Deep Learning based Medicinal Plants Leaf Recognition

by

Tonusri Mahalanabish
22173011

A project submitted to the Department of Computer Science and Engineering in
partial fulfillment of the requirements for the degree of
M.Engg. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
September 2022

Declaration

It is hereby declared that

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

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

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

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

<div style="margin-top:3em"></div>

_____

Tonusri Mahalanabish
22173011

Approval

The project titled "Deep Learning based Medicinal Plants Leaf Recognition" submitted by

Tonusri Mahalanabish (22173011)

Of Summer, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of M.Engg in Computer Science and Engineering on September 25, 2022.

Examining Committee:

Supervisor:
(Member)

Md. Golam Rabiul Alam, PhD
Professor
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Amitabha Chakrabarty, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

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

Abstract

Plants assume a significant part in Earth's nature by giving food, cover and keeping a solid environment.These plants contain some significant therapeutic qualities. Due to having fewer negative side effects and being more affordable than contemporary medicine, medicinal plants are receiving interest in the pharmaceutical business.In this work, I tried to classify the plant's images through classical methods and Deep neural network.30 medicinal plants leaves are represented by 1835 images in the proposed dataset.First, I applied CNN to classify the images and got 65.66% accuracy.Then I applied SVM with Normal features, GrayScale features, HOG features and combined features extraction and got 72.28% accuracy for Normal features,73.91% accuracy for GrayScale features, 79.34% accuracy for HOG features and 80.0% for Combined feature extraction.Next I applied the VGG-19 pre-trained model and got 96.74% accuracy.At last, I applied a GradCam explainable AI method to interpret the results generated from VGG19.From all these experiments, I got the best accuracy for the VGG19 pretrained model.That's why I used Grad Cam on the VGG19 results for getting the explanation for the predictions.

Keywords: Deep Neural Network; CNN; VGG19;Machine Learning; SVM; HOG Feature;GrayScale Feature;Grad-CAM.

Dedication

This work is dedicated to my beloved parents, loving husband and Dr. Golam Rabiul Alam sir for guiding and bearing with me during this period with love and patience.

Acknowledgement

Firstly, all praise to the Great God for whom this project have been completed without any major interruption.

Secondly, I am very much grateful to my supervisor Md. Golam Rabiul Alam, Associate Professor, Department of Computer Science and Engineering, BRAC University for his constructive suggestions, scholastic guidance, constant inspiration and kind co operation throughout the entire progress of this project work. It would have been impossible for me to accomplish my task without his invaluable advice.

Thirdly, I also would like to thank all the concerned teachers and stuffs of the department for their direct and indirect assistance at different events of the work.

Finally to my parents without their throughout support it may not be possible. With their kind support and prayer I am now on the verge of our post-graduation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Motivation

For all life on earth, plants are essential. The primary source of oxygen is plants. In addition to this, the plants are employed in a wide range of industrial purposes, such as herbs, biomass, medicines, and constituents in biofuels. Plants have long been used as a kind of traditional medicine. It is thought to be inexpensive and has no negative side effects. There are many different plant species on the planet, yet it can be challenging to identify them. There are numerous ways to identify a plant. Currently, taxonomists manually identify plants, which are prone to human mistake.Researchers have created automatic plant identification systems to prevent this. The plants known as medicinal plants are those that are utilized to treat and prevent particular human illnesses and diseases. The scientists and urban population are still unaware of the knowledge gathered by the peasants and tribal people regarding herbal medicine. Numerous plant species found in rural areas are on the edge of extinction and are listed as vulnerable.Rural populations are being displaced from their natural habitats as a result of deforestation, urbanization, and modernization, and their very expertise, particularly with regard to herbal medicines, is slowly vanishing [23].

The fundamental building block of medicine is plants. Without plants, life would not exist. Traditional medicinal herbs are the source of a number of significant medications that are still in use today. Ethnopharmacology and ethnobotany have become key sources of information in the search for new medicines. This has opened up new sources and classes of chemicals. Studies on structure-activity interactions are currently one of pharmacochemistry's most important and valued achievements, making it a leading member of the family of pharmaceutical sciences due to their influence on the design of innovative medications. The survival of the human race without plants on Earth is inconceivable.Since the beginning of the human race, humans have been dependent on plants. Medicine is frequently obtained from medicinal plants. The use of herbs in traditional medical practices like Ayurveda, Unani, and Chinese traditional medicine for the treatment of illnesses as well as for regenerating and bolstering bodily systems is supported by substantial evidence. Herbal use was originally intended to produce a favorable interaction with body chemistry [12].

By providing oxygen and water for living things, plants play a critical part in pro-

tecting life and maintaining biodiversity on earth.It is crucial to classify and identify medicinal plants in order to provide better care. Lack of understanding about the proper plants can seriously hurt or even kill a person. Because of this, accurate medicine requires that these medicinal plants be properly identified. Due to a dearth of knowledge in this field, accurate identification and classification of medicinal plants is a tedious task. Therefore, a fully automated approach for classifying medicinal plants is widely desired [29].

One of the most important types of plants, medicinal plants, play a significant part in protecting human physical and mental health. There are many different species of medicinal plants on the planet, however it is exceedingly challenging to identify them. The scientists and urban population are still unaware of a significant amount of information gathered by the tribal members and villagers on plant-based medicine. Typically, this knowledge is passed down from one generation to the next [23].

The essential question in leaf recognition, which is important for plant categorization, is whether the chosen traits are consistent and effectively distinguish different types of leaves. The recognition process takes a long time. Because of flawed models and ineffective representational strategies, computer-assisted plant recognition remains an extremely difficult task in computer vision. The evaluation of the Fourier moment-based and geometrical morphological leaf traits is the primary goal of plant recognition. This information is crucial for classifying the various plant types. An important step in the process of recognizing plants is the extraction of leaf traits from a plant.The technique of feature extraction presents a fresh problem in pattern recognition. The automatic computer capture of data from living plants has not been put into practice [2].

A range of stakeholders, including forestry services, botanists, taxonomists, physicians, pharmaceutical labs, organizations that protect endangered species, the government, and the general public, can all benefit greatly from proper identification of plant species. As a result, developing automated systems that can recognize distinct plant species is becoming more and more popular.There are thousands of different plant species in the globe, some of which are medicinal, some of which are endangered, and yet others which are detrimental to humans. Plants are not only a vital resource for humans; they also serve as the foundation of all food chains.Correct plant classification and study are essential for both practical usage and species protection. Unknown plant identification heavily depends on the innate knowledge of a skilled botanist. The most effective method for correctly and quickly identifying plants is a manual approach based on morphological traits. Thus many of the processes involved in identifying these plant species are 'dependent on information accumulation and talents of human beings'[19].

The first step in protecting a plant species is to identify it. To identify different species and prevent their extinction, an object recognition system is therefore absolutely necessary. To achieve this, it is possible to make use of recent breakthroughs in the field of computer vision. Recently, automated plant detection processes have been developed using computer vision and pattern recognition approaches [2].

Numerous plants are readily available in the area, but nobody is aware of them; only specialists are familiar with these medicinal plants. There is a need to properly record these species and prevent their extinction because a number of them have very beneficial medical characteristics. Herbs are both plants and parts of plants that are used in tablets, powders, capsules, and other medications for aroma, flavor, and other therapeutic purposes. These are used by people to improve their health and treat ailments. Some medicines are created from plants, while others are derived from herbal plants.Plants can be recognized by a number of fundamental physical characteristics, including size, shape, color, vein pattern, and texture. Some plants have similar shapes and appearances, making it challenging for laypeople to identify them correctly. In order to identify those plants, a suitable method is needed to store the data in a way that will allow leaf identification technology to utilize it in the future. For this reason, it is necessary to train a machine that can swiftly and accurately recognize such plant species [38].

By electronically capturing an image, computer vision has made it possible to duplicate human eyesight. Numerous methods for measuring leaf area have been developed as a result of the importance of calculating leaf area in plant sciences. These methods include grid counting, blue printing, photography, length and breadth correlation, and the use of electronic equipment. In computer-based solutions, a scanner or digital camera is often used to capture the image of the leaf, which is then analyzed by the computer using specially developed software. application of neural networks similar to those employed in the classification of blooming and cactus plants, statistical discriminant analysis, color-based clustering, and other methods. Using the CSS method and k-NN classifiers, several groups of chrysanthemum leaves were identified. Shape, edge, tip, base, surface, trichomes and arrangement on stem divisions of the main blade, color, texture, and venation are features that could be utilized to distinguish between distinct leaves. These qualities can be extracted using HSV color space and SIFT [2].

One of the most well-liked machine learning strategies for classifying plant leaves is neural networks. Artificial neural networks (ANN), probabilistic neural networks (PNN), convolutional neural networks (CNN), k-nearest neighbor (KNN), and support vector machines (SVM) are some of the neutral networks that are frequently employed. Some research have even integrated techniques for accuracy enhancement. Plant leaf classification seemed to perform better when a variety of preprocessing methods were used, as well as typical parameters in feature extraction [28].

Our ecosystem is mostly framed by the plants that are in our immediate environment. While certain plants provide edible foods for us above or below the ground, a large number of plants have been used in the Indian system of medicine known as Ayurveda. These plants are not like other plants. People could recognize the medicinal properties of these plants in treating numerous diseases in ancient times. Various characteristics of leaves, including texture, form, and color, can be used to categorize them [18].

1.2 Project Problem

To identify a medicinal plant can be a life and death problem, so high accuracy is a must. For this, various machine learning and deep learning algorithms can be used. But as they are mostly black box methods, it is not explained how and why the predictions are made. That's why there's an air of distrust in deep learning algorithms in such critical situations.

1.3 Aims and Objectives

Plant identification based on images of leaves. Some systems make use of botanist-style descriptions. However, it is challenging to extract and transfer those features automatically to a computer. The widespread use of digital cameras and handheld computers in the modern world has made the prospect of the development of this system a reality. Researchers are working to classify plants automatically thanks to studies in image processing and machine learning. The identification of distinguishing characteristics for diverse species presents a problem in the development of this system [22].The classification of plant leaves is a crucial procedure in botany as well as in the cotton, tea, and other industries. Additionally, the morphological traits of leaves are utilized to classify plants or to detect some plant diseases early on. Plant identification is a critical and challenging process [2].

The aim is to train a model which will classify the medicinal plants more accurately. As traditional medicines are inexpensive, nontoxic and have little to no side effects, recognition of these medicinal plants is very useful. So in summary, our propositions are:

1. The main objective of this project is to recognize medicinal plants so that they can be used properly for the right treatment and also explain why this prediction is made for reliability.

2. Machine learning and deep learning methods for identifying the medicinal plant correctly.

3. Explanation for the predictions of the model with the highest accuracy.

1.4 Organization of the Report

The report is structured in the following manner -

Chapter-1 describes the motivation, organization of the report structured, aims and objective of my project work.

Chapter-2 examines the background study that led to the creation of the current system, which includes a thorough description of the numerous medicinal plant leaves and tries to explain the concepts related to the subject matter in detail.

Chapter 3 describes a short walk through of the previous studies.

Chapter 4 briefly explains the data collection procedures and the developed system's system design. It offers an explanation of the many techniques that were employed in the system's design and development.

Chapter-5 explains the application's development in depth, including the analysis methods utilized for the plant recognition methodology based on image processing and discusses the project work's results.

Chapter-6 presents the conclusion and recommendations for the future while describing the system's weaknesses.

Chapter 2

Background Study

2.1 Importance of Medicinal Plants

Plants are the primary source of many different types of medication, and they have a substantial impact on both plant and human health. Many herbs and plants are widely known for their medicinal and healing properties. Herbs are environmentally friendly, free of negative effects, readily available locally, and generally safe because they are natural goods. There are numerous herbs that are utilized throughout the year as medicinal plants (MPs) for a variety of ailments. In order to preserve human lives, it is necessary to encourage the usage of therapeutic herbs. Currently, these organic products are safer alternatives to fungicides and pharmaceutical medications. These MPs are thought to be safe for both people and the environment.Although the relevance of natural products has been diminished by the synthetic products of the contemporary era, it is now necessary to promote them globally in cases of both human and plant pathology [42].

Alternative medicine(AM) refers to the use of plants to accomplish a medical goal. Nearly every civilization has employed alternative medicine, especially western and Asian societies. Unfortunately, the majority of people today still think that a dose form is necessary for a drug to be reliable and effective. Despite the fact that many tablets and capsules, including paclitaxel, digoxin, and aspirin, are made from plant chemicals. Our forefathers used plants and herbs to flavor and preserve food, alleviate headaches, lessen discomfort, and even prevent diseases like epidemics. Through the ages, human groups have shared knowledge of these plants' therapeutic powers.The biological characteristics of plant species used around the world for a variety of purposes, including the treatment of infectious diseases, are typically due to active chemicals created during secondary metabolism. When used as a remedy, medicinal plants have numerous qualities, including the following:

• Synergic medicine- Each plant contains a variety of chemicals that may interact simultaneously to either enhance or impair one another's functions or counteract any potential negative consequences.

• Support for conventional medicine – Chemicals and plant-based compounds can be combined to provide the required results.

• Preventive medicine - Certain plant compounds have been shown to be useful in preventing or lowering the risk of contracting specific diseases, which can assist to lessen the burden and expense of utilizing chemical therapies [30].

Many disorders in both human and plant pathology are treated with natural compounds derived from medicinal plants (MPs) and their plant extracts (PEs). Therefore, it is crucial to comprehend the ideal bond between man and his surroundings. The amount of research being done and the use of MPs and PEs in treating human and plant illnesses is growing daily. Medical plants were traditionally utilized as a therapeutic agent for the treatment of many ailments like typhoid, gonorrhea, measles, and cholera before the growth and civilization of the world [42].

Medical plants are now recognized as a crucial resource for the treatment and averting a variety of diseases. Each plant has a number of vital components that have medical applications and can be used to create a variety of medications. Many underdeveloped or even wealthy nations use herbal therapy to preserve human welfare, improve individual health, and treat specific diseases like the common cold [30].These plants include Alpinia Galanga (Rasna),Amaranthus Viridis (Arive-Dantu),Azadirachta Indica (Neem), Basella Alba (Basale) ,Brassica Juncea (Indian Mustard) and others.

2.1.1 Alpinia Galanga

Alpinia galanga is a member of the ginger family and produces a rhizome that is mostly utilized as a herb in unani medicine. The leaves are rectangular, 2 to 6 cm long, with a thin, silky surface and very little projections that resemble hairs on both sides. Powerful antioxidants that are abundant in Rasna lower oxidative stress and prevent damaging free radicals from reaching the renal tissues. Rasna is a fantastic natural treatment for protecting and enhancing renal functions to ensure the best possible system purification. Rasna leaf extracts have amazing anti-inflammatory properties that help to relieve pain and increase range of motion in the wrist, elbow, shoulder, and hip joints. When it comes to coughing, colds, wheezing, bronchitis, and asthma, rasna is helpful [54].

2.1.2 Amaranthus Viridis (Arive-Dantu)

Amaranthus viridis is a tall, annual plant with weak branches that grows to a maximum height of 20–90 cm. The leaves are oblong to trapezoid in form and have whole edges. Vitamin A is abundant in amaranth leaves. By encouraging osteoblastic activity, it increases bone mass. Additionally, it helps patients with Alzheimer's disease since it reduces the brain's neuronal damage. Amaranth leaves are given to people who are recovering from illness or who are fasting because they are gentle on the stomach. They are helpful in the management of diarrhea and hemorrhages. Digestion has been shown to benefit from regular ingestion [33].

2.1.3 Artocarpus Heterophyllus (Jackfruit)

An evergreen with a height range of 8 to 25 meters, Artocarpus heterophyllus(Jackfruit). It has extensive taproot. Jackfruit comes in two primary varieties. One is a little,

squishy thing. Although slimy, it has a pleasant flavor. The flesh of the other variety of jackfruit is crispy and crisp but less sweet. Ayurvedic practitioners treat wounds, Bell's palsy, sperm count, physical strength, diarrhea, skin conditions, and cases of poisoning with the Ayurvedic tree Artocarpus heterophyllus.As a good source of fiber, jackfruit can help us feel fuller for longer and encourage regular bowel movements. These sores within our stomachs may not form due to the natural chemicals in jackfruit. The potassium in this tropical fruit may help lower blood pressure, lowering risk for heart disease, stroke, and bone weakening. The high vitamin C content of jackfruit may help shield our skin from sun damage. Natural substances called phytonutrients, like those in jackfruit, may have anti-cancer effects by inhibiting the development of cancer cells in the body [48].

2.1.4   Azadirachta Indica (Neem)

Azadirachta indica is a member of the Meliaceae family of mahogany trees. Tropical and semi-tropical climates are where it is commonly grown. Neem includes compounds that may lower blood sugar levels, treat stomach ulcers, prevent pregnancy, eliminate bacteria, and stop plaque from accumulating in the mouth [52].

2.1.5   Basella Alba (Basale)

A member of the Basellaceae family, Basella alba is a perennial edible vine. It is frequently employed in treating mouth ulcers in south India. In cases of dysentery and anemia, Ayurveda suggests Basella. The plant is abundant in minerals including iron and calcium as well as vitamins A and C, proteins, and carbs. Leucine, isoleucine, arginine, lysine, threonine, and tryptophan are the necessary amino acids that are found in it. As a result, it is employed to assess child malnutrition. It is said to have antioxidant effects and is also high in mucilage. There is evidence that the mucilage found in Basella alba has hypoglycemic properties. The ability to heal wounds and repair ulcers as well as to reduce inflammation has all been attributed to it. As a safe laxative for youngsters and pregnant women, decoction of the leaves is advised [51].

2.1.6   Brassica Juncea (Indian Mustard)

The cruciferous plant family includes the annual herb known as Brassica juncea or Brown Mustard. Additionally, it is utilized medicinally and as fodder. It is a traditional treatment for lumbago, rheumatism, arthritis, and foot pain. The seed is used as a treatment for tumors in China. Mosquitoes may become insect-repellent after ingestion. Headaches are believed to be relieved by applying leaves on the forehead. The leaves are consumed in soups to alleviate bladder hemorrhage or inflammation. The seeds are used in Korea to cure stomach problems, lumbago, rheumatism, colds, and abscesses. To treat ulcers and skin outbreaks, use brown mustard oil. In Tanzania, cows are given the roots to encourage milk output.

2.1.7   Carissa Carandas (Karanda)

Apocynaceae is a family of flowering shrubs that includes Carissa carandas.The regular consumption of karonda fruit is said to be advantageous for enhancing mental

wellness. Tryptophan, vitamins, and the presence of magnesium all contribute to the stimulation of the neurotransmitter serotonin, which improves general mental health. Karonda has been used for treating fever for antiquity since it contains a significant amount of vitamin C.Due to its abundance in antioxidants, the delicious fruit aids in reducing fever by warding off illnesses. Fruit that is ripe or dried can greatly reduce fevers [54].

2.1.8 Citrus Limon (Lemon)

The Rutaceae family of flowering plants includes the species of little evergreen trees known as citrus limon. Tea and grilled meats and shellfish are both prepared with lemon tree leaves. Skin cells that contain oil are used to make lemon oil. Numerous cuisines incorporate preserved lemon. Pectin, a polysaccharide used as a gelling agent and stabilizer in food and other applications, can be made from peel [46].

2.1.9 Ficus Auriculata (Roxburgh fig)

Wide-spreading evergreens like the Roxburgh Fig can be found in tropical lowlands that receive plenty of moisture. A thick, 8–15 cm long petiole with alternating leaves that are grouped around the stem apex. Vietnam uses roots and leaves to treat poisoning and lessen edema. In Nepal and India, leaves are crushed, and the resulting paste is applied on wounds. Dysentery and diarrhea are also treated with leaves and roasted figs. Diarrhea, cuts, and wounds can all be treated with stem bark juice. Mumps, cholera, diarrhea, and vomiting can all be treated using root latex. The fig fruit's fiber promotes weight loss.

