



Special Issue of First International Conference on Information Technology, Computing & Applications (ICITCA 2021)

Pest Detection for Rice Using Artificial Intelligence

Madhavi.G¹, Dr. A. Jhansi Rani², Dr. S. Srinivasa Rao³

¹Assistant Professor, Dept.of ECE, MGIT, Gandipet, Hyderabad, Telangana, India

²Professor, Dept.of ECE, V.R.Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

³Associate Professor, Dept.of ECE, MGIT, Gandipet, Hyderabad, Telangana, India

gmadhavi_ece@mgit.ac.in¹, jhansi9rani@gmail.com², ssrinivasarao_ece@mgit.ac.in³

Abstract

Agriculture not only provides food for humans, but it is also a major source of revenue for any nation. Millions of dollars are spent every year to protect rice crops from insects and pests that cause damage during harvest and storage. Early pest detection, which allows the crop to be protected from pest attack, is one form of crop protection. The best way to learn about the health of a crop is to examine it regularly. If pests are discovered, adequate steps may be taken to prevent the crop from suffering a major loss of yield. Early detection will help to reduce the use of pesticides and direct the pesticide selection process. It has grown into a large field of science, with a lot of work being done around the world to detect pests automatically. The typical way of inspecting the fields is with the naked eye. A farmer must manually search and assess over a vast landscape of fields, risking overlooking different affected areas and conducting thorough research across large lots. To analyse the entire area, several human experts are needed, which is both costly and time-consuming. This proposed system is mainly intended to develop an Intelligent IT-driven system using various Artificial Intelligence and Computer Vision Algorithms for precision farming, enabling the delivery of information directly to the farmer's phone, providing the details of damage localization, crop health, and needs for fertilizer and pesticide application.

Keywords: *Object detection, Artificial Intelligence, K-means Clustering, Unsupervised Learning*

1. Introduction

1.1 Agriculture

Plant and livestock cultivation is both a science and an art. Agriculture [1] was a pivotal step in the evolution of human society. Developing farmers started planting wild grains about 11,500 years ago, after collecting them for at least 105,000 years.

1.2 Indian Agriculture

The backbone of the Indian economy [2] is Agriculture, and it will maintain for a long time. It would be able to sustain nearly 17% of the world's people by using just 2.3% of the world's land area and 4.2 % of the world's water supplies. Economic reforms implemented in the country during the

first decade of the 1990s have improved the economy's growth trajectory.

1.3 Agricultural Production and Productivity

Agricultural engineering is a discipline that aims to establish technologies that increase productivity while lowering cultivation costs [3]. Land operations and processing activities required the use of electricity. Agricultural engineering practises have increased as a result of the advent of electricity. To care for the projected population of 1.363 billion by 2025, productivity must increase by 100 % from its current level. The anticipated energy contribution to agriculture has increased from 1.3 to 2.4 KW/ha

The limitations of low agricultural production were recognised, and as a result, state and federal governments stressed the importance of agriculture's rapid growth. Farmers who use high-yielding varieties with additional fertiliser, pesticides, and assured irrigation by tube wells accelerate agricultural development. The adoption of better seeds, a higher fertiliser dosage, and plant protection chemicals, including irrigation, allowed for an increase in grain production.

1.4 Types of pests in Agriculture

We obtained images so twenty species of paddy field insect pests with ten shots per species from Google Images taken by the Department of Agricultural Biology, University of Jaffna, Sri Lanka. Fig. 1 shows some example images of these selected twenty species of insect pests primarily found in Jaffna. This image set has significant viewpoint changes, different backgrounds, arbitrary rotations, and scale differences within each class. Fig. 2 shows some of the intra-class variations in the image set and scale differences within each class

