DETECTION OF PESTICIDES IN LADYSFINGER USING HYPERSPECTRAL IMAGING

18EC810 Project

Submitted in partial fulfillment for the requirement of B.E. degree in Electronics and Communication Engineering of Anna University

Submitted by

Z. Sikkandar Basha
(Reg. No: 20D080)

B. Ganesh Shankar Ram
(Reg No: 20D027)

Guided by

Dr.S.Md. Mansoor Roomi
Professor,
Dept of ECE, TCE.

Department of Electronics and Communication Engineering

THIAGARAJAR COLLEGE OF ENGINEERING
(An Autonomous Institution Affiliated To Anna University, Chennai)

MADURAI – 625 015

May 2024

THIAGARAJAR COLLEGE OF ENGINEERING

(An Autonomous Institution Affiliated to Anna University, Chennai)

MADURAI – 625 015

BONAFIDE CERTIFICATE

This is to certify that the 18EC810 Project entitled "**DETECTION OF PESTICIDES IN LADY'SFINGER USING HYPERSPECTRAL IMAGING**", being submitted by ***Z. Sikkandar Basha*** (Register Number **20D080**), ***B. Ganesh Shankar Ram*** (Register Number **20D027**) in partial fulfillment for the requirement of Bachelor of Engineering Degree in Electronics and Communication Engineering, is a record of bonafide work done by them during the year 2023-2024 under my supervision. The results embodied in this report have not been submitted to any other university or institute for the award of any degree or diploma.

Dr.S.Md. Mansoor Roomi	Dr.S.Rajaram
Professor,	Professor and Head,
Department of ECE,	Department of ECE,
(Guide)	
Station: Madurai	Date:

Certified that the Candidate was examined in the Viva-Voce Examination held at Thiagarajar College of Engineering, Madurai on _______________.

INTERNAL EXAMINER   EXTERNAL EXAMINER   HOD-ECE

ACKNOWLEDGEMENTS

Gratitude cannot be put in words alone, as it has to be felt more than shown. But it is a mystery to us as to how to express it, as I have no choice but to put into words, which would never be sufficient to express our gratitude. We express our sincere thanks to our guide **Dr.S.Md. Mansoor Roomi**, Professor, Department of Electronics and Communication Engineering for his valuable suggestion and support during the course of the project work. We would like to express our gratitude to **Dr.S. Rajaram**, Head of Electronics Communication Engineering for permitting us to undertake this work at the department. I am deeply indebted to his continuous and consistent encouragement. We express our sincere thanks to **Dr.M. Palaninatharaja**, Principal, Thiagarajar College of engineering for allowing us to use the facilities and the help available in the college campus. And we would like to thank **Dr.B. Sathya Bama** for her guidance and we are proud to be associated with her. We wish to thank our friends for their help during the course of the project work. I also thank our family members who always stood with us.

Your's faithfully

SIKKANDAR BASHA Z

GANESH SHANKAR RAM B

LIST OF CONTENTS

Item	Page No.
Title Page	i
Certificates	ii
Acknowledgement	iii
List of contents	iv
List of figures	vii
List of tables	viii
List of abbreviations	ix
Abstract	x
CHAPTER.1 INTRODUCTION	
1.1 INTRODUCTION	1
1.1.1 TYPES OF PESTICIDES	1
1.1.2 PESTICIDE USAGE	2
1.2 STATISTICAL STUDY ON PESTICIDE USAGE IN INDIA	2
1.3 NEGATIVE IMPACT OF PESTICIDES	4
1.4 PENETRATION LEVEL OF PESTICIDES ON FRUITS AND VEGETABLES SURFACE	5
1.5 EXISTING METHODS FOR PESTICIDE DETECTION	6
1.6 IMPORTANCE OF PESTICIDE DETECTION	7
1.7 HYPER SPECTRAL IMAGING IN PESTICIDE DETECTION	8
1.8 SUMMARY	9
CHAPTER.2 LITERATURE REVIEW	
2.1 LITERATURE SURVEY	10
2.2 INFERENCE	16

CHAPTER.3 PROBLEM DESCRIPTION	
3.1 INTRODUCTION	17
3.2 PROBLEM DESCRIPTION	18
3.3 OBJECTIVES	19
3.4 SUMMARY	19
CHAPTER.4 DATABASE ACQUISITION AND DATBASE CREATION	
4.1 INTRODUCTION	20
4.2 HYPERSPECTRAL DATABASE	21
4.3 IMAGE ACQUISITION SYSTEM	22
4.3.1 SYSTEM RESONON	22
4.3.2 DATABASE CREATION	23
4.4 MOTIVATION	25
4.5 CREATION OF LADY'S FINGER DATABASE	26
4.5.1 IMAGE AUGMENTATION	27
4.6 CHALLENGING ASPECTS OF LADY'S FINGER DATABASE	28
4.7 CONCLUSION	29
4.8 SUMMARY	29
CHAPTER.5 METHODOOGY	
5.1 INTRODUCTION	31
5.2 MACHINE LEARNING PESTICIDE RESIDUE ESTIMATION	32
5.2.1 SAMPLE PREPARATION	33
5.2.2 WAVELENGTH SELECTION	34
5.2.3 FEATURE EXTRACTION	34
5.2.4 CLASSIFICATION	34
5.3 DEEP LEARNING ALGORITHM	38
5.3.1 INPUT DATA CUBE	39
5.3.2 REDUNDANCY ELIMINATION	39

5.3.3 SQUEEZED RESIDUAL CONNECTION	40
5.3.4 CLASSIFIER 3D CNN	41
5.3.4.1 ARCHITECTURE OF 3D CNN	41
5.4 SUMMARY	43
CHAPTER.6 RESULTS AND DISCUSSIONS	
6.1 INTRODUCTION	45
6.2 PREPROCESSING RESULT	45
6.2.1 ROI SELECTION	45
6.2.2 SPECTRAL SIGNATURE	46
6.3 MACHINE LEARNING RESULTS	48
6.3.1 FEATURE EXTRACTION	49
6.3.2 ACCURACY ASSESSMENT OF LADY’S FINGER	50
6.4 DEEP LEARNING RESULTS	51
6.5 AUGMENTED IMAGES OF LADYSFINGER	52
6.6.CNN RESULTS	53
6.7 SUMMARY	57
CHAPTER.7 CONCLUSION	
7.1 CONCLUSION	58
7.2 FUTURE WORK	58
7.3 CONTRIBUTION	58
REFERENCES	59

LIST OF FIGURES

Figure No.	Description	Page No.
3.1	Pesticide usage in India Vs. other countries	15
3.2	Adverse effects of pesticides	16
4.1	Hyperspectral Image samples	19
4.3	Images with and without pesticide	22
4.4	Resonon Hyperspectral camera	23
4.5	Augmented images of Lady's finger	24
4.6	Images of sample at different rotations	25
5.1	Flowchart of machine learning process	28
5.2	Block diagram of machine learning	29
5.3	Block diagram of deep learning	34
5.4	3D squeeze excitation residual spectral network	35
5.5	Block diagram of 3D CNN	37
5.6	Working of 3D convolution	38
5.7	Max pooling layer	38
6.1	ROI selection of lady's finger	42
6.2	Mean spectral curve of lady's finger	42
6.3	Mean spectral curve of lady's finger with and without pesticide	43
6.4	Augemented image of pure,low,medium and high class	47
6.5	Comparsion of training and validation loss	49
6.6	Network performance with and without SE	49
6.7	Cumulative explained variance	50

LIST OF TABLES

Figure No.	Description	Page No.
1.1	Pesticide application in Indian agriculture	2
1.2	Health disorder in padre village of Kerala	5
4.1	Dataset details	21
6.1	Feature extraction of Lady's finger for Lower Band	41
6.2	Feature extraction of Lady's finger for Higher Band	41
6.3	Classifier result for 4 classes of Lady's finger	45
6.4	Classifier result for 2 classes of Lady's finger	46
6.5	Lady's finger result for 4 classes	48

LIST OF ABBREVIATIONS

Abbreviation	Description
3D	3 Dimensional
SE	Squeeze Excitation
SERSN	Squeeze excitation residual spectral network
PCA	Principal component interest
ROI	Region of interest
VNIR	Very near infrared region
CNN	Convolution neural network
GMP	Global max pooling
R_{min}	Reflectance minimum
R_{max}	Reflectance maximum
CEV	Cumulative explained variance
RESNET	Residual network

ABSTRACT

As we see in India, the usage of pesticides have been increasing which has a boon of increasing the productivity at the same time it also has a bane of affecting our lives. In order to keep ourselves healthy we have to avoid consuming vegetables with pesticide. On considering this fact the proposed project detects the vegetables whether it is with or without pesticide. Lady's finger are chosen. Since hyperspectral images give more information the dataset are collected by capturing the images in Resonon Hyperspectral camera(Pika-L(400 – 1000 nm)). They are then classified into four classes : pure, insecticides with lower, medium and higher concentration. On choosing ROI and based on the mean spectral curve, the optimal wavelength is chosen. The minimum and maximum reflectance values are taken as required features for the four classes and trained through machine learning classifiers like Ensemble, Multi-class Support Vector Machine, K-nearest neighbor,Naive-Bayes and Support Vector Machine to know the classification accuracy. In deep learning the captured images were classified into 4 classes: pure, pesticides with lower, medium and higher concentrations. After band separation initially we had 616 images and these were augmented to 12,302 images. Then they are trained using 3D-SERSN(Squeezed Exicitation Residual Spectral Network) and the results are obtained accordingly.

The SE(Squeezed Excitation) block allows the network to learn channel-wise feature inter dependencies and adaptively recalibrate feature maps. Finally, the machine learning results were analysed, with KNN reporting the highest accuracy of 79.3% for the 4-class classification, and for the 2-class classification, the Decision Tree Classifier reported the highest accuracy of 84.6%. In deep learning, the 3D CNN with the SE layer reported an accuracy of 98.07%; without the SE layer, the accuracy was 77.83%. After comparing these results, we concluded that deep learning is better than machine learning in pesticide classification. Further, the results can be improved by increasing the data set.

CHAPTER 1
INTRODUCTION

1.1. INTRODUCTION

Pesticides are chemical substances that are meant to kill pests. In general, a pesticide is a chemical or a biological agent such as a virus, bacterium, antimicrobial, or disinfectant that deters, incapacitates, kills, pests. Pesticides are not recent inventions! Many ancient civilizations used pesticides to protect their crops from insects and pests. Ancient Sumerians used elemental sulfur to protect their crops from insects. Whereas, Medieval farmers experimented with chemicals using arsenic, lead on common crops. The Chinese used arsenic and mercury compounds to control body lice and other pests. While, the Greeks and Romans used oil, ash, sulfur, and other materials to protect themselves, their livestock, and their crops from various pests.

Meanwhile, in the nineteenth century, researchers focused more on natural techniques involving compounds made with the roots of tropical vegetables and chrysanthemums. In 1939, Dichloro-Diphenyl-Trichloroethane (DDT) was discovered, which has become extremely effective and rapidly used as the insecticide in the world. However, twenty years later, due to biological effects and human safety, DDT has been banned in almost 86 countries.

1.1.1 Types Of Pesticides:

Pesticides are commonly classified as insecticide, fungicide, herbicide, rodenticides, bactericide, fungicides, larvicides etc. These pesticides act against insects, rodents, bacteria, fungi, larvae and weeds which are harmful in agricultural or horticultural planting.

1.1.2 Pesticide usage:

The use of pesticides is so common that the term pesticide is often treated as synonymous with plant protection product. It is commonly used to eliminate or control a variety of agricultural pests that can damage crops and livestock and reduce farm productivity. Pesticides are essential for agricultural and horticultural crops production. Normally, farmers use the pesticides following the instruction written in the package. In most cases, the pesticides are mixed with water and sprayed over the plants. Basically, after spraying vegetables with pesticide, a period of 10 to 14 days is required to allow the chemical to degrade. However, the full degradation of pesticide is not always achieved. In recent years, some farmers ignored to use the pesticide correctly and rationally. In order to chase a better insecticidal effect and the economic interests, the phenomenon of using pesticide excessively, or selling the vegetables just after spraying the pesticide in few days are not difficult to see. Moreover, the pesticides overdosing also have the potential to contaminate the soil, air, and river.

1.2 STATISTICAL STUDY ON PESTICIDE USAGE IN INDIA:

Table 1.1 Pesticide application in Indian Agriculture

Reference year	Pesticide use (technical grade material) in ‘000’ tons
1950-51	2.35
1960-61	8.62
1970-71	24.32
1980-81	45.00
1990-91	75.00
1995-99	61.26
2000-10	57.00
2011-21	67.89

The Table 1.1[9] shows that application of synthetic pesticides has increased significantly from 1950-51 to the late twenties. Though there is a sharp decline during the later parts of the reference periods, especially in the nineties, pesticides use is still found to be high in India. Notably, total pesticides use was highest during the eighties as the period was undergoing the spiral effects of the green revolution. The continuous decrease after the eighties in the use of pesticides may be related to the fact that the farmers are increasingly aware of the adverse effects of such inputs.

