STUDIES ON COMPUTER VISION BASED IDENTIFICATION OF SPECIES AND MATURITY STAGES FOR SOME MEDICINAL PLANTS AND AGRICULTURAL PRODUCES

Thesis submitted by

GUNJAN MUKHERJEE

Doctor of Philosophy (Engineering)

DEPARTMENT OF PRINTING ENGINEERING

FACULTY COUNCIL OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

KOLKATA 700032, INDIA

2022

JADAVPUR UNIVERSITY
KOLKATA- 700032, INDIA

INDEX NO. 241/18/E

1. **Title of the Thesis** — **Studies on Computer Vision Based Identification of Species and Maturity Stages for Some Medicinal Plants and Agricultural Produces**

2. **Name, Designation &Institution of the Supervisor** —
Dr. Arpitam Chatterjee,
Associate Professor,
Department of Printing Engineering
Jadavpur University, Salt Lake Campus Sector-III, Block-LB, Plot-8,
Kolkata- 700098 , India
Phone:9836020770
Fax: 03323357254

3. **List of Publications** — **<u>Papers published in international Journals</u>**

1) Gunjan Mukherjee, Bipan Tudu , Arpitam Chatterjee. (2021) A convolutional neural network-driven computer vision system toward identification of species and aturity stage of medicinal leaves: case studies with Neem, Tulsi and Kalmegh leaves, Soft Computing, 25 (22), pp. 14119-14138.DOI: 10.1007/s00500-021-06139-9.

2) Gunjan Mukherjee, Bipan Tudu , Arpitam Chatterjee . (2022) A multi-channel convolution neural network driven computer vision system towards identification of Species and maturity stage of banana fruits: case studies with Martaman and Singapuri banana‟, *Int. J. Computational Intelligence Studies, Vol. 11, No. 1,pp.1–23.*

3) Gunjan Mukherjee, Arpitam Chatterjee, Bipan Tudu .(2022) Identification of the types of disease for tomato plants using amodified gray wolf optimization optimized Mobile NetV2 convolutional neural network architecture driven computer vision framework. *Concurrency Computaiot Pract Exper.;e7161. doi: 10.1002/cpe.7161*t:

4) Gunjan Mukherjee, Bipan Tudu, Arpitam Chatterje,TomNet: A deep neural network based computer vision towards simultaneous classification of species and maturity stages of tomato (Communicated) .

Papers published in International Conference Proceeding

1) Gunjan Mukherjee, Arpitam Chatterjee , Bipan Tudu.(2016) Study on the potential of combined GLCM features towards medicinal plant classification, 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC). IEEE
2) Gunjan Mukherjee, Arpitam Chatterjee, Bipan Tudu .(2017) Color feature based approach towards maturity discrimination ofmedicinal plants: Case study with Kalmegh (*Andrographis paniculata*), *Computational Science and Engineering – Deyasi et al. (Eds)© Taylor & Francis Group, ISBN 978-1-138-02983-5*.
3) Gunjan Mukherjee, Arpitam Chatterjee, Bipan Tudu. (2017) Morphological feature based maturity level identification of kalmegh and Tulsi leaves, third international conference on Research n computational intelligence and communication networks.IEEE.
4) Devraj. Vishnu, Gunjan. Mukherjee and Arpitam. Chatterjee, "A computer vision approach for grade ideaD. Vishnu. D, Mukherjee. G, and Chatterjee. A., "A computer vision approach for grade identification of rice bran," *2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2017, pp. 10-14, doi: 10.1109/ICRCICN.2017.8234473.ti 10.1109/ICRCICN.2017.8234473.
5) Gunjan Mukherjee, Arpitam Chatterjee, Bipan Tudu.(2022) A Computer Vision Approach Towards Maturity Stage Classification of Tomatoes Using Second Order Wavelet Features.In: Saraswat, M., Roy, S., Chowdhury, C., Gandomi, A.H. (eds) Proceedings of International Conference on Data Science and Applications . Lecture Notes in Networks and Systems, vol 288. Springer, Singapore. https://doi.org/10.1007/978-981- 16-5120-5_10

Statement of Originality

I, Gunjan Mukherjee, registered on the 25th April, 2018 do hereby declare that this thesis entitled "Studies on Computer Vision Based Identification of Species and Maturity Stages for Some Medicinal Plants and Agricultural Produces" contains literature survey and original research work done by the undersigned candidate as part of Doctoral studies.

All information in this thesis has been obtained and presented in accordance with existing academic rules and ethical conduct. I declare that, as required by rules and conduct, I have fully cited and referred all materials and results that are not original to this work.

I also declare that I have checked this thesis as per the "Policy on Anti Plagiarism, Jadavpur University, 2019",and the level of similarity as checked by iThenticate software is 5%.

Gunjan Mukherjee
03/11/2022

Gunjan Mukherjee
Index No. 241/18/E
Date:

Arpitam Chatterjee
03.11.22

Associate Professor
Department of Printing Engineering
Jadavpur University, Salt-lake Campus
Kolkata - 700 098

Dr. Arpitam Chatterjee
Associate Professor,
Department of Printing Engineering
Jadavpur University, Salt Lake Campus
Sector-III, Block-LB, Plot-8,
Kolkata- 700098 , India

CERTIFICATE FROM THE SUPERVISORS

Date:

This is to certify that the thesis entitled "Studies on Computer Vision Based Identification of Species and Maturity Stages for Some Medicinal Plants and Agricultural Produces" submitted by Shri Gunjan Mukherjee , who got his name registered on 25th of April,2018 for the award of Ph. D (Engg.) degree of Jadavpur University is absolutely based upon his own work under the supervision of Dr.Arpitam Chatterjee and that neither his thesis nor any part of the thesis has been submitted for any degree /diploma or any other academic award anywhere before.

Associate Professor
Department of Printing Engineering
Jadavpur University, Salt-lake Campus
Kolkata - 700 098

Dr. Arpitam Chatterjee

Associate Professor,

Department of Printing Engineering
Jadavpur University, Salt Lake Campus Sector-III, Block-LB, Plot-8,
Kolkata- 700098 , India

DEDICATED TO

TO

ALL MIGHTY GOD

ACKNOWLEDGMENTS

The success of performing any noble job and its real pleasure remain incomplete unless the people connected to the job as the integral part with their continuous support, supervision ultimately paves the way to make it happen. I , therefore would like to acknowledge with sincere gratitude to the kind assistance and supports provided by all concerned persons.

It is my great privilege to express my profound and sincere gratitude to Dr. Arpitam Chatterjee , Department of Printing Engineering, Jadavpur University, Kolkata for his continuous cooperation, supervision and precious guidance at every stage of the research works and preparation of thesis. The completion of this thesis would have become incomplete without his persistent supervision

I must also mention my heartfelt gratitude to Prof. Bipan Tudu, Department of Instrumentation & Electronics Engineering, Jadavpur University, Kolkata for his valuable suggestions, right direction and immense supports at each and every steps of the entire research program.

I am also grateful and thankful to Prof. Soumen Basak, Head of the Department of Printing Engineering, Jadavpur University, and Kolkata for extending his supports with all sorts of facilities for carrying out the research during the entire journey.

I would like to convey my sincere thanks to all the faculty members, technical and non technical staffs of the Department of Printing Engineering, Jadavpur University for their timely and prompt support and help towards the research program.

Finally, I would like to express my earnest gratitude to my mother, colleagues and friends to carry out this work. I would like to extend my warm thanks to my son and daughter for sharing their expected time to concentrate to my research work .Last but not the least, I would like to express my heartfelt and honest gratitude to my dear wife for her stern persistence, unconditional support , great motivation and profound encouragement during the entire research work.

03/11/2022

Gunjan Mukherjee

Abstract

We are familiar with different varieties of agricultural produces. The agricultural products comprise either different types of foods like fruits or vegetables or different leaves of medicinal plants used for the remedial of different critical and chronic diseases. All of these agricultural products show myriad food or medicinal values towards the human beings. The agro products neem, tulsi, kalmegh etc have been very familiar to us and has been proved to be the integral part of ayurveda from very ancient time with high potential of curing many different critical and chronic human diseases. Some notable contribution of these plants are natural healing powers towards many ailments like blood pressure, digestion disorder, blood purification , skin diseases and many other chronic diseases like cancer etc. The remedial qualities of such plants are attributable to the bio actives present in the plants. The bio actives or the biochemical adoptogen varies from one species to another and also from one maturity stage to another within the same species.

Like the medicinal plants, fruits like tomato or banana are also contributing majorly towards the wellbeing of the human health. Banana contains different minerals like potassium, magnesium, calcium, manganese etc which keeps the good health of human heart and apart from showing the remedial signs like fighting anemia, controlling blood pressure etc. The nutrients present in the fruit banana vary from one species to another and also varies from one stage of maturity to another stage with in the same species of the fruit group.

Like banana, tomato is also another widely consumed fruit and bear many different food values and remedial actions on the human health. It shows enough potentiality towards remedial of many chronic diseases like cancer, osteoporosis and other different types of cardiovascular

diseases. Presence of the bioactive lycopane in tomato shows typical remedial actions for the chronic diseases like reducing risk of cancer. In addition to such remedial qualities, it also shows curing potentials for diseases like osteoporosis, cardiovascular diseases etc. The action of such bioactive present can vary from one species to another and also from one maturity stage to another within the same species.

Despite having the beneficial aspects, plants show some degree of susceptibility towards the typical diseases and as a result undergo damages .Most of the diseases are caused by the bacterial or fungal infection and can spread at very rapid rate from one plant to the other. The need of identification of plant species and the disease types are being performed manually. The manual operations become subjective, high time taking and invasive in nature.

In the first phase of study of this thesis, the species and maturity discrimination of different medicinal plants have been explored. The leaves of the three categories of plants neem, tulsi and kalmegh in heir different maturity stages have been collected, preprocessed and finally grouped into the train and test set . The deep learning based CNN model has been formed and optimized with the modified GWO optimizer. The accuracy of classification for unknown leaf has been achieved at 99% approximately.

In the second phase of study, the species and maturity discrimination of the fruits and vegetables have been explored. The different species of fruits banana and the tomato were selected in three distinct maturity grades premature, mature and over mature. After preprocessing operations, the images were captured to make the train set. The dataset were finally fed to the deep learning based classifier CNN model optimized with the PSO or GWO optimizer. Depending on the performance criterions of the optimizer, the metrics were developed. The classification accuracies for both of the fruits were found very close to 99%.

In the third phase of study, seven different types of diseases of tomato plant leaves have been explored. The fungal or bacteria affected leaves have been collected and the captured images of those leaves have been formed into the training and test sets respectively. The training dataset have been utilized to train up the light weighted model mobileNet optimized with the modified GWO techniques. The classifier identifies the unknown leaves with the proper disease types with the accuracy value 98% approximately. The metric values and the accuracies obtained in respective phases of research studies vouches the fact the explored computer vision models can be potential tool towards classification of the different produces in terms of species and maturity levels.

Contents

List of Figures

List of Tables

List of Abbreviation

R&D	***Research and Development***
GC	***Gas Chromatography***
HPLC	***High Performance Liquid Chromatography***
NIRS	***Near Infra-RedSpectrography***
SVM	***Support Vector Machine***
PCA	***Principal Component Analysis***
LDA	***Linear Discriminant Analysis***
MRI	***Magnetic Resonance Imaging***
PH	***Potential of Hydrogen***
LBP	***Local Binary Pattern***
KNN	***K Nearest Neighbors Algorithm***
GA	***Genetic Algorithm***
CNN	***Convolutional Neural Network***
NMR	***Nuclear Magnetic Resonance***
ANN	***Artificial Neural Network***
PNN	***Probabilistic Neural Network***
BPNN	***Back Propagation Neural Network***

RGB	*Red Green Blue*
HSV	*Hue Saturaion Value*
SWNIR	*Short Wave Near Infra Red*
VIS NIR	*Visible Near Infra Red*
R-CNN	*Region Convolutional Neural Networks*
LBPH	*Local Binary Pattern Histogram*
ResNet	*Residual Network*
BNN	*Binarized Neural Network*
DNN	*Deep Neural Network*
VGGNET	*Visual Geometry Group*
DBN	*Deep Belief Network*
RNN	*Recurrent Neural Network*
RNTN	*Recursive Neural Tensor Network*
GWO	*Grey wolf Optimization*
GLCM	*Grey Level Co-occurrence Matrix*
PSO	*Particle Swarm Optimization*
NN	*Neural Network*
ReLU	*Rectified Linear Unit*
UV VIS	*Ultra Violet Visible Spectroscopy*
BPSO	*Binary Particle Swarm Optimisation*

TP	***True Positive***
TN	***True Negetive***
FP	***False Positive***
FN	***False Negetive***
ROC	***Receiver Operating Characteristics***
TPR	***True Positive Rate***
FPR	***False Positive Rate***
AUC	***Area Under the curve***
C_cost	***Computational cost***
D_C_cost	***Depth wise cost***
CR	***Depth wise convolution cost***
MSE	***Mean Square Error***
OS	***Operating system***

CHAPTER 1
INTRODUCTION AND SCOPE OF THE THESIS

1.1. Introduction

We are familiar to many varieties of agricultural products in our day to day lives. The conventional agro produces bear extensive impacts as the food item that provide the varying degree of food as well as medicinal values. These cultivars provide good immunity and nutrition as being edible along natural healing power for many diseases. There are some types of plants among the agro products which show strong therapeutic effects towards remedies of chronic diseases and are generally treated as the medicinal plants. The plants such as the *Neem*, *Tulsi*, *Kalmegh*, *Basaka*, *Turmeric*, *Ginger*, *Sarpagandha*, and many more such medicinal leaves are extensively used to cure number of common ailments of human beings. They are integral part of a popular alternative medicinal approach called Ayurveda which is originated in India but popular across the world [1] from the ancient age of human civilization. Blood pressure, digestion disorders, blood purification, different skin diseases and many other chronic diseases including cancer are being treated using medicinal plants. The biochemical composition varies from one medicinal leaf to another. For instance *Andrographolide* is the major constituent of *Kalmegh* and contributes towards the treatment of diseases like respiratory tract infection [2].The constituents *oleanolic acid, ursolicacid,rosemarinic acid, eugenol, carvacrol, linalool, β-caryophyllene* and*β-element* [3] are found in *Tulsi* as adaptogen. Presence of several chemicals like *nimbin*, *nimbini* and *nimbidin* [4] in the *Neem* leaves renders its' antiviral, antibacterial, antifungal and anthelmintic properties. The amount of bio-chemicals present in the medicinal plants plays the significant role in remedies of different diseases and varies with the maturity of leaves.

Like medicinal plants, the fruits such as tomato or banana are also major contributors of human health. Banana fruit is most popular among different fruits due to its world-wide availability. It contributes nearly 15% of the food production market [5] .Banana contains the high fiber which is responsible for improvement of the human digestive system. Moreover it also contain the minerals like *potassium, calcium, magnesium, manganese, iron, folate, niacin, riboflabin* and *B6*.These nutrients have the high impacts in maintaining overall body function [5].Some significant contribution of the banana fruits are keeping the good heart health, controlling blood pressure, fighting anemia and making ease in the overall digestion. All the nutrients present in banana and food values vary with different stages of maturity.

Like banana, tomato is also another popular and much widely consumed fruit. It contains the high amount of fiber along with *vitamin A* and *C* which contributes to the medicinal values [6]. Some of the important contributions of this fruit includes remedial potential for diseases like reducing risk of cancer, osteoporosis and other cardiovascular diseases [7], *Lycopane* [8] is the chemical compound present in tomato which renders the anticancer and antioxidant properties. Like medicinal leaves and bananas the nutritional values of tomato also varies depending on the species and maturity stages.

Despite several advantages the agro produces are also susceptible to damage due to different plant diseases. For instance, the leaf diseases in tomato are quite common and incur huge agricultural and economic loss. There are different diseases and each type of disease have different impacts. The diseases are mostly caused due to the fungal and bacterial organisms [9]. At the same time some of the plant diseases can spread from one plant to the other quickly which raise the need of early detection of plant diseases. Some of the major impacts of bioactive properties of agro produces in human life can be shown in Fig. 1.1.

The need of identification of species, maturity stages and diseases in the agro produces remain as utmost important tasks in agricultural and economic growth of any country. All these tasks are commonly performed manually by human experts however such manual operations are limited by high time consumption, subjective assessment and invasive in nature. Thus automated identification, classification and prediction operations becomes an impactful R&D field [10][11].This thesis explores the scope of the such automation in these three directions; i) identification of species and maturity stages for medicinal plants, ii) identification of species and maturity stages for fruits and vegetables and iii) early detection of plant diseases.

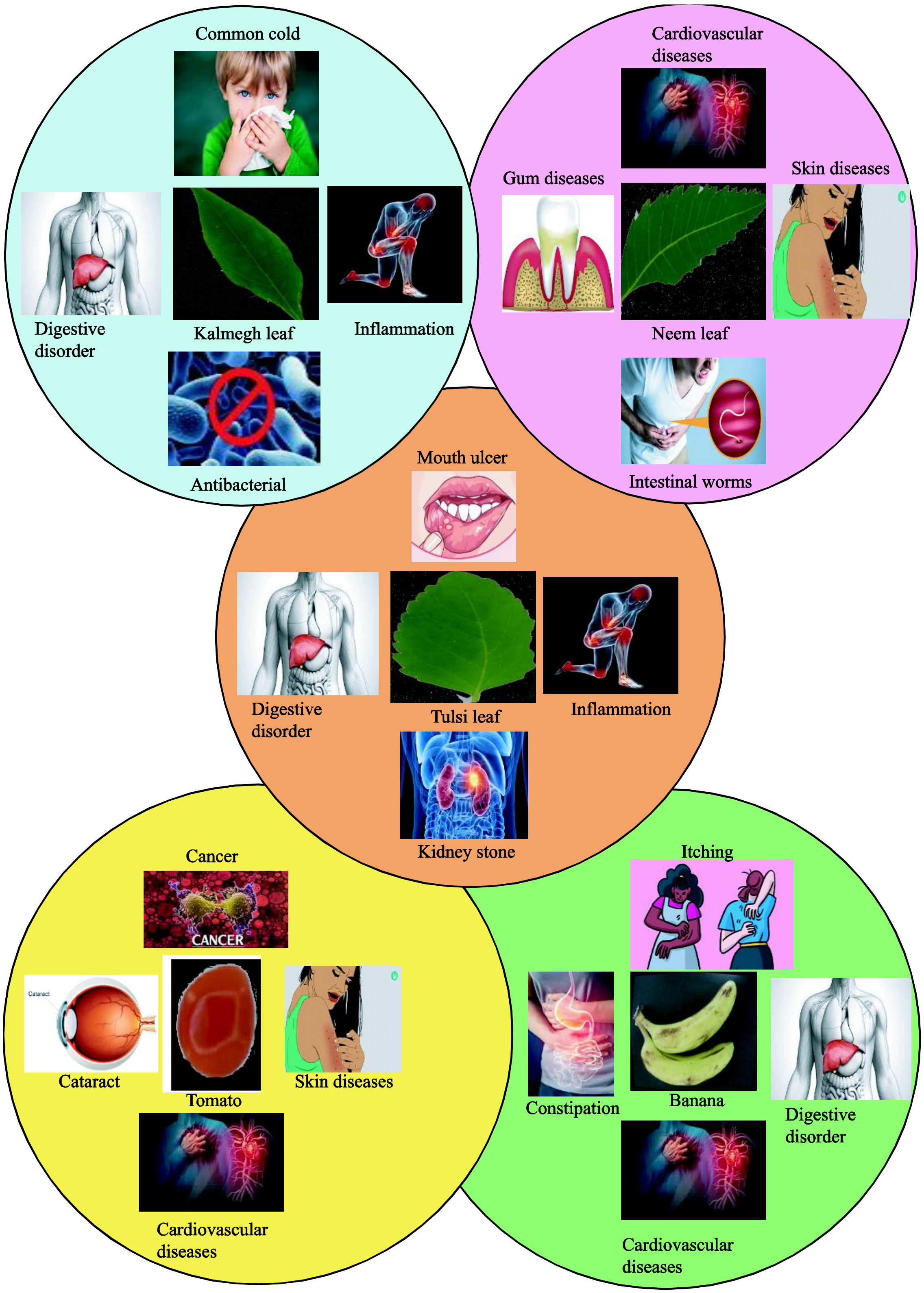


Fig. 1.1: *Some major contributions of agro produce towards human health issues and diseases.*

1.2 Conventional methodologies for agro produce grading

The grading and sorting of agro produces is an important task. It includes different identification, classification and prediction operation like detecting species and maturity stage since these two characteristics highly govern the medicinal and food values. Different analytical and instrumental methods are adopted for such grading apart from grading by human experts. Methods like gas chromatography (GC), high performance liquid chromatography (HPLC) , near infra-red spectrography (NIRS) are some of the popular techniques involved in this regard. Instrumental approach based on electronic nose and tongue, magnetic resonance in combination with chemometric method, optical method like absorption and detection property like Raman spectroscopy and NIR spectroscopy have as well been used. These fundamental operational principles of these methods have been discussed in the following sections.

1.2.1 Gas chromatography-mass spectroscope (GC-MS)

GC-MS, as shown in Fig.1.2 consists of two parts the gas chromatograph and mass spectrometer. It is a scientific diagnosis technique to detect special type of the compounds responsible for the plant growth called *phytochemicals*. This also attributes to the medicinal value of plants. The presence of bioactive in the plant renders the medicinal feature to the plant. The main objective of the GC-MS is to identify the unknown compounds from the permanent gas.

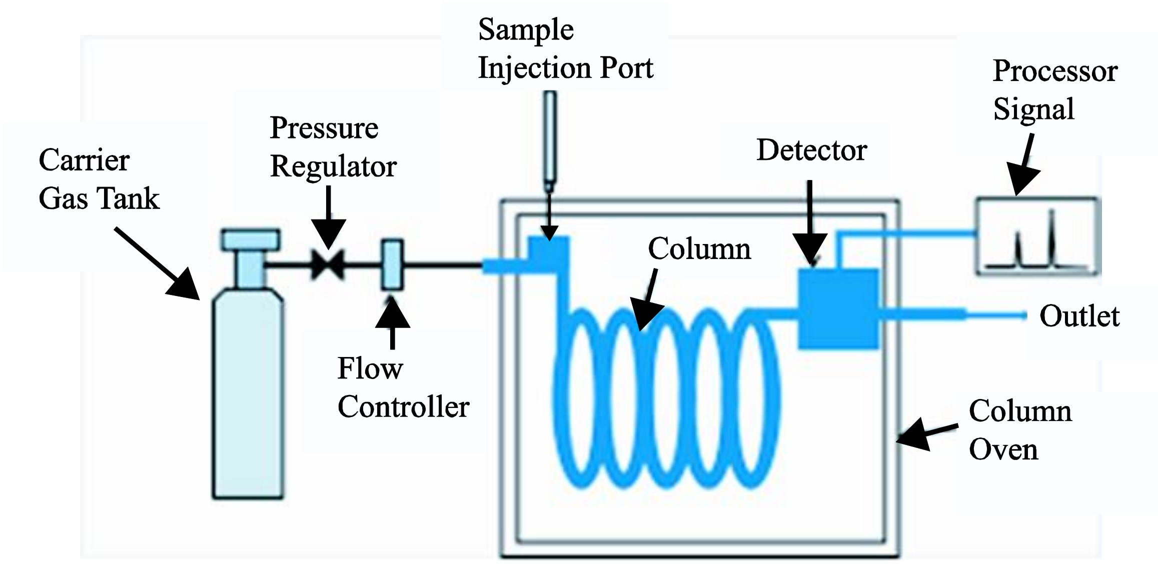


Fig 1.2.*Working principle of GC-MS.(Courtesy:Gas Chromatography Mass Spectrometry (GCMS) - Shimadzu |https://www.shimadzu.eu.com)*

1.2.2. High performance liquid chromatography(HPLC)

HPLC process is carried out in order to separate the compounds that are present in the sample and can be dissolved in liquid in trace concentration. HPLC process involves pumping the sample mixture or the analyte in a solvent under high pressure. Plant hormones regulate the overall growth of plant. The plant hormones are generally determined by application of HPLC techniques. The exact quantification of the hormones can be assessed by the reverse phase liquid chromatography. The basic principle of the HPLC, as shown in Fig.1.3 is dependent on the fact that some of the molecules move slowly compared to the rest of the molecules present in the chromatography column. This brings about the attraction force between the mobile phase (liquid and gases) with the solid phase (solid or liquid). The molecules experiencing more attraction force with the stationary phase take more time to pass through the column and vice versa.

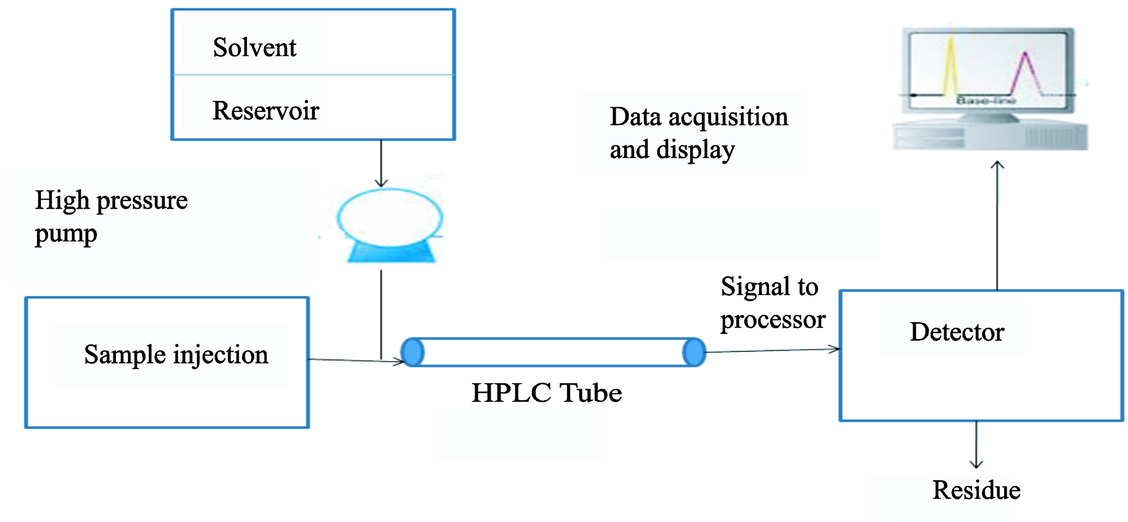


Fig 1.3.*Working principle of HPLC. (Courtesy:C:\Users\bwu\Downloads\ HPLC basics – Principles and parameters – KNAUER https:\www.knauer.net › Analytical-HPLC-UHPLC › HPL.)HPLC basics – Principles and parameters – KNAUERhttps://www.knauer.net › Analytical-HPLC-UHPLC › HPL.)*

1.2.3. Near infrared spectroscopy (NIR)

NIR is a popular technique involved in estimating the macro and micro nutrients present in the leaf tissues. NIR spectroscopy, as shown in Fig. 1.4, within the spectral range 780 to 2500 nm provides the intricacies of the information related to the vibration features as a result of the combinations of bond. It is often preferred over other instrumental methods for nondestructive, inexpensive and faster execution properties. The chemicals as well as the physical compositions of the substances under study can be revealed by means of typical patterns formed under the NIR spectra.

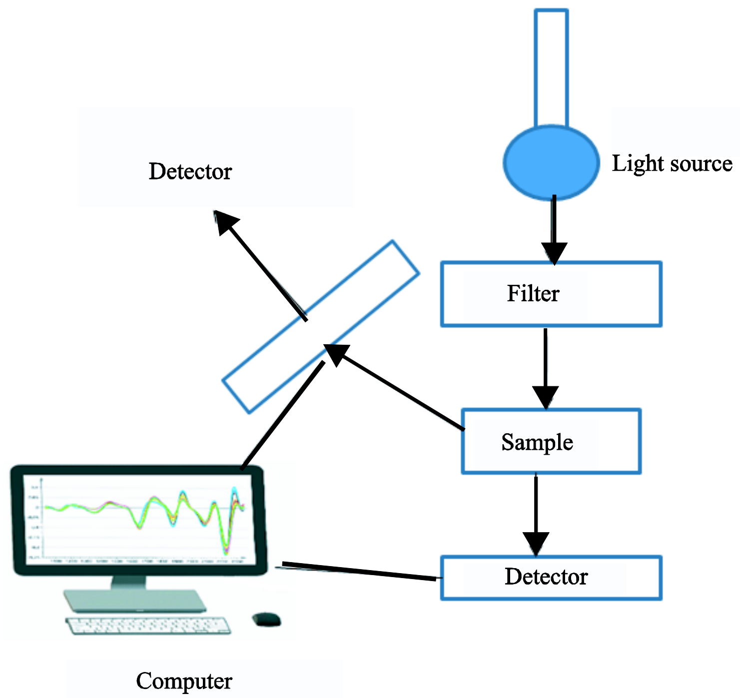


Fig 1.4.*Working principle of NIR spectroscopy.(Courtesy:Why FT-NIR spectroscopy? | Brukerhttps://www.bruker.com › ft-nir-spectrometers › what-is.)*

1.2.4 Electronic nose (E-nose)

Electronic nose, as shown in Fig. 1.5 is a device to monitor the volatile production of maturity states with the help of different metal oxide sensors. With the help of different computational methods like principle component analysis (PCA) and linear discriminant analysis (LDA). E-nose can perform food grading efficiently.

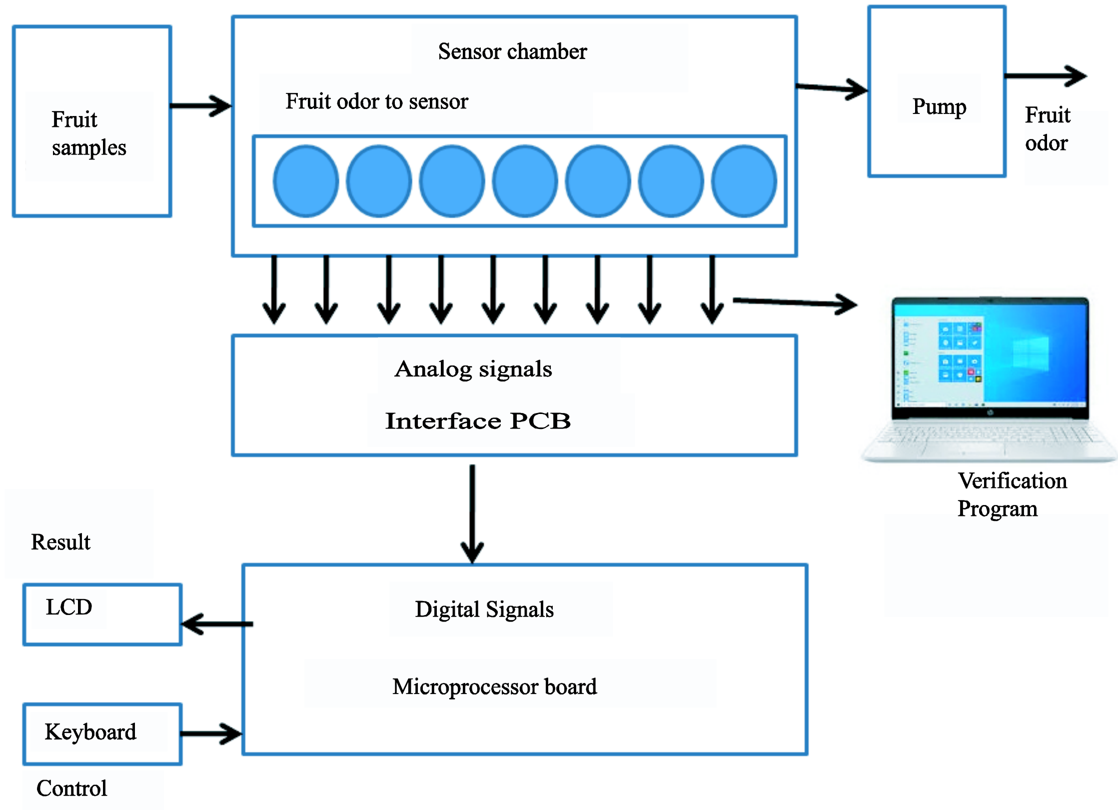


Fig 1.5*.Working principle of E-nose.(Courtesy:*C:\Users\bwu\Downloads\ What is Electronic Nose (enose): Working Principle https:\www.elprocus.com › electronic-nose-work*What is Electronic Nose (enose): Working Principle https://www.elprocus.com › electronic-nose-work*

1.2.5 Magnetic resonance with chemometrics

MRI system is made up of the permanent magnet fitted with the radio frequency coils. MRI system as shown in Fig. 1.6 is used to recognize the physio chemical changes in agro produces. The grading is performed using the magnetic resonance images of different grades with the help of qualitative and quantitative image analysis methods. Chemometrics involves the mathematical, statistical and other methods to design the optimal measurement procedure to provide the pertinent chemical information. This is also a nondestructive approach.

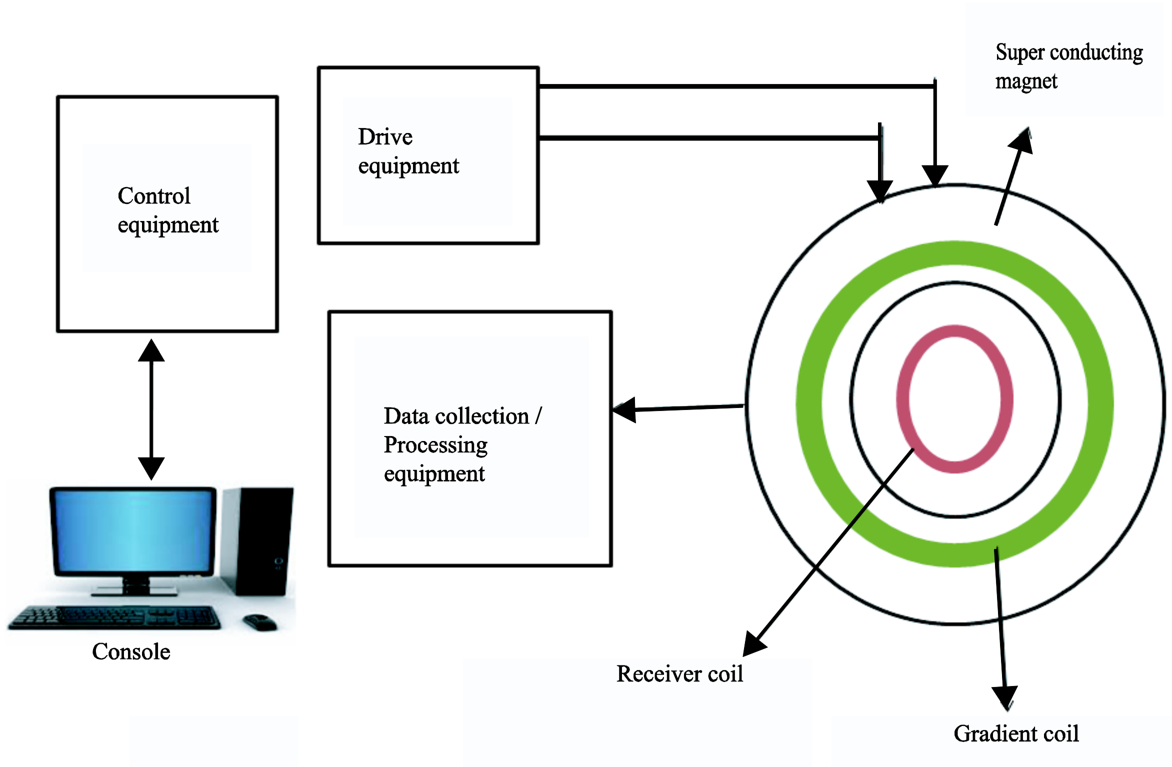


Fig1.6.*Working principle of magnetic resonance with chemometrics. (Courtesy:https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri)*

1.2.6. Chemical sensors

Chemical sensors, as shown in Fig. 1.7 convert the physical and chemical properties of analytes into measurable signals having the magnitude equivalent to the concentration of the subjected analytes. These work on sugar contents, PH, ethylene emission and absorption detection. There are two distinct components in chemical sensors; receptors and transducers. The receptor comes directly in contact with the analyte. Some of the receptors are active in bringing about the chemical reactions and some others single out the specific molecules. Chemical sensors based on the carbon nanostructure composites are used to assess the ripeness of tomato fruit. The change in the conductance and capacitance of the nanocomposite sensors are measured.

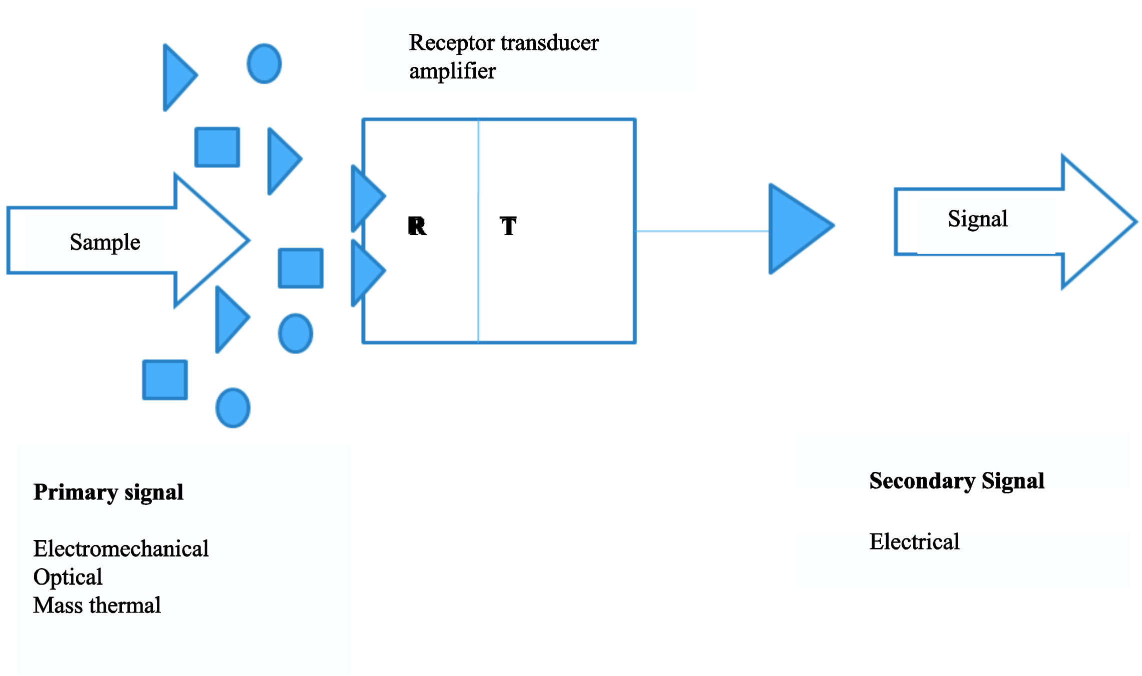


Fig 1.7.*Working principle of chemical sensors.(Courtesy:Chemical sensors and biosensors https://bme.lth.se › Courses › Kemiska_sensorer)*

1.3 Literature survey

Table 1.1 consolidates literature survey on some of the major applications of the previously stated instrumental and analytical techniques in quality grading of agro produces.