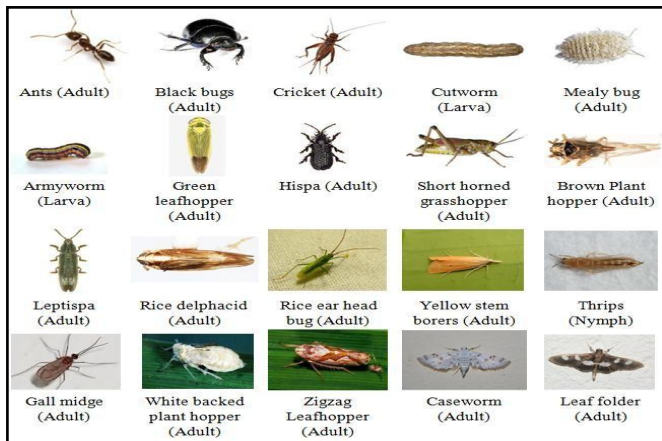


Fig.1.Rice pests set 1

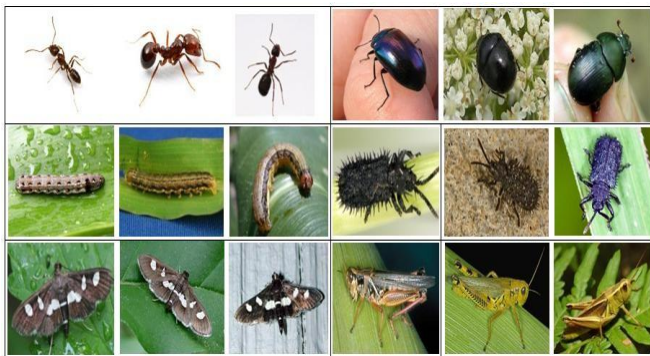


Fig.2. Rice pests set 2

2. Agriculture by using Artificial Intelligence (AI)

In agriculture, Artificial Intelligence (AI) advancements prove improvements in gaining yield and increasing and developing increasing crops. An AI algorithm [4] predicts the time takes for products like a tomato to be ready for picking, therefore increasing farming effectiveness. The advancement goes on, including Crop and Soil testing, Predictive Analytics, Agricultural Robots uses sophisticated algorithms and information collected to maintain and monitor crops' health. The growth of demand for food in the future due to the more population, they required no less than a 75% raise in production to maintain this demand from agriculture. People perceive that the alteration of these new techniques and AI will help reach that target.

The Artificial Intelligence in agriculture shown in Fig.3



Fig.3. Artificial Intelligence in Agriculture

2.1 Machine Learning

Machine learning (ML) is a subset of artificial intelligence (AI). Without being expressly programmed to do so, machine learning algorithms [5,6] made decisions based on a mathematical model supported sample data, referred to as "training data." Many programmes make use of machine learning algorithms. Computational statistics is what machine learning is said to be. The field of machine learning benefits from the study of mathematical optimization and theory, which provides methods and application domains. Data science is a related area of research that focuses on unsupervised learning for exploratory data analysis. Predictive analytics is a term used to describe machine learning algorithms.

2.2 Deep Learning

Deep Learning [6] can detect the existence of pests and disease in farms with great accuracy. CART, a machine learning algorithm, can reliably forecast the likelihood of potential illness and pest attacks. Regular human monitoring is unable to reliably predict the quantity and severity of pests and diseases attacking the plant, allowing for the application of appropriate and sufficient fertilizers, pesticides to eradicate the horde. As a result, simulated perception provides a corrective indicator of the amount of fertilisers and pesticides to spray at specific goal areas. This paper aims to assist farmers in protecting their farms from pest and illness attacks and removing them with no disrupting the soil's demureness or new plant sections. Farmers in India mostly rely on physical monitoring and a few apps that have large database restrictions and are only reliable for recognition. Since avoidance is preferable to heal, our paper aims to forecast pests/diseases in the future, allowing farmers to avoid such attacks. In the development of farms and agro-based industries, technology plays a critical role. In agriculture, automation is the most in-demand technology. Many companies have developed the most cutting-edge solutions in ML and AI, transform cultivation into Digital cultivation. The application of Deep Learning in cultivation is discussed and tested in this paper. For growers, diagnosis is normally a top priority. Farmers use pesticides/fertilizers throughout their whole farm because they are afraid of disease or pest attack, which can damage the soil and plants. This paper aims to teach farmers how to apply a small amount of pesticide/fertilizer only to a functional area where a pest/disease is present or may be present in the future. This assists farmers in preventing and eliminating such attacks on their farms by spraying in a small quantity and not polluting soil and other plant pieces. This has a number of advantages, including growing farmers' annual monetary income and reducing crop loss due to pest/disease attacks.