The government was proactive in providing liberal policy packages to agro-chemical industries and expectedly, the production increased and in special circumstances the government provided subsidies to farmers. For instance, during 2000-01 the Union government released a sum of RS 96.2 millions of which a major share of RS 74 millions was contributed by the Union government alone to combat Eriophyide mite on coconut palms in Karnataka. In the context of globalization and trade policies, production and application of chemical pesticides have increased in different parts of the world as observed by Dasgupta and others (2001) for Brazil. In India, both the Union and state governments (obviously for political reasons and 3 According to Gulati (1990) subsidy on agrochemical (fertilizer and pesticide) is largely shared by the industry and the farmers are not net subsidized but actually taxed. gains), subsidize the pesticide price especially to the resource poor farmers to the tune of 75 per cent (GOK, 2002). While application of chemical fertilizers is largely restricted to irrigated areas, chemical pesticides are used under both wet and dry land agricultural systems. The ignorance and strong faith on hearsay, resulted in wooing the farmers to apply dreaded chemicals for which the price is being paid now. In other words, when farmers seek to expand their production frontiers, they do so by increasing the inputs used which in turn is determined by their prices. This has contributed to an increase in the production and use of agro-chemicals.

1.2 NEGATIVE IMPACT OF PESTICIDES

In India, the Central Food Technological Research Institute's study shows that of 204 samples of cereals, pulses, milk, eggs, meat and vegetables, 108 were found to contain pesticide residues. A number of research studies in India have focused their attention on the impact of agro-chemicals in a more general perspective. Here, a couple of case studies on the adverse implications of pesticide spraying have been taken up for a closer examination.

A physician at the Indian Institute of Health Management Research in Jaipur, Rajasthan, reported that newborn children have neural tube defect (NTD), a deformity that results from the incomplete closure of the neural tube during early pregnancy. Alarmingly NTD takes a heavy toll of the order of half a million babies every year in the world and in Rajasthan alone about 8,000 babies are reportedly affected.

Once the pesticides are sprayed, the residues and particles directly enter into the humans through the food chain. Another notable observation is that around 500 times of MRL of endosulfan is noticed in animals' skin/tissue and these explain the maladies of chemical pesticides on the environment and living organisms, being a consequence of the governments sponsored environmental terrorism.

Another case of indiscriminate pesticide dumping by a state owned department resulted in serious environment and economic revelations in the highly literate state of Kerala). Way back in the 1970s, the Plantation Corporation of Kerala (PCK) began aerial spraying of pesticides on its 2,200 ha cashew gardens in Padre village in Kasargod district mainly to check pests during fleshing, flowering and fruit setting seasons. Continuous and indiscriminate application of synthetic pesticides affected the flora and fauna, and local people became victims of severe health problems like cancer.

Table 1.2 HEALTH DISORDERS IN PADRE VILLAGE OF KERALA

Disorders	No.of cases
Cancer	49
Mental retardation	23
Congenital anomalies	09
Psychiatric cases	43
Epilepsy	23
Suicides	09
Total	156

The Table 1.2 also indicates that apart from 9 suicides, a greater incidence of cancer, psychiatric problems , epilepsy are have been reported in the village of Kerala.This reveals how the environment and human beings suffer aerial spraying of Pesticide.

1.4. PENETRATION LEVEL OF PESTICIDES ON FRUITS AND VEGETABLES SURFACE:

The retention of pesticides depends on the physicochemical properties of the pesticide molecules as well as food.Pesticides may be introduced to fruits and vegetables during different phases of production. Some pesticides are used before blooming, some while fruits are growing and others after harvesting. Therefore the location of pesticides in the same fruit may be different. Very little quantity of systemic pesticides may be absorbed into flesh.

In fruits and fruit-type vegetables, the concentration of pesticide residue was higher in the fruit stalk and near the epidermis than in the sarcocarp or pericarp. In leaf vegetables, concentration of the pesticide residue was higher in outer leaves than in inner ones. The leaching of pesticides from the surface of fruits and leafy commodities is due to their solubility in water . The effectiveness of washing is also dependent and may be reduced for insecticides, specifically on synthetic pyrethroids, due to strong bonding between the insecticide molecules and waxy layer of fruit skin and also their non-systemic and non-translaminar movement characteristics.

The incidence of excessive pesticide residues may cause blindness, cancer, diseases of liver and nervous system etc. The long term effects could result in reduction of live sperm and fertility, increase in cholesterol levels, high infant mortality rates and several metabolic and genetic disorders.

1.5 EXISTING METHODS FOR PESTICIDE DETECTION

In recent years, there has been an increase in pesticide use to improve crop production due to the growth of agricultural activities. Currently, several different technologies such as gas chromatography (GC), high performance liquid chromatography (HPLC), thin-layer chromatography, supercritical fluid chromatography, chromatography-mass spectrometry, capillary electrophoresis, enzyme inhibition method, immunoassay method, and bio-sensor method are used to determine the concentration of pesticide residue. Chromatographic methods that are commonly used in determination and separation of target pesticides and herbicides include gas chromatography (GC) and liquid chromatography (LC). Mass spectrometry (MS) is used as an additional technique coupled with GC and LC in order to enhance the detection performance.

Optical detection for pesticide determination represents the highest percentage of flow methods designed. Spectrophotometry and luminescence are the most frequently employed. In particular, fluorescence is the most commonly used luminescence technique due to its higher selectivity and sensitivity, when compared to spectrophotometry. Optical flow sensors are based on the implementation of solid phase spectroscopy (SPS) in flow analysis. They are usually named flow-through opto sensors or flow opto sensors (FOs). Development of safe, fast, reliable and low-cost analytical methods for the determination of pesticide residue that avoids the use of organic solvents, and reduces the contact of operator with the toxic substances is growing interest at present.

1.6 IMPORTANCE OF PESTICIDE DETECTION:

Pesticides are an effective tool used by farmers all across the globe to prevent insects, rodents and other pests from eating or spoiling their crops. In recent years farmers are preferring pesticides rather than organic manure to increase the productivity of their crops in short time. As they use pesticides beyond the limit it affects the crop quality as well as consumer's health. So pesticide detection plays a vital role in consuming healthy vegetables. Unfortunately, there are not many pesticide detection methods. Thus, this makes pesticide detection an important thing to be noted and implemented. However, it can be difficult for communities to detect pesticides, and it is not as effective to test pesticide use after it has already been applied to a field or farm. For these reasons, there have been a number of international bodies that have joined together to create testing guidelines to ensure that the pesticides used in the U.S., Australia, Africa and anywhere else will not result in unwanted repercussions for humans nearby.

Additionally, the WHO Pesticide Evaluation Scheme pointed out that the reason using quality pesticides was important was because it is the only way to ensure safe use. "Good product quality is essential for effective and safe pesticide use. Impurities formed during manufacture of the pesticide or by interaction in unstable formulations can increase product toxicity to humans and the environment," the organization explained. When trying to regulate such a useful and potentially dangerous substance across every country in the world, the best tool that international bodies can turn to are well-respected standards providers for testing the quality of a pesticide to ensure that it will be safe and effective.

1.7 HYPER SPECTRAL IMAGING IN PESTICIDE DETECTION

Pesticide specific position. With hyper spectral imaging, a spectrum for each pixel can be obtained and a gray scale image for each narrow band can be acquired, enabling this system to reflect componential and constructional characteristics of an object and their spatial distributions.

In this project the hyper-spectral image of the samples collected will be acquired using the hyper-spectral spectral camera (Pika L (400 – 1000 nm). Hyper-spectral reflectance data will be acquired with a pixel size of 5.8 µm. Regarding spectral data processing, the original spectral data consisted of 300 spectral bands from 400 nm to 1000 nm with spectral resolution 3.2 nm.

This project aims at developing a spectroscopy based estimation and quantification of insecticides residues in Lady's finger, which is regarded as a potential method to solve the above problems.

1.8 SUMMARY:

In this chapter, agriculture production and pesticide usage in India is discussed. It also provides statistical analyses of the adverse effects of pesticide and its amount of usage in fruits and vegetables. In recent years pesticides have been used more and to reduce it's effect, our proposed model helps us to detect the level of insecticide concentration in lady's finger . Types of pesticides used to increase production is also discussed. Though the pesticides are harmful it can be used in critical situation to increase the production and also it is helpful for improvising the economic status of a farmer. This chapter emphasis the need for the pesticide detection and some of the existing methods to detect the pesticide in fruits and vegetables. It describes the role of Hyperspectral imaging techniques to detect the pesticide in fruits and vegetables. Thus in the upcoming chapters, the proposed model and it's usage to detect the pesticides is discussed in detail.

CHAPTER 2
LITERATURE SURVEY

2.1. LITERATURE SURVEY:

1. TITLE: Hyperspectral Imaging for Predicting the Internal Quality of Kiwifruits Based on Variable Selection Algorithms and Chemometric Models.

AUTHOR: Hongyan Zhu, Bingquan Chu,Yong He

PUBLISHED YEAR: 10 August 2017

PUBLICATION: Scientific Reports-Natrural Conference.

METHOD: The paper employs hyperspectral imaging to capture detailed data from kiwifruit samples and uses variable selection algorithms to identify key spectral features related to quality parameters like sugar content and firmness. These features are integrated into chemometric models like partial least squares regression (PLSR) or principal component regression (PCR) for prediction. The hyperspectral data undergo preprocessing for noise removal and normalization. Validation and calibration using cross-validation techniques ensure the models' accuracy and reliability. This approach provides a non-destructive and efficient method for assessing kiwifruit quality, offering valuable industry insights.

PESTICIDE: Nil

PARAMETERS: Determining firmness, soluble solids content (SSC), and pH in kiwifruits using hyperspectral imaging, combined with variable selection methods and calibration models.

HARDWARE: Hyperspectral imaging system.

ADVANTAGES: The results clearly demonstrated that hyperspectral imaging has the potential as a fast and non-invasive method to predict the quality attributes of kiwifruits.

2. TITLE: Hyperspectral near-infrared imaging for the detection of physical damages of pear.

AUTHOR: Wang-Hee Lee, Moon S. Kim, Hoonsoo Lee ,Stephen R. Delwiche.

YEAR: 9 January 2014

PUBLICATIONS: ELSEVIER

METHOD: In this study, near infrared hyperspectral imaging in the 950–1650 nm region was firstly used to investigate the feasibility of multispectral reflectance ratio imaging techniques for detection of bruise damages on ‘Shingo’ pear. Simple ANOVA classification, a novel way to analyze hyperspectral imaging, was explored to select waveband ratio and threshold values for optimal classification of bruised pears.

PESTICIDE: Control group, chlorpyrifos treatment group, carbendazim treatment group, chlorpyrifos and carbendazim combination group.

PARAMETERS: Investigate the detection of bruise damages on 'Shingo' pears. The study used a classification algorithm based on F-value to determine the optimal waveband ratio for discriminating between bruised and sound areas.

HARDWARE: ZOLIX HyperSIS-VNIR-PFH hyperspectral equipment.

ADVANTAGES: The optimal waveband ratio identified using F-value analysis achieved a bruise detection accuracy of 92%.

3. **TITLE:** Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning.

AUTHOR: Pengcheng Nie a, Jinnuo Zhang a, Xuping Feng a, Chenliang Yu b, Yong He a.

YEAR: 19 JUNE 2019

PUBLICATIONS: ELSEVIER

METHOD: Here, near-infrared hyperspectral imaging technology combined with deep learning was applied to classifying hybrid seeds. The hyperspectral images in the range of 975–1648 nm of a total of 6136 hybrid okra seeds and 4128 hybrid loofah seeds, which both contained six varieties, were collected. To establish discriminant analysis models, partial least squares discriminant analysis, support vector machine and deep convolutional neural network (DCNN) were used and their performances were compared among the different hybrid seed varieties. The discriminant analysis model based on the DCNN was the most stable and had the highest classification accuracy, greater than 95%.

VARIETIES: 2014HK2, 2014HK4, 2014HK16, baiguo, cuizhi, and danzhi, and the hybrid loofah seeds.

PARAMETER: The spectral curves of hybrid okra seeds and hybrid loofah seeds were quite different, owing to the differences in the composition of the seed epidermises. The spectral curves of hybrid loofah seeds were similar to those of lotus in the NIR region of 1000–1700 nm.

HARDWARE: Near-infrared hyperspectral imaging equipment.

ADVANTAGES: Rapid and efficient selection of eligible hybrid progeny.
Reduction in labor and time required for progeny selection
Accelerates progress in hybrid breeding research.

4. TITLE: Hybrid Okra Seed Identification

AUTHOR: Zeyu Yu, Hui Fang, Qiannan Zhangjin, Chunxiao Mi, Xuping Feng, Yong He

YEAR: 19 OCTOBER 2021

PUBLISHER: ELSEVIER

METHOD: The primary method used in the research involves hyperspectral imaging combined with deep learning algorithms, including Extreme Learning Machine (ELM), Back Propagation Neural Network (BPNN), Stacked Sparse Auto-encoder (SSAE), and Convolutional Neural Network (CNN). These methods were utilized to identify and classify hybrid okra seeds based on their spectral characteristics.

