty of the problem and how comprehensible the model is. With these approaches, they could make strong and constant classification systems for different purposes such as, detecting diseases of plants, image tagging or sound sentiment analysis. In the binary classification, model accuracy is one of the most important characteristics of the target being achieved, however, in machine learning and classification tasks, it would not be the only one and other parameters such as precision, F-Measure and specificity must be taken into account. For the two-class classification problem, we can also compute performance metrics like accuracy, precision, recall, F-measure, and specificity of a model. As most predictions are never correct, it is important to define how correct a model is over many predictions. Each positive prediction made by an algorithm model will encompass some errors and can be termed as a positive prediction but not true. In contrast, recall can be defined as a performance metric that measures the ability of an algorithm to find all the relevant cases within a dataset. The F-measure or F1 score is the metric which will always be somewhere between precision and recall and is calculated as harmonic mean of two metrics. However, specificity gives additionally the percentage of cases where actual Yes's were predicted as No's. These measures tell us vital information and how the model can possibly be improved and as a result even predict better in real world situations.

This chapter discussed the systematic process followed to identify diseases in Tulsi (*Ocimum sanctum*) leaves through an ensemble stacking classification model. The methodology involved several stages: image acquisition, preprocessing, feature extraction, model training, and evaluation. High-resolution images of Tulsi leaves were gathered and pre-processed in the form of normalization, resizing, and augmentation to increase model resilience. Feature extraction was done using pre-trained convolutional neural networks (CNNs) as a part of a transfer learning method to effectively represent disease patterns. To overcome limitations seen in earlier single-model strategies, an ensemble stacking classifier was utilized, combining diverse base models along with a meta-classifier to make the final prediction. This ensemble architecture largely enhanced classification performance, minimized overfitting, and maximized generalizability. The approach set the foundation for an efficient and scalable disease-detecting system specific to medicinal plants. The following chapter describes and discusses the results of experiments conducted with this approach.

Chapter 4
Results and Discussion

4.1 Tulsi leaf infections

In order to save the natural and cultural parts of the exalted Tulsi plant, it has to be properly grown, nurtured, protected and stored. Maintaining high quality and finding efficient methods of growing useful Tulsi leaves are important factors of sustaining quality of production and crop yield. The primary regions to cultivate tulsi plants are the temperate region as they require moist soil. Nevertheless, they are susceptible to an array of fungal, bacterial and nematode diseases. The manifestations of the infection includes exposure of the leaves to irregular black spots along with circular lesions having dark centres, yellowing of the foliage, stunted growth and general wilting of the plant and fungal infections of the stems. Yellowing of the leaves can be caused by nutrient deficits or seed contamination. Bacterial infections appear as dark or black water-soaked patches or angular lesions with streaks on the leaves and stems. Powdery mildew, seedling blight, and root rot are common diseases that inhibit Tulsi growth, affecting the leaves, stems, and roots. Fungal pathogens cause around 50% of illnesses, while insects, pests, and nematodes account for 40% and bacterial pathogens for 10%.

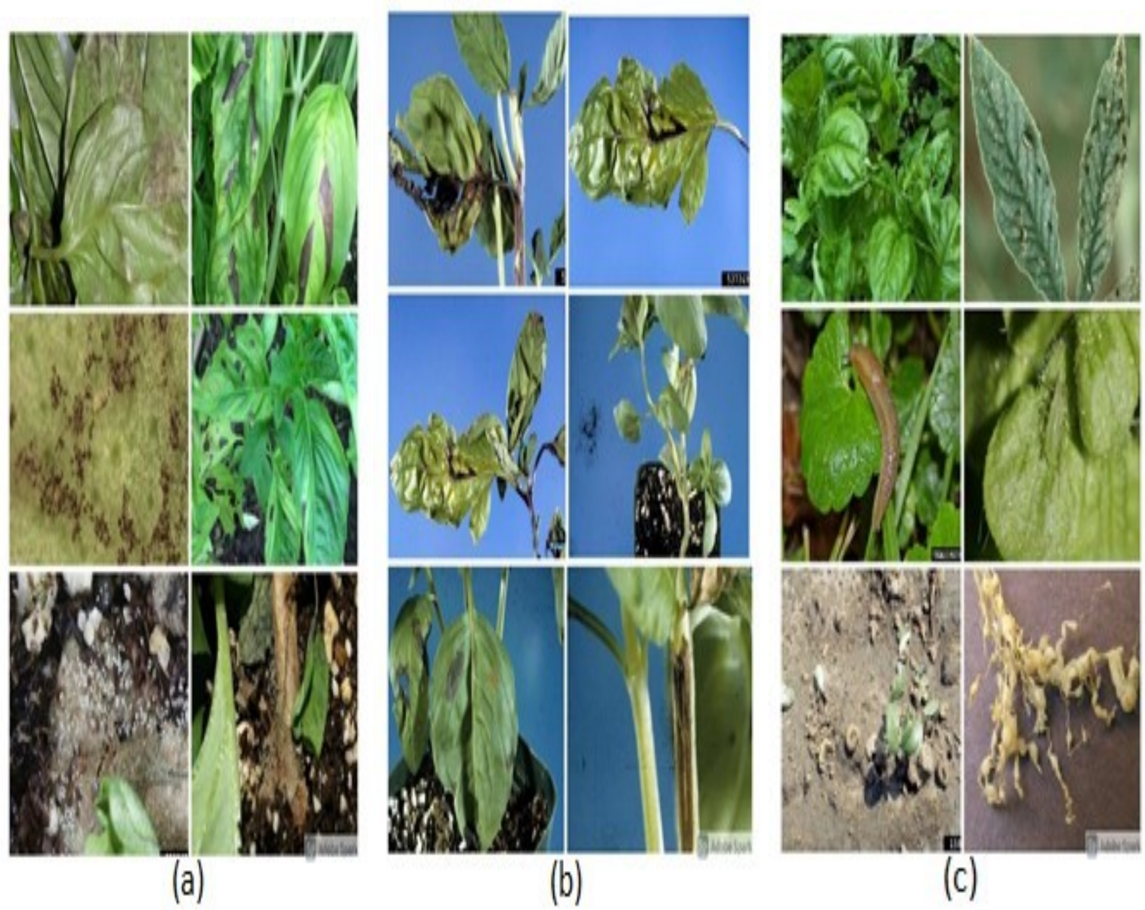


Figure 4.1. Fungal, bacterial, and insect-infected tulsi leaves.

Figure 4.1 depicts different types of infections and symptoms due to different types and causes of infections in various parts of tulsi herb.

4.2 Pre-processing

Image pre-processing involves enhancing photographs by removing unwanted distortions and emphasizing positive components. This usually entails cutting the leaf area and transforming it from RGB (red, green, blue) to greyscale. To improve image attributes, a variety of strategies can be used depending on the dataset. It summarizes common pre-processing methods and algorithms, emphasizing the features they improve. Successful image-based machine learning relies on accurate feature extraction during pre-processing. These strategies try to extract relevant aspects in images and improve visual data quality by removing undesired distortions or amplifying certain elements that are required for later processing. Scaling, enhancement, noise reduction, segmentation, grayscale conversion, and binarization are all examples of preprocessing procedures. Beginning with a raw image, the pre-processing stage produces an improved output image. Once leaf image faults have been resolved by enhancement and filtering, the next step is picture segmentation to identify the proper region. Figure 4.2 shows how our work extracts relevant image properties using a variety of pre-processing approaches such as enhancement, noise reduction, segmentation, color transformation, and binarization.

Figure 4.2. Pre-processing methods

Image pre-processing means removing image noise from the images in so that the vital features are lifted and worked with. It is worth noting that the type of device used in acquisition is of no relevance as all devices produce noisy images. The figures above depict the various pre-processed models of Jupyter Notebook of pre-processed tulsi leaves infected with fungus, bacteria and nymphi. In Image Figure 4.3, gray level spots, level of brightness, and changes of figure through histogram, elevated expositions, and other methods of image analysis are presented. Such a representation is a bar graph in which horizontal axis depicts the x-axis pixel intensity levels while the vertical shows the frequency of how many times the said values were recorded.

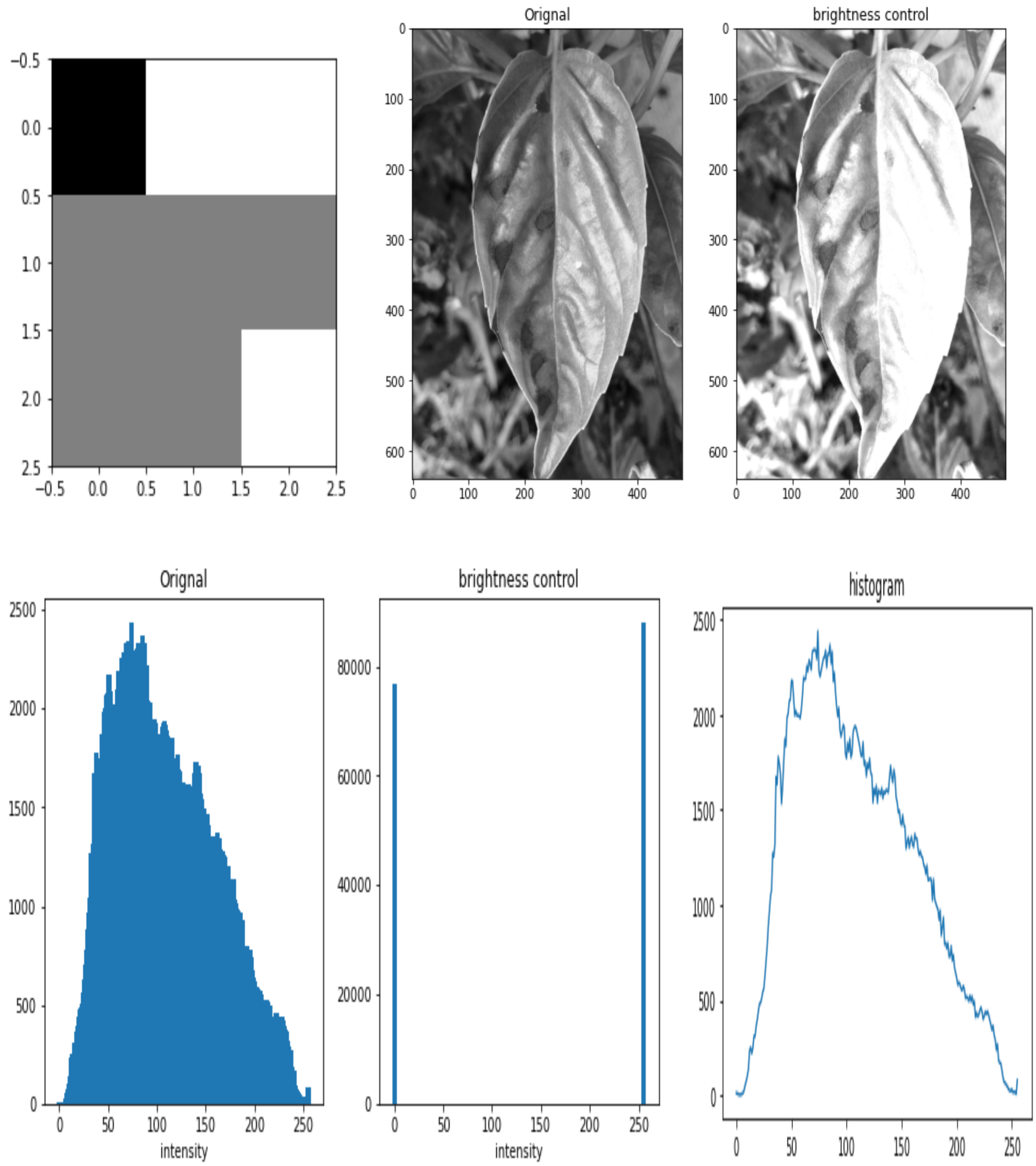


Figure 4.3. Histogram and brightness control for fungal infected leaf

Figure 4.4 depicts bacterially infected leaves and a histogram in terms of RGB, with red, green, and blue histograms representing different intensity levels.

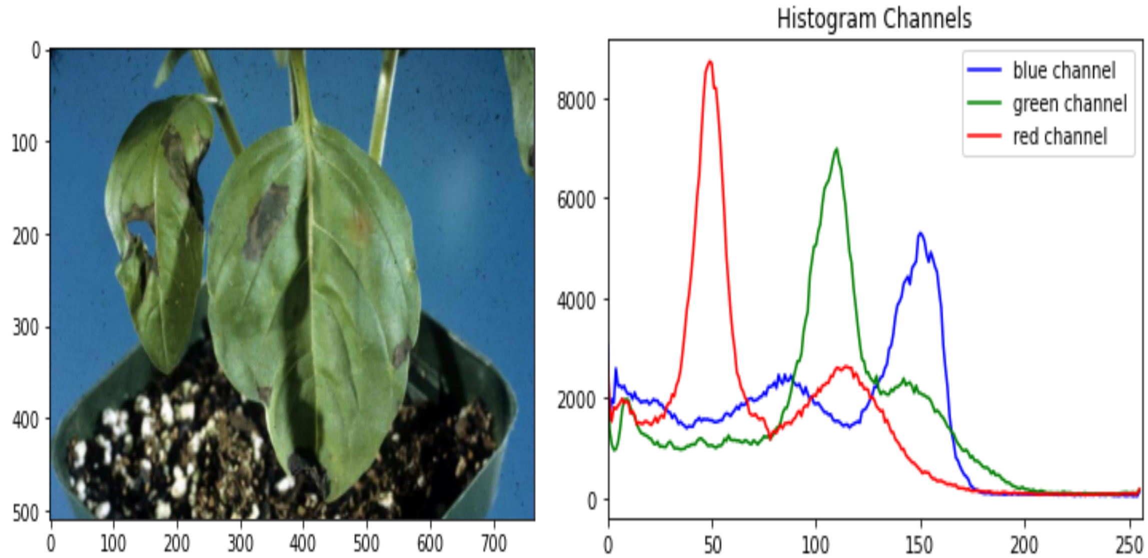


Figure 4.4. Histogram channels on bacterial infected leaves

Histogram Equalization is a technique that improves image contrast by widening the range of grayscale pixels by histogram flattening. Figure 4.5 shows the improved contrast of the nymph-infected leaf image after using the histogram equalization technique, which essentially extends out the histogram.

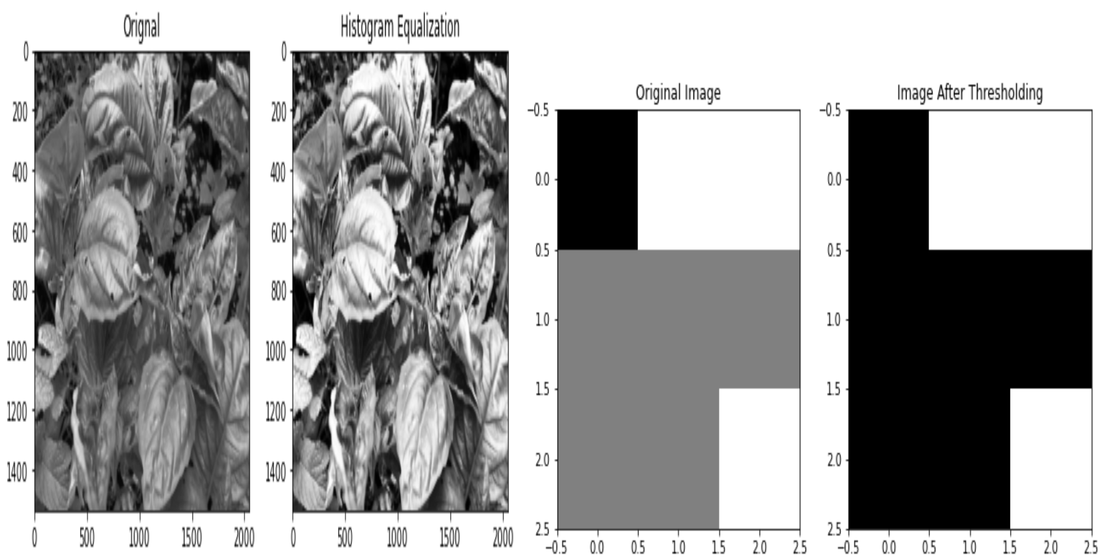


Figure 4.5. Histogram equalization on nymph infected leaves

Here an image is treated as a function and not just an etched array, that is instead of being flat it has three axes, x is the table row axis and y is the column table axis, which speaks volumes. The process of intensity modification is the custom of manipulating individual pixels of a certain colour. Let us assume we are looking at the linear transformation. The

equation g (x, y) = 2 f (x, y) + 1 which is to say g (x, y) on the basis, increases every pixel on a given image and also adds 1 to the new pixel value if any.
The variable r specifies grey level intensity, like histogram values. Notably, when intensities are transformed, darker intensities get brighter, brighter intensities become darker, and mid-range intensities remain constant. Parameters for controlling this transformation include α α for adjusting contrast and β for enhancing brightness.

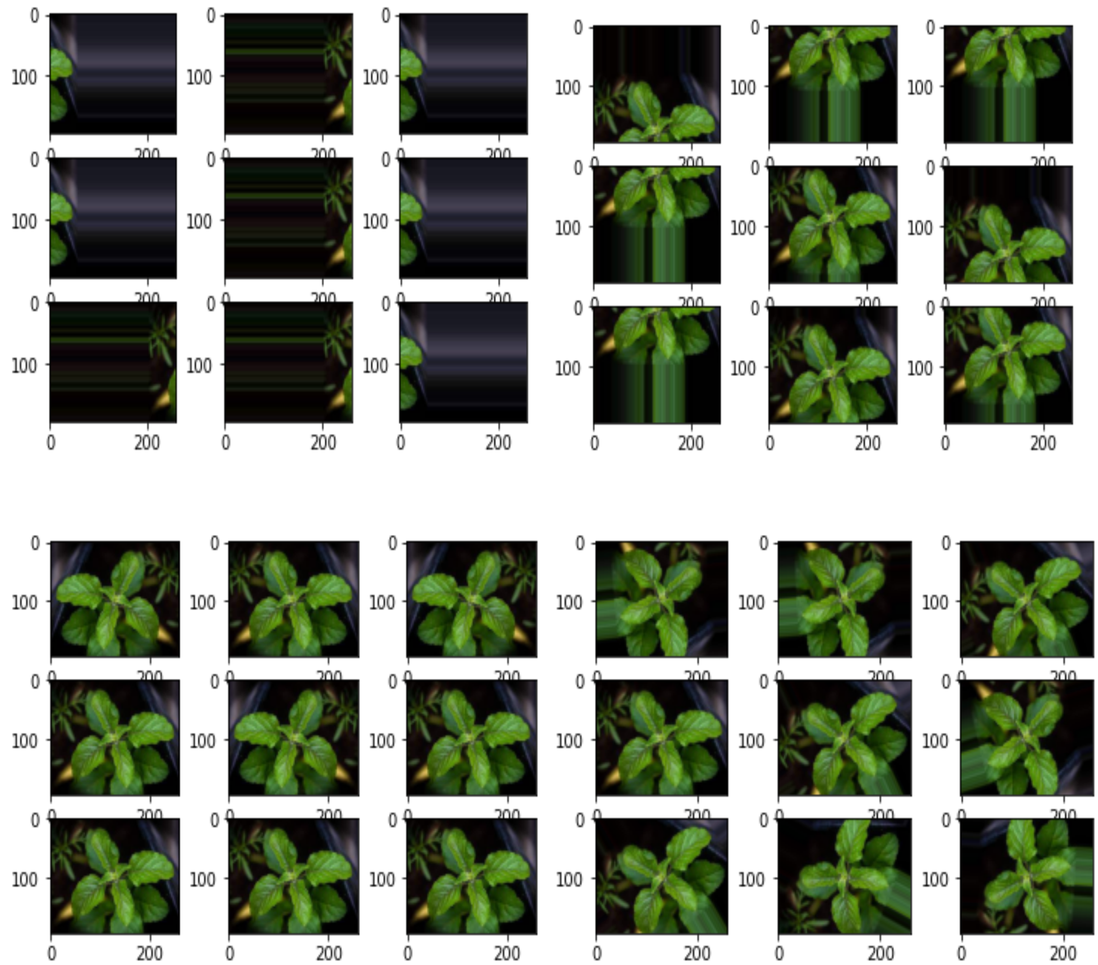


Figure 4.6. Data augmentation techniques

4.3 Image segmentation

In digital image analysis, image segmentation is a crucial process that often requires dividing the entire image to focus on certain problem areas. The intention is to remove foreground objects from the background and sometimes use edge or line detectors. Image segmentation is an important step in image analysis that plays a significant part in there being good accuracy of the desired results. It involves dividing or classifying an image into parts, the extent of the segmentation depending on the specifics of the problem. During segmentation, the image is divided in such a way that relevant objects are separated from the background in an effective way.

Image segmentation can be achieved using a number of techniques, ranging from basic thresholding to more complex colour segmentation techniques. These techniques usually emphasize characteristics that are easily discernible and appear to be separate objects in one photo. There are different sources for colour, edges and textures, but picture segmentation techniques utilize those sources in different pictures. Gray-level segmentation methods utilize two basic approaches. The first deals with the discontinuity of gray-level values where an image is divided along abrupt increases or decreases of gray levels. The second strategy is the clustering of pixels that possess approximately the same gray level by means of thresholding and region growing. The segmentation threshold which is assumed to be the same for all pictures of a set having the same exposure is often determined by the use of histogram techniques. For the purpose of leaf classification, edge detection, texture detection, and region based detection methods must be combined to obtain segmented images of good quality. These techniques help in the retrieval of the important features and therefore increase the accuracy of the image processing tasks.

4.4 Feature extraction

Features are definable factors which are extracted from various components of a picture including color, texture and geometric moments. In the construction of an aggregated dataset comprising of multiple feature types such as texture, shape and vein characteristics of a leaf, contour based or region based extraction methods have to be employed. Contour based acquiring employs length, width, aspect ratio and diameter of the leaf as distinguishing parameters. Region based methods define leaf characteristics by descriptors including shape, area, and compactness, rectangularity and eccentricity among others As previously stated, appropriate feature extraction is central to enhancing the accuracy and aiming of the automated classification models. In this case, the extraction process is made easy when relevant features are properly identified. The major parameters for the classification of leaf infections are texture, color and morphology. This approach to feature selection aids in dealing with large volume of data by retaining relevant information but also conserving resources. This reduction of the training/testing/validation data renders model building efficient and enhances the understanding of the process of machine learning. Linear Discriminant Analysis (LDA) is a supervised learning technique aimed at reducing dimensionality which Tunable LDA leaves information concerning the disease the leaves could have. advances in feature representation and the learning process of the leaves in the classification tasks.

