CHAPTER 2
An Analysis of Convolutional Neural Network (CNN) and Transfer-Learning Approach

Mr M. Mohammed Kasim
Nehru Institute of Engineering and Technology, India

Mrs Saranya. R
Nehru Institute of Engineering and Technology, India

Mr S. Mohan
Nehru Institute of Engineering and Technology, India

Mrs. Jebapaulin. M
Nehru Institute of Engineering and Technology, India

ABSTRACT

Agricultural production is extremely important to the economy. This is one of the reasons why disease detection in plants is vital in the agricultural area, as illness in plants is fairly natural. If sufficient care is not taken in this region, it has major consequences for plants, affecting product quality, quantity, or productivity. Plant disease detection using any automated technology is advantageous since it eliminates a considerable amount of monitoring work in large crop farms and detects disease signs at an early stage, i.e. when they occur on plant leaves. This study provides an image segmentation system for automated identification and categorization of plant leaf diseases. It also includes an overview of several disease categorization systems that may be used to detect plant leaf disease. Convolutional neural networks are used for image segmentation, which is vital for disease detection in plant leaf disease.

Keywords: *Convolutional neural networks, Deep Learning, Transfer Learning*

INTRODUCTION

Plant disease can directly lead to stunted growth causing bad effects on yields An economic loss of up to $20 billion per year is estimated all over the world . Diverse conditions are the most difficult challenge for researchers due to the geographic differences that may hinder the accurate identification. In addition, traditional methods mainly rely on specialists, experience, and manuals, but the majority of them are expensive, time- consuming, and labour-intensive with difficulty detecting precisely. Therefore, a rapid and accurate approach to identify plant diseases seems so urgent for the benefit of business and ecology to agriculture. Internet technologies, in particular the availability of multimodality data from various sensors including the Internet of things and sensor networks, have developed rapidly Agricultural productivity is something on which economy highly depends. This is the one of the reasons that disease detection in plants plays an important role in agriculture field, as having disease in plants are quite natural. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity is affected. For instance a disease named little leaf disease is a hazardous disease

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

found in pine trees in United States. Detection of plant disease through some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e. when they appear on plant leaves. This project presents an algorithm for image segmentation technique which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases classification techniques that can be used for plant leaf disease detection. Image segmentation, which is an important aspect for disease detection in plant leaf disease, is done by using CNN algorithm.

RELATED WORK

Compared with the definite classification, detection and segmentation tasks in computer vision , the requirements of plant diseases and pests detection is very general. In fact, its requirements can be divided into three different levels: what, where and how . In the first stage, “what” corresponds to the classification task in computer vision. As shown in Fig. 1, the label of the category to which it belongs is given. The task in this stage can be called classification and only gives the category information of the image. In the second stage, “where” corresponds to the location task in computer vision, and the positioning of this stage is the rigorous sense of detection. This stage not only acquires what types of diseases and pests exist in the image, but also gives their specific locations. As shown in Fig. 1, the plaque area of gray mold is marked with a rectangular box. In the third stage, “how” corresponds to the segmentation task in computer vision. As shown in Fig. 1, the lesions of gray mold are separated from the background pixel by pixel, and a series of information such as the length, area, location of the lesions of gray bold can be further obtained, which can assist the higher-level severity level evaluation of plant diseases and pests. Classification describes the image globally through feature expression,

EXISTING SYSTEM

The Existing System for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Agricultural productivity is something on which economy highly depends. This is the one of the reasons that disease detection in plants plays an important role in agriculture field, as having disease in plants are quite natural. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity is affected. Detection of plant disease through some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e. when they appear on plant leaves.

PROPOSED SYSTEM

(1) Image acquisition is the very first step that requires data set

(2) Preprocessing of input image to improve the quality of image and to remove the undesired distortion from the image. Clipping of the leaf image is performed to get the interested image region and then image smoothing is done using the smoothing filter. To increase the contrast Image enhancement is also done.