2.1.10 Ficus Religiosa (Peepal Tree)

The leaves of the Ficus religiosa (peepal) tree have historically been used to cure conditions such as diabetes, jaundice, fever, nasal bleeding, and heart disease. In order to control diarrhea, peepal is a useful herb. Due to its Kashaya and Sangrahi qualities, peepal bark powder aids in regulating water loss from the body and thickens stools. Because of its laxative properties, peepal leaf juice or pills might help treat constipation. The big intestine can easily evacuate waste materials thanks to it [53].

2.1.11 Hibiscus Rosa-sinensis

Hibiscus Rosa-sinensis is used as a natural remedy to treat high blood pressure, cholesterol, and the spread of cancer. It has been reported that hibiscus tea provides medical benefits, including anticancer effects. A number of cancer cell types have been proven to be inhibited in their development by hibiscus extracts. Recent research has discovered that polyphenols extracted using organic solvents from Hibiscus sabdariffa can suppress the proliferation of melanoma cells in particular [13].

2.1.12   Jasminum (Jasmine)

Jasmine has been used to treat liver illness, liver scar pain, and stomach pain brought on by extreme diarrhea. Additionally, it is used to treat cancer, calm the body, increase sexual desire, and prevent strokes [52].

2.1.13   Mangifera Indica (Mango)

Since Mangifera indica is a glucosy xanthone and a polyphenolic antioxidant, it has strong anti-oxidant, anti-lipid peroxidation, immunomodulatory, cardiotonic, hypotensive, wound healing, anti-degenerative, and antidiabetic effects. As well as being used to cure piles, leucorrhoea, piles, asthma, bronchitis, cough, hypertension, and insomnia, the plant's various parts can be used as a dentifrice, antiseptic, astringent, diaphoretic, stomachic, vermifuge, tonic, laxative, and diuretic. Abscesses, broken horns, tumors, rabid dog or jackal bites, snakebites, stings, datura poisoning, heat stroke, miscarriages, anthrax, blisters, mouth wounds, tympanitis, colic, diarrhea, glossitis, indigestion, bacillosis, bloody dysentery, liver disorders, and excessive urination are just a few of the conditions that all of its constituent [1].

2.1.14   Mentha (Mint)

There are thought to be between 13 and 24 different species, though the exact differences are uncertain. One of the oldest and best-known plants in the world, the Mentha species is commonly used in cooking, cosmetics, and complementary or alternative therapies, especially for the treatment of gastrointestinal diseases like flatulence, indigestion, nausea, vomiting, anorexia, and ulcerative colitis. Mentha spp. essential oil is also used as an antibiotic and a component of several analgesic lotions for the topical treatment of oral mucosal irritation. In addition to treating diverticulitis, secondary amenorrhea, oligomenorrhea, myalgia, neuralgia, irritable bowel syndrome, bile duct discomfort, and pain associated with irritable bowel syndrome, the oil from Mentha spp. has been licensed for use internally. Additionally, it is employed to treat irritable bowel syndrome [49].

2.1.15   Moringa Oleifera (Drumstick)

Moringa oleifera is widely cultivated for its immature seed pods and leaves, which are used in traditional herbal medicine and as vegetables. It is also used as a water filter. Moringa leaves are a rich source of the vitamins A, C, B1 (thiamin), B2 (riboflavin), B3 (niacin), B6, and folate.In addition, they are abundant in zinc, calcium, phosphorus, iron, magnesium, and iron. Amino acids, the components of proteins, are abundant in moringa leaves. They include 18 different types of amino acids, and each one of them is vital to human health. The onset of diabetes in people is triggered by persistently excessive blood sugar levels. In turn, diabetes can harm internal organs and result in heart issues. Keep the blood sugar levels under control to prevent this.Given that isothiocyanates are present in moringa leaves, they are an ideal source for that since they help to balance blood sugar levels.

2.1.16 Muntingia Calabura (Jamaica Cherry-Gasagase)

The emergence of several diseases, both infectious and non-infectious, has recently become a serious global issue. Unfortunately, all pharmaceutical medications have a variety of adverse effects despite being produced in large quantities and given as prescriptions to patients to treat illnesses. So the fruit can lessen your risk of developing other bacterial digestive troubles and enhance gut health thanks to its anti-inflammatory and antibacterial characteristics, which also help prevent and treat gastric ulcers by removing dangerous bacteria [45].

2.1.17 Murraya Koenigii (Curry)

Murraya koenigii is a vitamin-rich plant that is also used as a carminative, painkiller, and to cure stomach aches. Because bioactive components like carbazole alkaloid are present, it possesses anti-inflammatory, antioxidant, and anticancer activities [8].

2.1.18 Nerium Oleander (Oleander)

Despite the danger, nerium oleander seeds and leaves are used to make medication. Oleander is also used to treat heart problems, asthma, epilepsy, cancer, painful menstrual cycles, leprosy, malaria, ringworm, dyspepsia, and venereal illnesses in addition to being used to induce miscarriages. A constant combination of oleander leaf powdered extract, pheasant's eye fluid extract, lily of the valley fluid extract, and squill powdered extract has been used to treat mild heart failure.Oleander is occasionally administered topically to the skin to treat warts and other skin issues [39].

2.1.19 Nyctanthes Arbor-tristis (Parijata)

Nyctanthes arbor-tristis is utilized as an antibiotic, anti-inflammatory, and anthelmintic in addition to treating fever, fungal skin infections, and other conditions. Children are administered bitter leaf extract to get rid of their roundworms and threadworms. Leaf juice is used as an antidote for reptile venoms and snake bites, as well as for sciatica, rheumatism, and fever. The astringent and bitter blooms are used as a carminative and in the treatment of eye problems [14].

2.1.20 Ocimum Tenuiflorum (Tulsi)

The Ocimum tenuiflorum plant is well recognized in Ayurveda for its therapeutic and spiritual benefits, which include treating conditions like cough, asthma, diarrhea, fever, dysentery, arthritis, eye illnesses, indigestion, and gastrointestinal disorders. The numerous therapeutic benefits of tulsi, including its antibacterial, antifungal, antipyretic, antioxidant, antiseptic, and anticancer characteristics, are well established. The genes and enzymes in charge of producing 14 of the almost 40 secondary metabolites with therapeutic potential have been identified on the human genome. Tulsi has modest diuretic and detoxifying properties that aid in decreasing the body's uric acid levels. The acetic acid in holy basil aids in the removal of stones [15].

2.1.21 Betle (Betel)

The piper betel leaf is a source of phenolic substances with anti-bacterial, anti-mutagenic, anti-proliferative, and antioxidant activities. Betel leaves have been shown to have chemo-preventive potential against many forms of cancer in studies. Numerous anti-diabetic medications have long-term adverse effects on the liver and kidneys. In those with newly diagnosed type 2 diabetes mellitus, dried betel leaf powder has been shown in studies to lower blood glucose levels, and it has no negative side effects. Escherichia coli, Staphylococcus aureus, and Pseudomonas aeruginosa are only a few examples of the harmful bacteria that are resistant to the essential oil found in betel leaves' leaves. Additionally, the betel leaves' phenolic and phytochemical content aids in their ability to defend against both gram-positive and gram-negative germs [55].

2.1.22 Plectranthus Amboinicus (Mexican Mint)

Plectranthus amboinicus has a long list of health advantages, such as the capacity to enhance skin, detoxify the body, battle coughs and colds, lessen arthritis pain, lessen stress, and enhance digestion. It is frequently used to cure ailments like colds, asthma, constipation, headaches, coughs, fevers, and skin problems in traditional medicine [11].

2.1.23 Pongamia Pinnata (Indian Beech)

Pongamia pinnata has been used as a crude medicine to cure ulcers, tumors, piles, and skin conditions. The root can be used to treat vaginal and skin conditions as well as gonorrhea, as well as to clean teeth, gums, and ulcers. Pongamia pinnata leaves have gastrointestinal, laxative, and anthelmintic properties. They also work well for diarrhea, leprosy, dyspepsia, and cough. Flowers can help to relieve vata and kapha as well as diabetes-related dipsia [7].

2.1.24 Psidium Guajava (Guava)

Psidium guajava has a high concentration of both organic and inorganic substances, including secondary metabolites such as polyphenols, antioxidants, antiviral substances, and anti-inflammatory substances. Many of the substances in guava have anticancer properties. There are more vitamins and minerals in it. The guava contains significant amounts of phenolic substances such as flavonoids. Important antioxidants include lycopene and flavonoids. Myocardium inotropism may be impacted by guava. After receiving treatment for 21 days, guava skin extract can reduce blood sugar levels [26].

2.1.25 Punica Granatum (Pomegranate)

Punica granatum can reach heights of 5 to 10 meters. Numerous tribes and civilizations have employed pomegranates for thousands of years to treat a variety of illnesses. It has several health advantages as well as excellent nutritional value. Pomegranates are used in the treatment of cancer, osteoarthritis, and other illnesses. The pomegranate has been used in natural and holistic medicine to treat arthritis,

sore throats, coughs, urinary infections, digestive issues, skin issues, and tapeworms. Modern research, however, raises the possibility that pomegranates could help with the treatment of major illnesses like diabetes, osteoarthritis, prostate cancer, and skin cancer. According to clinical studies, pomegranates may help prevent heart disease, heart attacks, and strokes when included in a balanced diet. This is due to the fact that pomegranates may thin the blood, boost heart blood flow, lower blood pressure, remove plaque from the arteries, and lower bad cholesterol while raising good cholesterol. Syphilis is treated using a decoction of the seed. Juice used to treat diarrhea and jaundice. Flowers' juice is used to stop nosebleeds [5].

2.1.26   Santalum Album (Sandalwood)

Small tropical trees called Santalum album are the original source of sandalwood oil. The characteristics of sandalwood include diuretic, antiseptic, anti-scabietic, and antipyretic. The treatment of bronchitis, cystitis, dysuria, and illnesses of the urinary system is also successful with it. The primary component of sandalwood oil is -santalol, which has a variety of medicinal uses. White sandalwood is used in the treatment of the common cold, cough, bronchitis, fever, sore mouth, and throat. It is also used to treat gonorrhea, heatstroke, liver illness, gallbladder problems, and urinary tract infections in addition to conditions of the heart and blood vessels [43].

2.1.27   Syzygium Cumini (Jamun)

Syzygium cumini is regarded as a "magic tree" since it has multiple medical benefits from the ground up through the leaves, fruits, and even the bark. While the leaves of the tree are helpful for gum and tooth disorders, the pulp and seeds are important for treating diabetes. Jamun is a potent source of several different nutrients. It contains a wide range of vitamins, including vitamins A and C, calcium, iron, fiber, magnesium, potassium, glucose, protein, and other crucial minerals [47].

2.1.28   Syzygium Jambos (Rose Apple)

Syzygium jambos provide dietary fibers that lessen bloating and constipation. The rose apple's seeds also control gastrointestinal issues that result in diarrhea and vomiting. About 29 milligrams of calcium, a mineral crucial for healthy bones and overall wellbeing, may be found in a 100-gram serving of raw rose apples. Magnesium and potassium, which are also essential for bone health, are also present in rose apples in good quantities. The high water content of the rose apple aids in controlling body temperature. Damaged cells and bodily tissues can be repaired with its aid [56].

2.1.29   Tabernaemontana Divaricata (Crape Jasmine)

The root of Tabernaemontana divaricata is used to cure scabies, toothaches, headaches, and hypertension. To treat snake and scorpion venom, the roots, leaves, and petals are all used. The wood can be used for incense and perfume and the pulp surrounding the seed can be used to create red dyes for non-medical purposes. If utilized improperly, the plant's entire composition is lethal. The presence of Tabernaemontana

divaricata is significant because it acts as a host plant for the oleander hawk-moth caterpillars [35].

2.1.30   Trigonella Foenum-graecum (Fenugreek)

The Fabaceae family includes the annual plant fenugreek, which has three slender, obovate to oblong leaflets on each leaf. When drunk with lemon and honey, the cold water extract of T. foenum-graecum, popularly known as fenugreek tea, has historically been used to treat respiratory infections and to nourish the body when ill. It appears that fenugreek slows the stomach's absorption of sugar and increases insulin levels. For those with diabetes, both of these outcomes lower blood sugar. To increase interest in sex, fenugreek may also increase levels of estrogen and testosterone [52].

2.2   Deep Learning

A computer model learns to do classification tasks directly from images when it is part of the deep learning subclass of machine learning. Using a stacked hierarchy of concepts, each concept is defined in relation to simpler concepts, and more abstract representations are formed in less abstract ones, a particular sort of machine learning learns to describe the world. A computer model can directly learn to carry out categorization tasks from images, text, or voice via a process called deep learning. Deep learning models occasionally beat human performance in terms of accuracy. An enormous amount of labeled data and multi-layered neural network architectures are used to train the models [29].

2.2.1   Convolutional Neural Networks(CNNs)

As a subclass of Deep Neural Networks, CNNs are often used for the study of visual images because they can recognize and classify certain aspects in images. They can be used for a number of things, such as image and video recognition, classification of images, picture analysis for medical use, computer vision, and natural language processing. Given that a CNN uses 2D convolutional layers and mixes input data with learned features, it is especially well suited to processing 2D data, such as images.Using tens or hundreds of hidden layers, CNNs are taught to recognize various aspects in a picture. The complexity of the inferred image features grows with each concealed layer.For instance, the first hidden layer might pick up on edge detection, while the last one might pick up on more advanced form detection tailored to the geometry of the object we're trying to identify. CNN is an effective image processing method. Three different layers are available in convolutional neural networks.Here is the figure 2.1 which shows the CNN model basic architecture.

A. Convolutional Layer:

This is the first layer that is used to extract the various features from the input photos. At this layer, a mathematical operation called convolution is performed between the input picture and a filter with the dimensions MxM. By sliding the filter over the input image, the dot product is obtained between the filter and the elements of

14
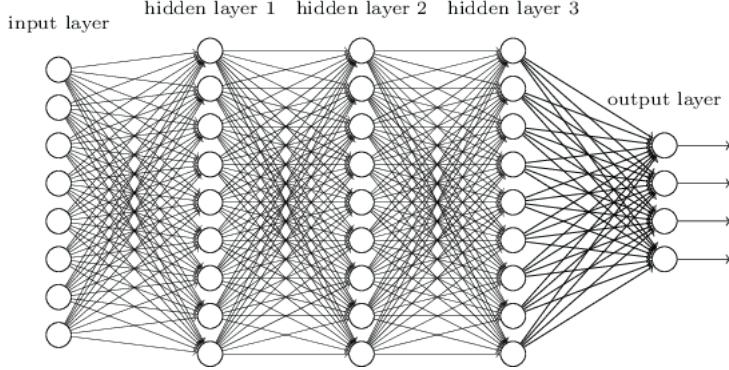
Figure 2.1: Convolutional Neural Networks(CNNs) Architecture

the input image in relation to the filter's size (MxM). This gives information about the image, including its corners and edges, and is known as the feature map. Later, additional layers are given access to this feature map so they can learn new features from the input image. Each input neuron in a conventional neural network is linked to the subsequent hidden layer.In CNN, the input layer neurons and hidden layer neurons are only loosely coupled.

B. Pooling Layer:

After a convolutional layer, a pooling layer is frequently applied. The major objective of this layer is to reduce the size of the convolved feature map in order to reduce computation costs. This is achieved by minimizing the connections between layers and working independently on each feature map. Depending on the technique used, there are many kinds of pooling procedures. It is essentially a summary of the features that a convolution layer produced. The feature map provides the largest contribution to Max Pooling. Through average pooling, the average of the components in a region of an image with a predetermined size is calculated. Sum Pooling calculates the components' cumulative sums inside the given section. Usually, the FC Layer and the Convolutional Layer are connected by the Pooling Layer. The dimensionality of the feature map is reduced using the pooling layer. There will be numerous activations and pooling layers inside the CNN hidden layer.

C. Fully Connected Layer:

The neurons as well as the weights and biases make up the Fully Connected (FC) layer, which connects the neurons between two layers. A CNN design frequently places the last few levels before the output layer. In this, the input image from the layers beneath is flattened and provided to the FC layer. The mathematical function's operations are often carried out in a few more FC levels after the vector has been flattened. At this point, the classification process gets started. Two layers are joined because they will function more effectively as a pair than as a single layer will. These CNN layers minimize human oversight.Fully Connected Layers are the network's final layers. The output of the previous pooling or convolutional layer is flattened and then sent into the fully connected layer as its input [44] .

2.2.2 VGG19 Pre-trained Model

VGG or Visual Geometry Group is a typical deep Convolutional Neural Network (CNN) architecture containing a number of layers. With VGG-16 or VGG-19 comprising 16 and 19 convolutional layers, respectively, the "deep" refers to the total number of layers. Groundbreaking object recognition models are based on the VGG architecture. It uses the 'Imagenet' dataset as the baseline dataset. The VGGNet, created as a deep neural network, outperforms baselines on numerous tasks and datasets besides ImageNet. Furthermore, it remains one of the most widely used image recognition architectures today [36].

VGG19 is a sophisticated CNN with pre-trained layers and a keen understanding of how shape, color, and structure create an image. A pre-trained model is one that was developed by another person to address a comparable issue. The extremely deep VGG19 has been trained on a massive amount of different images for challenging classification tasks. It is a variation of the VGG model with 19 layers in total (16 convolution layers, 3 Fully connected layers, 5 MaxPool layers and 1 SoftMax layer).This model architecture descriptions are given below:

- This network received a fixed-size (224 x 224) RGB picture as input, indicating that the matrix was shaped (224,224,3).That means RGB value of each pixel, calculated throughout the whole training set, was the only preprocessing that was carried out.

- They were able to cover the entirety of the image by using kernels that were (3 x 3) in size with a stride size of 1 pixel.

- To maintain the image's spatial resolution, spatial padding was applied and Sride 2 was used to conduct max pooling over a (2 x 2) pixel window.

- Rectified linear unit (ReLu) was used after this to add non-linearity to the model in order to enhance classification accuracy and computation time. As opposed to earlier models that used tanh or sigmoid functions, this one performed far better.

- Built three fully linked layers, the first two of which were 4096 in size, followed by a layer with 1000 channels for 1000-way ILSVRC classification, and a softmax function as the third layer [40].

2.2.3 Explainable AI Grad-CAM

A collection of procedures and techniques known as explainable artificial intelligence (XAI) enables human users to understand and believe the output and outcomes produced by deep learning algorithms. When describing an AI model, its anticipated effects, and any biases, the term "explainable AI" is employed. It contributes to the definition of model correctness, fairness, transparency, and results in AI-powered decision making. When putting AI models into production, a company must establish trust and confidence by using explainable AI. AI explainability aids a company's adoption of a responsible strategy for AI development [50].

Guided Grad-CAM, which produces high resolution detail of the target class in an image, is created by combining gradient weighted Class Activation Mapping (Grad-CAM) with pre-existing pixel-space gradient visualizations. To locate and draw attention to certain target regions in the image, it employs gradients of a specific target that pass through the convolutional network. The potential for Grad-CAM is enormous in computer vision applications like picture classification, object recognition, semantic segmentation, image captioning, and visual question answering, among others.

Grad-CAM Architecture uses CNN to forward-propagate the image and the class of interest, and task-specific computation to set the gradient of the desired class to 1 and make the rest of the gradients 0. The coarse Grad-CAM localization is then computed by back propagating this through rectified convolutional feature maps of interest (blue heatmap). To produce Guided Grad-CAM visualizations, the heat map is pointwise multiplied with guided backpropagation [31].

2.3   Machine Learning

Machine learning (ML), a kind of artificial intelligence, enables computer systems to learn from the past and adjust as necessary without the direct intervention of a programmer.Image classification is a common application for machine learning. Machine learning (ML) gives computers the ability to behave strikingly like people. In order to extract the essential information from the vast volume of data, ML can be used to construct algorithms based on data patterns and correlations. The application of ML can be advantageous in a variety of fields, including intrusion detection, bioinformatics, healthcare, marketing, gaming, and others. It allows the computers or machines to make data-driven decisions rather than being explicitly programmed to carry out a certain activity. These programs or algorithms are made in a way that, when they come into contact with new or uncovered data,, they learn and develop over time [41].