4.5 Feature selection

Image processing involves a number of activities. In constructing a classifier model, for example, feature selection is certainly one of the model activities, and it is important because irrelevant features are removed, which reduces the number of features to consider in such a large data set. The table contrasts filter, wrapper and embedded techniques for

feature selection in terms of efficiency and time taken, complexity, accuracy rate and performance of the model.
Filter approaches compute predictor importance independently of the prediction model, retaining those predictors that match requirements. Wrapper methods, on the other hand, begin with a given model and iteratively add or delete predictors in order to increase the effectiveness of that model. Though, wrapper methods are often claimed to be more accurate and precise, they might require intensive computations. This approach entails repeatedly training a model under different feature subsets and assessing the outputs according to validation accuracy.

On the other hand, when it comes to identifying the features, filter approaches guide compare feature sets against a proxy instead of accuracy achieved cross-validation approach. Despite being faster and not requiring the model to be trained from scratch at each round, filter approaches are prone to be less accurate in identifying the most predictive features for the model. Embedded techniques are a mixture of the filter and the wrapper methods that perform feature selection during the learning process rather than in independent feature selection stages. An instance is L1 regularization, which helps in boosting the sparsity of the model weights while performing automatic feature selection throughout the model training [80-81].
According upon the problem's particular requirements and constraints, filter or wrapper techniques can be employed to effectively choose features for leaf infection classification. Each method has advantages and disadvantages that must be balanced in order to achieve computing efficiency.

4.6 SVM Model Implementation

Predictive modelling tries to develop a model that would enable it to classify novel, previously unseen observations given test data based on pre-existing paradigms. The primary focus is on further dividing the diseases into healthy leaves and diseased leaves types. Such model can be developed using any free software or licensed machine learning application, and its effectiveness can always be gauged using different efficiency measures. Each time a feature or group of features is added to or removed from the dataset, the model of the system has to be trained and tested again. For automated and computationally fast machine learning algorithms, leaf images can be quickly processed to find out if the tree suffers from an infection.

A system that utilizes a multidimensional approach that employs several classifier models is advised in order to enhance recognition performance and accuracy. The prediction model system typically consists of five independent components by image preliminary analysis, object segmentation, feature development, disease detection, and specifics recognition (model training and testing). Image preprocessing is a basic approach for choosing the proper characteristics for the images. Images with noise or any other undesirable characteristic are considered not to be suitable for analysis regardless of the imaging system used (Naikwadi and Amoda, 2013) [82]. Chemical composition, noise, and environmental or external interference can all be reasons why image structures vary, and therefore different pre-processing treatments can be applied.

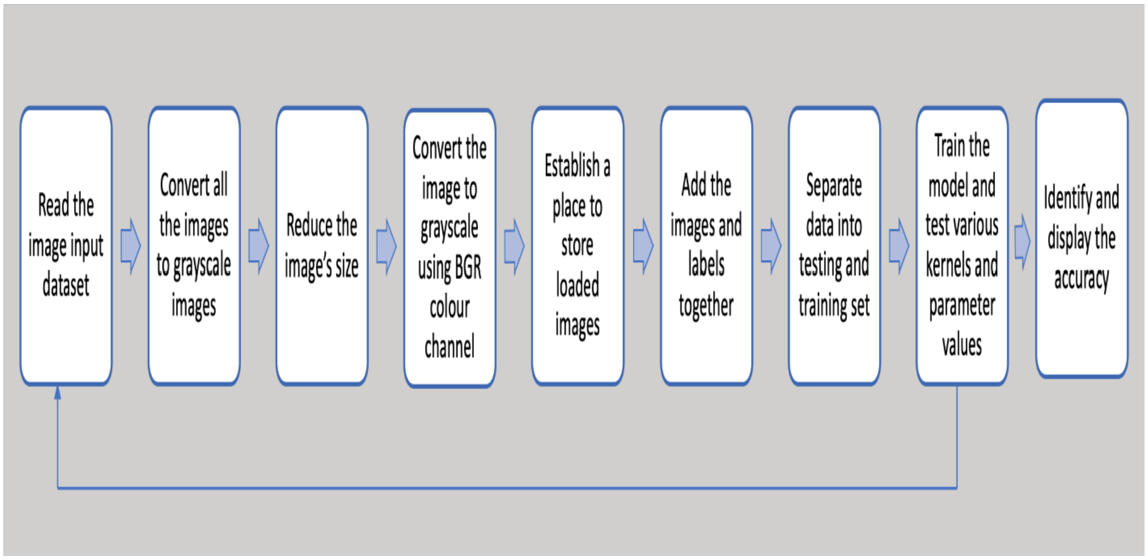


Figure 4.7. Implementation of SVM model.

Identifying leaf infection has been accomplished by information image scale, image augmentation along with segmentation, filtering of noise removal, greyscale conversion, thresholding and binarization. In Figure 4.7, the implemented steps of SVM for disease detection are depicted. SVM is one of the algorithms contained in ML, which constructs an optimal separation hyperplane for distinguishing new data points into various classes. In the specific configuration of integration of SVM model pointed out Azlah et al (2019), there is a possibility to adjust its regularization and kernel parameters for more increased axial [83]. The kernel parameter essentially defines how the separation is to be done, linearly or some other way. Regularization specifies the degree to which the simultaneous minimization of the error with respect to training samples and the maximization of the hyperplane in the margin is sought. The results of the SVM calculations of predicting one or another result incorporate the intelligence available in computers in a direct way, which is the consequence of mathematical modelling and a great deal of theoretical research, thus since completion and efficiency of the systems is improved the efficiency of the prediction increases as well. Further development of SVM made it possible to overcome overfitting, nonlinearity and local optima which previously were the problems with interpreting the results of a machine learning model.

Oo et al (2018) stated that accumulating evidence does not intensify the worry that activity observation might affect prediction accuracy [84]. In the process of classifying leaf infections, SVM is very effective as it can model and predict large datasets with high accuracy.

4.6.1 Experimental results and discussions

Our implementation work employed an imaging system called CV Learning Studio, which is known to be an easy and collaborative open-source computer vision tool. We managed our images into four folders, which contain leaves of fungi, bacteria, insects and healthy leaves on the CV Studio. The system was designed based on the data set made up of four classes and 10,169 images. Histogram of Oriented Gradients forms an important aspect in our implementation. HOG is used to add histograms for parts of the picture in order to make features. To generate a HOG feature, first, the image needs to be converted to grayscale. The image is then downscaled to a lower resolution quality to increase processing speed, then the image is converted to grayscale to minimize the number of colour components.

OpenCV work with photographs in BGR format, hence this colour channel was optimally exploited for greyscale conversion. In the case when HOG is used there is for example a set of features obtained that are used for training an SVM model. After HOG feature creation, the images are imported to CV Studio for training and testing set creation. In order to process the data, signed arrays that were marked as images were first concatenated vertically. The structure of the final array is [[1],[0],,[0]] whereby the corresponding figure represents a label. In the end, we combine the images and the descriptions and the data will be allocated into sets to be tested and trained according to the actions that Bao et al, 2020 provide [85]. This process guarantees that the machine learning model is well trained and validated with the prepared data set in the CV Learning Studio environment.

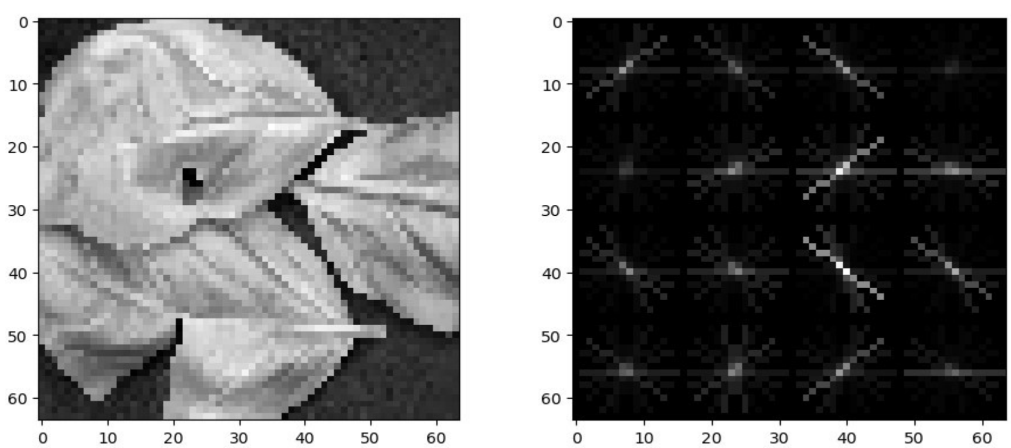


Figure 4.8. Image conversion to grayscale and HOG feature representation.

Hyperparameters play an important role while fine-tuning SVM (Support Vector Machine) models. For instance, one crucial hyperparameter is the choice of kernel, where typically most implementations use the radial basis function (RBF) kernel. In SVM, a regularization parameter C is used to motivate an optimal trade-off between the accurate classification of the training data and the broadness of the decision function's uncertainty region. The C parameter also determines the breadth of the margin as larger C is associated with smaller

M and vice versa. Schoellkopf et al. point out that a “large” C means that the function class should be narrow enough to separate all training data, thus leading to marginal misclassification [86]. Many times, there is no need for such fine-tuning and in such cases, a larger C leads to a wider margin which results in a sparse decision function. The RBF for example uses three parameters, cannot expand freely γ which changes the area of influence of the kernel and, hence the decision region in the SVM model. For example, if a low value of γ is used, choice borders will be broader (“far”), and as the values increase, the borders will become narrower (“near”) decision boundaries. The decision function might become overly sensitive to data points, resulting in excessive overfitting for high γ values. Furthermore, the parameter validation’s goal is to find the best suitable values of C and γ so that SVM models once trained can generalize well on unseen data. According to Mingqiang et al. (2020), performance and robustness of the SVM classifier can be improved by optimal adjustment of these hyperparameters [87].

Table 4.1. Comparison of SVM accuracy with various training set settings.

Train %	Accuracy %
80	96.71
75	94.53
70	85.52
65	81.34
60	80.92

Once the training phase is completed, new images can be classified using the predictions made by the trained SVM model. Their aim is to classify each image by feature extraction and pattern recognition based on decision boundaries which they have learnt. For instance, in the case of the sample image depicted in Figure 4.9, the model was able to correctly determine that the image is representative of the fungus category. This is to show how the adapted SVM model is capable of classifying and retrieving images according to their features and the learned patterns.

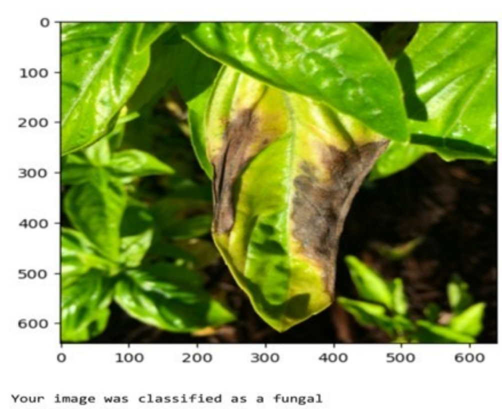


Figure 4.9. Test image correct classification illustration.

The present research highlights the effectiveness of using training and testing dataset combinations, rather than just image data, for automated diseases identification in outdoor environments. In our experiments, we used more training datasets, which improved disease prediction accuracy significantly [88-89]. Our results show the reliability and robustness of training datasets, which have been acknowledged for their accuracy and effectiveness even with a small number of training samples, when statistical learning theory is applied. Support Vector Machines (SVMs) are primarily binary classifiers, but they can be modified to handle the wide range of classification tasks encountered in remote sensing research. Our future efforts will centre on developing a breakthrough forecasting algorithm with the ultimate objective of increasing disease detection accuracy even more. This endeavour builds on the success of our previous study and attempts to develop automated disease detection approaches in field applications [90].

Table 4.2. Comparison Table of literature

Reference Paper	**Dataset Images**	**Training set**	**Preprocessing Features**	**Model Chosen**	**Tools**	**Average Accuracy**
Proposed (2023)	Tulsi	Hu Momemts, Texture, Harlick, BGR color histogram, HSV	Hu Momemts, Harlick Texture, color histogram	RF, SVM, KN N	Computer Vision Laboratory Setup, Open CV	KNN-99.6801 RF-99.89 98.2408-SVM
Anami et al. (2010) [16]	Medicinal Plants-900	HSV, YCbCr	Texture, Edge, Color,	SVM, RBENN	ANN Tool Matlab,	AA-90% Combine d
Naeem et al. (2021) [52]	Medicinal Plant Leaves-`6	Cropping: Leaf Region, RGB to GL.	Texture, Runlength, Multispect ral	MLP, LB, B, RF, SL	Open CV , Computer Vision Laboratory Setup,	99.10% tulsi-AA
Ananthi et al. (2014) [24]	Medicinal Leaves	RGB-Grayscale Binary	Morpholo gical, Filtering	Multilayer Perception	SPSS , Matlab	Recogini tion Rate70.87%

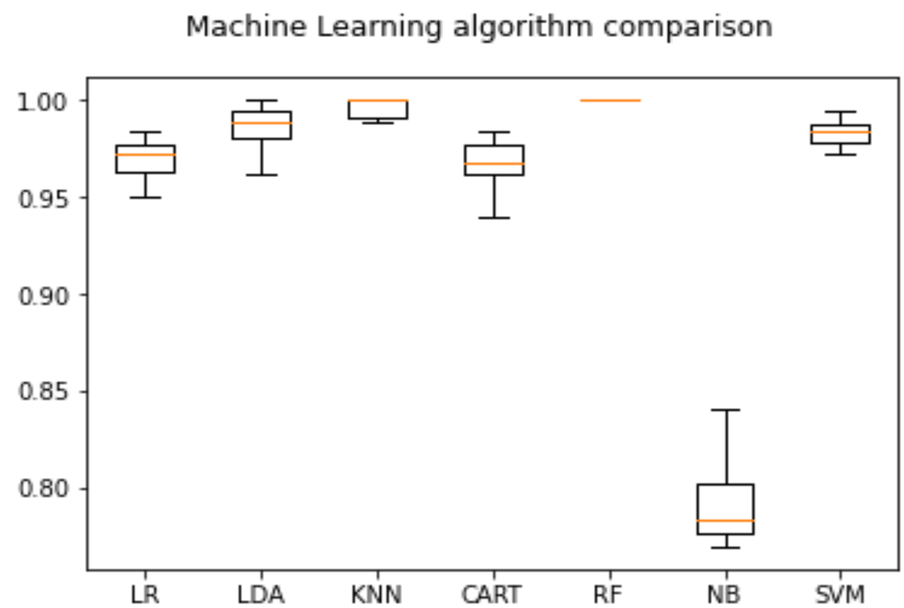


Figure 4.10. Box Plot Comparison for Machine Learning Algorithm

Table 4.3: Machine Learning Algorithm Comparison

Model	Accuracy	Execution Time
LR	96.5986	8122.75
LDA	98.5927	7631.92
KNN	99.6484	764.56
CART	96.3044	4458.58
RF	99.8264	14443.21
NB	78.8308	176.83
SVM	97.4207	3364.36

4.7 Proposed methodology for Best Classifier Model

One of the biggest challenges facing agriculture and plant pathology is the impregnation of a plant leaf disease, and it appears that the lack of a solid methodology for a leaf problem classification makes this task even more complex. Determining the effectiveness of your leaf disease classification method is contingent to several factors including the richness and variability of the dataset and the careful selection of deep leaning models and tuning of the Hyper Parameters. Moreover, there's an ever-increasing necessity for up keeping and accumulation of the data that will ensure that the functionality of the unit is preserved for a longer time [91-92].

Classification techniques employed enhance the performance of anti-virus systems by using pre-processing techniques employed in machine learning models. Such things have been well reported in the literature that exists and include rotation of images, cropping of images, noise elimination and grayscale transformation among other things. There are always multiple unwanted forms of noise in the data set that gets included during the collection process, so these comprehensive methods of bi pal inductively style can assist in its preprocessing by improving the core properties of the data set. With the correct model and appropriate parameters, we can develop reliable detection and recognition systems of various plant leaf diseases, which in turn will aid in better agricultural practices and the management of plant health [93-95].

4.7.1Otsu Thresholding

The software utilizes a specified cutoff value to translate a greyscale image to a binary image while treating any value over or under the cutoff as equivalent. In a grayscale picture, each pixel is assigned a brightness value that ranges from 255 at its most maximum to 0 at its very least. Thresholding is defined as the mapping of a pixel's intensity to an arbitrary defined threshold. All pixel values which are less than the threshold value will be assigned to 0 binary meaning black in colour, while all values greater than will be equal to one binary meaning white in colour. Such pixels can be selected or suppressed depending on the threshold value. After the cut off value is defined, the pixels can be changed in accordance with this cutoff so that some parts of the image can be maintained or removed. The general procedure is in place as follows: The first step involves starting off with an image which is then converted into a NumPy array which symbolises an RGB image as shown in Figure 4.10. Then the picture is converted to black and white (single colour image).Furthermore, Figure 4.11 shows the sample image in HSV (Hue, Saturation, Value) format, demonstrating the transformation of image representations for further analysis or processing. These stages are critical in image preprocessing because they identify relevant characteristics and prepare the data for further analysis or classification tasks.
Otsu's thresholding relies on the idea of selecting a threshold value that minimizes intra-class variability within both the foreground and background groups while maximizing inter-class variance, as seen in Figure 4.12. In layman's terms, it aims to construct a level that effectively distinguishes pixels belonging to the object of interest (foreground) and the rest of the image (background). This method can be particularly helpful when the backgrounds and objects in the image display distinct severity patterns of distribution, as it may build an optimum threshold without the need for manual parameter changes.
Otsu's thresholding technique is widely used in a variety of image processing applications, including image segmentation, entity detection, and detail analysis. Once the threshold is established, it is utilized to compute binary values, essentially binarizing [96].

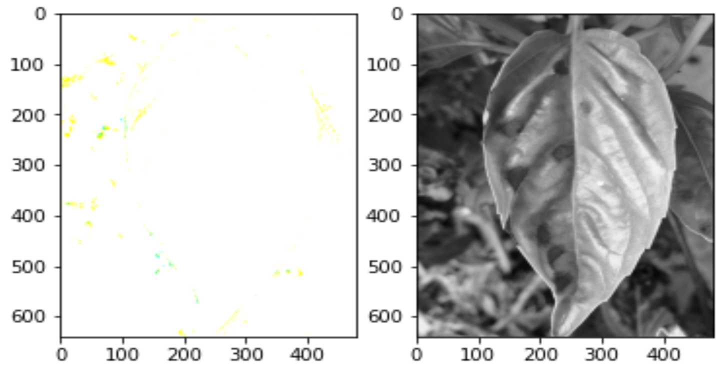
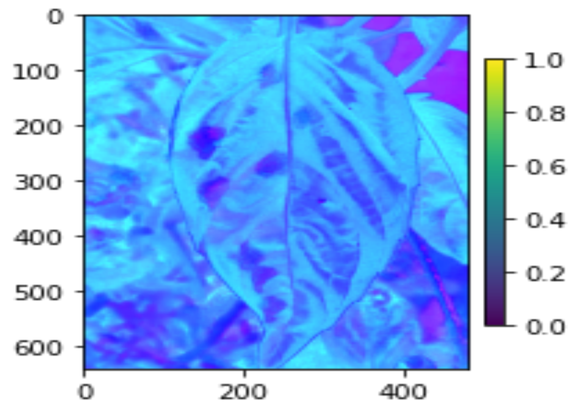


Figure 4.11. Reducing RGB input information for positioning inside the image creates an acceptable range (0–1 for floats and 0–255 for integers).



4.12. Image format for HSV

After thresholding the system, the total number of non-binarized Sauvola's local pixels is calculated. To decrease the background noise, a Gaussian filter is used, and local pixels are computed using a predefined algorithm. The geographic centre of a circle has been determined using specific coordinates, such as (220,110). The active contour area, as shown in Figure 4.13, is calculated from the provided image using the approach described. Active contour models, often known as snakes, are used for image segmentation tasks by delineating object boundaries using defined criteria and energy functions.