(3) Mostly green colored pixels, in this step, are masked. In this, we computed a threshold value that is used for these pixels. Then in the following way mostly green pixels are masked: if pixel intensity of the

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

green component is less than the pre-computed threshold value, then zero value is assigned to the red, green and blue components of the this pixel.

(4) In the infected clusters, inside the boundaries, remove the masked cells.

(5) Obtain the useful segments to classify the leaf diseases. Segment the components using CNN algorithm

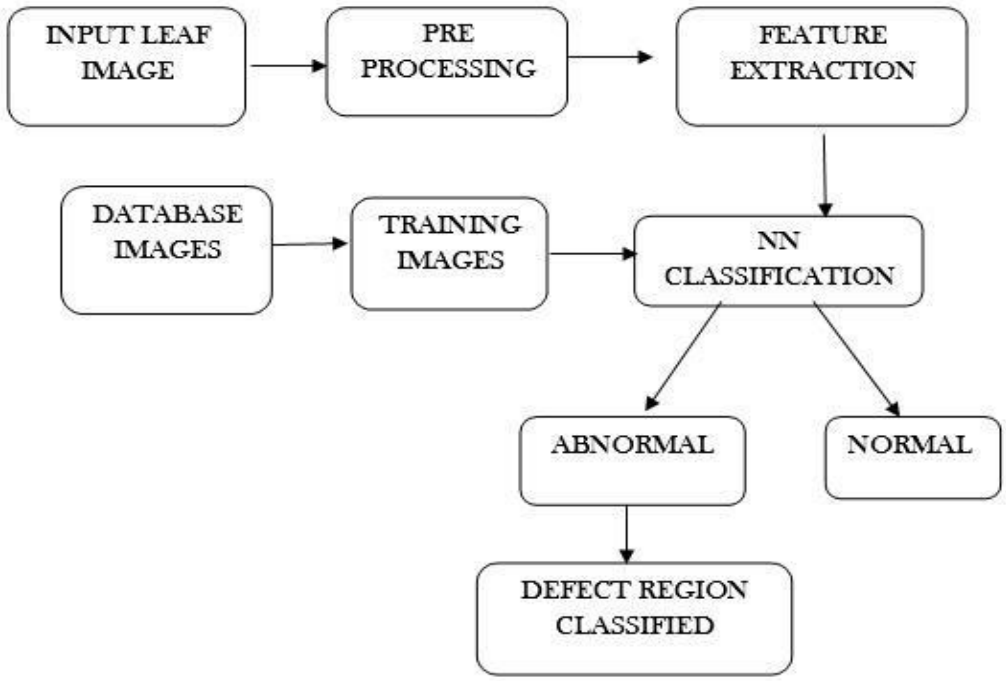


Figure 3.1 Proposed System

DATA SET DESCRIPTION

For the purpose of image-based identification which includes, training phase to evaluation phase where the performance of classification algorithms are evaluated, it is necessary to have huge data sets. Hence, the source of data is collected from Plant Village website. The images thus collated are labeled with four different categories-bacterial spot, yellow leaf curl virus , late blight and healthy(in order to differentiate healthy leaves from affected ones). Subsequently, there is a need to enhance the dataset by adding the images that are augmented. This paper further train the network to learn features that differentiates one class from other. Correspondingly, a database comprising of more than 5656 images are used to train and around 1889 images are further used to validate the same. Dataset consists of a total of 9430 labelled images. The 9430 labelled images are split into a training set (5656), a test set(1885) and a validation set (1889). The number of images per class are unbalanced with the two disease classes CMD and CBSD having 72% of the images.',

IMAGE ACQUISITION (EXAMINE AND UNDERSTAND DATASET)

“A dataset (or data set) is a collection of data, usually presented in tabular form. Each column represents a particular variable. Each row corresponds to a given member of the dataset in question. It lists values for

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

each of the variables, such as height and weight of an object. Each value is known as a datum. The dataset may comprise data for one or more members, corresponding to the number of rows. If your data set is messy, building models will not help you to solve your problem. In order to build a powerful machine learning system, we need to explore and understand our data set before we define a predictive task and solve it.