ADVANTAGES: Non-destructive and fast seed identification.
Provides both spectral and spatial information. Enables identification based on internal composition.
DL algorithms offer high precision and robustness.

ISSUES: Initial investment in equipment can be expensive.
Requires expertise in handling and maintenance. Sensitivity to environmental conditions.

HARDWARE: Hyperspectral Imaging Equipment

2.2 INFERENCE

1. In the above papers, there is not any proper mention of dataset for any hyperspectral images, in our project hyperspectral dataset will be created by using hyperspectral imaging using Resonon.
2. From the given literature survey, there is no method proposed to find the different concentration of pesticide in vegetables.
3. From the papers referred, it is inferred that those papers have just worked only on different techniques to differentiate and classify good quality fruits and vegetables from the bad ones whereas the proposed project works on detecting the pesticide level on fruits and vegetables and helps us eliminate the harmful ones.
4. In earlier discussed papers, datasets are collected in such a way that fruits and vegetables are stored in different temperatures in order to change it's texture and appearance so that using PCA, deep learning network and techniques it can separate the bad ones from good ones.
5. Also, in some papers, samples are collected in such a way that the fruits and vegetables are sprayed with pesticides, which sometimes change the texture and outer appearance, but the amount of pesticide is not known, and it gets classified under the bad ones, whereas the proposed project detects the amount of pesticide, whether it has resided on the outer surface alone or whether it has penetrated deep inside or not.

CHAPTER 3
PROBLEM DESCRIPTION

3.1 INTRODUCTION:

In recent years the amount of pesticides we consume through vegetables are high.According to reports, children consuming high level of pesticides leads to childhood cancers, Attention Deficit Hyperactivity Disorder(ADHD), autism and so on. Thus this proposal helps us find those vegetables which have high level of pesticides.

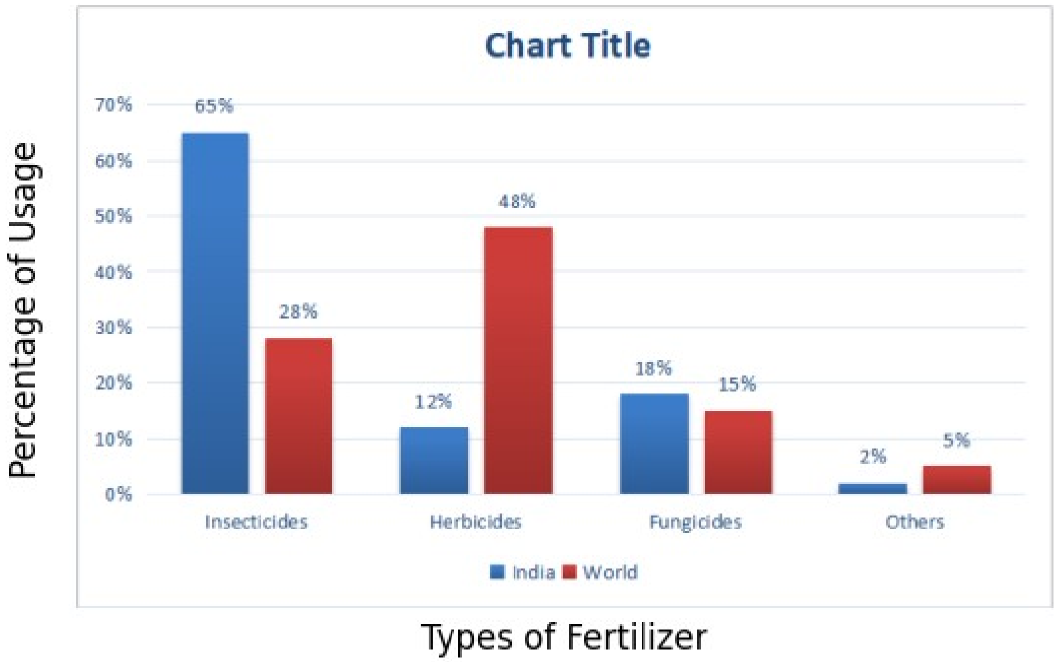

Fig.3.1 COMPARISION OF PESTICIDE USAGE IN INDIA WITH OTHER PARTS OF THE WORLD[10]

3.2 PROBLEM DESCRIPTION:

Nowadays, the use of pesticide in vegetables to retain their freshness, is enormously increasing beyond their actual content. This kind of actions may lead to cause adverse effects on the human health and cause several diseases like rashes, nausea, dizziness etc.

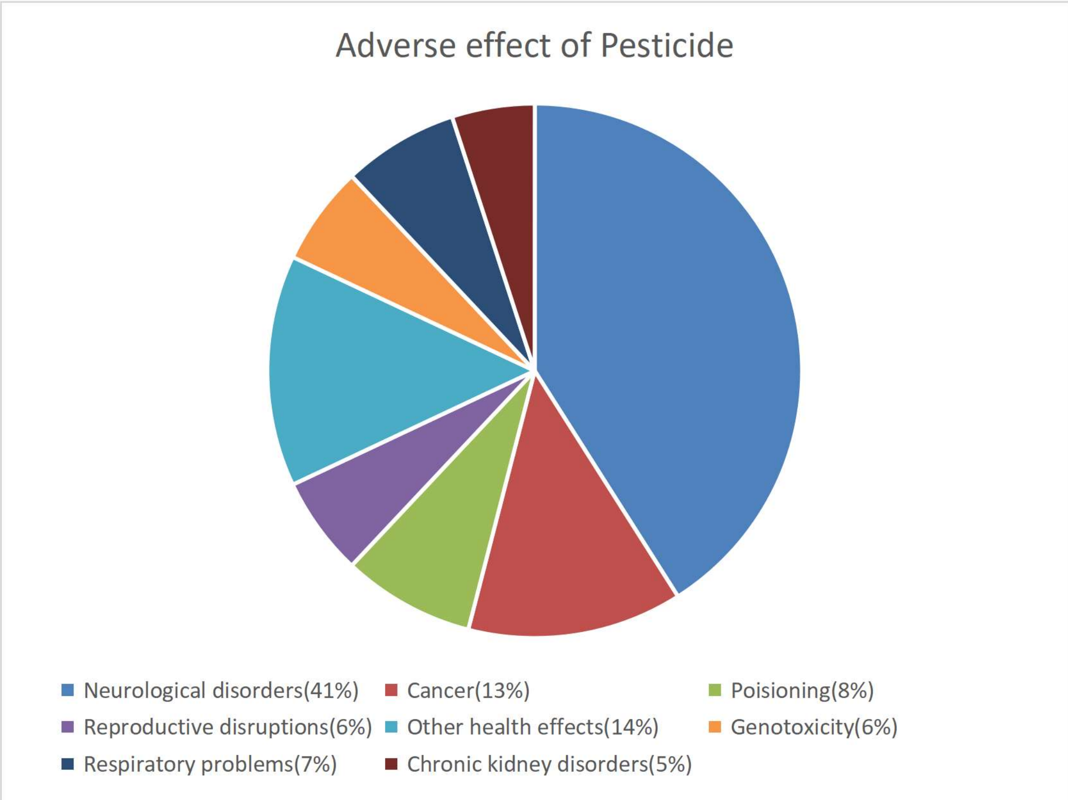

Fig.3.2 PERCENTAGE OF EFFECTS OF PESTICIDE

3.3 OBJECTIVES:

The overall objective of this research is to utilize hyperspectral imaging, for detection of pesticide residues in lady's finger.

The specific objectives for this project are:

- To prepare samples that contain different pesticide residues at different ratios and to create a hyperspectral image database.
- To develop Hyperspectral image analysis methods for estimating residues

of pesticides in lady's finger.

- To develop a machine learning and deep learning model that estimates the pesticide residue.

SUMMARY:

In this chapter the current scenario of pesticide usage in India was compared with the other parts of the world and discussed using bar graph. Pesticides and the diseases caused by and how far we are affected are also discussed using the pile chart. Objective of this proposed model was also discussed. The main objective is to avoid consuming unhealthy pesticide added lady's finger and also to know pesticide level in the lady's finger and also to detect whether the vegetable is with pesticide or not.

CHAPTER 4

DATABASE ACQUISITION AND DATBASE CREATION

In this digital era, image databases are highly essential to learn the features of an object and to assess the performance of object detection, classification and recognition techniques. This chapter discusses about existing benchmark image databases, need for lady's finger database, creation of lady's finger database, challenges involved in lady's finger database, and finally highlights its application towards scientific research.

4.1 INTRODUCTION:

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised. The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, and then that a network with a non-polynomial activation function with one hidden layer of unbounded width can on the other hand so be. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, whence the "structured" part. Therefore creating a relevant hyperspectral imaging database for the research work is the initial work.

4.2 HYPERSPECTRAL DATABASE:

As a first step we created a database for the lady'sfinger for image samples without pesticides and mixed with pesticides at low concentration, medium concentration and high concentration.Hyperspectral data delivers reflectance information in hundreds of bands, enabling in-depth examination and discrimination of material spectra of terrestrial features. Hyperspectral image classification is the task of classifying a class label to every pixel in an image that was captured using (hyper) spectral camera.

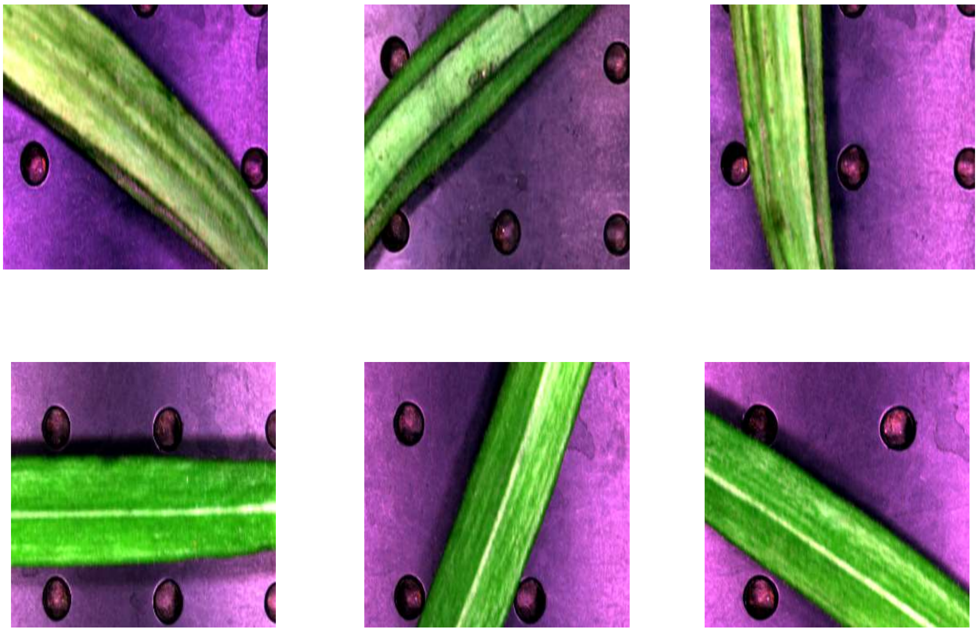

Fig.4.1 HYPERSPECTRAL IMAGE SAMPLES OF LADYSFINGER

4.3 Image Acquisition system:

4.3.1 System resonon:

Resonon's hyperspectral imagers are line-scan imagers (also referred to as push-broom imagers). Two-dimensional images are constructed by translating the sample relative to the camera. This is typically accomplished by placing the sample on a linear translation stage. Bench top hyper spectral imaging system is comprised of a Pika hyperspectral imaging camera, linear translation stage, mounting tower, lighting assembly, and software control system. The positions of the imager and lighting assembly are adjustable along the length of the tower.

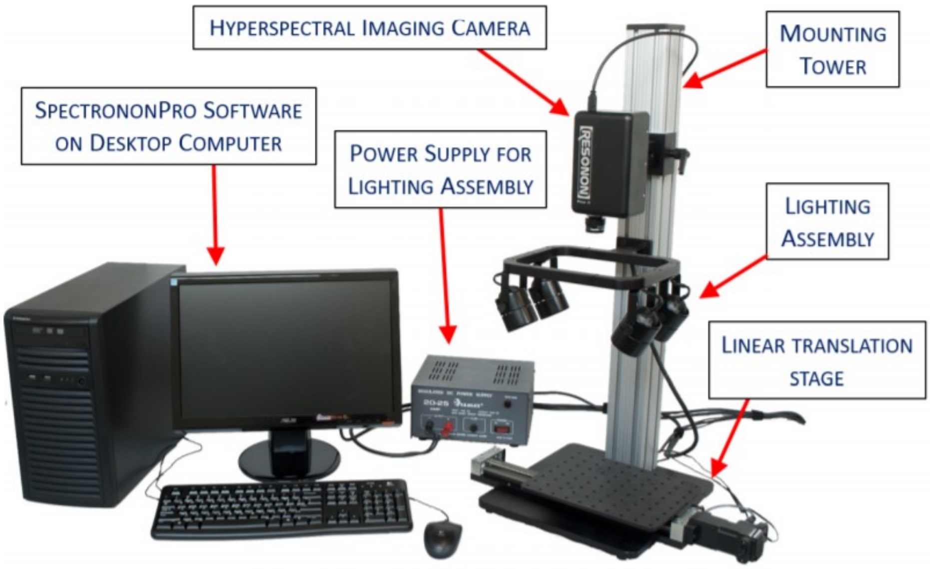

Fig. 4.4 RESONON HYPERSPECTRAL CAMERA

The Pika L hyperspectral camera is a Visible+Near-Infrared (VNIR) imager covering the 400 - 1000 nm spectral range. The Pika L weighs only 0.6 kg, making it ideal for drone-based remote sensing applications.The Pika L can also be used with laboratory and outdoor hyperspectral systems.