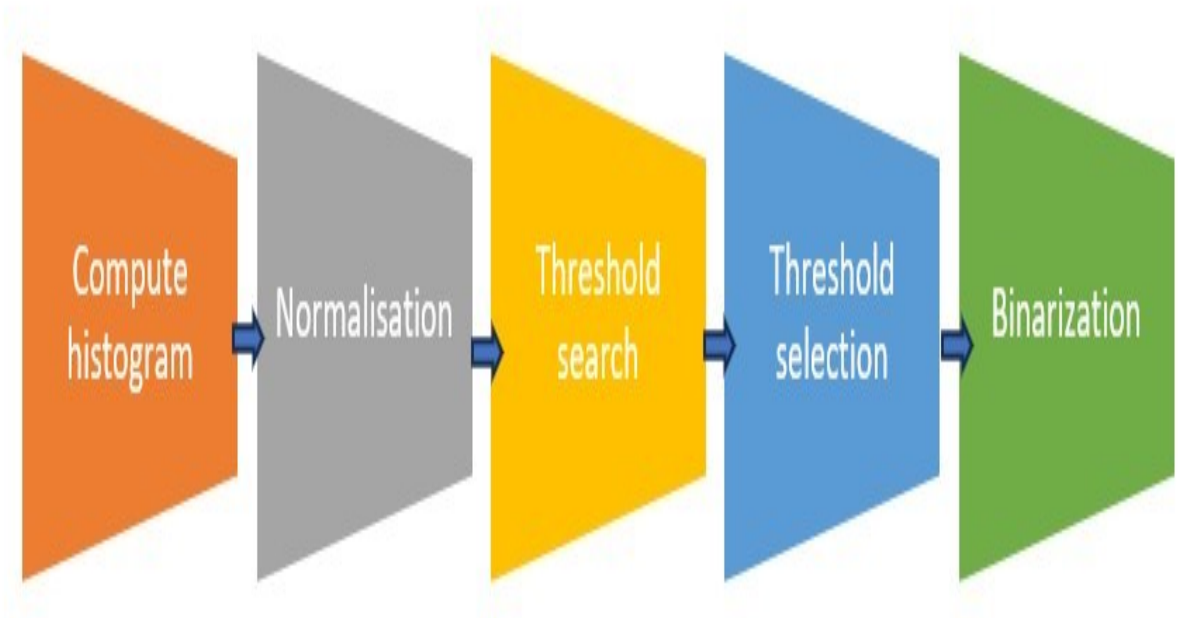


Figure 4.13. The Otsu's thresholding workflow

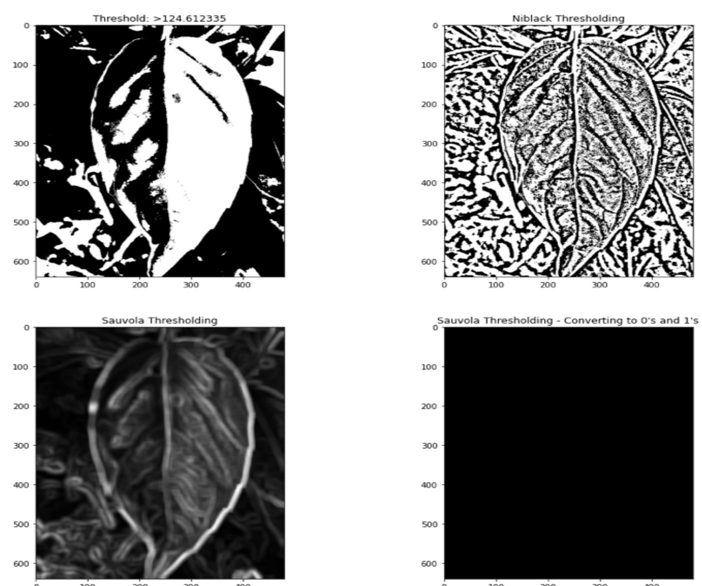


Figure 4.14. The Otsu's thresholding images for output

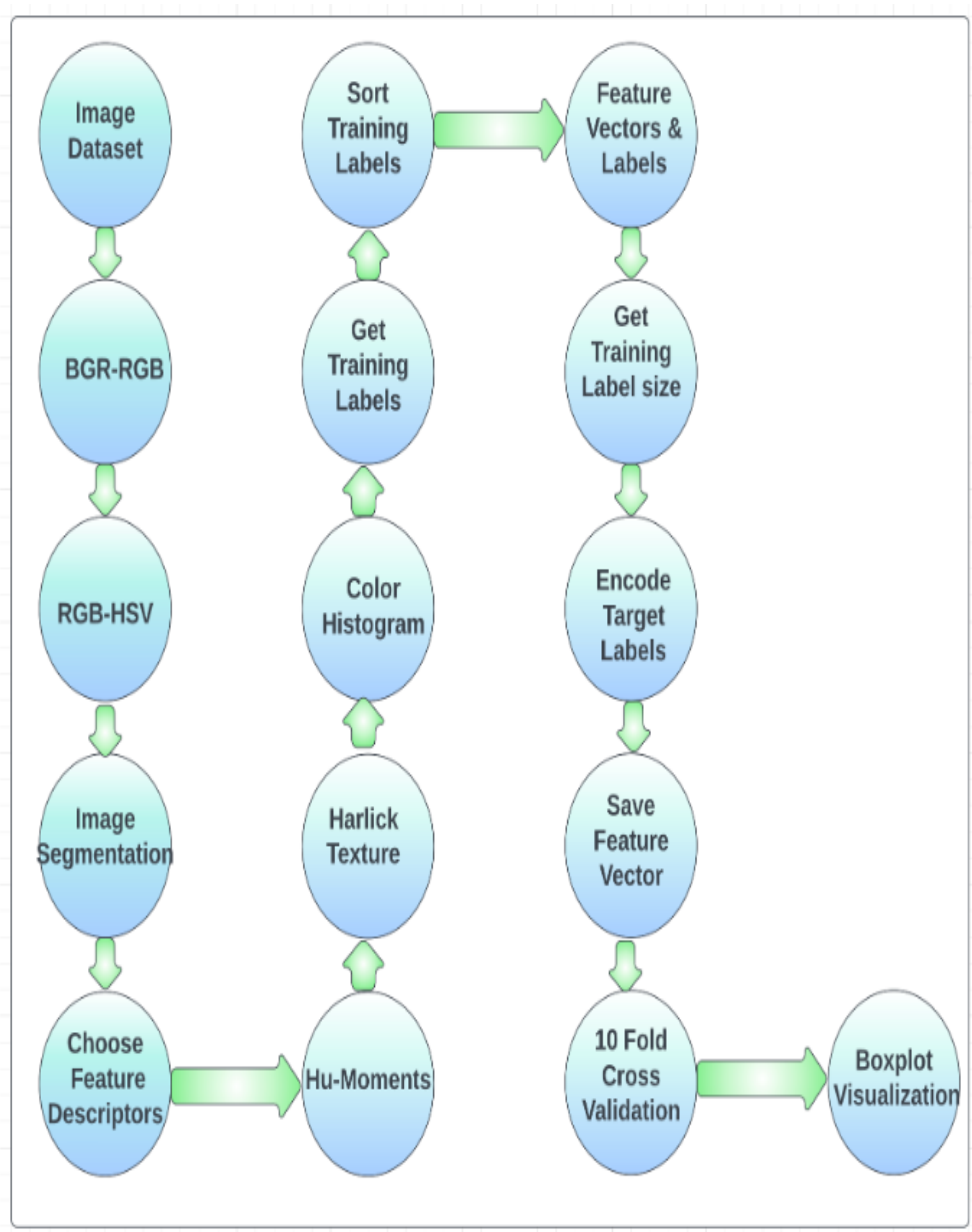


Figure 4.15. The Otsu's thresholding images for output workflow

The step-by-step procedure for Otsu's thresholding method is shown below:

i. Generate Histogram: Construct a histogram of image pixel's intensity levels for the given greyscale image. This histogram depicts how pixel values are allocated across the given image.

ii. Adjustment of The Histogram: Adjust the histogram with a goal to make the total sum of all elements equal to1 making it a probability density function. This proportion guarantees that the fact of every pixel number is determined adequately.

iii. Chosen Threshold Estimation: Go through all possible values for threshold (Usually 0-255 for 8-bit gray images). For each value of threshold, do the following: Within-Class Variance (Weighted Sum of Squared Variance). This finds the scattering of the pixel values within each one region (for example the foreground and background). It is calculated as the weighted averaged of the variances of both classes with weights.

iv. Inter-Class Variance: That looks at the variation found between the two classes (the bounded sets, for I, the foreground and background). It is calculated as the weighted mean of the square deviations of the two class means from their combined mean, with weights proportional to the class probabilities [97].

v. Threshold Technique: The maximum Inter-class variation (or the minimum within-class variation) achieved in isolating the portions of an image which exhibits the class image blocks from the background gives the optimal threshold. This kind of threshold is very helpful in separating the foreground from the background collections of pixels.

vi. It describes binarization which is also known as image thresholding in a simplistic context. For example, in this instance, pixel values are grouped into different segments with values above the threshold delineated as the object and values lower than the threshold marked as the background.

To produce detailed and accurate segmentation of an image, Otsu proposed automatic threshold selection as a result of which the intensity distributions of the pixels in the image get partitioned into two classes. Otsu's thresholding seeks to maximize class variance while minimizing intraclass variance of the thresholder image.

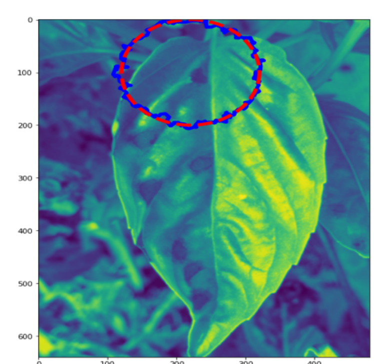


Figure 4.16. Otsu's thresholding image, with contour highlighted as output

4.7.2 Chan-Vese segmentation

The Chan-Vese segmentation strategy is when images are of low quality, those images may have noise, have objects with poor boundaries, or have poor noise itself. Even in setting with the greater internal thresholds, Chan-Vese can still perform. One important consideration is where active contours fail. This theory is a major improvement over other segmentation methods where the segmentation does not depend on the explicit definition of contour. Rather than working on the explicit definition of edges, It addresses the issue of segmenting images without clearly defined external edges which is often the case in image analysis. This approach uses level set theory, in which level sets are iteratively grown to minimize a well-defined energy function to constrain evolution. The energy function consists of weighted terms, which include:

i. Varying intensity averaged over a specified region; specifically, the difference between the average of the entire image and its pixels inside the segment
ii. Mean intensity deviation averaged inside a segmented region.
iii. An expression depicting the roundness or the finishing of the edges of the sphere boundary of the segmented region.

The Chan-Vese method of segmentation will be able to correctly demarcate the borders of objects by refinement of previously defined level sets based for smoothing and differences of intensity to ensure that the shapes of the objects are captured. This method tends to be used in the analysis of medical images, detection of objects.

The segmenting process in Chan-Vese begins with a contour or a curve that roughly describes the part of the image that is to be segmented. This initial contour can assume the shape of a circle or rectangle, or any structure that fits its surrounding. The classification of Chan-Vese has the purpose of retrieving the contour that best approximates the shape of the observed object by minimizing a functional, which is associated to the contours and the regions. The Region-Based Term (Similarity Term), which is part of a broad functional, measures the difference in brightness between a set of pixels inside a contour and a set of pixels outside the contour. It promotes homogenous brightness of pixels just along the contour which is expected to enhance the uniformity within the separated region. Therefore, it causes what is inside the contour to differ in intensity level and increases the distance between what is being presented and what surrounds it. The region specific issue appears to be important in this case. The Chan-Vese segmentation algorithm iteratively modifies the contour to minimize the energy functional, with the goal of properly defining the object's boundaries within a picture. This iterative approach involves updating the contour based on the gradient of the energy functional, which is modified with specific energy terms that direct the contour towards the object's border. The algorithm maintains this iterative refining process until a stopping condition is met, such as finishing a certain number of iterations, reaching an energy threshold measurement, or the contour ceases from moving significantly. At convergence, the final contour reflects the fragmented object's boundaries and with pixels inside the contour line classified as target and those outside as ground [98].

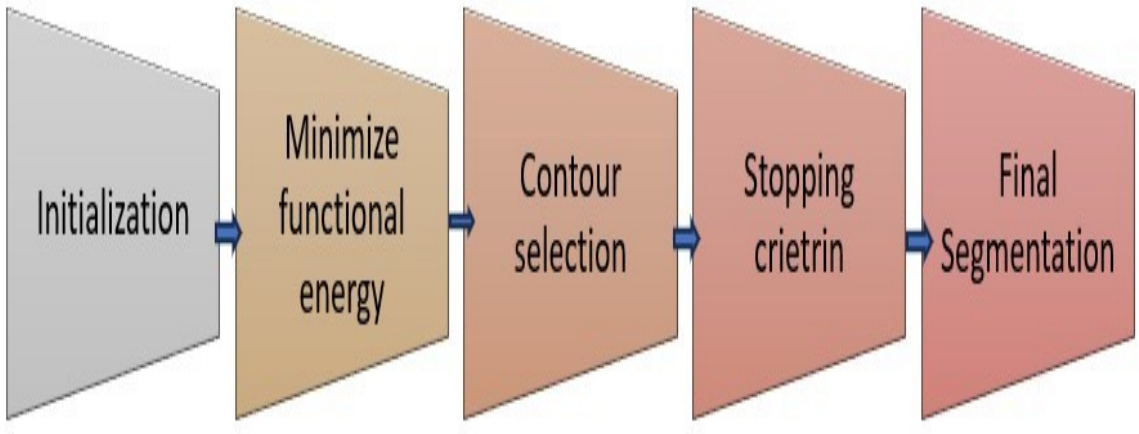


Figure 4.17. The Chan-Vese flow diagram

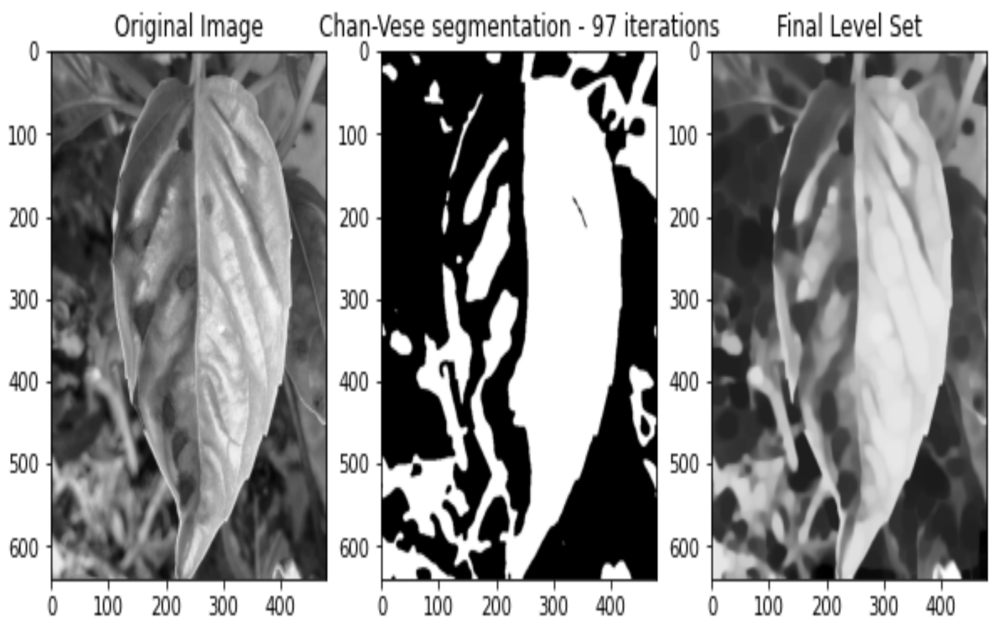


Figure 4.18. The Chan-Vese images for output

Chan-Vese segmentation provides various benefits:

i. It can successfully handle items with complex shapes and irregular edges.

ii. Unlike traditional active contours models, it is less affected by the contour's initial location.

iii. It works effectively with photos that have inhomogeneous intensities owing to shifting illumination.

However, there are limitations.

i. It may struggle to partition objects with similar intensities or significant texture differences.

ii. Level set techniques and numerical optimisation approaches are frequently utilised for Chan-Vese segmentation. This approach has been applied in medical image analysis, object tracking, and other domains that require extremely precise and resilient object categorisation from images.

4.7.3 Results and Discussions

The quality of segmentation is pivotal in crafting prediction models and the models are highly reliable and accurate when segregation is done with precision. Advanced models build precise binary masks that focus on key features and increase the model's segregation power, however poorly segmented images can impede accuracy as artefacts, noise or several mistakes might be embedded in the features when the Retrieval is performed. Depending on the task requirements and characteristics of the data, a decision must be made whether to delve into Otsu's thresholding and Chan-Vese segmentation. Otsu's thresholding deals with a situation in which the elements of interest must be distinctly defined in terms of their intensity which allows to achieve consistent results.

Contour-based concept: This phrase suggests, contour length should be minimized, but the smoothness should not be compromised. The subdivision ensures that it maintains homogeneity in intensity and that it is visible against the background. Chan-Vese segmentation method optimally seeks this term, thus, it allows constructing object contours by marked intensity changes across and outside the contour. The use of RIA for diagnosing agricultural diseases should tremendously improve agriculture through faster disease diagnosis and control, therefore reducing crop losses and ensuring proper growth of plants. However, sub optimal results have been. But, in order for the implementation to work in real agricultural scenarios, a lot of organization, adequate data collection and good model building have to be conducted. With regard to the type of data and to the needs of data, plant diseases can be classified by machine learning in many ways including effective and segmentation approaches. Such an approach ensures that the machine learning applications are specifically focused on the problems and goals pertinent to the agriculture resulting in better growth and output of the crop.

The identifying features of plant diseases can vary from simple thresholds moving on leaf's colour changes or discrete lesions to complex symptoms that require proficient segmentation approaches targeting the whole plant such as Chan-Vese and region growth algorithms. The determination of the segmentation mechanism is dependent on multiple factors which include the following:

i. Disease Symptoms: For diseases that have distinct visual evidence like lesions or colour changes, thresholding techniques are acceptable. On the other hand, areas of the plant that have a wider disease or do not exhibit strong symptoms will require advanced algorithms for segmentation.
ii. Labelled Data Availability: Other segmentation methods need to be labelled first through mask manuals for training and evaluation purposes which are expensive as

well as require a lot of time to obtain. It is important to ask if such model training labelled data exists.

iii. Computational Resources: Less powerful and resource demanding versions of segmentation techniques such as Chan-Vese should only be employed in virtual reality as they are extremely resource intensive. Make sure to take tabs on the computational resources available for various activities in picture processing.

In short, the segmentation techniques employed for the analysis of the plant diseases must be altered according to the peculiarities of the disease, the amount of labelled data and the computational feasibilities of the application. This holistic perspective ensures that the selected method is reasonable, useful and practical in agricultural settings.

Table 4.4. Otsu's Thresholding versus Chan-Vese segmentation

Factors	Otsu Thresholding	Chan-Vese
Purpose	Otsu's thresholding can be described as a method for performing binary image segmentation which classifies an image into two parts: a foreground and a background. The approach is effective when objects need to be separated from their backgrounds or when edges in an image are enhanced. Otsu's thresholding algorithm is suited to applications where accurate separation of the foreground and background is required, because it automatically determines the threshold for any given image based on the distribution of pixel brightness across the image.	Object segmentation problems come in different tasks, some are sophisticated and among them is the Chan-Vese segmentation specifically when dealing with objects with relatively ill-defined edges and a clear contour. This technique is effective for problems that require a precise segmentation of objects with complex geometry or with non-uniform intensity distribution across the object. In contrast to basic thresholding segmentation techniques, Chan-Vese segmentation can reliably identify intricate objects through successive refinement of the contour by alteration of the energy functional which is a combination of the no smoothness and the intensity of the contour. Such methods are useful when it is necessary to use more advanced segmentation to obtain more accurate object boundaries.

Methodology	Otsu's method efficiently determines a suitable threshold value which is obtained by maximizing the inter-class variance, which in turn minimizes the intra-class variance, of the pixel intensities. A single threshold is produced by this technique, which enables effective and efficient splitting of two classes in binary image segmentation, foreground and background. The threshold value is determined from the distribution of intensity values so as to maximize the separation between the object in the image and the background. This automated thresholding technique is ideal for the tasks requiring an easy distinction between the pixels belonging to the object and those that belong to the background.	The contour model based upon active contours used in Chan-Vese segmentation deforms with the iterations to fit the boundary of the object while minimizing an energy functional composed of the region based and the contour based terms. This successive approach is very accommodating of complex shapes of the objects and allows the correct tracing of the edges of the object more than the earlier approaches of segmentation. Chan-Vese segmentation on the other hand, for example, continually increasing energy functional combines several features smoothing and eliminating edge variance therefore making it possible for it to block out complex shapes and even deformable ones making it ideal for tasks requiring segmented images.
Complexity	For its ease of implementation and speedy results, Otsu's method is very efficient. In contrast to segmentation procedures like Chan-Vese, Otsu's method does not involve any iterative optimization techniques. Instead, the threshold is determined by the histogram of the image's intensities, which makes it straightforward to apply, and it doesn't require complex iterative approaches. Because of this efficiency and simplicity, it makes much more sense to use	In the case of Chan-Vese segmentation, the level set approaches or gradient descent optimization are available but they are much more costly than Otsu's thresholding. Chan-Vese's approach uses least squares estimates and was depicted as repeated practice necessitating many refinement stages to change the contour and decrease the energy functional. In time and efforts, this takes more than simple setting up a threshold value Otsu's approach is. Chan Vese Segmentation on the other hand is costly but allows for much greater

	Otsu's method where both speed and quality of binary image segmentation are required.	accuracy and control making it suitable for tasks that involve detailed segmentation in complicated images.
Accuracy	SVM- 97.76%, k-NN -99.54%, RF- 99.65%	SVM- 95.76%, k-NN -98.47%, RF- 98.65%

Disease classification algorithms accuracy is largely based on the outcome of the segmentation stages. In order to minimize wrong classification strokes, a segmentation technique which accurately delineates the diseased regions should be employed. When responsible for enigmas and lesions of varying shapes and approximately equal proportions, it is optimal to employ adaptive segmentation algorithms for structures with arbitrary shapes and sizes. As a matter of fact, the criteria for the selection of the segmentation approach are based on the disease categories being classified and characteristics of the dataset being used. For example basic thresholding can be adequate for images with clear disease features while Chan-Vese segmentation is appropriate for images which had minimal or complex signs of the disease.

4.8 Stacking Classifier Proposed Approach

Utilizing machine learning algorithms to study the multiple characteristics of leaves and search for any patterns related to certain diseases constitutes the stage towards constructing a model aimed at forecasting potential leaf diseases. Developing a model of forecast for leaf diseases involves the following main processes:

i. Data Collection: Create a repository of images of tagged leaves which documenting their normal or diseased views. Add leaves merged images depicting different types and stages of their diseases and health.
ii. Data Pre-processing: Remove noise, outliers, and missing values to make the dataset neater and more coherent. The leaf data set is built and prepared with the application of image transformations for example, scaling, normalization and others to enhance the data quality.
iii. Feature Extraction: Get pertinent characteristics of the leaf images through either manual picking of key features or the application of previously trained convolutional neural networks (MLP, for instance). Capture properties like color, texture, shape, and other relevant information [99].
iv. Model selection: Select an appropriate model for definition of artificial intelligence categories given the reconstructed features. Alternatives in this case include Decision Trees (DT), MLP, Random Forest (RF), SVM among others.
v. Model Training: Split the data collection into subsets for training and testing to ascertain the ability of the model. Train the model selected using the features gotten from the training set [100].

vi. Hyperparameter Tuning: Change the parameters of the model so that its performance can be improved. In order to find an optimal combination of hyperparameters, random or grid search techniques may be applied.
vii. Model Evaluation: Assess the effectiveness of the learned patterns using the testing set. The metrics for performance evaluation of the model include but not limited to accuracy, recall, precision and the F1 score.
viii. Iterate on the exemplary: by adjusting a set of measures and/or methods used in evaluation in environments as per feedback received. Continue optimizing the model to ensure a better output.
ix. Validation and Cross Validation: Employ cross validation techniques to check on the generalization capability and the validity of the model. Evaluate the model based on new data that has not been seen before to test its applicability in real life situations.

