Paddy Crop Disease Detection using Deep Learning Technique by using Digital Image Processing

G.Elakkiya[1], L.Isai[2], S.Kirthiga[3], Ms V.Sathiya[4]

[1,2,3]Dept. of Cse, Panimalar Engineering College, Chennai
[4]Professor, Dept. of Cse, Panimalar Engineering College, Chennai

---***---

Abstract - *In this project presents an automatic approach for early disease and nutrition deficiency detection in Paddy Crop. Agriculture is not only a major part in providing food for human existence, it also plays a major role in providing economy of any country. Insects , nutrition deficiency, plant disease and pests damage the crops and, thus, are very dangerous for the overall growth of the crop. One method to protect the crop is early disease detection and nutrition deficiency so that the crop can be protected. The best way to know about the health of the crop is the timely examination of the crop. If disease or nutrition deficiency are detected, appropriate measures can be taken to protect the crop from a big production loss at the end. Earlier detection of the disease is helpful for minimizing the usage of pesticides and also helps in providing guidance for the selection of the pesticides. It has become a wide area for research now a days and a lot of research has been carried out worldwide for automatic detection of diseases. Ancestors followed the traditional method of examination of the fields which is the naked eye examination but it is very difficult to have a detailed examination in large fields by this method. In order to examine the whole field, more number of human experts are needed.That is very expensive and time consuming. Hence an automatic system is required which can not only examine the crops to detect infestation but also can classify the type of disease on crops. Nowadays, computer vision techniques provide effective ways for analyzing the images of leaves. Support Vector Machine is used for classification of images with and without disease based on the image features. This technique is much simpler when compared to the other automated techniques and it alsom provides better results.*

***Key Words* - Pest infection, image processing, crop disease recognition, K- means clustering.**

1. INTRODUCTION

Agriculture is the mother of all cultures. The main focus is only on enhancing productivity, without considering the ecological impacts. This has resulted in environmental degradation. As diseases that are occuring to the plants are inevitable, detecting the disease that has occured to the plant plays a major role in the field of agriculture. Plant pathogens consist of fungi, organism, bacteria, viruses, phytoplasmas, etc. These components are necessary for the diseases to occur in any plant system, which infects all types of plant tissues including leaves, shoots, stems, crowns, roots, tuber, fruits, seeds and vascular tissues. Hence, detection and classification of diseases that occured to the plant is an important and urgent task. The naked eye observation of experts which our ancestors has followed is the main approach adopted in practice for detection and identification of plant diseases. However, this requires continuous monitoring of experts which might be prohibitively expensive in large farms. We can analyze the image of leaves which are infected by the disease by using computer image processing technology. Thus we can extract the features of disease spot according to color, texture and other characteristics from a quantitative point of view.

Due to which, the consulting experts even cost high as well as it is time consuming too. The suggested technique proves to be beneficial in monitoring large fields of crops, Plant disease identification by visual way is more laborious task and at the same time it is less accurate and can be done only in particular areas. In plant varities, some of the general diseases are brown and yellow spots, or early and late scorch, and other fungal, viral and bacterial diseases. Image processing is the major technique which is used for measuring affected areas of disease, and also to determine the difference in the color of the affected area.

Based on the interaction between the analyst and the computer during the classification there are two types of classification. They are i) Supervised and ii) unsupervised.

Machine based approaches for large varities of disease detection and classification of agricultural products have become an important part of civilization.

Food is one of the important and basic needs for the survival of human being. World population is also increasing day by day. So it is very important to grow sufficient amount of crops to feed such a huge population. But as the days are passing by, plants are affected with various kinds of diseases, which causes great harm to the agricultural plant productions.

Considering the current techniques of knowledge management, they have some common defects in efficiency, scalability, and applicability. When considering the specific features of crop diseases and pest data, this paper analyzed and classified the key techniques in the field of crop diseases and pest in recent years.

The scope of the project are special research is being conducted on how to reduce the dependence on data annotation. Its principle is to continuously improve the performance of the model by using massive unlabeled data, learn from small samples, and gradually learn new knowledge by self-exploration to form an interactive learning process

2. LITERATURE SURVEY

Title: Leaf Segmentation and Tracking Using Probabilistic Parametric Active Contours

Author: Jonas De Vylder1,, Daniel Ochoa1,2, Wilfried Philips1, Laury Chaerle3, and Dominique Van Der Straeten

Abstract: Active contours or snakes are widely used for segmentation and tracking. These techniques require the minimization of an energy function, which is generally a linear combination of a data fit term and a regularization term. This energy function can be adjusted to exploit the intrinsic object and image features. This can be done by changing the weighting parameters of the data fit and regularization term. There is, however, no rule to set these parameters optimally for a given application. This results in trial and error parameter estimation. In this paper, we propose a new active contour framework defined using probability theory. With this new technique there is no need for ad hoc parameter setting, since it uses probability distributions, which can be learned from a given training dataset.