Table 1.1. *A literature review*

Applications	Working Principle	Data analysis technique	Remarks	Ref
Recognizes different plant species.	Statistical features obtained from the digital images	Feeding of Statistical features to the neuro fuzzy classifier.	Leaves are fragile and prone to different environmental and biological factors. The procedure has been extended to address the small deformation.	[12]
Leaf shape , color and texture has been analyzed	The plant leaf is identified based on GIST and local binary pattern.	The binary particle swarm optimization technique as feature reduction techniques.	Out of the many machine learning approaches applied , the decision tree has got the better performance	[13]
Recognition of the plant leaves and species.	Morphological features like centroid, major axis , minor axis , perimeter , solidity , orientation	Use of KNN, decision tree and multilayer perception classifier on the open source data	The performance of classification has been enhanced by using the SVM with adaptive boosting techniques.	[14]

		set FLAVIA		
Classifying the documents into predefined categories according to their contents	Association rule mining techniques for correlation of the relationships among a large set of data items.	Use of the niching memticalgorithm and the genetic algorithm in hybridmode.	The improved GA algorithm contains rules inspired by the natural genetic populations to evaluate the problems.	[15]
Reviews the therapeutic uses of tulsi plant and its application in curing the diseases like cold, cough, malaria, dengue etc.	The bio organic constituents shows the antioxidant, anti atherogenic, anti-aging ,anti-inflammatory, antis tress and antihelmintic	Assessment of the anti inflamatory activities by using phyto chemical constituents .	Tulsi leaves show the excellent parricidal activities.	[16]
Evaluation of the medicinal qualities of kalmegh plant leaves	Origin, distribution, plant botany, taxonomy, genetics, edaphic and climatic requirements, planting geometry, nutrient management, irrigation management, disease and weed management, maturity and harvest	Assessment of the presence of bioactivesand other chemical properties due to continuous application of manures and better irrigation system.	The sharp decline in the availability of the drugs to the different medicine industries and escalation of price is attributable to the indiscreet collection from wide sources without paying any attention towards the study.	[17]

	period.			
Determination of the leaf based identification of soybean plant leaf cultivars	Implementation of the multi feature combined systems on the basis of multiple leaf morphological features like shape, texture and venation.	The CNN based algorithm with fine-tuned Xception network for the classification purpose.	All the soybean leaves were taken from each respective growth period of the plant.	[18]
Detection of the water deficiency in the leaf towards assessment of the draught as the abiotic stresss in the leaves of plants.	Use of the low field NMR (nuclear magnetic resonance) relaxometry as the development marker on the NMR relaxation spectra.	Determination of mild water stress by nitrogen rate in shoots and amino acids	The sensitivity of NMR was affected by the stage of development of the leaves. The impact of water was assessed on two different ranks of leaves.	[19]
Identification of Indian medicinal plants on the basis of leaf characteristics with the medicinal plants like Hibiscus, Betle, Ocimum, Leucas, Vinca, Murraya, Centella, Ruta and Mentha.	Treatment of the leaf characteristics like leaf color, area and edge features as the characteristic features.	The histogram of the color and edge for identifying different medicinal plants.	The acquired image characteristics obtained through the experiment were finally compared against the features of the images from the image database for identification purpose.	[20]

Review of different techniques for the classification and recognition of plant leaves	Neural network based approach like artificial neural network(ANN), Probabilistic neutral network(PNN), Convolutional neural network(CNN),K nearest neighbor(KNN) and support vector machine (SVM) and other combined techniques.	The thresholding of gray scale image and contour extraction and refinements as edge recognition by means of the Suzuki algorithm.	Leaf patterns for different plants are different. The most challenging part is to extract the distinctive features of the leaves by many statistical means.	[21]
To classify the plants by means of images captured by mobile phones.	The morphological features of the leaf like aspect ratio, rectangularity, convex area ratio, convex perimeter ratio, sphericity and circularly, eccentricity, form factor, regional moments of inertia and angle code histogram.	The classification of plant leaves on the basis of the nearest neighbor distance of the query leaf features from the median features of each species in the training set.	The speed of the live segmentation algorithm can be optimized in order to enhance the usability.	[22]
To identify the Leaf types automatically by	Recognition of the plant leaves by means of the	Deeplearning based techniques on the	The multiscale learning process has been adopted which	[23]

using the deep learning approach.	CNNarchitecturegoog lenet	Flaviadataset.	deforms several sizes by using the maximum and minimum sizes. The over fitting phenomena can be tackled by engaging the less learning data.	
To assess the internal quality of the tomato fruit in order to show the beneficial aspect of the chromacity values and lycopane contents of the tomato towards the maturity based classification.	Assessment of the chromaticity values by means of the image processing by engaging the calorimeter.	Use of the external visual colors for the classification of tomatoes of three sorted classes mature , green intermediate and advanced including six ripeness stages	The TSS and lycopene component of the tomato were squeezed out from the slurry of cut off portion of the tomato fruit.	[24]
To detect the tomato maturity stages by means of computer vision.	The computer vision based towards the determination of maturity of tomato based on the color features extracted from number of	Use of the classifier BPNN for reaching to the classified status of the tomato on the basis of colour features	The average hue values of the sub region were extracted as the feature color values for distinction between different	[25]

	concentric circles from the surface of tomato.		tomato fruits.	
Extraction of different features of fruits like color, size ,volume shape and textures.	Color mapping method and Montecarlo method followed by the Curvlet and gabor wavelet transform for extraction of the texture features along with discrete wavelet, fourier transform, scale space, chain code and clustering methods .	Use of both the external features and internal features such as aroma, test flavor , sweetness etc for extraction of features	Multiple methods are available for the extraction of features which are also dependent on the factors concerned.	[26]
To develop the model for automatic grading of the apple fruit.	machine vision and the artificial neural network model based classification system along with the back propagation ANN classifier .	Use of the physical parameters like size, color and external defects for apple classification purpose. Two categories of the parameters surface level apple quality and quality of apple are evident.	The experimental setup used the shorter number of parameters for achieving the shorter calculation time and with lower prediction error.	[27]
To detect the fresh tomato	The back propagation neural network	The feature extraction based	Three distinct maturity levels were	[28]

using computer vision.	basedclassification techniques based on the feature color values of two types of tomatoes roma and pear.	on the maximum inscribed circle on the surface of tomato and area followed by the subdivision into five concentric circles with color feature value in terms of hue color.	marked as the green, -orange and red tomatoes of two distinct species roma and pear.	
To evaluate the tomato maturity state by means of sensing the volatile production of ripeness states by a specific electronic node devises.	The electronic nose device PEN2 fitted with ten different arrays for detection of different ripening stages of tomato.	The LDA and PCA techniques in order to assess different ripening stages of tomatoes.	The PCA and LDA techniques have been studied in order to validate the differences among different maturity stages of tomato.	[29]
To evaluate the tomato maturity state using the magnetic resonance imaging	The assessment of the density of water contents in various maturity stage of tomato for assessing the maturity stages by subjecting the information to the MRI system.	The physiological changes of tomatoes in their different maturity stages captured through the MR images.	Water portion relaxation behavior is sensitive to the local environment and the material structure and thus used	[30]
Determination of the maturity by	Raman spectroscopy of the spatially offset	The assessment of the change of	The Raman peaks corresponding to the	[31]

means of the nondestructive approach through the Raman spectroscopy.	spectra studied in the specified wavelength range for selection of maturity of the tomato fruit.	Raman peak by means of spectral information divergence(SID).Tracking of the SID values with the gradual ripening of the tomato fruit.	carotenoids hasshown up along with the shift of the peaks has been observed.	
To estimate the maturity of tomato and to predict the textural properties of tomato species momotaro by means of he near infra-red spectroscopy	Near infrared spectroscopy to assess the maturity level of momotaro. The PCA and Soft independent modeling of class analogy concept for distinguishing among different maturities	The mature green, pink and red colored tomatoes subjected to the treatment for the assessment of the textural properties of the tomato by means of the partial least square regression techniques.	The bio yield force obtained from the puncture test shows that normalization pretreated spectra lying between 1100 – 1800 nm has the highest correlation coefficient value.	[32]
Nondestructive approach towards determination of the maturity of tomato with the help of hand held visible and near infrared instrument.	The handheld spectrometer to assess the maturity of tomatoes along with the NIR based statistical model calibrated and validated to carry out the operation.	The whole fruit interactance measurement carried out on the green tomatoes	The initial concentration of chlorophyll and the ripening in tomatoes are completely variety dependent. The NIR measurement follows the same pattern as observed in other fruits.	[33]

Image processing based algorithmic approach towards the identification of six different ripening or maturity stages of tomato.	The part of matlabsimulink application to carry out the assessment for identification process of the tomato.	The application of image processing algorithm on the tomatoes of different maturity grades.	The image processing algorithm can be used for other fruits such as apples, citrus fruits and mandarin oranges. The exposure of the greater surface area could be done by fitting a motor below the platform inorder to rotate the fruit.	[34]
To classify different maturity stages of tomato by using the multi class SVM based classification approach towards the assessment of ripeness.	Conversion of the color space from the RGB to HSV followed by the PCA application to generate the feature vectors followed by the feeding of the feature set to the multi class SVM classifier for identification of proper class.	The PCA along with the SVM for feature extraction and its analysis.	The color histogram of the tomato images thus formed has been subjected to the PCA plotting.	[35]
To recognize the object as tomato fruit and estimate the ripeness of	Fuzzy rule based classification for recognition and classification of the	The feature space has partitioning based on the principles of fuzzy	Six different ripeness stages have been detected. The detection process did	[36]

the fruit in pre harvest session by using automatic rule learning based decision tree.	fruit of Red green color differences followed by the use of red green color ratio for recognition and classification purpose.	logics into the linguistic variables with the help of learning algorithm. .	not interfere in the natural growth of tomato. The experiment carried out also maintained the color integrity while segmenting tomato from the background.	
To ascertain the ripeness stages of tomato by means of chemical sensors based on the carbon nanostructure composites.	Use of the chemical sensors fabricated by depositing the thin film of composites of carbon nano tubes, carbon nano coils and polyvinyl alcohol for assessing the changes in the value of conductance and capacitance.	Appointment of different sensors for detection of maturity categories of tomatoes	The approach is primarily focused on the quantitative measurement avoiding the need of the qualitative visual inspection.	[37]
To evaluate ripeness of tomato on the basis of optical absorption and scattering properties. The model was developed in order to find the ripeness stages	The hyper spectral imaging based spatially resolved instruments to measure the optical absorption and scattering coefficients.	The decrease of the absorption peak consistently with the increase of ripeness property.	The variation of spectra obtained confirms the clear distinction between different ripening stages of tomatoes.	[38]

of tomatoes.				
To classify the banana fruits based on the concept of deep learning.	A conventional neural network CNN model called as the ConvNets.	The deep learning approach has been carried out on kaggle dataset comprising of 1914 images divided in respective training and test set .	Three different classes of banana like normal banana, banana lady finger and banana red have been subjected to the model for experimental purposes.	[39]
To determine multiple ripeness stages of banana by using the artificial neural network.	The autonomous computer vision approach to identify the ripening stages of banana along with the artificial neural network based framework for classification.	Use of the features of banana namely color, development of brown spots and tamara statistical texture for the classification procedure	The results have been compared with various techniques like SVM, naïve Bayes and KNN, decision tree and discriminant analysis classifiers	[40]
To identify the maturity stages of banana fruit by means of image processing techniques	The maturity assessment based on the mean color intensity algorithm and area algorithm.	The attributes of banana like mean color intensity from histogram, area, perimeter, major axis length, minor axis length extracted from the calibration image of the fruit.	Analysis of variance between the different maturity stages plays the crucial role in assessment of maturity of the fruit.	[41]
To predict the	Parallel plate	Use of the	Green ripe banana	[42]

quality of banana during the ripening stage of the fruit using the capacitance sensing system	capacitor with the sine wave frequency generator for quality prediction .	dielectric property to assess the quality parameters of banana fruit during the ripening period.	fruit has the greater permittivity than the full ripe ones. The relative permittivity was correlated with the quality parameters of banana.	
To assess the sugar and starch in intact banana and mango fruit by SWNIR spectroscopy.	The prediction accuracy realization by collecting the visible short wavelength near infra-red spectra using the partial transmittance optical geometry.	The distinction of between the soluble and insoluble forms of carbohydrates using the short wave near infrared spectroscopy	The banana pulp was used instead to get the better results.	[43]
To predict the banana fruit quality using factors like pigment content, starch index, sugar content which are accountable for maturity classification.	Visible and near infrared spectroscopy in reflectance mode for nondestructive detection of the food chlorophyll and sugar contents. With the application of partial least square regression applied to auto scaled spectral data.	The chromatography method for individual sugar analysis.	First the VIS/NIR range was measured followed by the application of chemical methods.	[44]
To develop the hybrid system to	Computer vision concept and the	Ten different herb species samples	The accuracy value for different	[45]

classify the herbs and to detect the herbs plant diseases.	electronic nose The hybrid intelligent system involving the fuzzy inference system, naïve Baise system, probabilistic neural network along with SVM	subjected to the experimental hybrid system comprising three different separate classifier.	classifier used was appreciable and fell in the range of 95% to 99%.	
To develop a recognition method for cucumber diseases using leaf symptom images based on deep convolution network	Deep convolution model to recognize the cucumber diseases.	The CNN architecture made up of three distinct layers for extraction of features from the segmented image set.	The AlexNet has been used and was found to have outperformed the DCNN.	[46]
To identify tomato plant diseases of leaf images using the SqueezeNet model.	The model Squeeze Net based on the CNN to recognize the tomato plant disease.	The tomato plant leaves obtained from the vegetable crops research institute in Lambang for each class.	The study detected the tomato plant diseases through the leaf images.	[47]
To develop a procedure for the recognition and classification of maize leaf diseases out of the healthy	The convolution neural networks (CNN) based model for studying the recognition and classification of maize leaf diseases.	Neuroph was used to perform the training of the CNN.	The images of diseased leaves were all collected by means of the smart phone based camera.	[48]

leaves				
To detect the leaf diseases of tomato fruit by means of image processing approach	The CNN based model to classify the leaf diseases.	Classification of the tomato leaves of five different diseases namely bacterial spot, late blight, septoria leaf spot, and tomato mosaic, yellow curved alongside the healthy leaf based on CNN.	The average performance of classification is above 95% and the system has been proved to be much user friendly also.	[49]
To identify the different diseases of tomatoes and it's associated infected areas based on the deep convolutional neural networks.	The detection of different diseases and the infected area of the tomato fruit by means of the faster R-CNN and mask R–CNN architecture.	Implementation of the ROI align in the faster R-CNN andmask R-CNN architecture in order to minimize the equalization error and segment out the lesion spots.	Comparison between the different model has been carried out as has been found that the performance in case of mobile net outweighed that of the ResNet-101.	[50]
To identify the plant leaf diseases using the CNN model	The nine layer deep CNN model to identify different diseases of plant leaves. The deep CNN has been trained by using different mini batch sizes ranging from 64 to 160 with	The process image flipping, gamma correction, noise injection, principal component analysis, color augmentation, rotation and scaling	The performance of CNN has been compared to the popular model like AlexNet, VGG16, Inception V3 and ResNet.	[51]

	augmented datasets.			
To develop the automated system based on CNN architecture in order to diagnose the rice diseases.	The CNN based model for recognizing the rice blast diseases. The performance has been evaluated to assess the equality in potentiality of CNN with softmax and CNN with support vector machine.	The size of 128 X 128pixels has been extracted from the image with the help of moving window. The CNN, LBPH and Harr wavelet features of the plant leaves have been extracted.	The performance measurement has been carried out in hybrid mode by combining the CNN with SVM, LBPH and Harr wavelet and the performance assessment has been done.	[52]
To develop the system in detecting and estimating the severity estimation for diseases of tomato green house plants. The detection mechanism has been focused onto the generic features of plant diseases.	The UNet deep learning model depending on the principles of hybrid deep learning model.	The salient feature like the leaf shape, leaf surface and background content features for the UNet	The residual network Resent-50 with 50 layered architecture has been pretrained with the plant village datasets.	[53]
Identification of tomato plant leaf diseases at the	The statistical features of the image of plant leaves	The extraction of statistical function mean, median,	The difference between the healthy tomato plant and	[54]

early stage using the segmented features.	subjected to the CNN classifier.	range and standard deviation based features	diseased plant on the basis of histogram.	
To develop the deep learning model in order to detect the plant diseases.	ResNet-50, Xception, MobileNet,ShuffleNet,DenseNet121_Xception for the purpose of feature extraction.	The transfer Learning approach for the classification purpose.	Tomato diseased Detection application has become much effective in tomato pest control.	[55]
To develop the deep learning module for leaf disease detection based on convolution neural network by Attention Module.	The multistage BNN architecture integrated into ResNet-50 model for the process of leaf disease detection.	The squeeze and excitation network integrated with CNN model for carrying out the process of detection. and comprises 10 different categories of tomato leaf.	In order to cope up with the complex characteristics of tomato diseases, the study of multi scale extraction of disease features has been adopted.	[56]

The traditional approach towards determination of species and maturity of different agro products are done manually and accordingly suffer significant limitations. The sorting and grading process of the agro products require good skill sets of the sorter. In manual system, the sorting process can be lengthened and consumes considerable time. The chemical and analytical instrumental methods towards the species and maturity assessment of different agro products suffer some number of limitation like expensive and time consuming test process, sample preparation process and invasive nature of operation as well. Moreover the aforesaid procedure also suffers from the lack of mobility i.e. difficulty in the real time or on-spot detection due to stationary presence of the instruments in the lab. The computer vision system has shown the significant potentiality towards remedial of different sort of problems addressed above in the field of agriculture due to its non-destructive and non-invasive application.

1.4. Brief overview of computer vision

Computer vision is an emerging field of research which explores the way of sensing and understanding the information from images and videos by computer. It has the two distinct perspectives; biological and engineering. From the biological perspective, it is the computational model of human visual system on the other hand from the engineering perspective it is an automated system which can mimic the perception of human being. Some elemental tools for capturing image data include either single camera or multiple cameras and the data concerned can be the single dimensional or multi dimensions. The information extracted from the image has the multiple variance like identification or classification of the image, navigation in space and augmented reality applications. Some of the salient advantages of computer vision over established instrumental and analytical techniques include comparatively less expensive implementation, non-invasive methodology and faster decision making with competitive accuracy. The potential of computer vision has been explored in diverse applications. The advancements of machine learning algorithms and hardware configurations have further strengthened the potential of computer vision based deployments. The major building blocks of computer vision system include (i) image acquisition, (ii) object/pattern detection (ii) object/pattern recognition and (iii) results in terms of classification and prediction.

- ***Image acquisition:*** in this first step of computer vision framework the images are captured. The images can be single frame or multi frame (video capture). The images can also vary depending on the type of illumination and imaging conditions. There are applications where instead of visual images, feature images like thermal images are captured.

- ***Object/pattern detection:*** Once the images are captured the computer vision systems learns the inherent and distinguishable pattern in the images. For example, in case of object recognition the boundary of the object can be an important pattern that can distinguish between different types of objects present in an image. Similarly, colors and textures are also commonly used patterns that can distinguish between different objects. Such characteristics are commonly referred as features and the choice of features that can distinguish are entirely application dependent.
- ***Object recognition*:** Once the features are identified and extracted in the previous step they are subjected to the machine learning algorithms. These algorithms are mathematically developed and having the potential to learn from the examples. The degree to which the algorithm can learn varies widely but nevertheless by virtue of such learning mechanism all such algorithms can identify or detect any unknown sample of the same nature. One thing is important to note that the modern algorithms are capable of drawing soft and flexible boundaries between different classes unlike the hard and rigid boundaries obtained by many of the classical algorithms.
- ***Decision making:*** This is the stage where the computer vision declares its output against unknown test cases. There are two major types of declarations normally practices; classification and prediction. In former case the output is in terms of decided class for the subjected samples among many classes under consideration while in later case it is normally a predicted value which is often referred as regression. In many of the applications both of the applications are considered. For instance, in case of any agro industry application the classification task may be to identify the class of a particular agro produce and prediction task may be the prediction of how many days it can be stored prior to its rotten stage.

The widespread applications of computer vision can be consolidated as follows;

Special effect in the movie: Computer vision has been uniquely applied in many movies, especially the cartoon and animation movies. The movements of human actor have been simulated as digital actors. The exact positions of markers are accessed using the concept of computer vision.

Scene recognition: Photographs of any scene can be easily detected and recognized using this concept. It is widely used for digital surveillance and image-based searching over internet.

Digital biometrics: Computer vision approach can successfully be applied in detecting the face, expression and gesture recognition. This is widely used for biometrics where fingerprint and iris detections are also accomplished using computer vision.

Optical character recognition: Computer vision can be successfully used in recognizing the characters and numbers of many different patterns. This is widely used in document image processing and old document restorations.

Self-driving: Automation in car driving has been implemented by using the concept of computer vision. In many such applications this is also termed as robotic vision.

Augmented reality: This is another growing field of applications majorly used for computer and mobile gaming. This provides much reality in user experience which in turn increases the user engagement.

Virtual reality: In this area the position of a user and the position of all the objects around are getting connected to each other's and change in some realistic way.

Financial services: Computer vision principles have majorly been applied in the field of financial services. The banking and insurance sector have been enormously benefitted due to progress of computer vision. The fast banking services are provided to the customers by uploading the photo-id cards. The uploaded documents can easily be analyzed.

Health care: Computer vision can play crucial roles in diagnosis of diseases by means of machine learning and can reduce the clinical cost for economically growing countries.

Preventive maintenance in industry: The monitoring system of the industries has been based on the computer vision principles for producing reports on the status of infrastructure and processes. The alert systems for predictive maintenance are also developed using computer vision.

Social media: The enormous popularity of social media has become possible due to the advancement of research in the computer vision. Automatic tagging of user images and embedding of computer vision principles in the lens are results of advanced computer vision approaches.

Machine learning plays pivotal roles in computer vision. It brings the essence of human psychology and reasoning capability in computing. Table 1.2 presents some of the major works using computer vision in diverse research fields.

Table 1.2: *Literature review for applications of computer vision*

Applications	**Working principle**	**Data analysis technique**	**Remarks**	**Ref**
Preparing the authentic facial expression. Database based on the	The facial expression classification on the basis of machine	The model based face tracking algorithm	The two authentic databases namely the emotion database and the Cohen Kanade database has	[57]

spontaneous emotions available in the environment.	learning techniques.		been used for the purpose.	
To develop the machine learning algorithm for convolution feature selection.	The classifier CNN for feature selection.	A novel adaptive weights objective function approach for evaluation and selection of features	The optimization efficiency has been improved by introducing the quadratic programming methods.	[58]
To highlight the interaction between artificial intelligence, computer vision and machine learning.	Statistical, logical, conceptual, connectionists, neural network, genetic algorithms or evolutionary algorithms.	Some important strategies inductive deductive, analogical etc.	The current research paradigm has been based on the machine learning and computer vision.	[59]
To classify the Image Net with the help of deep convolution neural network	The neural network having 60 million parameters and 650,000 neurons	The Image Net based network		[60]

	consists of five convolution layers.			
To extract the multi text from the scene.	The extraction of meaningful group of regions.	The text group hypothesis by means of the proximity and similarity laws.	The datasets encompasses the texts in different orientations in two different languages.	[61]
To develop the object and action aware model for visual language navigation.	Object and action aware model to make the object centered and action centered instructions for the sake of visual perception.	The effective path loss to penalize the trajectories from the ground truth path.	It turns the general natural language instructions into robot agent actions	[62]
	CNN model architecture consisting of the convolution , pooling and fully connected layers has been used for	The CNN model as a classifier trained with the concerned known data sets to be trained and	The CNN model has been subjected to the regularization and optimization process in order to get the best accuracy in the result obtained.	[63]

	carrying out the classification task.	the trained model is used for unknown data.		
To highlight different CNN architectures and their application perspectives.	CNN is a class of DNN and consisting of the layers convolution, maxpooling and the fully connected.	The training tales place inthe CNN model followed by the classificatio n of test data.	Different variation of CNN are ResNet, VGGNET, LexNetetc	[64]
To highlight the significant role of artificial intelligence in the field of computer vision.	The statistical techniques and generic machine learning algorithms in different models like ANN, CNN etc.	The model ANN or CNN trained with the conducive datasets.	The models have been fed with the one and three dimensional data for training purpose.	[65]

1.5. Problem statement

The literature review conveys that determination of species and maturity is an important research field where different analytical and instrumental methods are mostly adopted. Despite appreciable accuracy such methods have drawbacks like expensiveness, lengthy and precision sample making process, invasive and destructive testing method and portability issues that limit their realization in different locations in the supply chain for on-the-spot detection. The sample making for obtaining the result is very much time taking and so for large scale assessment of maturity of agro products, these methods fail to meet the time constraints. The chemical reagents used for the assessment purpose are costly, too. These motivates towards development of noninvasive, non-destructive and chemical free approaches for food quality assessment. Literature review also reveals that computer vision frameworks have shown potential all such limitations associated with conventional methods. Machine learning is the effective means of tackling the problem of classification for different agro products between different species and maturity stages.

Among different machine learning algorithms deep learning algorithm is most emerging one. Convolutional new network (CNN) is one of most prominent deep learning mechanism in the field of computer vision. There are different reasons of its popularity. One major advantage is that it eliminates separate feature extraction step as it learns the high level features itself from the images through incremental steps.Often the comparatively high training time of CNN is projected as a drawback but once trained it can provide very fast decision with high accuracy. Earlier deep learning algorithms were found suitable for cases where large amount of data was present. But, some of the latest developments in learning mechanism have made application of CNN possible for small datasets with low variance. However, as the dataset size increases the performance potential also can increase. Another big advantage of CNN and other deep learning algorithms is its ability to generalize the prediction model to avoid over- or under-fitting problems often encountered with conventional algorithms for small dataset.

CNN architecture is based on the machine learning dynamics and exploits many layers of nonlinear information processing. It consists of convolution, pooling and fully connected layers. The prime benefit of CNN is that it does not need the hand engineered feature extraction. The cascaded model of CNN is effective towards the faster detection. Multi-channel CNN has been implemented by fitting two CNN models in parallel. CNN hyper parameter optimization is also another important boosting task towards achieving the greater accuracy of classification. The deep learning concept has come up in many variations like deep learning neural network(DNN),deep belief network(DBN),recurrent neural network(RNN) and recursive neural tensor networks(RNTN).Use of such machine learning based approach fully annulled the chances of damage of the leaves or plants on application of chemicals.

The detailed literature survey is presented in Table 1.1 and 1.2 confirms the possibility of applying machine learning particularly CNN driven computer vision framework for quality grading and detection/identification tasks in agro produces. The reviews can also be consolidated into following major categories of research directions.

a) Determination of different maturity stages of medicinal leaves.
b) Determination of different diseases of leaves of different agro plants.
c) Determination of different maturity stages of fruits and vegetables.
d) Application of deep learning concepts in the classification of different agro products.
e) Application of deep learning concepts in the classification of different diseases.

1.6. Objective and scope of the thesis

The medicinal plant leaves neem, tulsi and kalmegh , the fruits tomato and banana bears lot of medicinal qualities along with their food values. However, the bioactive present in those agro products vary with the maturity and also across different species. Similarly the diseases of plant leaves are also very threatening to the plant growth and agricultural economy. The lack of timely detection of diseases can lead to the damage and even death of the plant. In order to assess proper maturity stage of the plant leaves of different fruits or vegetables or to detect different sort of common plant diseases, many different invasive and destructive or chemical analyses have been carried out. In this thesis compute vision approach for classifications of different maturity stages of agro products and early detection of different diseases for the plants have been explored. The focus in machine learning aspect has been remained with CNN since that can provide considerable accuracy while reducing the separate feature selection step with conventional machine learning algorithms like support vector machine (SVM), random forest (RF) and linear or logical regression. The possibility of realizing the adopted CNN models for a hand held or mobile device development towards on-spot detection and prediction has also been focused while optimizing the CNN architectures. For building the generalized platform, all the external conditions and constraints have been taken into consideration and the associated errors arising from the samples have been addressed. The summary of the objectives have been represented below.

a) The detailed and extensive literature survey for finding the common medicinal plants and other agro products with high impacts on the human health.
b) To find different plants those are more susceptible to different diseases.
c) To apply the preprocessing techniques for removing noises in order to extract the enough analytical information from the samples.
d) To adopt the optimum deep learning model for classification and prediction.
e) To find the appropriate optimizing algorithm in order to increase the model performance.
f) To evaluate the models against standard metrics to assess the robustness.

1.7. Research questions

Considering the scope and objectives of the research topic in the thesis here, the following research questions have been framed.

RQ1. Can the deep learning based model be used to detect the maturity of medicinal leaves, fruits and vegetables?

RQ2. Can the deep learning based model be used to classify different diseases of the plant?

RQ3. Is it possible to develop any mathematical model from the computer vision data to predict the maturity stage or disease type for any medicinal plant leaves, fruits or vegetables?

RQ4. Can optimization of the deep learning model hyperparameters contribute towards performance improvement?

RQ5. Can the deep learning system be a suitable means to aid the existing instrumental and analytical methods of detection and prediction tasks in agro industry?

1.8. Thesis structure

In this thesis, the images of leaves from three medicinal plants namely, neem, tulsi and kalmegh and the fruits banana and tomato have been used in search of the possible answers to the research questions. As per the literature review it has been observed that the demands of getting the maximum medicinal benefits out of the leaves or fruits and detection of plant diseases in proper time is of utmost interest from the perspectives of economic benefits. There are several methods available for detection of maturity of agro products or diseases of different plants which are very costly and tedious to perform therefore the alternative method of using the deep learning approach has been proposed here. The deep learning model based on the modified CNN optimized with the appropriate optimizing algorithm has been attempted in this thesis work. In the perspectives of the above mentioned objectives, the entire work has been divided into six chapters as outlined below.

Chapter 1 introduces the scope of the thesis describing the origin of the problem and outlined the objectives of the thesis work. This elaborates the significance of assessment of maturity of medicinal plants and fruits along with classification of different diseases for plants. The existing methodologies for assessment of maturity and diseases have been discussed. The chapter also highlights the related literature survey. The chapter finally summarizes the scope, objectives and research questions of the thesis.

Chapter 2 provides the overview of medicinal plant and its significance in production of different medicines. It also highlights the experimental set up for sample collection, dataset preparation and model development. The optimization of the model has been carried out and has been presented analytically in this chapter. The associated results and discussion for assessment of maturity of leaves have also been included in this chapter. The chapter thus elucidates the scope of computer vision framework for identification of species and prediction of associated maturity stage that can contribute towards use of medicinal leaves at correct stage for preparing medicinal products.

Chapter 3 elucidates the species and maturity grading of vegetable. The experimental setups along with the process of sample collection, dataset preparation and model development have been elaborated. The experimental result with the familiar vegetable tomato has been included into this chapter. The result of experiment has been shown to be improved to higher grade with the application of hyperparameter optimization. The result of the experiment vouches the possible potentiality of the computer vision framework towards the identification and prediction tasks concerned to the species and maturity stages of vegetables.

Chapter 4 describes the significance of determination of species and the maturity grading of fruits. The experimental setups along with the process of sample collection, dataset preparation and model development have been highlighted. The unique feature of the model with separate channels for respective task of species and maturity determination has been illustrated. The experimental results with its associated metrics for popular fruit namely banana has been included in this chapter. The results of this chapter confirm the possibility of the computer vision framework for identification and prediction tasks related to species and maturity stages of fruits.

Chapter 5 elaborates the significance of early detection of the leaf diseases. This chapter also illustrates details of experimental setup including the sample collection, dataset preparation and model development. The Mobile net model and its corresponding optimization using modified GWO has been expressed in this chapter with the improvement reflected in the overall work. The associated result discussion has also been highlighted with the help of different metrics. The promising results of this chapter show that computer vision can be a potential proposition to avoid huge agricultural loss by means of early detection of plant diseases.

Chapter 6 summarizes the major findings of the thesis. The chapter also includes the limitations identified during the experiments under the presented scopes. The possible future extensions of the presented works have as well been presented in this concluding chapter.

1.9. Conclusion

This chapter provides the general overview of the thesis. It shows the importance of maturity and species detection task for agro produces by means of considerable literature survey. It also elaborates the potential of advanced machine learning driven computer vision systems to perform several engineering problems while addressing the limitations like invasive nature of operations, high instrument cost and long processing time of the conventional techniques. The diverse applications of computer vision have as well been presented in terms of literature survey. The motivation of exploring the scope of advanced machine learning driven computer vision frameworks for detection and prediction tasks in agro produces has been clearly elaborated. Objectives and scope of the thesis has been drawn in view of the background survey. The associated research questions have been clearly mentioned which is to be addressed in the conclusion section of the thesis. Lastly the thesis structure has been described with the flow diagram depicting various sectional topics of the thesis for a bird-eye view of the presented works.

References

1) Kumkar S., Dobos G.J., Rampp, T. (2017).The significance of ayurvedic medicinal plants. J Evid Based Complement Altern Med 22(3):494–501.
2) Verma H .et.al.(2019).Evaluation of an emerging medicinal crop Kalmegh [Andrographispaniculata (Burm. F.)Wall Ex. Nees] for commercial cultivation and pharmaceutical & industrial uses: a review. J PharmacognPhytochem 8(4):835–838.
3) Upadhyay R.K. (2017).Tulsi: a holy plant with high medicinal and therapeutic value. Int J Green Pharm (Suppl) 11(1):S1–S12.
4) Alzohairy M. (2016).Therapeutics role of Azadirachtaindica (Neem) and their active constituents in diseases prevention and treatment. Evidence-Based Complement Altern Med. https://doi.org/ 0.1155/2016/7382506.

5) Fatma M.,Ahmed N.(2019). Ripeness Classification of Bananas Using an Artificial Neural Network', Arabian Journal for Science and Engineering, King Fahd University of Petroleum &Minerals , Vol. 44, No. 8.

6) Bhowmik D., Kumar K.S., Paswan S., Srivastava, S. (2012).Tomato-a natural medicine and its health benefits. J. Pharmacogn. Phytochem.1(1), 33–43.

7) Saad A.M., Ibrahim A., El-Bialee N. (2016).Internal quality assessment of tomato fruits using image color analysis. Agric. Eng. Int.: CIGR J. 18, 339–352.

8) Suarez M., Rodríguez R., Elena M., Díaz C. (2008).Chemical composition of tomato (Lycopersiconesculentum) from Tenerife, the Canary Island. Food Chem., 106, 1046-1056.

9) Isleib J. (2018).Signs and symptoms of plant disease: Is if fungal, viral or bacterial? Retrieved from https://www.canr.msu.edu/news/signs_and_symptoms_of_plant_disease_is_it_fungal_viral_or_bacterial.

10) Rafiq A., Makroo H., Sachdev P., Sharma S. (2013). Application of Computer Vision System in Food Processing- A Review. International Journal of Engineering Research and Applications. 3. 1197-1205.

11) Faridi H., Aboonajmi M. (2017). Application of machine vision in agricultural products. In: Proceedings 4th iranian international NDT conference, Olympic Hotel, Tehran, Iran, Feb 26–27.

12) Chaki J., Parekh R., BhattacharyaS. (2015).Recognition of whole and deformed plant leaves using statistical shape features and neuro fuzzy classifier. In: IEEE Proceedings of 2nd international conference on recent trends in information system (ReTIS).

13) Keivani M., Mazloum J., Sedaghatfar E., Tavakoli M.B. (2020) .Automated analysis of leaf shape, texture, and color features for plant classification. Traitement du Signal 37(1):17–28.https://doi.org/10.18280/ts.370103.

14) Mahajan S., Raina A., Gao X-Z, Pandit K. (2021).A plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost. Symmetry 13:356. https://doi.org/10.3390/ sym13020356.

15) Santra A.K., Christy C.J.(2012).Genetic algorithm and confusion matrix for document clustering. Int J ComputSci 9(1):322–328.

16) Upadhyay R..K. (2017).Tulsi: a holy plant with high medicinal and therapeutic value.Int J Green Pharm (Suppl) 11(1):S1–S12.

17) Verma H., Negi M.S., Mahapatra B.S., PaulJ.(2019). Evaluation of an emerging medicinal crop Kalmegh [Andrographispaniculata (Burm. F.)Wall Ex. Nees] for commercial cultivation and pharmaceutical & industrial uses: a review. J PharmacognPhytochem 8(4):835–838.

18) Zhang Y., Peng J., Yuan X..et.al.(2021).MFCIS: An automatic leaf-based identification pipeline for plant cultivars using deep learning and persistent homology.Hortic Res8, 172 . https://doi.org/10.1038/s41438-021-00608-w.

19) Sorin C., Mariette F., Musse M., Leport L., CruzF.,et.al.(2018). Leaf development monitoring and early detection of water deficiency by low field nuclear magnetic resonance relaxation in NicotianatabacumPlants. Applied Sciences, MDPI, 8 (6), pp.943. 10.3390/app8060943. hal-01830627.

20) Kumar S.(2012). Leaf color area and edge features based approach for identification of Indian medicinal plants. Indian Journal of Computer Science and Engineering (IJCSE),Vol. 3 No.3 Jun-Jul.

21) Azlah M.A.F.,ChuaL.S., Rahmad F.R. , Abdullah F.I. , AlwiS.R. (2019).Review on Techniques for Plant Leaf Classification and Recognition, Computers , 8, 77.

22) Sangle S., ShirsatK., Bhosle V. (2013). Shape Based Plant Leaf Classification System Using Android, Int. J. of Engineering Research and Technology. 2(8), August.

23) Jeon W., RheeS.. (2017). Plant Leaf Recognition Using a Convolution Neural Network. IJFIS ;17:26-34. https://doi.org/10.5391/IJFIS.2017.17.1.26.

24) Saad A.M., Ibrahim A., El-Bialee N. (2016).Internal quality assessment of tomato fruits using image color analysis. Agric. Eng. Int.: CIGR J. 18, 339–352.

25) Van de Poel B. et.al. (2012).Model-based classification of tomato fruit development and ripening related to physiological maturity. Postharvest Biol. Technol. 67, 59–67. http://dx.doi.org/10.1016/j.postharvbio.2011.12.005.

26) Satpute M.R., JagdaleS.M.(2016).Color, size, volume, shape and texture feature extraction techniques for fruits: a review. Int. Res. J. Eng. Technol. 3, 703–708.

27) Bhatt A.K.,Pant D.(2015).Automatic apple grading model development based on backpropagation neural network and machine vision, and its performance evaluation.AISoc. 30, 45–56. http://dx.doi.org/10.1007/s00146-013-0516-5.

28) Wang P., Toudeshki A., Tan H., Ehsani R. (2018), A methodology for fresh tomato maturity detection using computer vision, Computers and Electronics in Agriculture 146 , 43–50.

29) Gómez A.H., Hu G., Wang J., PereiraA.G. (2006).Evaluation of tomato maturity by electronic nose.Comput Electron Agric 54(1),44–52.

30) Zhang L., McCarthyM.J.(2012).Measurement and evaluation of tomato maturity using magnetic resonance imaging. Postharvest BiolTechnol 67,37–43.

31) Qin J., Chao K., Kim M.S. (2012).Nondestructive evaluation of internal maturity of tomatoes using spatially offset Raman spectroscopy. Postharvest BiolTechnol 71, 21–31.

32) Sirisomboon P., Tanaka M., Kojima T., Williams P. (2012).Nondestructive estimation of maturity and textural properties on tomato ‘Momotaro’ by near infrared spectroscopy. J Food Eng 112(3),218–226.

33) Tiwari G., Slaughter D.C., Cantwell M. (2013).Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. Postharvest BiolTechnol86 , 221–229.

34) Rupanagudi S.R., Ranjani B.S., Nagaraj P., Bhat, V.G. (2014).A cost effective tomato maturity grading system using image processing for farmers. In: Contemporary Computing and Informatics (IC3I), International Conference on, IEEE, 7–12.

35) Lhariri E. et.al. (2014).Multi-class SVM based classification approach for tomato ripeness. In: Innovations in Bio-inspired Computing and Applications, Springer International Publishing, 175–186.

36) Goel N., Sehgal P. (2015).Fuzzy classification of pre-harvest tomatoes for ripeness estimation—an approach based on automatic rule learning using decision tree. Appl. Soft Comput. 36, 45–56.

37) Greenshields M.W.C.C. et. al. (2015).Non-destructive tomato ripeness determination using carbon nanostructure based chemical sensors: Application of tristimulus and principal component analysis methods. International Journal Electroactive Materials 3:10–14.

38) Zhu Q., He C., Lu R., Mendoza F., Cen H. (2015).Ripeness evaluation of ‘Sun Bright’ tomato using optical absorption and scattering properties. Postharvest BiolTechnol 103, 27–34.

39) Al-Daour A.F., AI Shawwa M.O., AbuNasir S.S. (2019).Banana Classification Using Deep Learning, International Journal of Academic Information Systems Research (IJAISR),ISSN: 2643-9026 ,Vol. 3, No. 12, pp. 6-11.

40) Mazen F.M.A., NashatA.A..(2019).Ripeness Classification of Bananas Using an Artificial Neural Network, Arabian Journal for Science and Engineering, King Fahd University of Petroleum & Minerals , Vol. 44, No. 8.

41) Surya P.D., SatheeshK,J. (2015).Assessment of banana fruit maturity by image processing technique,J.FoodSci.Technol,Vol.52,No.3,pp.1316–1327.[online] https://doi.org/10.1007/s13197-013-1188-3.

42) Soltani M., Alimardani R., Omid M. (2010).Prediction of banana quality during ripening stage using capacitance sensing system, Australian J. Crop Sci.,Vol. 4 ,No. 6,pp. 443–447.

43) Subedi P.P., Walsh K.B. (2011).Assessment of sugar and starch in intact banana and mango fruit by swnirspectroscopy ,Postharvest Biol. Technol, Vol. 62, No.3,pp. 238–245.

44) Zude M. (2003).Non-destructive prediction of banana fruit quality using VIS/NIR spectroscopy. Fruits 58, 135–142.

45) Mustafa M.S., Husin Z. , Tan W.K., Mavi M,F., Farook R.S.M.(2020),Development of automated hybrid intelligent system for herbs plant classification and early herbs plant disease detection. Neural Computing and Applications, 1-23.

46) Ma J., Du K., Zheng F., Zhang L. , Gong Z., Sun Z.(2018).A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network, Computers and Electronics in Agriculture, vol. 154, pp. 18–24.

47) Hidayatuloh A.,Nursalman M. , Nugraha E. (2018).Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model, International Conference on Information Technology Systems and Innovation (ICITSI), pp. 199-204, doi: 10.1109/ICITSI.2018.8696087.

48) Sibiya M.,Sumbwanyambe M. (2019).A Computational Procedure for the Recognition and Classification of Maize Leaf Diseases Out of Healthy Leaves Using Convolutional NeuralNetworks. AgriEngineering. 1:119 131.doi: 10.3390/agriengineering1010009.

49) Mim T.T., Sheikh M.H., Shampa R.A. , Reza M.S., Islam M.S.(2019).Leaves Diseases Detection of Tomato Using Image Processing, 8th International Conference System Modeling and Advancement in Research Trends (SMART), pp. 244-249, doi: 10.1109/SMART46866.2019.9117437.

50) Wang Q., Qi F., Sun M., Qu J., Xue J. (2019). Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and

ObjectDetectionTechniques, ComputationalIntelligenceandNeuroscience, vol. ArticleD 9142753, 15 pages, https://doi.org/10.1155/2019/9142753

51) Geetharamani R., Arun Pandian J. (2019).Identification of plant leaf diseases using a nine-layer deep convolutional neural network, Computers & Electrical Engineering, vol. 76, pp. 323–338.

52) Liang W.J., Zhang H. , Zhang G.F., Cao H.X. (2019).Rice Blast Disease Recognition Using a Deep Convolutional Neural Network. Sci Rep 9, 2869. https://doi.org/10.1038/s41598-019-38966-0.

53) Wspanialy P., Moussa M.A.(2020).A detection and severity estimation system for generic diseasesof tomato greenhouse plants. Comput.Electron. Agric. 178 doi: 10.1016/j.compag.105701.

54) Chahande M., DubeyA.K. , Mishra V., Mohanty M.N. (2020).Analysis of Segmented Features For Identification of Tomato Plant Leaf Diseases At Early Stage . International Journal of Advanced Science and Technology, 29(8s), 2730 - 2738. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/15198.

55) Hong H., Lin J., Huang F.(2020).Tomato Disease Detection and Classification by Deep Learning," International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), pp. 25-29, doi: 10.1109/ICBAIE49996.2020.00012.

56) Zhao S., Peng Y., Liu J.,Wu S.(2021).Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module. Agriculture , 11, 651. https://doi.org/10.3390/agriculture

57) Sebe N., Lew M.S.,Yafei S., Cohen I., Gevers T., Huang,T.S.(2007).Authentic facial expression analysis. Image and Vision Computing, 25(12): p. 1856-1863.

58) Zhang F.,Li W., Zhang Y., Feng Z.(2018).Data Driven Feature Selection for Machine Learning Algorithms in Computer Vision. IEEE Internet of Things Journal, 5(6): p. 4262-4272.

59) Esposito F.,Malerba M.(2001).Machine learning in computer vision. Applied Artificial Intelligence, 15(8): p. 693-705.

60) Krizhevsky A., Sutskever I., Hinton G.E.(2012).ImageNet Classification with Deep Convolutional Neural Networks. NIPS’12 Proc 25th IntConf Neural Inf Process Syst 1:1097–1105.

61) Gomez L., Karatzas D.(2013).Multi-script Text Extraction from Natural Scenes. In ICDAR, 2013.

62) Qi Y. , Pan Z. , Zhang S. ,Hengel A. ,Wu Q.(2020),Object-and-Action Aware Model for Visual Language Navigation. arXiv 2020, arXiv:2007.14626.