In averting and solving leaf diseases in agriculture, these structured approaches lead to the formulation of efficient and precise models that will enhance leaf disease prediction processes. The parameters of a stacking classifier model encompass numerous aspects for its structure and performance. The said parameters are clearly elucidated below:

4.8.1 Parameters of Base Classifiers:

i. Algorithm Specific Detailed Parameters: The parameters of kernel types in support vector machines (SVM), the number of trees in random forest, learning rates in k boosting gradient machines are examples of parameters that define each base classifier along the constituent algorithm.
ii. Hyperparameters: Those are key parameters as they form the overall behaviour of the basic classifiers and hence have a great impact on the performance of the model. A number of neurons in an artificial neural network, height of a decision tree and number of layers in a neural network.
iii. Feature Engineering Parameters: These include feature engineering activities undertaken on data received and before it is inputted into the base classifiers. This can include Feature selection algorithms.
iv. Transformation techniques (PCA, LDA etc.) Pre-Processing Configurations: Give instructions for the pre-processing of the data which include Normalisation and scaling techniques Image Data Augmentation Techniques [101].

4.8.2 Model Parameters for Stacking Method:

i. Stacking Technique: The type of the base classifiers informs the stacking technique to be used. Homogeneous Stacking: Every base classifier used is of same kind. Heterogeneous Stacking: There are various kinds of the base classifiers used.
ii. Multi-Layer Stacking: The stacking model consists of repeatedly stacking many layers of base classifiers.

iii. Aggregation Method: Determine how base classifier predictions are merged. Using basic classifier predictions as features in a meta-classifier. Weighted or simple averaging of projections

iv. Cross-Validation Configurations: Provide parameters for the cross-validation approach employed to evaluate and train models:
   a) Number of folds in cross validation
   b) Splitting approach (e.g., stratified or shuffled)

v. The stratification of the diversity in the ensemble: For the sake of achieving superior performance in the stacking model, it is necessary to ensure classifier diversity. This can be done through:
   a) The employment of several basic classifiers
   b) Differentiation in terms of hyper-parameter values and the feature engineering strategies of the basic classifiers [102].

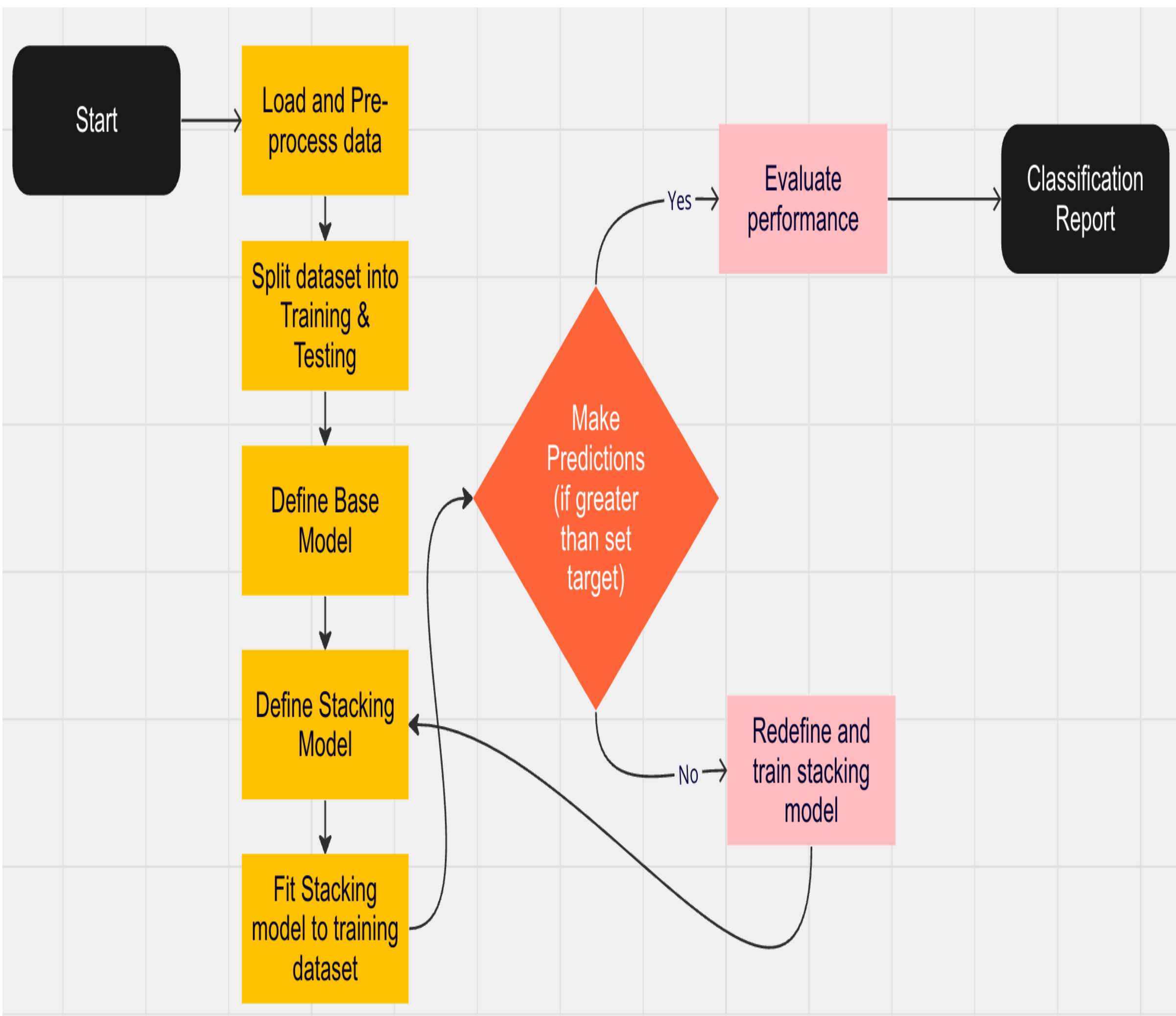


Figure 4.19. Flowchart: stacking classifier model

Those parameters, if controlled or supervised, can enhance and elevate the quality of stacking classifier models and improve prediction for diseased leaves. Factors such as feature behaviour, their performance, and the stacking model do play a very fundamental role in predicting models, more so, considering how models can range from weighted to basic averaging.

4.8.3 Classifier Model Comparisons

Below is a set of stacking classifier models which have been fit on a dataset of 20,000 tulsi leaves images broken down into four categories of bacterial, fungal, pests and healthy leaves. Out of the given data, 80% is used for training and 20% is allotted for testing. Thereafter, each classifier is trained separately on the respective input data. A combination of predictions from the given base estimators is used to create a stacking classification system. Further, the extra classifiers add one more layer to the final decision based on the integration of predictions provided by the base classifiers. This methodology is discussed in table 17 while figure 4.20 discusses the comparative results of five different models [103-105].

Table 4.5. Calculation of parameters of Classifier Model

Accuracy	Correctly classified data based on the total amount of data.	Accuracy = (TP + TN) / (TP + TN + FP + FN)	60%-70% poor 70%-80% good 80%-90% excellent 90%-100% Overfitting
Precision	Class predictions classified as positive; how many are positive.	Precision (P) = TP / (TP + FP)	0≤ P≤1
Recall	Correctly classified as True Positive based on the all-positive data.	Recall (R) = TP / (TP + FN)	0≤ R≤1
F-Measure	Single score/ Harmonic Mean for precision & Recall	F-Measure = (2 x P x R) / (P + R)	0≤ F ≤1
Specificity	True Negative Rate	Specificity=TN/(TN+FP)	0≤ S ≤1
ROC	Receiver Operating Characteristics	Plot showing Binary classifier as function of cut-off Threshold	AOC (0,0) to (1,1)

According to research, the type of data has cutting influence on the effectiveness of SVM and RF classification algorithms. SVM is very effective in high-dimensional spaces and is appropriate in cases when decision function is not always a single linear function. On the other hand, RF is very good at complex associations as well as interactions in the data. The combination of SVM and RF allows the model to embrace both types, linear and nonlinear, in order to understand the subtleties of the data quite well [108-115]. It was vital to run a variety of setups to determine which one was the best for this particular case. Quite excitingly, the Model B combination performed better than all others with respect to 98.25% of altogether accuracy as presented in Figure 4.21.

Table 4.6. Chosen parameters of Classifier Model

Proposed Prototypical	Assembling Estimators	Conclusive Estimator	Accurateness
Model A	Support Vector Machine, Random Forest	Random Forest	97.65%
Model B	Support Vector Machine, Random Forest	Support Vector Machine	98.25%
Model C	Support Vector Machine, Decision Tree, Random Forest	Decision Tree	95.38%
Model D	Support Vector Machine, Random Forest, Decision Tree	Convolutional Neural Network	97.12%
Model E	Support Vector Machine, Random Forest, Decision Tree, Convolutional Neural Network	Convolutional Neural Network	97.35%

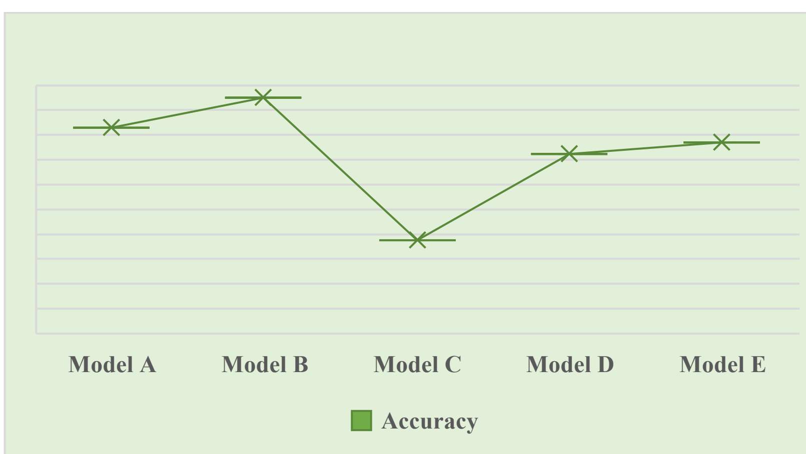


Figure 4.20. Accuracy comparison of model’s configuration

These two terms describe accuracy and precision which are key parameters for determining the success of a classification model as each captures a different aspect of the model's expectations. Accuracy equals the number of correct predictions positive and negative over the total number of predictions made [116-117].Interpretation parameter with high accuracy indicates that most of the expected outcomes will be closely matched with the real outcomes. This can, however, be noted as being a bit deceptive particularly in the instances where such data is unbalanced. For instance, if 95% of the data is about one class, that class will achieve 95% correct prediction rate by simply predicting that class. Precision evaluates the percentage of actual expected positives among the predicted positives. It shows the effectiveness of the model in predicting the positive class. It is critical for applications where the cost of a false positive is significant. For instance, in the case of infection diagnosis, it is necessary to eliminate any possibility that the model overestimates the value of a condition. Otherwise, it is more probable that an unnecessary treatment will be offered to somebody who is healthy. In the following charts, two models in an attempted precision evaluation are compared, as well as the accuracy and precision of all captured models and their standards [118-120]. Model A implements two strong ensemble/kernel-based approaches. Its accuracy is extremely high, which indicates very good performance on the task. The combination of SVM and Random Forest is most likely taking advantage of both SVM's usefulness in high-dimensional spaces and the ability of Random Forest to capture complicated interactions while preventing overfitting. With an accuracy of 98.25%, Model B achieves the best accuracy amongst all other Models. Like Model A, it utilizes Support Vector Machine and Random Forest. The accuracy increase of slightly over 1% compared to Model A (97.65%) indicates that perhaps the way these algorithms are applied, combined, or their hyperparameters are set differs enough as to result in this small but noticeable advantage. Given its lesser complexity and high accuracy, it stands out as a very strong performer. Compared to Model A, Model C adds a Decision Tree to the ensemble. Despite having more algorithms, its accuracy is lower than Model A. This may mean that the Decision Tree, in this specific ensemble, is perhaps adding some noise or otherwise failing to contribute positively to performance, or that the tuning on the combined model is not as good as Model A's. For the given dataset, the added complexity from the Decision Tree may not have been beneficial at all. We observe that Model D's accuracy of 97.12% is greater than that of Model E (97.35%) but lower than both Model A (97.65%) and Model B (98.25%). Both Model E and D share the same architecture of combining a Convolutional Neural Network with Support Vector Machine, Random Forest, and Decision Tree. It further reinforces the argument that while CNNs have a reputation for being powerful, their use is unnecessary if the problem at hand can be solved effectively using traditional ML algorithms, or if the CNN is poorly designed or trained. Model E includes all algorithms from Model C but adds a Convolutional Neural Network (CNN), making it the most complex model. Its accuracy of 97.35% is slightly lower than Model A's 97.65%. The inclusion of a CNN suggests that the data could be image-like or sequential in nature where CNNs perform best. Although Model E's accuracy is high, the additional complexity is not justified because the increase in performance over the simpler Model A is minimal.

Accuracy: 0.97

Classification Report:

	precision	recall	f1-score	support
bacterial	1.00	1.00	1.00	1024
fungal	0.96	0.97	0.96	977
healthy	0.97	0.97	0.97	995
pests	0.95	0.95	0.95	1004
accuracy			0.97	4000
macro avg	0.97	0.97	0.97	4000
weighted avg	0.97	0.97	0.97	4000

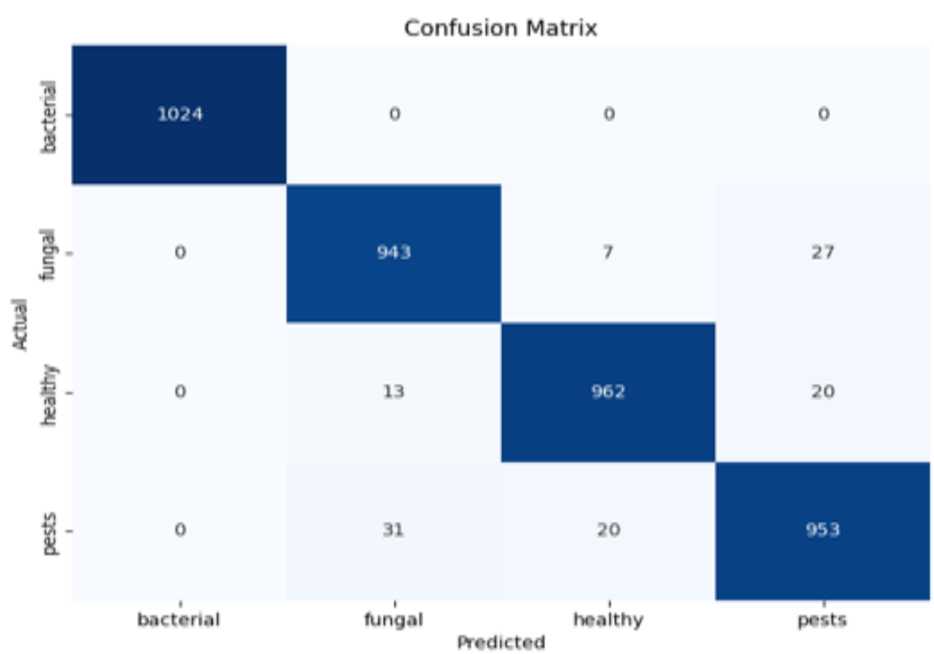


Figure 4.21. Model A Classifier result and confusion matrix

```
Accuracy: 0.98
Classification Report:
              precision    recall  f1-score   support

   bacterial       1.00      1.00      1.00      1024
      fungal       0.97      0.98      0.98       977
     healthy       0.97      0.98      0.97       995
       pests       0.96      0.95      0.96      1004

    accuracy                           0.98      4000
   macro avg       0.98      0.98      0.98      4000
weighted avg       0.98      0.98      0.98      4000
```

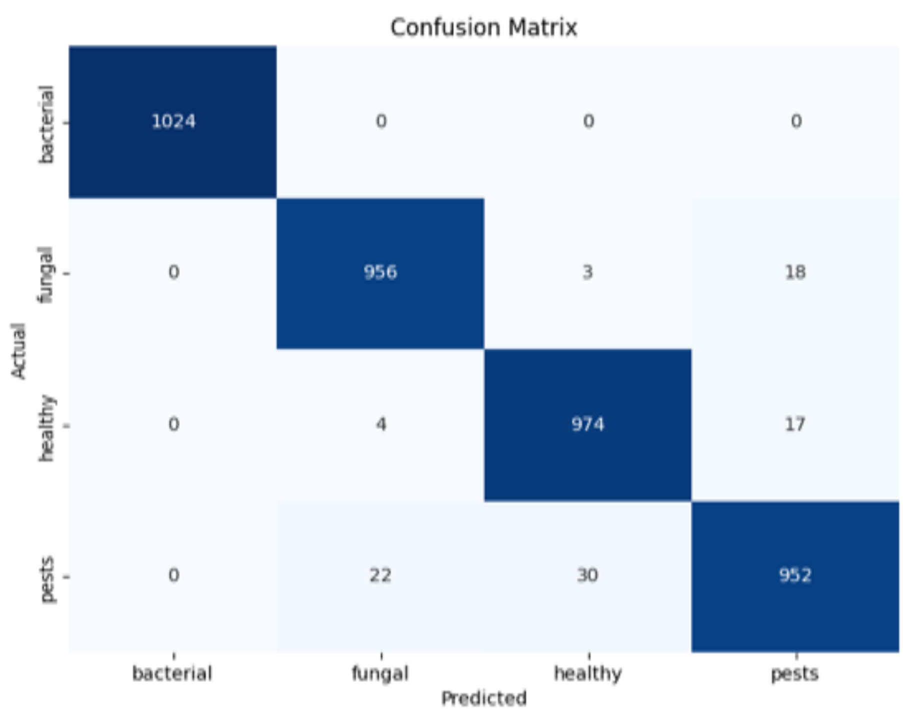


Figure 4.22. Model B Classifier result and confusion matrix

Accuracy: 0.95

Classification Report:

	precision	recall	f1-score	support
bacterial	1.00	1.00	1.00	1024
fungal	0.94	0.94	0.94	977
healthy	0.94	0.95	0.95	995
pests	0.93	0.93	0.93	1004
accuracy			0.95	4000
macro avg	0.95	0.95	0.95	4000
weighted avg	0.95	0.95	0.95	4000

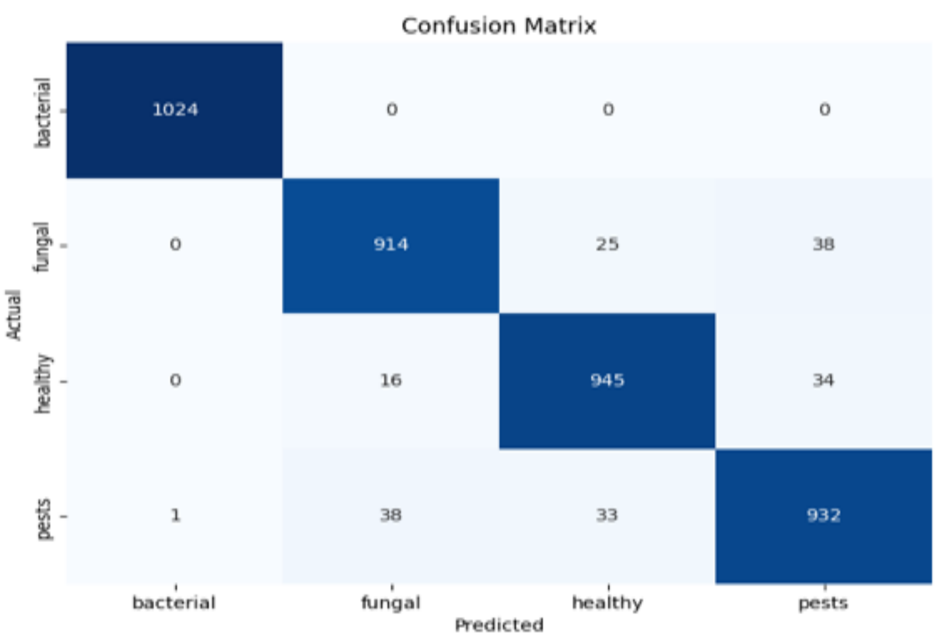


Figure 4.23. Model C Classifier result and confusion matrix

Accuracy: 0.97

Classification Report:

	precision	recall	f1-score	support
bacterial	1.00	1.00	1.00	1024
fungal	0.96	0.97	0.96	977
healthy	0.96	0.97	0.97	995
pests	0.97	0.95	0.96	1004
accuracy			0.97	4000
macro avg	0.97	0.97	0.97	4000
weighted avg	0.97	0.97	0.97	4000

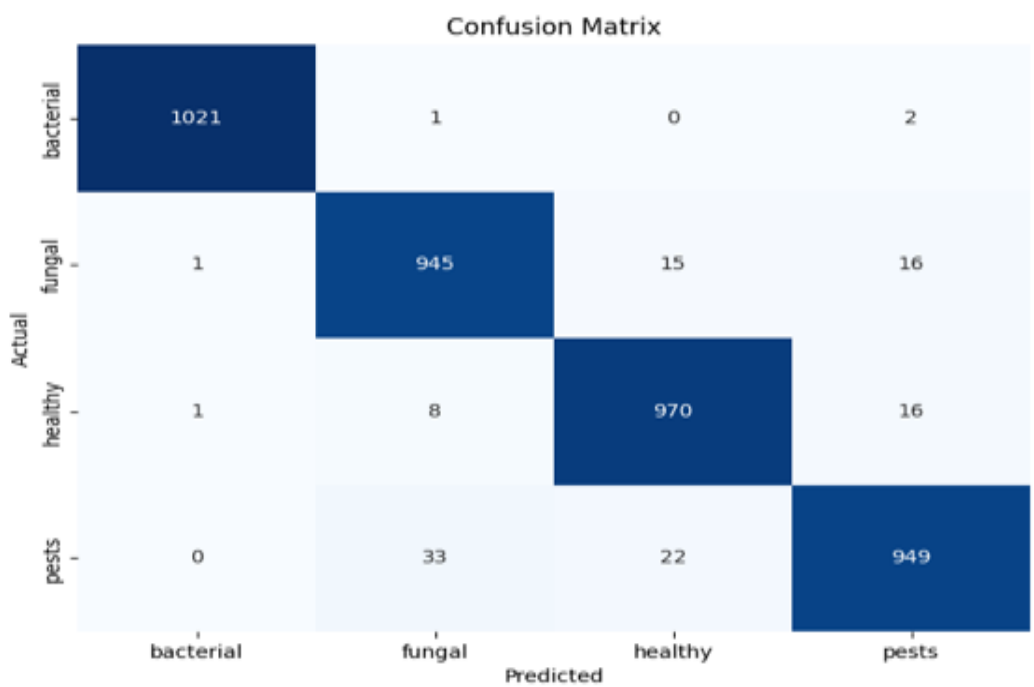


Figure 4.24. Model D Classifier result and confusion matrix

Accuracy: 0.97

Classification Report:

	precision	recall	f1-score	support
bacterial	1.00	1.00	1.00	1024
fungal	0.96	0.97	0.97	977
healthy	0.97	0.97	0.97	995
pests	0.96	0.95	0.96	1004
accuracy			0.97	4000
macro avg	0.97	0.97	0.97	4000
weighted avg	0.97	0.97	0.97	4000