4.3.2 DATABASE CREATION:

For the dataset collected without the use of pesticides, a total of 175 image samples of lady's finger were carefully captured using a high-tech hyperspectral camera, ensuring detailed and accurate imaging. In contrast, the dataset containing pesticide-treated samples involved a meticulous process of mixing the lady's finger samples with varying concentrations of pesticides, specifically at 1ml, 3ml, and 5ml per liter of water, representing low, medium, and high concentrations, respectively. This methodical approach aimed to simulate real-world scenarios where lady's finger crops are exposed to different pesticide levels to study their effects comprehensively.

In the pesticide-treated dataset, specific attention was paid to capturing image samples of lady's finger mixed with pesticides at low concentrations, totaling 220 images. Each image was expertly taken using the hyperspectral camera to ensure precise data collection for subsequent analysis and comparison. Similarly, for lady's finger samples mixed with pesticides at medium concentrations, a set of 100 image samples were diligently captured. The use of the hyperspectral camera facilitated the acquisition of detailed images that would serve as valuable data points in the research process.

Furthermore, the dataset included lady's finger image samples mixed with pesticides at high concentrations, comprising a total of 120 images. These images were meticulously captured through the hyperspectral camera to provide a thorough understanding of the impact of high pesticide concentrations on the lady's finger crops. The choice of the insecticide M Power for treating the lady's finger samples was deliberate, as it allowed for consistency and control in the experimental setup, ensuring reliable results and meaningful insights into the effects of pesticides on the crop.

Overall, the systematic approach taken in collecting and categorizing the image samples based on pesticide concentrations reflects a rigorous methodology designed to investigate the response of lady's finger crops to pesticide exposure. Through the detailed imaging process and thoughtful selection of pesticide concentrations, the dataset provides a rich resource for further analysis and research in the field of agricultural studies.

Table 4.1 DATASET DETAILS

LADYSFINGER IMAGES			
Without pesticides	Lower concentration	Medium concentration	Higher concentration
175	220	100	120
AFTER AUGMENTATION			
3400	4400	2000	2400

4.4 MOTIVATION:

In the existing works, there is not any proper mention of dataset for any hyperspectral images, in our project hyperspectral dataset will be created by using hyperspectral imaging using Resonon. From the papers referred, there is no method proposed to find the different concentration of pesticide in vegetables. It is inferred that those papers have just worked only on different techniques to differentiate and classify good quality vegetables from the bad ones whereas the proposed project works on detecting the pesticide level vegetables and helps us eliminate the harmful ones. In earlier discussed papers, datasets are collected in such a way that vegetables are stored in different temperatures in order to change it's texture and appearance so that using PCA, deep learning network and techniques it can separate the bad ones from good ones. Also in some papers samples are collected in such a way that the vegetables are sprayed with pesticides which sometimes changes the texture and outer appearance of them but it doesn't let us know the amount of pesticide and it gets classified under the bad ones whereas the proposed project detects and let us know the amount of pesticide whether it has been resided on the outer surface alone or it has penetrated deep inside or not.

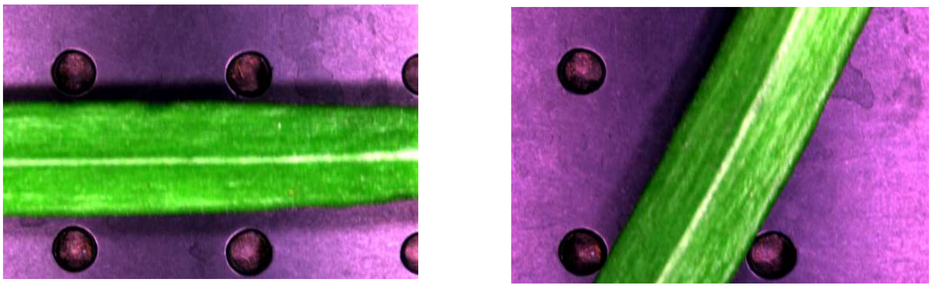

Fig.4.3 LADYSFINGER WITH AND WITHOUT PESTICIDE

4.5 CREATION OF LADY'S FINGER DATABASE:

This proposed project concentrates on the estimation of pesticide levels in lady's finger to avoid harmful pesticides. The key contributions of lady's finger database generation are:

- Hyperspectral image acquisition with proper setup and calibration.
- Acquisition of lady's finger samples with pesticides at different concentration level.
- Augmentation of the samples.
- Evaluation of the vulnerability of collected samples

4.5.1 IMAGE AUGMENTATION:

As the number of images collected from hyperspectral imaging camera are less in number that is for example initially we had 175 images in pure_lady's finger and to further increase the number of samples we augment the obtained samples. As a result of augmentation, we get 3500 images of samples without pesticides, 4400 images of samples at lower concentration, 2000 images of samples at medium concentration and 2402 images of samples at higher concentration, totally 12,302 images of lady's finger are obtained. Augmentation is a process done to increase the dataset. In the project after band separation there were 615 images initially and they are augmented to 12,302 images in each class. Each time while augmenting different features were added. It includes flip, rotate, noise, shear, auto-orient, resize and so on. The below images seen the augmented images with different features.

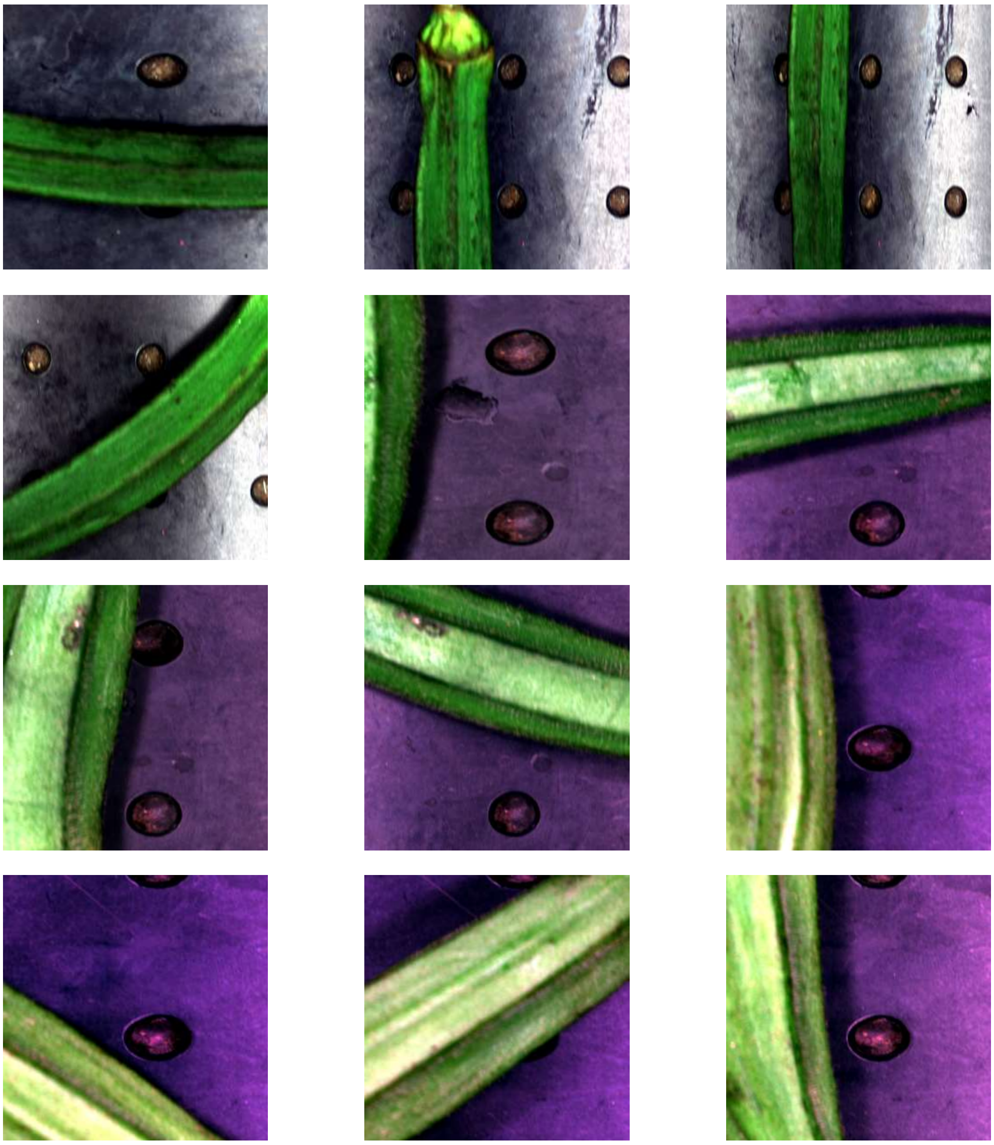

Fig.4.5 AUGMENTED IMAGES OF LADYSFINGER

4.6 CHALLENGING ASPECTS OF LADY'S FINGER DATABASE:

1. If an object is viewed under different view angle or rotation, then the object may look different. It is observed that when the images of lady's finger are viewed at different angle or position or pose variation occurs.

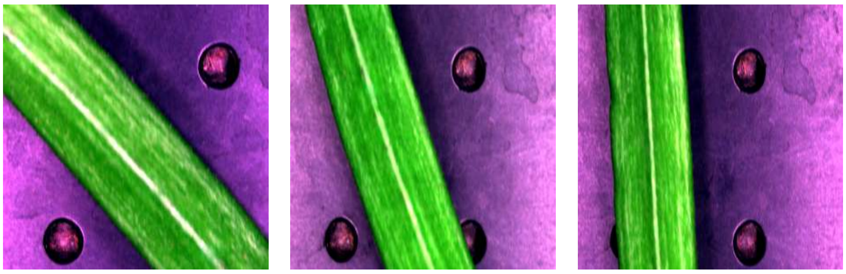

Fig.4.6 IMAGES OF SAMPLES AT DIFFERENT ROTATIONS

2. While taking the mean spectrum of the obtained images, band selection becomes quite challenging because the spectrum for the pesticide estimation is not recognised clearly.

4.7 CONCLUSION:

In this chapter, an effort has been put forward to generate a lady's finger database. The proposed database consists of 615 images for lady's finger and is highly structured and manageable.

4.8 SUMMARY:

In this chapter sample preparation and dataset collection were discussed in detail. The samples were created with three different concentration of pesticide for lady's finger. This chapter gives the complete description about the RESONON hyperspectral camera which is used to acquire hyperspectal image. Hyperspectral image obtained contists of 300 band images obtainted over the range of 400-1000nm. ROI is selected for the obtained image and band separation is done. The obtained hyperspectral image is augmented to increase the number of images.

CHAPTER 5
METHODOLOGY

5.1. INTRODUCTION:

The fields of reflectance spectroscopy are based on the principle of electromagnetic radiation being reflected from a material and then detected by a sensor. Reflectance spectroscopy can be applied for agricultural products across many different wavelengths of the electromagnetic spectrum, each with different strengths and weaknesses. The advantages of the optical range are that it is a passive technology; there is a range of sensors available, and it offers a cost effectiveness and repeatability not available from other technologies. The progression of improved spectral and spatial resolution has allowed for continued development in the application of reflectance spectroscopy in many areas. The development from multispectral to hyperspectral sensing has given users increased diagnostic power allowing for the detection and discrimination of more and more of the material's features. With these advantages the pesticide residue analysis system is proposed and given below.

Two methodologies were implemented, they are machine learning and deep learning. In the proposed model, the obtained data-set consisting of four classes (i.e) lady's finger without pesticide, with pesticide at three different level of concentration low, medium and high are trained using the Machine Learning and Deep Learning model.