3. Vegetation Index

A vegetation index [7-10] is a number that indicates how healthy the vegetation is. The physics of light reflection and absorption through bands is used to calculate the math for calculating a vegetation index. Good vegetation reflects light

strongly in the near-infrared band but not so much in the visible spectrum. As a result, a ratio of light reflected in the near-infrared to light reflected in the visible spectrum can be used to identify areas of potentially safe vegetation.

3.1 Normalized Difference Vegetation Index

The Normalized Difference Vegetation Index (NDVI) is a greenness index that ranges from 0 to 1, with 0 indicating minimal or no greenness and 1 indicating full naturalness. Over wide areas, NDVI is used as a quantitative proxy measure of vegetation health, cover, and phenology (life cycle stage).

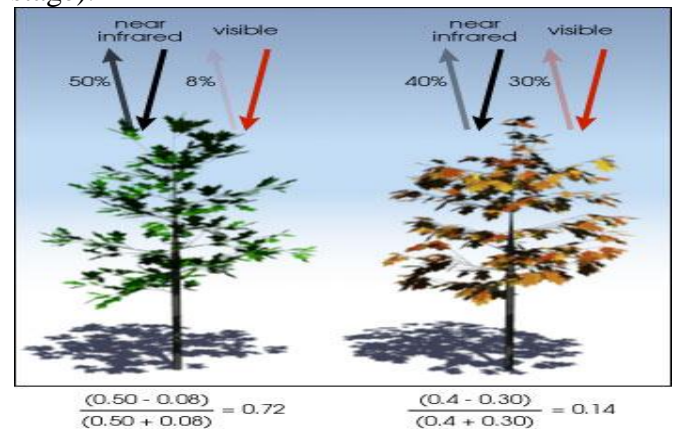


Fig.4. NDVI

The visible and near-infrared light reflected by vegetation is used to calculate the NDVI [11,12]. Healthy vegetation (left) absorbs the majority of visible light and reflects a large percentage of near-infrared light. Vegetation that is unhealthy (on the right) reflects more visible light and less near-infrared light.

To determine the density of green on a piece of ground, researchers must observe the different colours (wavelengths) of visible and near-infrared sunlight reflected by the plants. Relevant frequencies of this spectrum are absorbed, while others are reflected, as sunlight strikes the objects. Chlorophyll, the colour of plant leaves, absorbs a lot of visible light (0.4 to 0.7 m) for photosynthesis. Near-infrared light (0.7 to 1.1 m) is strongly reflected by the arrangement of cells in the plates. The more leaves the plant has, the more these wavelengths of light are affected.

This difference formula is used by almost all satellite Vegetation Indices to measure the density of plant growth on the Earth; the product of this formula is known as the Normalized Difference Vegetation Index (NDVI). Written mathematically

in Equation 1

$$NDVI = (NIR - VIS) / (NIR + VIS) \tag{1}$$

NIR- near-infrared radiation VIS- visible radiation the result of this formula is called the Normalized Difference Vegetation Index (NDVI).

NDVI for a given pixel results in a number that ranges from minus one (-1) to plus one (+1); however, no green leaves offer a value close to zero. A zero indicates no vegetation, and close to +1 (0.8- 0.9) indicates the highest possible density of green leaves.

4. Algorithm

4.1 Artificial Neural Networks Algorithm

With soil review, it detects the plant's welling (moisture content), burning sensation, disease, and pest. The dataset of plant leaves, pests, a variety of diseases, and soil images is categorised into multiple clusters using the MATLAB device, which organises multiple labels. The labels are distinguished from our enter check image and assigned to the closest corresponding label based solely on Euclidian distance. The fuzzy C-means algorithm [13,14] is used to detect pests and diseases on the farm. Convolutional Neural Networks (CNNs) are used to perform ideal analysis. Since the photograph used to enter is unknown, the classification method is Unsupervised Learning.