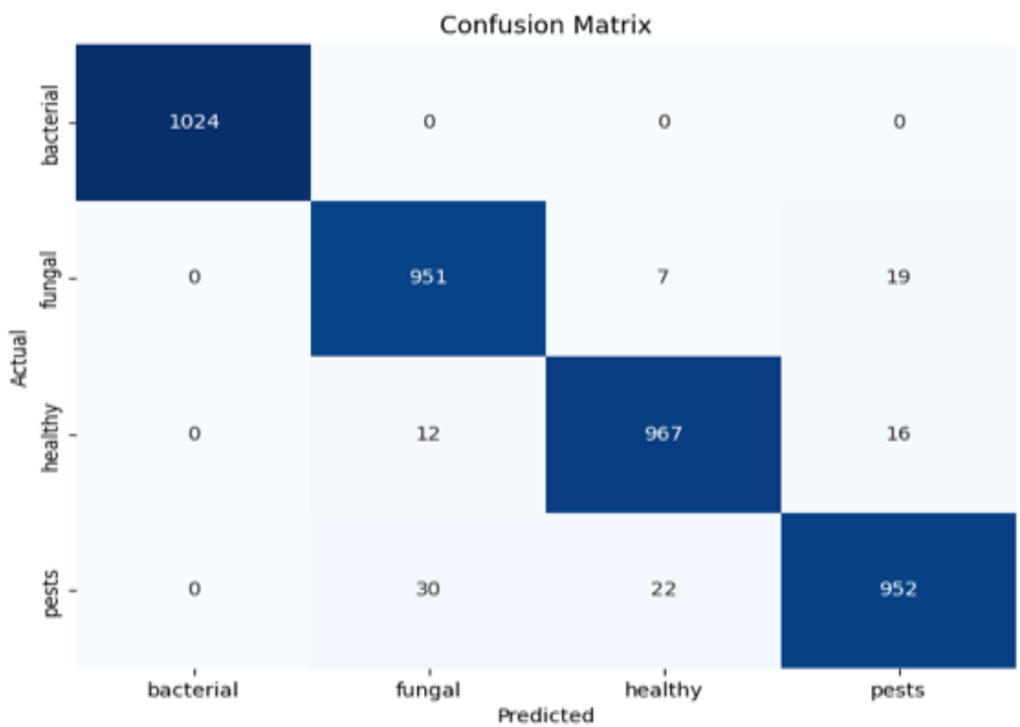


Figure 4.25. Model E Classifier result and confusion matrix

Table 4.7. Relating the classifier model to the current research on tulsi

Model	Stacking Estimators	Dataset	Accuracy
Model B (2024)	Stacking Estimators: SVM, RF Final Estimator: SVM	20,000 (4 classes)	98.25%
Qi et al. (2021)	SVM, Transfer Learning	7000 (7 classes)	SVM-97% Transfer learning-98%
Patil et al. (2022)	CNN	266 (2 classes)	75%
Sathiya et al. (2023)	CNN Inception V3 model	15220 (3 classes)	77.55%
Sathiya et al. (2022)	k-means clustering	1628 (3 classes)	0.59
Meenakshi et al. (2023)	Modified Logistic Regression	3,240 (3 classes)	84.23%
Lakshmanarao et al. (2024)	Hybrid Transfer Learning	8,000 (Medicinal leaves, 5 classes)	93.80%

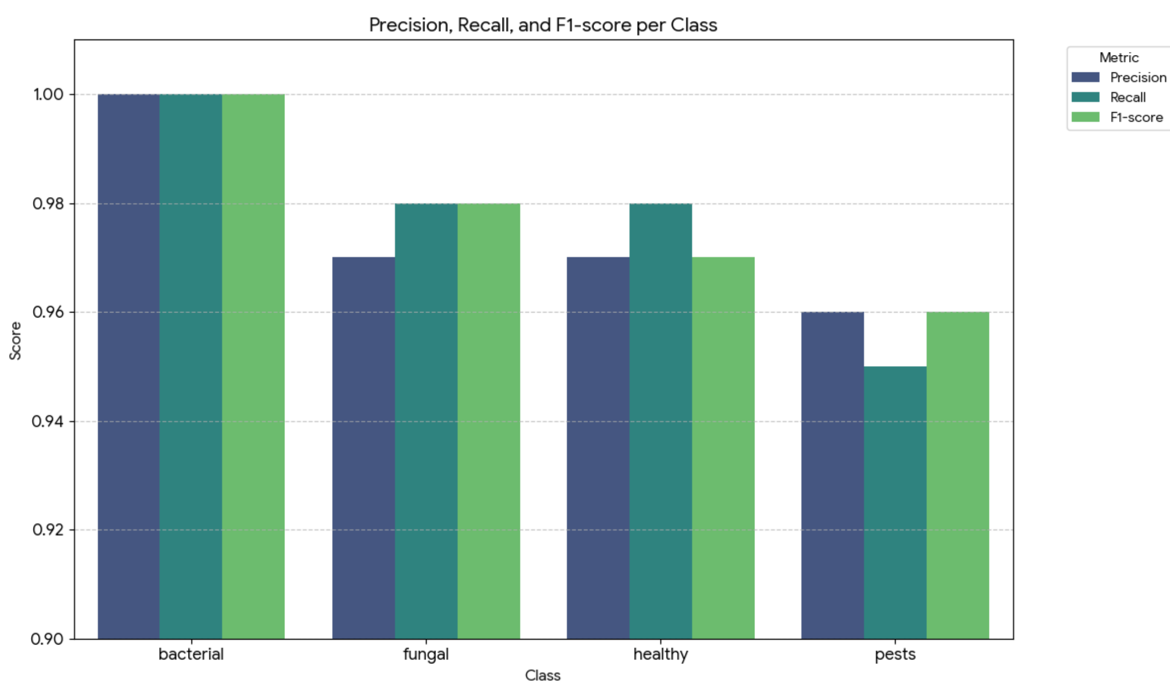


Figure 4.26. Class prediction of best model-B configuration

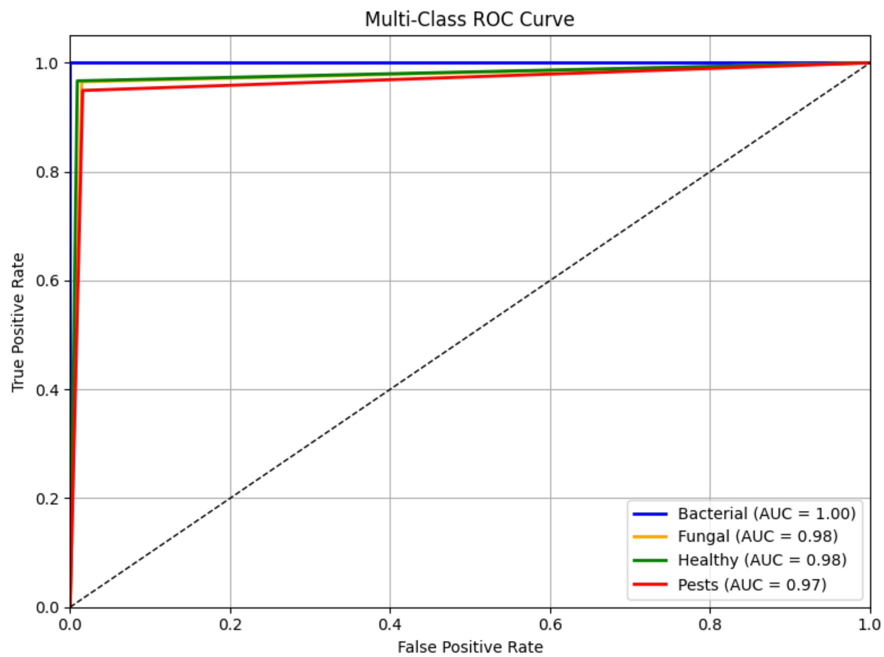


Figure 4.27. Model A ROC curve class analysis

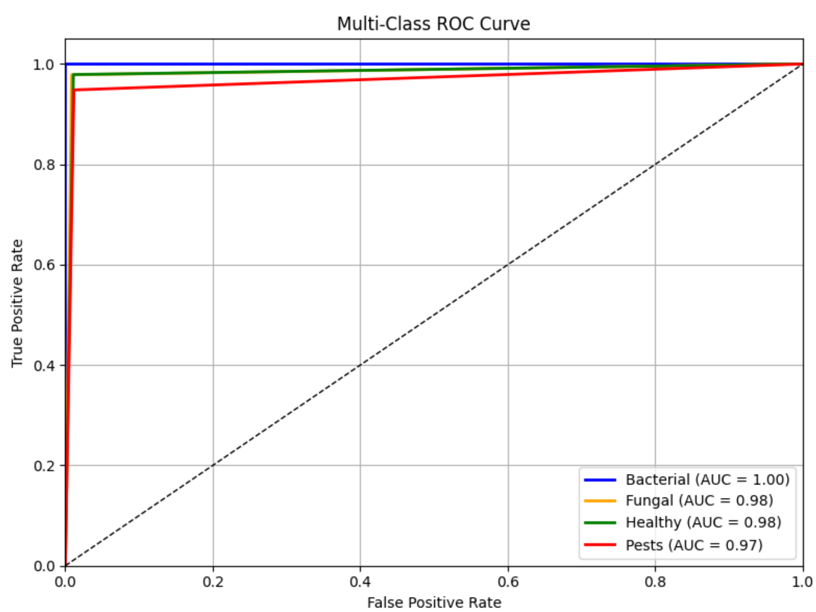


Figure 4.28. Model B ROC curve class analysis

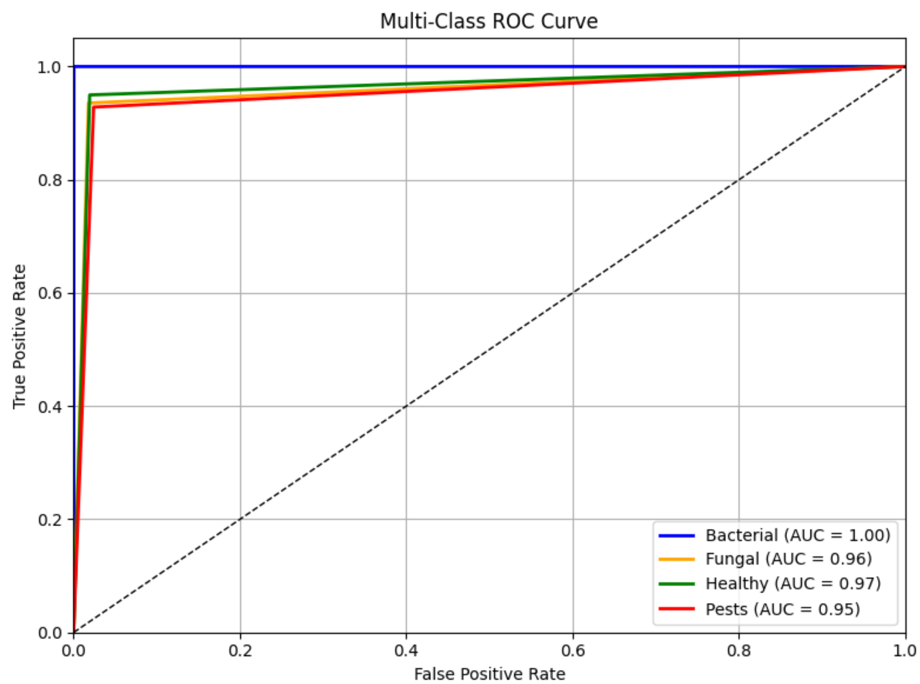


Figure 4.29. Model C ROC curve class analysis

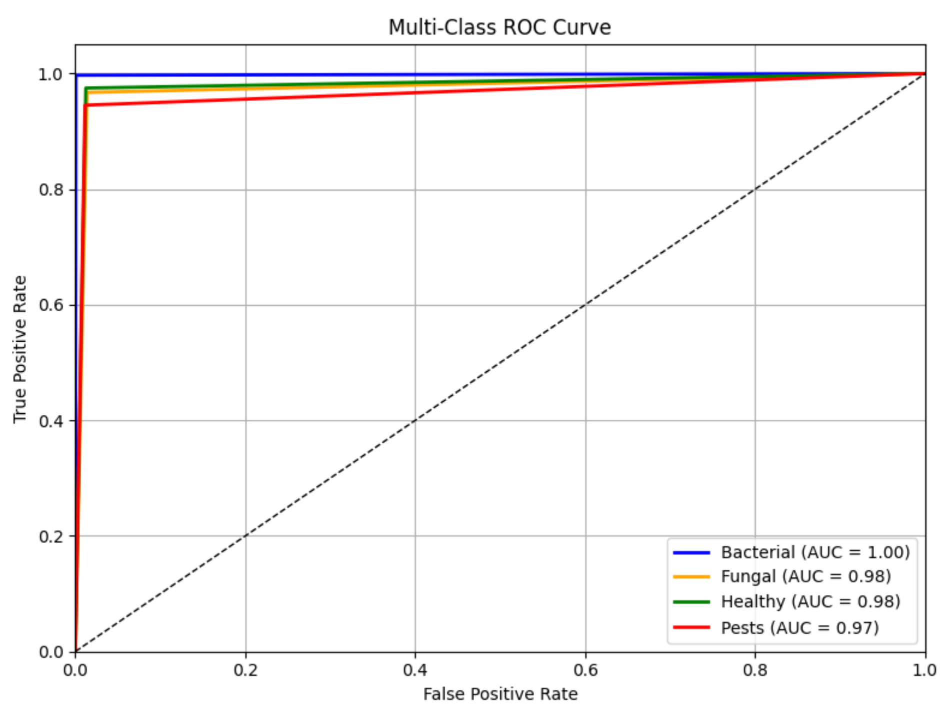


Figure 4.30. Model D ROC curve class analysis

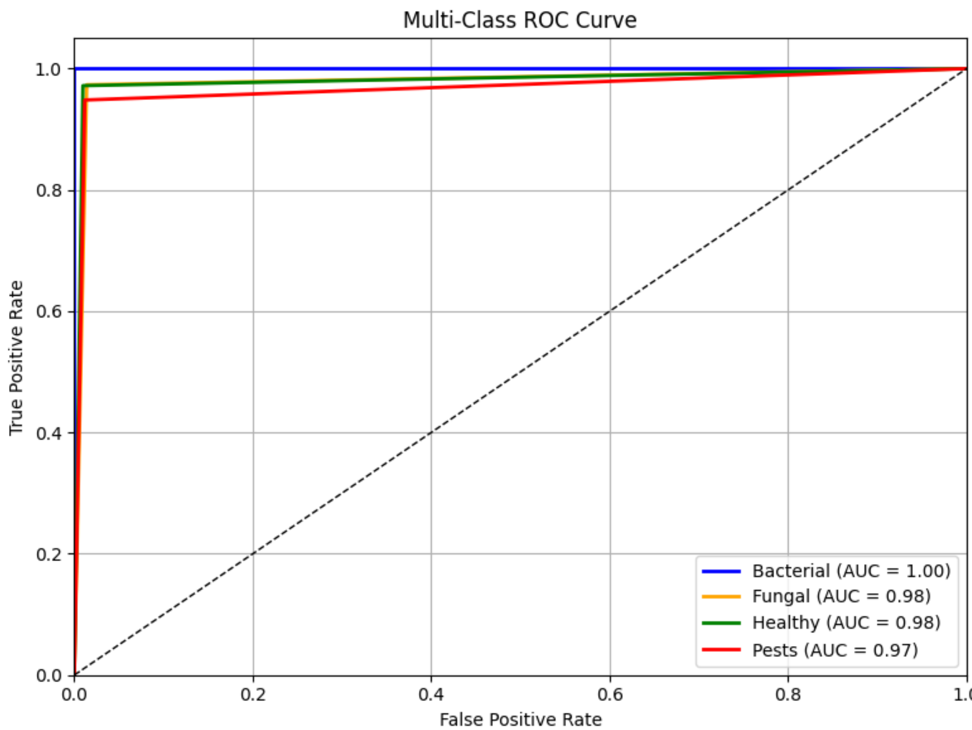


Figure 4.31. Model E ROC curve class analysis

Table 4.8. Performance analysis of designed models

Model	Class	Precision	Recall	F1-Score	Specificity
A	Bacterial	1	1	1	1
	Fungal	0.955	0.965	0.96	0.985
	Healthy	0.973	0.967	0.97	0.991
	Pests	0.953	0.949	0.951	0.984
	Macro Avg	**0.97025**	**0.97025**	**0.97025**	**0.99**
	Accuracy				**0.9705**
B	Bacterial	1	1	1	1
	Fungal	0.974	0.979	0.976	0.991
	Healthy	0.967	0.979	0.973	0.989
	Pests	0.965	0.948	0.956	0.988
	Macro Avg	**0.976**	**0.976**	**0.976**	**0.992**
	Accuracy				**0.9765**

C	Bacterial	0.999	1	0.9995	0.9997
	Fungal	0.9442	0.9355	0.9398	0.9821
	Healthy	0.9421	0.9498	0.9459	0.9807
	Pests	0.9283	0.9283	0.9283	0.976
	Macro avg	0.9534	0.9534	0.953375	0.984625
	Accuracy				**0.9538**
D	Bacterial	0.998	0.9971	0.9976	0.9993
	Fungal	0.9575	0.9682	0.9628	0.9861
	Healthy	0.9632	0.9749	0.969	0.9877
	Pests	0.9654	0.9452	0.9552	0.9887
	Macro Avg	0.971025	0.97135	0.97115	0.99045
	Accuracy				**0.9713**
E	Bacterial	1	1	1	1
	Fungal	0.9577	0.9734	0.9655	0.9861
	Healthy	0.97	0.9719	0.9709	0.9903
	Pests	0.9645	0.9482	0.9563	0.9883
	Macro Avg	0.97305	0.973375	0.973175	0.991175
	Accuracy				**0.9735**

As a result, the implementation of a model, including a collective results, and varies depending on the problem statement and dataset. To achieve resilience and dependability, the model's performance has been rigorously validated using distinct test sets and considering real-world application scenarios.

The model continuously functioned effectively across all evaluations, demonstrating the ensemble setup and models used are well-suited to tackling the specific problem at hand. This validation emphasizes the applicability and effectiveness of the strategy used to solve the specific challenge [108].

4.9 Discussions

The most powerful and upgrades improvement in recognition and description of leaf diseases is the union of AI and image processing technology. We have observed an increasing amount of attention being directed to this topic over the past 25 years but still advancement in this area has been rather sluggish to materialize. Several things influence this observation:

i. Disease recognition: This is a critical step, especially for plants such as Tulsi as infections in the leaves have to be accurately classified. The foremost efforts in identification should focus at automation of disease diagnosis and severity assessment.

ii. Dataset Quality: The limited availability of sufficient and high-quality datasets for both training and testing is a major bottleneck. Better results could be obtained instead by implementation ensemble techniques or examining other sets of data.
iii. Computational Complexity: A key requirement for any effective implementation is to be able to deal with and surmount the primary computational complexity and cost of these tasks.
iv. Implementation Efficiency: The results should be put into action in an efficient manner to be able to apply and incorporate the solutions within a more reasonable time frame.
v. Classification Models: There are improvements to be made by combining multiclass classifiers along with hybrid feature analysis and segmentation methods in an attempt to improve classification accuracy and performance.
vi. Performance Metrics: What can also be improved are metrics such as recognition rate, precision and overall classification accuracy. The three parameters applied in combination have the potential to render more uniform and reliable outputs.

In conclusion, even if some progress has been achieved with regard to machine intelligence, condition testing and computer vision technologies, these areas still continue to lag in their real time application endeavours. In solving the stated challenges, most of the approaches will be theoretical as well as the application of new strategies in the area of leaf infection development.

i. Precision agriculture improves disease diagnosis and management in crops by accurately classifying plant health situations.
ii. Plant Pathology Research: Helping researchers detect and understand diverse plant diseases.
iii. Crop management aims to reduce disease spread and optimise crop yield and quality.
iv. Integrating with automated farming systems allows for real-time disease monitoring and intervention.
v. Environmental monitoring: Going beyond agriculture to track and control plant health in natural settings.

Generally speaking, the studies devoted to stacking classifier models hold promise for increasing the efficiency of precision agriculture and plant pathology and thus for the development of more effective and sustainable agriculture practices [121-122]s.

i. Improving Model topologies: Check neural network architectures and model types of base models and the final estimator Model (CNN) to increase the predictive robustness of the model especially during the test period by optimization of topologies and hyperparameters tweaking.
ii. Data Expansion and Enhancement: Include various data sources and use data augmentation techniques to improve the dataset in terms of diversity as well as robustness.

iii. Real Time Monitoring System: Connecting the combined classifier model in an automated monitoring system for the Tulsi plants enables the farmers to receive plant health updates at necessary times to enable them get ready for imminent infections.
iv. Improving the understanding and the level of trust of the model by means of interpretable properties ensures that the outsiders are equipped with key features which are associated with various classes of infection.
v. Integration with Precision Agriculture Technologies: With the help of satellite images or other IoT based sensors, incorporating the model of stacking classifiers can provide a more integrated approach of monitoring and well-functioning of crops.
vi. Working closely with farmers and agricultural specialists during the validation and deployment phases guarantees that the model is realistic and applicable in real-world agricultural settings.