63) Khan S.(2018).A Guide to Convolutional Neural Networks for Computer Vision. Morgan & Claypool.

64) Upadhyay Y. (2019).Computer Vision: A Study on Different CNN Architectures and Their Applications.Medium,Alumn AI Academy, 29 Mar., medium.com/alumnaiacademy/introduction-to-computer-vision-4fc2a2ba9dc.

65) Karn A.(2021).ARTIFICIALINTELLIGENCE IN COMPUTER VISION, International Journal of Engineering Applied Sciences and Technology, 2021 Vol. 6, Issue 1, ISSN No. 2455-2143, Pages 249-254

CHAPTER 2

DETECTION OF SPECIES AND MATURITY FOR MEDICINAL PLANT LEAVES

2.1 Introduction

Ayurveda [1] has long been established as a popular field of alternative medicine in India and other countries in Asia. The healing power of the medicinal plants towards diseases like liver disease, skin diseases, digestion disorder, blood pressure and blood purification are few to mention among wide variety of lifestyle and chronic diseases. Leaf of the plants like Neem (*Azadirachtaindica*), Tulsi (*Ocimumtenuiflorum*) and Kalmegh (*Andrographispaniculata*) are used largely to cure up many such critical diseases. The biochemical constituents present in those leaves make it more efficient towards the remedies of many critical and chronic diseases. The amount of bio-actives present in the leaves vary with maturity stage of leaves. In order to get the significant medicinal effectiveness the medicinal plants requires the picking up of leaves at the correct maturity stages and sorting of leaves in terms of the species and maturity level.Commonly the maturity of the leaves is ascertained by the human experts by virtue of knowledge and experience.

The work presented in this chapter has been published as:

Mukherjee G., Tudu B., Chatterjee A. A convolutional neural network-driven computer vision system toward identification of species and maturity stage of medicinal leaves: case studies with Neem, Tulsi and Kalmegh leaves, (2021) Soft Computing, 25 (22), pp. 14119-14138.DOI: 10.1007/s00500-021-06139-9.

Many conventional instrumental and analytical methods have also been adopted for accurate detection of maturity stage and bio-chemical components in these leaves. Techniques like gas chromatography (GC) and high performance liquid chromatography (HPLC) are popular in this context. The efficiency of such methods is excellent but limited in terms of expensive instruments, precision sample preparation process, highly skilled manpower for operation and interpretation and mobility or portability of the instruments. The nature of testing is also invasive and destructive which in turn limits their usage in the automated process chain. Near infra-red (NIR) can address the portability issue but room for finding alternative solution considering real-time or on-the-spot detection is still open. The main objective of using the computer vision paradigm is to make it portable so that it could be used on-the-spot and also be realized in real life mobile app or a portable handheld device at the consumer or retailer end in the market. The non-invasive and non-destructive operations are the key merits of the computer vision based frameworks [2-3]. Some of the works utilizing various techniques in this direction have been consolidated in the Table 2.1.

Table 2.1: *Some of the reported techniques for detection of species and maturity for medicinal plant leaves.*

Base technique	**Method**	**Result**
The CNN model classifier for automatic classifications species of plants [4]	Custom built CNN model has been built up to classify the publicly available Flavia and Foliage datasets.	The precision accuracy of 98.69% and 98.75% has been achieved.
The computer vision backeddeep learning based approach [5]	The analysis of generated heatmaps on the image of leaf for recognition of leaves process.	The accuracy value varies 75% to 90%
Automatic extraction of the shape features and its analy-	The neuro fuzzy classification of the shape features	The accuracy value of 97.5% has been

sis through classifiers [6]		achieved.
Identification of plant leaf images through the GIST and Local Binary Pattern[7]	The pbest guided binary particle swarm optimisation methodology for reduction of features.	The highest value of the classification accuracy achieved was 96.85%
The leaf biometrics based plant identification model [8]	The particle swarm optimization as a pre-processing phases and grey wolf optimizer based leaf texture reduction.	The identification rateof 98.9% has been achieved.
The features like colour, vein ,fourier descriptor ,GrayLevel Co-occurrence Matrixetc have been adopted towards the identification of plants [9]	The extreme learning methods have been applied to learn the individual as well as the combined features of the leaf.	The evaluated accuracy of 99.10% have been achieved.
Classification ofMorphological features of plantwith the SVM along with the adaptive boosting techniques.[10]	Classification of the morphological features consisting of centroid, major axis length, minor axis length, solidity , perimeter and orientation using KNN, decision trees , multilayer perceptron	The precision rate of 95.85% has been achieved.

The neural network based approach has been proved to be effective in ascertaining the species and maturities of different medicinal plant leaves. The neural network model is made up of number of connected layers. The model is trained to adjust the connection weights as per the desired outputs. The main shortcoming of this popular soft computing model is that it requires a large number of different distinctive characteristic features to carry out the differentiation among the different medicinal leaves. Feature extraction out of the leaves is very much tedious, time taking and demands very close and meticulous observation of the leaves. These feature extraction part can be simplified by adopting the deep learning method CNN [11]. Different variations of deep learning[12] methods are deep belief network(DNN)[13],Convolution neural network(CNN)[14],Recurrent neural network(RNN)[15] and restricted recurrent neural tensor network(RRNTN)[16]Another very significant application area of the CNN is the recognition of plant leaf [17][18][19][20].

2.2. Computer vision framework for medicinal leaf species and maturity stage detection

As stated in previous chapter the computer vision framework has three major steps; image acquisition, image feature extraction and classification or prediction. Each of these steps as used for medicinal plant leaves have been detailed in following sections.

2.2.1. Sample collection and UV-vis. spectroscopy

Three popular medicinal plants, namely, neem, tulsi and kalmegh were considered for this work due to their ready availability in local market and domestic gardens. In terms of medicinal effectiveness these plants play pivotal role as well. The leaves were collected in the same season so that the changes due to climatic change can be avoided. Three maturity stages; labeled as premature, mature and over-mature were considered. It can be also noted that the samples of varying maturity stages were collected from the same tree. Initially, as a proof of fact that the important bio-chemical contents change with maturity an UV-vis. spectroscopy was performed in-house. The result of that is shown in Fig.2.1. For this experiment the leaf extracts were prepared using distilled water for proper functioning of spectroscopy. The characteristic plots of spectroscopy clearly show that the peaks of important bio-chemical considerably change with different maturity stage in the same species of leaves. The absorption spectra obtained were plotted against the wavelength and the peaks were obtained for Chlorophyll a, Chlorophyll b and Carotenoids occurred between the 400 and 700 nm. It was clearly visible that the peaks were varying with the corresponding maturity which in return indicates that the variation in the amount of bio-actives is attributable to the variation in maturity of plant.

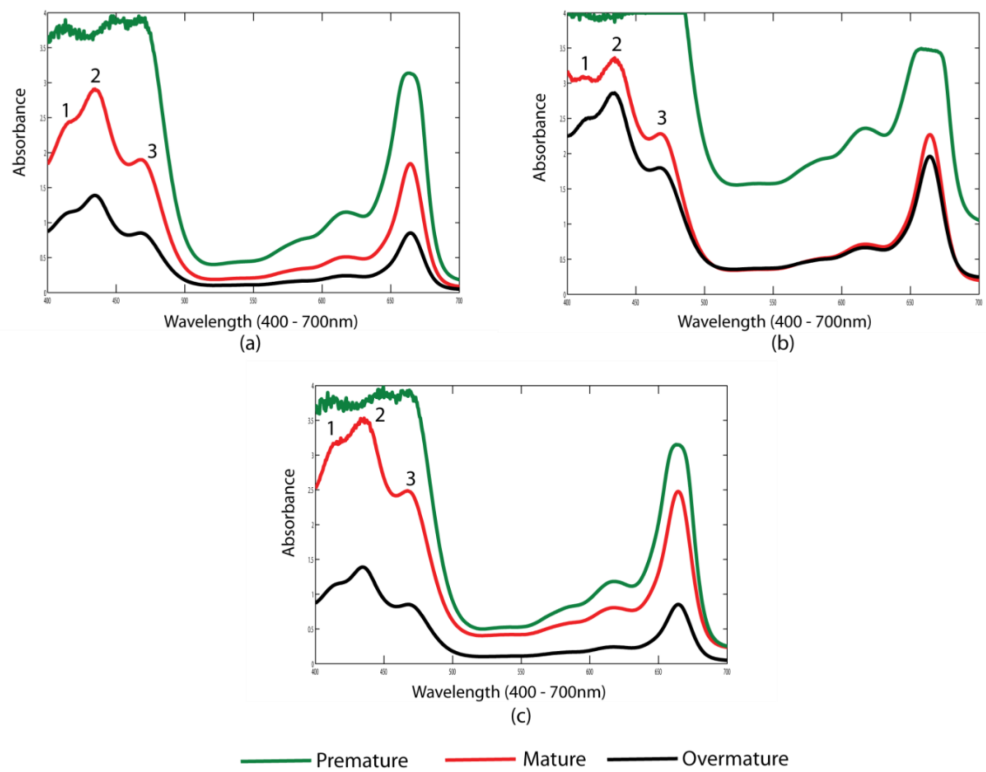


Fig.2.1: *UV-Vis spectroscopy for the (a) kalmegh (b) Neem and (c) Tulsi leaves. The labeling of peaks as 1, 2 and 3 in the graph indicate the concentrations of Chlorophyll a, Chlorophyll b and Carotenoids.*

Total 1500 leaves for each type of plant species were collected with 500 leaves for each maturity type. The maturity classes have been broadly classified into three categories named as premature, mature and over mature. Since the life span of the plants taken in this work largely varies, so the age concerned for grading of the plants is varying with the particular plant. The initial gradation of sample leaves was done by the human with expertise in Ayuervedic practice. The samples were carried in controlled temperature from the plucking point to the image acquisition point and subjected to the imaging chamber on the same day.

2.2.2. Imaging chamber

Illumination plays an important role for intensity images. Therefore an imaging chamber was developed in-house to control the illumination during image acquisition step. The photographs of actual chamber as used for this as well as all the other studies presented in this thesis work is shown in Fig. 2.2. The chamber casing was made of wood with the dimension of $0.5\,ft \times 0.5\,ft \times 1\,ft$. White led light were used for the illumination purpose and strips of LEDs were mounted at the four corners of the chamber. The adjustable base was made to adjust the distance between the camera lens and the object under consideration in order to obtain the desired focus. The operating current and the voltage for the chamber were 12V DC and 20 mA.The images were captured in JPEG format for its wide familiarity across different devices and mobile phones in spite of its lossy nature. The images were taken by using the camera SONY DSC –W30 and attached to the computer with the connector USB2.0. No camera filters, zooming options and external light sources like flash light of the camera were used during image acquisition. The images of actual imaging chamber are shown in Fig. 2.2. An in-house dataset of medicinal leaves with neem, tulsi and kalmegh leaves for different maturity stages was prepared for model development and experimentations.

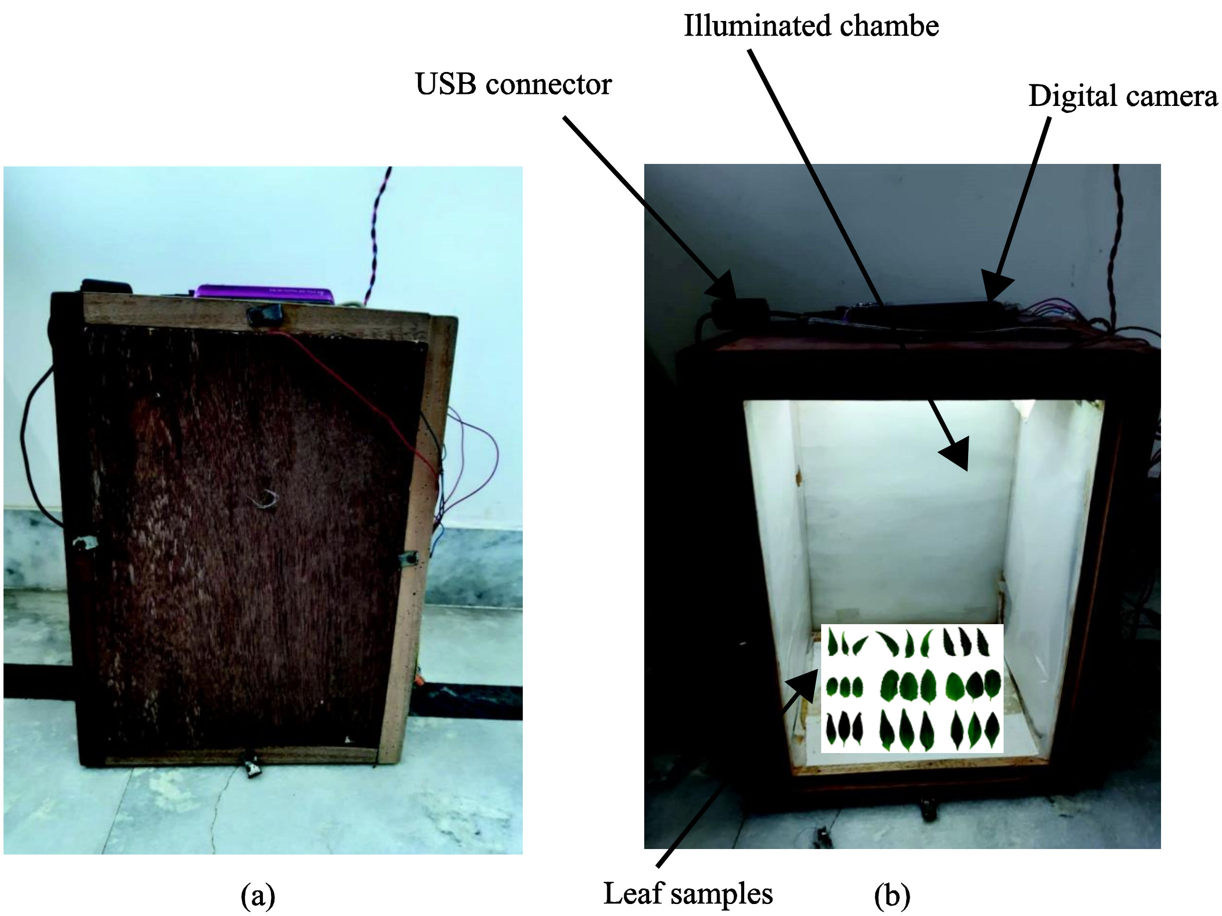


Fig. 2.2. The imaging chamber i) imaging mode and ii) loading mode

2.2.3 Leaf isolation

The images of leaves of any particular plant type have been captured inside the photographic chamber with the help of digital camera kept at a fixed height from the base supporting the leaf sample. Since the leaves of the plants under consideration were very tiny in size and large in number it was difficult to capture individual leaves. Hence an leaf isolation step was introduced so that multiple leaves can be captured together and followed by isolation of individual leaf. The leaf isolation algorithm as adopted in this work has been presented in Table 2.2 and illustrated in Fig. 2.3.

Table 2.2: *Leaf isolation algorithm*

Step 1: Leaves of different species under consideration of this research were arranged on the white sheet of paper in row maintaining the fixed gap.

Step 2:A horizontal scan line was drawn from left to right crossing all the individual images of leaves.

Step 3: Pixel values were scanned along the scan line tracking the number of transitions corresponding to a particular color channel between two consecutive pixels with respect to the fixed threshold value of 100. The value of 100 was arrived by Otsu's method for finding the adaptive thresholding value commonly used for image segmentation and binarization. The number of leaves was tracked by counting the number of transitions.

Step 4: The two extreme coordinates on the line were measured along the *X*-axis and the length of the line was been divided by number of leaves arranged on the white sheet. This gave the coordinate for bounding box

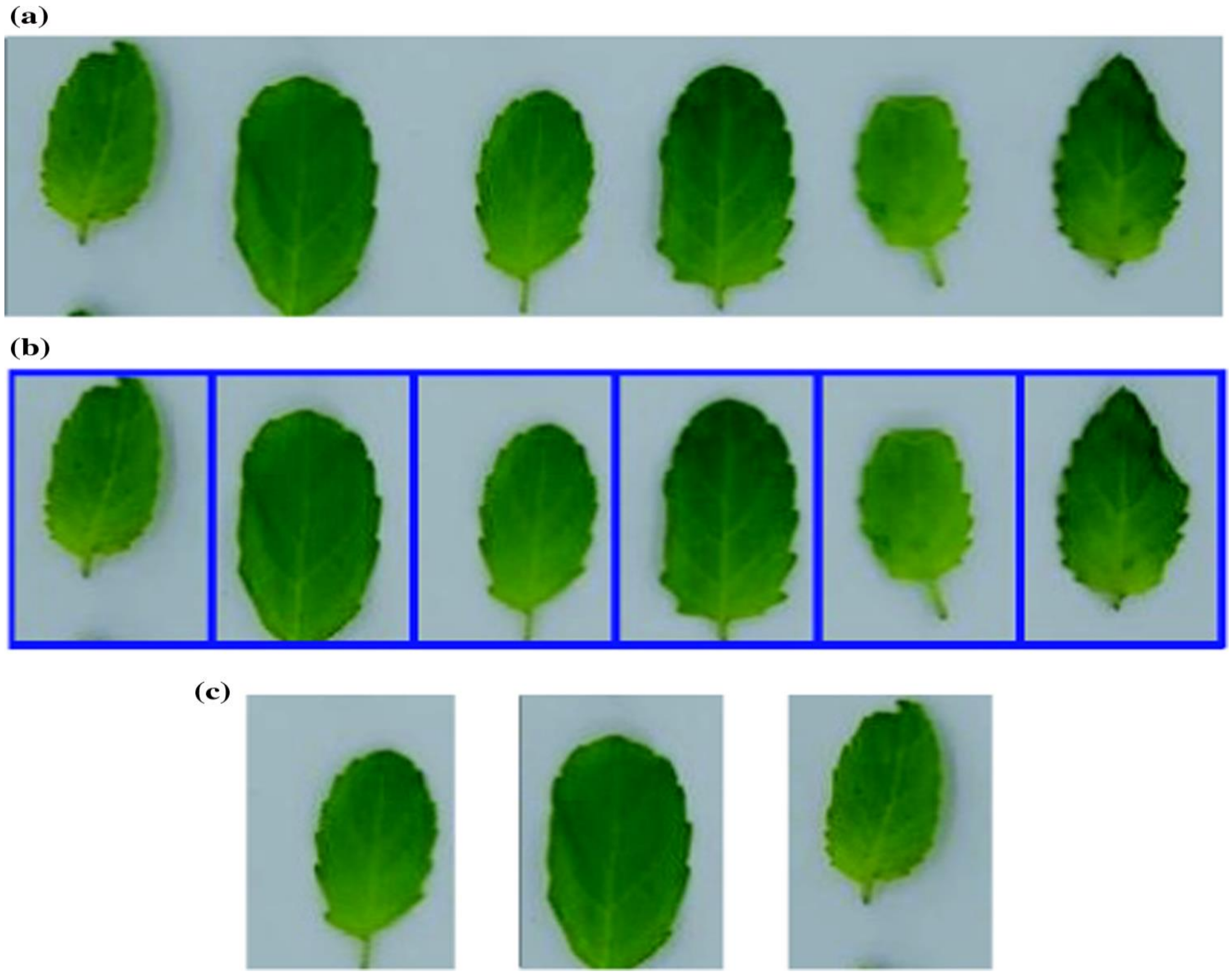


Fig.2.3:Leaf isolation process; (a) capturing multiple images, (b) laying bounding boxes around each leaf and (c) isolated leaf images.

2.3. Development of CNN model

The CNN [21]model has been implemented for the process of classification of different leaf species and their distinct maturity stages of medicinal leaf neem ,tulsi and kalmegh simultaneously. The CNN model comprises of three distinguished layers namely convolution, pooling and fully connected layers. The convolutional and pooling layers are engaged in extracting the information directly from the input images without going through the complex and critical steps of knowledge extraction process as used in many conventional algorithms. This also yields the good benefits in terms of time. The fully connected layer of the model is equivalent to the neural network (NN) architecture comprising of multiple hidden layers with distinct output layers made up of number of output nodes equal to the number of classes involved. Some of the popular pre-trained CNN models are AlexNet, GoogleNet and VGGNet [21]. Despite their considerable accuracy and wide-spread applications some of the limitations of these models include; high hardware configuration and hyperparameter optimization for specific application to obtain higher accuracy. Hence, for this work the CNN model has been developed from scratch to have a model which is lighter in architecture that can be realized in a mobile or handheld device with low-end hardware configurations. The dynamics of different layers of the model as used for this work have been described in following sections.

2.3.1. Convolution layer

Convolution layer comes after the input layer and responsible for producing the feature map from the input images. The extracted feature map moves forward to the pooling layer which appears as the intermediate interface between convolution and fully connected layers. The feature map is the direct result of operation on the images using the filter bank. The convolution layer is also attributed with number of hyper parameters like image size, filter size, number of filters, strides and padding. Stride implies the pixel numbers by means of which the filter traverses through the image and padding. The zero padding has been appended for this work. The feature maps obtained through the convolution layer has been realized by the neurons. Same types of neurons have been assigned for the same type of feature mapping and different types of neurons have been assigned for different feature mapping. The k^{th}output feature map can be expressed as Eq.(2.1).

$$Y_k = f(W_k * x) \qquad (2.1)$$

where,x implies the input image,W_k and Y_k denotethe convolution filter and the output feature map, respectively, corresponding to the k^{th} layer.'*' computes the inner product of the model at each location of the input image and *f(.)* denotes the non-linear representation of the activation function [22] used to extract the non-linear features from the input image. ReLU is one of the most commonly used activation function used in the mapping operation that can be expressed as Eq.(2.2)

$$R(z) = \max(0, z) \tag{2.2}$$

2.3.2 Pooling layer

Pooling layer immediately follows the convolution layer and is used to reduce the spatial resolution of the feature map. It also introduces translational invariance to the small shifts and distortions. This layer reduces the number of parameters used to be learned. Two moments mean and maximum (popularly acronym as *max*) values have been calculated within the convolved image. Max_pooling is the aggregation layer which propagates maximum values to the next layer [23]. Global average pooling is another variation of pooling layer operation also performing the down sampling from the size of $height \times size$ to the array of size 1×1 by means of averaging operation. The output of the pooling operation on the k^{th} feature map can be expressed as Eq.(2.3).

$$Y_{kij} = \max_{p,q \in R_{ij}} \left(X_{kpq} \right) \tag{2.3}$$

where, X_{kpq} denotes the element at the (p, q) location inside the pooling region. R_{ij} defines the receptive field around the location (i,j).

2.3.3 Fully connected layers

This is the final layer of the hierarchical CNN model and similar to the NN architecture. This layer is also named as *dense* layer in CNN realization. It contains one or more hidden layers consisting of numbers of functional nodes that are fully connected to each other. The nodes of last hidden layer are connected to the output layer which consists of output nodes same as the number of classes in subjected problem. The features of any image are extracted in the convolution layers and then down sampled by the pooling layers to get flattened by the concerned flatten layer and finally mapped by the subset of the fully connected layers to the final output of the network by means of high level reasoning [23]. The final output thus forms the *N* dimensional vectors where*N*is the number of classes.

The objective function in the dense layer can be of different types depending on the type of classification task. For multiple class problems, *softmax* is most commonly used one and it can be expressed as Eq.(2.4). It provides the distributed probability of belongingness among the defined classes where summation of all the probabilities results 1 .The class with the highest probability is chosen as the declared class. The nodes of each of the fully connected layers perform the non-linear mapping following the specified objective function which is ReLU in common practice as well as in this work.

$$p\left(y = j / \Theta^{(i)}\right) = \frac{e^{\Theta^{(i)}}}{\sum_{j=0}^{k} e^{\Theta_k^{(i)}}} \qquad (2.4)$$

where, $p(y=j)$ is the probability of belongingness to j^{th}class out of total k number of classes. $\Theta = \sum_{i=0}^{k} w_i x_i$. Schematic representation of the CNN model as used in this work is presented in Fig. 2.4.

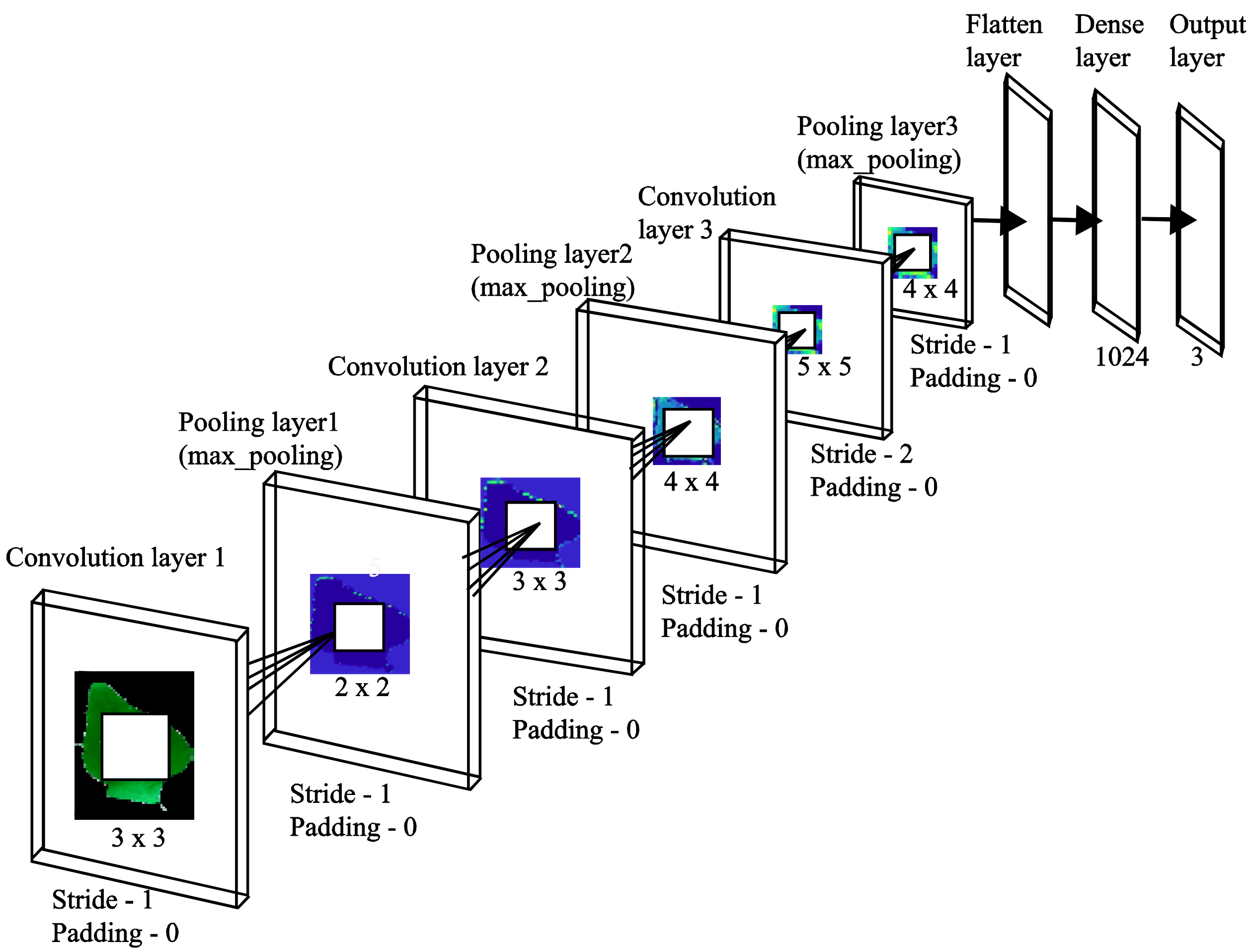


Fig. 2.4: Schematic representation of CNN used for medicinal plant leaf species and maturity classification.

2.4. Hyperparameter optimization

Driven by the layered architectures CNN consists of considerable number of hyperparameters. Setting of these plays crucial roles in model performance. This motivates towards optimization of hyperparameter settings. There are two prominent methods for that; i) starting with random values generated within a specified range for different parameters and employing some search heuristics for finding the optimum condition and ii) taking a pre-trained model and tuning the parameters. Since this work employs a custom CNN model the first approach was adopted.

The key parameters for any CNN architecture are input image size (considering number of rows and columns are equal), number of filters or kernels in individual convolution and pooling layers, kernel dimension, stride size, method of pooling (max or mean pooling) and number of nodes in fully connected layer. As all these are integer numbers the binary heuristics can be employed.

In this thesis three optimization algorithms have been used namely, particle swarm optimization (PSO)[24], gray wolf optimization (GWO)[25] and a proposed modified version of (GWO). However all these algorithms were implemented in binary domain since the problem has been converted into binary search space. One reason for such domain conversion is towards simplicity of the algorithms since binary domain eliminated many control parameters of continuous domain. PSO is one of the most common iterative search oriented optimization techniques and has been used widely including CNN parameter tuning [24, 26].GWO being a comparatively new algorithm but found to be competitive to PSO. Table 2.3 shows a comparative study between PSO and GWO. The modification of the conventional GWO is presented to address the average based position update phenomenon. The goal was to achieve improved performance by modified GWO optimized CNN. The comparative study between the PSO and GWO in terms of their performance potentiality has been represented in the Table 2.3.Table 2.4 expresses the description of different metrics with their corresponding expressions.

Table 2.3: *Comparative study between the PSO and GWO.*

Sl_no	PSO	GWO
1	PSO is the stochastic algorithm simulating the random movements of the swarms.	It is the met heuristic algorithm and simulated by the herds of wolves categorized as alpha, beta, delta and omega wolves as per the

		role played in herds while attacking any prey.
2	The best solution is achieved in the search space over which the swarms are flying around.The best optimum solution is achieved by the pbest or personal best and gbest or global best position of the swarms	The best solution is achieved in the search space over which the wolves are roaming in the attacking spree and it is achieved by the transformed average positional coordinate of the wolves.
3	PSO serves as the good multi objective optimiser	Simple GWO cannot serve towards the multi objective optimisationwell , the attachment of external archive to the optimizer makes it the good candidate for the optimization process.
4	It shows the tendency of early convergence over the short range data and need much memory space for updating velocity.	GWO shows the improved convergence over the PSO over the short range
5	It is prone to fall in local optimum in the space of higher dimension.	GWO is also prone to fall in the local optimum in case of complex problem.

Table 2.4 *Descriptions of the respective metrics*

Serial No	Metric	Expression	Implication
1.	Recall or True Positive Rate (TPR)	TP / (TP + FN)	This metric conveys how well the model can identify the positive cases that are actually positive. Hence higher value is desired.
2.	Specificity or True Negative Rate (TNR)	TN/(TN + FP)	This metric conveys how well the model can identify negative cases which are actually negative. Hence higher value is desired.
3.	Precision	TP/ (TP + FP)	This metric denotes how much the model is precise or accurate among all those predicted positive and how many of them actual positive.
4.	F1_Score	2*(Precision*Recall)/(Precision + Recall)	This metric is the measure to maintain balance between the precision and recall.

			This is important in case of uneven class distribution.
5.	Classification accuracy	Number of correct predictions /Total number of predictions made	This is meant for evaluating the classification model. This implies the fraction of predictions, the model makes.

2.4.1. Binary particle swarm optimization

BPSO has proven potential to reach the global minima with considerable speed and low computational complexity. Hence, in this work BPSO was opted. In brief the initialization, search and finding optimum solution in BPSO[25] follows the same path of PSO. The initial solutions are randomly generated in binary domain which is conceptualized as swarms existing in nature like bird flocks or fish swarms. Two best solutions are driving force of this algorithm, namely, local best (p_{best}) and global best (g_{best}). $p_{best}$is the best solution found by a group in its vicinity while g_{best} is the solution that is best among all the possible solutions in the entire search space. The term *best* is quantitatively assessed by evaluating each solution against a function called *fitness* function. Fundamentally BPSO mimic the food searching behaviors in natural swarms hence all the solutions other than local and global bests update their position and velocity based on those two parameters of the g_{best}andp_{best} solutions as expressed in Eq.(2.5) and (2.6), respectively.

$$v_i^{t+1} = \varpi v_i^t + c_1 r_1 \left(p_{best,i}^t - p_i^t\right) + c_2 r_2 \left(g_{best}^t - p_i^t\right) \tag{2.5}$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{2.6}$$

where, v_i^{t+1} and v_i^t are velocity of i^{th} solution at $(t+1)^{th}$ and t^{th} iteration while x_i^{t+1} and x_i^t are the position index for the same. ϖ is an inertia used for better convergence which is conventionally initiated at 0.9 and eventually reduced to 0.4.c_1 and $c_2$are the constant parameters generated in (1, 2) and r_1 and r_2 are the random numbers generated in (0, 1). Conceptually these random values bring the essence of cognitive and social aspects of natural swarms.

The update dynamics as presented in Eq.(2.5) and (2.6) result in continuous values like PSO which is probabilistically mapped to {0, 1} in case of BPSO as shown in Eq.(2.7).

$$s_{ij}^t = \frac{1}{1 + \exp\left(-v_{ij}^t\right)} \tag{2.7}$$

where, i and j are the solution and bit index, respectively. Since the solutions are binary coded the probability s_{ij}^t is used as possibility of changing the bit value from 0 to 1 or vice-versa. Hence the position update in BPSO is presented as Eq.(2.8).

$$x_{ij}^{t+1} = \begin{cases} 1, if \ q_{ij}^t < s_{ij}^t \\ 0, otherwise \end{cases} \tag{2.8}$$

where, x_{ij}^{t+1} is the j^{th} bit of i^{th} solution at $(t+1)^{th}$ iteration.

2.4.2. Hyperparameter optimization using GWO

Gray wolf optimization (GWO)is a comparatively newly introduced swarm intelligence algorithm that has shown significant potential in different optimization performance tasks and demands much credits for advantages over popular algorithms like particle swarm optimization (PSO) and artificial bee colony (ABC) in terms of optimal tunable parameters and faster convergence since it does not require any derived information in initial search which also helps in bypassing the local minima.

GWO is the bio-inspired metaheuristic designed algorithm which simulates the social behavior of the pack of wolves attacking any target prey in the hunting spree [27] The wolves in this algorithm are hierarchically classified into three groups; alpha(α), beta(β), delta(δ) and omega(ω) on the basis of their actions and motivated by the hunting behaviors of gray wolves. During the hunting process, the position updates of the wolves occur for performing the several tasks like searching, encircling and attacking the pray. Mathematically these steps are represented as Eq. (2.9 –2.10).

$$\overline{D} = |\overline{C} \cdot \overline{X}_p(t) - \overline{X}(t)| \tag{2.9}$$

$$\overline{X}(t+1) = |\overline{X}_p - \overline{A} \cdot \overline{D}(t)| \tag{2.10}$$

where, $\vec{A}$, $\vec{C}$ and $\vec{D}$ are the respective random vectors indicating the simulation of the encircling behaviors of the wolves. t stands for the current iteration, $\overline{A} = 2a \cdot \vec{r}_1 a$ and $\vec{C} = 2 \cdot \vec{r}_2$ .a is the control parameter expressed as in Eq. (2.11).

$$a = 2\left(1 - \frac{it}{N}\right) \tag{2.11}$$

where, the maximum number of iteration is limited by N, the random numbers in (0,1) are represented by $\bar{r}_1$ and $\bar{r}_2$. The position of the grey wolf and the prey vectors are denoted by $\vec{X}$ and

$\vec{X}_p$ respectively. The linear pattern of change in the variable 'a' has been tracked from 2 to 0. The relative updates are mathematically presented as Eq.s (2.12 – 2.14) where X_α, X_β and X_δ denote the position of α, β and δ. C_1, C_2 and C_3 implies set of random vectors in (0, 1). The current solution position has been denoted by $\overrightarrow{X}$.

$$\vec{D}_\alpha = |C_1 \cdot \vec{X}_\alpha - \vec{X}| \tag{2.12}$$

$$\vec{D}_\beta = |C_2 \cdot \vec{X}_\beta - \vec{X}| \tag{2.13}$$

$$\vec{D}_\delta = |C_3 \cdot \vec{X}_\delta - \vec{X}| \tag{2.14}$$

The Eq.s (2.15 – 2.18) represents the step size of the ϖ wolf where X_1, X_2 and X_3 are the final and updated position vectors of α, β and δ wolves, respectively. $\overrightarrow{A_1}$, $\overrightarrow{A_2}$ and $\vec{A}_3$ vectors denotes the random vectors in the range (0, 1). The vectors $\overrightarrow{A_1}$ and $\vec{C}$ play crucial roles in exploration and exploitation for the GWO algorithm. The exploration phase for wolves contains the conditions $|\overline{A}| < -1$ and $|\overline{A}| > 1$. The exploitation of the wolf consists of the condition $|\vec{C}| > 1$. The vector $\vec{A}$ is being influenced by the controlling parameter a.It also causes change in behavior of the ϖ wolves with respect to the prey. If $|\overline{A}| > 1$, the grey wolf may run away from the dominant i.e.The bigger search space or the global search space in optimization has been addressed. On the other hand, if $|\overline{A}| < 1$, the ϖ wolves approach the dominant i.e. the local search space has been addressed as a result of shrinking of the search space.

$$\vec{X}_1 = \vec{X}_\alpha - \overline{A}_1 \cdot \vec{D}_\alpha \tag{2.15}$$

$$\vec{X}_2 = \vec{X}_\beta - \overline{A}_2 \cdot \vec{D}_\beta \tag{2.16}$$

$$\vec{X}_3 = \vec{X}_\delta - \overline{A}_3 \cdot \vec{D}_\delta \tag{2.17}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{2.18}$$

2.4.3. Hyper parameter optimization using Modified GWO

GWO in its conventional form can be a potential optimizer as presented and adopted in previous chapters. In this chapter the conventional GWO has been modified to address the average-based movement dynamics of the wolves. The optimum solution in conventional GWO is reached by average position of the corresponding wolves which implies the average weight for all the wolves during their positional changes. But this is not the case in actual. The position of the α wolf is closest to the prey at the beginning and gradually it is getting away from the prey. On the other hand the position of β and δ wolves can vary keeping the position of β wolves less than or equal to that of α wolves and position of δ wolves less than or equal to that of the β wolves with respect to the dominant wolf or prey. This implies that α wolves are not always dominant and its weight can decrease with the increase in the weight values of β and δ wolves. On the basis of the hypothesis that the α wolves are closest to the prey, the α wolves are most important in attacking the prey and the $X_1 \succ X_2 \succ X_3$ the variable weight values ω_1, ω_2 and ω_3 have been introduced with ω_1 varying from 1 to 1/3 and ω_2, ω_3 varying from 0 to 1/3, respectively. The modification can be represented as Eq. (2.19) to accommodate such weighting strategy.

$$\vec{X}(t+1) = \omega_1\vec{X}_1 + \omega_2\vec{X}_2 + \omega_3\vec{X}_3 \text{ where } \textstyle\sum_{i=1}^{3} \omega_i = 1 \tag{2.19}$$

The weight parameters ω_1, ω_2 and ω_3 were formulated with following assumptions;

The random movement of α, β and δ wolves are influenced by the degree of reaction of the prey targeted by the wolves.

The variation of the weight values is limited between 1 and1/3 for ω_1 and 1/3 to 0 for ω_2 and ω_3 maintaining $\omega_1 > \omega_2 > \omega_3$.

The reaction of the prey is measured on the scale from 0 to 1 with 0 presenting the initial state of the prey when encircling of the wolves around the prey begins and 1 when the prey reacts back to the wolves with the maximum strength.

The α wolves retreat from the prey with ω_1 decreasing from 1 to1/3 where 1 represents the most active state of the wolves ie, the close position of the wolves with respect to the prey and 1/3 represents the passive state ie, the position farthest away from the prey under the influence of the reactionary response of the prey. The β and δ wolves come close to the prey and the associated weight values ω_2 and ω_3 change from 0 to 1/3respectively.

The initial value of the weight ω_1 for the α is 1 when weight values ω_2 and ω_3 of the β and δ wolves is 0.

The reactionary response of the prey wolf is the random number backed up by the probability function. The computation value of such probability function depends on the cost of anti-predator defense due to fear and is expressed by the function $f(k,v)$ of the Holling type II functional response [28] model represented as Eq.(2.20)

$$f(k,v) = \frac{1}{1+kv} \tag{2.20}$$

Where, v denotes the population of predators and $k \geq 0$ is the level of fear.

The respective range of variation of the weight value of the α wolves and the reactionary response of the prey measured by the probability function under the attacks of the grey wolves are mapped to the scale of range from 1 to 1/3 and0 to1, respectively. Any random state η of the prey can correspond to the momentary weight value m of the α wolves. The relation of the states has been given in Eq. (2.21).

$$
\begin{aligned}
&n = 3(1-m)/2;\ where\ 0 \leq n \leq 1 \\
&or \\
&m = (3-2n)/2;\ where\ 1/3 \leq m \leq 1
\end{aligned}
\tag{2.21}
$$

The value of m decides the weight value ω_1 of α wolves

It can be noted here that the optimization was performed in binary domain as in previous cases. The solutions were encoded as binary string as described in Eq. (8) in chapter 2 under section 2.4. The three different optimization processes have been applied to the individual works carried out in respective chapters. The discussion of the result of optimization has also been presented in consolidated format in result section under the comparative analysis section of each chapters.

The cost function plays pivotal role in any search based optimization algorithm. This function depicts the final optimal solution which is arrived by minimizing the cost function (or maximizing in some cases) in iterative manner. The cost function in this work was adopted from Rere et al**.** [29] as expressed in Eq.(2.22).

$$O = \frac{1}{2}\left(\frac{\sum_{N}(y-\hat{y})^2}{N}\right)^{1/2} \qquad (2.22)$$

where, y and $\hat{y}$ are the desired and actual classification accuracies, respectively. N is the number of training samples. The stopping criteria was either exhausting the specified number of iterations (100 in our case) or meeting the value of O at zero for three consecutive iterations.

The values for hyperparameters were encoded in binary and the examples of encoded representations as considered for this work are shown in Table 2.5. It can be seen that each solution is 30 bit long binary code with the specified range of values. A pictorial example of the optimization using BPSO dynamics has been shown in Fig. 2.5.