2.3.1   Feature Extraction

Features are elements or patterns that assist identify an object in an image. The main step in plant recognition is feature extraction, which involves identifying different leaf traits. Specific plant leaf features have been classified based on their various sizes and forms [37].I mainly used the normal feature,grayscale feature, HOG feature and combined feature.

A.Gray Scale Feature:

Grayscale is applied to the image. The grayscale photos were processed using image contrast and intensity enhancement methods, and after that, they were stacked as slices [22]. Grayscale refers to the spectrum of monochromatic (gray) tones, from pure white at the lightest end to pure black at the darkest end. Since there is no color information in a grayscale image, the highest luminance is white and the lowest luminance is black; everything in between is a shade of gray. As a result, grayscale

photographs only feature various degrees of gray rather than color.Achromatic is another name for grayscale. The only part of an image that is not eliminated is the grayscale; without it, the image would always be pure black, regardless of the amount of color information present. Red, green, and blue (RGB)-colored groupings of three pixels, also known as channels in digital imaging, make up a digital image. Each channel also has a brightness value that describes the color's lightness or darkness. When the color information from each channel is removed to make a grayscale image, leaving only the luminance values, the image changes into a pattern of light and dark patches lacking in color, producing a black and white image. Even the most basic digital photography software packages include the ability to grayscale an image. This is especially important for printing because color printing requires all three print colors (cyan, magenta, and yellow) while monochrome printing just uses black ink [24].

B.Histogram of Oriented Gradients (HOG)Feature:

In computer vision and image processing, the histogram of oriented gradients is a feature descriptor for object detection. This technique counts the instances of gradient orientation in the focused region of a picture. An overlapping local contrast normalization technique is used to compute HOG over a dense grid of cells that are evenly spaced out.Calculating the gradient values is the initial stage in the calculating process. The most popular method is the one-dimensional point discrete derivative mask in one or both of the horizontal and vertical axes. In the second phase, cell histograms are created using the values from the gradient computation, and each pixel within the cell casts a weighted vote for an orientation-based histogram channel. A larger, physically connected block is created by combining these cells. To accommodate for differences in light and contrast, the gradient strengths must be locally normalized. The vector for the HOG feature is then constructed using the normalized cell histograms from each block region. Each cell may contribute to the final description more than once as a result of these blocks' potential for overlap [23].

The hog features descriptors, which count orientation in certain areas of an image, and are used for item detection. This method is similar to the scale-invariant feature, but it varies in that it computes on a grid of cells with a dense number of cells and uniform spacing and employs overlapping local contrast normalization for more accuracy [22].

2.3.2 Support Vector Machine (SVM)

Support Vector Machines (SVMs) are a special type of classifier that are a group of similar supervised learning techniques used specifically for classification. The space's separating hyperplane, produced by SVM, will widen the gap between the two data sets. The creation of two parallel hyperplanes, one on either side of the separating hyperplane between the two data sets, serves as the border. Since each data point in this model is represented by a p-dimensional vector, it is required to investigate whether SVM can split such points with a p-1dimensional hyperplane. A linear classifier is what this is.Support vector machines are linear classifiers, and because

of their capacity to find the best hyper plane to enhance the separation between patterns, SVMs have the potential to be useful for classification tasks. Training (F1) and testing (F2) data are originally split up randomly from the total collection of data. The features in F1 are trained, and those in F2 are predicted, using an SVM classifier [3].

The idea of decision planes, which specify decision boundaries, serves as the foundation for support vector machines. A decision plane is a diagram that distinguishes between a collection of objects with various class memberships. A binary linear non-probabilistic classifier is created using an SVM training method that assigns new instances to one of two categories. New instances are then projected to belong to a category and mapped into the same space based on which side of the gap they are on. It can be used for tasks like classification and regression, among others. The hyperplane with the longest distance to the closest training data point of any class is able to achieve a respectable separation since, generally speaking, the bigger the margin, the smaller the generalization error of the classifier [22].

Chapter 3

Related Works

Begue, A. et al. [19] discovered that, with an accuracy of 90.1%, the random forest classifier outperformed other machine learning methods such as the k-nearest neighbor, naive Bayes, support vector machines, and neural networks. A database of tropical island Mauritius has 30 pictures of 24 different plant types. Numerous leaf characteristics, such as the circumference, number of vertices, length, width, circumference and area of the hull, area, and color, were retrieved from the B. Additionally, derived features were extracted from these fundamental traits.Future research will focus on employing a larger dataset and high-performance computer resources to study the effectiveness of deep learning neural networks to detect medicinal plants used in primary healthcare. These results are highly encouraging.

Dileep M.R. and Pournami P.N.[29] proposed the AyurLeaf Convolutional Neural Network (CNN) model, which uses deep learning to classify medicinal plants based on characteristics of its leaves, such as size, shape, color, and texture. This paper proposes a shared dataset for medicinal plants, which are widely distributed in Kerala, a state on the southwest coast of India. There are leaf samples from 40 medicinal plants in the proposed dataset. Finally, Softmax and SVM classifiers are used to perform the classification. On the AyurLeaf dataset, the model had a classification accuracy of 96.76% after five cross validations. AyurLeaf offers a simple way to recognize and categorize medicinal plants and aids in the preservation of the traditional medical knowledge passed down from our ancestors.

Amala Sabu et al. [23] used a machine learning and computer vision technique to categorize the medicinal plants present in the Western Ghats. The method classified the collected SURF and HoG features using the kNN classification algorithm. The leaf veins are modeled using HoG features, whereas the twenty different places of interest on the leaves are represented using SURF feature descriptors. This method uses the k-NN classification algorithm to classify leaves with k values equal to 1. Even though this model's feature extraction technique offers accuracy of more than 96%, it requires a lot of processing power.

Priya Pinder Kaur and Sukhdev Singh [37] talked about an automated system that calculates many aspects of leaf morphology. They used two classifiers, the SVM model (support vector machine) and the KNN model, to accurately label the plants (k-nearest neighbor).

The goal of Muhammad Azfar Firdaus Azlah et al. [28] was to study and analyze the performance of various plant classification approaches.

For the aim of classification, ArunPriya C. and Balasaravanan T. [3] employed an effective machine learning method. On 12 leaf features that were orthogonalized by Principal Components Analysis, they employed the SVM classifier (PCA).

The automatic technique utilized by Gopal, A. et al. [2] identified medicinal species. 100 leaf photos, ten from each of the ten plant species, were included in the dataset. Identification based on boundary- and moment-based features with a classification method using the optical color of leaves. 50 medicinal leaves of various types were tested, and the results of the experiment indicated 92% accuracy, with 46 of them being accurately identified.

An ANN-based model was created by R. Janani et al. [6] to categorize different kinds of medicinal plants based on the color, texture, and form of their leaves. A total of 63 leaves were used by the model: 20 for testing, 7 for validation, and 36 for training. Eight minimally prevalent traits were found to categorize the leaves out of the 20 various features of leaves. Compactness, eccentricity, skewness, kurtosis, energy, correlation, sum-variance, and entropy are the eight key properties. This method's accuracy was 94.4%.

A computer vision-based method was put forth by D. Venkatarama et al. [18] to identify the feature set needed to categorize the medicinal plant leaves that were provided and to extract their medicinal properties. To determine the leaf class, this technique employs a probabilistic neural network classifier. Preprocessing, feature extraction, categorization, and retrieval of medicinal values are the phases that make up the technique. The calculation of the feature vector and similarity matching are required for classification.

In order to detect and categorize herbal medicinal plants, Vijayshree, T. et al. [27] developed an automatic approach. Utilizing color, texture, and shape, 500 leaves and 21 traits with 50 distinct species were retrieved. The results of the experiment demonstrate that while using all features results in 99.2% accuracy, using neural networks for texture feature identification and classification achieved 93.3% accuracy.

Dahigaonkar, T.D., et al. [25] identified the ayurvedic medicinal plants that used them based on the form, color, and texture of the leaves. Using characteristics like solidity, eccentricity, entropy, extension, standard deviation, contrast, equivalent diameter, mean, and class, the total number of unique plants is 32. Using a support vector machine classifier, they were able to achieve an accuracy of 96.66%.

Kan.H.X, et al. [21] created an automated method for classifying medicinal plants that uses their leaves. They take 10 shape features and 5 texture features, and then they use an SVM classifier to categorize the plant. The dataset contains 240 unique plant leaves from the medicinal plant specimen library of Anhui University

of Traditional Chinese Medicine, with a 93.3% recognition rate in an experiment.

De Luna, R. G. et al. [20] developed a system for categorizing Philippine herbal medicine plants based on leaf characteristics using ANN.The data set includes a total of 600 images from 12 plants, with 50 images from each plant, and includes a number of fundamental characteristics that define the shape, venation, and structure of leaves, including length, width, area, circumference of the convex hull, circularity, solidity, and convexity. Data analysis and classification are done using Python and Matlab. In trials using a range of methods, such as logistic regression, naive bayes, linear discriminant analysis, KNN, classification and regression trees, SVM, and ANN, a 98.6% recognition rate for the identification of herbal plants was attained.

An intelligent technique for classifying plant leaves for medicinal use was put forth by B. Sekeroglu et al. in their paper [16]. Back propagation neural networks were used to identify and recognize leaves from 27 different types. The test employed three groups of four sets of images, and the first group had a recognition rate of 97.2%, the second had a recognition rate of 91.7% under high noise settings, and the third had a recognition rate of 79.6% under conditions of extremely high noise rate. The ANN classification of plants produced the desired outcomes.

Purohit, S. et al. [10] presented a method for automatically classifying plants using a machine learning methodology. To automatically recognize plants by their various parts, such as stems, flowers, leaves, fruits, and leaf scans, "Plantector" was developed. The Image llef dataset has been utilized with 50 different plant species. Plant classification was done using SVM and KNN classifiers. 10 plant classes were used in the experiment, which was divided into five categories: fruit, flower, stem, leaf, and leafScan. Flower category accuracy was 73.18% using a color, shape, and texture-based method, while fruit accuracy was 67.33%, leaf accuracy was 69.17%, stem accuracy was 76.57%, and leaf accuracy was 98% using a SIFT-based technique.

Jamil, N., et al. [9] created an automated technique in which Malaysian leaves were used to identify herbal medicinal plants. For the experiment, 455 herbal medicinal leaves from different classes and 65 from within each class were gathered. Color, texture, segmentation, shape-based fractal analytic features, and scale invariant feature transform were all used. Twelve geometric changes were applied to images of various sizes and orientations. Three of the aforementioned traits combined resulted in a 94% recognition rate, while just one texture feature was enough to identify herbal plants with 92% accuracy. 70% of the leaves were used for training, with 30% being used for testing.

Images taken from the outer edge of the Western Ghats were employed in an automated system for medicinal plant leaf recognition by Arun, C.H. et al. [4]. 250 leaves from five different therapeutic plants are included. Gray level, GTSDM, and L BP for plants are utilized as textural characteristics for classification, and several classifiers such SVM, DT, ET, SG D, KNN, and RF are used. The results of these classifiers are then compared. SGD, DT, and K NN were used to categorize plants with a maximum accuracy of 94.7% using only 75 leaf pictures for testing.

Chapter 4

Methodology

Methodology refers to the general plan and purpose of a project.It entails researching the theories and ideas underlying the methodologies employed in the project field in order to create a strategy that is aligned with the goals. The methodology section outlines the procedures to be followed to investigate the problems and the justification for the application with specific procedures or techniques used to identify, select, process, and analyze information applied to understand the problem, allowing the readers to critically evaluate a study's overall validity and reliability. As a result, methodology provides a number of benefits, including being a simple method for learning and making it easier to go back to prior steps when there are errors or problems.

4.1   Dataset Description

Utilizing medicinal plants is seen to be relatively safe as there are either no or very few side effects. These medications have a strong affinity towards nature, which is their main benefit. I gathered Kaggle picture datasets of leaves from medicinal plants. Here's the link: https://www.kaggle.com/datasets/riteshranjansaroj/segmented-medicinal-leaf-images. The dataset comprises of thirty species of healthy medicinal herbs such as Santalum album (Sandalwood), Muntingia calabura (Jamaica cherry), Plectranthus amboinicus / Coleus amboinicus (Indian Mint, Mexican mint), Brassica juncea (Oriental mustard), and many more. The dataset consists of 1835 images of thirty species.There are between 50 and 130 high-resolution images for each species. The folder names match the species' botanical and scientific names. The leaves came from a variety of the same species of plants that were discovered in Indian neighborhood gardens. The leaves are healthy and ripe. Figure 4.1 shows the Pie Chart of Datasets.
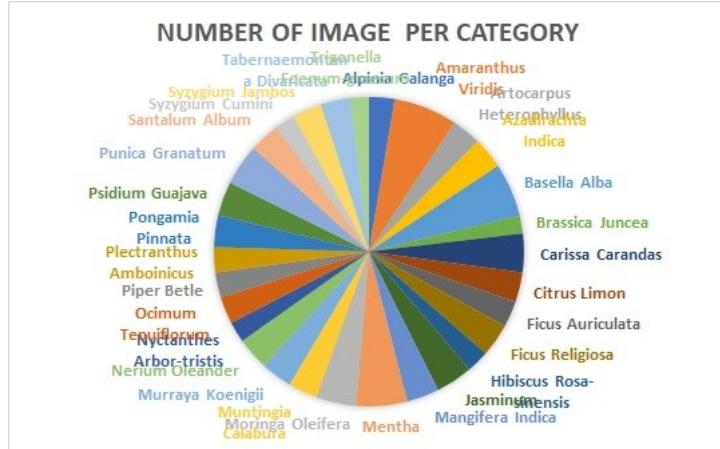
Figure 4.1: Pie Chart of Datasets

4.2 Proposed Method

As I have used 2 approaches, namely deep learning and machine learning approach for the classification of leaf images, I'll explain them in the following section:

4.2.1 Deep Learning Approach
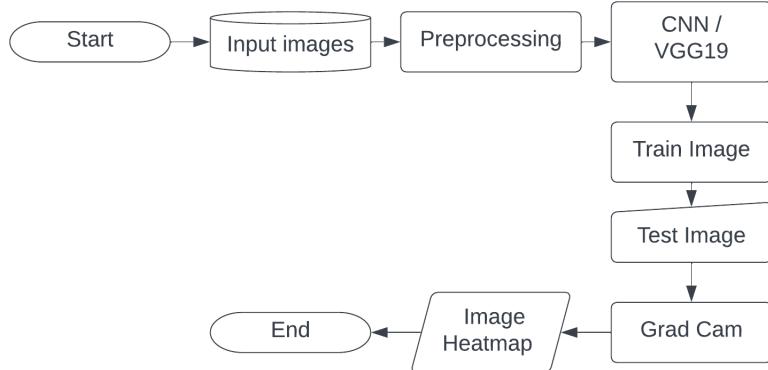
Figure 4.2: Deep Learning Proposed Method Architecture

Figure 4.2 shows the deep learning proposed method architecture.First, I input the leaf image dataset and preprocess them. For CNN, I split the dataset into train and test subsets. After that, I prepared the CNN model and trained it with the train dataset. Next, I inferred the model with the test image. For VGG19, I split the dataset into three subsets: train, validation and test datasets. Next I prepared the model with a VGG19 pretrained model and trained it with the train and validation dataset. The train dataset is for training the model and validation set is to validate the training. Next I evaluated the model with the test dataset. Then I implemented Grad Cam on the first 15 images of the predicted test images. This was done to

generate the heatmaps for checking the class activation visualization of the images.

4.2.2 Machine Learning Approach
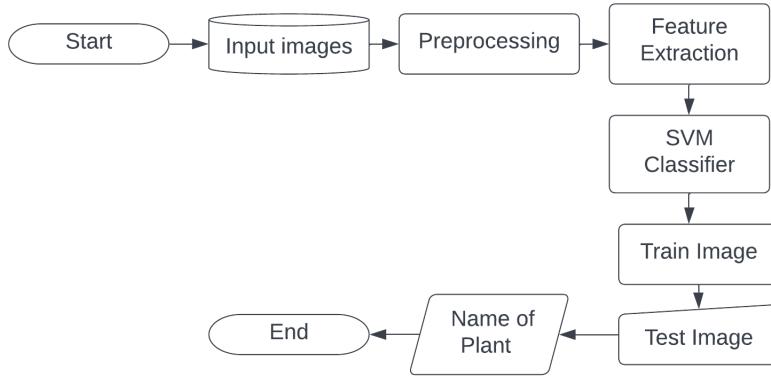
Figure 4.3: Machine Learning Proposed Method Architecture

Figure 4.3 shows the machine learning proposed method architecture. First, I input the leaf image dataset and preprocess them. Next, I extracted the features from the image and saved them in a feature dataset. Then I split the feature dataset into train and test subsets. After that, I trained the SVM classifier with the train dataset. Next, I inferred the model with the test dataset. Finally, I checked a test image to ensure the correctness of the prediction.