5.2 MACHINE LEARNING PESTICIDE RESIDUE ESTIMATION:

FLOW CHART

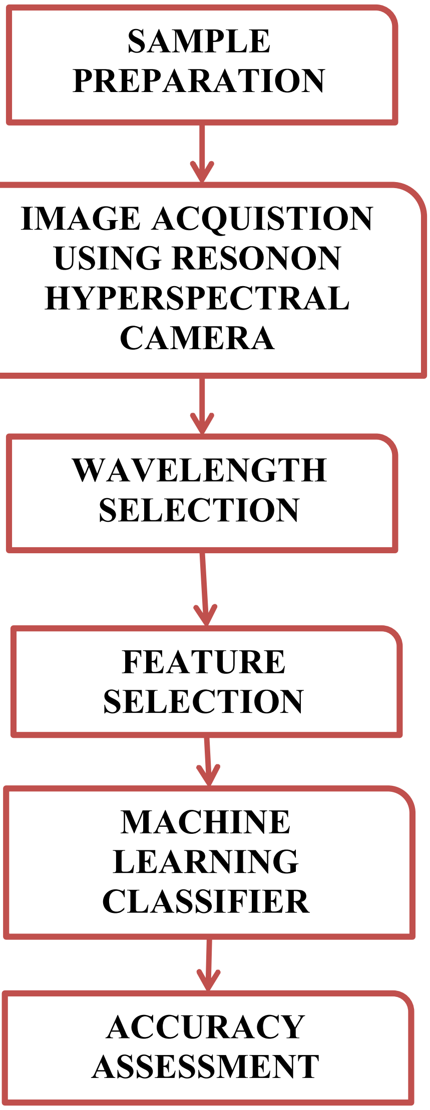

Fig.5.1 FLOWCHART OF MACHINE LEARNING PROCESS

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and

adjust actions accordingly. Machine learning classification can be categorized into two types : Supervised classification and Unsupervised classification.

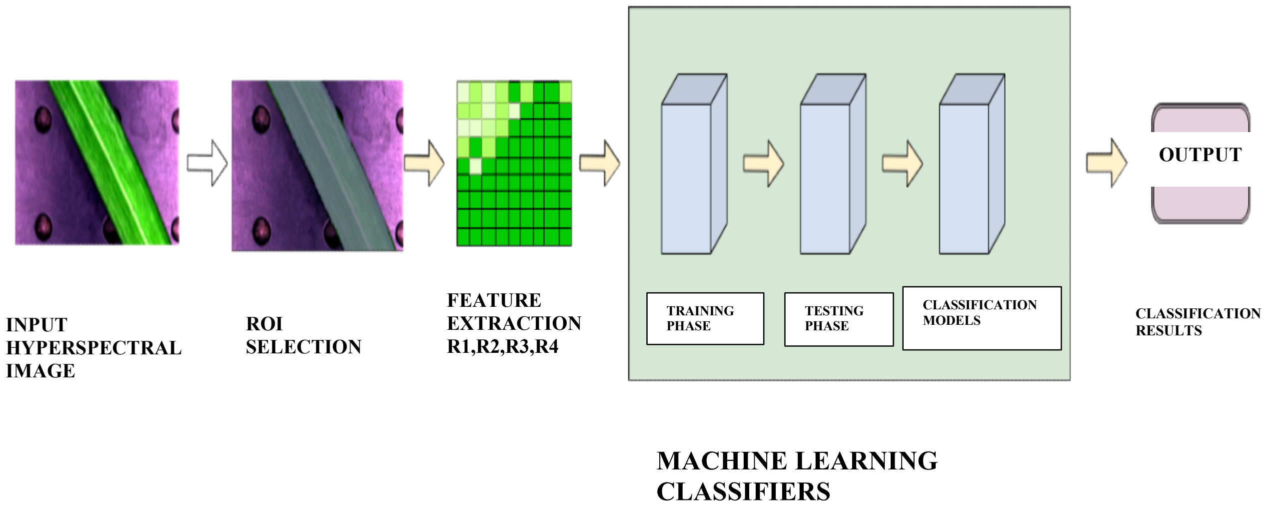

Fig.5.2 BLOCK DIAGRAM OF MACHINE LEARNING

Fig.5.1 Features extracted from ROI of hyperspectral images is given as input to the machine learning classifier model and classification results are obtained.

5.2.1 SAMPLE PREPARATION:

A pesticide sample is prepared with M power being used as an insecticide. For lower concentrations, 1L of water and 1ml of pesticide are added; for medium concentrations, 1L of water and 3ml of pesticide are added; and for higher concentrations, 1L of water and 5ml of pesticide are added for both samples.

5.2.2 WAVELENGTH SELECTION:

Wavelength selection involves, choosing Region Of Interest(ROI) for the hyperspectral images of Lady's finger and finding mean spectrum. On comparing the mean spectral curves for the images with and without pesticide a mild difference is noticed in visible and near IR region. So the optimal wavelength chosen are for Lady's finger the region between 388.85 to 660nm (Visible region) and 720 to 1000nm(NIR region) were selected.

5.2.3 FEATURE EXTRACTION:

After wavelength selection form the mean spectrum the features were obatined for the corresponding reflectance values. The obtained feature consists of the minimum and maximum reflectance between the selected wavelength region for both the upper and lower band regions. On comparing the reflectance values R_{min} and R_{max} values for without pesticide and with pesticide(three concentrations lower, medium and high) are chosen as required feature is extracted. At first 100 features for lady's finger for four class classification were extracted out of which 80 features were used for training and remaining 20 were used for testing.

5.2.4 CLASSIFICATION:

The extracted features are given to 5 classifiers and the accuracy results are assessed. The five classifiers are:

1. Ensemble classifier
2. Support Vector Machine classifier
3. K-Nearest Neighbour classifier
4. Naive Bayes classifier
5. Logistic Regression

1) ENSEMBLE CLASSIFIER:

Ensemble can boost or bag decision tree learners or discriminant analysis classifiers. The function can also train random subspace ensembles of KNN or discriminant analysis classifiers. For simpler interfaces that fit classification and regression ensembles, instead use fitcensemble and fitrensemble, respectively. Also, fitcensemble and fitrensemble provide options for Bayesian optimization.

2) SUPPORT VECTOR MACHINE CLASSIFIER:

Support Vector Machine basically helps in sorting the data into two or more categories with the help of a boundary to differentiate similar categories. Decision Boundary is the main separator for dividing the points into their respective classes. The equation of the main separator line is called a hyperplane equation. The hyperplane equation dividing the points (for class

$$H: w^T(x) + b = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.1$$

Here: b = Intercept and bias term of the hyperplane equation

3) K - NEAREST NEIGHBOUR CLASSIFIER:

k- Nearest Neighbor (or kNN) is a supervised machine learning algorithm useful for classification problems. It calculates the distance between the test data and the input and gives the prediction according. kNN calculates the distance between data points. For this , we use the simple Euclidean Distance formula.

$$d(p,q)=d(q,p)=\sqrt{}(q1-p1)^2+(q2-p2)^2+\ldots.+(qn-pn)^2$$
$$=\sqrt{}\sum(qi-pi)^2 \quad (i=1 \text{ to } n) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.2$$

The above formula takes in n number of dimensions or here we can say them as our features in machine learning.The data point which is located at the minimum distance from the test point is assumed to belong to the same class. The above formula works the same in n number of dimensions and therefore it can be used with n number of features.

4) NAIVE BAYES CLASSIFIER:

A group of classification algorithms based on Bayes Theorem are known as naive Bayes classifiers. It is actually a family of algorithms rather than a single method, and they are all based on the same principle—that is, each pair of features being classified stands alone.

Bayes Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation:

$$P(A/B) = (P(B/A) * P(A))/P(B) \ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.3$$

Where A and B are events and $P(B) \neq 0$

Applications of the Naive Bayes method include classification issues. In the classification of texts, it is often employed. Data in text classification tasks has a high dimension since each word in the data represents a feature. It is applied to rating categorization, sentiment analysis, spam filtering, and other tasks. The quickness of naive Bayes is an advantage. With a large number of data dimensions, it is quick and simple to make predictions.

5) LOGISTIC REGRESSION:

For classification problems where the objective is to predict the likelihood that an instance belongs to a specific class or not, one supervised machine learning approach that is utilized is called logistic regression.

When using the sigmoid function, which accepts input as independent variables and outputs a probability value between 0 and 1, logistic regression is utilized for binary and multi class classification. The ratio of something happening to nothing happening is the odd. It differs from probability since the latter is the ratio of the likelihood of an event to all possible outcomes. so peculiar will be:

$$p(x)/(1 - p(x)) = e^{z}$$

Applying natural log on odd, then log odd will be:

$$\log[p(x)/(1 - p(x)] = z$$

$$\log[p(x)/(1 - p(x)] = w \,.\, X + b$$

$$[p(x)/(1 - p(x)] = e^{w.X + b} \text{ ....Exponentiate both sides}$$

$$p(x) = e^{w.X + b}.(1 - p(x))$$

$$p(x) = e^{w.X + b} - e^{w.X + b} \,.\, p(x)$$

$$p(x) = e^{w.X + b}/(1 + e^{w.X + b})$$

then the final logistic regression equation will be:

$$p(X; b, w) = 1/(1 + e^{-w.X + b}) \text{ .... } 5.4$$

5.3 DEEP LEARNING ALGORITHM:

FLOW CHART

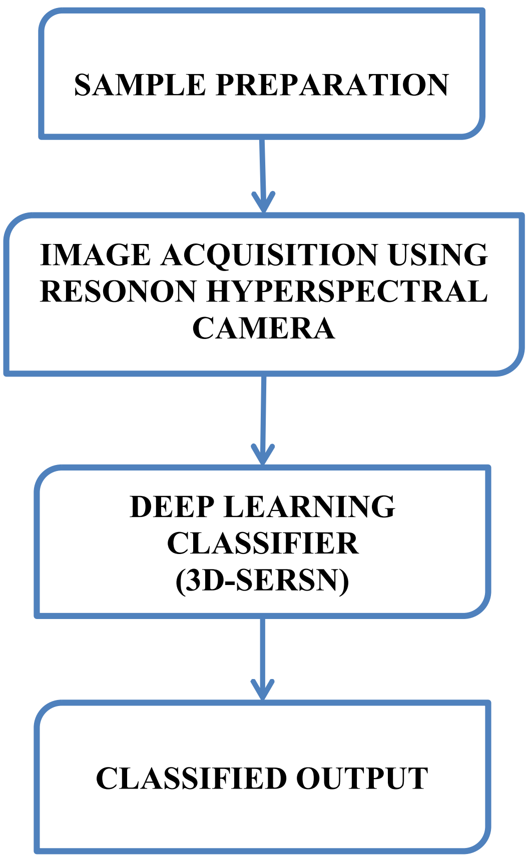

Fig.5.3 BLOCK DIAGRAM OF DEEP LEARNING

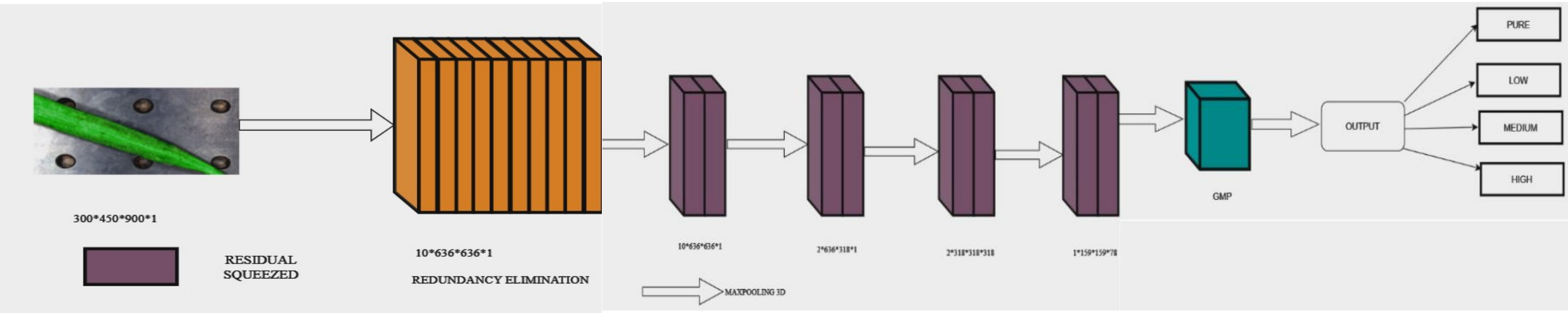

Fig.5.3 The above figure explains the procedure of proposed 3D-SERSN network

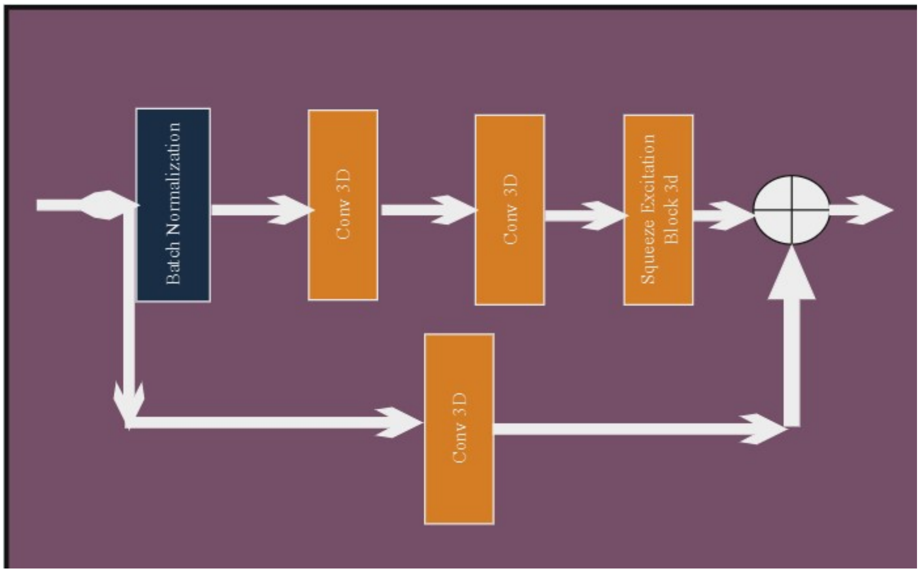

Fig.5.4 3D Squeeze Excitation Residual Spectral Network

5.3.1 Input data cube

The HSI Lady's finger data cube was captured using a Resonon Pika L hyperspectral imager with 300 spectral channels across a spectral range of 400 nm to 1000 nm, resulting in a spectral resolution of approximately 3.3 nm. One data cube was extracted with a size of 450x900x300 from the original HSI and is resized to 256x384x300 for easier computation.