Table 2.5: *The hyperparameter coding for BPSO algorithm*

Layer type	Variables	Range	Bit length	Encoded example
Convolution	Image size	{32,64,128,256}	2	00 (for 32)
	Im_row = im_column			
	No. of kernels	[4,128]	7	1000000 (64)
	Kernel size	[1,8]	3	110 (6)
	Stride size	[1,4]	2	00 (1)
	Total bit length		**14**	
Pooling	Kernel size	[1,8]	3	100 (4)
	Stride size	[1,4]	2	01 (2)
	Method 1. maxpooling – 0 2. meanpooling - 1	{0,1}	1	1 (1)

	Total bit length		**6**	
Dense (fully connected layer)	No. of nodes	[64,1024]	10	1100100000 (800)
	Total bit length		**10**	
	Model bit length		**30**	

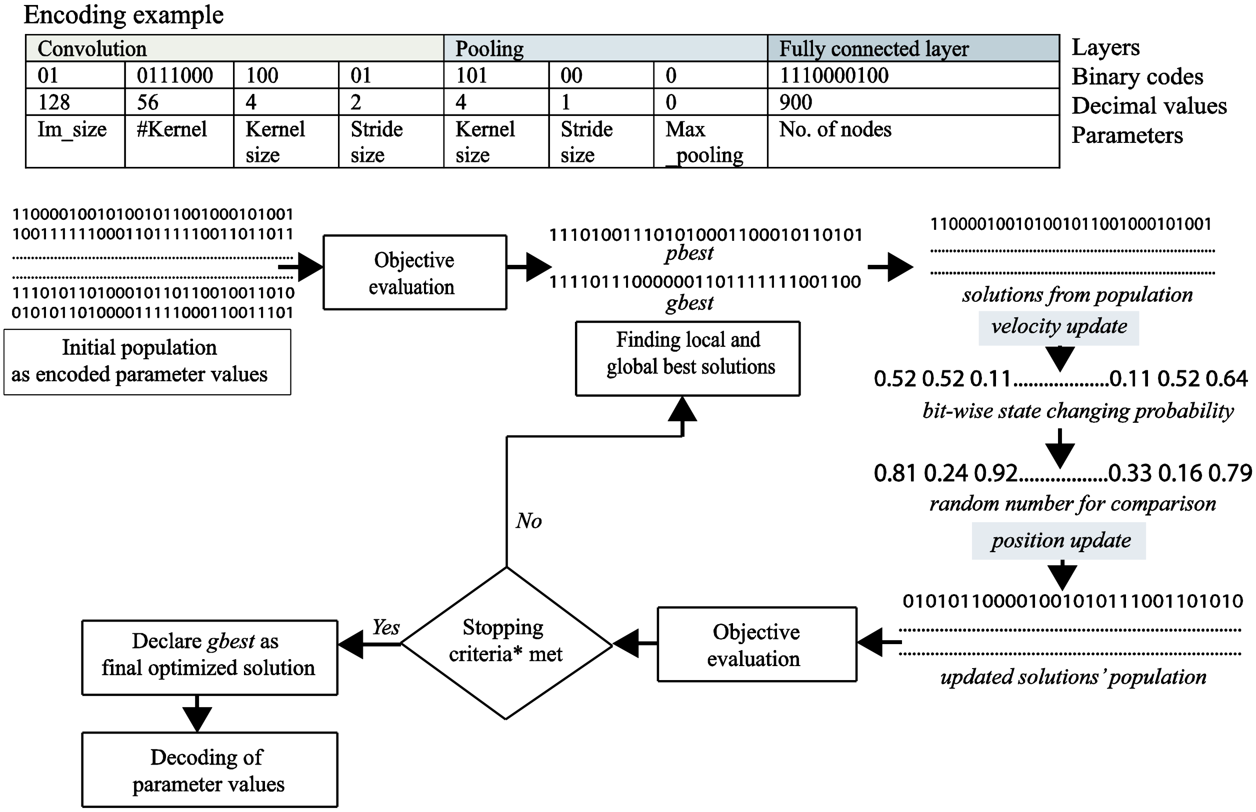


Fig. 2.5: *Schematic representation of the BPSO algorithm.*

2.5. Model performance evaluation

One of the important advantages of machine learning based implementation compared to the conventional instrumental method is the rigorous performance evaluation of the model prior actual implementations against different internal and external metrics. K-fold validationis one of the efficient methods to assess the model performance and consistency. In this case the validation set is divided into equal k-fold and k-1 folds are subjected to training while rest fold is used for testing. In this manner it is guaranteed that each data is subjected to training and testing for at least once. In this thesis 10-fold cross validation has been used where 9 folds have been used for training and remaining 1 fold for testing. However, the 10-fold cross validation has been performed with number of times with different randomly validation dataset from the entire dataset. If the model can perform consistently well in this test with small standard deviation between the accuracies obtained with different folds then it can ascertained that the model is consistent in performance.

In terms of classification confusion matrixis one of the most popularly used method where the matrix is drawn with the actual and resulted true positive (TP), true negative (TN), false positive (FP) and false negative (FN) cases. As the name says TP reflects how many of classified true cases are actually true and TN reflects how many of classified false cases are actually false. The 'true' and 'false' is conventionally used but those can be directly mapped to a case 'belonging' to a particular class and 'not belonging' to that particular case. Classically the confusion matrix was reported for binary classifications (where there are only a positive and a negative case) as shown in Fig. 2.6 but it can be seamlessly extended to multi-class problem taking one class as true and rest all as false.

		Predicted	
		Positive	Negative
Actual	Positive	TP	FP
	Negative	FN	TN

Fig. 2.6: *Confusion matrix for binary classification*

Different important metrics are derived from the confusion matrix. Some of the important metrics as used throughout this thesis have been discussed in Table 2.6.

Table 2.6: *Performance evaluation metrics derived from confusion metrics*

Serial No	Metric	Expression	Implication
1.	Recall or True Positive Rate (TPR)	TP / (TP + FN)	This metric conveys how well the model can identify the positive cases that are actually positive. Hence higher value is desired.
2.	Specificity or True Negative Rate (TNR)	TN/(TN + FP)	This metric conveys how well the model can identify negative cases which are actually negative. Hence higher value is desired.
3.	Precision	TP/ (TP + FP)	This metric denotes how much the model is precise or accurate among all those predicted positive and how many of them actual positive.

4.	F1_Score	2*(Precision*Recall)/(Precision + Recall)	This metric is the measure to maintain balance between the precision and recall. This is important in case of uneven class distribution.
5.	Classification accuracy	Number of correct predictions /Total number of predictions made	This is meant for evaluating the classification model. This implies the fraction of predictions, the model makes.

As CNN does not use previously extracted features Cohen Kappa scoreis an important measure that can reflect the suitability of the automatically extracted features by CNN. It measures the inter-rater reliability of different classes and a higher value indicates betterment of extracted features.

The receiver operating characteristic (ROC)curve also plays important role in assessing the model potential. In this case the true positive rate (TPR) is plotted against false positive rate (FPR). The plot along with the area under the curve (ROC-AUC) depicts the performance. Commonly it is plotted against a no-skill naïve classifier as a diagonal line. Higher ROC-AUC conveys better performance in terms of classification accuracy.

2.6. Results and Discussions

The CNN classifier model has been implemented to carry out the classification process for different plant species and their maturity grading. The CNN has been implemented in python platform using the scikitlearn [30], keras [31] libraries and Tensorflow backend [32]. 30% of entire dataset was randomly selected for hyperparameter optimization using BPSO, GWO and the modified GWO algorithm. The convergence characteristicsare shown in Fig. 2.7 for 100 iterations. The plot of Fig. 2.7 shows that after 80 iterations the BPSO converges with more than 97% accuracy, the GWO converges with a little more than 98% accuracy and for modified GWO, the convergence varies with nearly 99% accuracy value. The hyperparameter finally arrived with modified GWO and as used for the presented results are consolidated in Table 2.7.

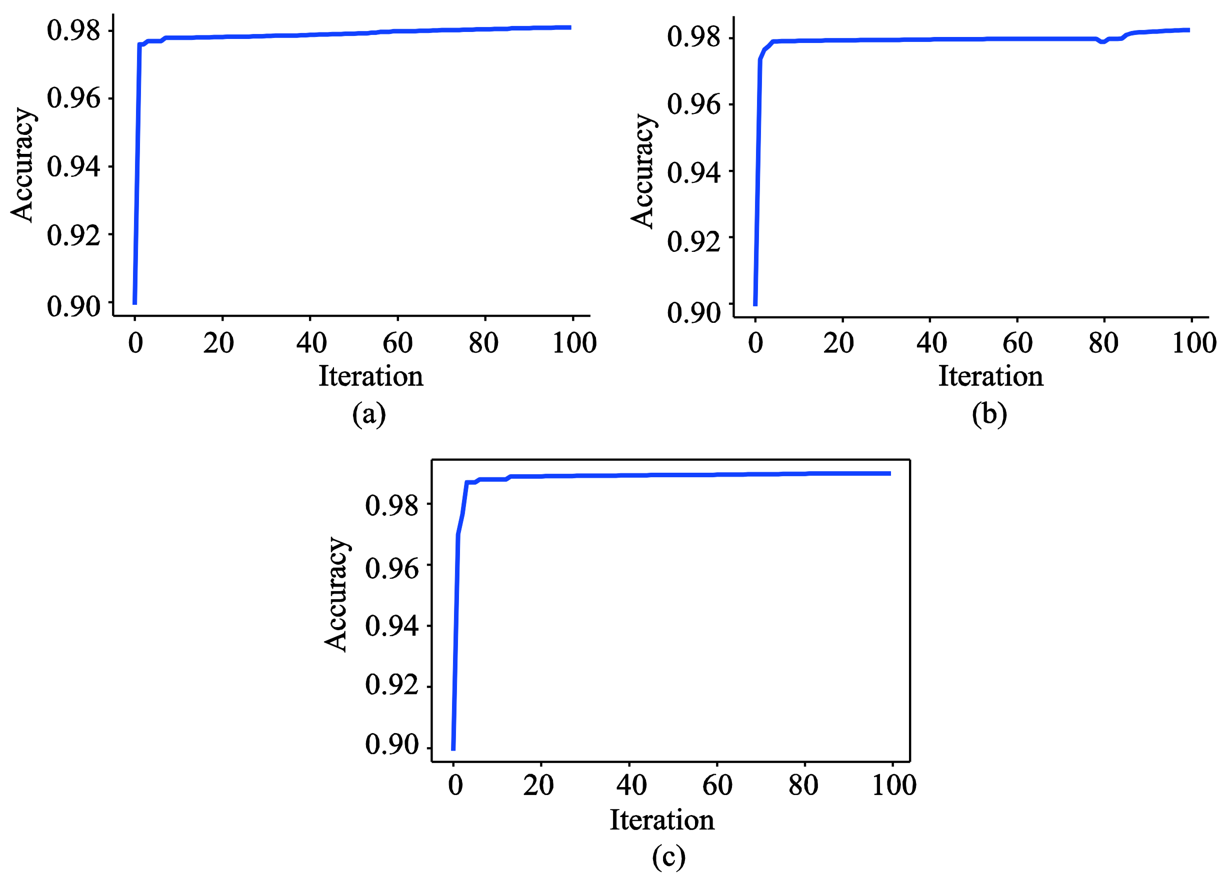


Fig. 2.7:*Convergence plot for (a) BPSO,(b) GWO and (c) modified GWO optimizer*

Table 2.7: *Optimized hyperparameter settings of CNN*

Layer types	Hyperparameters	Parameter values
Convolution 1	Image size	256 × 256
	No. of kernels	32
	Kernel size	3 × 3
	Stride size	1
Pooling 1	Kernel size	2
	Stride size	1
	Method	0
Convolution 2	Image size	128 × 128
	No. of kernels	64
	Kernel size	3 × 3
	Stride size	1
Pooling 2	Kernel size	4
	Stride size	1
	Method	0
Convolution 3	Image size	64 × 64
	No. of kernels	128

	Kernel size	5 × 5
	Stride size	2
Pooling 1	Kernel size	4
	Stride size	1
	Method	0
Flatten	Dimension	6
Dense	No of nodes	1024
Output	No. of nodes	3

Based on the convergence characteristics as given in the Table2.7, the modified GWO optimized CNN model has been chosen as the classifier for carrying out the species and maturity stage classification of the medicinal plants and all other metrics have been measured based on this chosen best.

The CNN by its convolution and pooling layer extracts the inherent features from the images. Also as the input image proceeds from outer to deeper layers more abstraction in feature extraction can be observed. This can be observed in Fig. 2.8 where the image from each layer of the model has been presented as a matter of proof that the model is working towards extracting the deeper features from the input images. As it can be seen in the input layer the image is visible as a plant leaf but the convolution layer and pooling layer following the input layer extracts the feature which is more about the shape and contour of the leaf. As the image proceeds through the network the extraction of features is more towards color and pixel correlation based as can be seen in the right end image.

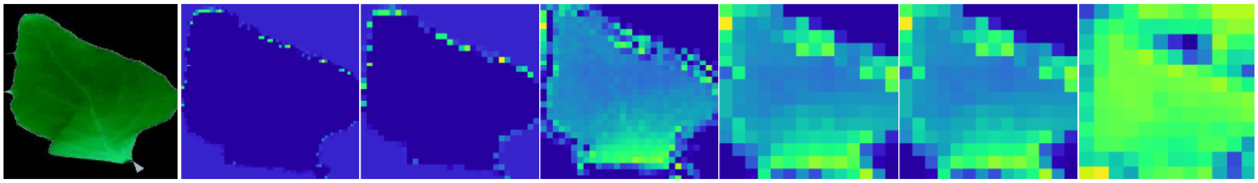


Fig. 2.8: Examples of feature extracted in different layers of CNN

2.6.1. Model validation

The model was validated using 10-fold classification [33]. It can be noted here that the entire dataset was partitioned in 60:20:20 ratios for training, validation and testing purpose. Hence, the all the validation results shown here are with that validation set which disjoint from training and testing set. The results of 10-fold cross validation are shown in Table 2.8.

Table 2.8: *Results of 10-fold cross validation*

Number of folds	Training Data		Testing Data	
	Loss	Accuracy (%)	Loss	Accuracy (%)
1	0.0114	98.31	0.0304	98.28
2	0.0319	99.94	0.0180	97.75
3	0.0154	99.91	0.0307	98.62
4	0.0072	98.97	0.0347	98.06
5	0.0313	99.60	0.0164	99.07
6	0.0338	98.28	0.0125	98.87
7	0.0157	98.84	0.0192	98.75
8	0.0048	99.83	0.0201	98.10

9	0.0097	99.58	0.0240	98.24
10	0.0147	99.91	0.0271	98.17
Average	0.018	99.32	0.023	98.39
Maximum	0.034	99.91	0.035	99.07
Minimum	0.005	98.31	0.013	97.75
Std. Dev.	$\pm$ 0.011	$\pm$ 0.662	$\pm$ 0.007	$\pm$ 0.417

Table 2.8 reflects the consistency in the model performance. The accuracy for both training and testing set is in the tune of 98% with a considerably small (<1) standard deviation. At the same the loss values are also in the tune of 10^{-2} scale which has gone to as low as 10^{-3} scale in some cases. Such small loss is significant in terms of acceptability of model potential. Another important observation is the performance of the model with testing data which was not included in training is visibly close to the training dataset. This reflects the favorable generalization potential of the model as well.

Accuracy and loss are two important parameters to judge the model potential. In machine learning implementations plotting of those two parameters for training and validation sets is a popular measure to understand. The coherence of the plots between training and validation data reflects that the model is not biased and can address the over- and under-fitting issues of model. The over- and under-fitting problems are generally associated for cases where the model can give high accuracy to the training data but fails to perform with considerably accuracy when unknown test data is subjected. The accuracy and loss plots for the entire model as well for classification tasks for individual leaf species (neem, tulsi and kalmegh) under consideration have been shown in Fig. 2.8 for 50 epochs.

Fig. 2.9 conveys that the plots for accuracy and loss with training and validation dataset are similar in nature and trend. The observation remains same for loss and accuracy plots for individual species of leaves. However, the degree of accuracy and loss is varying between species. In none of the cases the gap between the plots at the end of epochs is not visibly high which reflects that the model is not biased. This visibly good in case of tulsi and kalmegh where the plots for training and validation are almost converging. In case of neem there is gap but it is considerably less. In terms of values in all the cases the accuracy is in the range of 96 – 98% while loss values are in the tune 10^{-2}. Overall the validation plots vouch the generalization potential of the model like the 10-fold cross validation results.

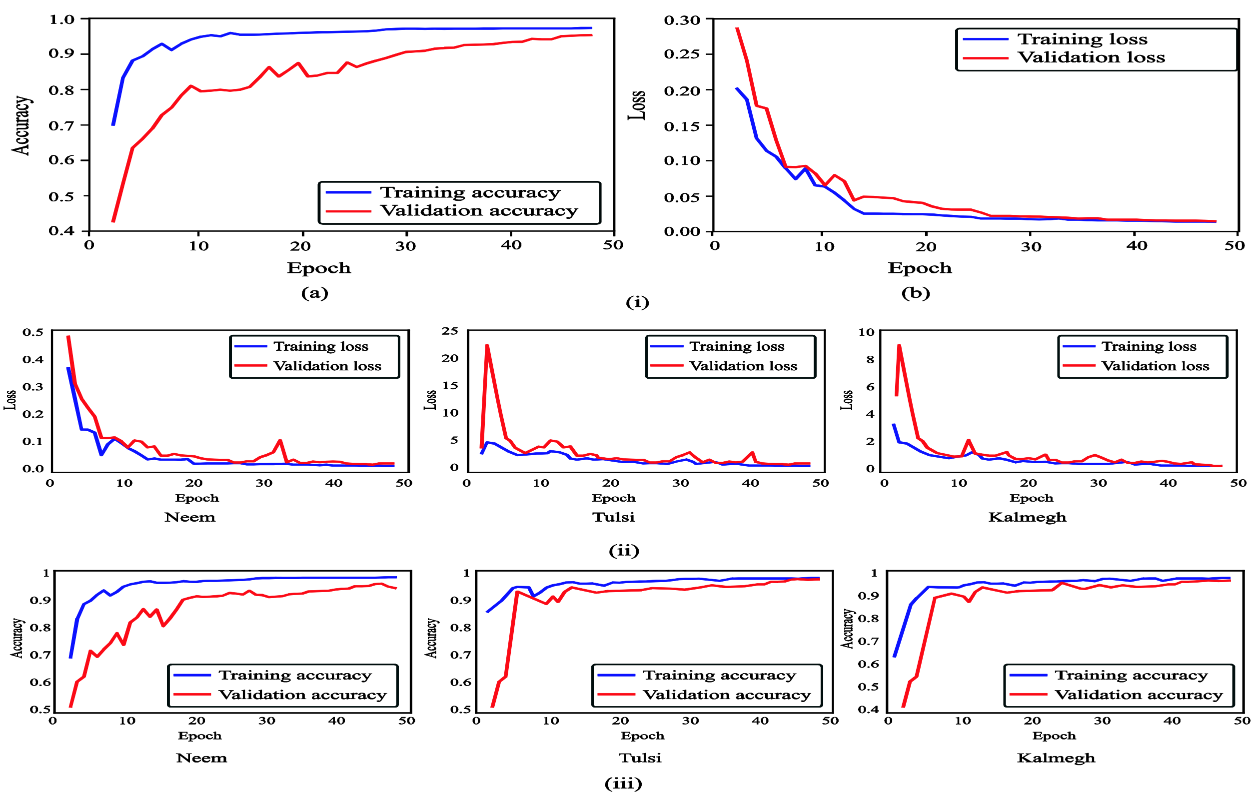


Fig.2.9:*Model validation plots for (i) plots with entire training and validation set, (ii) loss plots for individual species and (iii) accuracy plots for individual species.*

2.6.2. Evaluation against confusion matrix

The confusion matrices [34] for model performance with entire dataset and individual leaf species have been consolidated in Fig. 2.10. The values for derived metric described in section 2.3 have been presented in Table 2.9 and 2.10 for leaf species and maturity classification, respectively.

Actual \ Predicted	Kalmegh	Neem	Tulsi
Kalmegh	498	1	1
Neem	2	496	2
Tulsi	1	2	497

Leaf species classifier

Actual \ Predicted	PM	M	OM
PM	100	0	0
M	1	98	1
OM	1	0	99

Maturity stage classifier for Kalmegh

Actual \ Predicted	PM	M	OM
PM	100	0	0
M	0	100	0
OM	1	0	99

Maturity stage classifier for Neem

Actual \ Predicted	PM	M	OM
PM	99	0	1
M	0	100	0
OM	0	0	100

Maturity stage classifier for Tulsi

Fig. 2.10: *Confusion matrix for leaf species classification and maturity classification for individual leaf species. The maturity stages premature, mature and over-mature have been abbreviated as PM, M and OM respectively.*

Table 2.9: *Performance evaluation for leaf species classification*

Leaf Classes	Precision	Sensitivity/ recall	FMeasure	Specificity
Kalmegh	1.00	1.00	0.99	
Neem	0.99	0.98	0.98	
Tulsi	0.99	1.00	0.99	
Micro avg	1.00	0.99	0.99	
Macro avg	0.99	0.99	0.99	0.98
Weighted avg	1.00	0.99	0.99	

Table 2.10: *Performance evaluation for maturity stage classification*

Leaf types	Maturity type	Precision	Sensitivity/ recall	FMeasure	Specificity
Kalmegh	Premature	0.99	1.00	1.00	
	Mature	0.99	0.98	0.98	
	Overmature	0.99	1.00	0.99	
	Micro avg	0.99	0.99	0.99	
	Macro avg	0.99	0.98	0.99	0.98
	Weighted avg	1.00	0.99	1.00	

Neem	Premature	1.00	0.99	0.99	
	Mature	1.00	0.99	0.99	
	Overmature	0.98	0.99	0.98	
	Micro avg	0.99	0.99	0.99	0.99
	Macro avg	0.99	0.99	0.99	
	Weighted avg	0.99	0.99	0.99	
Tulsi	Premature	0.99	0.98	0.99	
	Mature	1.00	0.99	0.99	
	Overmature	0.99	1.00	1.00	
	Micro avg	0.99	0.99	0.99	0.99
	Macro avg	0.99	0.99	0.99	
	Weighted avg	1.00	1.00	0.99	

The performance measures of the confusion matrix show that the model can visibly perform for species identification between three medicinal leaf species. In all the cases more than 495 unknown (not included in the training or validation sets) samples out of 500 samples have been correctly identified. Similarly in case of maturity stage identification about 99 samples out 100 have been correctly identified in terms of their maturity stage. The metric values presented in Table 2.9 and 2.10 echo the same facts. The overall accuracy in the range of 98 - 99% confirms the ability of the model to provide promising performance for both species and maturity classification. The precision and specificity values in the tune of 98% also convey that the model can identify the actual species and maturity of the subjected samples with considerable accuracy.

It can be observed that in some cases the accuracy and other metrics have reached even 100% it is because for some of the plant species the visual different is very prominent between the maturity levels. For instance, in case of neem leaves the greenness in the leaves considerably decrease when it turns to mature from premature but not such visual difference is present when it turns to over-mature in terms of color. However, the shape and creases occurs when it turns to over-mature. As CNN captures not only a single feature but an abstracted holistic features in cases where the classes are not visibly apart it can distinguish considerably. The F-Measure is also an important metric that provide the balance between sensitivity and specificity. In case the F-Measure is low the model turns to be a biased one which can either detect the positive cases with high accuracy but not the negative cases or vice-versa. But, in this case the F-Measure value is also more than 0.98 which conveys that in 98% cases the model can correctly identify the species and maturity stages of the leaves accurately.

2.6.3. Cohen kappa coefficient score

Cohen kappa [35] is one of the metric that assesses the inter rater reliability of different classes. The high value for the coefficient (near to 1) reflects goodness of the feature selection. The value of this coefficient is represented in Table 2.11 for the species classification and maturity classification. The estimated value for the coefficient is found to be 0.99 across all the species and maturity classes which convey the suitability of the extracted features by CNN for the subjected classification tasks.

Table 2.11:*Evaluation of Cohen Kappa score for Leaf species classifier and Maturity classifier*

Classifier	**Classes taken for measurement**	**Cohen Kappa score**
Leaf species classifier	Kalmegh, Neem and Tulsi leaves	0.9988
	Pre-mature, mature and over-mature	
Maturity stage classifier	Kalmegh	0.9998
	Neem	0.9988
	Tulsi	0.9980

2.6.4. Receiver operating characteristics (ROC) curve

The ROC plots [36] for different classification task are shown in Fig. 2.11. It can be observed that for species the area under coverage is considerably high for all the three medicinal plants under consideration. The lowest ROC-AUC is also in the tune of 0.98 which conveys that all the classifiers can classify the actual species of medicinal leaf among different leaf species with more than 98% accuracy. In case of maturity stage classification, the area coverage for different maturity stages among different plants is consistently in the range of 0.98 – 1. That reflects the potential of the presented CNN model towards correct identification of leaf maturity. One thing can be noticed here that for the all the cases the ability of identifying premature leaves is almost 100% which convey the fact the visual difference between premature and mature or over-mature is very certain but the potential of the model is reflected by high accuracy to distinguish between mature and over-mature which are comparatively close in terms of their visual appearance and difficult to be segregated by human vision.

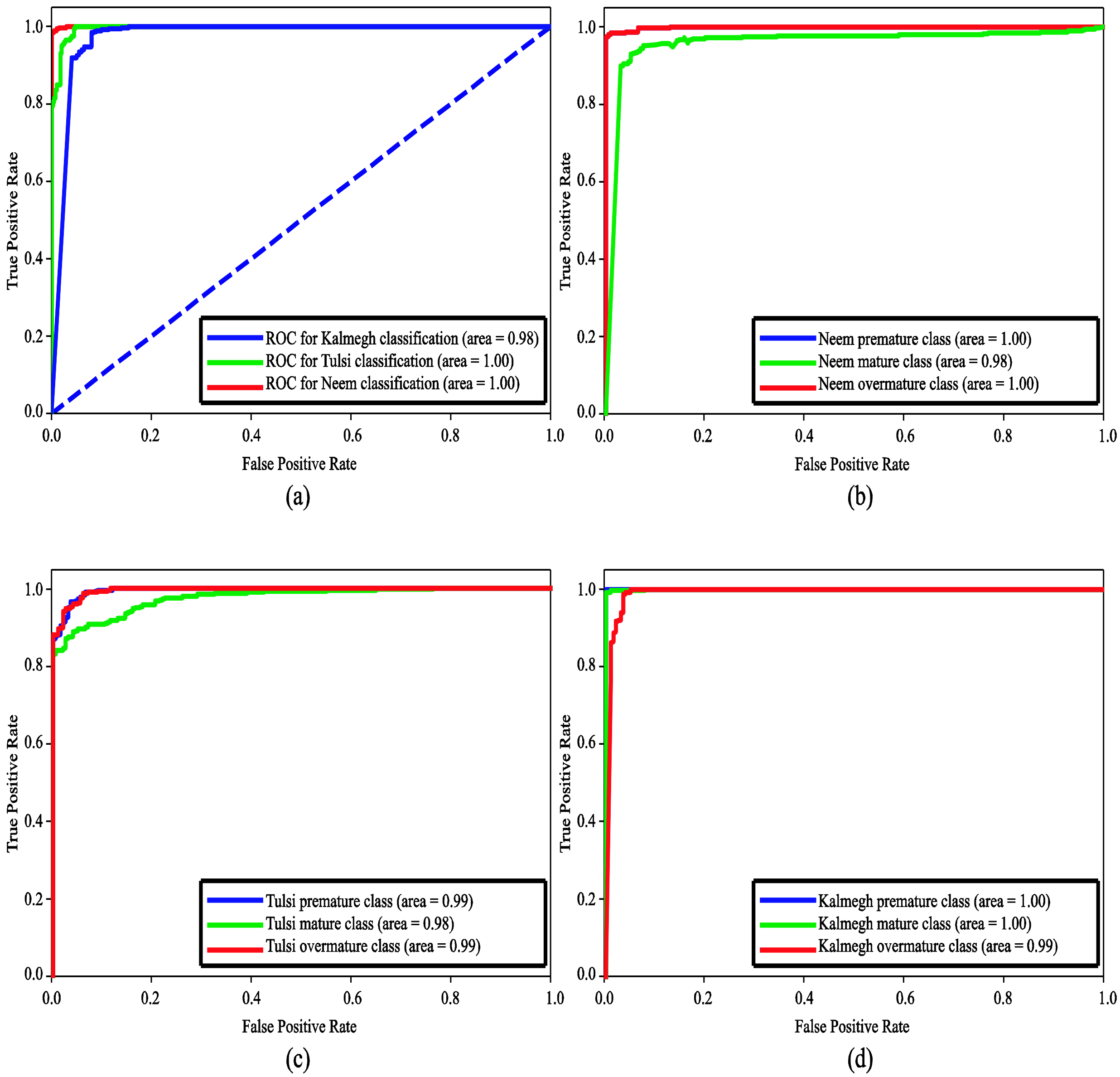


Fig. 2.11:ROC curves of classification of (a) species and maturity stages for (b) neem, (c) tulsi and (d) kalmegh leaves.

2.6.5. Comparative analysis

Table 2.12 presents a comparative analysis on the performance of CNN and other conventional classification algorithms, namely, artificial neural network with back propagation multilayer perceptron architecture (BPMLP) [37], support vector machine (SVM) [38] and random forest (RF) [39]. Unlike CNN all these algorithms need prior feature extraction. Hence, the statistical parameters mean, median, standard deviation, minimum and maximum values for individual color channels i.e. R, G, B were extracted as features. This resulted feature set dimension of 4500×5 which was again splitted into training, validation and testing sets with 60:20:20 ratios as CNN implementation. Apart from the conventional models some of the pre-trained CNN models namely AlexNet [40], VGG 19[41] and ResNet 50 [42] have also been included in the comparative analysis.

Table 2.12: *Performance comparison of presented CNN model with other classifiers*

		Species classification/Maturity stage classification				
Method	**Classification**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**FMeasure**
BP MLP	Species	0.97	0.98	0.96	0.97	0.96
	Maturity	0.96	0.97	0.96	0.97	0.96
SVM	Species	0.96	0.95	0.96	0.96	0.95
	Maturity	0.96	0.95	0.96	0.96	0.95
RF	Species	0.95	0.95	0.97	0.95	0.95
	Maturity	0.95	0.95	0.96	0.94	0.96
Deep learning based						

AlexNet	Species	0.96	0.94	0.95	0.96	0.96
	Maturity	0.94	0.94	0.96	0.97	0.94
VGG19	Species	0.93	0.93	0.94	0.95	0.96
	Maturity	0.94	0.93	0.95	0.96	0.96
ResNet50	Species	0.95	0.95	0.93	0.93	0.94
	Maturity	0.93	0.94	0.94	0.94	0.95
Presented CNN model optimized with BPSO	Species	0.99	0.99	0.98	1.00	0.99
	Maturity	0.99	0.99	0.99	0.99	0.99
Presented CNN model optimized with GWO	Species	0.99	0.99	0.98	1.00	0.99
	Maturity	0.99	0.99	0.99	0.99	0.99
Presented CNN model optimized with modified GWO	Species	0.99	0.99	0.98	1.00	0.99
	Maturity	0.99	0.99	1.00	1.00	1.00

Table 2.13 shows that CNN outperforms other classifiers under consideration. SVM performs similarly for both species and maturity stage detection tasks and performance is not good as BPMLP and RF. BPMLP and RF performs equivalently, but considering the computational complexity RF may be favored. The values computed by respective metrics carried out by the

deep learning-based classifiers AlexNet,VGG19 and ResNet50 are not also as good as those presented in the CNN model developed in this work. CNN not only performs better than the other classifiers in comparison but also eliminates the feature extraction step. Hence, presented CNN based classifier can be considered as more potential method. It is clearly visible from the performance presentation of the optimized CNN that BPSO optimized CNN and the GWO optimized CNN are analogous in terms of both the species and maturity classification. The modified GWO has slightly higher impact on performance metric of the CNN model compared to the GWO approach.

2.6.6. GUI presentation

The GUI implementation for the species as well as the maturity classification of the leaves neem, tulsi and kalmegh has been presented in this chapter.The Fig 2.12 presents the species wise and maturity wise identification for any unknown leaf sample belonging to any of the respective species classes,neem, tulsiandkalmegh and to any of the three different maturity stages premature,mature and overmature. The prediction regarding the correct species or the maturity state has been expressed by the corresponding prediction vector. The prediction vector for the GUI has been expressed as the combination of 0 and 1. The species classification activity for any unknown sample of the medicinal plants has involved the prediction vectors [1.0.0], [0.1.0] and [0.0.1] for leaves kalmegh, neem and tulsi respectively. The maturity identification process involving the determination of premature , mature and overmature states for all different medicinal leaves respectively has included the prediction vectors [1.0.0], [0.1.0] and [0.0.1] .

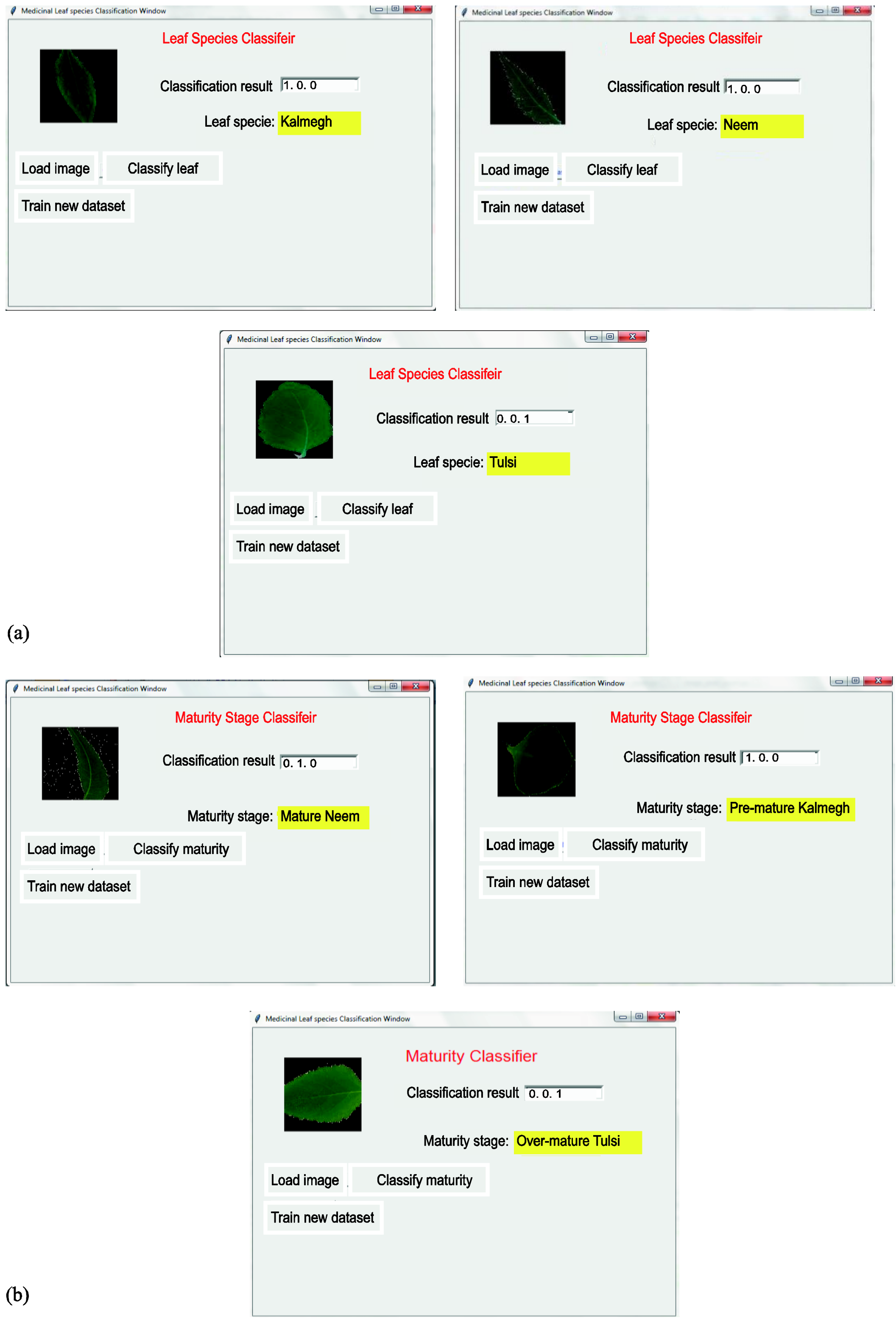


Fig. 2.12:*Screenshot of GUI windows corresponding to the medicinal leaf (a) species and (b) maturity identification for the unknown samples.*

2.7. Conclusion

This chapter has presented a promising computer vision framework based on CNN architecture towards classification of species and maturity of medicinal leaves. The details of variation of different bio-actives present in the medicinal lave and their accountability towards the maturity wise discrimination has been examined in justification to the problem. The development of an indigenous imaging chamber and its application in capturing images of plant leaves have been elaborated in this chapter. The chapter includes the procedure starting from the sample collection to CNN model development. It also highlights the use of BPSO,GWO and modified GWO optimization for hyperparameter optimization in order to get more generalization and performance of the CNN model. The performance of the modified GWO has shown slight improved performance towards the accuracy of classification. The validation of the model and classification results with unknown testing set has been evaluated against standard metrics which show that the presented method can achieve up to 98% accuracy. The comparative analysis has also been presented towards justification of using CNN over some of the conventional classifiers. On an overall assessment, by virtue of significant performance potential of the presented framework, it can be considered as an addition to the existing instrumental and analytical screening and sorting methods of plants and medicinal leaves.

Reference

1) Manjunath B.S., OhmJ.R.,Vasudevan V.V.,Yamada A.(2001). Color and texture descriptors, IEEE Transactions On Circuits and Systems for Video Technology, 11(6), pp. 703- 715.

2) Rafiqa A.et.Al. (2013). Application of computer vision system in food processing- A review. J Eng Res Appl 3(6):1197–1205.

3) Faridi H., Aboonajmi M. (2017). Application of machine vision in agricultural products. In: Proceedings 4th iranianinternational NDT conference, Olympic Hotel, Tehran, Iran, Feb 26–27.
4) Yalcin H., Razavi S.(2016). Plant classification using convolutional neural networks, *Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, pp. 1-5, doi: 10.1109/Agro-Geoinformatics.2016.7577698.Wilfa P et al (2016), Computer vision cracks the leaf code. PNAS 113(12):3305–3310.
5) Chaki J., Parekh R., Bhattacharya S. (2015). Recognition of whole and deformed plant leaves using statistical shape features and neurofuzzy classifier. In: IEEE Proceedings of 2nd international conference on recent trends in information system (ReTIS): 2015.
6) Keivani M., Mazloum J., Sedaghatfar E., Tavakoli M.B. (2020). Automated analysis of leaf shape, texture, and color features for plant classification. Traitement du Signal 37(1):17–28. https://doi.org/10.18280/ts.370103.
7) Eid H.F., Abraham A. (2018). Plant species identification using leaf biometrics and swarm optimization: a hybrid PSO GWO, SVM Model. Int J Hybrid IntellSyst 14(2):1–11.
8) Turkoglu M., Hanbay D. (2019). Recognition of plant leaves: an approach with hybrid features produced by dividing leaf images into two and four parts. Appl Math Comput 352:1–14.
9) Mahajan S., Raina A., Gao X.Z., PanditA.K. (2021). A plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost. Symmetry 13:356. https://doi.org/10.3390/ sym13020356.
10) Bhuiyan M.M., Abdullahil-Oaphy M., KhanamR.S., IslamM.S. (2020).MediNET: A Deep Learning Approach to Recognize Bangladeshi Ordinary Medicinal Plants Using CNN.
11) LeCun Y., Bengio Y., Hinton G.R. (2015). Deep learning. Nature 521(7553):436–444.
12) Hinton E., Osindero S., TheY.W.(2006). A fast learning algorithm for deep belief networks. Neural Comput 18(7):1527–1554.
13) Jaswal D., SowmyaV., Soman K.P. (2014). Image classification using convolutional neural networks. Int J Adv Res Technol 3(6):1661–1668.

14) Lipton Z.C., Berkowitz J. (2015). A critical review of recurrent neural networks for sequence learning, arXiv:1506.00019v4[cs.LG] 17 Oct, 2015.

15) Salle A., Villavicencio A. (2018). Restricted recurrent neural Tensor networks: Exploiting word frequency and compositionality. In: Proceedings of the 56th annual meeting of the association for computational linguistics (Short Papers): 8–13, Melbourne, Australia, July 15 – 20.

16) Sardogan M., Tuncer A., Ozen Y. (2018). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In: 2018 3rd International conference on computer science and engineering (UBMK) (pp. 382–385). IEEE.

17) Deepalakshmi P., LavanyaK., Srinivasu P.N. (2021). Plant leaf disease detection using CNN algorithm. Int J InfSyst Model Des (IJISMD) 12(1):1–21.

18) Karthik R., Hariharan M., Anand S. et al. (2019). Attention embedded residual CNN for disease detection in tomato leaves. Appl Soft Comput. https://doi.org/10.1016/j.asoc.2019.105933.

19) Jeon W.S., Rhee S.Y. (2017). Plant leaf recognition using a convolution neural network. Int J Fuzzy Log IntellSyst 17:26–34.

20) Sapijaszko G., MikhaelW.B.(2018). An overview of recent convolutional neural network algorithms for image recognition. In:2018IEEE 61st International midwest symposium on circuits and systems (MWSCAS), pp. 743–746. https://doi.org/10.1109/MWSCAS.2018.8623911.

21) Yamashita R.,Nishio, M. , Do, R.K.G. (2018). Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629.

22) Traorea B.B., Foguema B.K., Tangara F. (2018). Deep convolution neural network for image recognition. Eco Inform 48:257–268.

23) Wang G.,Guo J., Chen Y., Li Y. , Xu Q. (2019). A PSO and BFO-based learning strategy applied to faster R-CNN for object detection in autonomous driving.IEEE Access. https://doi.org/10.1109/ACCESS.2019.2897283.

24) ShahrakiM.H.N., Taghian S., Mirjalili S.(2020). An improved grey wolf optimizer for solving engineering problems, Expert Systems with Applications, vol. 166, p. 113917.

25) Mirjalili S. , Mirjalili S.M., Lewis A. (2014). Grey wolf optimizer. AdvEngSoftw 69:46–61.

26) Wang Y., Zhang H., Zhang G. (2019).cPSO-CNN: An efficient PSO based algorithm for fine-tuning hyper-parameters of convolutional neural networks. Swarm EvolComput 49:114–123.

27) Wang X.,Zanette L., Zou X.(2016). Modelling the fear effect in predator–prey interactions, Mathematical Biology, J. Math. Biol. 73:1179–1204.