4.3 Process Flow

4.3.1 CNN

The flow diagram of proposed Medicinal Plant recognition using CNN is shown in figure 4.4. Here I used leaf images of these plants for recognition. At first,I loaded the data. The dataset consists of 1835 images of thirty species. Among the 1835 images, I used the one single image as a sample image. As the images are of various sizes, I resized the images to (180,180,3). Then split the dataset into two parts as training set and Validation set. It is splitted by the ratio of 80% and 20%. That means the train dataset has 1468 images and the validation dataset has 366 images.The first 16 images for training dataset and validation dataset are respectively shown in figure 4.5 and figure 4.6.
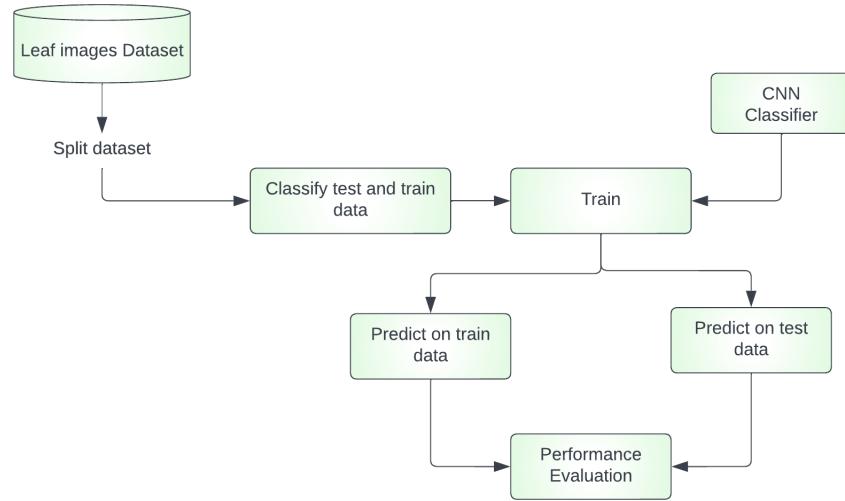
Figure 4.4: Flow Diagram of Proposed Medicinal Plant Recognition using CNN
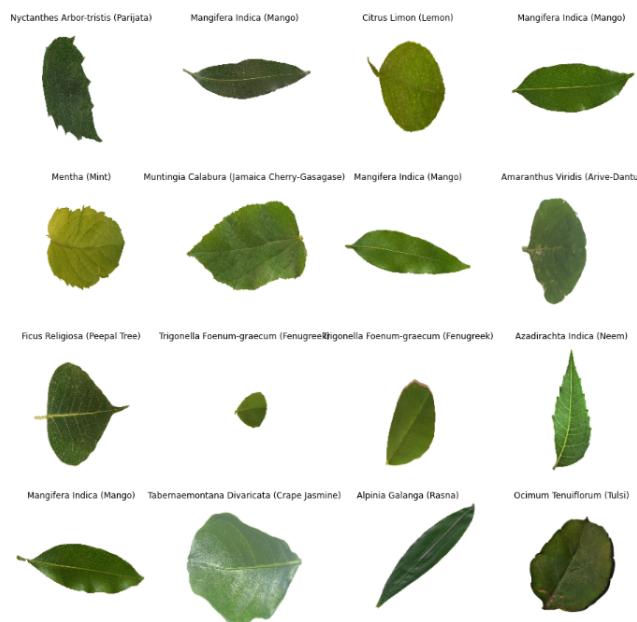
Figure 4.5: Sample images for training dataset

Figure 4.6: Sample images for validation dataset

4.3.2 SVM

The flow diagram of proposed Medicinal Plant recognition using SVM is shown in figure 4.7. Here I used leaf images of these plants for recognition. At first,I loaded the data. The dataset consists of 1834 images of thirty species. Here I extracted 3 types of features and used SVM on them individually. Here, the features are normal features which is just the image as a numpy array, grayscale features and HOG features. Next, I combined the three features and again ran SVM with the combined features.
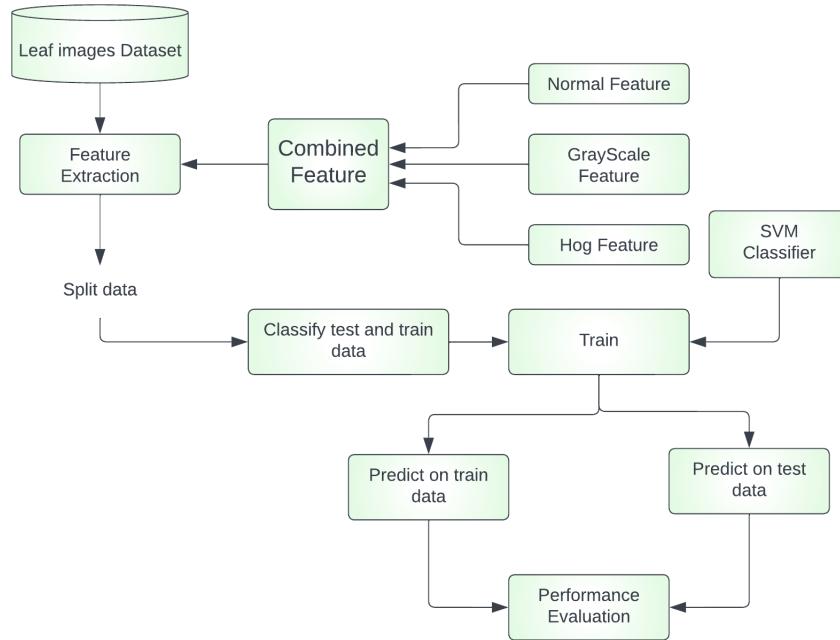
Figure 4.7: Flow Diagram of Proposed Medicinal Plant Recognition using SVM

4.3.3 VGG19

The flow diagram of proposed Medicinal Plant recognition using VGG19 pre-trained model is shown in figure 4.8. Here I used leaf images of these plants for recognition. At first,I loaded the data. The dataset consists of 1834 images of thirty species. As the images are of various sizes, I resized the images to (224,224,3) because the VGG19 set only takes input of (224,224,3). Then split the dataset into two parts as training set and test set. It is splitted by the ratio of 90% and 10%. Then from the train dataset, I splitted 20% of it to form a validation set. That means the train dataset has 1320 images,the validation dataset has 330 images and the test dataset has 184 images.

4.3.4 VGG19 with GradCam

The flow diagram of proposed Medicinal Plant recognition using VGG19 pre-trained model with GradCam is shown in figure 4.9.After running VGG19 as the previous, I ran GradCam explainable AI to get a better understanding of the predictions. Here I ran it on the top 15 predicted images to generate a heatmap of the activated areas that are responsible for the predictions.
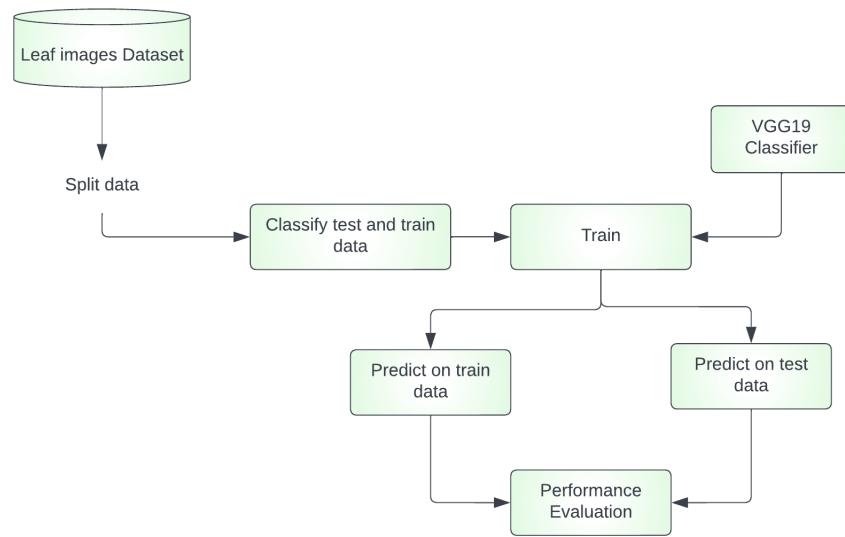
Figure 4.8: Flow Diagram of Proposed Medicinal Plant Recognition using VGG19
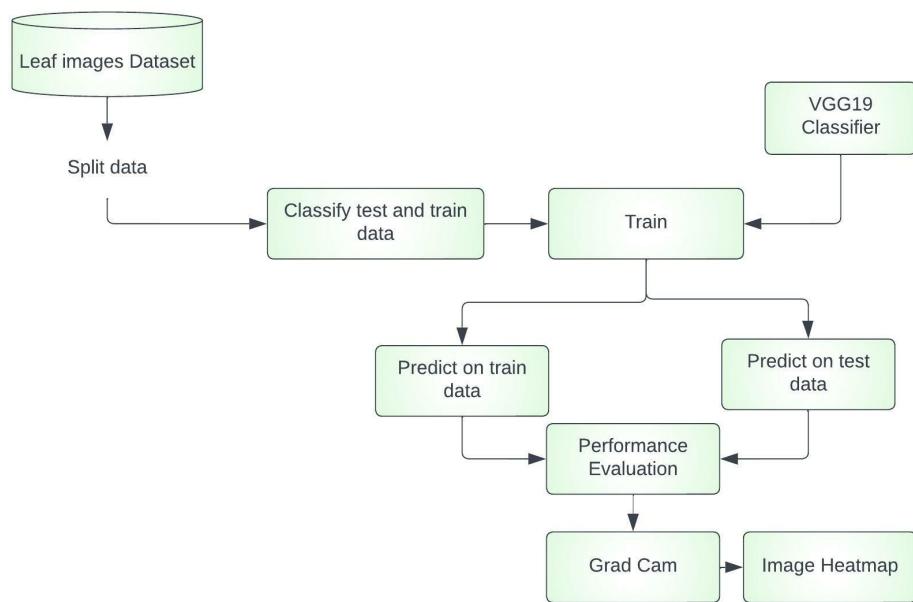
Figure 4.9: Flow Diagram of Proposed Medicinal Plant Recognition using VGG19 with GradCam

Chapter 5

Implementation & Result Analysis

5.1 Performance Evaluation

Performance evaluation is a useful process for gauging a project's efforts and outcomes. It is the methodical evaluation of a learning goal. The ideas of True Positive, True Negative, False Positive, and False Negative are the foundation of Accuracy, Precision, Recall, and F1-Score [34].

5.1.1 True Positive (TP)

These are the accurately predicted positive values, indicating that both the actual class value and the predicted class value are both true.

5.1.2 True Negatives (TN)

These are the accurately predicted negative values, indicating that both the actual and predicted classes have false values.

5.1.3 False Positives (FP)

when the predicted class is true but the actual class is false.

5.1.4 False Negatives (FN)

False Negative (FN) results occur when the classifier predicted a negative result when the actual result was positive.

5.1.5 Accuracy

Accuracy is the most intuitive performance measure. Additionally, it is just the proportion of accurately predicted observations to all observations. It is computed using the following formula:

Accuracy = TP+TN/(TP+TN+FP+FN)

5.1.6 Precision

The precision is defined as the proportion of accurately predicted positive observations to all expected positive observations. When false positives have a significant cost, precision is helpful. With the following formula, precision can be determined:

Precision = TP/TP+FP

5.1.7 Recall

Recall is the percentage of pertinent instances that were successfully recovered. It is also known as sensitivity. Recall is beneficial when the expense of false negatives is significant [17]. The formula below can be used to determine recall:

Recall = TP/TP+FN

5.1.8 F1-score

A test's accuracy is evaluated by the F1-score. It is calculated using the test's precision and recall, where the precision is the ratio of true positive results to all positive results, including those that were incorrectly identified, and the recall is the ratio of true positive results to all samples that should have been classified as positive. An F1-score can have a number as high as 1.0, which denotes perfect precision and recall, or as low as 0, which occurs when either precision or recall are zero. For evaluating search, document classification, and query classification performance, the F1-score is frequently employed in the field of information retrieval [46]. The following formula can be used to determine the F1-score:

F1 Score = 2 x(Recall x Precision) / (Recall + Precision)

5.1.9 Confusion matrix

Confusion matrices are a widely used measurement when attempting to solve classification issues. Both binary classification and multiclass classification issues can be solved with it. Confusion matrices show counts between expected and observed values. Sometimes, to show classifier performance based on the four variables, a confusion matrix is employed (TP, FP, TN, FN). A table called a confusion matrix is used to describe how well a classification system performs.

A matrix that shows how a classification model performed on a set of test data is what is referred to as the confusion matrix, also known as the error matrix. The potential of a classifier can be accurately assessed using a confusion matrix. Each diagonal element represents a successfully classified result. The off diagonals of the confusion matrix show the misclassified results [32].

5.2 Implementation Details

5.2.1 CNN

I created the CNN model. The input shape is the same as the resized image that is (180,180,3). It has 5 layers. 1 input layer, 3 hidden layers and 1 output layer. In the first hidden layer I used 16 neurons and the activation function was relu.In the second hidden layer I used 32 neurons and the activation function was also relu.In the third hidden layer I used 64 neurons and the activation function was the same. The output layer gives the class number of the image. Here Softmax is used as the activation function. A total of 3,992,510 parameters are trained. In the Input layer, 448 parameters are trained. A total of (4670+18496+3965056) 3988222 parameters are trained in the hidden layers. And 3870 parameters are trained in the output layer.Here is the figure 5.1 which shows the CNN model summary.

```
Layer (type)                 Output Shape              Param #
=================================================================
rescaling_1 (Rescaling)      (None, 180, 180, 3)       0

conv2d (Conv2D)              (None, 180, 180, 16)      448

max_pooling2d (MaxPooling2D  (None, 90, 90, 16)        0
)

conv2d_1 (Conv2D)            (None, 90, 90, 32)        4640

max_pooling2d_1 (MaxPooling  (None, 45, 45, 32)        0
2D)

conv2d_2 (Conv2D)            (None, 45, 45, 64)        18496

max_pooling2d_2 (MaxPooling  (None, 22, 22, 64)        0
2D)

flatten (Flatten)           (None, 30976)             0

dense (Dense)                (None, 128)               3965056

dense_1 (Dense)              (None, 30)                3870

=================================================================
Total params: 3,992,510
Trainable params: 3,992,510
Non-trainable params: 0
```

Figure 5.1: Summary of the CNN model

Now, For training the model, I executed 10 epochs. From every single epoch I calculated the loss accuracy, validation accuracy, and validation loss. Here in Figure 5.2 shows the training loss and training accuracy for the training set and validation accuracy, and validation loss for the validation set.

```
Epoch 1/10
46/46 [==============================] - 15s 77ms/step - loss: 3.4803 - accuracy: 0.0688 - val_loss: 3.1817 - val_accuracy: 0.1284
Epoch 2/10
46/46 [==============================] - 1s 24ms/step - loss: 2.6337 - accuracy: 0.2663 - val_loss: 2.1574 - val_accuracy: 0.4317
Epoch 3/10
46/46 [==============================] - 1s 24ms/step - loss: 1.5746 - accuracy: 0.5620 - val_loss: 1.4417 - val_accuracy: 0.5765
Epoch 4/10
46/46 [==============================] - 1s 24ms/step - loss: 0.9946 - accuracy: 0.7037 - val_loss: 1.0809 - val_accuracy: 0.6721
Epoch 5/10
46/46 [==============================] - 1s 23ms/step - loss: 0.6185 - accuracy: 0.8140 - val_loss: 0.9602 - val_accuracy: 0.6858
Epoch 6/10
46/46 [==============================] - 1s 23ms/step - loss: 0.4281 - accuracy: 0.8590 - val_loss: 0.8913 - val_accuracy: 0.7432
Epoch 7/10
46/46 [==============================] - 1s 23ms/step - loss: 0.2497 - accuracy: 0.9217 - val_loss: 0.9101 - val_accuracy: 0.7486
Epoch 8/10
46/46 [==============================] - 1s 24ms/step - loss: 0.1645 - accuracy: 0.9482 - val_loss: 0.9415 - val_accuracy: 0.7732
Epoch 9/10
46/46 [==============================] - 1s 24ms/step - loss: 0.1068 - accuracy: 0.9700 - val_loss: 0.9223 - val_accuracy: 0.7842
Epoch 10/10
46/46 [==============================] - 1s 24ms/step - loss: 0.0547 - accuracy: 0.9850 - val_loss: 0.9701 - val_accuracy: 0.7842
```

Figure 5.2: 10 Epochs iteration of CNN model

5.2.2 SVM

Normal Feature

At first I resized the images to (100,100,3) and converted them into a list of numpy arrays. Then I split the dataset into a training set and Validation set with a ratio of 90% for the training data and 10% for test data. Then I used SVM classifier with polynomial kernel and C=1 to train the train dataset. Figure 5.3 shows the sample image for normal feature extraction.
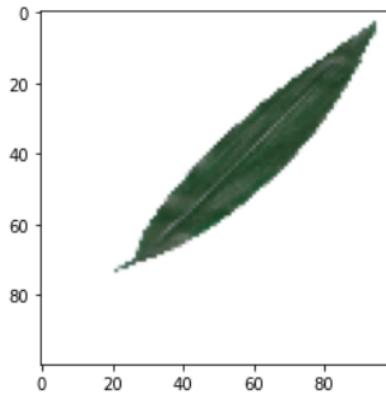
Figure 5.3: Sample images for Normal Feature

Grayscale Feature

At first I resized the images to (100,100,1).Here 1 is for the grayScale version of the image. converted them into an list of numpy array. Then I split the dataset into a training set and Validation set with a ratio of 90% for the training data and 10% for test data. Then I used SVM classifier with a polynomial kernel and C=1 to train the train dataset.Figure 5.4 shows the image with its grayscale features.
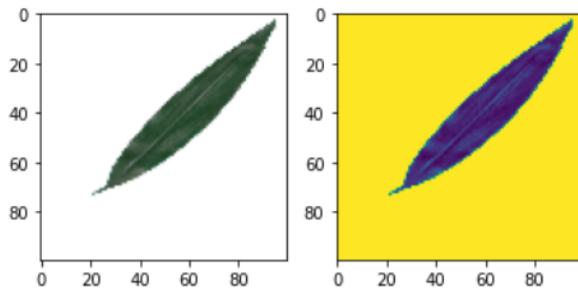
Figure 5.4: Normal Image and it's Grayscale Feature

HOG Feature

At first I resized the images to (100,100,1).Here 1 is for the grayScale version of the image. Then I extracted the HOG features of the images and stored them in a list. Then I split the dataset into a training set and Validation set with a ratio of 90% for the training data and 10% for test data. Then I used SVM classifier with

33

a polynomial kernel and C=1 to train the train dataset.Figure 5.5 shows the image with its HOG feature.
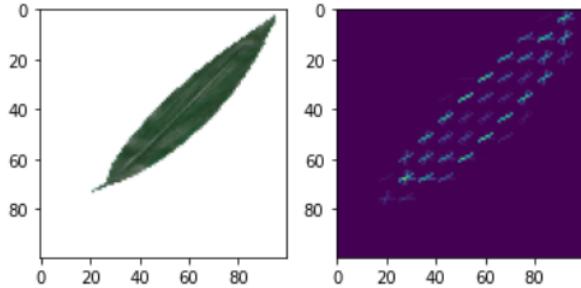
Figure 5.5: Normal Image and it's HOG Feature

Combined Feature

At first I resized the images to (224,224,3) and converted them into a numpy array. Then I resized it again to (224,224,1). Here 1 is for the grayScale version of the image. Then I converted them into a list of numpy arrays. Then I extracted the HOG features of the images. Then I combined the three features and stored them in a list. Then I split the dataset into a training set and Validation set with a ratio of 90% for the training data and 10% for test data. Then I used SVM classifier with a polynomial kernel and C=1 to train the train dataset.