IMAGE PRE-PROCESSING

Image pre-processing or data cleansing is a crucial step and most of the Deep Learning engineers spend a good amount of time in data pre-processing before building the model. Some examples for data pre-processing includes outlier detection, missing value treatments and remove the unwanted or noisy data.

Similarly, Image pre-processing is the term for operations on images at the lowest level of abstraction. These operations do not increase image information content but they decrease it if entropy is an information measure. The aim of pre- processing is an improvement of the image data that suppresses undesired distortions or enhances some image features relevant for further processing and analysis task.

PIXEL BRIGHTNESS TRANSFORMATIONS(PBT)

Brightness transformations modify pixel brightness and the transformation depends on the properties of a pixel itself. In PBT, output pixel's value depends only on the corresponding input pixel value. Examples of such operators include brightness and contrast adjustments as well as colour correction and transformations. Contrast enhancement is an important area in image processing for both human and computer vision. It is widely used for medical image processing and as a pre-processing step in speech recognition, texture synthesis, and many other image/video processing applications

There are two types of Brightness transformations and they are below.

1. Brightness corrections

2. Gray scale transformation

The most common Pixel brightness transforms operations are

1. Gamma correction or Power Law Transform

2. Sigmoid stretching

3. Histogram equalization

Two commonly used point processes are multiplication and addition with a constant.

$g(x)=\alpha f(x)+\beta$

The parameters $\alpha>0$ and β are called the gain and bias parameters and sometimes these parameters are said to control contrast and brightness respectively.

cv.convertScale Abs(image, alpha=alpha, beta=beta)

for different values of alpha and beta, the image brightness and contrast varies.

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

The goal of using filters is to modify or enhance image properties and/or to extract valuable information from the pictures such as edges, corners, and blobs. A filter is defined by a kernel, which is a small array applied to each pixel and its neighbors within an image Some of the basic filtering techniques are

1. Low Pass Filtering (Smoothing): A low pass filter is the basis for most smoothing methods. An image is smoothed by decreasing the disparity between pixel values by averaging nearby pixels

2. High pass filters (Edge Detection, Sharpening): High-pass filter can be used to make an image appear sharper. These filters emphasize fine details in the image

– the opposite of the low-pass filter. High- pass filtering works in the same way as low-pass filtering; it just uses a different convolution kernel.

3. Directional Filtering: Directional filter is an edge detector that can be used to compute the first derivatives of an image. The first derivatives (or slopes) are most evident when a large change occurs between adjacent pixel values. Directional filters can be designed for any direction within a given space

4. Laplacian Filtering: Laplacian filter is an edge detector used to compute the second derivatives of an image, measuring the rate at which the first derivatives change. This determines if a change in adjacent pixel values is from an edge or continuous progression. Laplacian filter kernels usually contain negative values in a cross pattern, centered within the array. The corners are either zero or positive values. The center value can be either negative or positive.

IMAGE SEGMENTATION

Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image. Image segmentation could involve separating foreground from background, or clustering regions of pixels based on similarities in colour or shape.

FEATURE EXTRACTION

OVERALL ARCHITECTURE DIAGRAM

The most important advancement of deep learning over traditional machine learning is that its performance improves with the increasing of the amount of data. A 4 convolutional layers' model (4-Conv CNN) was built from scratch for this project. The structure of CNN is shown in Figure 1. This 4-Conv CNN has 4 Conv2D layers, 2 MaxPooling layers, and 5 Dropout layers, and a fully connected layer is the following.

RESULTS AND DISCUSSION

- The input test image is acquired and pre-processed in the next stage and then it is converted into array form for comparison.

- The selected database is properly segregated and pre-processed and then renamed into proper folders.

- The model is properly trained using CNN and then classification takes place.

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

- The comparison of the test image and the trained model take place followed by the display of the result.