4.2 Confusion Matrix

To understand the merits of a classification algorithm, a confusion matrix is used. When there is an unequal variety of findings in each classification or more than two training in the dataset, the accuracy of classification on my own can be deceiving. A confusion matrix's desired form is shown in Fig.5

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Fig.5. Confusion Matrix

In the figure.5, TP indicates True Positive, FP indicates False Positive, FN indicates False Negative, and TN indicates True Negative.

4.2.1 True Positive

An output where the model correctly predicts the positive class.

4.2.2 True Negative

An output where the model correctly predicts the negative class

4.2.3 False Positive

An output where the model incorrectly predicts the positive class.

4.2.4 False Negative

An outcome where the model incorrectly predicts the negative class. Besides, some terms define the model, such as sensitivity, accuracy, precision, specificity.

4.3 Accuracy

The ratio of exact predictions to the total predictions made is known as accuracy.

$$Accuracy = (TP+TN) / (TP+TN+FP+FN) \tag{2}$$

4.4 Precision

It is sometimes also known as positive predictive value. It is defined as the ratio of correct optimistic predictions to the total predicted positives.

$$Precision = TP / (TP+FP) \tag{3}$$

4.5 Sensitivity

The model's sensitivity is a measure of how well it can detect events in the positive category. It's calculated by dividing the number of true positives by the total number of positive events in the dataset, i.e. the positive class events that were correctly predicted. Positive class events were expected incorrectly.

$$Sensitivity = TP / (TP+FN) \tag{4}$$

4.6 Specificity

It assesses how precise the positive class assignment is. It's calculated by dividing the number of true negatives by the total number of adverse events in the dataset, with the negative class events correctly predicted and the adverse class events incorrectly predicted.

$$Specificity = TN / (FP+TN) \tag{5}$$

4.7 Image Processing Algorithm

There are several advantages to convert snap shots to grayscale and understanding how to do so. The most obvious advantage of convert a shaded image to grayscale is that takes up considerably fewer spaces. If you're unfamiliar with how pictures are saved on a computer, this is a good place to start. An RGB image [15] is made up of three images stacked on top of each other: a red scale, a green scale, and a blue scale image, each with eight bits

per pixel (intensity value ranging 0 - 255). A single pixel of an RGB image needs 8 bits for each of the three colours, for a total of 24 bits per pixel.