5.2.3   VGG19

Then I created the CNN model with VGG19 as the base model . The input shape is the same as the resized image that is (224,224,3). It has a total of 22 layers with 19 layers from the VGG19 pre-trained and another two hidden layers for processing the VGG19 output and 1 output layer. Here the output layer of the VGG19 has 512 neurons.I then ran the output of the VGG19 through a dense layer with 128 neurons and ReLU as the activation function. The 2nd hidden layer has the same configuration. And the output layer has 30 neurons with each representing 1 class. I used SoftMax as the activation function of the output layer. Here a total number of 20110430 parameters are there with 20024384 non trained and 86046 parameters trained.Here is the figure 5.6 which shows the VGG19 model summary.

```
input_1 (InputLayer)           [(None, 224, 224, 3)]      0

block1_conv1 (Conv2D)          (None, 224, 224, 64)       1792

block1_conv2 (Conv2D)          (None, 224, 224, 64)       36928

block1_pool (MaxPooling2D)     (None, 112, 112, 64)       0

block2_conv1 (Conv2D)          (None, 112, 112, 128)      73856

block2_conv2 (Conv2D)          (None, 112, 112, 128)      147584

block2_pool (MaxPooling2D)     (None, 56, 56, 128)        0

block3_conv1 (Conv2D)          (None, 56, 56, 256)        295168

block3_conv2 (Conv2D)          (None, 56, 56, 256)        590080

block3_conv3 (Conv2D)          (None, 56, 56, 256)        590080

block3_conv4 (Conv2D)          (None, 56, 56, 256)        590080

block3_pool (MaxPooling2D)     (None, 28, 28, 256)        0

block4_conv1 (Conv2D)          (None, 28, 28, 512)        1180160

block4_conv2 (Conv2D)          (None, 28, 28, 512)        2359808

block4_conv3 (Conv2D)          (None, 28, 28, 512)        2359808

block4_conv4 (Conv2D)          (None, 28, 28, 512)        2359808

block4_pool (MaxPooling2D)     (None, 14, 14, 512)        0

block5_conv1 (Conv2D)          (None, 14, 14, 512)        2359808

block5_conv2 (Conv2D)          (None, 14, 14, 512)        2359808

block5_conv3 (Conv2D)          (None, 14, 14, 512)        2359808

block5_conv4 (Conv2D)          (None, 14, 14, 512)        2359808

block5_pool (MaxPooling2D)     (None, 7, 7, 512)          0

global_average_pooling2d (G    (None, 512)                0
lobalAveragePooling2D)

dense (Dense)                  (None, 128)                65664

dense_1 (Dense)                (None, 128)                16512

dense_2 (Dense)                (None, 30)                 3870

=================================================================
Total params: 20,110,430
Trainable params: 86,046
Non-trainable params: 20,024,384
_____
```

Figure 5.6: Summary of the VGG19 model

Now, For training the model, I executed 10 epochs. From every single epoch I calculated the loss accuracy, validation accuracy, and validation loss. Here in Figure 5.7 shows the training loss and training accuracy for the training set and validation accuracy, and validation loss for the validation set.

```
Epoch 1/10
42/42 [==============================] - 57s 1s/step - loss: 2.3208 - accuracy: 0.4424 - val_loss: 1.1075 - val_accuracy: 0.6939
Epoch 2/10
42/42 [==============================] - 38s 918ms/step - loss: 0.6458 - accuracy: 0.8379 - val_loss: 0.5063 - val_accuracy: 0.8515
Epoch 3/10
42/42 [==============================] - 39s 927ms/step - loss: 0.2931 - accuracy: 0.9311 - val_loss: 0.3564 - val_accuracy: 0.8939
Epoch 4/10
42/42 [==============================] - 39s 921ms/step - loss: 0.1592 - accuracy: 0.9689 - val_loss: 0.2398 - val_accuracy: 0.9303
Epoch 5/10
42/42 [==============================] - 39s 926ms/step - loss: 0.1212 - accuracy: 0.9720 - val_loss: 0.2047 - val_accuracy: 0.9364
Epoch 6/10
42/42 [==============================] - 41s 984ms/step - loss: 0.0596 - accuracy: 0.9955 - val_loss: 0.1725 - val_accuracy: 0.9424
Epoch 7/10
42/42 [==============================] - 40s 948ms/step - loss: 0.0382 - accuracy: 0.9985 - val_loss: 0.1423 - val_accuracy: 0.9545
Epoch 8/10
42/42 [==============================] - 40s 960ms/step - loss: 0.0286 - accuracy: 1.0000 - val_loss: 0.1323 - val_accuracy: 0.9606
Epoch 9/10
42/42 [==============================] - 40s 956ms/step - loss: 0.0197 - accuracy: 1.0000 - val_loss: 0.1119 - val_accuracy: 0.9697
Epoch 10/10
42/42 [==============================] - 39s 935ms/step - loss: 0.0154 - accuracy: 1.0000 - val_loss: 0.1131 - val_accuracy: 0.9636
```

Figure 5.7: Epochs iteration of VGG19 model

5.2.4   VGG19 with GradCam

Figure 5.8 shows the VGG19 model architecture.After running the CNN model
with VGG19 pretrained model, I implemented a model that mapped the images
that are used as an input to the activation of the last convolutional layer and also
the predictions. Next the gradient of the top 15 predictions were calculated with
respect to the last convolutional layer of the VGG19 model which is block5_conv4.
The calculated gradient is the gradient of the neurons of the output layer with respect
to the last convolutional layer. The output is a vector where each entry is the mean
intensity of the gradient over a specific feature map channel.Then we multiply the
channels with regards to the importance of the channel. Then I summed all the
channels to get the activation heatmaps. Then I superimposed the heatmap onto
the real images.
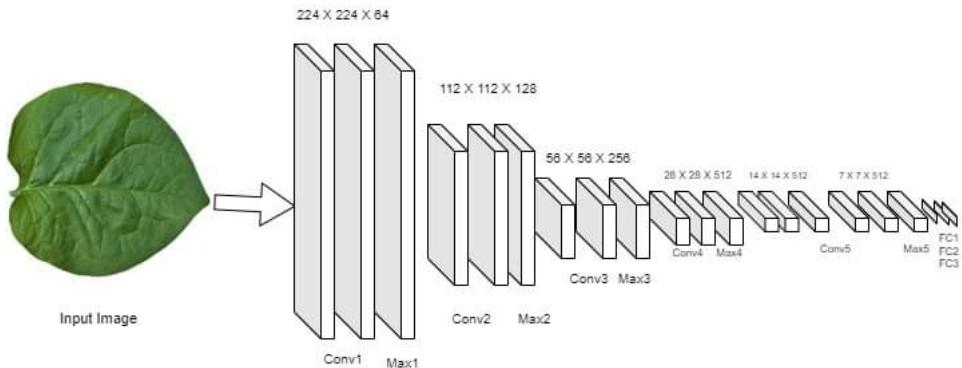
Figure 5.8: VGG19 Model Architecture

5.3   Results Analysis

Traditional medicine and global ethnomedicine have long made use of medicinal
plants. However, accurate plant species identification provides significant advantages
for a wide range of stakeholders, including consumers, forestry services, botanists,

taxonomists, doctors, pharmaceutical laboratories, organizations working to save endangered species, the government, and the general public.

5.3.1 Training Accuracy and Validation Accuracy Curve

The dataset formed from the initial dataset and used to build the models is normally the biggest subset, and this is called the training set. Next, in order to identify the best model, the models are assessed using the validation set.Training accuracy is defined as using the same images for both training and testing.When comparing test and training accuracy, test accuracy denotes that the trained model correctly detects independent images that were not used in training. The algorithm's performance is measured using validation accuracy, which may be easily understood.The training accuracy and validation accuracy for the training set are displayed in Figure 5.9 and 5.10 here.
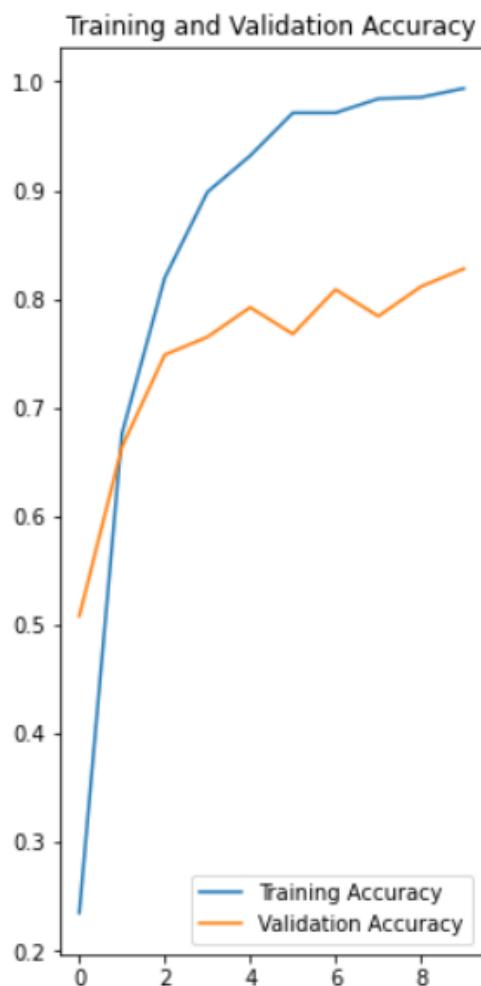
Figure 5.9: CNN Model Training Accuracy and Validation Accuracy Curve
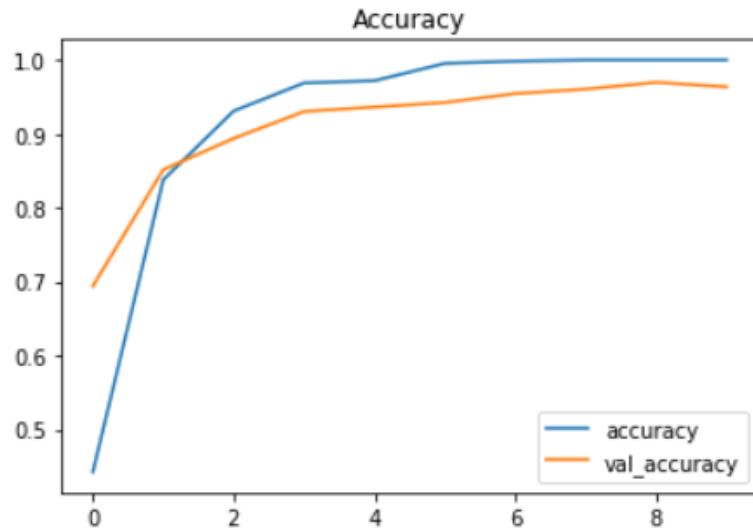
Figure 5.10: VGG19 Model Training Accuracy and Validation Accuracy Curve

5.3.2 Training Loss and Validation Loss Curve

The sum of errors for every example in the training set is used to determine the training loss. Similar to the training loss, the validation loss is computed by adding the mistakes for each sample in the validation set.The validation loss shows how well the model fits new data, while the training loss shows how well it matches training data. The training loss and validation loss for the training set are displayed in Figure 5.11 and 5.12 here.
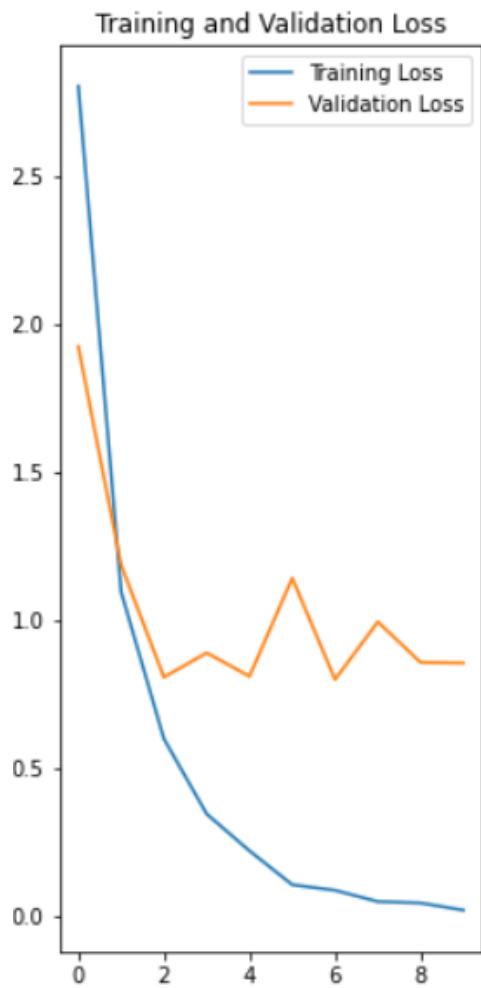
Figure 5.11: CNN Model Training loss and Validation loss Curve
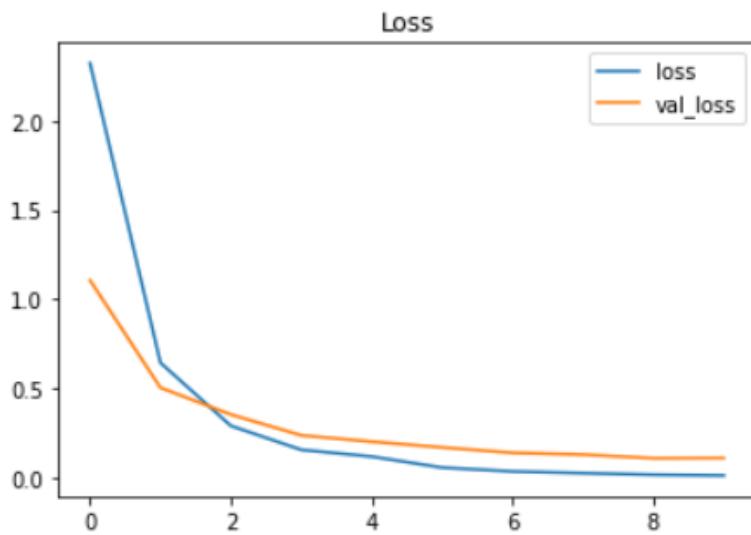
Figure 5.12: VGG19 Model Training loss and Validation loss Curve

5.3.3 Accuracy & Miss-classification in Confusion Matrix

1. CNN

The performance of my proposed method has been considered based on the following evaluation metrics.I used a confusion matrix whose dimensions is 30x30 for multi-class classifications problem with 30 number of target classes. Here is the figure 5.13 which shows the CNN model Confusion Matrix.
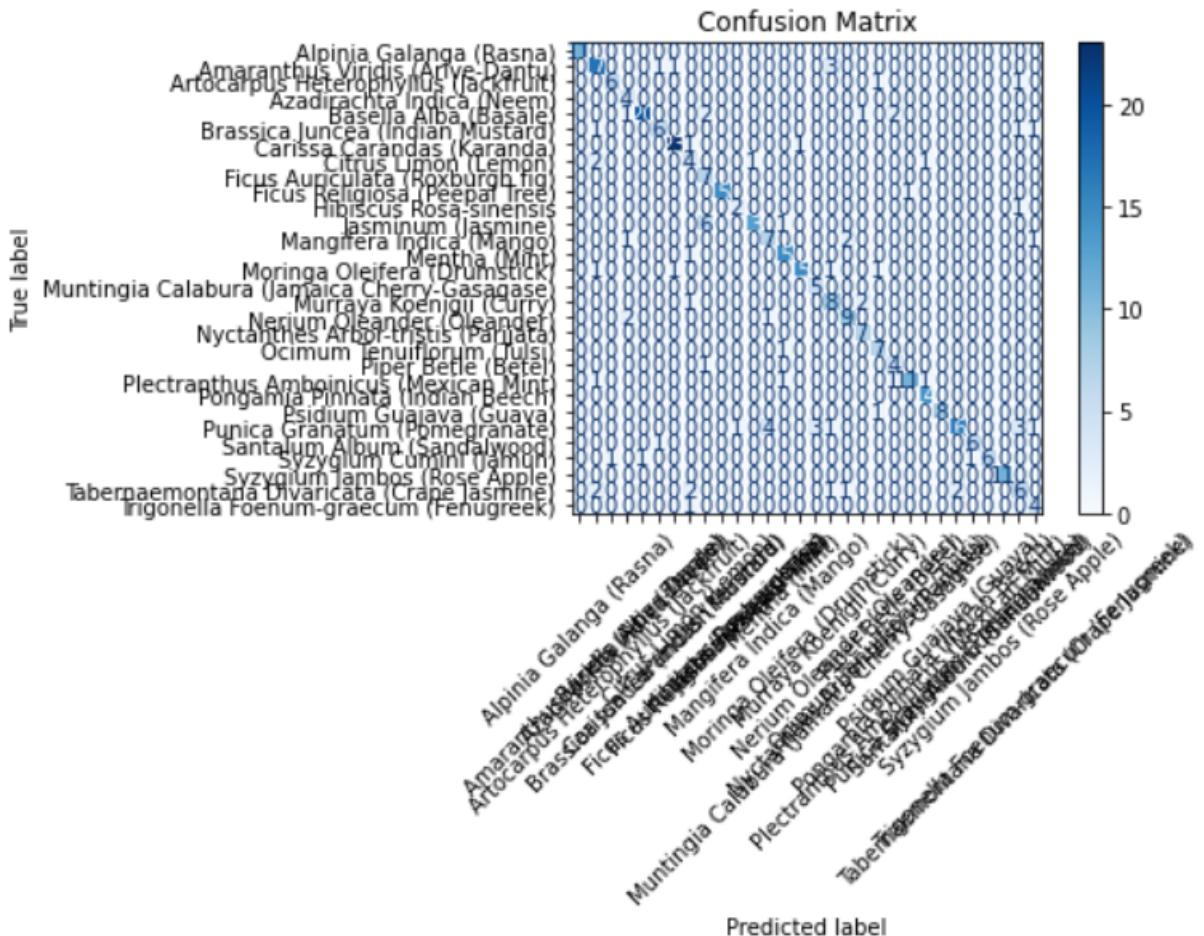
Figure 5.13: CNN Model Confusion Matrix

2. SVM

An N x N matrix called a confusion matrix is used to assess the effectiveness of a classification model, where N is the total number of target classes. In the matrix, the actual goal values are contrasted with those that the machine learning model anticipated. I used a confusion matrix whose dimensions is 30x30 for a multi-class classifications problem with 30 numbers of target classes. In this confusion matrix, the predicted label on x axis and the actual label y axis.

Normal Feature

The performance of the SVM classifier with normal features is visualized and summarized.It indicates the performance for all 30 classes. That means how many classes

are classified correctly and how many are mis-classified.Here is the figure 5.14 which shows the Confusion Matrix for normal feature extraction.
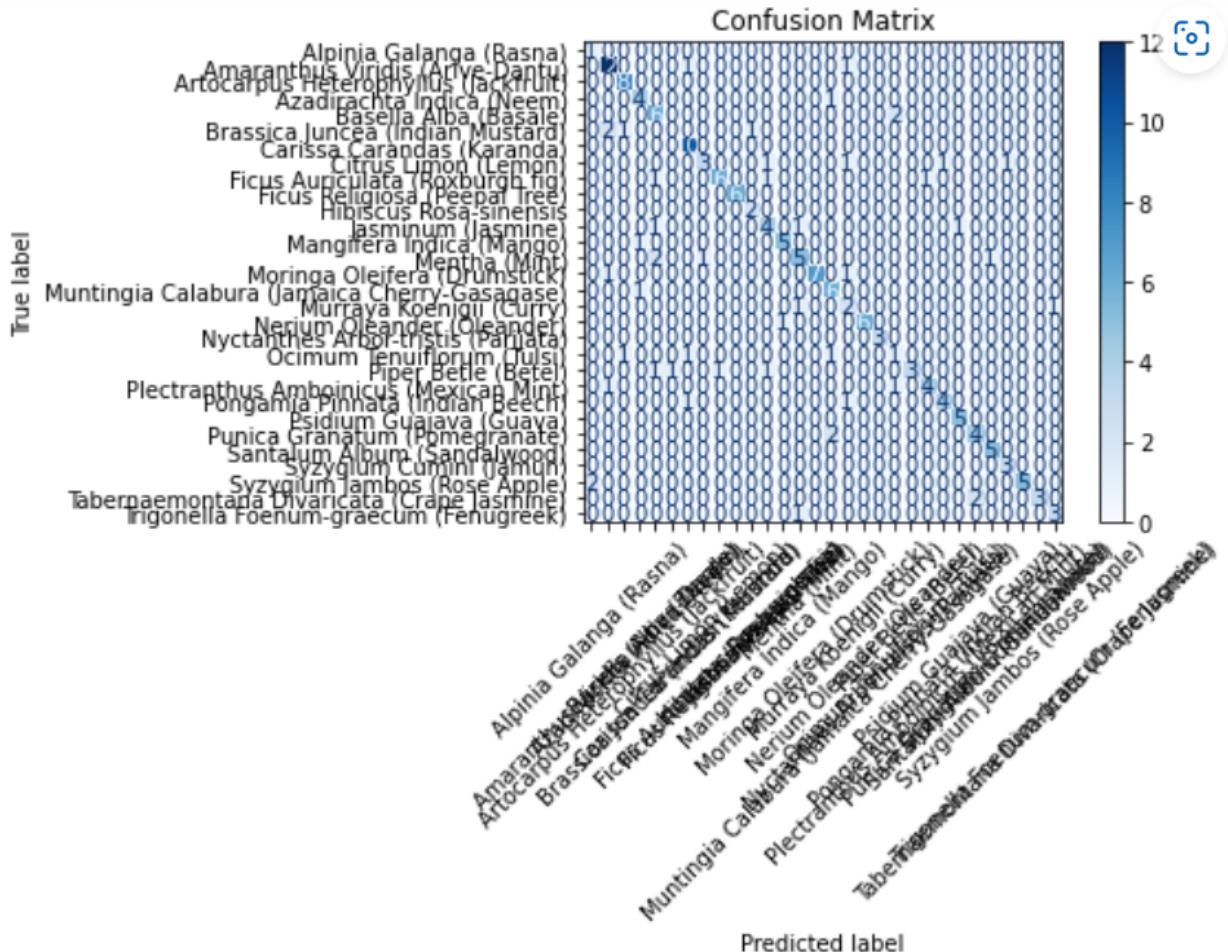
Figure 5.14: Confusion Matrix for Normal Feature