In sum, the development of a stacking classification model has potential in predicting tulsi plant infections. Improvement in task specific model development, model training and real time farming deployment will make the work more beneficial to farming and crop health management. Implementing the stacking classifier algorithm may greatly improve presumptions and appreciably assist the practice of precision agriculture and implementation of application of pests and diseases to Tulsi leaves. To enhance the prediction, structural optimization of the constituent models and meticulous calibration of the hyper parameters both for base and final CNN estimator. Further augmentation of the dataset by including more images with angle differences and image task enhancing can increase the strength of this approach. Applying this technique as part of a monitoring system for the Tulsi crops would ensure rapid response to agricultural diseases. Providing interpretability of the models, fusion with precision agriculture techniques such as satellite images and IoT sensors, and collaboration with farmers will make the model applicable in reality and for the management of crops.

Hence, this chapter described the implementation of methodological framework adopted for the classification of Tulsi (Ocimum sanctum) leaf diseases based on a strong ensemble-based approach. The methodology was organized in terms of data preprocessing, feature extraction, and deployment of several ensemble models to enhance classification accuracy and generalization. The ensemble stacking approach utilized the strengths of traditional machine learning algorithms and deep learning models complementarily. Interestingly, Model B showed the best accuracy (98.25%), which represents the efficiency of SVM as a decisive estimator in association with conventional ensemble learners. The methodological platform described in chapter 3 offered a safe, scalable, and high-precision framework specifically designed for Tulsi leaf disease classification and serves as the basis for the performance testing covered in this chapter.

Chapter 5
Summary and Conclusions

5.1 Synthesis of Results

This thesis work involves the model formulation and implementation of the disease prediction system for Tulsi (Ocimum sanctum) plants. The focus of the study was to meet the great need for early detection of diseases in the cultivation of Tulsi so as to curb losses in the production and also manage the crop throughout the farming seasons. The study employs modern technologies such as computer vision and machine learning to automate disease prediction using leaf images as the most important diagnostic feature. Various image processing methods, feature extraction methods and machine learning algorithms were examined in order to increase the accuracy and the efficiency of disease diagnosis. The disease prediction and classification model was aimed at diagnosing and classifying the plant diseases visible on the tulsi plant by systematically modelling the major symptoms of the diseases on the tulsi plant. The extent of having a strong predictive model claimed correct passive identification and classification of some of the common diseases which quickly affected Tulsi plants through observing the appearance of the leaves. The results of the study have some significant implications for advancing precision agriculture and monitoring systems for growing and development of plants including expectations for researchers, practitioners and other stakeholders involved in Tulsi growing and development of agricultural technology.

In the concluding remarks, we stated that the tulsi plants disease prediction model is one of the major milestones in agricultural technology trends focusing on increased crop security and yield. Successful application of computer vision and science machine learning algorithms attested to the prospects of machine-based disease detection based on leaves images which cuts across the preventive measures and treatment options. The work was recognized as finding significant challenges and possibilities in forecasting plant diseases with an emphasis on the interdisciplinary nature of the solution. Some of the areas that may be pursued in the future include improving and fine tuning the models for prediction, utilising the models along with the corresponding systems for live observations and extending the technology to cover other farming practices. Lastly, the approach to the prediction of disease occurrence and spread among the Tulsi plants proposed within the framework of this thesis report is likely to change the prevailing concepts in the management of the relevant diseases by providing farmers with efficient tools for accurate diagnosis of the disease and consistent management of it, thus enhancing sustainability and robustness of agricultural systems. In conclusion, we mention that the algorithm for predicting infections specially designed for Tulsi plants is indeed a great achievement in the agricultural sciences, providing applicable mechanisms of improving the health and production of crops.

Automation of imagery based intervention and management techniques along with accurate disease diagnosis has been made possible by the combination of computers, vision, and machines knowledge. These findings of the study stress out the significance of integrated approach in agriculture, technology, and data sciences. By employing advanced technologies, researchers and practitioners are in a position to enhance precision agriculture practices and to promote more eco-friendlier farming methods. Future research should involve fine-tuning and optimizing the model of predicting diseases, augmenting its functionalities with real-time monitoring systems, as well as integrating it into larger agricultural technology.

The findings of the training demonstrate the importance of underdisciplined in agriculture, technology, and data science. Utilizing modern technologies, researchers and practitioners should be able to improve precision agriculture practices and market more environmental friendly farming methods. Future research activities involve the refinement and optimization of the disease prediction model, its expansion with real time monitoring systems, and its incorporation with other agricultural technologies.

These improvements will provide farmers with better facilities for early illness surveillance and advanced disease management, thereby enhancing the resilience and sustainability slash security of agriculture. We also believe that as these methods get re-evaluated and modified, they will lead to more successful and efficient efforts in the management of crop health that are beneficial to agriculture, researchers and other concerned parties in using viable agricultural production systems.

The primary objective of the research is to design an early sickness detection model focussing on Tulsi production systems This aspect is very important in the reduction of losses due to output and in efficient management of the health of crops This study integrates state of the art technologies including Ahn et al 2019:

i. Computer Vision: This technology enables the model to examine images of leaves to identify possible visible signs of diseases.
ii. Machine Learning: Several AI techniques are used to provide automated predictions of diseases based on data from images of the leaves taken.

The study looks at different applicable interventions associated with

i. Image Processing Techniques: These are techniques that increase the clarity of the images so that they can be easily examined.
ii. Feature Abstraction Methods: These techniques are aimed at finding salient features in leaf images which indicate the presence of disease.

With this systematic approach, the study formulates a robust prediction algorithm that would assist in accurate diagnoses and classification of most common major diseases afflicting Tulsi plants by looking at images of the leaves. The results of the study have a meaningful impact in the area of precision agriculture. By macro standing over the disease prediction task, the technique enhances the chances of accurately tracking the health of the plant. The research provides significant insights into Scholars can use the results to formulate and enhance disease prediction algorithms.

i. The model provides farmers and agricultural professionals with a better disease management strategy in the Tulsi farming.
ii. The information offered will be useful to people and entities that have investing interests in agricultural technology and production.

Agricultural procedures considering the findings of the research may be enhanced, especially in relation to the management of Tulsi plants. Eco-friendlier variables are promoted because it provides early detection of diseases thus enhancing productivity and minimizing losses The entire work also stresses the role of technology into agricultural development, owing to its importance into particular diseases control practices such as for the case of Tulsi crops, which in turn has both economic and environmental consequences on the farming activities.

Such changes will develop new possibilities for agriculture through the creation of tools for the DTs' diagnosis, disease surveillance, and disease management enhancing DTs farming, agriculture, and eco-systems resilience. If we continue inventing and developing these technologies, we can lay the foundations for new approaches for crop management that would enable farmers, researchers, and other stakeholders to effectively solve the problems that crop farming efforts face while enhancing productivity.

5.2 Directions for Future Work

Although the present study was able to effectively prove the efficacy of ensemble stacking models for Tulsi leaf disease classification, there are many avenues to push and deepen this research to expand its practical relevance and scientific value.

i. Increasing Dataset Diversity and Size-The validity and applicability of machine learning algorithms largely rely on the size and diversity of the training dataset. For future work, one can attempt to gather a larger and more diverse dataset of Tulsi leaves, in terms of various cultivars, growth stages, and images obtained in diverse environmental and lighting conditions. This would improve model robustness and performance in real-world situations.

ii. Development for Real-Time and Mobile Applications-Implementing the suggested model on a user-friendly mobile app or an edge computing device would allow real-time disease detection for farmers and herbalists. This would help in early intervention as well as minimize crop loss. Optimization methods such as model pruning or quantization could be utilized to minimize computational needs for such implementation.

iii. Multi-Disease and Multi-Plant Classification-The present work centered mainly on binary classification of healthy vs. diseased Tulsi leaves. Future research could extend this to multi-class classification of various diseases that impact Tulsi, and other medicinal herbs. This would give a complete diagnostic tool to the agricultural and pharmaceutical sectors.

iv. Integration of Explainable AI Techniques-To increase trust and interpretability in automated disease diagnosis, including explainability techniques like Grad-CAM, SHAP, or LIME would assist in visualizing what regions of the leaf impacted the model's predictions. This is especially important for end-users who might need expert validation.

v. Cross-Domain Transfer Learning-Utilizing knowledge from extensively researched crops (such as tomato, potato) with transfer learning might hasten the process to build reliable models for medicinal plants. Studies might investigate domain adaptation methods for the transfer of learned features from large-scale datasets onto smaller Tulsi datasets.

vi. Integration of Environmental and Soil Conditions-Future models may be enhanced by the inclusion of additional data such as environmental conditions (temperature, humidity) and indicators of soil quality, which affect the formation of disease. Multimodal data fusion would enable more precise and context-aware disease prediction.

vii. Collaboration with Agricultural Experts-Collaboration with botanists and plant pathologists for model prediction validation and the enhancement of labelling accuracy would enhance the scientific strength and practical utility of future research

Reference

[1] Bhamra, S., Heinrich, M., Howard, C., Johnson, M., & Slater, A. (2015). DNA authentication of tulsi (Ocimum tenuiflorum) using the nuclear ribosomal internal transcribed spacer (ITS) and the chloroplast intergenic spacer trnH-psbA. *Planta Medica*, 81(16).

[2] Gautam, M. K., & Goel, R. K. (2014). Toxicological study of Ocimum sanctum Linn leaves: hematological, biochemical, and histopathological studies. *Journal of toxicology*.

[3] Hemalatha, R., Babu, K. N., Karthik, M., Ramesh, R., Kumar, B. D., & Kumar, P. U. (2011). Immunomodulatory activity and Th1/Th2 cytokine response of Ocimum sanctum in myelosuppressed swiss albino mice. *Trends Med Res*, 6, 23-31.

[4] Jamshidi, N., Cohen, M.M.(2017). The Clinical Efficacy and Safety of Tulsi in Humans: A Systematic Review of the Literature. *Evidence-Based Complementary and Alternative Medicine* 9217567, 13. https://doi.org/10.1155/2017/9217567.

[5] Kulkarni, K.V., Adavirao, B.V. (2018). A review on: Indian traditional shrub Tulsi (*ocimum sanctum*):The unique medicinal plant. *Journal of Medicinal Plant Studies* 6(2),106-110.

[6] Mohan, Lalit, M. V. Amberkar, and Meena Kumari. (2011). Ocimum sanctum linn.(TULSI)-an overview. *Int J Pharm Sci Rev Res* 7.1, 51-53.

[7] Roopashree, S.,Anitha, J. (2020). Enrich Ayurveda knowledge using machine learning techniques. *Indian Journal of Traditional Knowledge 19*(4), 813-820.

[8] Tiwari, A. K., Mishra, R., Chaturvedi, A., Sadan, A., Arogyadham, J. R. D., & Ayurveda, I. (2014). Traditional uses of Ocimum sanctum (Tulsi). *Int. J. Glob. Sci. Res*, 1(2), 126-131.

[9] Suryawanshi, S. K., & Chouhan, U. (2018). Computational approaches for the prediction of antimicrobial potential peptides from Ocimum tenuiflorum. *Asian J. Pharm. Clin. Res*, 11, 398-401.

[10] Lavanya, V., Ganapathy, D., Visalakshi, R.M. (2019). Antioxidant and free radical scavenging activity of Ocimum basilicum - An in vitro study. *Drug invention Today* 12(5), 1004-1007.

[11] Shah, G.-M., Ahmad, F., Panwar, S., Khadka, M. S., Ali, A., & Bisht, S. (2019). Ocimum Sanctum [Tulsi]—An Alternative Additional Livelihood Option for the Poor and Smallholder Farmers. *Sustainability* 11(1), 227. http://dx.doi.org/10.3390/su11010227.

[12] Rhoades, H. (April 2021). *Diseases And Problems With Growing Basil.* https://www.gardeningknowhow.com/edible/herbs/basil/basil-diseases.htm.

[13] Gowans M. (April 2021). *Brownish Yellow Spots on a Holy Basil Plant*. https://homeguides.sfgate.com/brownish-yellow-spots-holy-basil-plant-71343.html

[14] Sharma, Suresh & Verma, Deepak & Kumar, Yogesh & Kumar, Sanjay & Srivastava, Ramesh. (2019). Comparative economic study of tulsi and paddy cultivation in Sitapur district of Uttar Pradesh. *Agricultural Science Digest - A Research Journal*. doi:10.18805/ag.D-4890.

[15] Singh, P., and Prakash, D. (2014). Disease management in Tulsi (Ocimum sanctum L.) – A Holy medicinal plant. International Journal of Current Microbiology and Applied Sciences, 3(10), 106-115.

[16] Padalia, R.C., Verma, R.S., and Chauhan, A. (2011). Chemical composition of essential oil from different parts of Ocimum sanctum L. cultivated in India and its antimicrobial activity.

[17] Singh, S.K., Patra, D.D., and Mishra, S.K. (2013). Disease management in Tulsi (Ocimum sanctum L.): A review. International Journal of Advanced Research, 1(8), 236-246.

[18] Pandey, A.K., Singh, P., and Tripathi, N.N. (2003). Chemistry and bioactivities of essential oils of some Ocimum species: An overview. Asian Journal of Chemistry, 15(4), 1955-1963.

[19] Singh, S., Singh, A., and Gaurav, S.S. (2012). Antifungal activity of different essential oils against fungal pathogens of Ocimum sanctum L. and their phytochemical analysis. International Journal of Drug Development and Research, 4(3), 165-172.

[20] Bharti, P., and Sengar, R.S. (2015). Detection of leaf spot disease of Tulsi (Ocimum sanctum L.) and its management strategies. International Journal of Agriculture Sciences, 7(2), 302-305.

[21] Tandon, S., and Dixit, S.N. (1991). Leaf spot disease of Tulsi (Ocimum sanctum L.) and its management. Indian Phytopathology, 44(2), 253-255.

[22] Verma, R.S., Rahman, L., and Verma, R.K. (2010). Chemical composition of Ocimum sanctum L. essential oil and its efficacy as a fungitoxicant against storage fungi. International Journal of Agriculture, Environment and Biotechnology, 3(3), 379-384.

[23] Sah, A.N., Dixit, S.N., and Tandon, S. (1985). Fungi associated with the leaf spot disease of Tulsi (Ocimum sanctum L.). Current Science, 54(24), 1298-1299.

[24] Padalia, R.C., Verma, R.S., and Chauhan, A. (2012). A review on Ocimum sanctum (Tulsi). International Journal of Pharmaceutical Sciences Review and Research, 13(1), 91-96.

[25] Prakash, P. A. G. N., & Gupta, N. (2005). Therapeutic uses of Ocimum sanctum Linn (Tulsi) with a note on eugenol and its pharmacological actions: a short review. *Indian journal of physiology and pharmacology*, 49(2), 125.

[26] Anami, B.S., Nandyal, S.S., Govardhan, A. (2010). A combined color texture and edge features-based approach for identification and classification of Indian medicinal plants. *International Journal of Computer Applications* 6(12), 45-51.

[27] Pattnayak, P., Behera, P., Das, D., & Panda, S. K. (2010). Ocimum sanctum Linn. A reservoir plant for therapeutic applications: An overview. *Pharmacognosy reviews* 4(7), 95–105. https://doi.org/10.4103/0973-7847.65323.

[28] Valliammal, N., Geethalakshmi , S.N. (2011). Automatic recognition system using preferential image segmentation for leaf and flower images. *Comput. Sci. Eng. Int. Journal* 1(4), 13-25.

[29] Tewari, D., A N, S., H K, P., H S, M. (2012). A Review on Phytoconstituents of Ocimum (Tulsi). *International journal of Ayurvedic medicine.* 3(1), 1-9. doi:1-9. 10.47552/ijam.v3i1.103.

[30] Jabal, M. F. A., Hamid, S., Shuib, S. & Ahmad, I. (2013). LEAF FEATURES EXTRACTION AND RECOGNITION APPROACHES TO CLASSIFY PLANT. *Journal of Computer Science* 9(10), 1295-1304. https://doi.org/10.3844/jcssp.2013.1295.1304.

[31] Kumar, A., Rahal, A., Chakraborty, S., Tiwari, R., Latheef S.K.,Dhama, K. (2013). *Ocimum sanctum* (Tulsi): a miracle herb and boon to medical science - A Review. *International Journal of Agronomy and Plant Production* 4(7), 1580-1589.

[32] Larese. M.G., Namías, R., Craviotto, R.M., Arango, M.R., Gallo C., Granitto, P.M. (2014). Automatic classification of legumes using leaf vein image features. *Pattern Recognition* 47 (1), 158-168.

[33] Nilamudeen, M., Girija V.K., Nandakumar, C. (2013). Colletotrichum leaf spot of holy basil (Ocimum tenuiflorum L.)- A new report from Kerala. *Progressive Horticulture* 45(2), 383-384.

[34] Ananthi, C., Periyasamy, A., & Muruganand, S. (2014). Pattern Recognition of Medicinal Leaves Using Image Processing Techniques. *Journal of Nanoscience and Nanotechnology*, 2(2), 214-218.

[35] Cohen, M.M. (2014). Tulsi - Ocimum sanctum: A herb for all reasons. *Journal of Ayurveda & Integrative Medicine* 5(4), 251-259.

[36] Dubey, S., Jalal, A.S. (2015). Application of Image Processing in Fruit and Vegetable Analysis: *A Review. Journal of Intelligent Systems*, 24(4), 405 - 424. doi:10.1515/jisys2014-0079.

[37] Gomathi, M. & Vijayashree, T. (2015). Development of novel neural networks for the identification of types of Tulsi leaves. *Research Journal of Pharmaceutical, Biological and Chemical Sciences* 6 (3). 1252-1259.

[38] Vijayashree, T., Gopal, A. (2015). Classification of Tulsi Leaves Based on Texture Analysis. *Middle-East Journal of Scientific Research 23 (Sensing, Signal Processing and Security).* 102-107.

[39] Wang, Z., Sun, X., Zhang, Y. (2016). Leaf recognition based on PCNN. *Neural Computing & Applications* 27, 899–908. https://doi.org/10.1007/s00521-015-1904-1.

[40] Yun, S., Xianfeng, W., Shanwen, Z., Chuanlei, Z. (2015). PNN based crop disease recognition with leaf image features and meteorological data. *International Journal of Agricultural & Biological Engineering* 8(4), 60-68. doi: 10.3965/j.ijabe.20150804.1719

[41] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience* 2016, 3289801, 11. https://doi.org/10.1155/2016/3289801.

[42] Tsaftaris, S., Minervini, M., Scharr, H. (2016). Machine Learning for Plant Phenotyping Needs Image Processing. *Trends in plant science* 21(12), 989–991. https://doi.org/10.1016/j.tplants.2016.10.00.

[43] Zhou, C., Ye, H., Xu, Z., Hu, J., Shi, X., Hua, S., Yue, J., et al. (2019). Estimating Maize-Leaf Coverage in Field Conditions by Applying a Machine Learning Algorithm to UAV Remote Sensing Images. Applied Sciences, 9(11), 2389. http://dx.doi.org/10.3390/app9112389.

[44] Muneer, A., Fati, S.M. (2020). Efficient and Automated Herbs Classification Approach Based on Shape and Texture Features using Deep Learning. *IEEE Access 8*, 196747-196764. doi: 10.1109/ACCESS.2020.3034033.

[45] Ganatra, N., Patel, A. (2020). A Multiclass Plant Leaf Disease Detection using Image Processing and Machine Learning Techniques. *International Journal on Emerging Technologies* 11(2), 1082-1086.

[46] Gautam, S., Gautam, A., Chhetri, S., & Bhattarai, U. (2020). Immunity Against COVID-19: Potential Role of Ayush Kwath. *Journal of Ayurveda and integrative medicine*. Advance online publication. https://doi.org/10.1016/j.jaim.2020.08.003

[47] Ji, M., Zhang, L., Wu, Q. (2020). Automatic grape leaf diseases identification via United Model based on multiple convolutional neural networks. *Information Processing in Agriculture* 7(3), 418-426. https://doi.org/10.1016/j.inpa.2019.10.003.

[48] Malvankar, Shubham. (2020). Plant Disease Detection using Image Processing. International *Journal for Research in Applied Science and Engineering Technology* 8(5). 2685-2690. doi:10.22214/ijraset.2020.5450.

[49] Mishra, T., Rai, A. (2021). Antimicrobial potential of Ocimum santum and Adhatoda vasica: The Medicinal Herbs. *International Journal of Pharmacy and Biological Sciences* 1191, 17-28. doi:10.21276/ijpbs.2021.11.1.3.

[50] Manzano, S.,Garrido-Cardenas JA, Manzano-Agugliaro F. (2020). Worldwide Research Trends on Medicinal Plants. *International Journal of Environmental Research and Public Health* 17(10), 3376. doi: 10.3390/ijerph17103376.

[52] Ngugi, L.C., Abelwahab, M., Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. *Information Processing in Agriculture* 8(1), 27-51. https://doi.org/10.1016/j.inpa.2020.04.004.

[53] Saleem, M. H., Potgieter, J., Arif, K. M. (2020). Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers. *Plants*, 9(10), 1319. doi:10.3390/plants9101319.

[54] Sharma, P., Berwal, Y.P.S., Ghai, W., (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. *Information Processing in Agriculture* 7(4), 566-574. https://doi.org/10.1016/j.inpa.2019.11.001.

[55] Sun, J., Yang, Y., He, X., Wu, X. (2020) Northern Maize Leaf Blight Detection Under Complex Field Environment Based on Deep Learning. *IEEE Access* 8, 33679-33688. doi: 10.1109/ACCESS.2020.2973658.

[56] Wang, S., Li, Y., Yuan, J., Song, L., Liu, X., Liu,X. (2020). Recognition of cotton growth period for precise spraying based on convolution neural network. *Information Processing in Agriculture,* 2214-3173. https://doi.org/10.1016/j.inpa.2020.05.001.

[57] Yang, X., Chen, A., Zhou, G., Wang, J., Chen, W., Gao, Y., Jiang, R. (2020). Instance Segmentation and Classification Method for Plant Leaf Images Based on ISC-MRCNN and APS-DCCNN. *IEEE Access* 8, 151555-151573. doi:10.1109/ACCESS.2020.3017560.

[58] Zeng, Q., Ma, X., Cheng, B., Zhou, E., Pang, W. (2020) GANs-Based Data Augmentation for Citrus Disease Severity Detection Using Deep Learning. *IEEE Access* 8, 172882-172891. doi: 10.1109/ACCESS.2020.3025196.

[59] Zhang, Y., Cui, J., Wang, Z., Kang, J., & Min, Y. (2020). Leaf Image Recognition Based on Bag of Features. Applied Sciences 10(15), 5177. doi:10.3390/app10155177

[60] Zhang, W. (2020). imageprocAnal: A novel Matlab software package for image processing and analysis. *Network Pharmacology* 5(1-2), 1-32.