5.3.2 Redundancy elimination

Spectral-based classifiers suffer from feature redundancy and noise in the classification performance due to the large number of spectral bands in the HSI data, especially when the number of training samples is small. The redundancy problem in the spectral channels and the pixel inconsistency in the spatial neighborhoods is solved by using PCA-based Explained variance. It is a statistical measure of how much variation in a dataset can be attributed to each of the principal components (eigenvectors) generated by the principal component analysis (PCA) method. Explained variance can be represented as a function of the ratio of related eigenvalues and the sum of eigenvalues of all eigenvectors. The larger the variance explained by a principal component, the more important that component is. The Cumulative Explained Variance (CEV) shows the accumulation of variance for each principal component number. Fig.6.7 shows the CEV plot obtained for the acquired lady's finger HSI datacube. Here the first 10 PCs account for 99% of the cumulative and accumulate much information, these 10 bands (10x256x384) are used for further processing.

The details of the layers of the proposed 3D SERSN are as follows. The proposed network has four Residual Squeeze blocks. After applying PCA, the input HSI datacube dimension is reduced to 10x256x384x1. At the beginning, we use 32 filters to compress the image to 10x256x384x32 by performing 3D convolution with 32 filters of size 3x3x3x32. Similarly in consequent RS stages 32,64,96 filters have been used. Here, the number of residual blocks and compression channels are adjustable. Following the residual blocks, a Global Max Pooling is used to transform the feature map into a one-dimensional vector. Finally, through softmax , the prediction labels corresponding to each category are obtained.

5.3.3 SQUEEZED RESIDUAL CONNECTION

A residual connection adds a shortcut by identity mapping, forcing the network to learn the residual function to restore the original non-linear transformation. Residual connections can excellently improve the flow of data between the top and bottom of the network and can reduce the over-fitting problem. In addition, residual networks are easier to optimize. In each residual squeeze block, batch normalization, 3D convolution and Squeeze -Excitation(S are performed. 3D-CNN uses 3D kernels for the 3D convolution operation and can extract spatial features and spectral features simultaneously. In the RS block, the Squeeze-Excitation block is plugged in after the final convolutional layer, in the block before the addition of the residual in the skip connection. The intuition behind this is to keep the skip connection branch as clean as possible to make learning the identity easy. Motivated by the idea of re-calibration of the SE structure, the SE trains the network to suppress or motivate features at a certain position, which can effectively resist noise interference and improve the classification result.

5.3.4 Classifier – 3D CNN:

An extension of the conventional CNN, a three-dimensional convolutional neural network (3D CNN) is made to process three-dimensional input, such as volumetric images or videos. A 3D CNN processes three-dimensional data with filters that also slide over the depth dimension, allowing it to capture spatiotemporal features in videos or 3D structures in volumetric images. A standard CNN processes two-dimensional data (like images) using filters that slide over the width and height dimensions.

5.3.4.1 Architecture of 3D-CNN:

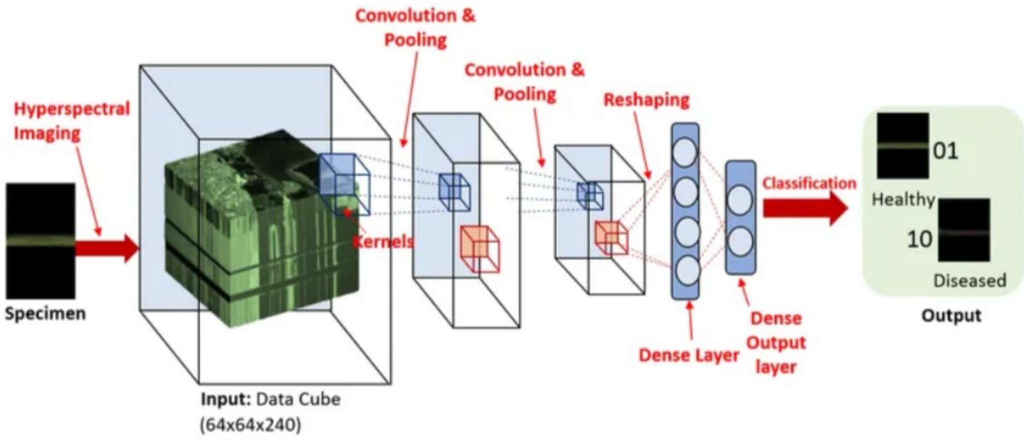

Fig.5.5 BLOCK DIAGRAM OF 3D CNN

The architecture of a 3D Convolutional Neural Network (3D CNN) can vary based on the specific application and complexity of the task at hand. However, a general architecture can be outlined with the following components:

1) 3D Convolution:

3D CNNs use 3D convolutional layers to extract features from volumetric data. These layers slide a 3D kernel (a cube) over the input volume to detect patterns in all three dimensions.

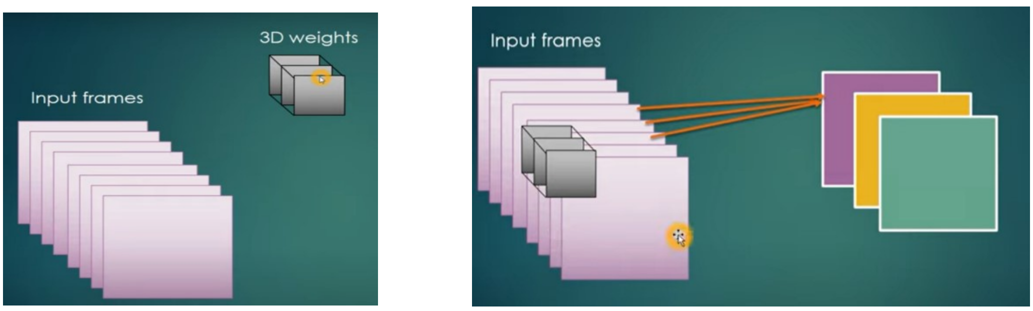

Fig.5.6 WORKING OF 3D CONVOLUTION

2) Pooling and Striding:

3D CNNs use 3D max-pooling layers and strides to downsample the spatial dimensions of the data, reducing the computational load. Max pooling takes the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling.

MAX POOLING:

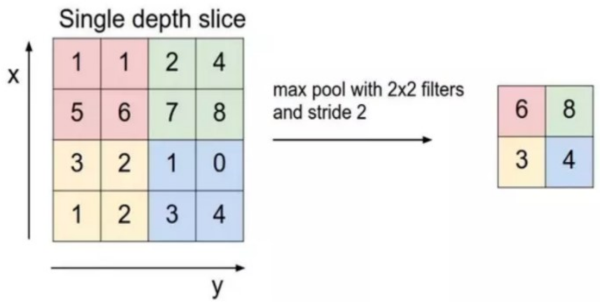

Fig.5.7 MAX POOLING LAYER

3) Skip Connections:

Skip connections, similar to those used in 2D U-Net architectures, can be applied to 3D CNNs to improve segmentation accuracy.

SUMMARY:

Machine learning and deep learning algorithms were discussed in detail. In machine learning with the flow chart the process is explained step by step. Optimal wavelength is chosen on comparing the mean spectral curves of 2 samples one with pesticide and the other without pesticide. Visible region and NIR region are chosen as optimal wavelength region and mean spectral curves are converted into text format and wavelength and corresponding reflectance values are copied in an excel sheet. In optimal wavelength region maximum and minimum reflectance values (R_{min} and R_{max}) are extracted as required features and given to 5 classifiers and results are assessed.

In deep learning, we developed a SERSN(Squeezed Excitation Residual Spectral Network) and the images were trained and tested using this algorithm and the results are obtained accordingly. Architecture of 3D CNN and it's layers are also discussed in detail in this chapter.

CHAPTER 6
RESULTS AND DISCUSSION

6.1 INTRODUCTION:

The results and their implications are covered in this chapter. If necessary, hyperspectral images are pre-processed. Then, as was previously mentioned, two algorithms—machine learning and deep learning—are put into practice. Five classifiers' accuracy and error percentage are acquired as outcomes for machine learning, and they are examined. Similarly, four classes of training accuracy and validation accuracy are produced for deep learning, and the outcomes are compared and examined.

6.2 PREPROCESSING RESULT

Hyperspectral cameras are used to take pictures of materials, and the resulting images are categorized into four classes: pure, pesticides in varying concentrations (low, medium, and high). If the deliberation is not done correctly, the images are cropped to the appropriate size. Subsequently, the four classes' hyperspectral pictures are processed as follows.

6.2.1 ROI Selection:

The Region of Interest (ROI) is initially selected in order to obtain the mean spectral curve. As seen in Fig. 6.1, the vegetable region is selected as the ROI. The corresponding mean spectral curve is then produced by right-clicking the ROI mean spectrum. Similarly, ROI is selected for each sample, and mean spectral curves are produced in accordance.

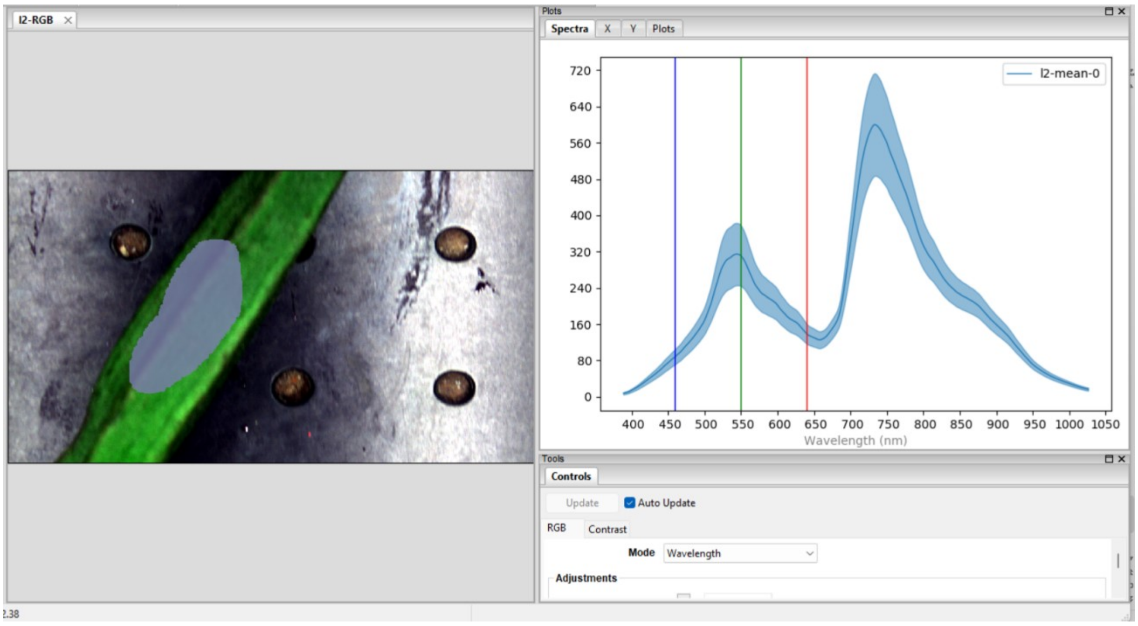

Fig 6.1.ROI SELECTION OF LADYSFINGER

6.2.2 Spectral signature

On hiding the spectral deviations, clear mean spectral curves are obtained as seen below.