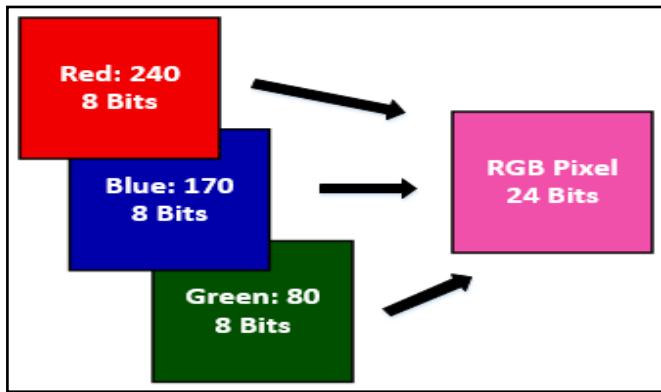


Fig.6. RGB to Greyscale conversion

5. Results and Discussion

The training phase has done on the image dataset. We also used K-Means clustering to classify the given input dataset into its corresponding labels and trained an algorithm to do so. We used a built-in function in MATLAB called Bag of Features to categorise the images. The moisture content of Edge Detection is also completed, as is the number of days the crop can withstand without external attacks during the post-harvest season. This section is beneficial to agro-based industries because, according to a survey, 20 percent of crop post-harvest is lost due to moisture levels that are not kept at optimal levels. In comparison to a perfect moisture leaf, Figure.7 depicts the moisture content in the given input test picture. Via an edge detection algorithm, it has established a math relation for determining moisture levels. We estimated the approximate moisture levels from the input levels, as well as the number of days the crop will survive without external infections, using trial and error math. If the moisture is more than usual, the leaf will swell; if humidity is less, it will drain. The leaf will swell if there is more moisture than normal, and it will drain if there is less. Based on the current humidity content, this parameter is used to predict disease attacks. There's a risk of fungal attacks if there's a lot of moisture. As a result, this is one of the most important factors to consider when predicting foreign agent attacks. Figure .8.shows how to extract features using the MATLAB built-in function collection -Of-Features. K-Means Clustering is used by the algorithm to classify the images. The number of

features was 53055, and the number of clusters was 500, thanks to the algorithm. Following the development of training fundamentals, the clusters are ready for training and labelling for the detection process.

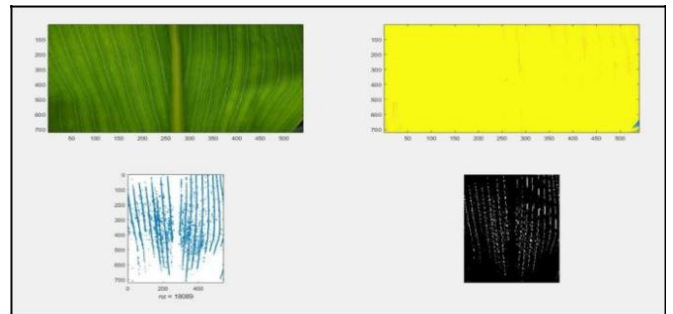


Fig.7. Moisture level extraction with sparsity levels

```

Command Window
New to MATLAB? See resources for Getting Started.

Creating Bag-Of-Features.
-----
* Image category 1: Alternaria Alternata
* Image category 2: Anthracnose
* Image category 3: Bacterial Blight
* Image category 4: Cercospora Leaf Spot
* Image category 5: Healthy Leaves
* Selecting feature point locations using the Grid method.
* Extracting SURF features from the selected feature point locations.
** The GridStep is [8 8] and the BlockWidth is [32 64 96 128].

* Extracting features from 25 images...done. Extracted 73472 features.

* Keeping 80 percent of the strongest features from each category.

* Balancing the number of features across all image categories to improve clustering.
** Image category 4 has the least number of strongest features: 10611.
*** Using the strongest 10611 features from each of the other image categories.

* Using K-Means clustering to create a 500 word visual vocabulary.
* Number of features      : 53055
* Number of clusters (K)  : 500

* Initializing cluster centers...100.00%.
* Clustering...completed 29/100 iterations (~0.42 seconds/iteration)...converged in 29 iterations.

* Finished creating Bag-Of-Features
  
```

Fig.8. Bag of Features

The category names have been given as the disease names of the plants. The images get into those corresponding labels for identification purposes after the clustering. The SVM's X-Y axes are formed by this. We developed a Perception in which image extraction is performed in hidden layers. The number of iterations is measured in Epochs. We've been through 500 iterations. Based on the size of the image dataset, the algorithm will automatically stop when it reaches the most accurate value.

```

Training an image category classifier for 5 categories.
-----
* Category 1: Alternaria Alternata
* Category 2: Anthracnose
* Category 3: Bacterial Blight
* Category 4: Cercospora Leaf Spot
* Category 5: Healthy Leaves

* Encoding features for 25 images...done.

* Finished training the category classifier. Use evaluate to test the classifier on a test set.

Predicting categories using a multi-class classifier. Categories are:
-----
* Category 1: Alternaria Alternata
* Category 2: Anthracnose
* Category 3: Bacterial Blight
* Category 4: Cercospora Leaf Spot
* Category 5: Healthy Leaves

* Processing 25 images...done.
* Finished processing all image sets.
  