28) Dupke C., Peters A., Morellet N. *et al.* (2021).Holling meets habitat selection: functional response of large herbivores revisited. *Mov Ecol* **9**, 45 (2021). https://doi.org/10.1186/s40462-021-00282-6

29) Rere L.M., Fanany M.I., ArymurthyA,M,(2016), Metaheuristic Algorithms for Convolution Neural Network. *Computational Intelligence and Neuroscience*, *2016*, [1537325]. https://doi.org/10.1155/2016/1537325.

30) Buitinck. L. et. al. (2013). API design for machine learning software: experiences from the scikit-learn project. In: European conference on machine learning and principles and practices of knowledge discovery in databases. arXiv:1309.0238.

31) Chollet F. (2015).Keras. https://github.com/fchollet/keras, Accessed on 10 January 2020.

32) Abadi M.et.al.(2016). TensorFlow: A system for large-scale machine learning. In:12th USENIX Symposium on operating systems design and implementation (OSDI 16), USENIX Association:265–283.

33) Berrar D. (2018). Cross-validation. Encyclopedia of Bioinformatics and Computational Biology, 1, Elsevier, 542 -545.

34) Santra A.K., Christy,C.J.(2012). Genetic algorithm and confusion matrixfor document clustering. Int J ComputSci 9(1):322–328.

35) McHugh M.L. (2012). Interrater reliability: the kappa statistic. Biochem Med 22(3):276–282.

36) Fawcett T. (2006). An introduction to ROC analysis. Pattern RecognLett 27:861–874.

37) Achkar R., Owayjan M., Harajli B., Khazaal D., Dbouk M.,Magnifico G. (2019), Brain Tumor Classification Using Back Propagation Algorithm in MLP,*2019 Fourth International Conference*

on Advances in Computational Tools for Engineering Applications (ACTEA), pp. 1-5, doi: 10.1109/ACTEA.2019.8851109.

38) Cristianini N, Shawe-Taylor J. (2000), An Introduction to Support Vector Machines and other kernel-based learning methods. Cambridge University Press, Cambrigde, Book website: www.support-vector.net.
39) Breiman L.(2001). Random Forests. Machine Learning, 45, 5-32. http://dx.doi.org/10.1023/A:1010933404324.
40) Xiao L., Yan Q. , Deng S. (2017). Scene classification with improved AlexNet model, *12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE),* pp. 1-6, doi: 10.1109/ISKE.2017.8258820.
41) Bansal M., Kumar M., Sachdeva M., Mittal A. (2021). Transfer learning for image classification using VGG19: Caltech-101 image data set. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-021-03488-z.
42) Celano G.G.A. (2021).A ResNet-50-Based Convolutional Neural Network Model for Language ID Identification from Speech Recordings. In Proceedings of the Third Workshop on Computational Typology and Multilingual NLP, pages 136–144, Online. Association for Computational Linguistics.

CHAPTER 3

SPECIES AND MATURITY GRADING OF VEGETABLES – CASE STUDY WITH TOMATOES

3.1Introduction

Vegetables are essential part of our daily food habit [1]. These provide required nutrients to our body that enhances immunity against several diseases [2]. Among all the vegetables, tomato isone of themost widely consumed vegetables across the world including India. Tomato has the unique capacity of reducing the risk of cancer, osteoporosis, and other different types of cardiovascular diseases [2].The presence of the chemical compound like *lycopene* inside tomato renders its antioxidant properties [3].

Tomato is available in different species and maturity stages. Pear and roma are two popular varieties of tomato in India. In case of tomato fruit, the change in the maturity is expressed in terms of change in the peel color along with the textural change of the fruit skin. Since both the agro produces under consideration show biochemical reactions after harvesting it is important to monitor their maturity and assess the number of days it can be stored which is commonly referred as *days-to-rot*.

The work presented in this chapter has been published as:

Mukherjee G, Chatterjee A, Tudu B..(2022). A Computer Vision Approach Towards Maturity Stage Classification of Tomatoes Using Second Order Wavelet Features.In: Saraswat, M., Roy, S., Chowdhury, C., Gandomi, A.H. (eds) Proceedings of International Conference on Data Science and Applications . Lecture Notes in Networks and Systems, vol 288. Springer, Singapore. https://doi.org/10.1007/978-981- 16-5120-5_10

The different stages of ripening can be divided into three stages; namely, premature, mature and over-mature. Premature corresponds to unripe, mature with slightly ripen and over-mature with the fully ripen having the trends of growth towards the rotten stage at slow pace.

The detection of species and maturity plays pivotal role in terms of human health as well as economy of a country. For instance, if the days-to-rot can be predicted the harvested fruits can be well utilized in the supply chain to the supermarket avoiding the loss of rotten fruits. Some of the works utilizing various different techniques in this direction have been consolidated in the Table 3.1.

Table 3.1: *Some of the reported techniques for automatic grading of tomatoes*

Base technique	**Method**	**Result**
Portable electronic nose device attached to 10 metal oxide sensors [4]	The PCA and LDA techniques	The accuracy value of 100% has been achieved
The internal structures of the tomato fruits has been measured by means of the two dimensional magnetic resonance images.[5]	The MR image information clubbed with the classical analysis techniques in tracking structural and physiochemical changes during the maturation of fruit.	The accuracy value of 90% has been obtained in this process.
Raman spectra in determination of the internal maturity of the tomato[6]	The method of self modelling mixture analysis for separating the Raman spectra corresponding to the outer pericarp and Teflon layer.	The decrease in the SID value with ripening of tomatoes is distinctly found in the Raman spectra.
Near infrared spectroscopy technique in classification in terms of maturity and to predict about the textural	The PCA and soft independent modelling of class analogy(SIMCA) .	The classification accuracy value of 96.85% has been achieved.

properties of the fruit [7]		
The visible and near infrared spectroscopy for determining the maturity of the green tomatoes[8]	The Bayesian and joint variety global classification model in predict the tomato maturity	The procedure carried out here resulted with the accuracyvalue of 85 %.
The fuzzy rule based classification approach to determine the ripeness of tomato. [9]	Mamdani fuzzy inference systems for classifying tomato into six different classes.	The classification accuracies has been found to be 94.29%.
Chemical sensors based analysis depending on the carbon nanotubes and carbon nanocoils for assessing the ripeness stages of tomato[10]	Tristimulus and PCA methods	Improved response of PCA vouching for the electronic distinctions of the ripening stages tomatoes

The instrumental methods include limitations like expensive instruments, time consuming and precise sample preparation process and lack of portability. Most of these techniques are destructive and invasive as well. On the other hand manual inspection is tedious, time consuming and subjective. Computer vision can be a promising proposition to address all such limitations of current instrumental and manual methods. The advancement in CNN has further enhanced the potential of computer vision models in different detection and prediction tasks for agro produces. This chapter elaborates a CNN based computer vision framework that can be used for species and maturity classification for tomatoes.A new swarm intelligence method called gray wolf optimization (GWO) has been adopted in this work to optimize the hyperparameters in order to enhance performance of the model in terms of accuracy as well as consistency.

3.2. Dataset preparation

3.2.1. Sample collection

Among variety of vegetables tomato was selected for this work due to its world-wide consumption and ready availability throughout the year. There are different species of tomatoes. In this work two of them namely,roma and pear were selected due to their ready availability in local market. All the samples of tomato under consideration were collected in the same season in order to avoid the seasonal effects. Three maturity stages were labeled as i.e. premature, mature and over-mature depending on their ripeness and days-to-rot. It is worth mentioning here that lowering the ripening stage makes longer days-to-rot in normal storage condition. Total 15000 tomatoes comprising of 7500 for each species were collected. For each of the maturity stages 2500 samples were collected. While procuring the maturity stage classification was done by botanists and the samples were preserved at 20ºC until they were subjected for imaging.

3.2.2. Image acquisition

The indigenous imaging chamber as described in section 2.2.2 was used for this work. It can be noted here that, there are publicly available datasets for tomato classification but the motivation for dataset preparation was towards inclusion of more custom distortions and better observations as of those which can provide a more generalized classification model.

Due to the huge numbers of samples, in this case also multiple samples were captured in single imaging run and followed by automated image isolation process as described in Table 3.2. After isolation the background subtraction process was carried in order to increase the region of interest towards improving the model accuracy. Some of the results of the automated isolation and background subtraction are shown in Fig. 3.1.

Table 3.2: *Tomato and banana image isolation algorithm*

Step 1: The fruits were arranged in a row on the white sheet of white paper maintaining a fixed gap between samples.

Step 2: A horizontal scan line was drawn from left to right crossing all the individual images of samples.

Step 3: The image was decomposed in R, G and B color channels. Pixel values were scanned along the scan line tracking the number of transitions corresponding to a particular color channel between two consecutive pixels with respect to the adaptive threshold value obtained using Otsu's algorithm. The selection of color channel and finding the threshold value has been described in detail in section 3.2.2.1. The number of fruits was tracked by counting the number of transitions.

Step 4: The horizontal length between the two extreme coordinates(left to right) along the scan line was measured along the *X*-axis and the measured length was divided by the number of fruits laid on the sheet to retrieve the location of the bounding box around each of the fruit.

Step 5: Each individual fruit along the line were surrounded by the respective bounding boxes based on the coordinate position detected in step 4.

Step 6: Finally the image of respective fruit under each bounding box was cropped into separate images.

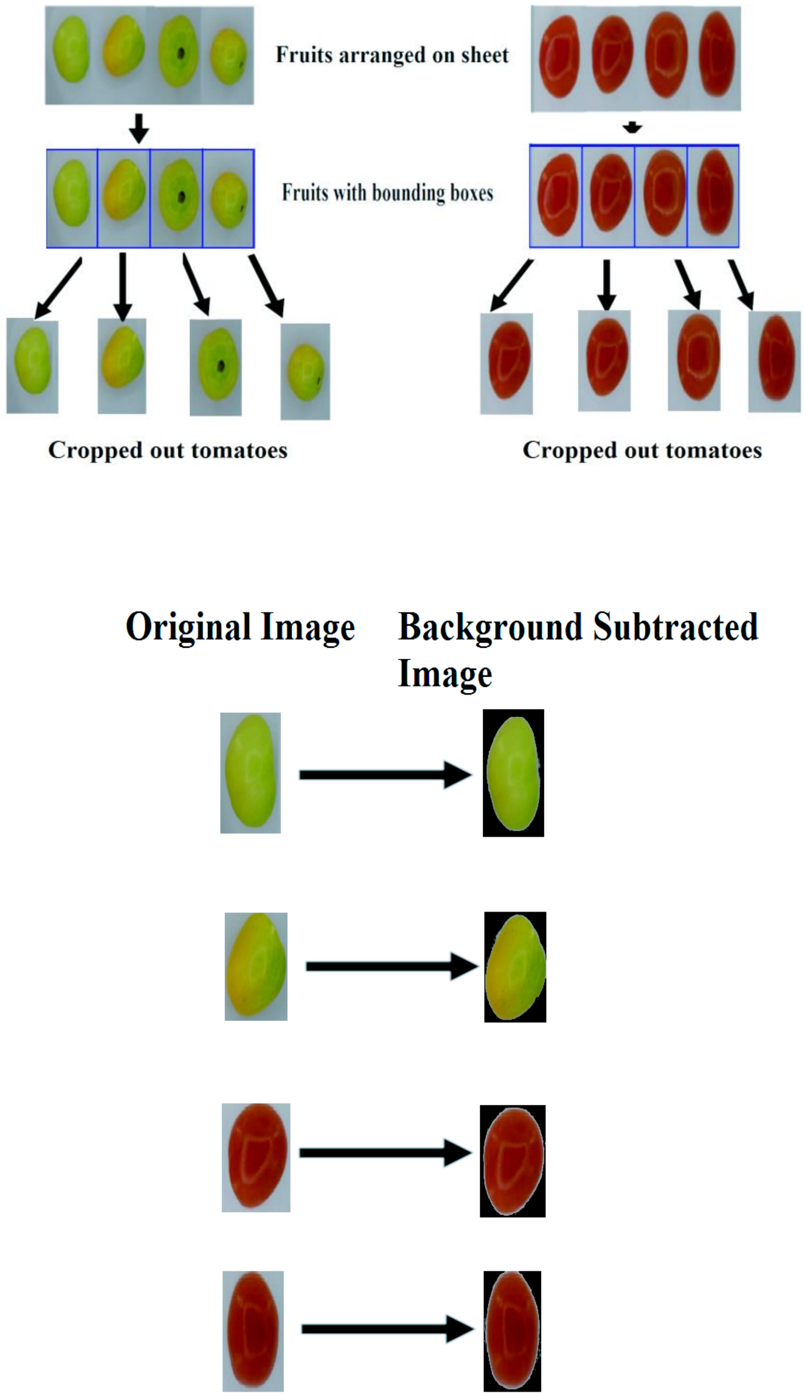


Fig. 3.1: *Examples tomato isolation and background subtraction output*

3.2.2.1. Channel selection and thresholding for image isolation

The color channel separation for one of the tomato image is shown in Fig. 3.2. As it can be seen that the color channels (R, G and B) are not similar in nature as that largely depends on the color of tomato skin. It can also be seen that the histogram of B channel is bimodal with a considerable gap between peaks. Hence this channel was subjected to the Otsu's segmentation. To automate this process randomly 100 images were selected for each species and maturity stage of tomatoes. The mean of adaptive threshold values was chosen as the threshold value for that species and maturity stage. For example, in this way the threshold value was 100 for pear mature class the value was taken as 80.

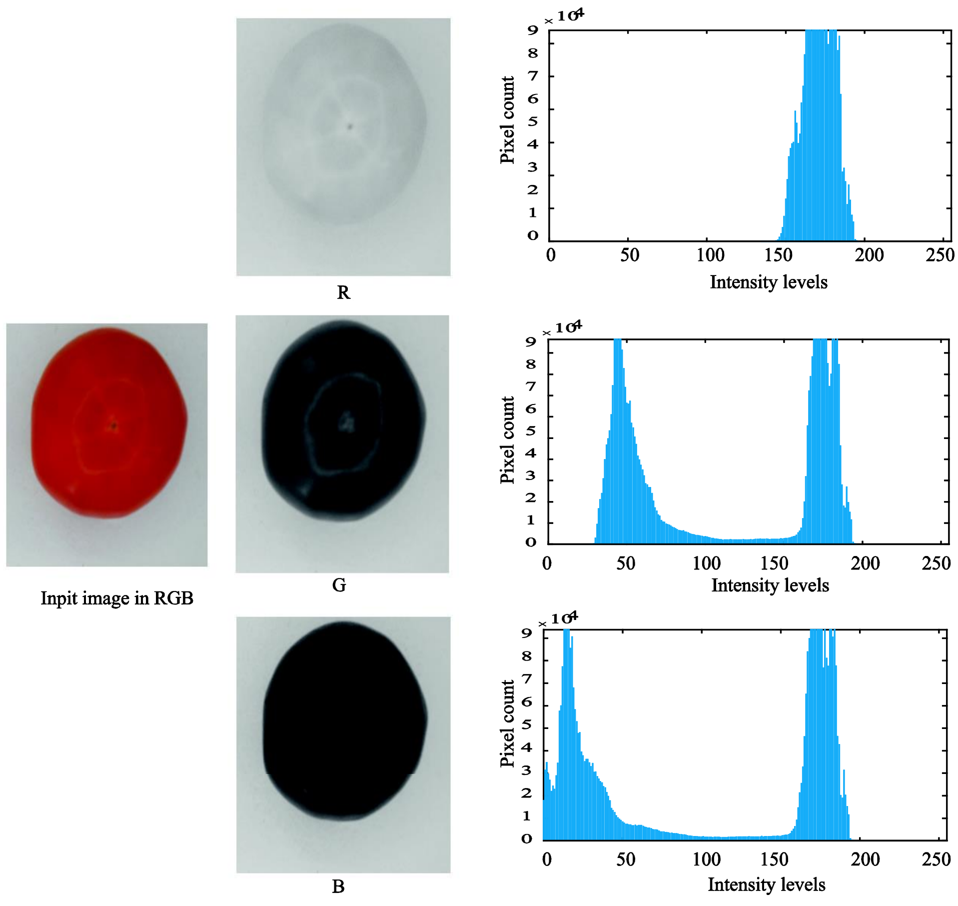


Fig. 3.2: Color channel dynamics for tomato images

3.3. TomNet – CNN model for species and maturity stage classification of tomatoes

This section presents the CNN model developed for tomato grading in terms of their species (roma or pear) and maturity stage (either of premature, mature and over-mature). For convenience purpose the model has been called TomNet throughout this thesis. Conventionally the CNN model consists of three distinct layers convolution layer, pooling layer and the final fully connected layer. The CNN model for this work consisted of 7 convolution layer each followed by max pooling layer. The output layer was connected to 3 fully connected layers. The filter size for convolution and max_pooling layer was chosen as 3×3 and 2×2, respectively. The schematic representation of the model is shown in Fig. 3.3. The layer dynamics of TomNet are presented in Table 3.3. As it can be seen TomNet uses the BatchNormalization (batch_norm) function which significantly improves the performance of CNN models by faster training, improved regularization and avoiding internal covariate shifts. All these contribute towards higher potential of the model to avoid over-fitting problem.

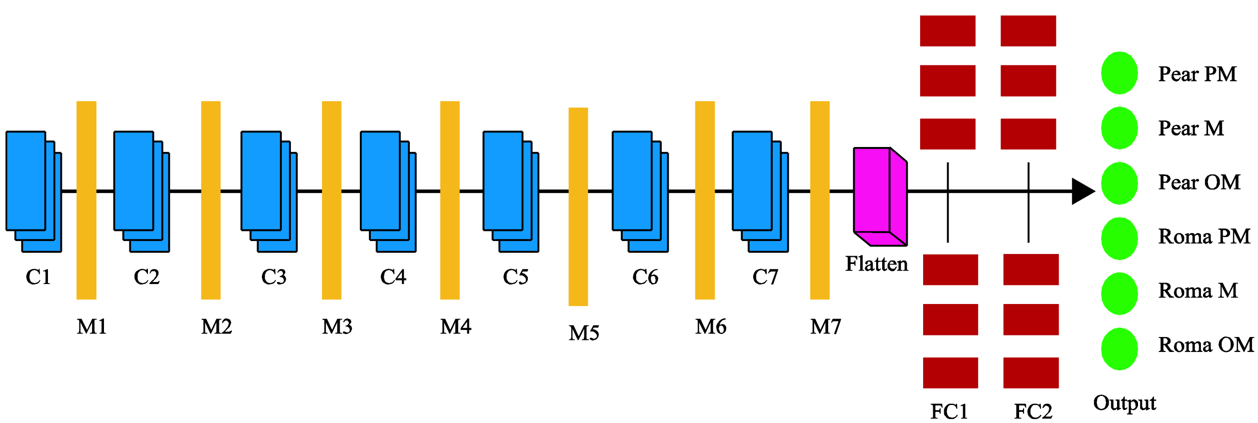


Fig. 3.3: *Schematic representation of TomNet architecture*

In training stage the learning rate was initiated at 0.005 and the batch size was maintained at 64. Due to the selection of *Adam* as optimization function no manual modulation of training rate was needed. . Values for exponential decay rate in respect to first and second moment were 0.9 and 0.98, respectively. Examples of layer-wise output from different convolution layers have been shown in Fig. 3.4.

Table 3.3:*TomNet layer dynamics*

Layer type	Purpose	Mathematical expression
Convolution layer Conv_2d	Feature extraction through the non-linear activation function [11]	$Y_k = f(W_k * x)$ Y_k - feature map of k^{th} filter W_k - convolutional filter x – input image $f(\cdot)$- RelU $R(z) = \max(0, z)$
Pooling layer max_pooling2d	Reduces the spatial resolution of the feature maps followed by the introduction of the translation invariance to the small shifts and distortions which results into the reduced number of parameters to be learned [12].	$Y_{kij} = \max_{p,q \in R_{ij}} (X_{kpq})$ Y_{kj} - max pooled output k^{th} feature map X_{kpq} - element at the (p, q) location contained in the pooling region R_{ij} - the receptive field around the location (i, j)
Dense layer	Fully connected layer for final decision with non-linear *softmax* activation function.[13]	$p(y = j \mid \Theta^{(i)}) = \frac{e^{\Theta^{(i)}}}{\sum_{j=0}^{k} e^{\Theta_k^{(i)}}}$ $p(i, j)$- probability of belongingness to j^{th} class among total k classes $\Theta = \sum_{i=0}^{k} w_i x_i$

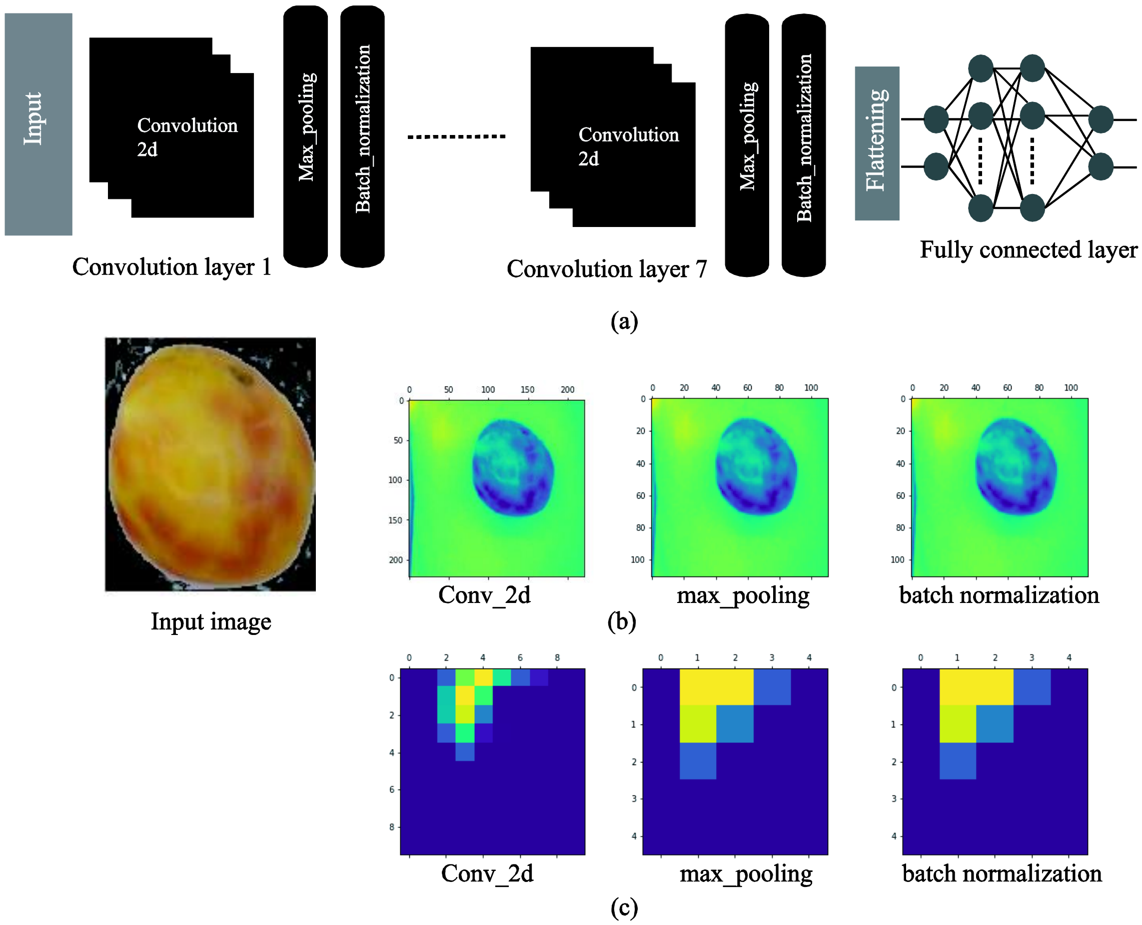


Fig. 3.4: *Visualizing the output of TomNet model*

Fig. 3.4 reflects the advantage of CNN architecture in extracting the abstract features from the images automatically which many of the conventional algorithms lack. For instance, Fig. 3.4 (b) shows the images from the initial convolution, max_pooling and batch normalization layers. It can be seen that in all these three images the shape of tomatoes can be identified. As the images forwards through the deeper layers such clear visibility of shape does not remain same and more inherent color, texture and pixel corellational features get extracted as can be seen in Fig. 3.4 (c). In this way TomNet can achieve feature extraction crucially that contain both surface level features like shape and more deeper levels of features like color distribution, pixel correlation and texture. Such deeper sensing mechanism enables TomNet towards higher accuracy and generalization while avoiding many distortions like blurring and rotation during image capture stage.

3.3.1. Hyperparameter optimization of the TomNet model

The hyperparameters of the TomNet model have been optimized by using the three most promising methods BPSO, GWO and the modified GWO . The architectural description of the appointed models has been described in the section 2.4 in details. The comparative analysis of the three optimizer have been carried out in terms of the convergence curves which is provided in the result and discussion section of this chapter.

3.4. Results and Discussions

For the classification of tomato fruit and its identification for maturity, the model Tomnet has been used. Total 7500 different samples of individual species of tomato with 2500 samples for each of the maturity grade have been considered for the presented results. For hyperparameter optimization 30% of entire dataset was randomly chosen which makes total 4500 samples with 2250 for each species and 750 samples of each maturity stage. The convergence plot based on the BPSO, GWO and modified GWO for 100 iterations are shown in Fig. 3.5.

The plot of Fig. 3.5 shows that after 80 iterations the BPSO [14] converges with more than 99% accuracy, the GWO [15] converges with a little more than 98% accuracy and for modifiedGWO , the convergence varies with nearly 98.5% accuracy value. The hyperparameter finally arrived with the BPSO as used for the presented results are consolidated in Table 3.4.

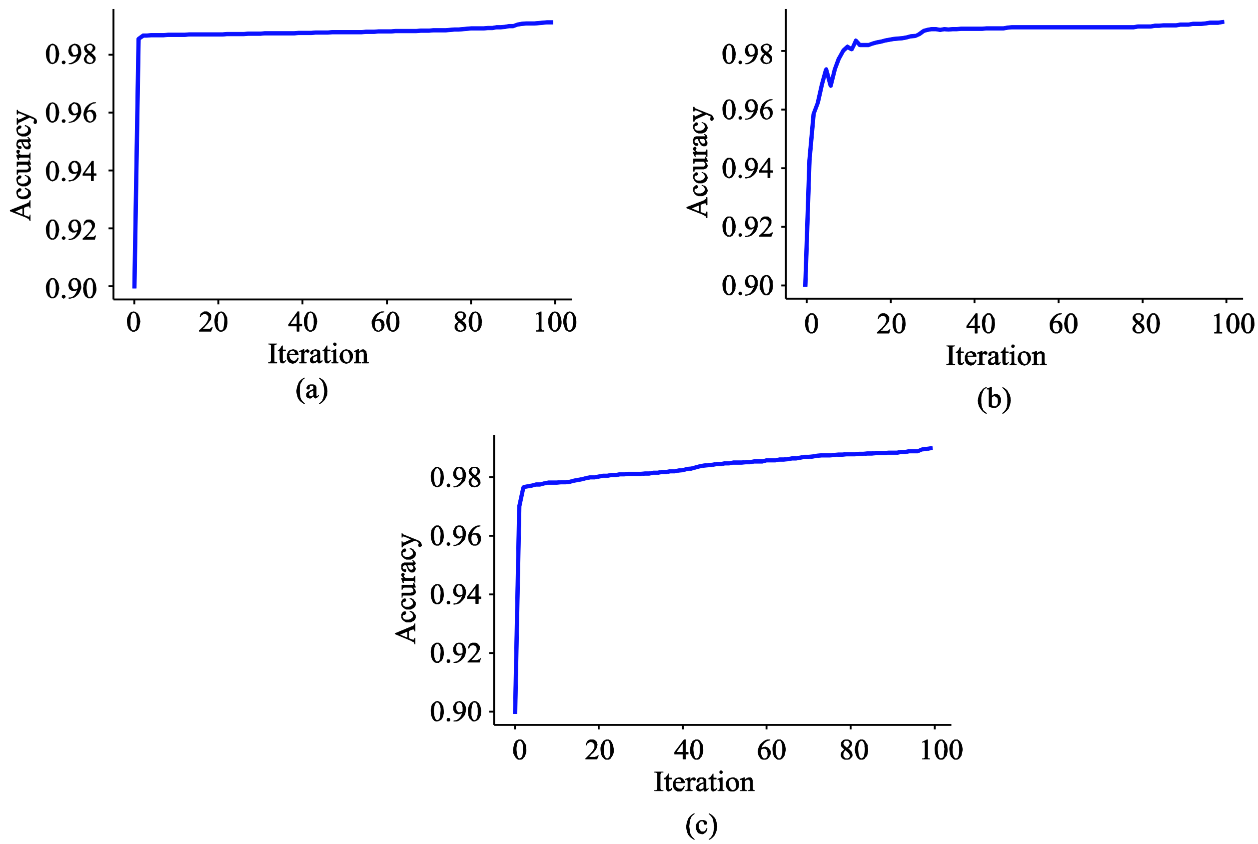


Fig. 3.5:Convergence plot for (a) BPSO ,(b) GWO and (c) modified GWO optimizer

Table 3.4:Hyperparameter settings of TomNet

Layer	Output shape	Number of parameters
conv2d_1 (Conv2D)	(510, 510, 32)	896
max_pooling2d_1	(255, 255, 32)	0
batch_normalization_1	(255, 255, 32)	128
conv2d_2 (Conv2D)	(253, 253, 32)	9248
max_pooling2d_2	(126, 126, 32)	0
batch_normalization_2	(126, 126, 32)	128
conv2d_3 (Conv2D)	(124, 124, 64)	18496
max_pooling2d_3	(62, 62, 64)	0

batch_normalization_3	(62, 62, 64)	256
conv2d_4 (Conv2D)	(60, 60, 64)	36928
max_pooling2d_4	(30, 30, 64)	0
batch_normalization_4	(30, 30, 64)	256
conv2d_5 (Conv2D)	(28, 28, 64)	36928
max_pooling2d_5	(14, 14, 64)	0
batch_normalization_5	(14, 14, 64)	256
conv2d_6 (Conv2D)	(12, 12, 128)	73856
max_pooling2d_6	(6, 6, 128)	0
batch_normalization_6	(6, 6, 128)	512
conv2d_7 (Conv2D)	(4, 4, 128)	147584
max_pooling2d_7	(2, 2, 128)	0
batch_normalization_7	(2, 2, 128)	512
flatten_1 (Flatten)	512	0
dense_1 (Dense)	256	131328
activation_1 (Activation)	256	0
dropout_1 (Dropout)	256	0
dense_2 (Dense)	84	21588
Output	6	510
activation_2 (Activation)	6	0

Total parameters: 479,410

Trainable parameters: 478,386

Non-trainable parameters: 1,024

Based on the convergence characteristics as given in the Fig 3.5 , the BPSO optimizedTomNet model has been chosen as the classifier on the convergence characteristics performance shown in fig 3.5 for carrying out the species and maturity stage classification of the fruit tomato and all other metrics have been measured based on this chosen best.

3.4.1. Model validation

The model validation was performed using accuracy and loss plot with validation set which were non-overlapping 20% samples of the entire dataset. Validation plots have been shown in Fig. 3.6. The plots for accuracy show that for training set the model can reach upto 98% accuracy while for validation set it is about 96%. This reflects that the model can identify the unknown samples quite accurately in terms of their species and maturity stages. The loss values are also visibly less for validation set. Further the gap at the end between plots for training and validation set are also not big and both have similar trend. Hence, this can be considered as good generalization potential of the model to avoid possible over- and under-fitting problems.

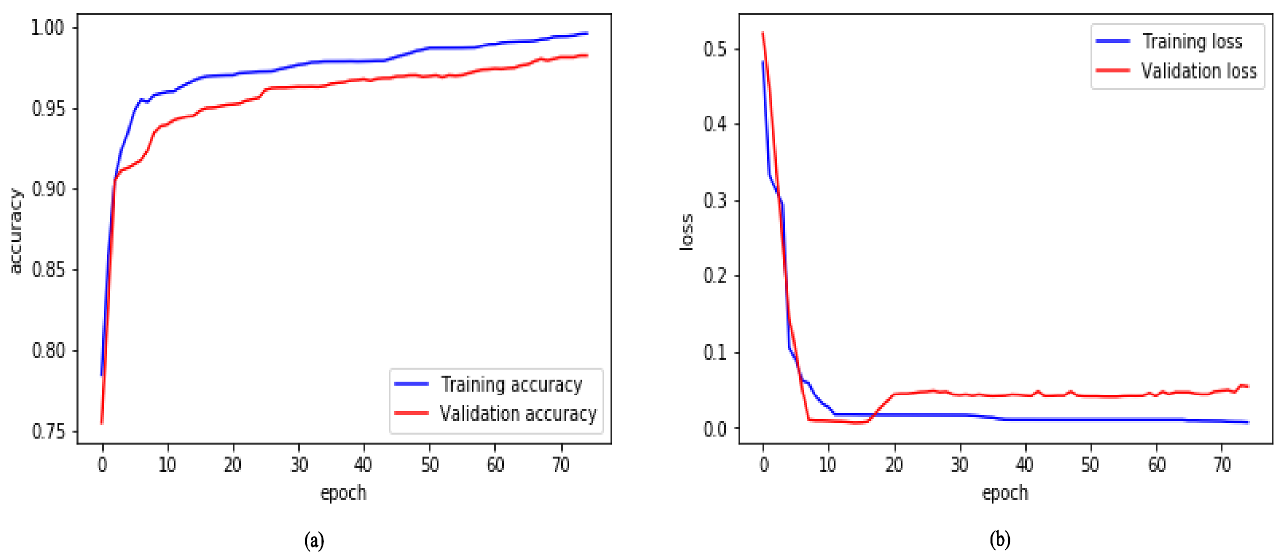


Fig. 3.6:*TomNet validation using (a) accuracy and (b) loss plots.*

The advantages of 10-fold cross-validation [16] have been stated previously in section 2.4.1. The result of this validation for TomNet model has been presented in Table 3.5. It conveys that the TomNet is consistent in performance. The average accuracy for training set is about 99% while it is in the tune of 98% for testing set. Another interesting observation is not much difference between minimum and maximum values for both the training and testing sets. It along with low standard deviation values (<1) vouches the consistency of the model.

Table 3.5: *10-fold cross-validation results for TomNet model*

10-fold cross validation run	**Training data**		**Validation data**	
	Loss	**Accuracy**	**Loss**	**Accuracy**
1	0.0199	0.9897	0.053	0.9801
2	0.0126	0.9934	0.043	0.9834
3	0.0235	0.9965	0.034	0.9856
4	0.0156	0.9943	0.054	0.9833
5	0.0246	0.9974	0.044	0.9792
6	0.0136	0.9899	0.051	0.9821
7	0.015	0.9979	0.042	0.9784
8	0.0166	0.9985	0.043	0.9863
9	0.0126	0.9943	0.057	0.9833
10	0.0172	0.9945	0.054	0.9892
Average	0.1712	0.9947	0.475	0.9831
Maximum	0.0246	0.9985	0.057	0.9892
Minimum	0.0126	0.9897	0.034	0.9784
Std. Dev.	0.0043	0.0031	0.007	0.0034

3.4.2. Performance evaluation against confusion matrix

The confusion matrix [17] of the model performance with unknown (not been used for training and validation) testing set is presented in Fig. 3.7. As the testing set contains 20% non-overlapping samples of entire dataset each of the 6 classes under consideration has 500 samples.

Actual \ Predicted	Pear premature	Pear mature	Pear overmature	Roma premature	Roma mature	Roma overmature
Pear premature	495	0	0	2	2	1
Pear mature	0	490	0	6	1	3
Pear overmature	0	1	496	0	3	0
Roma premature	2	0	0	495	1	2
Roma mature	0	4	2	2	492	0
Roma overmature	1	2	2	0	0	495

Fig. 3.7: *Confusion matrix for tomato species and maturity stage classification using TomNet*

Fig. 3.7 shows that for each class most of the samples were correctly identified in terms of species and maturity stage. Class-wise maximum 4 samples out of 500 have been wrongly classified in case of roma mature which has been identified as pear mature. In all other cases the class-wise misclassification is not more than 2. Table 3.6 presents the metric values derived from the confusion matrix.

Table 3.6:*TomNet performance evaluation using metrics of confusion matrix*

Specie	**Maturity stage**	**Precision**	**Recall**	**F1-score**	**Specificity**
Pear	Premature	0.99	0.96	0.99	0.98
	Mature	0.99	0.98	0.98	0.98
	Overmature	0.99	1.00	1.00	0.98
Roma	Premature	0.99	0.97	0.99	0.99
	Mature	0.99	0.99	0.98	099
	Overmature	0.97	1.00	1.00	0.99
Overall model performance	Micro average	0.99	0.99	0.99	0.98
	Macro average	0.99	0.99	0.99	0.98
	Weighted average	1.00	0.99	0.99	0.99

Table 3.6 shows that the precision and recall[18][19] values for maturity stage classification of both the species are varying between species. The range of values are 0.96 – 1 which shows the significant potential of the model. As it can be seen that for premature cases the recall values are comparatively lower which indicates the chances of some of the samples which does not belong to premature stages have been identified as premature. However, in terms of percentage this is not more than 4%. On the other hand the F1-score[20] is 98% and more which reflects good balance between specificity and sensitivity of the model. It can vouch the fact the model is able to classify species and maturity stages simultaneously with 98% accuracy.

3.4.3. Cohen kappa coefficient score

Cohen kappa metric value [21] is one of the indicators to judge the suitability and goodness of the automatically extracted features by the CNN model. The score for TomNet across species and maturity stage classifications have been shown in Table 3.7. It shows that the Cohen Kappa values for both the classification tasks is in the tune of 0.98 which ensures the appropriateness of the automatically extracted features of TomNet.

Table 3.7: *Cohen kappa scores for TomNet*

Classification types		Cohen Kappa Score
Tomato classes	**Maturity type**	
Pear	Premature	
	Mature	0.98
	Overmature	
Roma	Premature	
	Mature	
	Overmature	

3.4.4. Receiver operating characteristics (ROC) curve

The ROC curves [22] for different tasks performed by TomNet have been consolidated in Fig. 3.8.This curve also reflects that the model can potentially classify between maturity classes of different species with more than 97% accuracy. It can be noted that the classification potential for over-maturity stage of both the species are maximum which is significant for inventory and supply chain management. The over-mature fruits can be subjected for immediate consumption and food processing industries for minimizing the agro loss.

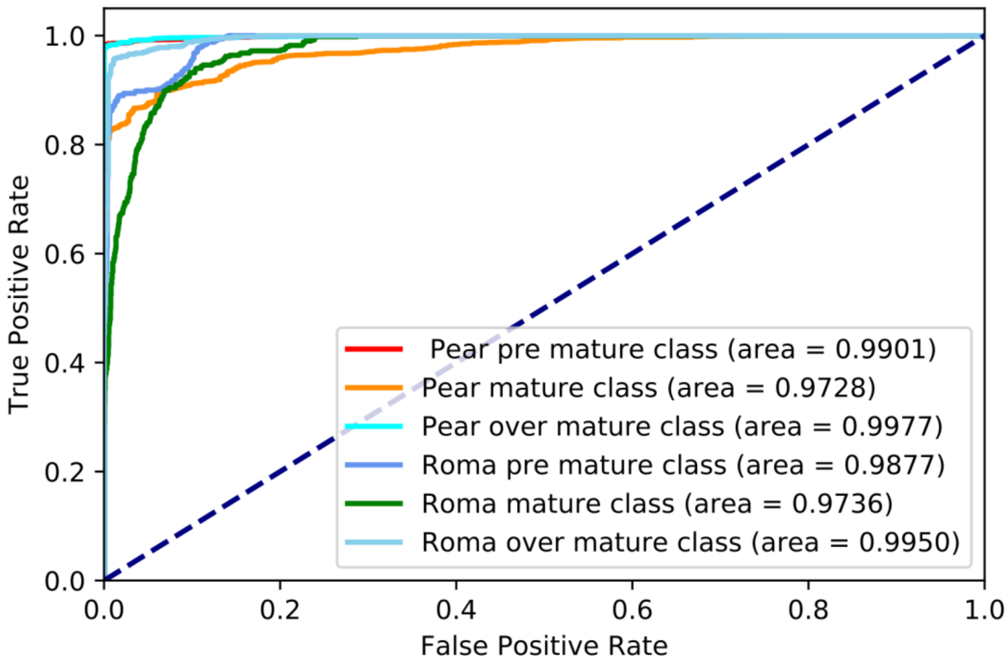


Fig. 3.8: ROC-AUC for different classification tasks usingTomNet

3.4.5. Comparative analysis

The comparative analysis between the performance of TomNet and some of the conventional classifiers like artificial neural network with backpropagation multilayer perceptron architecture (BPMLP)[23] ,support vector machine (SVM)[24] and random forest (RF) along with the deep learning principles based models AlexNet[25],VGG19[26],ResNet50[27] as well as pre-trained CNN models has been presented in Table 3.8. Like the medicinal leaves here also the color features of different color channels were used as featureset for conventional classifiers that need hand engineered features. Each of the models have been used comparatively for the classification of species and maturity both simultaneously.