- If there is a defect or disease in the plant the software displays the disease along with the remedy.

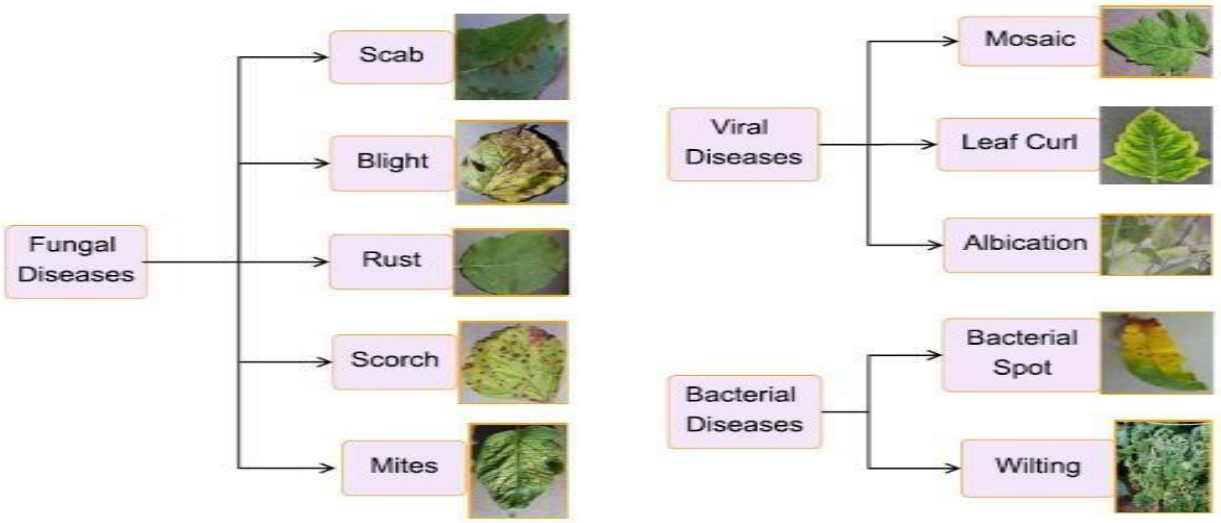
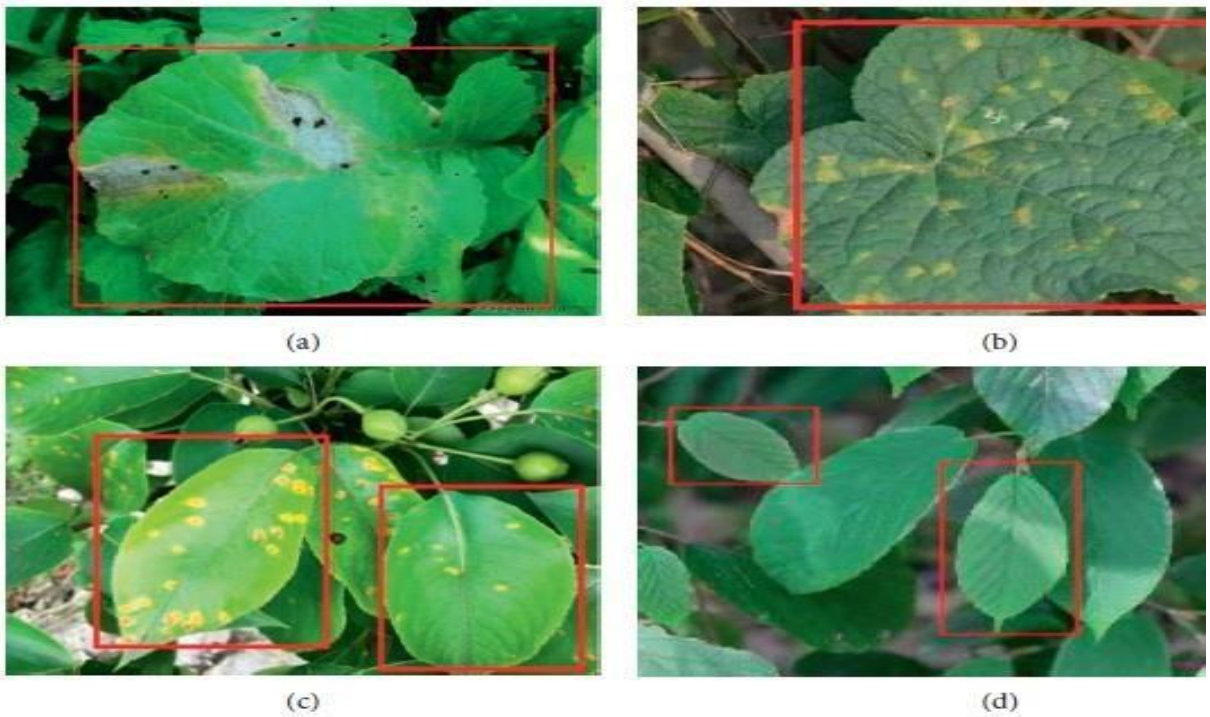