[61] Ji, M., Yang, Y., Zheng, Y., Zhu, Q., Huang, M., Guo, Y. (2021). In-field automatic detection of maize tassels using computer vision. *Information Processing in Agriculture* 8(1), 87-95. https://doi.org/10.1016/j.inpa.2020.03.002.

[62] Naeem, S., Ali, A., Chesneau, C., Tahir, M. H., Jamal, F., Sherwani, R. A. K., & Ul Hassan, M. (2021). The Classification of Medicinal Plant Leaves Based on Multispectral and Texture Feature Using Machine Learning Approach. *Agronomy*, 11(2), 263. http://dx.doi.org/10.3390/agronomy11020263.

[63] Tavakoli, H., Alirezazadeh, P., Hedayatipour A., Banijamali Nasib, A.H., Landwehr, N. (2020). Leaf image-based classification of some common bean cultivars using discriminative convolutional neural networks. *Computers and Electronics in Agriculture*. 181, 105935.

[64] Zhou, C., Zhou, S., Xing, J., Song, J (2021). Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network. *IEEE Access* 9, 28822-28831. doi: 10.1109/ACCESS.2021.3058947.

[65] Vishnoi, V.K., Kumar, K., & Kumar, B.M. (2020). Plant disease detection using computational intelligence and image processing. *Journal of Plant Diseases and Protection, 128*, 19-53.

[66] Vishnoi, V.K., Kumar, K., & Kumar, B.M. (2022). A comprehensive study of feature extraction techniques for plant leaf disease detection. *Multimedia Tools and Applications, 81*, 367-419.

[67] Madhavan, M. V., Thanh, D. N. H., Khamparia, A., Pande S, M. R., & Gupta, D. (2021). Recognition and classification of pomegranate leaves diseases by image processing and machine learning techniques. *Computers, Materials & Continua*, *66*(3), 2939-2955.

[68] Ashwinkumar, S., Rajagopal, S., Manimaran, V., & Jegajothi, B. (2022). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. *Materials Today: Proceedings*, *51*, 480-487.

[69] Bhujel, A., Kim, N. E., Arulmozhi, E., Basak, J. K., & Kim, H. T. (2022). A lightweight Attention-based convolutional neural networks for tomato leaf disease classification. *Agriculture*, *12*(2), 228.

[70] Ravi, V., Acharya, V., & Pham, T. D. (2022). Attention deep learning-based largescale learning classifier for Cassava leaf disease classification. *Expert Systems*, *39*(2), e12862.

[71] Subramanian, M., Shanmugavadivel, K., & Nandhini, P. S. (2022). On fine-tuning deep learning models using transfer learning and hyper-parameters optimization for disease identification in maize leaves. *Neural Computing and Applications*, 1-18.

[72] Singh, R. K., Tiwari, A., & Gupta, R. K. (2022). Deep transfer modeling for classification of Maize Plant Leaf Disease. *Multimedia Tools and Applications*, 1-17.

[73] Khan, M. A., Alqahtani, A., Khan, A., Alsubai, S., Binbusayyis, A., Ch, M. M. I., ... & Cha, J. (2022). Cucumber Leaf Diseases Recognition Using Multi Level Deep EntropyELM Feature Selection. *Applied Sciences*, *12*(2), 593.

[74] Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A., & Alabdali, A. M. (2022). Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction. *Sensors*, *22*(2), 575.

[75] Dhingra, G., Kumar, V., Joshi, H.D. (2017). Study of digital image processing techniques for leaf disease detection and classification. *Multimed Tools Appl* 77, 19951–20000. https://doi.org/10.1007/s11042-017-5445-8.

[76] Kumar, M., Gupta, S., Gao, X., Singh, A. (2019). Plant Species Recognition Using Morphological Features and Adaptive Boosting Methodology. *IEEE Access 7*, 163912163918. doi: 10.1109/ACCESS.2019.2952176.

[77] Malhotra, N., Rani, G.(2020). Plant's Leaf Analysis Using Pattern Recognition Algorithm. *Journal of Computational and Theoretical Nanoscience*, 17, 5167-5173. https://doi.org/10.1166/jctn.2020.9359.

[78] Barbedo, J. G. A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, 2(1), 1-12.

[79] Jasim, M. A., & Al-Tuwaijari, J. M. (2020). Plant leaf diseases detection and classification using image processing and deep learning techniques. In 2020 *International Conference on Computer Science and Software Engineering (CSASE)* (pp. 259-265). IEEE.

[80] Sakhamuri, S., & Kompalli, V. S. (2020). An overview on prediction of plant leaves disease using image processing techniques. *In IOP Conference Series: Materials Science and Engineering* (Vol. 981, No. 2, p. 022024). IOP Publishing.

[81] Vishnoi, V. K., Kumar, K., & Kumar, B. (2021). Plant disease detection using computational intelligence and image processing. *Journal of Plant Diseases and Protection*, 128(1), 19-53.

[82] Naikwadi, S., & Amoda, N. (2013). Advances in image processing for detection of plant diseases. *International Journal of Application or Innovation in Engineering & Management,* 2(11).

[83] Azlah, M. A. F., Chua, L. S., Rahmad, F. R., Abdullah, F. I., & Wan Alwi, S. R. (2019). Review on Techniques for Plant Leaf Classification and Recognition. *Computers*, 8(4), 77. http://dx.doi.org/10.3390/computers8040077.

[84] Oo, Y.M., & Htun, N. (2018). Plant Leaf Disease Detection and Classification using Image Processing. *International Journal of Research and Engineering* 5, 516-523.

[85] Bao,Q.T., Kiet, N.T.T., Dinh, T.Q., Hiep, H.K. (2020) Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks. *Journal of Information and Telecommunication* 4(2), 140-150. doi: 10.1080/24751839.2019.1666625.

[86] Schölkopf, B., & Smola, A. (2005). Support vector machines and kernel algorithms. *In Encyclopedia of Biostatistics* (pp. 5328-5335). Wiley.

[87] Mingqiang, Ji., Yang, Yu., & Zheng, Y., Zhu, Q., & Huang, M., & Guo, Ya. (2020). In-field automatic detection of maize tassels using computer vision. *Information Processing in Agriculture* 8, 87-95. https://doi.org/10.1016/j.inpa.2020.03.002.

[88] Kaur, S., Kaur, P. (2019). Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques. *Journal of Multimedia Systems* 6(2), 49-60. https://doi.org/10.33851/JMIS.2019.6.2.49.

[89] Konstantinos P. F. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture* 145, 311-318. https://doi.org/10.1016/j.compag.2018.01.009.

[90] Rahmani, M.E., Amine, A., Hamou, R.M. (2020). Supervised machine learning for plants identification based on images of their leaves. *Cognitive Analytics: Concepts Methodologies Tools and Applications Hershey PA,* 1314-1330.

[91] Sarker, *I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions*. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592x

[92] S. Ramesh et al., "*Plant Disease Detection Using Machine Learning*," 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 2018, pp. 41-45, doi: 10.1109/ICDI3C.2018.00017.

[93] Begue, A., Kowlessur, V., Singh, U., Mahomoodally, F., Sameerchand, P. (2017). Automatic Recognition of Medicinal Plants using Machine Learning Techniques. International Journal of Advanced Computer Science and Applications 8(4), 166-175. doi:10.14569/IJACSA.2017.080424

[94] Sarkar, C., Gupta, D., Gupta, U., & Hazarika, B. B. (2023). Leaf disease detection using machine learning and deep learning: Review and challenges. Applied Soft Computing, 110534

[95] Roopashree, S., Anitha, J., Mahesh, T. R., Kumar, V. V., Viriyasitavat, W., & Kaur, A. (2022). An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach. *Measurement,* 200, 111484.

[96] Tangtisanon, P., & Kornrapat, S. (2020, February). Holy basil curl leaf disease classification using edge detection and machine learning. In Proceedings of the 2020 12th *International Conference on Computer and Automation Engineering* (pp. 85-89).

[97] Patil, S. S., Patil, S. H., Bhall, A., Rajvaidya, A., Sehrawat, H., Pawar, A. M., & Agarwal, D. (2022, December). Tulsi Leaf Disease Detection using CNN. In 2022 *IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)* (pp. 1-4). IEEE

[98] Bantan, R.A., Ali, A., Naeem, S., Jamal, F., Elgarhy, M., Chesneau, C. (2020). Discrimination of sunflower seeds using multispectral and texture dataset in combination with region selection and supervised classification methods. *Chaos Interdiscip. J. Nonlinear Sci*. 30, 113142.

[99] Jiang, P., Chen, Y., Liu, B., He, D., & Liang, C. (2019). Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks. *IEEE Access* 7, 59069-59080.

[100] Zhang, S., Wu, X., You, z., Zhang, L. (2017). Leaf image based cucumber disease recognition using sparse representation classification. *Computers and Electronics in Agriculture* 134, 135-141. https://doi.org/10.1016/j.compag.2017.01.014.

[101] Zhang, X., Qiao, Y., Meng, F., Fan, C., Zhang, M. (2018). Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks. IEEE Access 6, 3037030377. doi: 10.1109/ACCESS.2018.2844405

[102] Azim, M. A., Islam, M. K., Rahman, M. M., & Jahan, F. (2021). An effective feature extraction method for rice leaf disease classification. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *19*(2), 463-470

[103] Kavitha, R., Kavitha, M., Srinivasan, R., Rajalakshmi, N. R., & Dhayanidhi, R. (2022, November). Basil Leaf Diseases Detection using Deep Learning architectures. In 2022 *IEEE 19th India Council International Conference (INDICON)* (pp. 1-5). IEEE.

[104] Sathiya, V., Josephine, M. S., & Jeyabalaraja, V. (2023). Plant Disease Classification of Basil and Mint Leaves using Convolutional Neural Networks. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2), 153-163.

[105] Kaur, M., Singh, S., & Gehlot, A. (2023). Machine Learning Classification Of Infection In Ocimum Tenuiflorum Using Predictive Modelling. *European Chemical Bulletin*, 12, 2097-2104.

[106] Kaur, M., Singh, S., & Gehlot, A. (2023). Optimized prediction model using support vector machine for classification of infection in Tulsi leaf. *Recent Advances in Computing Sciences*, 206-210.

[107] L. Li, S. Zhang and B. Wang, "*Plant Disease Detection and Classification by Deep Learning—A Review*," in *IEEE Access*, vol. 9, pp. 56683-56698, 2021, doi: 10.1109/ACCESS.2021.3069646.

[108] Chakraborty, S., Paul, S. and Rahat-uz-Zaman, M., 2021, January. *Prediction of apple leaf diseases using multiclass support vector machine*. In 2021 2Nd international conference on robotics, electrical and signal processing techniques (ICREST) (pp. 147151). IEEE.

[109] Patil, B.M. and Burkpalli, V., 2022. *Segmentation of cotton leaf images using a modified chan vese method*. Multimedia Tools and Applications, 81(11), pp.15419-15437.

[110] Kaur, P., Harnal, S., Gautam, V., Singh, M.P. and Singh, S.P., 2023. *Performance analysis of segmentation models to detect leaf diseases in tomato plant*. Multimedia Tools and Applications, pp.1-25.

[111] Alom, M., Ali, M.Y., Islam, M.T., Uddin, A.H. and Rahman, W., 2023. *Species classification of brassica napus based on flowers, leaves, and packets using deep neural networks*. Journal of Agriculture and Food Research, 14, p.100658.

[112] Meenakshi, T., 2023. *Automatic detection of diseases in leaves of medicinal plants using modified logistic regression algorithm*. Wireless Personal Communications, 131(4), pp.2573-2597.

[113] Mahum, R., Munir, H., Mughal, Z.U.N., Awais, M., Sher Khan, F., Saqlain, M., Mahamad, S. and Tlili, I., 2023. *A novel framework for potato leaf disease detection using an efficient deep learning model. Human and Ecological Risk Assessmen*t: An International Journal, 29(2), pp.303-326.

[114] Singh, J. and Kaur, H., 2019. Plant disease detection based on region-based segmentation and KNN classifier. In Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB) (pp. 16671675). Springer International Publishing.

[115] Yaswanth, T. N. S., Lakshmi, K. V., Vamsi, K. G., Supraja, D., & Vani, K. S. (2023, July). Classification of Medicinal Leaves using SVM. In 2023 World Conference on Communication & Computing (WCONF) (pp. 1-7). IEEE

[116] Haagsma, M., Page, G. F., Johnson, J. S., Still, C., Waring, K. M., Sniezko, R. A., & Selker, J. S. (2021). Model selection and timing of acquisition date impacts classification accuracy: A case study using hyperspectral imaging to detect white pine blister rust over time. Computers and Electronics in Agriculture, 191, 106555

[117] Vijaya Durga Reddy, P., & Amudha, V. (2023, May). Comparison of tulasi leaf diseases identification using different convolutional neural network layers. In AIP Conference Proceedings (Vol. 2602, No. 1). AIP Publishing

[118] Azadnia, R., Al-Amidi, M. M., Mohammadi, H., Cifci, M. A., Daryab, A., & Cavallo, E. (2022). An AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling. Agronomy, 12(11), 2723

[119] Patil, S. S., Patil, S. H., Bhall, A., Rajvaidya, A., Sehrawat, H., Pawar, A. M., & Agarwal, D. (2022, December). Tulsi Leaf Disease Detection using CNN. In 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI) (pp. 1-4). IEEE

[120] Sathiya, V., Josephine, M. S., & Jeyabalaraja, V. (2023). Plant Disease Classification of Basil and Mint Leaves using Convolutional Neural Networks. International Journal of Intelligent Systems and Applications in Engineering, 11(2), 153-163

[121] Sathiya, V., Josephine, D., & Jeyabalaraja, D. (2022). Identification And Classification Of Diseases In Basil And Mint Plants Using Psorbfnn. J. Theor. Appl. Inf. Technol. 100, 21

[122] Bharathan, K., & Deepasree, V. P. (2018, December). Tulsi Leaves Classification System. In 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), 1-5. IEEE

[123] Muthunayagam Muthulakshmi, Nagasubramanian Aishwarya, Rajes Kumar Vinesh Kumar, & Babu Rakesh Thoppaen Suresh (2024). *Potato Leaf Disease Detection and Classification With Weighted Ensembling of YOLOv8 Variants*. Journal of Phytopathology, 172(6), e13433.

[124] Sharma, R., Srivastava, R., Goyal, S., & Vinita (2024). *Leaf Disease Detection of Medicinal Plants Using Convolutional Neural Network*. In P. K. Jain, Y. N. Singh, R. P. Gollapalli, & S. P. Singh (Eds.), Advances in Signal Processing and Communication Engineering (ICASPACE 2023) (pp. 275–283).

[125] Kumar, A., & Kalaiselvi, K. (2024). A Hybrid Framework for Plant Leaf Disease Detection and Classification Using Convolutional Neural Networks and Vision Transformer. *Complex & Intelligent Systems*, 11, Article 142. https://doi.org/10.1007/s40747-024-01764-x

[126] Kumar, A., & Kalaiselvi, K. (2024). Performance Analysis of Multiple Leaf Disease Detection in Plants Using CNN Model. In *Proceedings of Third International Conference in Mechanical and Energy Technology (ICMET 2023)* (pp. 1–15). Springer. https://doi.org/10.1007/978-981-97-2716-2_1

[127] Lakshmanarao, A., Kumar, P. S., Chauhan, D. S., & Venkat, M. S. (2024). Medicinal Plant Classification Using Transfer Learning Through Hybrid Machine Learning and Image Processing Techniques. In *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS).*

[128] Deshmukh, M. (2024). Deep Learning for the Classification and Recognition of Medicinal Plant Species. *Indian Journal of Science and Technology*, 17(11), 1070–1077.

APPENDICES
APPENDICE-A
PUBLICATIONS and CONFERENCES

Proposed Objective	**Status of Outcome**
Objective 1: To identify and categorize types of infections in Tulsi Herb.	Presented paper at RACS conference Taylor & Francis Group, Scopus Indexed Kaur, M., Singh, S., & Gehlot, A. (2023). Optimized prediction model using support vector machine for classification of infection in Tulsi leaf. *Recent Advances in Computing Sciences* (pp. 206-210). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003405573-36/optimized-prediction-model-using-support-vector-machine-classification-infection-tulsi-leaf-manjot-kaur-someet-singh-anita-gehlot Journal: Unveiling the Potential: Machine Learning and Image Processing for Early Disease Detection in Tulsi Herb https://journal.esrgroups.org/jes/article/view/3333
Objective 2: To apply image processing for classification of infection in leaves	From Pixels to Prognosis: Machine Learning for Infection Segmentation and Disease Prediction in Tulsi leaf "*5th International Conference on Intelligent Circuits and Systems (ICICS-2023)*" October 12-13th, 2023, Taylor and Francis Books India Pvt Ltd
Objective 3: To develop the training model for prediction of infection in leaves.	Kaur, M., Singh, S., & Gehlot, A. (2023). Machine Learning Classification Of Infection In Ocimum Tenuiflorum Using Predictive Modelling. *Eur. Chem. Bull.* 2023, 12(Special Issue 10), 2097 – 2104 DOI: 10.48047/ecb/2023.12.si10.00251, Publisher: Deuton-X Ltd., ISSN 2063-5346
Objective 4: Testing of the designed prediction model.	"Enhancing Precision in Tulsi Leaf Infection Classification: A Stacking Classifier Ensemble Strategy" https://ijisae.org/index.php/IJISAE/article/view/5123#:~:text=Researchers%20and%20farmers%20have%20been,leaf%20diseases%20in%20this%20work.

International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

"Enhancing Precision in Tulsi Leaf Infection Classification: A Stacking Classifier Ensemble Strategy"

Manjot Kaur[1*], Someet Singh[2], Anita Gehlot[3]

Submitted: 16/01/2024 **Revised**: 24/02/2024 **Accepted**: 02/03/2024

Abstract: Tulsi, often known as Holy Basil (*Ocimum sanctum*), is a herb with cultural significance and health benefits. Researchers and farmers have been concerned about the prevalence of illnesses that harm tulsi leaves in recent years. In order to determine the best model for early identification and intervention, we do a thorough evaluation of many stacking classifier algorithms for the prediction of tulsi leaf diseases in this work. The collection of tulsi leaf imagery in the dataset is broad and includes labels designating various disease conditions. We investigate the efficacy of stacking classifiers by employing a blend of foundational models, such as Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), and Multi-layer Perceptron (MLP) Classifier. Every base model offers a different perspective on the traits connected to both healthy and infected tulsi leaves. We compare various stacking classifier setups based on their accuracy, precision, recall, and F1 score. We take into account differences in the makeup of base models and how model performance is affected by hyperparameter adjustment. Furthermore, we use cross-validation methods to evaluate the five models' generalizability. Farmers and researchers can rely on the model B that is found to have the best predictive performance with an average accuracy of 98.25%, since it provides a strong means of early disease diagnosis and management. This study advances precise farming methods, encourages tulsi cultivation that is sustainable, and guarantees the plant's continued use in traditional medicine.

Keywords: *Stacking Classifier, Precision Agriculture, Machine Learning, Image Classification*

1. Introduction

In India, tulsi is a commonly grown herb with both medicinal and cultural value [1]. It is produced in farms and backyard gardens frequently, and it's used in traditional Ayurvedic treatment. In India, the amount of tulsi produced varies according to demand, climate, and local farming methods. India was one of the world's top producers of tulsi as of 2021[2]. The plant is cultivated in several states throughout the nation, and both conventional and innovative agricultural techniques have been used to enhance its yield. Precision farmers can target specific portions of a field that exhibit disease indicators by implementing classifier models. Utilising resources like pesticides, water, and fertilisers more effectively is made possible by this targeted strategy. Cost reductions may come from early detection and focused treatment based on classifier predictions. It minimises the requirement for treatments to be applied widely throughout the field, maximising resource utilisation and lowering environmental impact. Large leaf image datasets may be processed quickly and effectively by automated systems driven by classifier models, which makes them scalable for crop monitoring over vast agricultural regions [3].

Classifier models provide an impartial and standardised evaluation of leaf health by generating predictions based on pre-established patterns and attributes. This guarantees a more standardised review procedure and lessens the unpredictability brought about by human observers. For extensive crop health monitoring, classifier models can be incorporated into remote sensing technologies like drones or satellite photos [4]. A more comprehensive perspective of agricultural landscapes is possible with this method. Classifier models that examine patterns and trends in leaf health data over time might produce insightful results. Making educated judgements regarding resource allocation, crop management, and disease prevention techniques is possible with the use of this knowledge. Internet of Things (IoT) devices, like field-positioned sensors and cameras, can be combined with classifier models [5]. The model can interpret real-time data from various devices so that decisions can be made right away. In conclusion, classifier models are essential for agricultural plant disease detection and leaf health prediction. They facilitate data-driven decision-making, focused interventions, and early detection, all of which lead to more effective and sustainable farming methods.

[1*]*Lovely Professional University, School of Electrical and Electronics Engineering, Phagwara, India*
[2,1]*Lovely Professional University, School of Electrical and Electronics Engineering, Phagwara, India*
[3]*Uttaranchal University, Uttaranchal Institute of Technology, Dehradun, India*
manjotkaur.hpk@gmail.com[1], someet.17380@lpu.co.in[2], anita.ri@uttaranchaluniversity.ac.in[3]
**Corresponding Author: Manjot Kaur*
[*]*Lovely Professional University, School of Electrical and Electronics Engineering, Phagwara, India*

J. Electrical Systems 20-5s (2024): 2954-2974

[1]Manjot Kaur

[2]Someet Singh

[3]Anita Gehlot

[4]Rajesh Singh

Unveiling the Potential: Machine Learning and Image Processing for Early Disease Detection in Tulsi Herb

Abstract: - Tulsi (Ocimum tenuiflorum) herb is very much predisposed to infections that influence the growth of the plant and impact the farmer's ability to learn about the environmental factors affecting the plant. To discover any type of plant infection at a very preliminary stage, a prediction model employing machine learning and image processing techniques can be developed to accelerate the method of disease detection and classification with high-performance metrics. The deployment of various variable preprocessing methods and distinctive factors in the feature extraction process appeared to enhance the implementation of infection recognition and categorization. This article intends to assess and explore the application and implementation of numerous approaches and developments regarding leaf infection categorization and classification. A thorough analysis is provided for disease infection and classification implementation upon examination of formerly recommended avant-garde methods. Finally, challenges and some commendations in this space are considered for the real-time implementation of numerous image computational algorithms for disease detection and recognition of Ocimum tenuiflorum.