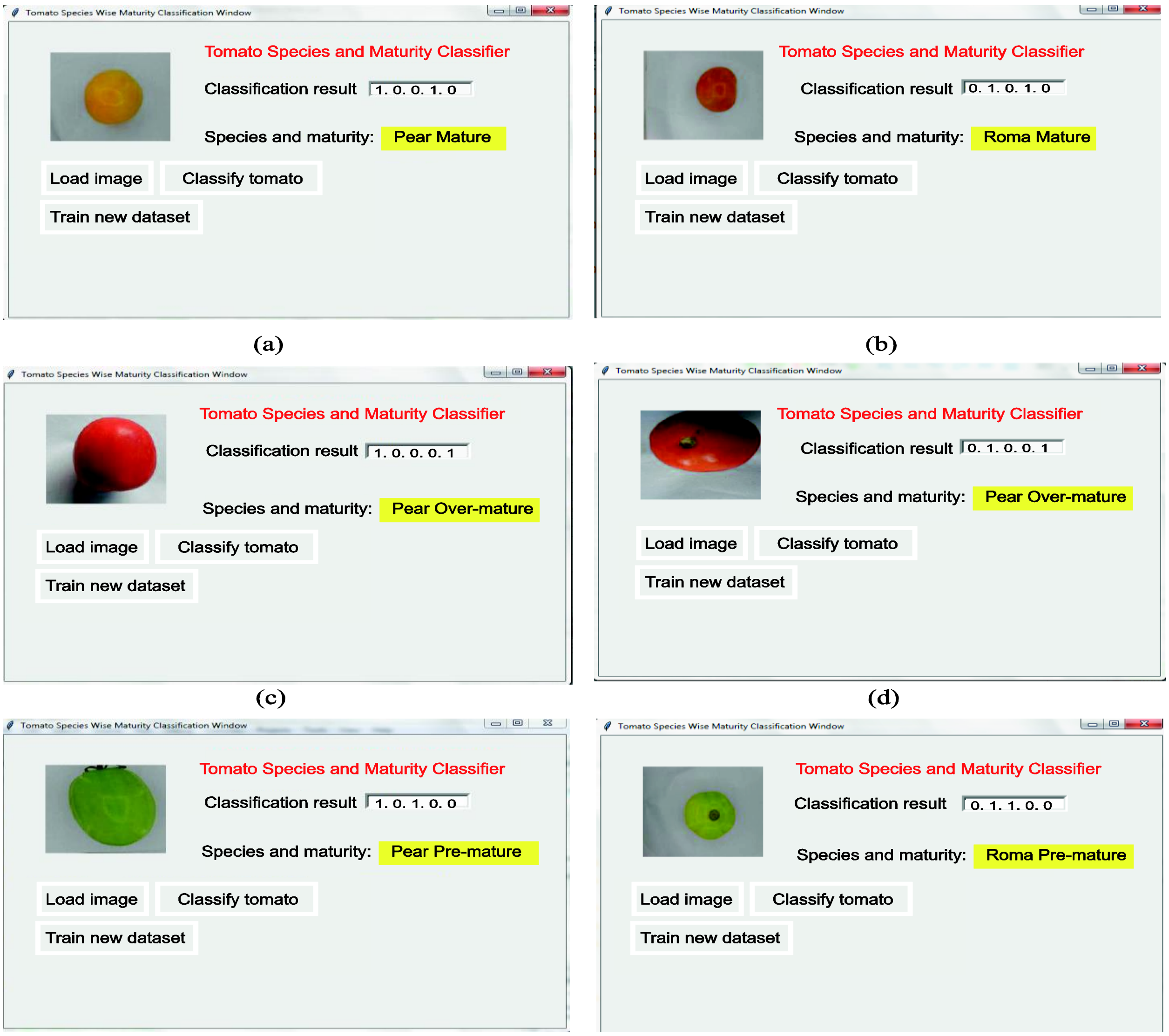
Table 3.8: *Comparative performance evaluation of TomNet*

		Species classification/Maturity stage classification				
Method	**Classification**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**FMeasure**
BP MLP	Species and Maturity	0.97	0.85	0.96	0.96	0.97
SVM	Species and Maturity	0.96	0.95	0.88	0.96	0.95
RF	Species and Maturity	0.95	0.92	0.94	0.95	0.95
Deep learning based						
AlexNet	Species and Maturity	0.75	0.80	0.75	0.76	0.79
VGG19	Species and Maturity	0.88	0.77	0.87	0.82	0.78
ResNet50	Species and Maturity	0.76	0.79	0.77	0.79	0.80
TomNet model optimized with GWO	Species and Maturity	0.98	0.97	0.97	0.98	0.98
TomNet model optimized with modified GWO	Species and Maturity	0.98	0.98	0.98	0.98	0.98
TomNet model optimized with BPSO	Species and Maturity	0.99	0.99	0.98	0.99	0.99

Table 3.8 shows that the TomNet model outperforms other classifiers under consideration. The performance of the RF model is not as good as BPMLP and SVM.BPMLPP and SVM performs equivalently .The values of respective metrics ,carried out by the deep learning based classifiers AlexNet,VGG19 and ResNet50 are not also as good as those presented in the model TomNet.The CNN model TomNet not only outperforms the other models it also eliminates the feature extraction step. Hence the CNN based modelTomNet has been considered as the more potential in comparison to the others. It is clearly visible from the performance presentation of the optimized TomNet, that GWO optimized TomNet and the modified GWO optimized TomNet are analogous in terms of both the species and maturity classification. The BPSO has slightly higher impact on performance metric of the CNN based TomNet model compared to the GWO and the modified GWO approach.

3.4.6. GUI presentation

Fig. 3.9 presents GUI implementation of the species-wise maturity information for unknown tomato sample belonging to any of the six categories namely premature, mature and overmature types of roma and peer tomatoes. The prediction regarding the correct species wise maturity states has been expressed by the corresponding prediction vectors.The prediction vectors for premature,mature and overmature states of the pear and roma species are expressed as [1.0.1.0.0],1.0.0.1.0], [1.0.0.0.1], [0.1.1.0.0], [0.1.0.1.0] and [0.1.0.0.1]respectively for the unknown sample.



(a) (b) (c) (d) (e) (f)

Fig. 3.9: GUI windows corresponding to the tomato class and maturity identification for any unknown sample identified as (a) premature pear (b) mature pear (c) overmature pear (d) premature roma (e) mature roma and (f) overmatureroma.

3.5. Conclusions

This chapter has presented a promising computer vision framework based on CNN architecture towards classification of species and maturity of tomato fruit. The details of variation of different bio-actives and nutrients present in the fruit and their accountability towards the maturity wise discrimination has been examined in justification to the problem. The development of an indigenous imaging chamber and its application in capturing images of tomato fruit have been elaborated in this chapter. The chapter includes the procedure starting from the sample collection to CNN model TomNet development. It also highlights the use of BPSO, GWO and modified GWO optimization for hyperparameter optimization in order to get more generalization and performance of the TomNet model. The performance measurement of the BPSO algorithm has shown the slight improved performance towards the accuracy of calculation. The validation of the model and classification results with unknown testing set has been evaluated against standard metrics which show that the presented method can achieve up to 99% accuracy. The comparative analysis has also been presented towards justification of using the model TomNet over some of the conventional classifiers including deep learning based models. On an overall assessment, by virtue of significant performance potential of the presented framework, the same can be considered as an addition to the existing instrumental and analytical screening and sorting methods of tomato fruits of different species and maturity grades.

References

1) Bhowmik D., Kumar K.P.S., Paswan S.. Srivastava S. (2012), Tomato-a natural medicine and its health benefits. J. Pharmacogn. Phytochem.1(1), 33–43.

2) Saad A.M.Ibrahim A., El-Bialee N. (2016). Internal quality assessment of tomato fruits using image color analysis. Agric. Eng. Int.: CIGR J. 18, 339–352.

3) Suarez M., Rodríguez R., Elena M., DíazC. (2008). Chemical composition of tomato (Lycopersiconesculentum) from Tenerife, the Canary Island. Food Chem., 106, 1046-1056.
4) Gómez A.H., Hu G., Wang J., Pereira A.G. (2006). Evaluation of tomato maturity by electronic nose. Comput Electron Agric 54(1),44–52.
5) Zhang L., McCarthy M.J.(2012). Measurement and evaluation of tomato maturity using magnetic resonance imaging. Postharvest BiolTechnol 67,37–43.
6) Qin J., Chao K., Kim M.S.(2012). Nondestructive evaluation of internal maturity of tomatoes using spatially offset Raman spectroscopy. Postharvest BiolTechnol 71, 21–31.
7) Sirisomboon P., Tanaka M., Kojima T., Williams P. (2012). Nondestructive estimation of maturity and textural properties on tomato ‘Momotaro’ by near infrared spectroscopy. J Food Eng 112(3),218–226.
8) Tiwari G., Slaughter D.C., Cantwell M. (2013). Nondestructive maturitydetermination in green tomatoes using a handheld visible and nearinfrared instrument. Postharvest BiolTechnol86 , 221–229.
9) Goel N., Sehgal P. (2015). Fuzzy classification of pre-harvest tomatoesfor ripeness estimation—an approach based on automatic rule learning using decision tree. Appl. Soft Comput. 36, 45–56.
10) Greenshields M.W.C.C. et. al. (2015). Non-destructive tomato ripeness determination using carbon nanostructure based chemical sensors: Application of tristimulus and principal component analysis methods. International Journal Electroactive Materials 3:10–14.
11) Indolia S., Goswami A.K., MishraS.P.,AsopaP. (2018). Conceptual understanding of convolutional neural network – a deep learning approach, International Conference on Computational Intelligence and Data Science (ICCIDS), Procedia Computer Science, Vol. 132, pp.679–688, DOI: doi.org/10.1016/j.procs.2018.05.069.
12) Yamashital R., Nishio M., DoR.K.G., Togashi1K.(2018), Convolutional neural networks: an overview and application in radiology, Insights into Imaging ,pp.611–629.

13) Traorea B.B., Foguema B.K.,Tangara F.(2018). Deep convolution neural network for image recognition, Ecological Informatics, Vol.48. pp..257-268.

14) Wang Y., Zhang H., Zhang G. (2019),cPSO-CNN: An efficient PSO based algorithm for fine-tuning hyper-parameters of convolutional neural networks. Swarm EvolComput 49:114–123.

15) Nadimi-Shahraki M.H., Taghian S.,Mirjalili S.(2020). An improved grey wolf optimizer for solving engineering problems, Expert Systems with Applications, vol. 166, p. 113917.

16) Berrar D. (2018). Cross-validation. Encyclopedia of Bioinformatics and Computational Biology, 1, Elsevier, 542 -545.

17) Santra A.K., Christy C.J. (2012). Genetic algorithm and confusion matrix for document clustering. International Journal of Computer Science Issues (IJCSI), 9(1), 1694-0814.

18) Hossin M., Sulaiman M. (2015).A review on evaluation metrics for data classification evaluations. Int. J. Data Min. Knowl. Manage. Process, 5(2).

19) Silva V.A., Bittencourt II., Maldonado J.C. (2019). Automatic question classifiers: asystematic review. IEEE Trans. Learn. Technol.

20) Sharma K.K., Seal A.(2019). Modeling uncertain data using Monte Carlo integration method for clustering. Expert Syst. Appl. 137, 100–106.

21) McHugh M.L. (2012).Interrater reliability: the kappa statistic, BiochemiaMedica. 22(3), 276-82.

22) Fawcett T. (2006).An introduction to ROC analysis. Pattern Recognition Letter, 27, 861–874.

23) Kanungo D.P., Naik B., Nayak J., Baboo S., Behera H.S. (2015). An improved PSO based back propagation learning-MLP (IPSO-BP-MLP) for classification, Computational

24) TianY. ShiY.,LiuX.(2012).,Recent advances on support vector machines research ,Technological and Economic Development of Economy, Vol. 18, No. 1, pp.5–33 [online] ISSN: 2029-4913, [print] ISSN: 2029-4921.

25) Liu Y., Li Z., Chen X., Gong G., Lu H.(2020). Improving the accuracy of SqueezeNet with negligible extra computational cost, *International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 1-6, doi: 10.1109/HPBDIS49115.2020.9130577.

26) Zhang X., Zhou X., Lin M., Sun J., (2018), ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices, Conference, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

27) Dhankhar P. (2019). ResNet-50 and VGG-16 for recognizing Facial Emotions.International Journal of Innovations in Engineering and Technology (IJIET) http://dx.doi.org/10.21172/ijiet.134.18.

CHAPTER 4

SPECIES AND MATURITY GRADING OF FRUITS – CASE STUDY WITH BANANAS

4.1 Introduction

Fruits are the major contributor of human nutrients. It is consumed across the world for protection against diseases and boosting the immunity. About 75 % of world economy depends on export and import of fruits. Apart from raw consumption, fruits are also used for processed foods and juices. They provide vitamins, minerals and carbohydrates. Different fruits have different bio-chemical compositions and they help human health accordingly. Grading of fruits in terms of species and maturity stage thus is an important research area. Among different fruits this work focuses on bananas which are one of the most popular fruits across the world. Bananas are available in large varieties and cultivated in most of the countries. However, Asian countries like India, Srilanka, Japan and China are some of the major suppliers of bananas.

Banana has about 15% share of food production market according Food and Agriculture Organization of the United Nation (FAO) [1]. The high fibre content of bananas helpsin improved functioning of digestive system. It is also rich in different minerals like potassium, calcium,manganese, magnesium, iron, folate, niacin, riboflavin and vitamin B6 those too contribute towards improved body functioning [2].

The work presented in this chapter has been published as:

Mukherjee G, Tudu B , Chatterjee A.(2022). A multi-channel convolution neural network driven computer vision system towards identification of Species and maturity stage of banana fruits: case studies with Martaman and Singapuri banana, Int. J. Computational Intelligence Studies, Vol. 11, No. 1,pp.1–23.

It has proven protective characteristics against heart, blood pressure related diseases and diseases like anemia. Bananas belong to Musa family and available in number of varieties. Among several species Singapuri and Martaman are two widely consumed varieties in India.These variants are visually different for instance Singapuri appears green while Martaman appears yellow however the tonal changes happen with the ripening stages for both the species.

The bio-chemical composition of bananas changes with maturity like all other fruits and vegetable. Hence, identification and grading of bananas plays important role in terms of market supply chain and inventory management. Another important factor for banana is it has different usages for different maturity stages, for example, unripen bananas are often consumed in fried snacks form while ripen and slightly over-ripen are consumed directly. Over-ripen but not rotten bananas are used for different processed food preparation like confectionaries and bakery items. One of the most visible change occurs in peel color of banana with change of maturity stage. Brown spots due to enzymatic reactions on the banana skin are also an important indicator of maturity and days-to-rot. Although manual method of grading is mostly used in local market it is limited by its grading and sorting speed and subjective nature of operations. Some of the reported methods to address this manual process have been briefly reviewed in Table 4.1.

Table 4.1: *Some of the reported techniques for automatic grading of bananas*

Base technique	Method	Result
The Computer vision techniques for ascertaining ripening stages of banana. [1]	Four class of homemade dataset of the banana have been used based on the artificial neural network backed frame work. The feature set used for the experiment were thecolour,	The accuracy value for the classification process of different maturity grades of the banana has been reported at 97.75%

	brown spots and tamura statistical textural features.	
Image processing techniques in order to get the different maturity stages of-banana on the basis of their colour and size values. [3]	The under mature, mature and over mature banana fruits have been subjected to the classifier on the basis of size and colour value features.	The mean colour intensity algorithm has been reported at 99.1% and 85% accurate at differentiating the under mature banana.
Visible short wavelength near infrared spectra for the assessment of sugar and starch content in the banana. [4]	Banana pulp was used using the partial transmittance optical geometry.	The accuracy value of 94% has been reported for the green ripening ripe and overripe stages of banana.
BET monolayer procedure in studying Hygroscophic behaviour of the banana for the desorption analysis.[5]	Nanicao banana flours were used for the study of adsorption and desorption isotherms property by means of the BET and GAB model.	The estimated value for the coeficient of determination has been found to be 0.97 ie the efficiency of the model has been expressed to a higher degree.
The Convolutional Neural () Network(CNN) in identification of ripeness of the banana [6]	The four classes(unripe, through ripe, almost ripe and ripe) of ripening stages of the cavendishbanana has been identified with the CNN architecture with the SGD parameter optimisation.	The 94.12% accuracy has been achieved for training data and 72% accuracy value has been achieved for the test data.
partial least square discrimination analysis for the maturity of banana . [7]	The hyperspectral imaging of the green bananas in different maturity stages were subjected to the partial least square regressions and interval PLS	The reported accuracy value for the maturity prediction of the banana fingers was nearly 100%.

	methods.	
Arduino device backed automated tool has been appointed to assess the maturity state of the banana fruit[8].	The barlin banana with four different types of maturity levels have been used to detect different maturity stages by the principles of image processing.	The accuracy prediction has been found to be nearly 100%.
The Convolution Neural Network(CNN) in classification of the different ripening stages of the banana fruit[9].	The data of different ripening stages of the banana fruit has been used for classification using the CNN architecture.	The accuracy value for the determination of maturity value has been found very close tobe 92%
zNose in classification of banana fruits based on the ripening and maturity of the fruit [10].	The PCA technique has been used to identify the ripening stages of the banana fruit. The fruit quality parameters were tracked by using the Multiple Linear Regression(MLR)	The coefficient of determination value was estimated to be at 96%.

Table 4.1 shows that computer vision due to its non-invasive, less-expensive and portable nature of operations can be a possible alternative for fruit grading as well. Despite several works the room of improvement and employment of CNN [11] architecture has become different from ones that have been reported are still open. This work explored the scope of a custom designed CNN architecture for simultaneous classification of species and maturity stages of bananas. The experimentation have been performed with an indigenous dataset prepared with banana samples procured from local market. The hyperparameter optimization has as well been presented. The model validation and performance analysis of the model for banana classification and grading have been presented in details.

4.2. Dataset preparation

4.2.1. Sample collection

Bananas of two species under consideration were procured from local market and in the same season to avoid any major seasonal effects. For each species 1500 bananas were procured comprising of 500 bananas for each of the three maturity stages namely, premature, mature and over-mature. This made the total samples of 300 bananas. Like tomatoes in this case also the bananas were preserved in controlled environment at 20ºC till they were subjected to imaging chamber.

4.2.2. Image acquisition

The indigenous imaging chamber as described in section 2.2.2 was used for this work. For better observations and experimentation flexibility the publicly available datasets were not used. The imaging policy was same as that for tomatoes as described in 3.2.2 and the same isolation and background elimination method described in Table 3.1 was adopted for this case as well. Nevertheless banana has separate color channel dynamics hence the adaptive thresholding process for this case has been elaborated in following section. Some of the examples of captures and isolated banana images have been presented in Fig. 4.1.

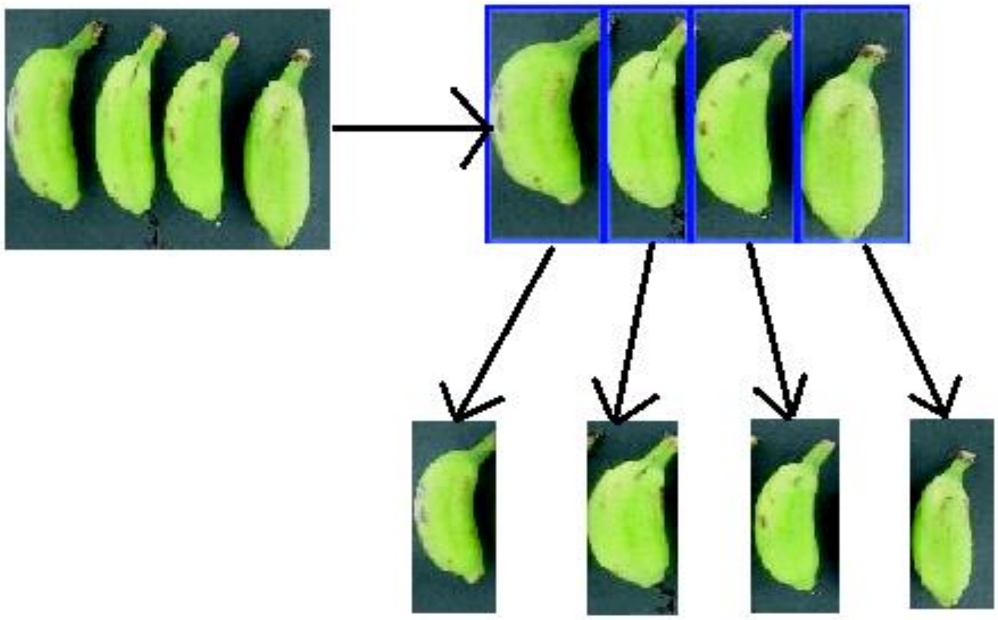


Fig. 4.1: *Examples of isolated and background subtracted banana images*

4.2.2.1. Channel selection and thresholding for image isolation

The skin color of bananas is yellow dominated. Hence, the R and G channels contain the major information while blue (B) channel is expected give much differentiable peaks. This is evident from the channel-wise images of bananas shown in Fig. 4.2(a) and (b) for Martaman and Singapori, respectively. The color-channel images have been studies using the histogram distributions and shown beside the respective color channels. It can be seen that R and G channel histograms can be suitable candidate to differentiate the background from bananas as these histograms are visibly bimodal. Between R and G, later one show better valley between the peaks while in case R channel the rise of second peak is not that well separable. Hence, for this work the G channel was chosen and subjected to adaptive thresholding using Otsu's algorithm. Randomly 600 samples were chosen with normal distribution and the average of the adaptive threshold values were almost 100. Thus for this work the threshold val ue 100 were used in the process described in Table 3.1 Step 3.

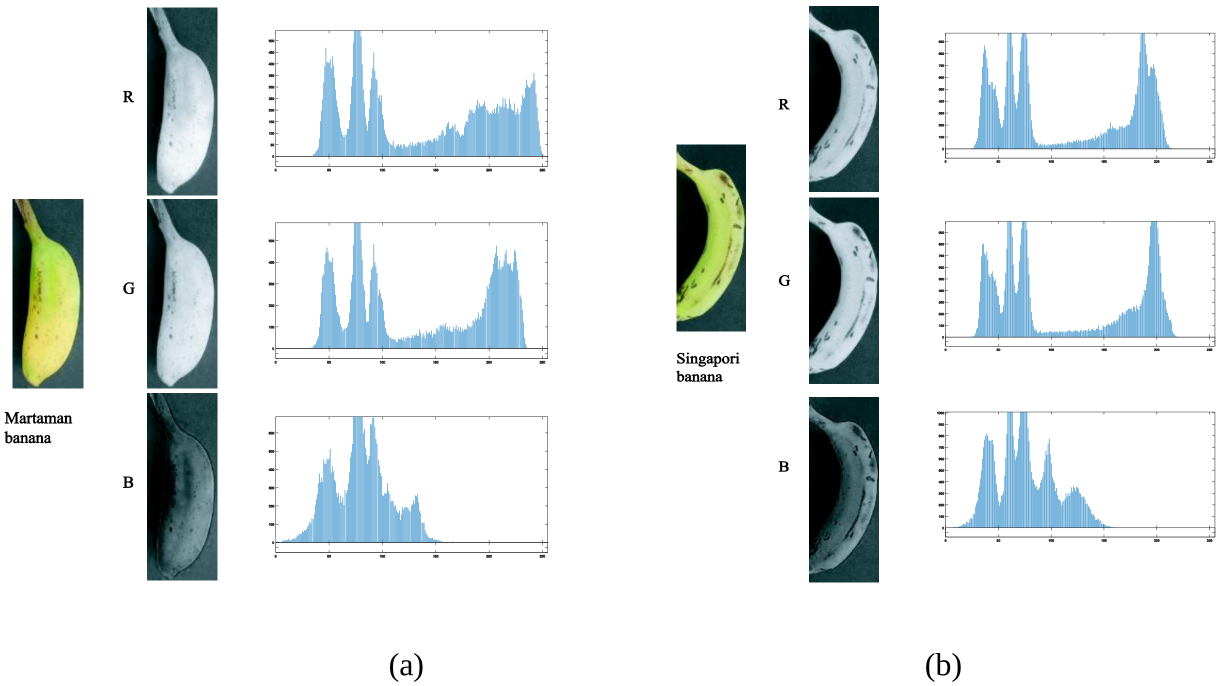


Fig. 4.2: *Color channel dynamics for (a) Martaman and (b) Singapori banana*

4.3. CNN model for species and maturity stage classification of bananas

In this work a multi-channel convolution network has been developed and explored for simultaneous detection of species and maturity stages. In case of TomNet there were 6 different nodes for 6 classes where 3 nodes were assigned for 3 maturity stages for each of the 2 species. Here, 2 different CNN architectures have been collaborated in channel-fashion, thus one channel provides the classification of species and another maturity stage as presented schematically in Fig. 4.3. The layer characteristics of convolution, pooling, batch normalization and fully connected remains same as described in section 2.3. This model uses another important mechanism called 'dropout' which is an important regularization technique to avoid over-fittingproblem. In this model these layers have been placed after the convolution layer but they can be accommodated after the pooling layer as well. In this technique during training some of the nodes' outputs are discarded in forward pass hence not updated during the weight update of backward pass. This provides faster training with higher generalization towards identification of unknown samples that were not included during training of the CNN [11]. The mathematical dynamics of the layers have been consolidated in Table 4.2 and the layer-wise feature extractions can be visualized as given in Fig. 4.4.

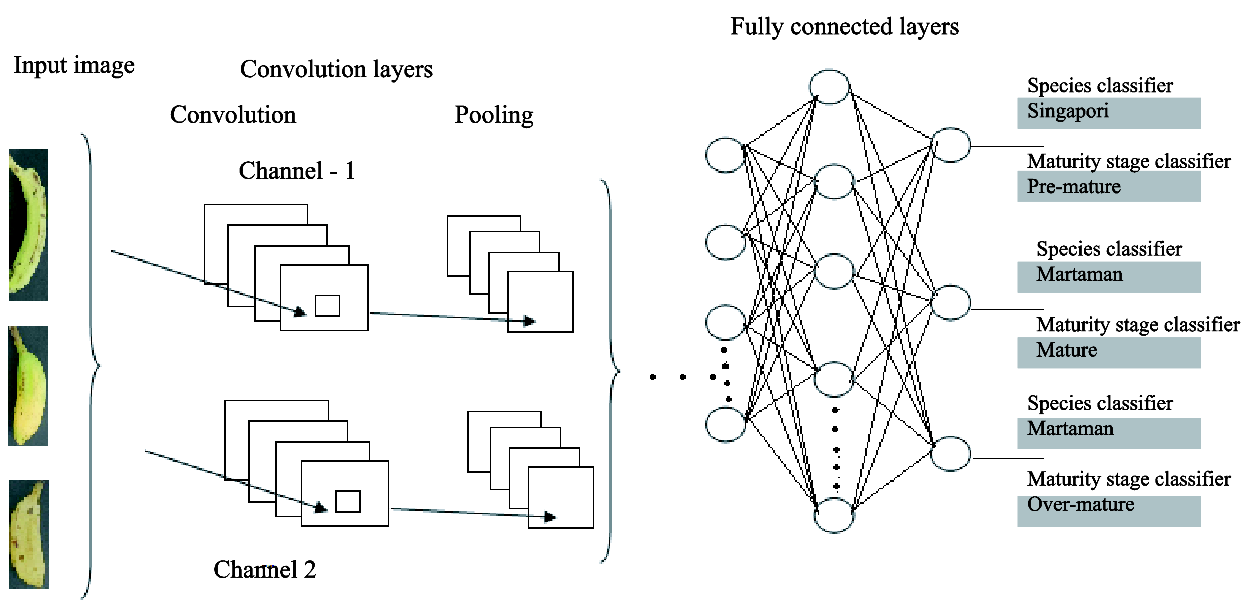


Fig. 4.3: *Schematic representation of multi-channel CNN for banana species and maturity stage classification.*

Table 4.2:*Multi-channel CNN layer dynamics (Common layer presentation for both the channels)*

Layer type	Purpose	Mathematical expression
Convolution layer Conv_2d	Feature extraction through the non-linear activation function [12]	$Y_k = f(W_k * x)$ Y_k - feature map of k^{th} filter W_k - convolutional filter x – input image $f(\cdot)$- RelU $R(z)=\max(0,z)$
Pooling layer max_pooling2d	Reduces the spatial resolution of the feature maps followed by the introduction of the translation invariance to the small shifts and distortions which results into the reduced number of parameters to be learned [13].	$Y_{kij} = \max_{p,q \in R_{ij}}(X_{kpq})$ Y_{kj} - max pooled output k^{th} feature map X_{kpq} - element at the (p,q) location contained in the pooling region R_{ij} - the receptive field around the location (i,j)
Dense layer	Fully connected layer for final decision with non-linear *softmax* activation function[14].	$p(y=j \mid \Theta^{(i)}) = \frac{e^{\Theta^{(i)}}}{\sum_{j=0}^{k} e^{\Theta_k^{(i)}}}$ $p(i,j)$- probability of belongingness to j^{th} class among total k classes $\Theta = \sum_{i=0}^{k} w_i x_i$

4.3.1. Hyperparameter optimization

The hyperparameters of the multi channel CNN model have been optimized by using the three most promising methods BPSO, GWO and the modified GWO . The architectural description of the appointed models has been described in the section 2.4 in details. The comparative analysis of the three optimizer has been carried out in terms of the convergence curves which isprovided in the result and discussion section of this chapter.

4.4. Results and discussions

Following the convention in this case also the entire dataset of 3000samples were divided into 60:20:20 ratios for training, validation and testing purposes. For optimization 30% of entire dataset images were randomly selected with normal distribution. The convergence plot based on the BPSO, GWO and modified GWO for 100 iterations are shown in Fig. 4.4.

The plot of Fig. 4.4 shows that after 80 iterations the BPSO [15] converges with more than 99% accuracy, the GWO[16] converges with a little more than 98% accuracy and for modified GWO, there is no visible improvement in the convergence characteristics notified. The hyperparameter finally arrived with BPSO as used for the presented results are consolidated in Table 4.3.

The optimized hyperparameter settings have been presented in Table 4.3 and these have been used for all the results presented in this section. The layer-wise output have been consolidated in Fig. 4.5 which shows that as the image progresses through the deeper layer more abstract color and texture features are extracted.

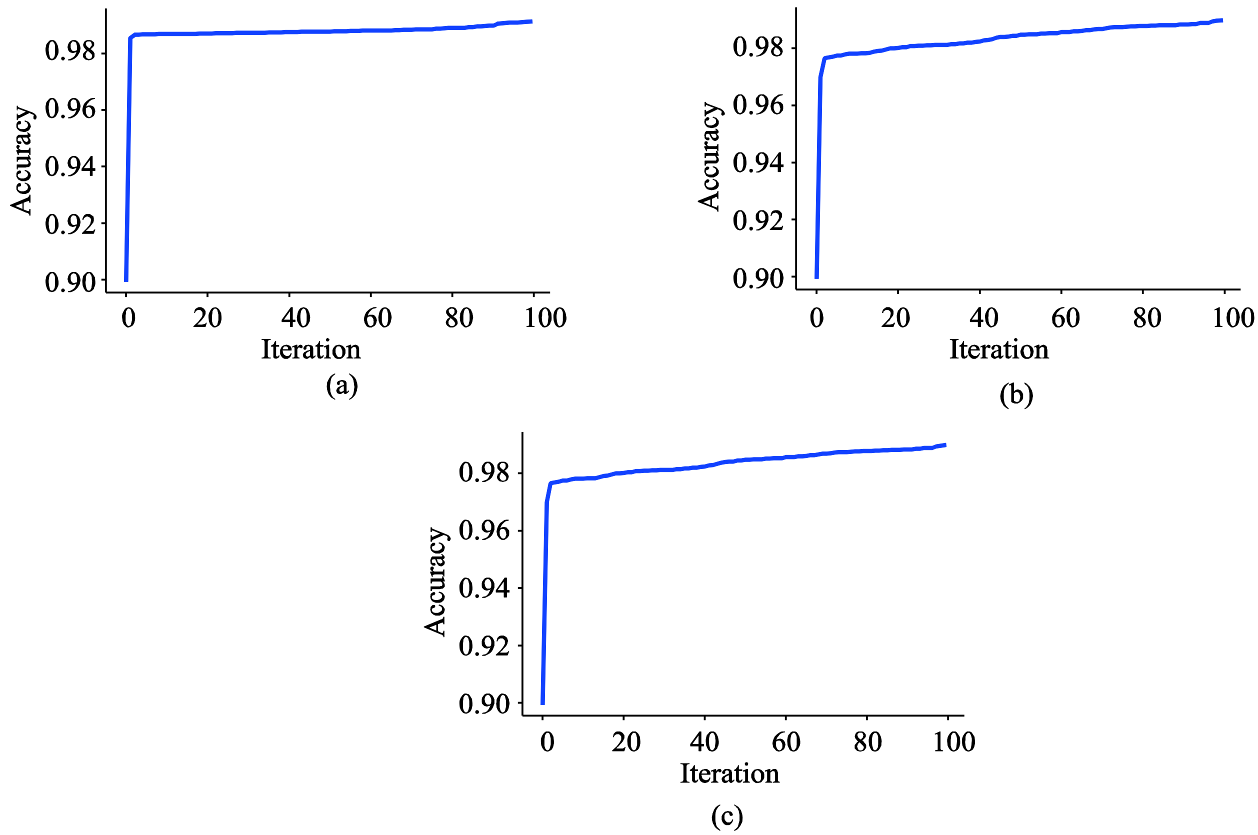


Fig. 4.4: *Convergence plot for (a) BPSO,(b) GWO and (c) modified GWO optimizer*

Table 4.3: *Hyperparameter settings of multichannel CNN*

Channel type	Layer(type)	Output shape	Number of parameters
Channel -1	**Channel1 (InputLayer)**	**(None, 32, 32, 3)**	**0**
Species classification	conv2d_1 (Conv2D)	(None, 30, 30, 32)	896
	dropout_1 (Dropout)	(None, 30, 30, 32)	0
	max_pooling2d_1 (MaxPooling2D)	(None, 15, 15, 32)	0

	batch_normalization_1 (BatchNormalisation)	(None, 15, 15, 32)	128
	conv2d_2 (Conv2D)	(None, 13, 13, 64)	18496
	dropout_2 (Dropout)	(None, 13, 13, 64)	0
	max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 64)	0
	batch_normalization_2 (BatchNormalization	(None, 6, 6, 64)	256
	flatten_1 (Flatten)	(None, 2304)	0
Channel-2	**Channel2 (InputLayer)**	**(None, 32, 32, 3)**	**0**
Maturity stage classification	conv2d_3 (Conv2D)	(None, 30, 30, 32)	896
	dropout_3 (Dropout)	(None, 30, 30, 32)	0
	max_pooling2d_3 (MaxPooling2D)	(None, 15, 15, 32)	0
	batch_normalization_3 (BatchNor	(None, 15, 15, 32)	128
	conv2d_4 (Conv2D)	(None, 13, 13, 64)	18496
	dropout_4 (Dropout)	(None, 13, 13, 64)	0
	max_pooling2d_4 (MaxPooling2D)	(None, 6, 6, 64)	0
	batch_normalization_4 (BatchNormalization	(None, 6, 6, 64)	256
	flatten_2 (Flatten)	(None, 2304)	0

concatenate_1 (Concatenate)	(None, 4608)	0
dense_1 (Dense)	(None, 128)	589952
dense_2 (Dense)	(None, 3)	387

Total params: 629,891

Trainable params: 629,507

Non-trainable params: 384

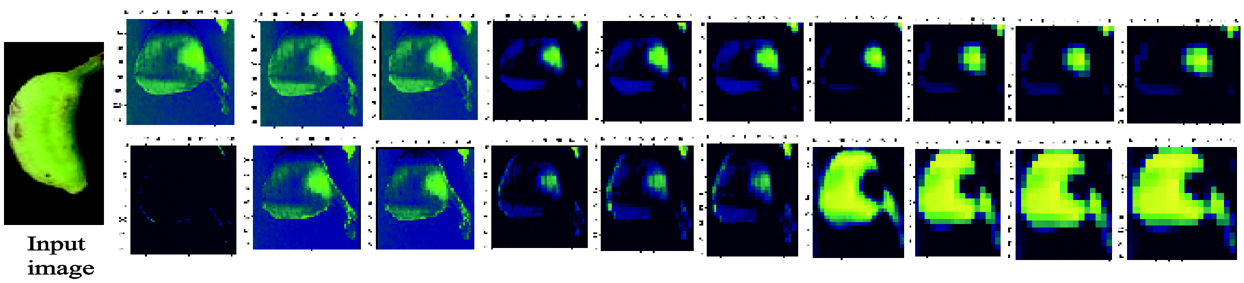


Fig. 4.5: *Visual representation of feature extractions from different CNN layers.*

Based on the convergence characteristics as given in the Fig 4.4 , the BPSO optimized TomNet model has been chosen as the classifier on the convergence characteristics performance shown in fig 4.4 for carrying out the species and maturity stage classification of the fruit banana and all other metrics have been measured based on this chosen best.

4.4.1. Model validation

Model validation was performed using validation dataset (3000 samples) and the accuracy and loss plots for training and validation have been presented in Fig. 4.6. Since there are two distinct channel the validation of both channels have been presented separately. Fig. 4.6(a) and (b) corresponds to species classification while Fig. 4.6(c) and (d) corresponds to maturity classification channel. It can be observed that both the channels provide about 98% accuracy with training set while 96% and 94% accuracy for species and maturity stage, respectively. In terms of loss species classifier can provide low value ($<10^{-2}$) and maturity classifier (~ 0.10) for training set. The values for validation sets are also considerably low in the tune of 0.10 and 0.15 for species and maturity classifier, respectively. It can be also seen that the difference between training and validation plots are considerably less at the end of epochs which convey the generalization potential of the model.

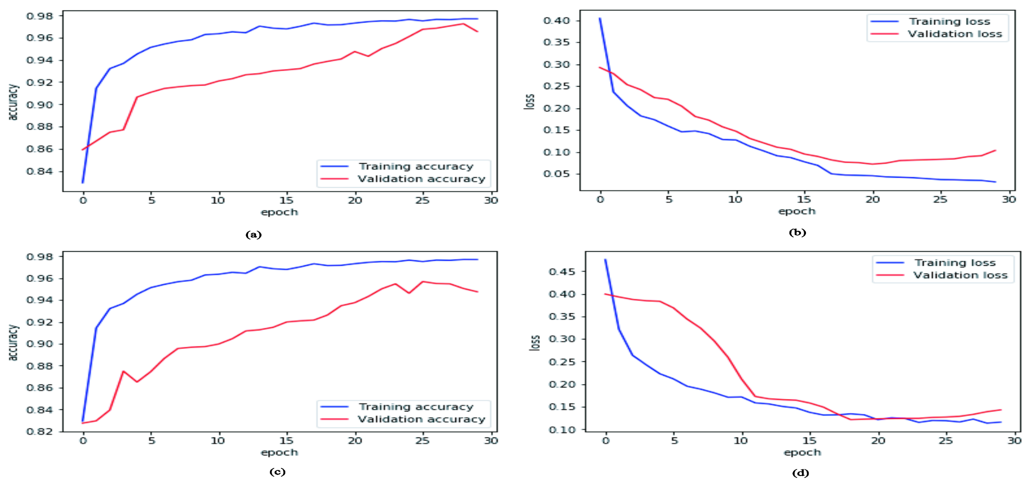


Fig. 4.6: *Model validation for testing and validation set using accuracy and loss values*

The 10-fold cross-validation [17] results have been provided in Table 4.4 and 4.5 for species and maturity channels, respectively. This time also the validation results show consistency between different folds with an average of 99% and about 98% accuracy for training and validation set, respectively. In case of loss values as well the performance potential is visible with low average values of 0.019 and 0.045 for training and validation set, respectively. The consistency of the model performance is reflected in the small standard deviation values for loss and accuracy in case of the training and validation data. The standard deviation is in the tune of 10^{-3} across all the cases which is significant towards the promising potential of the model across different folds.

__Table 4.4:__ 10-fold cross-validation results for species classification

Number of folds	Training data		Validation data	
	Loss	**Accuracy**	**Loss**	**Accuracy**
1	0.0199	0.9888	0.053	0.9766
2	0.0276	0.9934	0.033	0.9734
3	0.0235	0.9889	0.034	0.9756
4	0.0156	0.9843	0.054	0.9788
5	0.0246	0.9945	0.044	0.9792
6	0.0136	0.9803	0.051	0.9779
7	0.0175	0.9899	0.032	0.9785
8	0.0166	0.9985	0.043	0.9755
9	0.0126	0.9875	0.056	0.9753
10	0.0167	0.9945	0.047	0.9752
Average	0.0188	0.9901	0.0447	0.9766
Maximum	0.0276	0.9985	0.056	0.9792
Minimum	0.0126	0.9803	0.032	0.9734
Std. Dev	0.0049	0.0054	0.0091	0.0019

Table 4.5: *10-fold cross-validation results for maturity classification*

Number of folds	**Training Data**		**Validation data**	
	Loss	**Accuracy**	**Loss**	**Accuracy**
1	0.0301	0.9877	0.053	0.9771
2	0.0150	0.9899	0.033	0.9794
3	0.0135	0.9898	0.034	0.9697
4	0.0215	0.9937	0.054	0.9798
5	0.0204	0.9881	0.044	0.9699
6	0.0253	0.9965	0.051	0.9761
7	0.0175	0.9823	0.032	0.9784
8	0.0102	0.9993	0.043	0.9786
9	0.0166	0.9984	0.056	0.9798
10	0.0235	0.9885	0.047	0.9789
Average	0.0194	0.9914	0.0447	0.9767
Maximum	0.0301	0.9993	0.056	0.9798
Minimum	0.0102	0.9823	0.032	0.9697
Std. Dev.	0.0059	0.0054	0.0091	0.0039

4.4.2. Performance evaluation against confusion matrix

The confusion matrix [18] for model performance with unknown test dataset has been presented in Fig. 4.7. In this case also two different confusion matrices have been presented to evaluate the performance of the model in individual channel. It can be seen that for species classification only 14 samples out of 600 unknown samples have been misclassified. The misclassification rate for both the species are almost equal and that is only 2% hence about 98% cases the species classifier channel could classify between species accurately. In case of maturity stage classifier the performance is equally promising. For premature cases the misclassification is highest but that too only 3% while for mature and over-mature that is 1% and 2%, respectively. Such low misclassification rates show about 98% potential of the maturity stage classifier to classify the maturity stage accurately. The different metrics in association to the confusion matrix have been consolidated in Table 4.6 and 4.7 for species and maturity stage channel, respectively.

Predicted

Actual

	Martaman	**Singapuri**
Martaman	294	6
Singapuri	8	292

Species classification

Predicted

Actual

	PM	**M**	**OM**
PM	97	1	2
M	0	99	1
OM	1	1	98

Maturity stage classification

Fig. 4.7: *Confusion matrix for species and maturity stage classification*

Table 4.6: *Classification performance of species classifier channel*

	Precision	Sensitivity/recall	FMeasure	Specificity
Martaman	0.98	0.99	0.98	0.98
Singapuri	0.99	0.98	0.99	
Micro avg	0.98	0.98	0.98	
		0.98		
Macro avg	0.98	0.98	0.98	
Weighted avg	0.99	0.99	0.98	

Table 4.7: *Classification performance of maturity stage classifier channel*

	Precision	Sensitivity/ Recall	FMeasure	Specificity
Premature	0.98	1.00	0.98	0. 98
Mature	1.00	0.98	0.98	
Overmature	1.00	1.00	1.00	
Micro Avegrage	0.99	0.99	0.99	
Macro Average	0.99	0.99	0.99	
Weighted avg	1.00	0.99	0.99	

Table 4.6 conveys that for both the species the classifier performs with high sensitivity and specificity values that reflect the ability to identify the correct specie of banana as well as identification of different species of banana mixed with one species. It can also be noted that the precision of identifying Singapori is slightly higher while recall is slightly lower compared to Martaman. The reason is the mature stage of both the banana species when both the bananas turns yellow so color features do not contribute much in classification. Nevertheless, the precision and recall values are never below 98% which vouches the potential of the multi-channel CNN in species classification task. Similarly for maturity classification the performance presented in the Table 4.7 is considerably appreciable. At over-mature stage the peel characteristics of the bananas for these two species are quite different hence about 100% classifications could be achieved. For premature and mature stages the distinguishability is quite difficult particularly for Martaman bananas as in both the stages the overall appearance is yellow. Despite such difficult distinguishability the CNN model could identify the maturity stage correctly for more than 98% cases.