Figure 6.1 Various Biotic Infections in plants with their categories in various crops



Copyright © 2023, AETS Global.. Pages: 16- 24, September 2023
ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

Figure 6.2 The Result of leaf Identification (a) Black rot disease; (b) Bacteria plaque disease (c) rust disease: (d) Healthy leaf

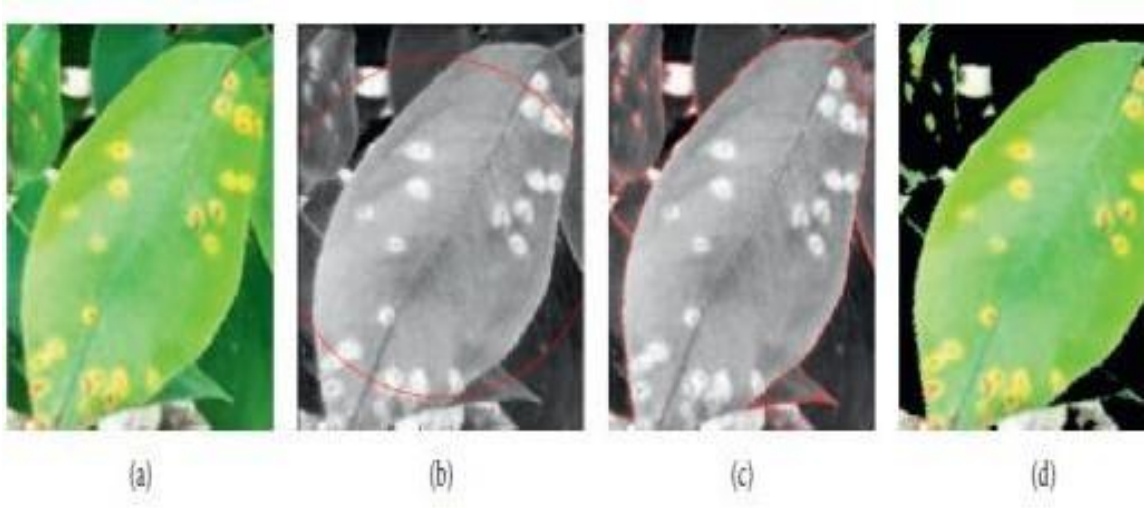


Figure 6.3 (a) Image capture; (b) Initial zero level set (c) contour image after 500 iterations (d) segmentation results

Figure 6.4 Sample leaf training data sets

Initial step for any image processing based project is acquiring proper database which is valid. Most of the time the standard database is preferred but in certain circumstances we do not get proper database .So in such conditions we can collect the images and can form our own database. The database is accessed from crowdAI which is plant disease classification challenge. Data available here is not labelled .So the first task is to clean and label the database. There is a huge database so basically the images with better resolution and angle are selected. After selection of images we should have deep knowledge about the different leaves and the disease they have.