Keywords: Avant-garde methods; Disease detection; Image processing technique; Machine learning; Plant infection; Ocimum tenuiflorum

1. Introduction

Ocimum tenuiflorum, 'The Embodiment of Herbaceous plant', the prominent 'Exceptional one' 'Tonic of Life', stands as an annual delicate herb [1]. It is grown at a large scale in the tropical climate and warm regions of India. It is an aromatic shrub, grown throughout the eastern world tropics. It is widely planted all over India in kitchen gardens and as indoor plants for religious and traditional medicine purposes and is famous for its essential oils. India remains a developing country, and the domain of agriculture is the backbone of the country's continuous growth and development. The field of agriculture faces lots of hurdles like a huge loss in crop production and the traditional method of plant leaf disease identification is also exceedingly difficult in the agricultural field [2].

The traditional manual method of identifying the infection in plants is variable and unpredictable; hence it cannot be considered a reliable systematic strategy to crop control. The consumer challenge for increased protection and superiority in agricultural products is what's required right now. Therefore, it will be crucial in the future to develop an automated, accurate, inexpensive, and effective model to identify plant infection in leaf, stem, and root. [3-4].

The investigators have proposed image processing and classification model work on various medicinal plants *Ocimum tenuiflorum*, bael(*Aegle marmelos*), peppermint*(Mentha × piperita*), catnip(*Nepeta cataria*), lemon balm(*Melissa officinalis*), and stevia(*Stevia rebaudiana)* [5-8]. *Ocimum tenuiflorum* is an important medicinal plant in day-to-day life and to identify the infections at an early stage for increasing the production quality and quantity, novel image processing and machine learning classifier models are required. The prediction model for plant leaf infection recognition and classification utilizes fast computer vision software and machine learning methods [11-13]. The symptoms of *Ocimum tenuiflorum* infection can decide the placement of dataset of infected leaves in training model under appropriate categorization labels. In image processing features like shape, texture, colour, vein, and many more are extracted and processed [14-15]. Machine learning classifiers like k-nearest neighbour, support vector machine, random forest, decision tree and artificial neural network are compared for better accuracy model [16-17].

[1] [1]Lovely Professional University, School of Electrical and Electronics Engineering, Phagwara, India manjotkaur.hpk@gmail.com

[2]Lovely Professional University, School of Electrical and Electronics Engineering, Phagwara, India someet.17380@lpu.co.in

[3]Uttaranchal University, Uttaranchal Institute of Technology, Dehradun, India anita.ri@uttaranchaluniversity.ac.in

[4]Uttaranchal University, Uttaranchal Institute of Technology, Dehradun, India, rajeshsingh@uttaranchaluniversity.ac.in

MACHINE LEARNING CLASSIFICATION OF INFECTION IN *OCIMUM TENUIFLORUM* USING PREDICTIVE MODELLING

Manjot Kaur[1*], Someet Singh[2], Anita Gehlot[3]

Abstract

India's economy relies heavily on agricultural production and is a major source of employment. The early identification of plant leaf diseases is crucial for maximizing revenue and crop productivity. There are more tulsi (*Ocimum Tenuiflorum*) plant products produced in India than anywhere else in the world. Early methods of observing solely through visual inspection were time-consuming and inaccurate. The present work identifies and categorizes leaf diseases by using a variety of image-processing approaches. The present study demonstrates comprehensive methods for identifying and classifying infections in medicinal plants utilizing image processing and machine learning. The plant village input image dataset having three different types of infected *Ocimum Tenuiflorum* leaf and healthy leaf is amassed as the basis for this dataset. Through the use of a computer vision lab framework, the image datasets are augmented, pre-processed, segmented, extracted and validated with certain features. Five machine learning classifiers are evaluated using an optimized dataset of infected leaves of *Ocimum Tenuiflorum*. including logistic regression, linear discriminant analysis, k nearest neighbour, classification and regression trees, random forests, naive bayes, and support vector machines. According to the results, the random forest classifier outperforms the others with an accuracy of 99.86%, followed by the linear discriminant analysis with 98.59%, and the support vector machine with 97.42%.

Keywords: Augmentation, Image Processing, Machine Learning, *Ocimum tenuiflorum*, Predictive Modelling

[1*]School of Electrical & Electronics Engineering, Lovely Professional University, Phagwara-144411, India, manjotkaur.hpk@gmail.com
[2]School of Electrical & Electronics Engineering, Lovely Professional University, Phagwara-144411, India, someet.17380@lpu.co.in
[3]Uttaranchal Institute of Technology, Uttaranchal University, Dehradun-248007, India, anita.ri@uttaranchaluniversity.ac.in

***Corresponding Author:** Manjot Kaur
School of Electrical & Electronics Engineering, Lovely Professional University, Phagwara-144411, India, manjotkaur.hpk@gmail.com

DOI: 10.48047/ecb/2023.12.si10.00251

36 Optimized prediction model using support vector machine for classification of infection in Tulsi leaf

Manjot Kaur[1,a], *Someet Singh*[1,b], *and Anita Gehlot*[2,c]

[1]School of Electrical & Electronics Engineering, Lovely Professional University, Punjab, India

[2]Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, India

Abstract

Tulsi (*Ocimumtenuiflorum*) herb is highly susceptible to diseases that can hinder plant growth and impact the farmer's ability to learn about the environmental factors affecting plant development. A prediction model integrating machine learning and image processing techniques can be constructed to expedite the approach of disease detection and classification with high-performance indicators to find any type of plant infection at a very initial stage. Machine learning and computer vision are developing technologies that enable computers to recognize and comprehend information from digital images. The purpose of this research is to assess and investigate the use and implementation of supervised machine learning image classifier models support vector machine (SVM) utilizing histograms of gradients as feature extraction and for categorization and classification of Tulsi leaf diseases. Finally, computations from the confusion matrix are used to compare the two models for better accuracy. The accuracy of the classification of the leaf disease using SVM is calculated. The proposed prediction model performs well with an increased training dataset with an accuracy of 96.714%.

Keywords: Image processing, machine learning, *Ocimumtenuiflorum*, support vector machine

Introduction

Productive recognition and identification of leaf infection is an enduring research work in computer vision (CV) due to its substantial applications in agribusiness and agrarian frugality as per Barbedo (2013). Numerous infections exist in horticulture that has an impact on the development and nature of plants. Many of these infections are decided according to a specialist in this space considering their side effects. However, it is expensive due to the difficulty in reaching specialists and higher costs as described by Jasim et al. (2020). In this way, the processing scientists have developed several calculations for the automated discovery of infections in plants in collaboration with agricultural experts. To distinguish the infections in a few types of plants, leaf side effects are a valuable source of information. The major herb plant *Ocimumtenuiflorum* is well known for its nutritional benefits. But the attack of numerous infections damages both its creation and quality. Therefore, it is crucial to develop an automated framework for early phase symptom finding and categorizing of *Ocimumtenuiflorum* leaf symptoms. Finding the lesion spot is an active research subject in computer vision Vishnoi et al. (2021) and numerous methods for plant disease diagnosis using image processing and machine learning algorithms have recently been introduced Dhingra et al. (2018). Characteristics like color, texture, and form are crucial for classifying leaf diseases since the plant lesion spots are typically evaluated by their appearance. A prediction model for plant leaf infection detection and classification using quick computer vision technology and machine learning techniques are proposed widely.

Need for *Ocimumtenuiflorum* Infection Classification

Like any restorative plant, ideal development, collecting, safeguarding, and stockpiling strategies are expected for *Ocimumtenuiflorum's* therapeutic and profound qualities as mentioned by Mishra et al. (2021). It is essential to guarantee the quality principles and cycles for solid *Ocimumtenuiflorum* leaves according to Roopashree et al. (2020) It turns out to be critical to recognize the sound and contaminated infection passes on and take fundamental estimations to further develop the creation quality and harvest yield. The *Ocimumtenuiflorum* plant is filled normally in sodden soil all over globe. It is filled in mild environment, and it tends to be effectively impacted by contagious, bacterial microorganisms, and nematodes. Unpredictable dull spots and round spots on leaves with faint light habitats, yellow, withering leaves, staggered development, dying leaves dropping from plant, sores on *stem* are caused because of parasite. Yellowing leaves are regularly viewed as supplement lack; however, they are spread by tainting of seeds. Sporadic brown or dark water-drenched spots or rakish spots on leaves, streaks on stems are caused because of microbes as illustrated

[a]manjotkaur.hpk@gmail.com; [b]someet.17380@lpu.co.in; [c]anita.ri@uttaranchaluniversity.ac.in

64. From Pixels to Prognosis: Machine Learning for Infection Segmentation and Disease Prediction in Tulsi Leaf

Manjot Kaur[1], Someet Singh[1], and Anita Gehlot[2]

[1]*School of Electrical and Electronics Engineering, Lovely Professional University, Phagwara, India*
[2]*Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, India*

Abstract: Infections in the tulsi plant account for one-fourth of India's agricultural crop production loss on average. A significant component of the economy is dependent on agricultural output. Leaf diseases are every country's top agricultural issue, as worldwide food consumption is rapidly increasing due to population increase. Prenatal disease detection remains difficult due to limitations in lab infrastructure and capability. The initial and detailed diagnosis of leaf infections is critical to prevent their spread. Image processing techniques that use mathematical equations and transformations can be used to diagnose diseases. When viewed with human eyes, we can extract specific information from an image using these colours, but modern computers store images in a mathematical framework. In this study, we evaluate the possibility of computer vision approaches for scalable and early detection of plant-borne illnesses in tulsi leaf using picture segmentation techniques. The low availability of suitable large-scale datasets remains a significant restriction for enabling vision-based plant disease identification. As a result, we present TulsiDoc, a dataset for detecting visual plant illnesses in the tulsi plant. The collection, which included data augmentation in annotating the PlantVillage dataset, contains over 2500 images and as many as three classes of infections: fungal, bacterial and pests. To demonstrate the utility of the dataset, we create two models for the task and employed segmentation techniques using thresholding and intensity segmentation and compared the outcome of the methods. The connection between plant health and human health consumption is complicated and multifaceted. It is largely concerned with the quality and safety of the food we eat, which is directly affected by the health and well-being of the plants employed in agriculture. In order to do accurate analysis, feature extraction and personalised treatment, segmentation is a crucial stage in disease prediction. The paper aids in early diagnosis, lessens false positives and improves our comprehension of how

ICICS
International Conference on
Intelligent Circuits and Systems

LOVELY PROFESSIONAL UNIVERSITY
Transforming Education Transforming India

NAAC GRADE A++

Certificate No. 294681

Certificate of Paper Presentation

This is to certify that **Prof./Dr./Mr./Ms. Manjot Kaur** of **Lovely Professional University, Punjab, India** has presented the paper titled **From Pixels to Prognosis: Machine Learning for Infection Segmentation and Disease Prediction in tulsi leaf** in the **"5th International Conference on Intelligent Circuits and Systems (ICICS-2023)"** held on **October 12-13th, 2023**, organized by School of Electronics and Electrical Engineering at Lovely Professional University, Punjab.

Date of Issue : 30-10-2023
Place : Phagwara (Punjab), India

Prepared by (Administrative Officer-Records)	**Dr. Bhaveshkumar C. Dharmani Organizing Secretary (ICICS-2023)**	**Dr. Charanjit Singh Convenor (ICICS-2023)**	**Dr. Gaurav Sethi Conference Chair (ICICS-2023)**

LOVELY PROFESSIONAL UNIVERSITY
Transforming Education Transforming India

RACS
SCHOOL OF COMPUTER APPLICATION

Certificate No. 256866

Certificate of Paper Presentation

This is to certify that **Dr./Mr./Ms. Manjot Kaur** of **Lovely Professional University, Punjab** has presented a paper entitled "**Optimized prediction model using support vector machine for classification of infection in Tulsi leaf**" in the **1st International Conference on Recent Advances in Computing Sciences (RACS-2022)** held from 4th to 5th November 2022, organized by School of Computer Application at Lovely Professional University, Punjab.

Date of Issue : 30-11-2022
Place : Phagwara (Punjab), India

Prepared by (Administrative Officer-Records)	**Convener (RACS-2022)**	**Conference Chair (RACS-2022)**	**Session Chair (RACS-2022)**

भारत सरकार / GOVERNMENT OF INDIA
कॉपीराइट कार्यालय / Copyright Office
बौद्धिक संपदा भवन,प्लॉट संख्या 32, सेक्टर 14, द्वारका, नई दिल्ली-110078 फोन: 011-28032496
Intellectual Property Bhawan, Plot No. 32, Sector 14, Dwarka, New Delhi-110078 Phone: 011-28032496

Copyright Office

रसीद / Receipt

PAGE No : 1

To,	**RECEIPT NO**	**: 156802**
Monica-Gulati **Registrar-Lovely professional University-**	**FILING DATE**	**: 22/12/2024**
	BRANCH	**: Delhi**
(M)-9888160523 : E-mail- dip@lpu.co.in	**USER**	**: Lovelyprofessionaluniversity**

S.No	Form	Diary No.	Request No	Title	Amount (Rupees)
1	Form-XIV	40570/2024-CO/L	173756	Enhancing Infection Diagnosis with Tulsi: AI-driven Stacking Classification and Prediction Model	500
Amount in Words		Rupees Five Hundreds			500

PAYMENT MODE	Transaction Id	CIN
Online	C-0000192049	2212240002595

(Administrative Officer)

***This is a computer genereated receipt, hence no signature required.**

***Please provide your email id with every form or document submitted to the Copyright office so that you may also receive acknowledgements and other documents by email.**

APPENDICE-B
COURSES and DEVELOPMENT PROGRAMS

Title	Type	Date	Duration	Platform
Image Data Augmentation with Keras	Course	8 June 2022	Project	Coursera- Grade Achieved-100%
Complete Python Based Image Processing and Computer Vision	Course	8 June 2022	5.5 hours	Udemy
Design Patterns in Python	Course	27 Oct 2021	9 hours	Udemy
Introduction to Computer Vision and Image Processing-IBM	Course	14 Jul 2021	15 hours	Coursera-Grade Achieved-85.21%
Medical Diagnosis using Support Vector Machines	Course	28 June 2021	1 day	Coursera- Grade Achieved-100%
Energy Management 4.0	Course	02 June-4 July 2021	108 hours	Lovely Professional University
Python for Data Science	Course	Jan-Apr 2020	4 weeks	NPTEL

Title	Type	Date	Duration	Platform
Advances in Electronics and Communication Engineering for Industrial Applications	FDP	26-30 Sept 2021	5 days	AICTE ATAL Academy, SCETW, Hyd
“Artificial Intelligence & Conceptualization of Algorithms using Python	FDP	12-16 July 2021	5 days	AICTE ATAL Academy, BIT, Bangalore
Data Science and Machine Learning using Python	FDP	5-9 July 2021	5 days	AICTE ATAL Academy, Manipal University, Jaipur

Smart Cities	FDP	24-28 August 2020	5 days	AICTE ATAL Academy, PEC Chandigarh
Hands on Artificial Intelligence 5 days Workshop	FDP	27-31 July 2020	5 days	BITS Pilani, Hyderabad, Brainlabs
Python for Image Processing and IoT	FDP	23-28 Nov 2020	6 days	Geethanjali Institute of Science & Technology


coursera project network

Jun 8, 2022

Manjot Kaur

has successfully completed

Image Data Augmentation with Keras

an online non-credit project authorized by Coursera Project Network and offered through Coursera

Amit Yadav
Instructor
Machine Learning and Data Science

PROJECT CERTIFICATE

EDUCATION FOR EVERYONE coursera PROJECT CERTIFICATE

Verify at:
https://coursera.org/verify/D7DSTRB5VSBD
Coursera has confirmed the identity of this individual and their participation in the project.

udemy

Certificate no: UC-cd55d9fa-c69d-4e35-9c61-f2a10aed161a
Certificate url: ude.my/UC-cd55d9fa-c69d-4e35-9c61-f2a10aed161a
Reference Number: 0004

CERTIFICATE OF COMPLETION

Complete Python Based Image Processing and Computer Vision

Instructors **Minerva Singh**

Manjot Kaur

Date **June 8, 2022**
Length **5.5 total hours**

This certificate is computer generated and can be verified by scanning the QR code given below. This will display the certificate from the NPTEL repository, https://nptel.ac.in/noc/

Roll No: NPTEL20CS36S1750495

To
MANJOT KAUR
HNO 595320 RADHA REGAL ROWS YAPRAL
HYDERABAD
TELANGANA-500087
PH. NO :9849666550

No. of credits recommended by NPTEL:1
An additional 1 credit may be awarded if the University deems it fit, based on the actual student effort involved.

NPTEL Online Certification

(Funded by the Ministry of HRD, Govt. of India)

Covid-19 impacted January 2020 semester

This certificate is awarded to
MANJOT KAUR
for passing the course
Python for Data Science
*with Score** **79** %

Devendra Jalihal
Prof. Devendra Jalihal
Chairman
Centre for Continuing Education, IITM

Jan-Feb 2020
(4 week course)

Prof. Andrew Thangaraj
NPTEL Coordinator
IIT Madras

Indian Institute of Technology Madras

swayam

*Continuous online assessment score

To validate and check scores: https://nptel.ac.in/noc udemy

Certificate no: UC-dee14ada-f5e1-4697-aad2-c7e3b5567b33
Certificate url: ude.my/UC-dee14ada-f5e1-4697-aad2-c7e3b5567b33
Reference Number: 0004

CERTIFICATE OF COMPLETION

Design Patterns in Python

Instructors **Dmitri Nesteruk**

Manjot Kaur

Date **Oct. 27, 2021**
Length **9 total hours**

IBM

COURSE CERTIFICATE

14-Jul-2021

Manjot Kaur

has successfully completed

Introduction to Computer Vision and Image Processing

an online non-credit course authorized by IBM and offered through Coursera

EDUCATION FOR EVERYONE coursera COURSE CERTIFICATE

Aije Egwaikhide
Senior Data Scientist
IBM

Joseph Santarcangelo
Senior Data Scientist
IBM

Verify at coursera.org/verify/KXFXPQ95Q6CG
Coursera has confirmed the identity of this individual and their participation in the course.

coursera project network

COURSE CERTIFICATE

Jun 28, 2021

Manjot Kaur

has successfully completed

Medical Diagnosis using Support Vector Machines

an online non-credit course authorized by Coursera Project Network and offered through Coursera

Daniel Romaniuk
Project Instructor
Computer Science

Verify at coursera.org/verify/DQYHN2T4JCT2
Coursera has confirmed the identity of this individual and their participation in the course.

Italian National Agency for New Technologies, Energy and Sustainable Economic Development

Certificate No.226568

Certificate of Participation

This is to certify that **Ms. Manjot Kaur** D/o Sh. Daljeet Singh, Assistant Professor, Department of Electronics, Stanley College of Engineering and Technology, Hyderabad participated in **Short Term Course on Energy Management 4.0** organized by Lovely Professional University w.e.f. **June 02, 2021** to **July 04, 2021** and obtained "**O**" Grade.

Date of Issue : 04-07-2021
Place : Phagwara (Punjab), India

Prepared by
(Administrative Officer-Records)

Dr. Savita Gupta
Professor
Human Resource Development Center
Lovely Professional University

Dr. Sunaina Ahuja
Professor and Associate Dean
Head- Human Resource Development Center
Lovely Professional University

Dr. Dharam Buddhi
Executive Dean
Division of Innovation and Entrepreneurship
Lovely Professional University

No: ATAL/2021/1625484660

ATAL

STANLEY

ALL INDIA COUNCIL FOR TECHNICAL EDUCATION

Nelson Mandela Marg, Vasant Kunj, New Delhi – 110 070

AICTE Training and Learning (ATAL) Academy

Certificate

This is certified that **Manjot Kaur**, **Assistant Professor** of **Stanley College of Engineering and Technology for Women** participated & completed successfully AICTE Training And Learning (ATAL) Academy **Online Elementary** FDP on "**Advances in Electronics and Communication Engineering for Industrial Applications**" from **26/07/2021** to **30/07/2021** at **Stanley College of Engineering and Technology for Women**.

Advisor-I, ATAL Academy
Dr. Mamta Rani Agarwal

Coordinator

No:ATAL/2021/1624200895

ATAL

ALL INDIA COUNCIL FOR TECHNICAL EDUCATION

Nelson Mandela Marg, Vasant Kunj, New Delhi – 110 070

AICTE Training And Learning (ATAL) Academy

Certificate

This is certified that **Manjot Kaur**, **Assistant Professor** of **Stanley College of Engineering and Technology for Women** participated & completed successfully AICTE Training And Learning (ATAL) Academy **Online Elementary** FDP on "**Artificial Intelligence and Conceptualization of Algorithms using Python**" from **2021-7-12** to **2021-7-16** at **Bangalore Institute of Technology, Bangalore**.

Adviser-I, ATAL Academy

Coordinator

No:ATAL/2020/1597170615

ATAL

AICTE

PEC
EXPLORE INNOVATE EXCEL

ALL INDIA COUNCIL FOR TECHNICAL EDUCATION

Nelson Mandela Marg, Vasant Kunj, New Delhi – 110 070

AICTE Training And Learning (ATAL) Academy

Certificate

This is certified that **Manjot Kaur**, **Assistant Professor** of **Stanley College of Engineering and Technology for Women** participated & completed successfully AICTE Training And Learning (ATAL) Academy **Online** FDP on "**Smart Cities**" from **2020-8-24** to **2020-8-28** at **Punjab Engineering College (Deemed to be University), Chandigarh**.

Director ATAL Academies

Coordinator

GEETHANJALI

Institute of Science & Technology

(Approved by AICTE, Permanently Affiliated to J.N.T.University, Anantapur, 2(f) & 12(B) Status of UGC)
Gangavaram, Nellore, Andhra Pradesh – 524 137, India

Certificate of Participation

IXQZYK-CE000046

This is to certify that Manjot Kaur, Assistant Professor from Stanley College of Engineering and Technology for Women has participated in One Week AICTE-ISTE Refresher/Induction Program (Online-Phase-I) on **"Python for Image Processing and IoT"** Organised by Department of Electronics and Communication Engineering during 23rd to 28th November, 2020.

28/11/2020
Convenor & Co-ordinator

28/11/20
Co-Convener

Principal

Made for free with Certify'em