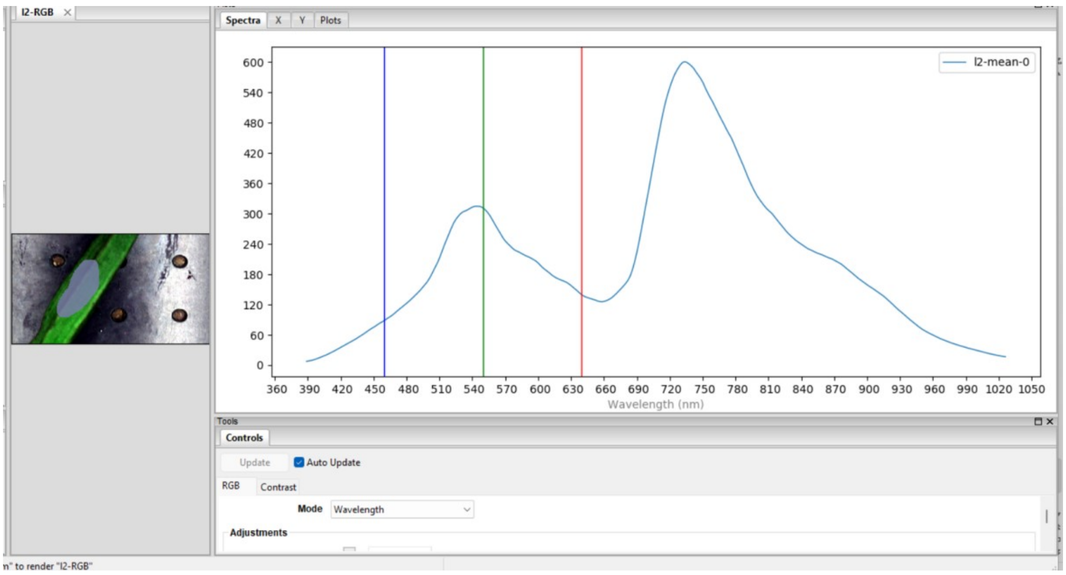

Fig.6.2 MEAN SPECTRAL CURVE OF LADYSFINGER

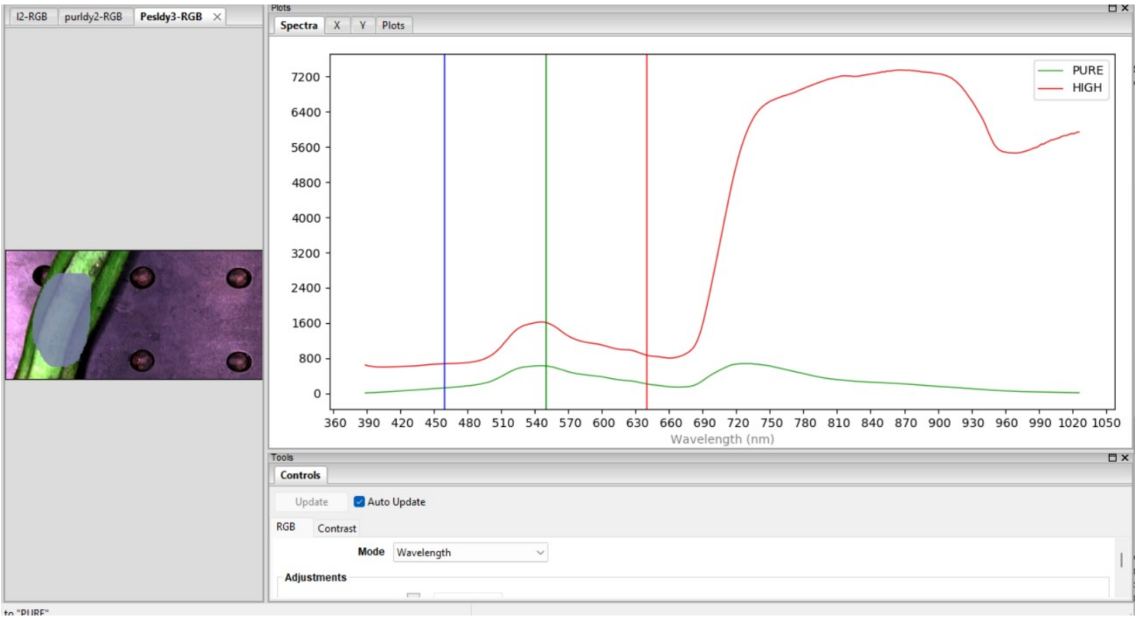

Fig.6.3 MEAN SPECTRAL CURVE OF LADYSFINGER WITH AND WITHOUT PESTICIDE

The above figure depicts the mean spectral curve for 2 samples of Lady's finger one with pesticide and the other without pesticide. Green colour curve is the mean spectral curve of Lady's finger without pesticide whereas red colour curve is the one with pesticide.There is difference of reflectance between the curves at the RGB region around 450-630nm and between 720-930nm at the NIR region. Thus the visible region and NIR region is chosen as optimal wavelength. Likewise mean spectral curves are obtained for all the samples of Lady's finger with and without pesticide.

6.3 MACHINE LEARNING RESULTS

As was previously mentioned, the mean spectral curves are transformed into text format and copied into an Excel sheet as part of the machine learning procedure. It consists of reflectance values that correlate to wavelengths. The maximum and minimum reflectance values (R_{max} and R_{min}) are taken as necessary features from the ideal wavelength region in both the upper band (NIR region) and the lower band (visible region). These values are then fed to five machine learning classifiers, and the outcomes are evaluated.

6.3.1 Feature Extraction

Table 6.1 FEATURE EXTRACTION OF LADYSFINGER FOR LOWER BAND

S.NO	Lower Band					
	Without Insecticide			With Insecticide		
	R_{min}	R_{max}	R_{avg}	R_{min}	R_{max}	R_{avg}
1)	292.46	1658.13	965.295	337.194	1895.02	1116.107
2)	326.33	1840.66	1083.495	307.742	1596.09	951.916
3)	300.45	1624.83	962.64	713.083	2565.99	1639.53
4)	333.407	1899.79	1116.5985	343.047	1819.69	1081.36
5)	487.85	2488.42	1488.135	488.433	2488.42	1488.42

Table 6.2 FEATURE EXTRACTION OF LADYSFINGER FOR HIGHER BAND

S.NO	Higher Band					
	Without Insecticide			With Insecticide		
	R_{min}	R_{max}	R_{avg}	R_{min}	R_{max}	R_{avg}
1)	5126.15	7226.85	6176.5	5278.78	7204.18	6241.48
2)	5245.28	7400.85	6323.06	5155.97	7125.79	6140.88
3)	5073.21	7182.95	6128.08	6294.98	8729.31	7512.14
4)	5262	7227.35	6244.67	3106.98	5678.97	4392.97
5)	4151.27	5184.55	4667.91	4151.27	5247.14	4699.20

Features taken from the lady's finger where the difference between using and not using pesticide was noted are included in Table 6.1. The R_{min} and R_{max} in the upper and lower bands were retrieved using these features, and the results are supplied as an input to the ML models. First, using the previously specified features, two classes—one with and one without pesticide—are carried out for Lady's finger. After that, there are four classes: pure, medium, and high concentrations of pesticide, and pure. The classification accuracy is displayed in the table below when the classifiers are fed these extracted characteristics.

6.3.2 Accuracy Assessment Of Lady's Finger

In this case the number of samples chosen for two classes with and without pesticide are 120 and 205 , with 100 being low , 55 being medium and the rest being high respectively.

Table 6.3 CLASSIFIER RESULTS FOR FOUR CLASSES OF LADYSFINGER

S.No	Classifier	Accuracy(%)	Error(%)
1	ENSEMBLE	0.76	0.24
2	KNN	0.79385	0.20615
3	DECISION TREE CLASSIFIER	0.76923	0.23077
4	SVM	0.60615	0.39385
5	NAIVE BAYES	0.75692	0.24308

Table: 6.3 shows the accuracy of several classifiers for the finger of the lady. Four characteristics (R_{min} and R_{max} for the upper and lower bands for the samples with and without pesticide) were extracted for each sample for four classes, and the results were provided. Of the 325 samples, 265 were utilized for testing, and the remaining 60 for training.Out of this, KNN and the Decision tree classifier have the highest accuracy with accuracy being 79.3% and 77% respectively.

Table 6.4 CLASSIFIER RESULTS FOR TWO CLASSES OF LADYSFINGER

S.No	Classifier	Accuracy(%)	Error(%)
1	KNN	0.83692	0.16308
2	SVM	0.48308	0.51692
3	ENSEMBLE	0.83077	0.16923
4	DECISION TREE CLASSIFIER	0.84615	0.15385
5	NAIVE BAYES	0.83077	0.16923

Table: 6.4 shows the accuracy of several classifiers for the finger of the lady. Four characteristics (R_{min} and R_{max} for the upper and lower bands for the samples with and without pesticide) were extracted for each sample for two classes, and the results were provided. Of the 325 samples, 265 were utilized for testing, and the remaining 60 for training.Out of this, Decision tree and KNN have reported the highest accuracy of 84.6% and 83.6% respectively. This is for 2 classes.

6.4 DEEP LEARNING RESULTS:

We tried to increase the dataset by augmentation Each image is rotated,shear,flipped and resized with each images being augmented into 20 images respectively. Hence ,we got about 12,320 augmented images.

6.5 AUGMENTED IMAGES OF LADYSFINGER

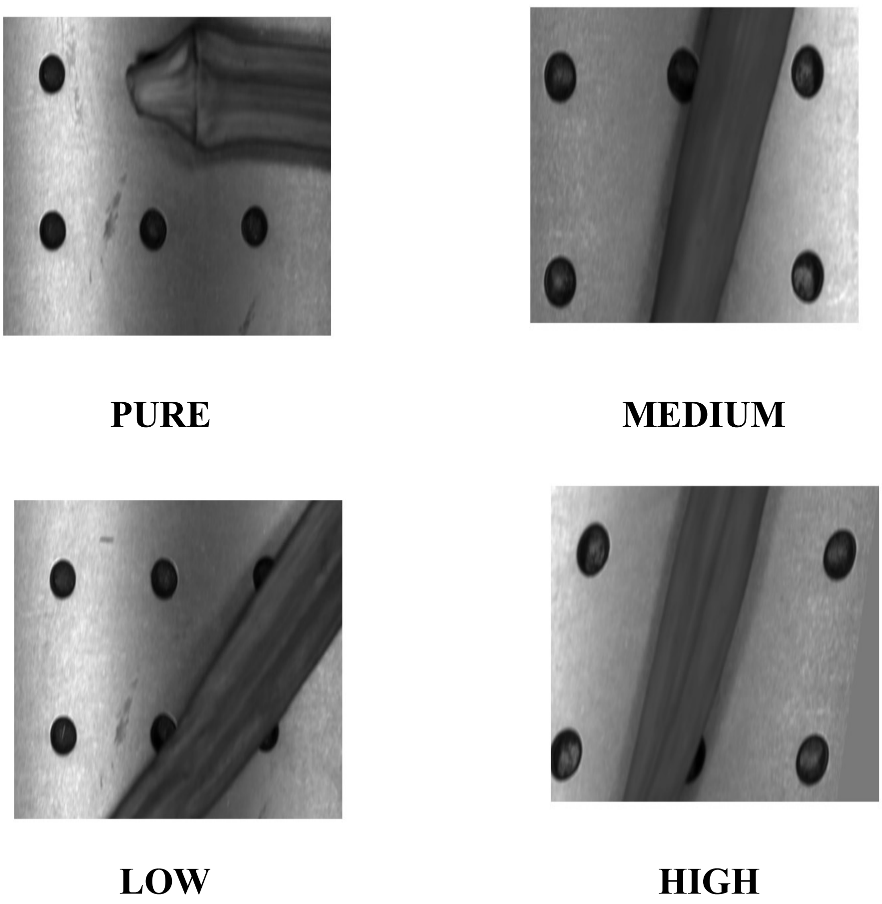

Fig 6.4 AUGEMENTED IMAGE OF PURE,MEDIUM,LOW AND HIGH CLASS LADY'S FINGER

6.6 CNN RESULTS:

The results were carried out for 5 and 10 epochs for models **with Squeezed excitation layer and without squeezed excitation layer.**The results are shown below.

TABLE 6.5 LADYSFINGER RESULT FOR 4 CLASSES WITH SQUEEZED EXCITATION LAYER

NO.OF.IMAGES				EPOCH	ACCURACY	PRECISION
PURE	**LOW**	**MEDIUM**	**HIGH**			
175	220	100	120	10	98.07%	100%
175	220	100	120	5	85.03%	100%

TABLE 6.6 LADYSFINGER RESULT FOR 4 CLASSES WITHOUT SQUEEZED EXCITATION LAYER

NO.OF.IMAGES				EPOCH	ACCURACY	PRECISION
PURE	**LOW**	**MEDIUM**	**HIGH**			
175	220	100	120	5	90.83%	100%

The above two tables show the result of the classification of 4 classes of lady's finger with and without squeezed excitation layer. The absence of the layer has resulted in the accuracy being affected heavily with the accuracy being reduced to 78% from 98% .This shows the importance of the layer.The further analysis are shown below.