Grayscale Feature

It presents a synthesis of the SVM classifier's performance using grayscale features. It displays the results for all 30 classes. In other words, how many classes are correctly categorized and how many are not.Here is the figure 5.15 which shows the Confusion Matrix for normal feature extraction.
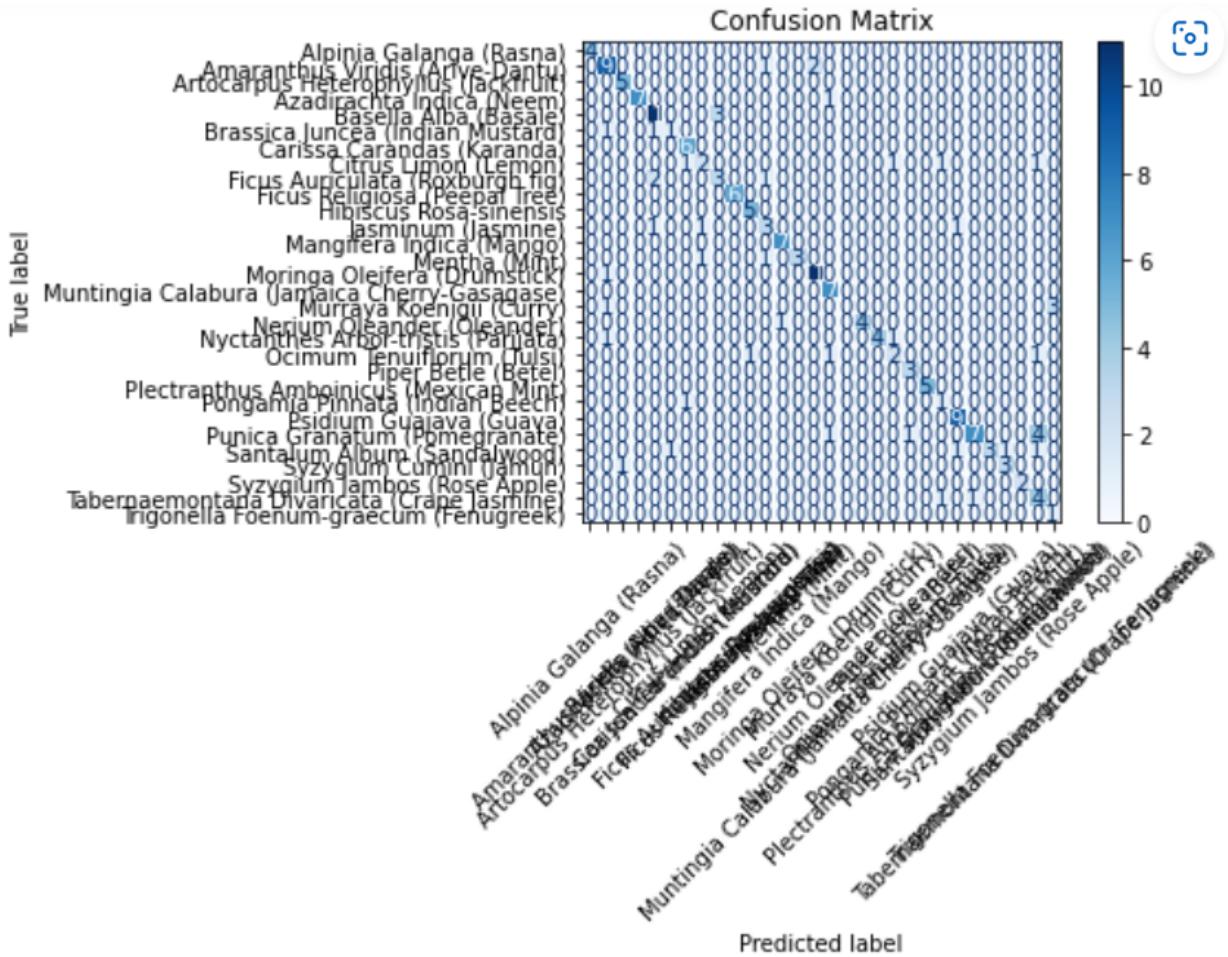
Figure 5.15: Confusion Matrix for Grayscale Feature

HOG Feature

It provides a summary of the performance of the SVM classifier utilizing HOG features. The outcomes for each of the 30 classes are shown. In other words, the number of classes that are correctly classified vs those that are not. Here is the figure 5.16 which shows the Confusion Matrix for HOG feature extraction .
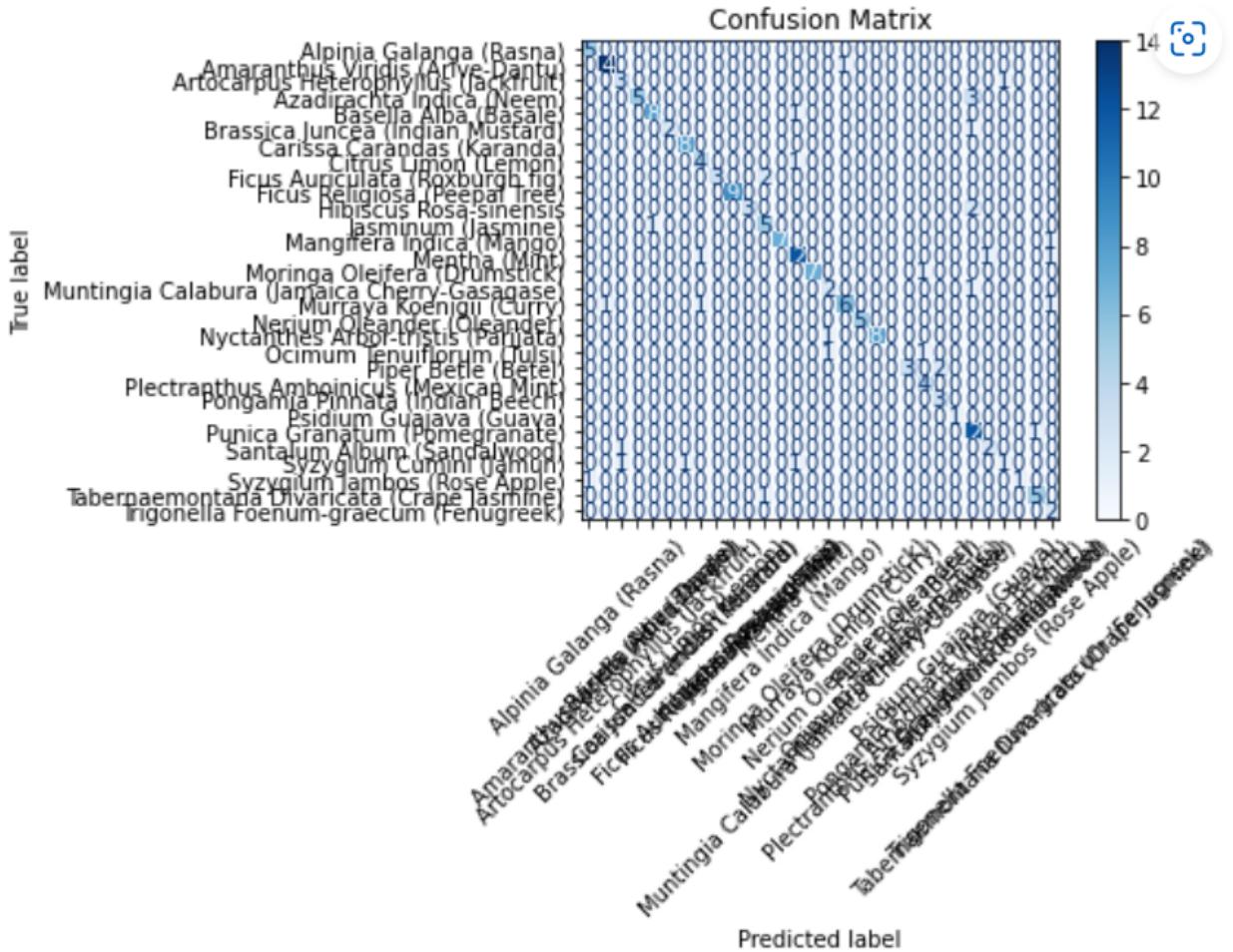
Figure 5.16: Confusion Matrix for HOG Feature

Combined Feature

The effectiveness of the SVM classifier with Combined features is displayed and evaluated. It displays the results for all 30 classes. In other words, how many classes are correctly classified and how many are not.Here is the figure 5.17 which shows the Confusion Matrix for combined feature extraction .
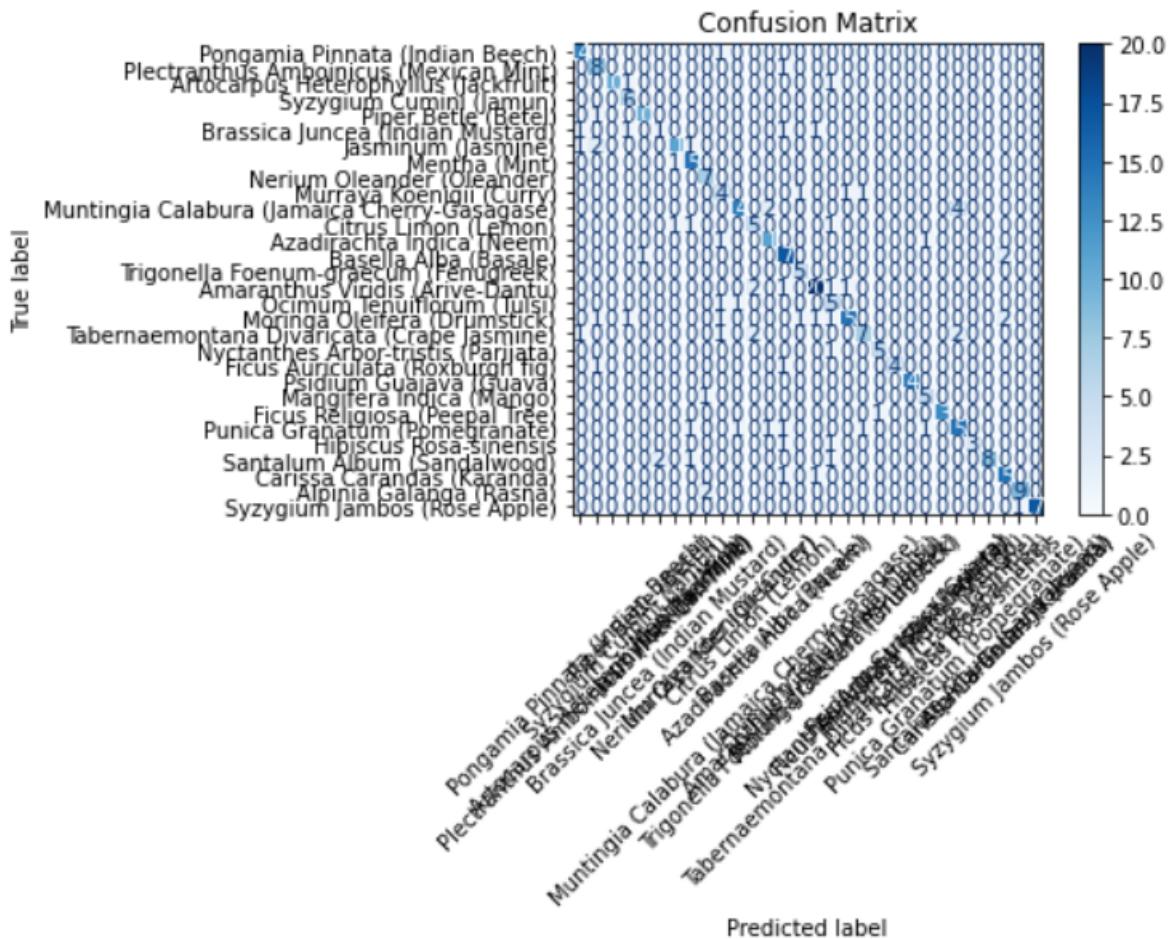
Figure 5.17: Confusion Matrix for Combined Feature

3. VGG19

I used a confusion matrix whose dimensions is 30x30 for multi-class classifications problem with 30 number of target classes.The diagonal shows the actual results for all the classes.Here is the figure 5.18 which shows the VGG19 model Confusion Matrix.
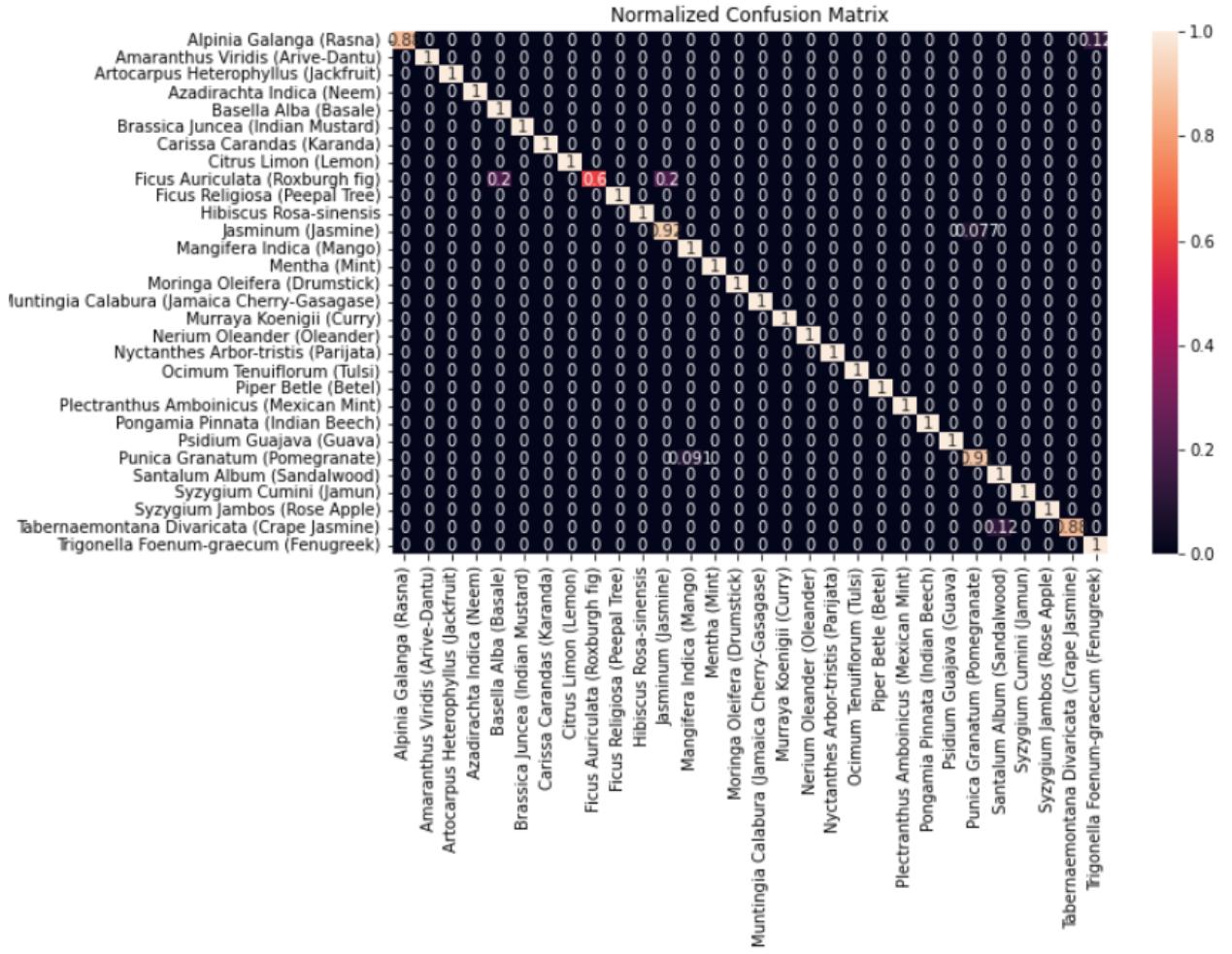
Figure 5.18: VGG19 Model Confusion Matrix

5.3.4 Confusion Matrix Performance (TP,FN,FP,TN)

True Positive (TP) results occur when the model correctly predicts the positive class. A True Negative (TN) outcome occurs when the model correctly predicts the negative class. False Positive (FP) is a result where the model incorrectly predicts the positive class. False Positive (FP) is a result where the model incorrectly predicts the positive class.Here Table 5.1 represents the CNN model Confusion Matrix Performance (TP,FN,FP,TN).

Table 5.1: CNN Model Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	0.900000	0.100000	0.005618	0.994382
Amaranthus Viridis	0.888889	0.111111	0.022989	0.977011
Artocarpus Heterophyllus	0.777778	0.222222	0.000000	1.000000
Azadirachta Indica	0.777778	0.222222	0.002801	0.997199
Basella Alba	0.863636	0.136364	0.005814	0.994186
Brassica Juncea	0.857143	0.142857	0.005571	0.994429
Carissa Carandas	1.000000	0.000000	0.008721	0.991279
Citrus Limon	0.875000	0.125000	0.008380	0.991620
Ficus Auriculata	0.625000	0.375000	0.017143	0.982857
Ficus Religiosa	0.875000	0.125000	0.002857	0.997143
Hibiscus Rosa-sinensis	0.333333	0.666667	0.002778	0.997222
Jasminum	0.714286	0.285714	0.011364	0.988636
Mangifera Indica	1.000000	0.000000	0.011173	0.988827
Mentha	0.782609	0.217391	0.000000	1.000000
Moringa Oleifera	1.000000	0.000000	0.011299	0.988701
Muntingia Calabura	0.857143	0.142857	0.005571	0.994429
Murraya Koenigii	0.666667	0.333333	0.005747	0.994253
Nerium Oleander	0.909091	0.090909	0.005634	0.994366
Nyctanthes Arbor-tristis	1.000000	0.000000	0.005587	0.994413
Ocimum Tenuiflorum	1.000000	0.000000	0.011111	0.988889
Piper Betle	0.583333	0.416667	0.000000	1.000000
Plectranthus Amboinicus	1.000000	0.000000	0.002817	0.997183
Pongamia Pinnata	0.941176	0.058824	0.000000	1.000000
Psidium Guajava	0.888889	0.111111	0.000000	1.000000
Punica Granatum	0.695652	0.304348	0.005831	0.994169
Santalum Album	0.777778	0.222222	0.000000	1.000000
Syzygium Cumini	0.833333	0.166667	0.002778	0.997222
Syzygium Jambos	0.833333	0.166667	0.002825	0.997175
Tabernaemontana Divaricata	0.800000	0.200000	0.011236	0.988764
Trigonella Foenum-graecum	0.857143	0.142857	0.002786	0.997214

Confusion Matrix Performance for SVM classifier with Normal features is shown below:

Table 5.2: SVM with Normal Feature Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	1.000000	0.000000	0.000000	1.000000
Amaranthus Viridis	0.700000	0.300000	0.028736	0.971264
Artocarpus Heterophyllus	0.800000	0.200000	0.011173	0.988827
Azadirachta Indica	0.571429	0.428571	0.011299	0.988701
Basella Alba	1.000000	0.000000	0.011236	0.988764
Brassica Juncea	0.500000	0.500000	0.005556	0.994444
Carissa Carandas	1.000000	0.000000	0.005682	0.994318
Citrus Limon	0.666667	0.333333	0.011236	0.988764
Ficus Auriculata	0.600000	0.400000	0.000000	1.000000
Ficus Religiosa	1.000000	0.000000	0.005650	0.994350
Hibiscus Rosa-sinensis	1.000000	0.000000	0.000000	1.000000
Jasminum	0.875000	0.125000	0.000000	1.000000
Mangifera Indica	0.800000	0.200000	0.011173	0.988827
Mentha	1.000000	0.000000	0.011696	0.988304
Moringa Oleifera	0.600000	0.400000	0.000000	1.000000
Muntingia Calabura	0.666667	0.333333	0.005525	0.994475
Murraya Koenigii	0.666667	0.333333	0.011050	0.988950
Nerium Oleander	0.875000	0.125000	0.000000	1.000000
Nyctanthes Arbor-tristi	0.800000	0.200000	0.000000	1.000000
Ocimum Tenuiflorum	1.000000	0.000000	0.005556	0.994444
Piper Betle	1.000000	0.000000	0.005525	0.994475
Plectranthus Amboinicus	0.857143	0.142857	0.005650	0.994350
Pongamia Pinnata	0.888889	0.111111	0.000000	1.000000
Psidium Guajava	1.000000	0.000000	0.000000	1.000000
Punica Granatum	0.818182	0.181818	0.005780	0.994220
Santalum Album	0.714286	0.285714	0.005650	0.994350
Syzygium Cumini	0.833333	0.166667	0.000000	1.000000
Syzygium Jambos	1.000000	0.000000	0.000000	1.000000
Tabernaemontana Divaricata	0.800000	0.200000	0.016760	0.983240
Trigonella Foenum-graecum	0.500000	0.500000	0.000000	1.000000