4.4.3. Cohen kappa score

The Cohen Kappa [19] score for different channels are presented in Table 4.8. It shows that in both the cases the scores are in the tune of 0.99. As this metric is one of such metrics that convey the suitability of the automatically extracted features by CNN, it can be assessed that despite two different tasks the single model could extract features that are suitable for different channels.

Table 4.8: *Cohen Kappa scores for different channels*

Classifier	Classes taken for measurement	Cohen Kappa score
Fruit species classifier	Martaman and Singapuri	0.9899
Maturity stage classifier	Pre-mature, mature and over-mature	0.9886

4.4.4. Receiver operating characteristics (ROC) curve

The ROC curves [20] for species and maturity channels have been consolidated in Fig. 4.8. In this figure the ROCs for maturity classification for Martaman and Singapuri have been shown separate for better visibility and observations. Fig. 4.8(a) shows that the species classifier channel has AUC about 99% which reflects significant potential of the model in species classification channel. The performance is in the same tune for maturity stage classification channels except the Singapuri premature stage where AUC is about 96% and the reason is the green component as the mature stage of this species as mentioned earlier.

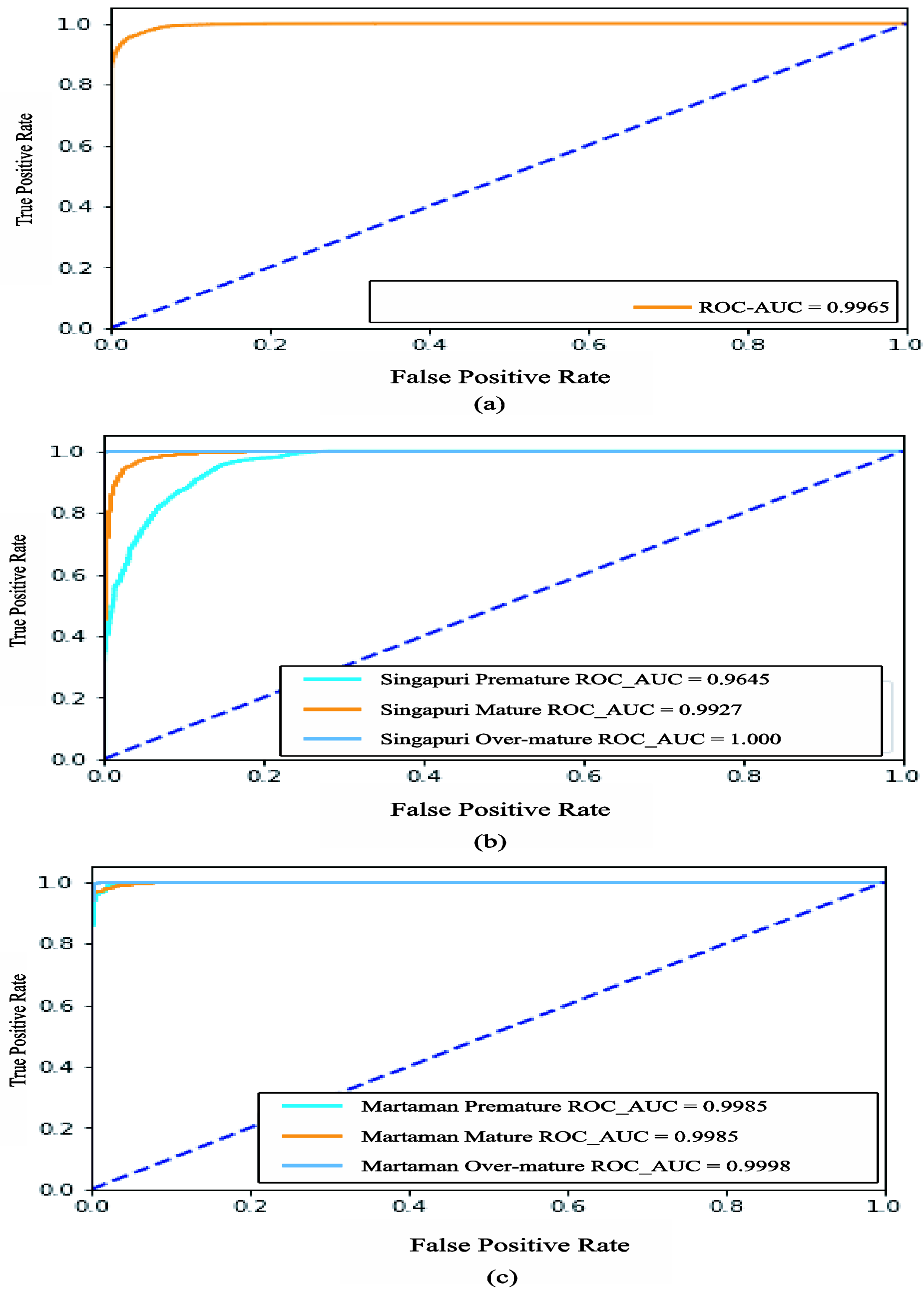


Fig. 4.8: *The ROC and ROC-AUC for (a) species classifier channel and maturity stage classifier channel for (b) Singapuri and (c) Martaman.*

4.4.5. Comparative analysis

The comparative analysis of the presented multichannel CNN model with other pre-trained CNN models and some of the conventional classifiers have been presented in Table 4.9. As in the case of tomatoes and medicinal leaves the channel-wise (R, G and B) statistical characteristics have been considered for preparation of dataset to use in case of conventional classifiers that need to be extracted as feature sets.

Table 4.9: *Performance comparison of presented model*

Species classification/Maturity stage classification						
Method	**Classification**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**FMeasure**
BP MLP	Species and Maturity	0.97	0.96	0.96	0.95	0.95
SVM	Species and Maturity	0.96	0.95	0.96	0.94	0.94
RF	Species and Maturity	0.95	0.92	0.94	0.95	0.95
AlexNet	Species and Maturity	0.97	0.96	0.96	0.95	0.95
VGG19	Species and Maturity	0.97	0.97	0.97	0.96	0.95
ResNet50	Species and Maturity	0.96	0.96	0.97	0.96	0.96
CNN	Species and Maturity	0.97	0.97	0.97	0.96	0.96

Multichannel CNN (Species and maturity classifier) optimized with GWO	Species and Maturity	0.98	0.98	0.98	0.98	0.98
Multichannel CNN (Species and maturity classifier) with modified GWO	Species and Maturity	0.98	0.98	0.98	0.98	0.99
Multichannel CNN (Species and maturity classifier) optimized with BPSO	Species and Maturity	0.99	0.98	0.98	0.99	0.99

Table 4.9 shows that the multichannel CNN outperforms the other conventional and deep learning based classifiers. The performance of BPMLP [21] and SVM [22] model is not as good as those of the deep learning models AlexNet [23], VGG19 [24] and ResNet50 [25] used in the classification process. It is also prominently visible from the table that the performance accuracy of the multichannel CNN model has also superseded that of the species and maturity classifier model and eliminates the steps of feature extraction and the corresponding time to considerable degree. Hence the multichannel species and maturity classifier model has the more potentiality over the other models. It is clearly visible from the performance presentation of the optimized multi channel CNN model, that GWO optimized CNN and the modified GWO optimized CNN

are analogous in terms of both the species and maturity classification. The BPSO has slightly higher impact on performance metric of the CNNmodel compared to the GWO and the modified GWO approach.

4.4.6. GUI representation

The GUI implementation for the multichannel CNN model with channel 1 for species classification of martaman and singapuri banana and the channel 2 for maturity classification of each of the banana species as premature , mature and overmature have been represented in this chapter. Fig. 4.9 presents the species wise maturity information for any unknown banana sample belonging to any of the six categories namely premature, mature and overmature types of martaman and singapuri banana. The prediction regarding the correct species of banana has been determined by the channel -1 and correct maturity information of the banana sample has been determined by the channel -2. The prediction vectors corresponding to the species martaman and singapuri has been expressed as [1.0] and [0.1] respectively. Prediction vectors for premature, mature and overmature states of the martaman and singapuri banana species are expressed as [1.0.0], [0.1.0], [0.0.1], [1.0.0], [.0.1.0] and [0.0.1] respectively.

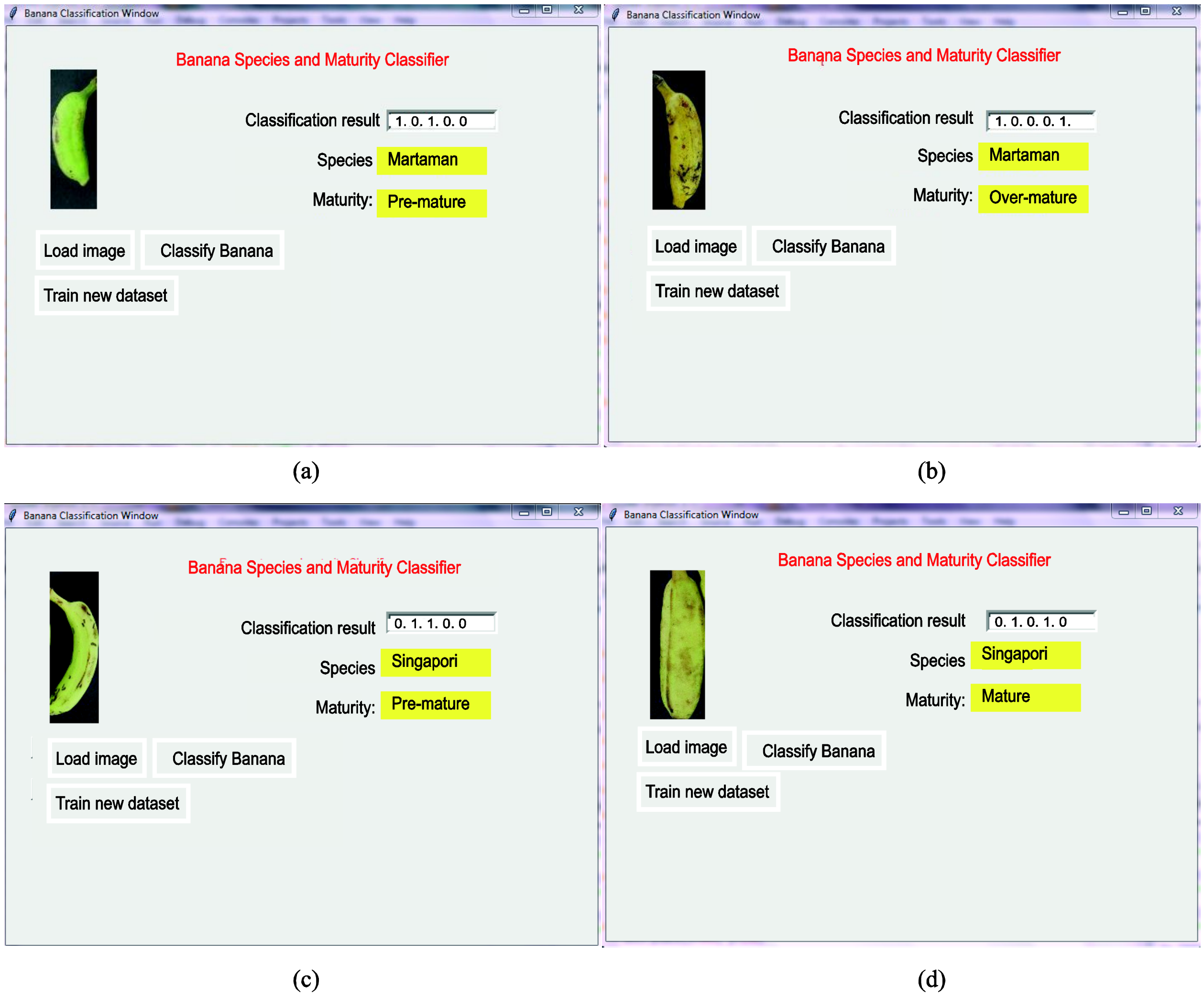


Fig. 4.9: *GUI windows corresponding to the banana class and maturity identification for any unknown sample identified as (a) premature martaman (b) over-mature martaman (c) pre-mature singapori and (d) mature singapori.*

4.5. Conclusions

This chapter presented a computer vision framework for multi channel model based on the CNN architecture towards the classification of species and maturity of the banana fruit using respective channel.The variation of different bio-actives and the nutrients present in the banana fruit and also their accountability towards the maturity based classification has been experimentally assessed in justification of the problem. Indigenous imaging chamber has been appointed for the purpose of capturing images of banana fruits which has been elaborated in this chapter. This chapter focuses on the development of the multi-channel CNNmodel starting from the sample collection using the photographic chamber. It also highlights the use of BPSO, GWO and modified GWO optimization for hyperparameter optimization in order to get more generalization and performance of the multichannel CNN model. The performance measurement of the BPSO algorithm has shown the slight improved performance towards the accuracy of calculation. The model has been subjected to the validation and the classification result of the unknown testing set has been evaluated against the standard metrics. The comparison shows that the presented method can achieve upto 99% accuracy. The comparative analysis has also been presented towards the justification of using the model over some existing conventional and deep learning based models. On the all round assessment of the performance of the model on a significant scale confirms the presented framework to be the addition to the conventional instrumental and analytical screening and sorting methods ofthe banana fruits of different species and maturity grades.

References

1) Fatma M. Nashat A.A. (2019). Ripeness classification of bananas using an artificial neural network, Arabian Journal for Science and Engineering, Vol. 44, No. 8, pp.6901–6910, King Fahd University of Petroleum & Minerals.
2) Ahmed F.A.,MohammedO.A.,Samy S.A.N.(2019).Banana Classification Using Deep Learning, International Journal of Academic Information Systems Research (IJAISR),ISSN: 2643-9026 ,Vol. 3, No. 12, pp. 6-11.
3) Prabha D.S., Kumar J.S. (2015). Assessment of banana fruit maturity by image processing technique,J.FoodSci. Technol.,Vol.52, No. 3, pp.1316–1327 [online] https://doi.org/10.1007/s13197-013-1188-3

4) Subedi P.P.,WalshK.B. (2011). Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy, Postharvest Biol. Technol., Vol. 62, No. 3, pp.238–245.

5) CardosoJ.M. , Pena R.D.S. (2014).,Hygroscopic behavior of banana (Musa ssp. AAA) flour in different ripening stage, Food Bioproducts Process, Vol. 92, No. 1, pp.73–79.

6) AndasRamadhan Y., DjamalE.C.,KasyidiF.(2020).Identification of Cavendish Banana Maturity Using Convolutional Neural Networks ,Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management Detroit, Michigan, USA, August 10 - 14.

7) Chu X. , Miao P.,ZhangK., Wei H. , Fu H., Liu H., Jiang H. , Ma Z. (2022).Green Banana Maturity Classification and Quality Evaluation Using Hyperspectral Imaging. Agriculture , 12, 530. https:// doi.org/10.3390/agriculture12040530.

8) Sutan S.M., Prayogi I.Y., DamayantiR.,Djoyowasito G.(2020). Design to prediction tools for banana maturity based on image processing, International Conference on Green Agro-industry and Bioeconomy ,IOP Conference Series Earth and Environmental Science 475(1):012010.

9) Leelavathy B., Sri Datta Y.S.S., Rachana Y.S. (2021). Quality Assessment of Orange Fruit Images Using Convolutional Neural Networks.Proceedings of International Conference on Computational Intelligence and Data Engineering. Lecture Notes on Data Engineering and Communications Technologies, vol 56. Springer, Singapore. https://doi.org/10.1007/978-981-15-8767-2_33.

10) Rajkumar P., ChandrasekaranI.R., Wang N., Raghavan V.G.S. (2015). Classification of banana fruit maturity using zNose. AgricEngInt: CIGR Journal, 17(4):389-396.

11) Khan A., Sohail A.,Zahoora U., Saeed A. (2021). Basic Overview of CNN and its Architectural Survey.

12) Indolia S., Goswami A.K., Mishra S.P., Asopa P. (2018). Conceptual understanding of convolutional neural network–a deep learning approach, International Conference on Computational Intelligence and Data Science (ICCIDS) Procedia Computer Science, Vol. 132,pp.679–688, DOI: doi.org/10.1016/j.procs.2018.05.069.

13) Yamashital R., Nishio M., Do R.K.G. ,TogashilK.(2018) Convolutional neural networks: an overview and application in radiology, Insights into Imaging ,pp.611–629.

14) Traorea B.B., Foguema B.K. ,Tangara F.(2018). Deep convolution neural network for image recognition, Ecological Informatics, Vol.48. pp..257-268.

15) Wang Y., Zhang H., Zhang G. (2019).cPSO-CNN: An efficient PSO based algorithm for fine-tuning hyper-parameters of convolutional neural networks. Swarm EvolComput 49:114–123.

16) Nadimi-Shahraki M.H., Taghian S., MirjaliliS.(2020). An improved grey wolf optimizer for solving engineering problems, Expert Systems with Applications, vol. 166, p. 113917.

17) Berrar D. (2018). Cross-validation. Encyclopedia of Bioinformatics and Computational Biology, 1, Elsevier, 542 -545.

18) Santra A.K., ChristyC.J.(2012).Genetic algorithm and confusion matrix for document clustering. International Journal of Computer Science Issues (IJCSI), 9(1), 1694-0814.

19) McHugh M.L. (2012).Interrater reliability: the kappa statistic, BiochemiaMedica. 22(3), 276-82.

20) Fawcett T. (2006).An introduction to ROC analysis. Pattern Recognition Letter, 27, 861–874.

21) Kanungo D.P., Naik B., Nayak J., Baboo S., Behera H.S. (2015). An improved PSO based back propagation learning-MLP (IPSO-BP-MLP) for classification, Computational.

22) TianY. Shi Y., Liu X. (2012).Recent advances on support vector machines research ,Technological and Economic Development of Economy, Vol. 18, No. 1, pp.5–33 [online] ISSN: 2029-4913, [print] ISSN: 2029-4921.

23) Liu Y., Li Z., Chen X., Gong G., Lu H.(2020). Improving the accuracy of SqueezeNet with negligible extra computational cost, *International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 1-6, doi: 10.1109/HPBDIS49115.2020.9130577.

24) Zhang X., Zhou X., Lin M., Sun J., ShuffleNet: (2018), An Extremely Efficient Convolutional Neural Network for Mobile Devices, Conference, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

25) Dhankhar P. (2019). ResNet-50 and VGG-16 for recognizing Facial Emotions. International Journal of Innovations in Engineering and Technology (IJIET) http://dx.doi.org/10.21172/ijiet.134.18.

CHAPTER 5

PLANT LEAF DISEASE DETECTION – CASE STUDY WITH TOMATO LEAVES

5.1 Introduction

Early detection of plant diseases is an important step towards prevention of huge agricultural and economic loss [1]. It is difficult to perform manually in huge farming land but most of the diseases are caused by the fungal or bacterial organisms [2] and spread across the plants rapidly due to their infectious nature. Hence, computer vision has been a popular measure for this task which can be a part of electronic surveillance for farming areas. Different plants have different diseases, In our case the study has been presented with tomato plants since the species gradation and maturity stage identification have been already presented for tomatoes and availability of diseased leaves database in public domain. At the same point of time tomato is one of widely consumed agro produces across the world, therefore majorly contributes in the agro economy of producing countries.

The work presented in this chapter has been published as:

Mukherjee G, Chatterjee A, Tudu B .(2022) Identification of the types of disease for tomato plants using a modified gray wolf optimization optimized Mobile NetV2 convolutional neural network architecture driven computer vision framework. ***Concurrency Computaiot Pract Exper.;e7161. doi: 10.1002/cpe.7161*****t:**

Out of different tomato leaf diseases, the seven main salient types are bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, spider mites and target spots. Some of the sample images of the diseased leaves as collected from Plant Village database [3] have been presented in Fig. 5.1. Some of the major works reported in this direction have been consolidated in Table 5.1.

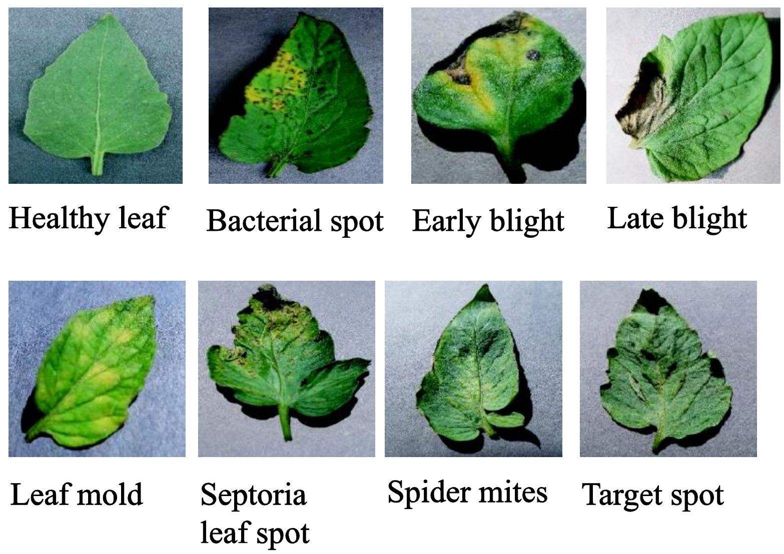


Fig. 5.1: *Examples of tomato leaf diseases from Plant Village database*

Table 5.1: *Some of the reported techniques for automatic grading of bananas*

Base technique	**Method**	**Result**
Hybrid intelligent system in categorization and classification of diseases. The hybrid system comprises of the Fuzzy inferences naive Bayes, probabilistic neural network and support vector machine classifiers.[4]	Ten different her diseased samples has been subjected to the hybrid intelligent system.	The accuracu of classification is varying between the 94% to 98% on the basis of classifiers.
The deep convolutional neural network in detection of four different cucumber	The symptom images has been separated from the diseased leaves of cucumber under some field condition and the	The accuracy value of 93.4% has been obtained as the good

diseases.[5]	set of images have been subjected to the DCNN.	recognition of DCNN.
The CNN based SqueezeNet architecture in identification of the tomato plant leaf diseases. [6]	The seven different diseases along with the healthy tomato leaves have been chosen as the candidates for the Squeezenet classifier based on the keras deep learning framework	The accuracy value of 87% has been achieved in the classi-fication process.
The CNN based model in recognising the maize leaf diseases.[7]	The neuroph based on the common neural network has been used to train the model with the diseased maize leaves.	The accuracy value upto 99.9% has been achieved through this process for the northern con leaf blight.
The RCNN and the mask RCNN model for the detection of tomato leaf diseases.[8]	Four different deep convolution models have been combined based on the two different object detection models. The open data source has been used for the experimental purpose.The mean average precision value has been estimated	The procedure carried out here resulted with the accuracyvalue of 99.64%.
CNN architecture in high level features of the rice diseases diagnosis.[9]	The selected high level features has been proved to have the higher edges over the locallocal binary patterns and the wavelet transform.	The classification accuracies has been found to be 95%.

Table 5.1 shows that the non-invasive and inexpensive natures of computer vision can be a promising solution for early detection of plant diseases. As elaborated in previous chapters CNN can play pivotally in this case as well. Some of the CNN based applications using pre-trained networks like AlexNet, GoogleNet, ResNet and others have been used and found to be significantly potential in this application. However, the possibility of employing a comparatively

lighter CNN model towards hand held mobile device based realization is an open room to explore. This motivated to explore the scope of MobileNetV2[10] for tomato leaf disease detection. MobileNetV2 is comparatively lighter in architecture and have proven potential for faster performance in different problems. For this work the CNN model has been developed using the MobileNetV2 layer dynamics.

Optimization of hyperparameters[11]for MobileNet V2 [12] is important to obtain optimum performance and generalization. In this work a modified gray wolf optimization (GWO) optimization [13] paradigm has been adopted for that. The optimized model has shown significant performance in terms of identifying the tomato diseases in the tune of 98% accuracy. The lightness of the model has also been compared with many pre-trained models and found to be promising for realization with low-configuration hardware setups.

5.2. MobileNet V2 model for tomato leaf disease detection

5.2.1. Dataset preparation

As stated in previous section for this work the Plant Village dataset for tomato leaf diseases has been used. But, the images were subjected to the background subtraction which can contribute towards better understanding of region of interest and that can results into improved performance of classification model (MobileNet V2 in this case). The background segmentation involved the similar approach based on Otsu's thresholding as described in Table 3.2 of chapter 2 but as all are tomato leaves here and does not involve any species-wise classification the images were converted into hue, saturation and value (HSV) color space from the native RGB color space. The reasons of using HSV color space to obtain distinguishing color channel information as in RGB all the channels are almost equivalent and another important observation is the shadow present in most of the images which could not be avoided with RGB channels as shown in Fig.

5.2. Fig. 5.2 reveals interesting fact that the using the hue (H) channel of HSB color space serves two purposes; avoiding the shadow distortion present in the images since shadows are always dark (black) in terms of hue which provides significant difference in hue map from the hue of the object of interest and easier to find the real contour of the object. Although, CNN by its dynamics can handle different distortions but such preprocessing can accelerate the potential especially for cases where the architecture is lighter and requirement is for faster decision.

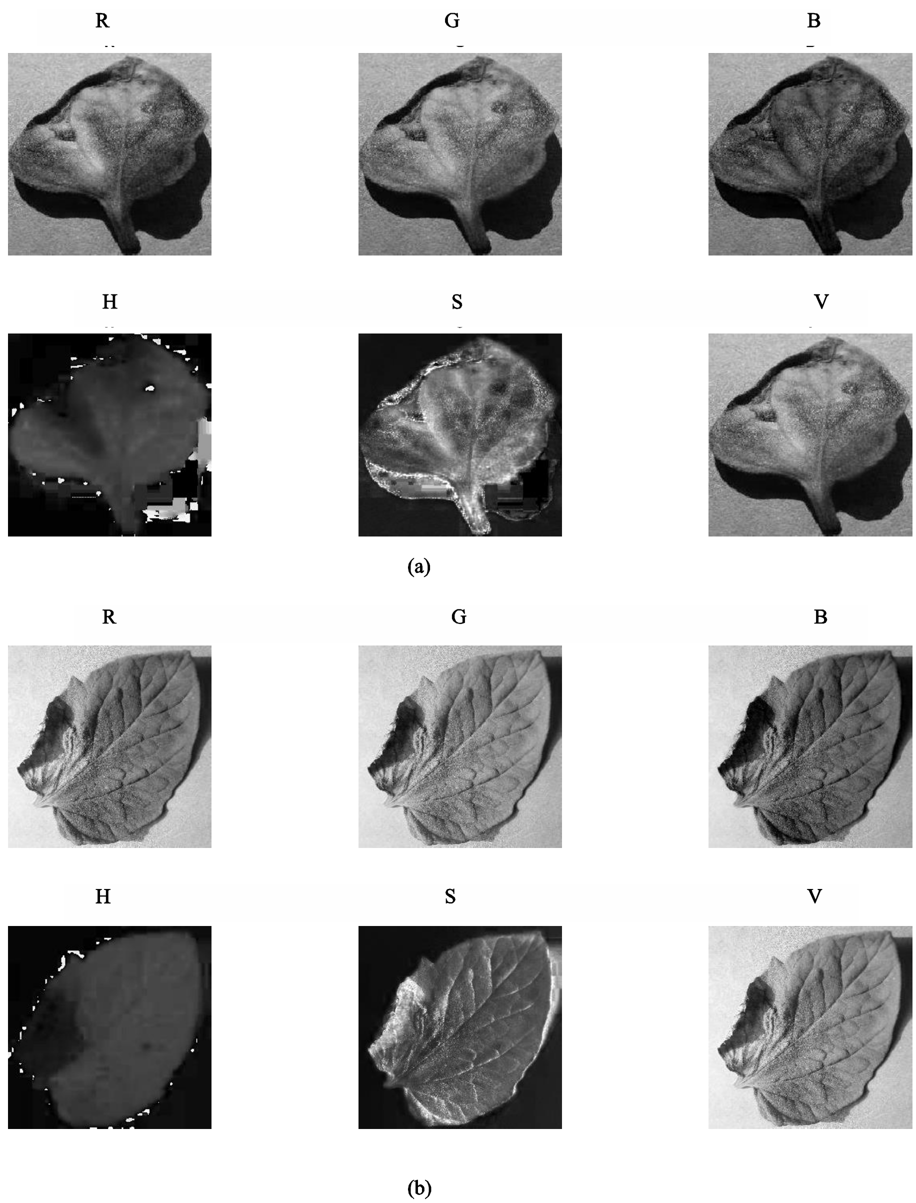


Fig. 5.2: *Color channel separations for diseased leaves with (a) early blight and (b) target spot diseases.*

5.2.2. MobileNet V2 classification model development

MobileNetV2 is originated from MobileNet[14] architecture that is designed with 13 convolution blocks. The convolution block consists of convolution layer, pooling layer and batch normalization (batchNorm) layer. RelU activation function has been used to achieve the non-linearity in the model. The fully connected layer is similar to previous discussions and the output layer consists of output nodes equal to the number of classes under consideration (8 in our case i.e. healthy and 7 types of diseases tomato leaves).

In MobileNetV2 the cost of computation in convolution layers has been reduced by the use of depth wise separable convolution blocks. Each convolution blocks consists of the two distinct convolution layers; depth-wise convolutional layer (*D*) and point-wise convolutional layer (*P*).The*D* layer is concerned to the application of single filter to each of the input channels. The *P* layer then applies a filter to the combination of the output produced from the entire depth wise convolution. Figure 5.3 shows the split of convolution blocks which results into considerable reduction of computational expenses.

Considering *X* and *Y* as the input and output of convolution block feature map, *P* and *Q* as the input and output depth, respectively the input and output feature maps can be represented as $D_x \times D_x \times P$ and $D_y \times D_y \times Q$ where D_x and D_y are the spatial dimension of input and output feature maps, respectively. If the convolution is performed with the filter dimension $D_t \times D_t \times P \times Q$ with square kernel of size D_t the output feature map produced by convolution operation with padding and stride can be represented as Eq. (5.1) and the associated computation cost can be presented as Eq. (5.2).

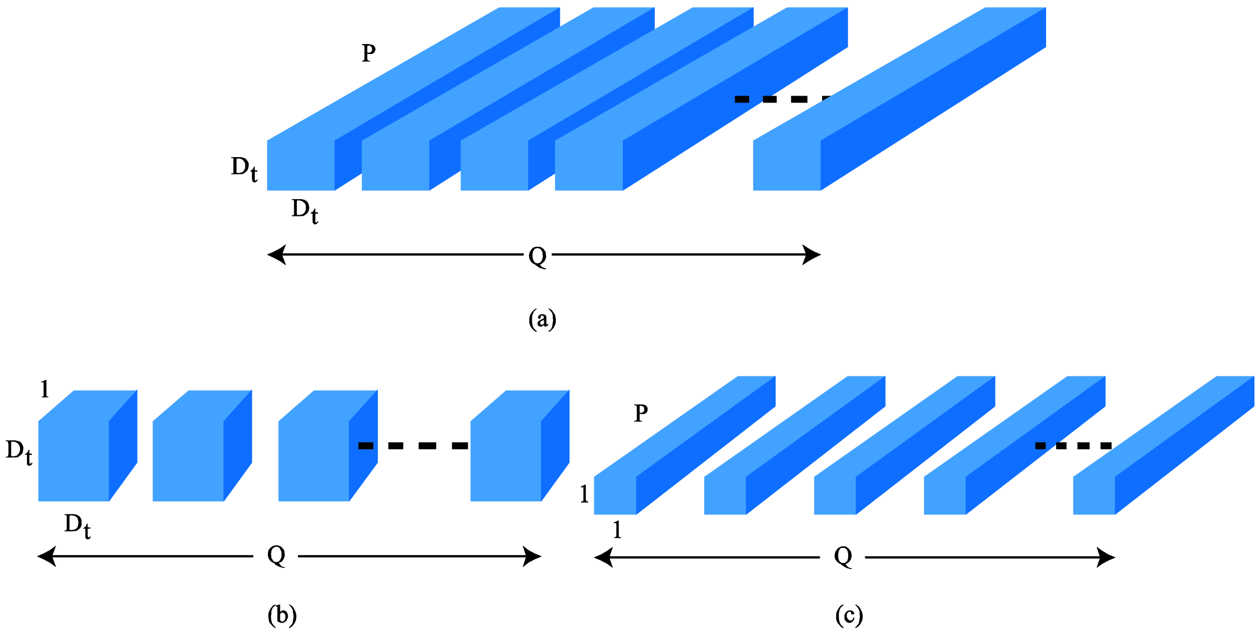


Fig. 5.3: *Split in convolutional block(a) standard convolution filters (b) depth-wise and (c) point-wise convolution filters.*

$$Y_{k,l,n} = \sum_{i,j,m} T_{i,j,m,n} X_{k+i-1,l+j-1,m} \quad (5.1)$$

$$C\_cost = D_t \times D_t \times P \times Q \times D_x \times D_x \quad (5.2)$$

where, *C_cost* depends on the spatial dimension *Dx* and *Dy* The calculation of depth-wise convolution in each block/channel is performed using Eq. (5.3) where $\hat{T}$ is the kernel of dimension $D_t \times D_t \times P$ and m_{th} is filter in $\hat{T}$ that is applied to the mth channel in X in order to produce the mthchannel of the filtered output feature map $\hat{Y}$. The *C_cost*for depth-wise convolution is performed as Eq. (5.4).

$$\hat{Y}_{k,l,n} = \sum_{i,j,m} \hat{T}_{i,j,m,n} \hat{X}_{k+i-1,l+j-1,m} \quad (5.3)$$

$$C\_cost_{depth-wise} = D_t \times D_t \times P \times D_x \times D_x \quad (5.4)$$

The point-wise convolution is performed to combine the output of multiple depth-wise convolution channels and produce new features with *1 x 1* convolution. The output of point-wise convolution is finally combined to that of depth-wise convolution [9] to form the depth-wise separable convolution results. The cost of depth-wise separable convolution can be calculated as Eq. (5.5).

$$D\_C\_cost = D_t \times D_t \times P \times D_x \times D_x + D_x \times D_x \times P \times Q \tag{5.5}$$

The cost ratio (CR) as calculated by Eq. (5.6) and (5.7) can be used to measure the reduction of cost of computation in MobileNetV2. The value of Q is usually greater than 10 which conveys that computation cost on using the depth-wise convolution is much less (approximately 8 to 9 times less) compared to the standard convolution with nominal drop of inaccuracy value produced.Width parameter and resolution parameter are two important hyperparameters of MobileNet architecture to achieve balance between efficiency and accuracy of the model. Including α and ρ as width and resolution parameter, respectively, the cost of depth-wise separable convolution can be presented as Eq. (5.8).

$$CR = \frac{D_t \times D_t \times P \times D_x \times D_x + D_x \times D_x \times P \times Q}{D_t \times D_t \times P \times Q \times D_x \times D_x} \tag{5.6}$$

$$CR = \frac{1}{Q} + \frac{1}{D_t^2} \tag{5.7}$$

$$D\_C\_cost = D_t \times D_t \times \alpha P \times \rho D_x \times D_x + \rho D_x \times D_x \times \alpha P \times \alpha Q \tag{5.8}$$

The back propagation principle of propagating error from the output layer is adopted for learning. The weight updates are achieved using mean square error (MSE) calculated as Eq. (5.9) where L and N are the number of layers and the number of outputs in the output layer, t^I and $[o_1^L,...,o_N^L]$ are the target vector and the output vectors corresponding to the input vector I , respectively.

$$ER_I = \sum_{k=1}^{N} (o_k^L - t_k^I)^2 \tag{5.9}$$

Linear bottleneck layers are another important functional part of MobileNet V2 architecture.

The bottleneck layer consists of the three distinct convolution layers. The first layer has the dimension 1 X 1 and is concerned to the channel reduction. The next layer having 2 X 3 convolution is concerned to the extraction of spatial features. The last layer with the dimension 1

X 1 is concerned to the channel expansion. The bottleneck structure can further be developed by widening the channels in each of the convolutional layer. The inverted residual structure is following the converse idea of the classic bottleneck structure. It introduces the shortcuts between the linear bottlenecks. It plays a prime role in optimizing the model complexities to the higher degree. The inverted residual block has been introduced in the later version of the MobileNet. It takes as input the compressed tensors of low dimension and converts the same into higher dimension by means of the point wise convolution. The low dimensional feature tensors have been produced inside the inverted residual block by subjecting the structure to the depth wise convolution followed by the point wise convolution.

The schematic representation of MobileNetV2[15] architecture comprising of the one convolution layer, 17 bottleneck residual blocks ,one regular convolution layer ,global average pooling layer and finally classification layer as used for tomato leaf disease classification is shown in Fig. 5.4. The implementation of MobileNetV2 for leaf disease classification is shown in Fig. 5.5.

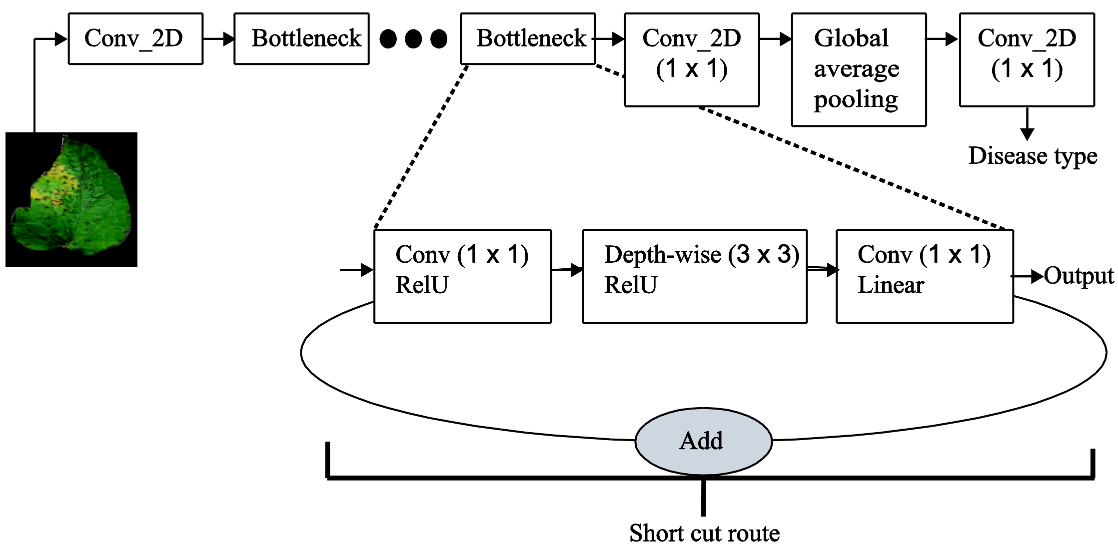


Fig. 5.4: *The MobileNet V2 architecture used for tomato leaf disease classification.*

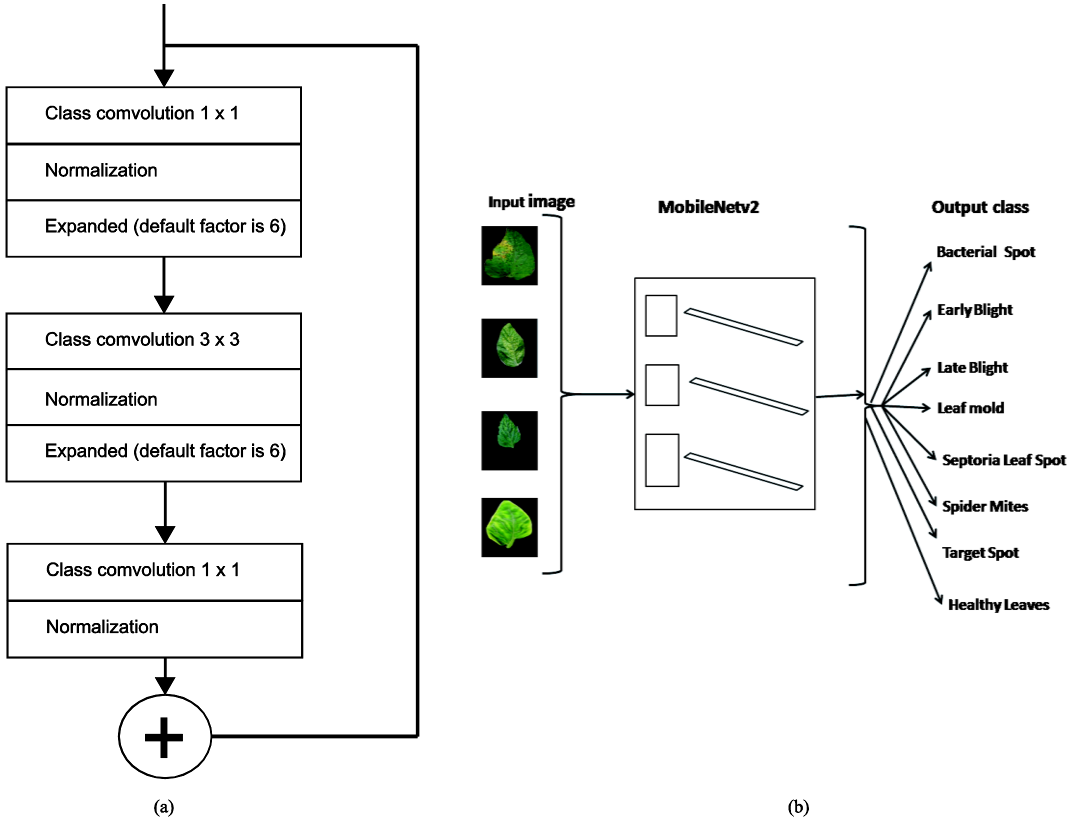


Fig. 5.5: *Leaf disease classification using MobileNet V2 architecture (a) layered architecture and (b) classification operation.*

The final layer has been exempted from the nonlinearity and finally feeds into the *softmax* layer for classification of different types of diseases. The first layer and the depth-wise convolution layers all maintain the strided convolution and the spatial resolution has been reduced down to 1 before reaching the fully connected layer by means of the final average pooling. MobileNetV2 preserves altogether 28 layers by counting the depth-wise and point-wise convolutions separately. The fully connected layer of the network has been entirely removed and all the weights of the model have been frozen up except the last four layers. Kernels and weights were updated by the gradient descent optimization algorithm. The degree of compatibility among different output predictions are measured by the loss function. Cross-entropy and the mean-squared error(MSE) are two most popular loss functions. Gradient is the partial derivative of the loss with respect to each learnable parameter. The weight updates occur following Eq(5.10).

$$w = w - \alpha * \frac{\partial L}{\partial w} \qquad (5.10)$$

Where, α is the learning rate and *L* and *w* stand for the loss function and each learnable parameter, respectively.