```

Fig.9. Future Training Data

5.1 Artificial Intelligence Algorithm Output

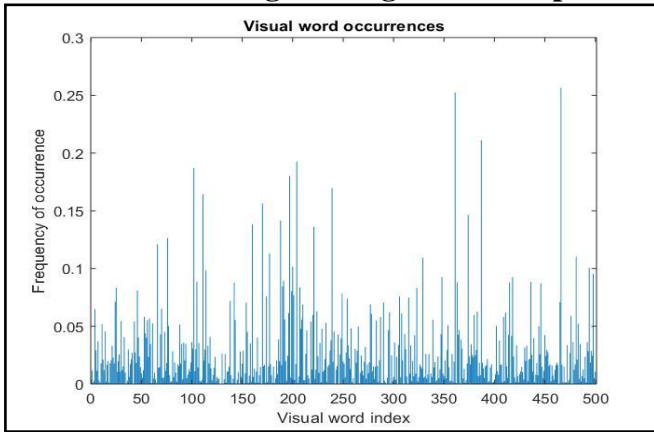


Fig.10. Output 1

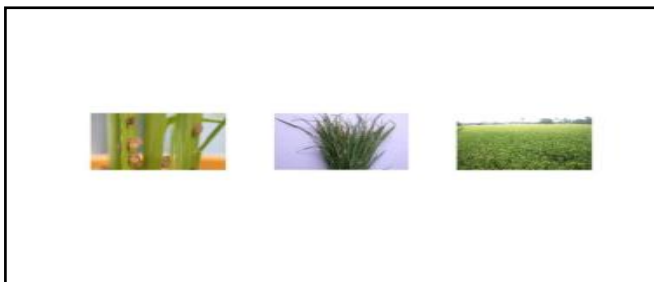


Fig.11. Output 2

5.2 Image Processing Algorithm Output

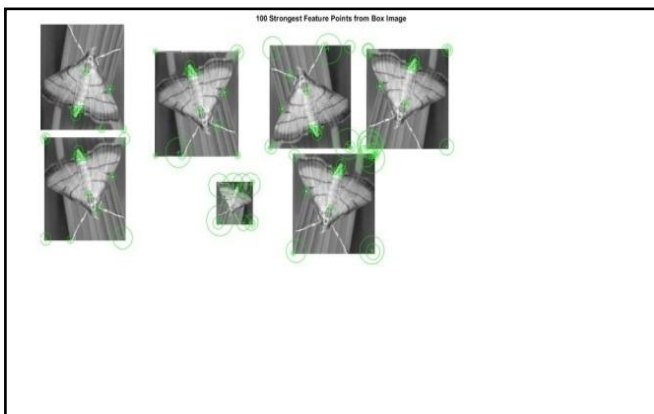


Fig.12. Output 1



Fig.13. Output 2

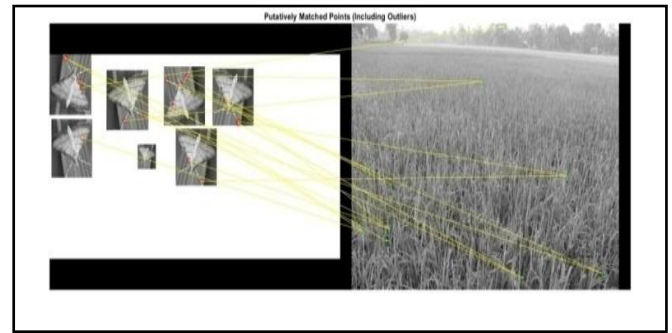


Fig.14. Output 3

Conclusions

The algorithm is implemented in MATLAB after Image categorization, Feature Extraction, and Training Data have been completed. Statistics and Machine Learning Toolbox, Neural Network Toolbox, and Image Processing Toolbox are among the toolboxes used. The coaching statistics in the form of photograph categories, picture classification using K-Means clustering, and moisture quality, as well as predicting withstanding, are the outputs. The algorithm is used to classify the photograph dataset and apply coaching information. The visual representation contrasts with the informed data for identification and prediction. For increased precision, unsupervised learning is used. Take ordered statistics from Indian rice flowers and check the box to enter as an African rice plant. Because of the minor difference in appearance, the accuracy will be compromised. As a result, we're concentrating on Unsupervised Learning. Using the Fuzzy C-Means algorithm, the instance will achieve 99% accuracy. The records may be hazy, but they will be becoming more precise. As a result, we are refraining from using Supervised Learning methods.