Here in the below table, the TPR,FNR,FPR, TNR for SVM classifier with grayscale features is shown:

Table 5.3: SVM with Grayscale Feature Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	0.875000	0.125000	0.000000	1.000000
Amaranthus Viridis	0.769231	0.230769	0.017544	0.982456
Artocarpus Heterophyllus	1.000000	0.000000	0.000000	1.000000
Azadirachta Indica	0.875000	0.125000	0.005682	0.994318
Basella Alba	0.571429	0.428571	0.022599	0.977401
Brassica Juncea	0.333333	0.666667	0.000000	1.000000
Carissa Carandas	1.000000	0.000000	0.011173	0.988827
Citrus Limon	0.200000	0.800000	0.016760	0.983240
Ficus Auriculata	0.666667	0.333333	0.011236	0.988764
Ficus Religiosa	1.000000	0.000000	0.005587	0.994413
Hibiscus Rosa-sinensis	0.750000	0.250000	0.005556	0.994444
Jasminum	0.428571	0.571429	0.011299	0.988701
Mangifera Indica	0.333333	0.666667	0.011050	0.988950
Mentha	0.888889	0.111111	0.000000	1.000000
Moringa Oleifera	0.600000	0.400000	0.011494	0.988506
Muntingia Calabura	0.857143	0.142857	0.005650	0.994350
Murraya Koenigii	0.500000	0.500000	0.027778	0.972222
Nerium Oleander	1.000000	0.000000	0.016760	0.983240
Nyctanthes Arbor-tristis	0.833333	0.166667	0.000000	1.000000
Ocimum Tenuiflorum	0.500000	0.500000	0.017045	0.982955
Piper Betle	0.833333	0.166667	0.000000	1.000000
Plectranthus Amboinicus	1.000000	0.000000	0.000000	1.000000
Pongamia Pinnata	0.500000	0.500000	0.000000	1.000000
Psidium Guajava	1.000000	0.000000	0.011111	0.988889
Punica Granatum	0.545455	0.454545	0.017341	0.982659
Santalum Album	1.000000	0.000000	0.005587	0.994413
Syzygium Cumini	1.000000	0.000000	0.016393	0.983607
Syzygium Jambos	1.000000	0.000000	0.000000	1.000000
Tabernaemontana Divaricata	0.428571	0.571429	0.016949	0.983051
Trigonella Foenum-graecum	0.833333	0.166667	0.005618	0.994382

The TPR, FNR, FPR, and TNR for the SVM classifier with HOG features are displayed in the table below:.

Table 5.4: SVM with HOG Feature Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	0.800000	0.200000	0.000000	1.000000
Amaranthus Viridis	0.818182	0.181818	0.023121	0.976879
Artocarpus Heterophyllus	1.000000	0.000000	0.005525	0.994475
Azadirachta Indica	0.333333	0.666667	0.000000	1.000000
Basella Alba	0.916667	0.083333	0.017442	0.982558
Brassica Juncea	0.600000	0.400000	0.000000	1.000000
Carissa Carandas	0.600000	0.400000	0.011494	0.988506
Citrus Limon	0.600000	0.400000	0.000000	1.000000
Ficus Auriculata	0.875000	0.125000	0.005682	0.994318
Ficus Religiosa	1.000000	0.000000	0.000000	1.000000
Hibiscus Rosa-sinensis	0.666667	0.333333	0.000000	1.000000
Jasminum	0.750000	0.250000	0.033333	0.966667
Mangifera Indica	0.750000	0.250000	0.005682	0.994318
Mentha	0.714286	0.285714	0.011299	0.988701
Moringa Oleifera	0.818182	0.181818	0.005780	0.994220
Muntingia Calabura	0.750000	0.250000	0.005556	0.994444
Murraya Koenigii	1.000000	0.000000	0.011050	0.988950
Nerium Oleander	1.000000	0.000000	0.011299	0.988701
Nyctanthes Arbor-tristi	1.000000	0.000000	0.000000	1.000000
Ocimum Tenuiflorum	0.333333	0.666667	0.000000	1.000000
Piper Betle	0.857143	0.142857	0.000000	1.000000
Plectranthus Amboinicus	1.000000	0.000000	0.000000	1.000000
Pongamia Pinnata	1.000000	0.000000	0.000000	1.000000
Psidium Guajava	1.000000	0.000000	0.005587	0.994413
Punica Granatum	0.888889	0.111111	0.022857	0.977143
Santalum Album	0.750000	0.250000	0.022222	0.977778
Syzygium Cumini	0.500000	0.500000	0.000000	1.000000
Syzygium Jambos	0.833333	0.166667	0.005618	0.994382
Tabernaemontana Divaricata	1.000000	0.000000	0.005495	0.994505
Trigonella Foenum-graecum	1.000000	0.000000	0.005556	0.994444

Here the performance matrix for SVM classifier with Combined features is displayed below:

Table 5.5: SVM with Combined Feature Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	0.933333	0.066667	0.008523	0.991477
Amaranthus Viridis	0.888889	0.111111	0.013966	0.986034
Artocarpus Heterophyllus	0.833333	0.166667	0.000000	1.000000
Azadirachta Indica	1.000000	0.000000	0.008310	0.991690
Basella Alba	0.909091	0.090909	0.002809	0.997191
Brassica Juncea	0.200000	0.800000	0.005525	0.994475
Carissa Carandas	0.733333	0.266667	0.005682	0.994318
Citrus Limon	0.937500	0.062500	0.011396	0.988604
Ficus Auriculata	1.000000	0.000000	0.008333	0.991667
Ficus Religiosa	0.571429	0.428571	0.013889	0.986111
Hibiscus Rosa-sinensis	0.700000	0.300000	0.005764	0.994236
Jasminum	0.625000	0.375000	0.011142	0.988858
Mangifera Indica	0.846154	0.153846	0.008475	0.991525
Mentha	0.850000	0.150000	0.020173	0.979827
Moringa Oleifera	1.000000	0.000000	0.005525	0.994475
Muntingia Calabura	0.800000	0.200000	0.008772	0.991228
Murraya Koenigii	0.714286	0.285714	0.019444	0.980556
Nerium Oleander	0.652174	0.347826	0.005814	0.994186
Nyctanthes Arbor-tristi	0.538462	0.461538	0.005650	0.994350
Ocimum Tenuiflorum	0.833333	0.166667	0.002770	0.997230
Piper Betle	0.666667	0.333333	0.000000	1.000000
Plectranthus Amboinicus	1.000000	0.000000	0.000000	1.000000
Pongamia Pinnata	0.833333	0.166667	0.005540	0.994460
Psidium Guajava	0.928571	0.071429	0.000000	1.000000
Punica Granatum	0.714286	0.285714	0.017341	0.982659
Santalum Album	0.600000	0.400000	0.000000	1.000000
Syzygium Cumini	0.666667	0.333333	0.000000	1.000000
Syzygium Jambos	0.882353	0.117647	0.011429	0.988571
Tabernaemontana Divaricata	0.818182	0.181818	0.002809	0.997191
Trigonella Foenum-graecum	0.944444	0.055556	0.000000	1.000000

Below is a depiction of the Confusion matrix for the VGG19 model:

Table 5.6: VGG19 Model Confusion Matrix Performance(TP,FN,FP,TN)

Medicinal plant	TPR	FNR	FPR	TNR
Alpinia Galanga	0.875000	0.125000	0.000000	1.000000
Amaranthus Viridis	1.000000	0.000000	0.000000	1.000000
Artocarpus Heterophyllus	1.000000	0.000000	0.000000	1.000000
Azadirachta Indica	1.000000	0.000000	0.000000	1.000000
Basella Alba	1.000000	0.000000	0.006897	0.993103
Brassica Juncea	1.000000	0.000000	0.000000	1.000000
Carissa Carandas	1.000000	0.000000	0.000000	1.000000
Citrus Limon	1.000000	0.000000	0.000000	1.000000
Ficus Auriculata	0.600000	0.400000	0.000000	1.000000
Ficus Religiosa	1.000000	0.000000	0.000000	1.000000
Hibiscus Rosa-sinensis	1.000000	0.000000	0.000000	1.000000
Jasminum	0.923077	0.076923	0.006897	0.993103
Mangifera Indica	1.000000	0.000000	0.003135	0.996865
Mentha	1.000000	0.000000	0.000000	1.000000
Moringa Oleifera	1.000000	0.000000	0.000000	1.000000
Muntingia Calabura	1.000000	0.000000	0.000000	1.000000
Murraya Koenigii	1.000000	0.000000	0.000000	1.000000
Nerium Oleander	1.000000	0.000000	0.000000	1.000000
Nyctanthes Arbor-tristis	1.000000	0.000000	0.000000	1.000000
Ocimum Tenuiflorum	1.000000	0.000000	0.000000	1.000000
Piper Betle	1.000000	0.000000	0.000000	1.000000
Plectranthus Amboinicus	1.000000	0.000000	0.000000	1.000000
Pongamia Pinnata	1.000000	0.000000	0.000000	1.000000
Psidium Guajava	1.000000	0.000000	0.000000	1.000000
Punica Granatum	0.909091	0.090909	0.002653	0.997347
Santalum Album	1.000000	0.000000	0.004310	0.995690
Syzygium Cumini	1.000000	0.000000	0.000000	1.000000
Syzygium Jambos	1.000000	0.000000	0.000000	1.000000
Tabernaemontana Divaricata	0.875000	0.125000	0.000000	1.000000
Trigonella Foenum-graecum	1.000000	0.000000	0.004310	0.995690

5.3.5 Classification Report

Performance evaluation produces a classification report with precision, recall, F1-score, and support values as well as a confusion matrix that displays the number of plants that fall into each group. The precision, recall, f1-score, and support values indicate that the model performs reasonably well with the dataset. To get some understanding while also optimizing the dataset in order to enhance the model.Here Table 5.7 represents the CNN model Classification Report.

Table 5.7: CNN Model Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	0.82	0.90	0.86	10
Amaranthus Viridis	0.67	0.89	0.76	18
Artocarpus Heterophyllus	1.00	0.78	0.88	9
Azadirachta Indica	0.88	0.78	0.82	9
Basella Alba	0.90	0.86	0.88	22
Brassica Juncea	0.75	0.86	0.80	7
Carissa Carandas	0.88	1.00	0.94	22
Citrus Limon	0.70	0.88	0.78	8
Ficus Auriculata	0.62	0.62	0.62	16
Ficus Religiosa	0.93	0.88	0.90	16
Hibiscus Rosa-sinensis	0.67	0.33	0.44	6
Jasminum	0.71	0.71	0.71	14
Mangifera Indica	0.67	1.00	0.80	8
Mentha	1.00	0.78	0.88	23
Moringa Oleifera	0.75	1.00	0.86	12
Muntingia Calabura	0.75	0.86	0.80	7
Murraya Koenigii	0.86	0.67	0.75	18
Nerium Oleander	0.83	0.91	0.87	11
Nyctanthes Arbor-tristis	0.80	1.00	0.89	8
Ocimum Tenuiflorum	0.60	1.00	0.75	6
Piper Betle	1.00	0.58	0.74	12
Plectranthus Amboinicus	0.92	1.00	0.96	11
Pongamia Pinnata	1.00	0.94	0.97	17
Psidium Guajava	1.00	0.89	0.94	9
Punica Granatum	0.89	0.70	0.78	23
Santalum Album	1.00	0.78	0.88	9
Syzygium Cumini	0.83	0.83	0.83	6
Syzygium Jambos	0.91	0.83	0.87	12
Tabernaemontana Divaricata	0.67	0.80	0.73	10
Trigonella Foenum-graecum	0.86	0.86	0.86	7
Overall	0.85	0.83	0.83	366

In the table below, the classification report for SVM classifier with Normal features is displayed. Here, the overall precision for this classifier is 85%, Recall is 83%, and f1 score is 83%.

Table 5.8: SVM with Normal Feature Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	1.00	1.00	1.00	7
Amaranthus Viridis	0.58	0.70	0.64	10
Artocarpus Heterophyllus	0.67	0.80	0.73	5
Azadirachta Indica	0.67	0.57	0.62	7
Basella Alba	0.75	1.00	0.86	6
Brassica Juncea	0.67	0.50	0.57	4
Carissa Carandas	0.89	1.00	0.94	8
Citrus Limon	00.67	0.67	0.67	6
Ficus Auriculata	1.00	0.60	0.75	5
Ficus Religiosa	0.88	1.00	0.93	7
Hibiscus Rosa-sinensis	1.00	1.00	1.00	5
Jasminum	1.00	0.88	0.93	8
Mangifera Indica	0.67	0.80	0.73	5
Mentha	0.87	1.00	0.93	13
Moringa Oleifera	1.00	0.60	0.75	5
Muntingia Calabura	0.67	0.67	0.67	3
Murraya Koenigii	0.50	0.67	0.57	3
Nerium Oleander	1.00	0.88	0.93	8
Nyctanthes Arbor-tristis	1.00	0.80	0.89	5
Ocimum Tenuiflorum	0.80	1.00	0.89	4
Piper Betle	0.75	1.00	0.86	3
Plectranthus Amboinicus	0.86	0.86	0.86	7
Pongamia Pinnata	1.00	0.89	0.94	9
Psidium Guajava	1.00	1.00	1.00	5
Punica Granatum	0.90	0.82	0.86	11
Santalum Album	0.83	0.71	0.77	7
Syzygium Cumini	1.00	0.83	0.91	6
Tabernaemontana Divaricata	1.00	1.00	1.00	3
Tabernaemontana Divaricata	0.57	0.80	0.67	5
Trigonella Foenum-graecum	1.00	0.50	0.67	4
Overall	0.77	0.72	0.72	184

The classification report for the SVM classifier with Grayscale features is shown in the table below. This classifier's overall precision is 76%, recall is 74%, and f1 score is 74%.

Table 5.9: SVM with Grayscale Feature Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	1.00	0.88	0.93	8
Amaranthus Viridis	0.77	0.77	0.77	13
Artocarpus Heterophyllus	1.00	1.00	1.00	8
Azadirachta Indica	0.88	0.88	0.88	8
Basella Alba	0.50	0.57	0.53	7
Brassica Juncea	1.00	0.33	0.50	3
Carissa Carandas	0.71	1.00	0.83	5
Citrus Limon	0.25	0.20	0.22	5
Ficus Auriculata	0.67	0.67	0.67	6
Ficus Religiosa	0.83	1.00	0.91	5
Hibiscus Rosa-sinensis	0.75	0.75	0.75	4
Jasminum	0.60	0.43	0.50	7
Mangifera Indica	0.33	0.33	0.33	3
Mentha	1.00	0.89	0.94	9
Moringa Oleifera	0.75	0.60	0.67	10
Muntingia Calabura	0.86	0.86	0.86	7
Murraya Koenigii	0.29	0.50	0.36	4
Nerium Oleander	0.62	1.00	0.77	5
Nyctanthes Arbor-tristis	1.00	0.83	0.91	6
Ocimum Tenuiflorum	0.57	0.50	0.53	8
Piper Betle	1.00	1.00	0.91	6
Plectranthus Amboinicus	1.00	1.00	1.00	4
Pongamia Pinnata	1.00	0.50	0.67	2
Psidium Guajava	0.67	1.00	0.80	4
Punica Granatum	0.67	0.55	0.60	11
Santalum Album	0.83	1.00	0.91	5
Syzygium Cumini	0.25	1.00	0.40	1
Syzygium Jambos	1.00	1.00	1.00	7
Tabernaemontana Divaricata	0.50	0.43	0.46	7
Trigonella Foenum-graecum	0.83	0.83	0.83	6
Overall	0.76	0.74	0.74	184

The classification report for the HOG features-based SVM classifier is shown in the table below. Overall, this classifier's precision is 84%, recall is 79%, and f1 score is 79%.

Table 5.10: SVM with HOG Feature Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	1.00	0.80	0.89	5
Amaranthus Viridis	0.69	0.82	0.75	11
Artocarpus Heterophyllus	0.75	1.00	0.86	3
Azadirachta Indica	1.00	0.33	0.50	6
Basella Alba	0.79	0.92	0.85	12
Brassica Juncea	1.00	0.60	0.75	5
Carissa Carandas	0.75	0.60	0.67	10
Citrus Limon	1.00	0.60	0.75	5
Ficus Auriculata	0.88	0.88	0.88	8
Ficus Religiosa	1.00	1.00	1.00	7
Hibiscus Rosa-sinensis	1.00	0.67	0.80	3
Jasminum	0.33	0.75	0.46	4
Mangifera Indica	0.86	0.75	0.80	8
Mentha	0.71	0.71	0.71	7
Moringa Oleifera	0.90	0.82	0.86	11
Muntingia Calabura	0.75	0.75	0.75	4
Murraya Koenigii	0.60	1.00	0.75	3
Nerium Oleander	0.78	1.00	0.88	7
Nyctanthes Arbor-tristi	1.00	1.00	1.00	5
Ocimum Tenuiflorum	1.00	0.33	0.50	9
Piper Betle	1.00	0.86	0.92	7
Plectranthus Amboinicus	1.00	1.00	1.00	3
Pongamia Pinnata	1.00	1.00	1.00	7
Psidium Guajava	0.83	1.00	0.91	5
Punica Granatum	0.67	0.89	0.76	9
Santalum Album	0.43	0.75	0.55	4
Syzygium Cumini	1.00	0.50	0.67	4
Syzygium Jambos	0.83	0.83	0.83	6
Tabernaemontana Divaricata	0.67	1.00	0.80	2
Trigonella Foenum-graecum	0.80	1.00	0.89	4
Overall	0.84	0.79	0.79	184

The SVM classifier using Combined features' classification report is presented in the table below. Overall, this classifier's precision is 82%, recall is 80%, and f1 score is 80%.

Table 5.11: SVM with Combined Feature Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	0.82	0.93	0.87	15
Amaranthus Viridis	0.62	0.89	0.73	9
Artocarpus Heterophyllus	1.00	0.83	0.91	12
Azadirachta Indica	0.67	1.00	0.80	6
Basella Alba	0.91	0.91	0.91	11
Brassica Juncea	0.33	0.20	0.25	5
Carissa Carandas	0.85	0.73	0.79	15
Citrus Limon	0.79	0.94	0.86	16
Ficus Auriculata	0.70	1.00	0.82	7
Ficus Religiosa	0.44	0.57	0.50	7
Hibiscus Rosa-sinensis	0.88	0.70	0.78	20
Jasminum	0.56	0.62	0.59	8
Mangifera Indica	0.79	0.85	0.81	13
Mentha	0.71	0.85	0.77	20
Moringa Oleifera	0.71	1.00	0.83	5
Muntingia Calabura	0.87	0.80	0.83	25
Murraya Koenigii	0.42	0.71	0.53	7
Nerium Oleander	0.88	0.65	0.75	23
Nyctanthes Arbor-tristi	0.78	0.54	0.64	13
Ocimum Tenuiflorum	0.83	0.83	0.83	6
Piper Betle	1.00	0.67	0.80	6
Plectranthus Amboinicus	1.00	1.00	1.00	14
Pongamia Pinnata	0.71	0.83	0.77	6
Psidium Guajava	1.00	0.93	0.96	14
Punica Granatum	0.71	0.71	0.71	21
Santalum Album	1.00	0.60	0.75	5
Syzygium Cumini	1.00	0.67	0.80	12
Syzygium Jambos	0.79	0.88	0.83	17
Tabernaemontana Divaricata	0.90	0.82	0.86	11
Trigonella Foenum-graecum	1.00	0.94	0.97	18
Overall	0.82	0.80	0.80	367

The classification report for the VGG19 pretrained model is presented in the table below. Here, the classifier's overall precision is 97%, recall is 97%, and f1 score is 97%.