5.2.3. Hyperparameter optimization using Modified GWO

The hyperparameters of the multi channel CNN model have been optimized by using the three most promising methods BPSO, GWO and the modified GWO . The architectural description of the appointed models has been described in the section 2.4 in details. The comparative analysis of the three optimizer has been carried out in terms of the convergence curves which are provided in the result and discussion section of this chapter.

5.3. Results and discussions

The MobileNetV2 model was trained using *Tensorflow* backend [16] with *Adam* [17] optimizer with the asynchronous gradient descentlike the InceptionV3 [18].Total 8000 images were taken for the experimentations consisting of 1000 leaves for each of the 8 classes. The optimization of hyperparameters was performed with 30% randomly taken samples and the convergence plot based on the BPSO,GWO and modified GWO for 100 iterations is shown in Fig. 5.6 and the features extracted from different layers of the model have been consolidated in Fig. 5.7.

The plot of Fig. 5.6 shows that after 80 iterations the BPSO [19] converges with more than 98% accuracy, the GWO[20] converges with a little more than 98% accuracy and for modified GWO , the convergence varies with nearly 98.5% accuracy value. The hyperparameter finally arrived with the BPSO as used for the presented results are consolidated in Table 5.2.

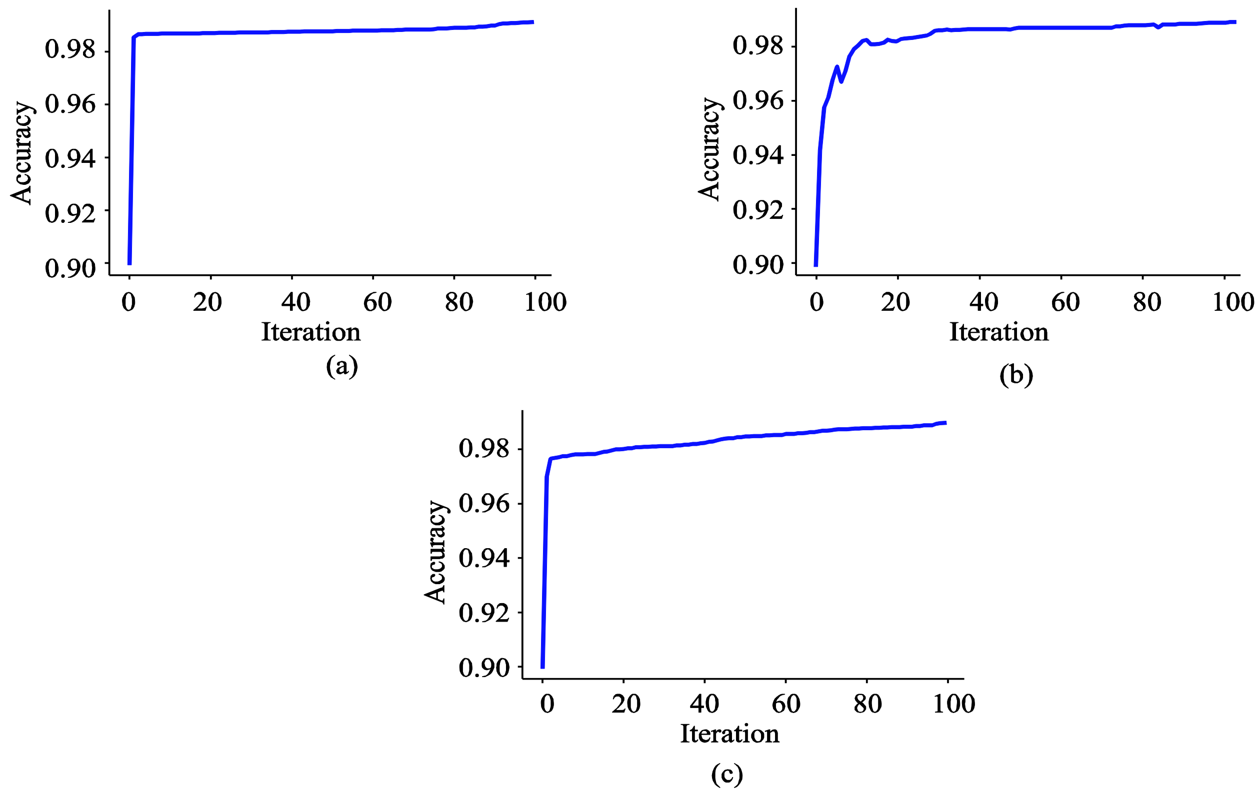


Fig. 5.6: *Convergence plot for (a) BPSO ,(b) GWO and (c) modified GWO optimizer*

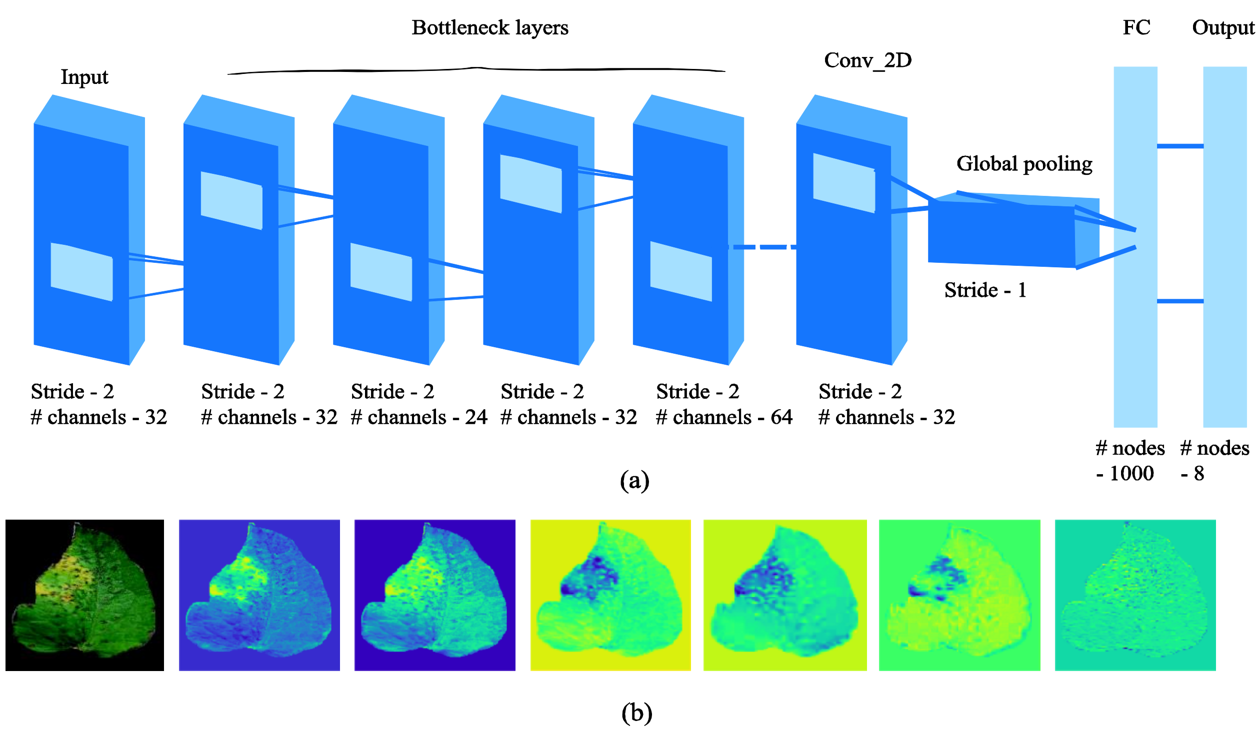


Fig. 5.7: *Feature extraction by MobileV2 (a) layers and (b) images extracted from layers.*

Figure 5.6 shows that at convergence modified GWO results about 98% accuracy with the optimized parameter settings. The functioning of the architecture in terms of deeper feature extraction is evident from Fig. 5.7. It can be seen that in the initial layers the image is visible as leaf but as the image progresses through the deeper layers more abstract features are getting extracted. This ensures automatic extraction and combination of different features like shape, color and textures along with their derived features by the split convolution layers. The optimum hyperparameter settings as arrived by the optimization and employed for all the results shown in chapter have been presented in Table 5.2.

Table 5.2:*MobileNet V2 optimized parameter settings*

Layer type	**Variable**	**Range**
Conv2D	Image size	224X 224
	Number of kernels	32
	Kernel size	3 X 3
	Stride size	2
	Number of channels	32
Bottleneck	Image size	112 X 112
	Number of kernels	16
	Kernel size	3X 3
	Stride size	2
	Number of channels	16
Bottleneck	Image size	56 X 56
	Number of kernels	24

	Kernel size	3X 3
	Stride size	2
	Number of channels	24
Bottleneck	Image size	28 X 28
	Number of kernels	32
	Kernel size	3X 3
	Stride size	2
	Number of channels	32
Regular Conv2D	Image size	7 X 7
	Number of kernels	320
	Kernel size	3X 3
	Stride size	1
	Number of channels	1280
Global average pooling	Image size	1280
	Number of kernels	32
	Kernel size	3X 3
	Stride size	-
	Number of channels	1

Global average pooling(Prediction dense layers)	Number of nodes	1000
dense (Dense)layer	Number of nodes	8

Based on the convergence characteristics as given in the Fig 5.6 , the modified GWO optimized MobileNet2 model has been chosen as the classifier on the convergence characteristics performance shown in Fig 5.6 for plant leaf disease detection and all other metrics have been measured based on this chosen best.

5.3.1. Model validation

Following the convention of the previous chapters and common practice in machine learning application domain here also the entire experimental dataset (8000 samples/images) was divided into 60:20:20 ratios for training, validation and testing. It can be noted that these sets were disjoined so testing set was not known to the trained and validated model. The accuracy and loss plots with the validation set to judge the generalization potential of the model is presented in Fig. 5.8. It shows the generalization potential of the employed optimized MobileNetV2 architecture. The plots for validation and training are following the same trend and have very small gap at the ends. These small gaps convey the avoidance of biasness to particular class(es) by the model. The plots also confirm the potential to avoid under-fitting problems hence was expected to perform well when subjected to unknown testing set.

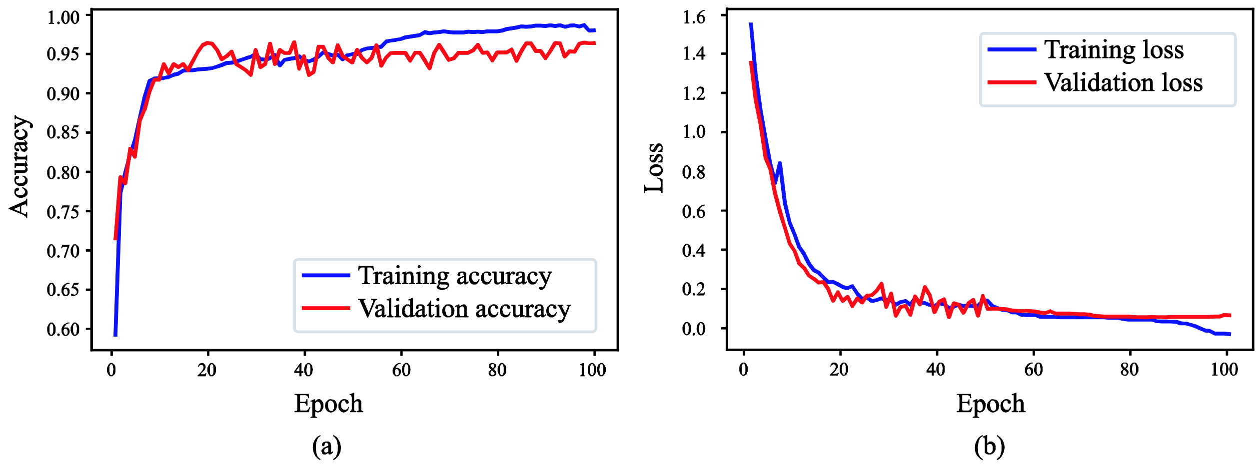


Fig. 5.8: *Model validation plots for (a) accuracy and (b) loss.*

The 10-fold cross-validation results are shown in Table 5.3. It echoes the similar observation as Fig. 5.8 in terms of generalization potential as the accuracy with training and validation sets are very close (both in the tune of 98%). The loss value for validation set is higher than that of training set which also convey practical realization ability and generalization potential of the model. Nevertheless, the loss value with validation set visibly small. The consistency of the model is reflected in the accuracy vs epoch plots which are also in same magnitude however the loss value for validation set is reflected by small gap between minimum and maximum values for both loss and accuracy along with considerably small standard deviation values.

Table 5.3: *Results of 10-fold cross validation*

Number Offolds	**Training data**		**Validation data**	
	Loss	**Accuracy**	**Loss**	**Accuracy**
1	0.0199	0.9888	0.053	0.9766
2	0.0276	0.9834	0.033	0.9834
3	0.0235	0.9889	0.034	0.9856
4	0.0156	0.9843	0.054	0.9888

5	0.0246	0.9845	0.044	0.9892
6	0.0136	0.9803	0.051	0.9879
7	0.0175	0.9899	0.032	0.9858
8	0.0166	0.9985	0.043	0.97655
9	0.0126	0.9875	0.056	0.97653
10	0.0167	0.9845	0.047	0.9852
Average	0.0188	0.98606	0.0447	0.983558
Maximum	0.0276	0.9899	0.0560	0.9892
Minimum	0.0126	0.9803	0.0320	0.9634
Std. Dev.	± 0.0049	± 0.0031	± 0.0091	± 0.0051

5.3.2. Performance evaluation against confusion matrix

The confusion matrix with the classification results obtained with the model has been shown in Fig. 5.9. For convenience of presentation the diseases have been abbreviated. The confusion matrix show that to the maximum 5 out of 200 leaves have been wrongly classified in case of bacterial spot diseased leaves while in case of target spot it is as low as 2 out of 200 leaves. On an average 3 out of 200 leaves have been misclassified for most the diseases and healthy leaf classes. This results into 98.5% overall classification accuracy which is considerably promising as many of the leaf diseases have same kind of appearance hence difficult to distinguish. The performance is further vouched by the metric values that have been listed in Table 5.4.

Predicted

Actual	BS	EB	LB	LM	SLS	SM	TS	HL
BS	195	0	2	1	0	1	1	0
EB	1	196	0	1	1	0	1	0
LB	1	0	198	0	0	1	0	0
LM	1	0	0	197	0	2	0	0
SLS	0	1	1	0	198	0	0	0
SM	1	0	0	0	1	198	0	0
TS	1	0	0	0	0	1	196	0
HL	0	0	1	0	0	1	1	197

Fig. 5.9: *The confusion matrix for the optimized MobileNetV2 classifier for test dataset (BS,EB,LB,LM,SLS,SM, TS and HL indicates bacterial spot, early blight, late blight, leaf mold, Septoria leaf spot, spider mites, target spot and healthy leaves, respectively)*

Table 5.4: *Classification performance by MobileNet V2 architecture*

Diseasetype	Precision(%)	Sensitivity(%)	FMeasure(%)	Specificity(%)
Bacterialspot	98	99	98	
Early blight	97	98	99	
Late blight	97	97	97	
Leaf mold	96	98	98	

Septoria leaf spot	97	97	97	
Spider mites	96	97	98	
Target spots	98	98	99	
Healthy leaves	97	98	98	98
Micro avg.	98	98	98	
Macro avg.	98	98	98	
Weighted avg.	99	99	98	

Table 5.4 conveys 98% sensitivity which indicates that the model can correctly identify the diseased leaves. The specificity value of 98% indicated that the chance of identifying a healthy leaf as diseased ones is very less (only about 2%). The overall accuracy is nearly 99% which further confirms the potential of the model. Finally, the F-measure values between 98-99% vouches the good balance between recall and precision which in-turns reflect the potential of the model to avoid over- and under-fitting while subjected to unknown samples.

5.3.3. Cohen Kappa score

The suitability and quality assessment of the automatic feature extraction with the presented model was evaluated using Cohen Kappa [21] measure. The result is shown in Table 5.5 and with the value of nearly 99% the suitability of extracted feature for the subjected goal can be confirmed.

Table 5.5: *Cohen Kappa score for tomato leaf disease classifier*

Classifier	Classes takenfor measurement	Cohen Kappa Score(%)
MobileNetV2	Bacterial spot, Early blight, Late blight, Leaf mold, Septoria leaf spot , Spider mites, Target spots, Healthy leaves	98.87

5.3.4. Receiver operating characteristics (ROC) curve

Fig. 5.10 shows the ROC curves [22] resulted with the presented model for different classes along with the corresponding ROC-AUC values. It can be seen that the AUC values are consistently in the range of (0.98 – 1). The AUC value of 1 in case of healthy leaves is important to note which conveys the fact the model can distinguish between healthy and diseased leaves with considerable accuracy. Apart from healthy leaves for diseases like light blight, leaf mold, spider might and septoria leaf spot the AUC is 1 as well. On an overall consideration the high AUC values can be important parameter to convey the significant potential of the MobileNetV2 classifier.

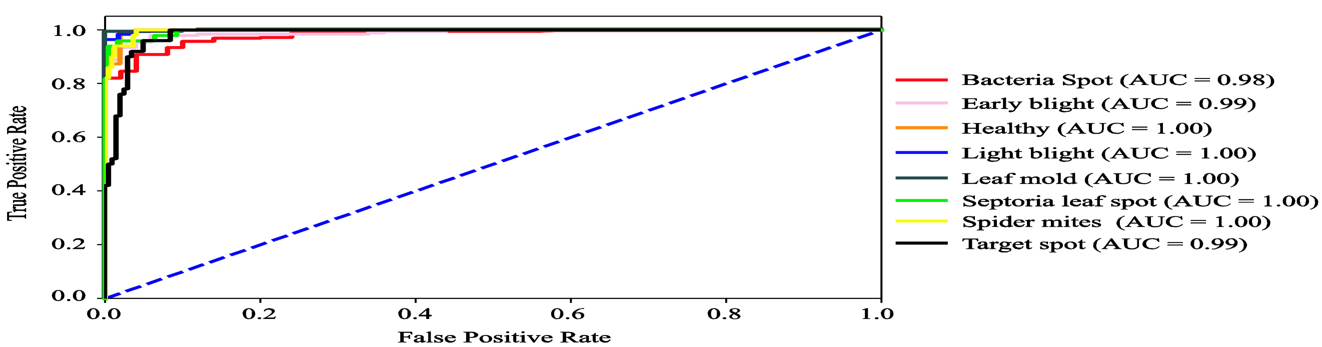


Fig. 5.10: *ROC curves and ROC-AUC values for different classes resulted by optimized MobileNetV2 classifier.*

5.3.5. Comparative analysis

Table 5.6 shows the comparative analysis between the performance of MobileNetV2 based model and other conventional as well as pre-trained CNN models. Since many of diseased leaves have spots and other marks in this case with the color features the contour features also have been extracted for classification using techniques that need hand engineered features.

Table 5.6: *Performance comparison of presented model*

Classification of the diseased tomato leaves						
Method	**Classification**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**FMeasure**
BP MLP	Diseased tomato leaves	0.94	0.94	0.93	0.95	0.95
SVM	Diseased tomato leaves	0.95	0.95	0.95	0.94	0.94
RF	Diseased tomato leaves	0.96	0.94	0.95	0.95	0.96
AlexNet	Diseased tomato leaves	0.98	0.97	0.97	0.97	0.96
VGG19	Diseased tomato leaves	0.95	0.96	0.96	0.97	0.96
ResNet50	Diseased tomato leaves	0.95	0.95	0.96	0.96	0.97
		0.97	0.98	0.97	0.97	0.97

CNN						
MobileNetV2	Diseased tomato leaves	0.98	0.97	0.97	0.96	0.98
BPSO optimized MobileNetV2	Diseased tomato leaves	0.98	0.98	0.98	0.98	0.98
GWO optimized MobileNetV2	Diseased tomato leaves	0.98	0.99	0.98	0.98	0.99
Modified GWO optimized MobileNetV2	Diseased tomato leaves	0.99	0.98	0.98	0.99	0.99

The Table 5.6 shoes that the optimized MobileNetv2 architecture has got the higher score over the other deep learning and conventional classifiers. The performance of the BPMLP ,SVM and RF are not little less compared to the other deep learning based models AlexNet [23], VGG19 [24] and ResNet50 [25] used in the process of disease classification. It is also much evident from the table 5.6 that performance accuracies of the modified GWO optimized MobileNetV2 architecture has better performance over the CNN MobileNetV2 architecture optimized with the simple GWO and the BPSO. This model also reduced the time of computation to a considerable degree as it eliminates the steps of feature extractions. Hence the GWO optimized MobileNetV2 architecture has the more potential capability towards classification of different diseased leaves.

5.3.6. GUI representation

The GUI implementation of the MobileNetV2 model with modified GWO for detection of the diseased tomato leaves has been represented in this chapter. The Fig.5.11 presents the identified tomato plant diseases of unknown category. The prediction regarding the disease for the tomato sample has been done by the optimized MobileNetV2 model. The prediction vectors for different tomato leaf diseases under consideration are expressed as [0.0.0.0.0.0.0] , [1.0.0.0.0.0.0] , [0.1.0.0.0.0.0] , [0.0.1.0.0.0.0] , [0.0.0.1.0.0.0] , [0.0.0.0.1.0.0], [0.0.0.0.0.1.0] and [0.0.0.0.0.0.1] respectively for the healthy leaf, bacterial spot, Early blight, Late blight, leaf mold , septoria leaf spot, spider mites and Target spot.

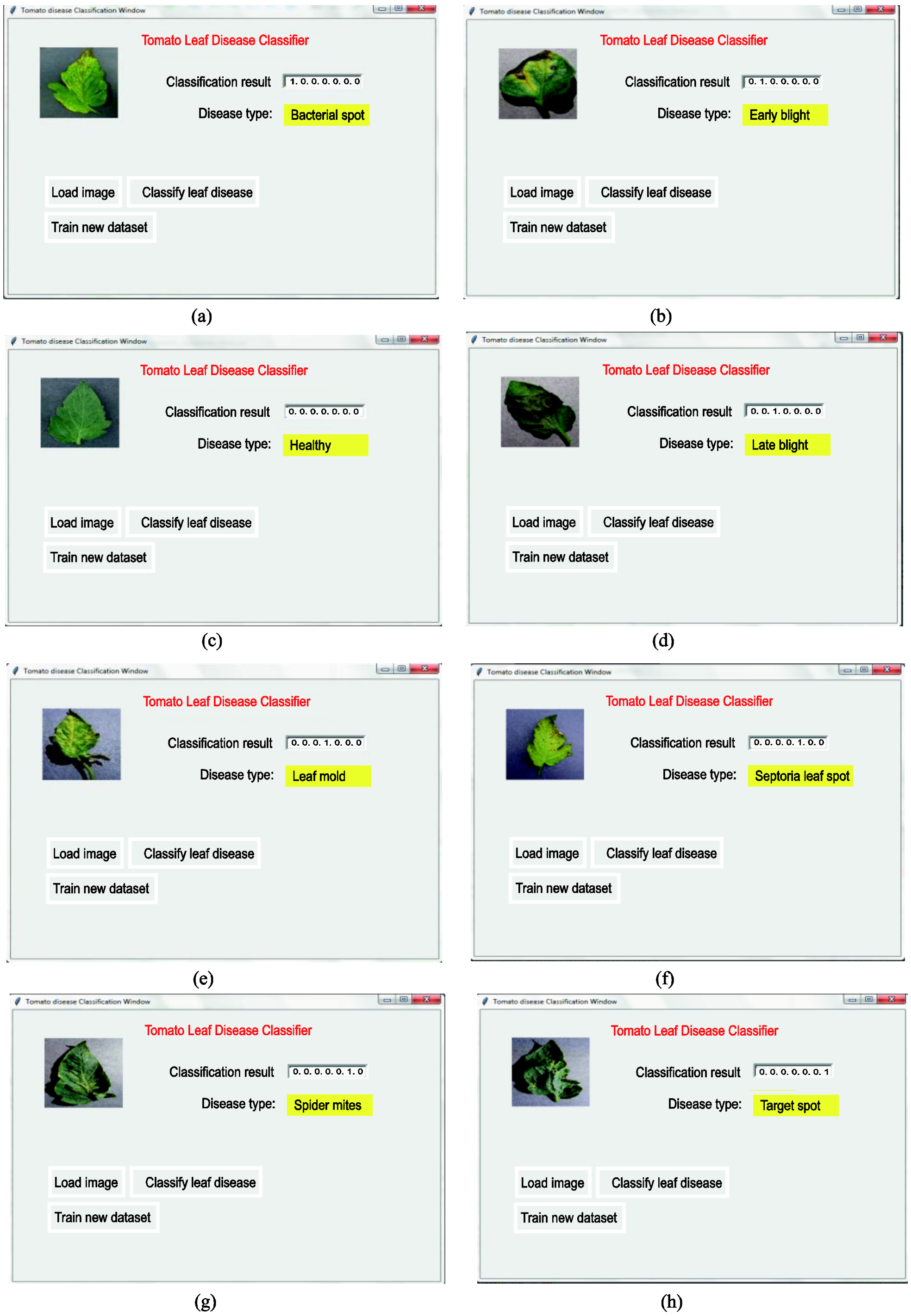


Fig. 5.11: *GUI windows corresponding to the tomato leaf disease identification for any unknown sample identified as (a) Healthy leaf (b) Bacterial spot (c) Early blight (d) Late blight (e) Leaf mold (f) Septoria leaf spot (g) Spider mites and (h) Target spot.*

5.4 Conclusion

This chapter has presented the unique work on the computer vision based framework MobileNetV2 model optimized by the modified GWO architecture towards the detection of diseased leaf and the associated disease types. MobileNetV2 is the light weight variation of CNN model with relatively low complexities and functionalities but with similar computing capability. It also highlights the use of BPSO, GWO and modified GWO optimization for hyperparameter optimization in order to get more generalization and performance of model . The performance measurement of the modified GWO algorithm has shown the slight improved performance towards the accuracy of calculation.. The tomato plant leaves affected with different types of diseases has been experimentally assessed in justification to the problem. The performance of the optimized model has been validated with the test image sets comprising of ground truth image datasets of the diseased leaves obtained from the plant village database. The evaluation performance of the classifier against different metrics confirms the consistency of the model with considerable generalization potentiality. The comparison confirms the achievement of 98.5% accuracy. The optimized MobileNetV2 model has been compared to other existing conventional and deep learning based models. The overall performance assessment of the model confirms the presented framework to be the addition to other existing conventional instrumental methods for detection of any sort of leaf diseases for different plant species.

References

1) Williams S.D., BoehmM.J.(2017). Plants Get Sick Tool, Retrieved from https://ohioline.osu.edu/factsheet/plpath-gen-1.
2) Isleib J. (2018).Signs and symptoms of plant disease: Is if fungal, viral or bacterial? “Retrieved from,https://www.canr.msu.edu/news/signs_and_symptoms_of_plant_disease_is_it_fungal_viral_or_bacterial.
3) Vinay P., Santhosh K.G., Aravind K.U. (2021). Classification of Tomato Plant Leaf Disease Using Neural Network” Proceedings of the International Conference on Innovative Computing & Communication. (ICICC)
   SSRN: https://ssrn.com/abstract=3747953 or http://dx.doi.org/10.2139/ssrn.3747953
4) Mustafa M.S., Husin Z., TanW.K., MaviM.F. ,Farook R.S.M., (2020). Development of automated hybrid intelligent system for herbs plant classification and early herbs plant disease detection”. Neural Computing and Applications, 1-23.
5) Ma J., Du K., Zheng F., Zhang L., Gong Z., Sun Z.(2018).A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network,Computers and Electronics in Agriculture, vol. 154, pp. 18–24.
6) Hidayatuloh A., Nursalman M.,Nugraha E.(2018) Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model, *International Conference on Information Technology Systems and Innovation (ICITSI)*, 2018, pp. 199-204, doi: 10.1109/ICITSI.2018.8696087.
7) Sibiya M., Sumbwanyambe M. (2019). A Computational Procedure for the Recognition and Classification of Maize Leaf Diseases Out of Healthy Leaves Using Convolutional NeuralNetworks . AgriEngineering. 1:119–131.doi: 10.3390/agriengineering1010009. [CrossRef] [Google Scholar]
8) Mim T.T., Sheikh M.H., Shampa R.A., Reza, M.S., IslamM.S.(2019). Leaves Diseases Detection of Tomato Using Image Processing, *2019 8th International Conference System Modeling and Ad-*

vancement in Research Trends (SMART), pp. 244-249, doi: 10.1109/SMART46866.2019.9117437.

9) Liang W., ZhangH., Zhang G.F., Cao,H.(2019). Rice Blast Disease Recognition Using a Deep Convolutional Neural Network. Sci Rep 9, 2869. https://doi.org/10.1038/s41598-019-38966-0.

10) Sandler M. , Howard A., Zhu M., ZhmoginovA.,Chen L.C. (2018). Mobilenetv2: Inverted Residuals and Linear Bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, 18-23 June 2018, 4510-4520. https://doi.org/10.1109/CVPR.2018.00474.

11) Wang Y., Li H., Cai Z., Tang X.,Yang Y. (2020). CNN Hyperparameter Optimization Based on CNN Visualization and Perception Hash Algorithm", *19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, 2020, pp. 78-82, doi: 10.1109/DCABES50732.2020.00029.

12) Sinha D.,Sharkawy M.E.(2019). Thin MobileNet: An Enhanced MobileNet Architecture, IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 0280-0285, doi: 10.1109/UEMCON47517.2019.8993089.

13) Mirjalili S., Mirjalili S.M., Lewis A. (2014). Grey wolf optimizer. AdvEngSoftw 69:46–61.

14) Sinha D., El-Sharkawy M. (2019). Thin MobileNet: An Enhanced MobileNet Architecture, *IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0280-0285, doi: 10.1109/UEMCON47517.2019.8993089.

15) Howard A.G., ZhuM., ChenB., KalenichenkoD., WangW., WeyandT., Andreetto M ., Adam H.(2017). mMobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

16) Abadi M., Barham P., Chen J., Chen Z., DavisA., DeanJ., Devin M., Ghemawat S., Irving,G. ,IsardM. et.al.(2016).TensorFlow: A system for large-scale machine learning. *OSDI*,16,265-283.

17) Kingma D.P.,BaJ. (2015)..Adam: A Method for Stochastic Optimization. *CoRR, abs/1412.6980*.

18) Bankar J. ,Gavai N.R. (2018). Convolutional Neural Network based Inception v3 Model for Animal Classification, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 7, Issue 5.

19) Berrar D. (2019). Performance measures for binary classification. In: Ranganathan S, Gribskov M, Nakai K, Scho¨nbach C (eds)Encyclopedia of bioinformatics and computational biology. Academic Press, Oxford, pp 546–560. https://doi.org/10.1016/ B978-0-12-809633-8.20351-8.

20) Mirjalili S. , MirjaliliS.M,.LewisA. (2014). Grey wolf optimizer. AdvEngSoftw 69:46–61.

21) McHugh M.L. (2012). Interrater reliability: the kappa statistic. Biochem Med 22(3):276–282.

22) Fawcett T. (2006), An introduction to ROC analysis, Pattern Recognition Letters Vol. 27,pp. 861–874.

23) Liu Y., Li Z., Chen X., Gong G., LuH.(2020). Improving the accuracy of SqueezeNet with negligible extra computational cost, *International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 1-6, doi: 10.1109/HPBDIS49115.2020.9130577.

24) Zhang X., Zhou X., Lin M., Sun J. (2018).ShuffleNet:. An Extremely Efficient Convolutional Neural Network for Mobile Devices, Conference, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

25) Sarwinda D., Paradisa R.H., Bustamam A., Anggia P. (2021).Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer,Procedia Computer Science,Volume 179, ,Pages 423-431,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2021.01.025.

CHAPTER 6

Concluding Remarks

6.1 Introduction

Literature survey has revealed the importance of automatic identification, grading and screening of agro produces. However, it has also revealed the room of developing computer vision techniques particularly using emerging convolutional neural network (CNN) methodology. The computer vision technology can be advantageous in terms of low cost, easier implementation and high portability. The thesis presented detailed studies on different classification and prediction tasks in agro produces using computer vision framework and different deep learning architectures.

The species and maturity of medicinal plants were first to consider in this work. Medicinal plants have considerable contribution and healing power for human health due to different bio-actives present in them. However, the concentration of these bio-activates varies between species and maturity stages. This reveals the importance of identification of those two aspects for medicinal leaves. The work in this thesis presented a convolution neural network (CNN) based computer vision model which has shown high accuracy in grading of medicinal leaves, particularly three of them i.e. neem, tulsi and kalmegh, in terms of species as well as maturity stages.

The investigation on the scope of computer vision for screening species and maturity was carried forward for fruits and vegetables as those two are again major contributors of human nutrition. Among different fruits and vegetables, banana and tomato were chosen for experimentation. Maintaining the focus of the thesis in this case also CNN was used but CNNs were different in terms of architecture and parameter settings. In this case also visible model performance has been seen. Finding the right maturity of the medicinal leaves or the fruits for different species of the fruits tomato and bananas are much essential in order to get the maximum medicinal benefits.

Once the three major groups of agro produces, namely, medicinal plants, fruits and vegetables have been approached the work has been extended to identification of diseases in agro produces by means of identifying the leaf diseases. The motivation of the work was to early prediction of diseases which can in turn help to avoid huge agricultural and economical loss. In this case tomato leaves were chosen for experimentation and seven different diseases were considered. A different CNN architecture called MobileNet V2 has been explored to realize lighter yet efficient model. The classification results remain significant in this case as well.

Apart from the above mentioned implementations, the thesis also provide in-house dataset preparation technique using an indigenous imaging chamber, optimization of CNN parameters using soft computing methodologies and also some modifications of the existing conventional soft computing technique for enhanced performance. The highlights of the thesis can be listed as below;

i) Development of the convolution neural network driven computer vision methodology incorporating the tuning of different layer dynamics of the CNN model.
ii) Exploring the feature extractions by individual CNN layers and making required changes in layer parameters for betterment.

iii) Collection and imaging of different agro produces samples (except the leaf disease one) and making the image dataset with accurate labeling.

iv) Developments of possible layer configuration in CNN, for example, multi-channel implementation for simultaneous identification of species and maturity stages.

v) Optimization of the CNN models for improved accuracy and consistency of the models using computational intelligence methods like binary particle swarm optimization (BPSO), gray wolf optimization (GWO) and its modified version.

vi) Exploring the scope of lighter CNN model, namely MobileNet V2 for possible implementation in real-life cases in order to obtain faster results.

vii) Performance evaluation of all the models using standard metrics like confusion matrix and it's associated metrics, Cohen Kappa score, ROC_AUC, etc.

This chapter consolidates the major findings of the thesis, the specific attainment to the research questions framed in the 'Introduction' chapter and some major future works which can be possible extensions of this work.

6.2. MajorFinding

The major findings obtained during the research works on the detection of species of medicinal plant leaves, fruits and their corresponding maturity information can be presented as follows.

- While detecting the species and the maturity information of the medicinal leaves, the UV –Vis spectroscopy has been carried out for the leaves neem, tulsi and kalmegh. The wavelength vs absorption plots has been obtained which shows the peaks for chlorophyll a , chlorophyll b and carotenoids which occur between the wavelength range 400 and 700nm. It is obvious from the plots that the peaks are varying with maturity implying the variation of containment of important biochemical constituents. The maturity wise variation of peaks in the curve provides distinctive estimation of the bio-actives and indicates the prospects of classification on the basis of maturities of leaves. However, the UV-vis spectroscopy and other spectroscopic methods are expensive and time taking despite their accuracy and precision. The computer vision model driven by the BPSO optimized CNN has shown about 99% accuracy. The presented model also shown considerable competitiveness while comparing to some of the most popularly used conventional classifiers and pre-trained CNN models.

- Detection of distinct species and maturity stages for fruits and vegetables also have shown about 98% accuracy. The results also have shown competitiveness against the popular conventional classifiers and some of the popular pre-trained CNN models. This study show that computer vision based model can be an important inclusion to the existing screening methods in retail chain of fruits and vegetable due to its easy implementation ability and faster yet efficient decision.
- The final experimentation work in terms of leaf disease identification also reveal promising potential of CNN driven computer vision framework. It has shown about 99% accuracy when tested with public domain leaf disease dataset. This can be of particular interest as realization of this has been done using a lighter CNN model, called MobileNetV2 and it is possible to integrate with the existing electronic surveillance system to avoid huge agricultural and economical loss by means of early predictions of leaf/plant diseases. Nevertheless, in this case as well the comparison of the presented model with other classification methods found motivating and competitive.
- In all the cases inclusion of CNN parameter optimization has played pivotal role. CNN is associated with considerable number of tunable parameters which affect the performance

of the CNN. Optimization not only improves performance but also causes improvement in terms of consistency. In this connection modification of BPSO and GWO found to be efficient while modification of GWO to address the movement of different wolf in hierarchical levels towards the prey instead of conventional averaging method.

- The thesis also throws lights on image capturing in an inexpensive indigenous imaging chamber along with pixel-scanning based image isolation method. This can be helpful for crops and fruits which are of individual small size. This can also be helpful for real-life implementations where numbers of samples can be pictured together and then can be isolated automatically for feeding into the computer vision classification model.
- Easier and simpler GUI development possibility using Python makes the presented method more realizable in real-world applications. It also provides better scaling possibilities considering IoT based application of computer vision in agro industry.

6.3. Answer to research questions

At the end of this thesis, the answers to the research questions, raised in the first chapter can be addressed as follows.

➢ The concept of computer vision has been proved to be the successful means for determination of maturity of agro items in the noninvasive and nondestructive manner. CNN is an emerging approach of computer vision and demands no feature extraction from the objects like agro items in noninvasive manner. More over the deep learning model can easily be trained in supervised mode with the pool of indigenous datasets to reach considerable accuracy compared to the other existing classifiers as are shown in the current research results. This serves as the answer to the research question 1

- Deep learning based model is the good fit as a machine learning tool used for the classification of any objects on the basis of the image features. Diseased leaves are represented in different form having different colors and textural traits, however visible differences is often insufficient to distinguish between types of diseases by conventional classifiers. Deep learning tool is adept enough to classify different diseased leaves very efficiently. This can be of utmost important in terms of early prediction of rapidly spreading diseases in plant to avoid agricultural and financial loss. The researches carried out using this very principle offers good accuracy on public domain database towards the disease detection of different tomato plant leaves and therefore serve as the answer to the research question2.
- The regression model can be defined on the maturity values of leaves or fruits individually. The model thus formed serves as the predictor model which could be used to predict the maturity stage of any unknown leaf or fruit .The research question 3 therefore can be answered on this basis.
- The optimization of the deep learning models has been carried out by engaging the BPSO, GWO and modified GWO algorithms. The precision and accuracy result obtained with the differently optimized classifier models vary for the different types of datasets employed. The accuracy result with BPSO sometimes outweighs the accuracy results obtained from the other optimizer or vice versa. The research question 4 thus has been answered.
- The agro industry is very vast and deals with various kinds of agro products. Machine learning can be a boon for agro industry for inspection of maturities and species to get the best of the yields. Study of the deep learning model has been proved as a potential means of analysis for different agro produces for their maturity assessment and identification of

species, maturity stages as well diseases in agro produces. It is therefore possible to develop the model in user friendly manner to address such task in industries. The research question 5 has been thus answered.

6.4. Future scopes

The promising potential of the presented methods open up different research avenues as the future research scopes. Some of the major of those can be consolidated as follows.

- In this work only three medicinal plants namely, neem, tulsi and kalmegh has been considered for experimentation due to their ready availability and high medicinal values. This can be extended to many other such plants as a more holistic application.
- Like medicinal plants in case of fruits and vegetables as well the work can be extended to fruits other than banana and vegetables other than tomatoes. That will make the presented models more comprehensive in terms of wide-spread applications in cultivation and supply chain of agro industries.
- The architectures of CNN is a huge area to explore as there can numbers of variants in different aspects. Hence, exploring more on this front can also be a significant future research area.
- Considering the real-life scenario images with more distortions can be included in the experimental dataset so that the model becomes more robust and generous.
- One of the major future work directions is development of hand-held mobile device which can detect maturity and species of the agro produces on-the-spot without requirement of much instrumental setup. This can even be developed as an app in different mobile phone OS platforms. This can address the portability issues of most of the instrumental and analytical methods currently involved in food screening.

- Future works can also be directed towards development of more light yet efficient CNN model which will in-turn provide better convenience in terms of realization and time consumption in decision making.
- The study about identification of plant diseases has been carried out only for the tomato plant leaves in this thesis. The deep learning model has been used for the same. The similar model can be appointed for carrying out identifications of different disease types for different plants. There is a scope for assessing the gravity of diseases for any plant in future. Making the database of such diseases also can be a potential contribution to this domain as many local or regional fruits, vegetables and agro produces may not have publicly available dataset like in case of tomatoes.

6.5. Conclusion

The thesis explored the potential of deep learning methodologies driven computer vision models towards species and maturity classes grading of agro produces. Three major segments, namely, medicinal plants, fruits and vegetables, in the huge varieties of agro produces have been focused in this work. Alongside the potential of computer vision driven models in identification of leaf diseases have as well been studies and found promising. The optimizations of parameter settings also have shown significant contribution to the competitive performances. Considering the high accuracy and possible future developments this work can be a potential step towards development of hand-held devices that can detect species and maturity stages of the agro produces in production line or on-spot of buying. It can also be realized as an app for mobile devices. Scaling and integration of this work in IoT framework can be one of the future trends in emerging automation in food grading and screening.

03.11.22
Associate Professor
Department of Printing Engineering
Jadavpur University, Salt-lake Campus
Kolkata - 700 098

03/11/2022