CONCLUSION

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

This project shows that the plant disease recognition model based on deep learning has the characteristics of unsupervised, high accuracy, good universality, and high training efficiency. However, there are many challenges in accuracy practicability of plant disease detection in the complex environment. In order to solve these problems and optimize the identification method, this project proposes a recognition model integrating CNN algorithm, which can effectively solve the problem of plant disease identification in the complex environment. The model not only adapts to complex environments, but also increases the accuracy of identification. Compared with the traditional model, the model proposed in this paper not only guarantees the robustness of the convolutional neural network, but also reduces the number and quality requirements of the convolutional neural network on the data set and obtains better results. The proposed system is based on python and gives an accuracy of around 88%.The accuracy and the speed can be increased by use of Google's GPU for processing better results.

REFERENCES

[1] Redha Ali, Russell Hardie. "A Leaf Recognition Approach to Plant Classification Using Machine Learning." NAECON 2018 - IEEE National Aerospace and Electronics Conference,2018.

[2] S¸ekeroglu, Boran, and Y ˘ucel ¨Inan. "Leaves recognition system using a neural network." Procedia Computer Science 102 (2016): 578-582.

[3] Narayanan, Barath Narayanan, Ouboti Djaneye-Boundjou, and Temesguen M. Kebede. "Performance analysis of machine learning and pattern recognition algorithms for malware classification." In Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS),2016 IEEE National, pp. 338-342. IEEE, 2016.

[4] Khairnar, Khushal, and Rahul Dagade. "Disease Detection and Diagnosis on Plant using Image Processing." International Journal of Computer Applications 108, no. 13 (2014): 36-38.

[5] Wang, Jinjun, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas Huang, and Yihong Gong. "Locality-constrained linear coding for image classification."In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 3360-3367. IEEE, 2010.

[6] Bosch, Anna, Xavier Mu˜noz, and Robert Mart´ı. "Which is the best way to organize/classify images by content?." Image and vision computing 25, no. 6 (2007): 778-791.

[7] Neto, Joao Camargo, George E. Meyer, David D. Jones, and Ashok ˜K. Samal. "Plant species identification using Elliptic Fourier leaf shape analysis." Computers and electronics in agriculture 50, no. 2 (2006):121-134.

[8] Lazebnik, Svetlana, Cordelia Schmid, and Jean Ponce. "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories." In null, pp. 2169-2178. IEEE, 2006.

[9] Mikolajczyk, Krystian, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and Luc Van Gool. "A comparison of affine region detectors." International journal of computer vision 65, no. 1-2 (2005): 43-72.

[10] Harris, Chris, and Mike Stephens. "A combined corner and edge detector." In Alvey vision conference, vol. 15, no. 50, pp. 10-5244. 2000.

[11] Y. Ampatzidis, L. De Bellis, and A. Luvisi, "iPathology: robotic applications and management of plants and plant diseases," Sustainability, vol. 9, no. 6, p. 1010, 2017.

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing

[12] A. Breukers, D. L. Kettenis, M. Mourits, W. V. D. Werf, and A. O. Lansink, "Individual-based models in the analysis of disease transmission in plant production chains: an application to potato brown rot," Academy of Sciences, vol. 90, no. 1–3, pp. 112–131, 2006.

[13] S. Ghosal, D. Blystone, A. K. Singh,B. Ganapathysubramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," Proceedings of the National Academy of Sciences, vol. 115, no. 18, pp. 4613–4618, 2018.

[14] E.-C. Oerke, "Crop losses to pests," ,e Journal of AgriculturaL Science, vol. 144, no. 1, pp. 31–43, 2006.

[15] X. E. Pantazi, D. Moshou, and A. A. Tamouridou, "Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers," Computers andElectronics in Agriculture, vol. 156, pp. 96–104, 2019.

ISBN: 978-93-5786-261-5, www.aetsjournal.com ―Handbook of research in Embedded, Real time Systems and High Performance Computing