Table 5.12: VGG19 Model Classification Report

Medicinal plant	precision	recall	f1-score	support
Alpinia Galanga	1.00	0.88	0.93	8
Amaranthus Viridis	1.00	1.00	1.00	15
Artocarpus Heterophyllus	1.00	1.00	1.00	2
Azadirachta Indica	1.00	1.00	1.00	5
Basella Alba	0.90	1.00	0.95	9
Brassica Juncea	1.00	1.00	1.00	4
Carissa Carandas	1.00	1.00	1.00	6
Citrus Limon	1.00	1.00	1.00	1
Ficus Auriculata	1.00	0.60	0.75	5
Ficus Religiosa	1.00	1.00	1.00	5
Hibiscus Rosa-sinensis	1.00	1.00	1.00	3
Jasminum	0.92	0.92	0.92	13
Mangifera Indica	0.75	1.00	0.86	3
Mentha	1.00	1.00	1.00	12
Moringa Oleifera	1.00	1.00	1.00	3
Muntingia Calabura	1.00	1.00	1.00	8
Murraya Koenigii	1.00	1.00	1.00	5
Nerium Oleander	1.00	1.00	1.00	6
Nyctanthes Arbor-tristis	1.00	1.00	1.00	11
Ocimum Tenuiflorum	1.00	1.00	1.00	2
Piper Betle	1.00	1.00	1.00	3
Plectranthus Amboinicus	1.00	1.00	1.00	4
Pongamia Pinnata	1.00	1.00	1.00	4
Psidium Guajava	1.00	1.00	1.00	8
Punica Granatum	0.91	0.91	0.91	11
Santalum Album	0.86	1.00	0.92	6
Syzygium Cumini	1.00	1.00	1.00	5
Syzygium Jambos	1.00	1.00	1.00	6
Tabernaemontana Divaricata	1.00	0.88	0.93	8
Trigonella Foenum-graecum	0.75	1.00	0.86	3
Overall	0.97	0.97	0.97	184

5.3.6 Output of Grad-CAM

Grad-CAM produces a heatmap visualization with a certain class label as its result. Here in the figure 5.19, I am showing the top 15 predicted images with their activation heatmap superimposed onto them. From these I can clearly see which regions are activated in the last convolutional layer to make the prediction.
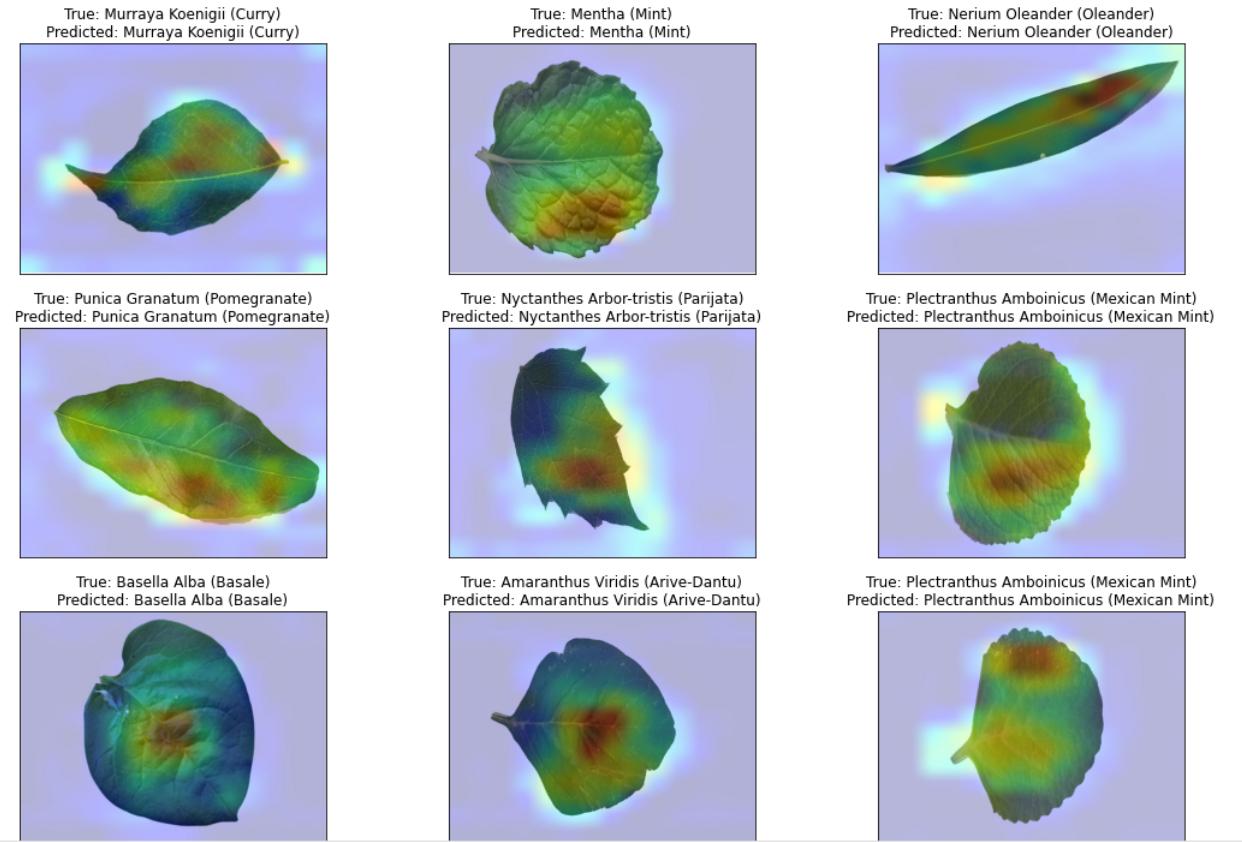
Figure 5.19: Heatmaps generated from GradCam of the top 15 predictions

From all the results above, I can clearly see the VGG19 achieves supremacy over other algorithms used in the project with the highest accuracy, precision, recall and F1 score.For unknown images, I got an accuracy of 96.74% accuracy from this VGG19 pretrained model. Then I implemented GradCam to check the optical clarification from gradient based localization.

Chapter 6

Conclusion & Future Work

6.1 Conclusion

Many useful chemicals and medications are mosty derived from medicinal plants. According to a recent WHO (World Health Organization) assessment, 80% of people globally rely on herbal remedies for some part of their primary healthcare requirements. The WHO estimates that 21,000 plant species have the potential to be used as medicines. To recognize a pre-trained medicinal plant from its leaf, the suggested method uses computer vision and machine learning techniques.Several Machine learning algorithms and deep learning algorithms were then used to classify the leaves from 30 different plant species into their appropriate categories. I tried the CNN classifier and applied SVM classifier with the normal feature, the Grayscale features, the HOG features and the combined features. Then I tried a pretrained model VGG19. Of all the experiments, the highest accuracy of 96.74% was obtained from the VGG19. In this project only this pretrained model was used to gain the most accuracy. Then GradCam was applied on the VGG19 top outputs to get an explanation for the predictions.

Limitations

Firstly in this work, only JPG and PNG images were considered. Other image format was not considered and they might show an error. Secondly, the dataset I used in the project was very small. For this, only 30 classes are used where there are countless other medicinal plants. Also It will not show any good result for any other images other than leafs.

6.2 Future Work

A lot more images and classes as well as other recognizing features such as flowers and fruits can be used for future scope of this work. Furthermore other ML algorithms such as KNN, ANN, Decision tree, Random forest and other pre-trained deep learning models with more features can be explored for more precise classification. APIs can be developed based on this work and could be used as backend of any frontend software for recognizing the proper medicinal plant.

Bibliography

[1] K. A. Shah, M. B. Patel, R. J. Patel, and P. K. Parmar, *Mangifera indica (mango)*, Jan. 2010. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3249901/#:~:text=Various%5C%20parts%5C%20of%5C%20plant%5C%20are,%5C%2C%5C%20leucorrhoea%5C%2C%5C%20haemorrhage%5C%20and%5C%20piles..

[2] A. Gopal, S. P. Reddy, and V. Gayatri, "Classification of selected medicinal plants leaf using image processing," in *2012 International Conference on Machine Vision and Image Processing (MVIP)*, IEEE, 2012, pp. 5–8.

[3] C. A. Priya, T. Balasaravanan, and A. S. Thanamani, "An efficient leaf recognition algorithm for plant classification using support vector machine," in *International conference on pattern recognition, informatics and medical engineering (PRIME-2012)*, IEEE, 2012, pp. 428–432.

[4] C. Arun, W. S. Emmanuel, and D. C. Durairaj, "Texture feature extraction for identification of medicinal plants and comparison of different classifiers," *International Journal of Computer Applications*, vol. 62, no. 12, pp. 1–9, 2013.

[5] D. Bhowmik, H. Gopinath, B. Kumar, D. S, G. Aravind, and K. Kumar, "Medicinal uses of punica granatum and its health benefits," *J Pharmacog Phytochem*, vol. 1, pp. 28–35, Jan. 2013.

[6] R. Janani and A. Gopal, "Identification of selected medicinal plant leaves using image features and ann," in *2013 international conference on advanced electronic systems (ICAES)*, IEEE, 2013, pp. 238–242.

[7] S. L. Badole and K. Y. Patil, "Chapter 34 - pongamia pinnata (linn.) pierre and inflammation," in *Polyphenols in Human Health and Disease*, R. R. Watson, V. R. Preedy, and S. Zibadi, Eds., San Diego: Academic Press, 2014, pp. 463–465, ISBN: 978-0-12-398456-2. DOI: https://doi.org/10.1016/B978-0-12-398456-2.00034-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123984562000347.

[8] V. Mani, A. Ahmad, K. Ramasamy, S. M. Lim, and A. B. A. Majeed, *Murraya koenigii leaves and their use in dementia*, Nov. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780124078246000975.

[9] N. Jamil, N. A. C. Hussin, S. Nordin, and K. Awang, "Automatic plant identification: Is shape the key feature?" *Procedia Computer Science*, vol. 76, pp. 436–442, 2015.

[10] S. Purohit, R. Viroja, S. Gandhi, and N. Chaudhary, "Automatic plant species recognition technique using machine learning approaches," in *2015 International Conference on Computing and Network Communications (CoCoNet)*, IEEE, 2015, pp. 710–719.

[11] G. Arumugam, M. K. Swamy, and U. R. Sinniah, *Plectranthus amboinicus (lour.) spreng: Botanical, phytochemical, pharmacological and nutritional significance*, Mar. 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6274163/.

[12] M. S. Aslam and M. S. Ahmad, "Worldwide importance of medicinal plants: Current and historical perspectives," *Recent Adv Biol Med*, vol. 2, no. 2016, p. 909, 2016.

[13] K. H. Goldberg, A. C. Yin, A. Mupparapu, E. P. Retzbach, G. S. Goldberg, and C. F. Yang, *Components in aqueous hibiscus rosa-sinensis flower extract inhibit innbsp;vitro melanoma cell growth*, Feb. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2225411016000213?via%5C%3Dihub..

[14] V. Hiremath, B. S. Hiremath, S. Mohapatra, and A. K. Das, *Literary review of parijata (nyctanthus arbor-tristis linn.) an herbal medicament with special reference to ayurveda and botanical literatures*, Dec. 2016. [Online]. Available: shorturl.at/ejor3.

[15] R. Prasad, *Medicinal properties of tulsi unravelled*, Mar. 2016. [Online]. Available: https://www.thehindu.com/sci-tech/health/medicinal-properties-of-tulsi-unravelled/article7593813.ece.

[16] B. Şekeroğlu and Y. İnan, "Leaves recognition system using a neural network," *Procedia Computer Science*, vol. 102, pp. 578–582, 2016.

[17] E. Solutions, *Accuracy, precision, recall amp; f1 score: Interpretation of performance measures*, Nov. 2016. [Online]. Available: https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/.

[18] D. Venkataraman and N. Mangayarkarasi, "Computer vision based feature extraction of leaves for identification of medicinal values of plants," in *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, IEEE, 2016, pp. 1–5.

[19] A. Begue, V. Kowlessur, U. Singh, F. Mahomoodally, and S. Pudaruth, "Automatic recognition of medicinal plants using machine learning techniques," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 4, pp. 166–175, 2017.

[20] R. G. De Luna, R. G. Baldovino, E. A. Cotoco, *et al.*, "Identification of philippine herbal medicine plant leaf using artificial neural network," in *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, IEEE, 2017, pp. 1–8.

[21] H. Kan, L. Jin, and F. Zhou, "Classification of medicinal plant leaf image based on multi-feature extraction," *Pattern recognition and image analysis*, vol. 27, no. 3, pp. 581–587, 2017.

[22] A. Sabu and K. Sreekumar, "Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach," in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, IEEE, 2017, pp. 145–149.

[23] A. Sabu, K. Sreekumar, and R. R. Nair, "Recognition of ayurvedic medicinal plants from leaves: A computer vision approach," in *2017 Fourth International Conference on Image Information Processing (ICIIP)*, IEEE, 2017, pp. 1–5.

[24] Techopedia, *What is grayscale? - definition from techopedia*, Aug. 2017. [Online]. Available: https://www.techopedia.com/definition/7468/grayscale.

[25] T. D. Dahigaonkar and R. Kalyane, "Identification of ayurvedic medicinal plants by image processing of leaf samples," *International Research Journal of Engineering and Technology (IRJET)*, vol. 5, no. 5, pp. 351–355, 2018.

[26] S. Naseer, S. Hussain, N. Naeem, M. Pervaiz, and M. Rahman, "The phytochemistry and medicinal value of psidium guajava (guava)," *Clinical Phytoscience*, vol. 4, no. 1, p. 32, Dec. 2018, ISSN: 2199-1197. DOI: 10.1186/s40816-018-0093-8. [Online]. Available: https://doi.org/10.1186/s40816-018-0093-8.

[27] T. Vijayashree and A. Gopal, "Identification of herbal medicinal plant leaves using image processing algorithm," *RESEARCH JOURNAL OF PHARMACEUTICAL BIOLOGICAL AND CHEMICAL SCIENCES*, vol. 9, no. 4, pp. 1221–1228, 2018.

[28] M. A. F. Azlah, L. S. Chua, F. R. Rahmad, F. I. Abdullah, and S. R. Wan Alwi, "Review on techniques for plant leaf classification and recognition," *Computers*, vol. 8, no. 4, p. 77, 2019.

[29] M. Dileep and P. Pournami, "Ayurleaf: A deep learning approach for classification of medicinal plants," in *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, IEEE, 2019, pp. 321–325.

[30] A. H. Mohammed, "Importance of medicinal plants," *Research in Pharmacy and Health Sciences*, vol. 5, no. 2, pp. 124–125, 2019.

[31] Amit, *Explainable ai grad-cam*, Jun. 2020. [Online]. Available: https://medium.com/@aimldl1984/explainable-ai-grad-cam-8acd04dd2dd3.

[32] A. Kulkarni, D. Chong, and F. A. Batarseh, "5 - foundations of data imbalance and solutions for a data democracy," in *Data Democracy*, F. A. Batarseh and R. Yang, Eds., Academic Press, 2020, pp. 83–106, ISBN: 978-0-12-818366-3. DOI: https://doi.org/10.1016/B978-0-12-818366-3.00005-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128183663000058.

[33] Opinion, *The health benefits of amaranthus hybridus and amaranthus viridis*, Mar. 2020. [Online]. Available: https://punchng.com/the-health-benefits-of-amaranthus-hybridus-and-amaranthus-viridis/.

[34] *Performance evaluation: Definition, characteristics, steps*, Jul. 2020. [Online]. Available: https://www.iedunote.com/performance-evaluation.

[35] *Tabernaemontana divaricata*, Jul. 2020. [Online]. Available: shorturl.at/ceJU8.

[36] G. Boesch, *Vgg very deep convolutional networks (vggnet) - what you need to know*, Dec. 2021. [Online]. Available: https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/.

[37] P. P. Kaur and S. Singh, "Classification of herbal plant and comparative analysis of svm and knn classifier models on the leaf features using machine learning," in *Soft Computing for Intelligent Systems*, Springer, 2021, pp. 227–239.

[38] P. P. Kaur, S. Singh, and M. Pathak, "Review of machine learning herbal plant recognition system," in *Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020*, Proceedings of the International Conference on Innovative Computing & Communication (ICICC), 2021.

[39] *Oleander: Health benefits, side effects, uses, dose precautions*, Jun. 2021. [Online]. Available: https://www.rxlist.com/oleander/supplements.htm.

[40] D. I. Sec., *Vgg-19 convolutional neural network*, Mar. 2021. [Online]. Available: https://blog.techcraft.org/vgg-19-convolutional-neural-network/.

[41] A. Tripathi, A. K. Singh, K. K. Singh, P. Choudhary, and P. C. Vashist, *Machine learning architecture and framework*, Apr. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128212295000057.

[42] F. Ur Rehman, M. Kalsoom, M. Adnan, *et al.*, "Importance of medicinal plants in human and plant pathology: A review," *Int. J. Pharm. Biomed. Res*, vol. 8, pp. 1–11, 2021.

[43] *White sandalwood: Health benefits, side effects, uses, dose precautions*, Jun. 2021. [Online]. Available: https://www.rxlist.com/white_sandalwood/supplements.htm.

[44] *Basic cnn architecture: Explaining 5 layers of convolutional neural network*, Aug. 2022. [Online]. Available: https://www.upgrad.com/blog/basic-cnn-architecture/.

[45] K. Jagran, *Jamaican cherry: Is it good for health?* May 2022. [Online]. Available: https://krishijagran.com/health-lifestyle/jamaican-cherry-is-it-good-for-health/.

[46] *Lemon*, Aug. 2022. [Online]. Available: https://en.wikipedia.org/wiki/Lemon.

[47] S. Basu, *Jamun: Medicinal uses, therapeutic benefits for skin, diabetes and supplements.* [Online]. Available: https://www.netmeds.com/health-library/post/jamun-medicinal-uses-therapeutic-benefits-for-skin-diabetes-and-supplements.

[48] S. Booth, *Jackfruit: Nutrition, health benefits, amp; how to prepare.* [Online]. Available: https://www.webmd.com/food-recipes/health-benefits-jackfruit.

[49] *Chemical composition and biological activities of mentha species.* [Online]. Available: https://www.intechopen.com/chapters/54028.

[50] *Explainable ai.* [Online]. Available: https://www.ibm.com/watson/explainable-ai.

[51] *Less known facts about the health benefits of basella alba.* [Online]. Available: http://ccras.nic.in/content/less-known-facts-about-health-benefits-basella-alba.

[52] *Neem: Overview, uses, side effects, precautions, interactions, dosing and reviews.* [Online]. Available: https://www.webmd.com/vitamins/ai/ingredientmono-577/neem.

[53] *Peepal : Benefits, precautions and dosage.* [Online]. Available: https://www.1mg.com/ayurveda/peepal-198.

[54] *Rasna: Health benefits, ayurvedic uses, formulations, dosage and side effects.* [Online]. Available: https://www.netmeds.com/health-library/post/rasna-health-benefits-ayurvedic-uses-formulations-dosage-and-side-effects.

[55] *Top 10 medicinal benefits of betel leaves.* [Online]. Available: https://www.medindia.net/dietandnutrition/top-10-medicinal-benefits-of-betel-leaves.htm.

[56] C. Wahome, *Rose apples: How they can benefit your health.* [Online]. Available: https://www.webmd.com/diet/health-benefits-rose-apples.