Future Scope

Furthermore, we plan to transform this project from a concept to a complete end-user product. This can be accomplished by using the Tensor Flow library function in the Python IDE with a large number of processors. The product will be capable of both predicting and identifying disease/pest attacks. For the education network, a more comprehensive collection of data will be provided. The entire algorithm will be created using Tensor Flow to create a higher-level framework open CV for Image analytics, similar to the Image Processing Toolbox in MATLAB.

References**Journals**

- [1].Johnny L. Miranda, Bobby D. Gerardo, and Bartolome T. Tanguilig III, "Pest Detection and Extraction Using Image Processing Techniques," *International Journal of Computer and Communication Engineering*, Vol. 3, No. 3, pp. 189-192, May 2014
- [2].V Suresh, D Gopinath, M Hemavarthini, K Jayanthan, Mohana Krishnan, "Plant Disease Detection using Image Processing," *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181, Vol. 9 Issue 03, pp. 78-82, March-2020
- [3].A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "A novel feature extraction for robust EMG pattern recognition," *Journal of Computing*, vol. 1, pp. 71-80, Dec. 2009.
- [4].R. K. Samanta and I. Ghosh, "Tea insect pests classification base on artificial neural networks," *International Journal of Computer Engineering Science*, vol. 2, June 2012.
- [5].L. Zhu and Z. Zhang, "Insect recognition based on integrated region matching and dual tree complex wavelet transform," *Journal of Zhejiang University-SCIENCE C (Computers & Electronics)*, 2011.
- [6].M. T. Do et al., "A test of a pattern recognition system for identification of spiders," *Bulletin of Entomological Research*, 1999.
- [7].Savita N. Ghaiwat, ParulArora, "Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review", *International Journal of Recent Advances in Engineering and Technology*, ISSN: 2347-2812, Volume 2, Issue 3, 2014.
- [8]. Prof. Sanjay B. Dhaygude and Mr. Nitin P. Kumbhar, "Agricultural plant Leaf Disease Detection Using Image Processing" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, Volume 2, Issue , 2013.
- [9].Mr. Pramod and S. landge , "Automatic Detection and Classification of Plant Disease through Image Processing", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 3, Issue 7, ISSN: 2277 128X, 2013.
- [10].Anand H. Kulkarni and Ashwin Patil R. K, "Applying image processing technique to

detect plant diseases", *International Journal of Modern Engineering Research (IJMER)*, Volume 2, Issue.5, pp-3661-3664 ISSN: 2249-6645, 2012

Conference Proceedings

- [11].G. Pavithra, K. S. Palanisamy, "Computer based Diagnosis of Leaf Rust Disease using Multiclass SVM," *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181, Vol.5, Issue.17, pp.1-7 RTICCT - 2017 Conference Proceedings
- [12].Haiguang Wang, Guanlin Li, Zhanhong Ma, Xiaolong Li, "Image Recognition of Plant Diseases Based on Back propagation Networks", 5th International Congress on Image and Signal Processing (CISP), 2012

Chapter in a Book

- [13].Resource Manual on Integrated Production and Pest Management (IPPM) in Rice, World Education (INGO) Philippines, Inc., ch. 6, pp. 113-115, 2005
- [14].F. A. Carino, P. E. Kenmore, and V. A. Dyck, "A FARMCOP suction sampler for hoppers and predators in flooded rice fields," *The International Rice Research Newsletter*, vol. 4, ch. 5, pp. 21-22, 1979.
- [15].Sampling insect pests, their damage and beneficial organisms Philippine rice self-sufficiency plan (PRSSP). [Online]. Available:http://pinoyrkb.com/main/resources/publications-and-figures/rice-science-for-decisionmakers/doc_download/517-sampling-insect-pests