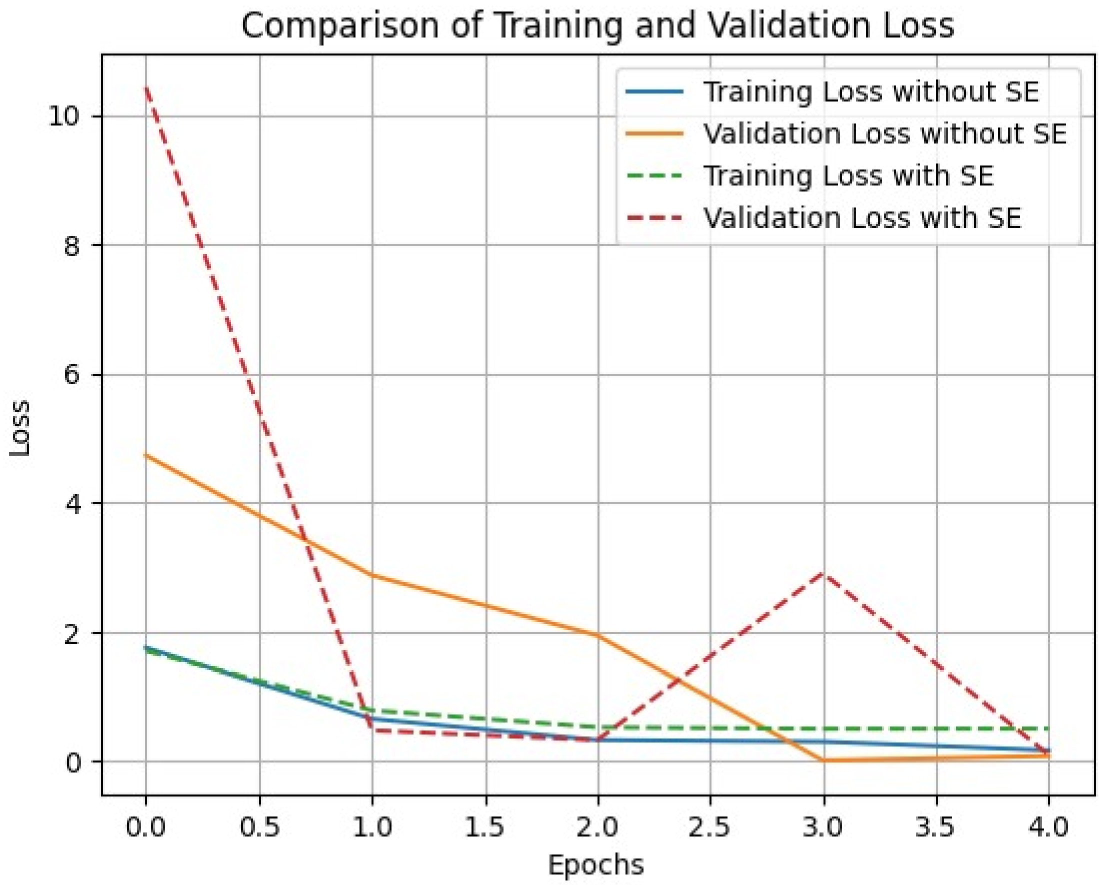

FIG 6.5 COMPARISION OF TRAINING AND VALIDATION LOSS

The above graph compares the training and validation loss with and without the squeezed excitation layer. The training loss for both the parameters remains the same .But the validation loss with SE is more at the start and gradually gets reduced to 0 after 4 epoch.The reason for intial loss is that The Squeeze-and-Excitation (SE) module might introduce additional capacity to the model, allowing it to capture more complex patterns from the training data. This increased capacity could lead to overfitting, especially at the beginning of training when the model has not yet learned to generalize well.Overall, the decreasing validation loss for the model with the SE module suggests that it is effectively learning to generalize from the training data, despite the initial overfitting. This behavior is common in deep learning models and underscores the importance of monitoring both training and validation metrics to assess model performance accurately.

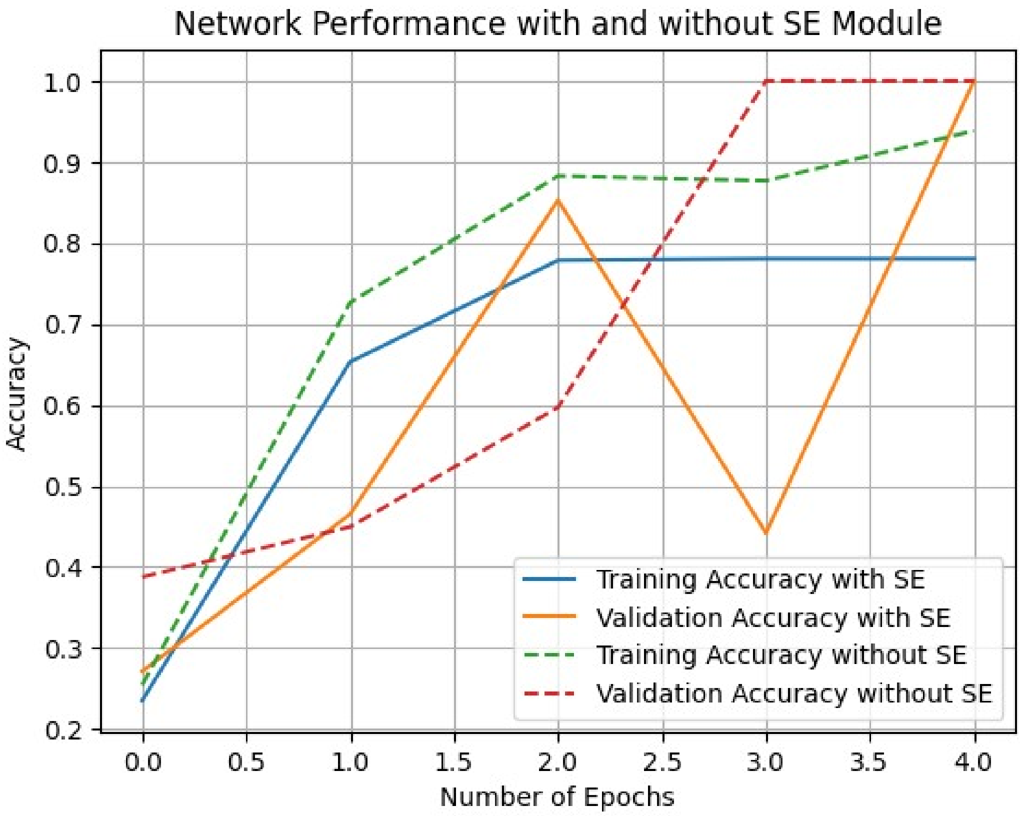

FIG 6.6 NETWORK PERFORMANCE WITH AND WITHOUT SE MODULE

The above graph explains the performance with and without SE module.We can see training accuracy without SE rises rapidly because of the less number of epoch and the training accuracy takes time to train the data set and hence the accuracy reaches only to 78% till 5 epoch and after that the growth is rapid and hence the accuracy reaches 98% after 10 epoch.The validation accuracy remain the same after 5 epoch for both with and without the squeezed excitation layer,with the validation accuracy seeing a dip slightly after 3 epoch.The learning rate used for training both models might not be optimal for the model with SE. A higher learning rate could lead to more aggressive updates to the model weights, which might cause instability and slower convergence initially. and the other hyper-parameters such as batch size, optimizer choice, and regularization techniques could also influence the training dynamics and convergence speed of the models.

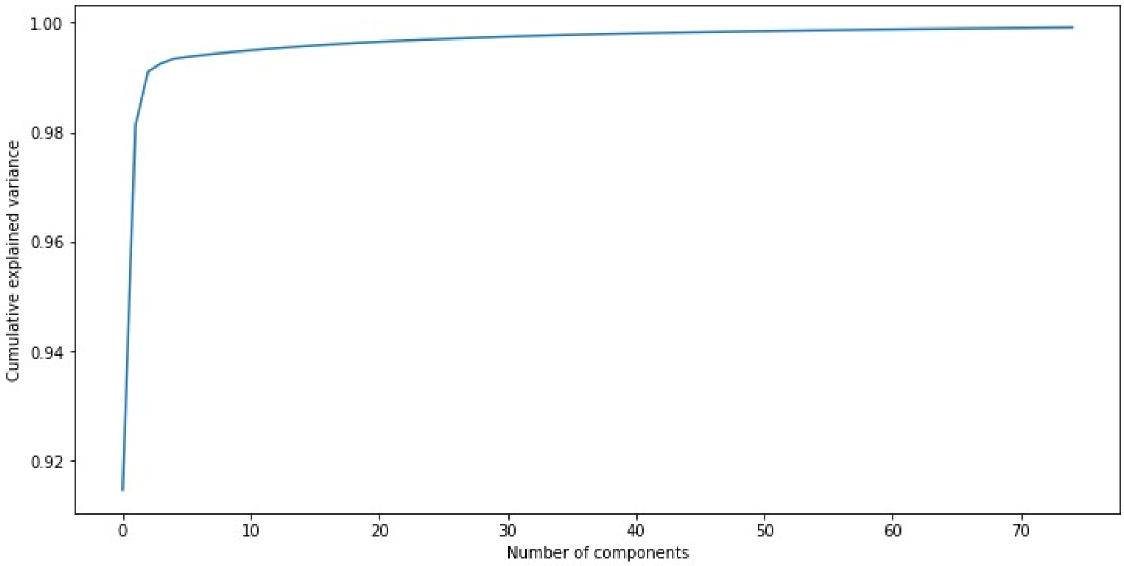

FIG 6.7 CUMULATIVE EXPLAINED VARIANCE

The above plot shows the increase in cumulative explained variance for increase in the number of components .If the cumulative explained variance increases as the number of components increases in a cumulative explained variance plot, it suggests that each additional component contributes to explaining more of the total variance in the dataset. This behavior is expected in principal component analysis (PCA) or other dimensionality reduction techniques .Each component in PCA captures a certain amount of variance in the original data. When we add more components, you're including more information from the original data. The cumulative explained variance represents the total amount of variance explained by all the components up to a certain point. As you add more components, the cumulative explained variance continues to increase because each component contributes some amount of variance.

6.7 SUMMARY

In this chapter the results and inference were discussed. Both machine learning and deep learning algorithms were recalled and it’s step by step implementation and their results were analysed. In machine learning how ROI is chosen and how will it look like is shown as well as the mean spectral curves were also shown. The extracted reflectance features were tabulated as a sample to show how it will be grouped and minimum and maximum values will be selected. In deep learning, 3D RESNET was used.

CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1 CONCLUSION

The suggested project assists in identifying which vegetables contain pesticides at varying concentrations. Both deep learning and machine learning have been created for the proposed project. The reflectance values (R_{max} and R_{min}) are employed as input features in machine learning, which yields classification accuracy. Following an averaging process, CNN (deep learning) is used to train and evaluate the augmented images, yielding results.Although there are several ways to tell if something contains pesticides or not, the suggested model uses hyperspectral images, which provide a wealth of information in a variety of wavelengths and aid in our in-depth analysis.

7.2 FUTURE WORK

Similar to vegetables, fruits can also be included in the work to determine if pesticides are used or not.The accuracy rate of the method can be increased by expanding the dataset and fine-tuning the algorithm. It is possible to estimate the various pesticide types that are applied on fruits and vegetables. It is possible to identify the disorders brought on by several pesticide applications.

7.3 OUR CONTRIBUTION

We have created a hyperspectral image dataset of lady's fingers with and without pesticides at varying concentrations. Additionally, we have built an appropriate feature extraction technique for the analysis and identification of pesticides using various machine learning techniques. and the creation of a lady's finger pesticide residue estimation system based on deep learning.

REFERENCES

[1] Zhang Chun-hua. Optimization Problems in Support Vector Machines. Beijing: China Agricultural University, 2004

[2] Deng Nai-yang,Tian Ying-jie. A new means of Data Mining ü Support Vector Machine.Beijing: Science Press. 2006

[3] Jiao Yue, Zhang Xin-feng, Zhou Li, Weighted Support Vector Machine for Classification of Tongue Images, Measurement & Control Technology, 2010,29(5).pp:1-5.

[4] Qin Zhen, Zhao Jian-yong, Yan Yi., PatternRecognition of Elevator Traffic Mode Based On Multi-value Classification Support Vector Machine, Computer Engineering. 2011, 37(9), pp: 201-206.

[5] Liu Jing, Guan Xiao, Application of SVM in the classification of meat freshness, Science and Technology of Food Industry,2011,32(4),pp:112-116.

[6] Guo-Fang Pang. Compilation of OfficialMethods of Analysis for Pesticide Residues andVeterinary Drug Residues. Beijing: Standardspress of China. 2009

[7] Belhumeur, P. N., Hespanha, J. P., and Kriejman, D. J. 1997. Eigenfaces vs. Fisherfaces: Recognition using class−specific linear projection. IEEE Trans. on PAMI 19(7): 711−721.

[8] Talukder, A., and D. Casasent. 1998. A general methodology for simultaneous representation and discriminant of multiple object classes. Optical Eng. (special issue on advances in recognition techniques) 37(3): 904−913.

[9] Dr.S. Rajendran, ENVIRONMENT AND HEALTH ASPECTS OF PESTICIDES USE IN INDIAN AGRICULTURE(2013), pp: 353 – 373.

[10] Anket Sharma, Worldwide pesticide usage and its impacts on ecosystem, Springer Volume 1,article number 1446,(2019)