Title : Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in Arabidopsis thaliana

Author: Xu Zhang,* Ronald J. Hause, Jr.,†,‡ and Justin O. Borevitz*,‡,1

ABSTRACT: Leaf growth and development determines a plant's capacity for photosynthesis and carbon fixation. These morphological traits are the integration of genetic and environmental factors through time. Yet fine dissection of the developmental genetic basis of leaf expansion throughout a growing season is difficult, due to the complexity of the trait and the need for real time measurement. In this study, we developed a time-lapse image analysis approach, which traces leaf expansion under seasonal light variation. Three growth traits, rosette leaf area, circular area, and their ratio as compactness, were measured and normalized on a linear timescale to control for developmental heterogeneity. We found high heritability for all growth traits that changed over time. Our study highlights a cost-effective, high-throughput phenotyping approach that facilitates the dissection of genetic basis of plant shoot growth and development under dynamic environmental conditions.

Title: Multi-modality imagery database for plant phenotyping

Author: Jeffrey A. Cruz · Xi Yin · Xiaoming Liu · Saif M. Imran3 · Daniel D. Morris · David M. Kramer· Jin Chen

Abstract :Among many applications of machine vision, plant image analysis has recently began to gain more attention due to its potential impact on plant visual phenotyping, particularly in understanding plant growth, assessing the quality/performance of crop plants, and improving crop yield. Despite its importance, the lack of publicly available research databases containing plant imagery has substantially hindered the advancement of plant image analysis. To alleviate this issue, this paper presents a new multi-modality plant imagery database named "MSU-PID," with two distinct properties. First, MSU-PID is captured using four types of imaging sensors, fluorescence, infrared, RGB color, and depth. Second, the imaging setup and the variety of manual labels allow MSU-PID to be suitable for a diverse set of plant image analysis applications, such as leaf segmentation, leaf counting, leaf alignment, and leaf tracking. We provide detailed information on the plants, imaging sensors, calibration, labeling, and baseline performances of this new database.

Title: Tracking Rhythms in Plants, an automated leaf movement analysis program for circadian period estimation

Author: KathleenGreenham, Ping Lou1, Sara E Remsen, Hany Farid2 and C Robertson McClung

Abstract Background: A well characterized output of the circadian clock in plants is the daily rhythmic movement of leaves. This process has been used extensively in Arabidopsis to estimate circadian period in natural accessions as well as mutants with known defects in circadian clock function. Current methods for estimating circadian period by leaf movement involve manual steps throughout the analysis and are often limited to analyzing one leaf or cotyledon at a time. Results: In this study, we describe the development of TRiP (Tracking Rhythms in Plants), a new method for estimating circadian period using a motion estimation algorithm that can be applied to whole plant images. To validate this new method, we apply TRiP to a

Recombinant Inbred Line (RIL) population in Arabidopsis using our high-throughput imaging platform. We begin imaging at the cotyledon stage and image through the emergence of true leaves. Trip successfully tracks the movement of cotyledons and leaves without the need to select individual leaves to be analyzed.

3. EXISTING SYSTEM

Traditional reasoning based on logical rules promotes the automation process of reasoning for pest diagnosis and control measures to a certain extent, but it experiences obvious defects, such as insufficient learning ability, low data utilization rate, and inaccurate rate to be improved, which do not meet the requirements of practical application. Considering the increasing data on diseases and insect pests, problems are inevitable, such as missing information and prolonged time.

This method is not effective for knowledge map with a large number of relationships and sparse knowledge.

The existing system determines crop disease for a specific crop. Results are not accurate in the existing system.

4. PROPOSED SYSTEM

Our proposed system concentrates the rapid development of computer vision and image processing technology; it is possible to improve the accuracy and efficiency of the conventional agriculture work such as disease detection and early warning. In this system, we detect various parameters of the leaf. The parameters of the leaves are monitored with the help of digital image processing techniques. The controller is used for alerting or indication purposes.

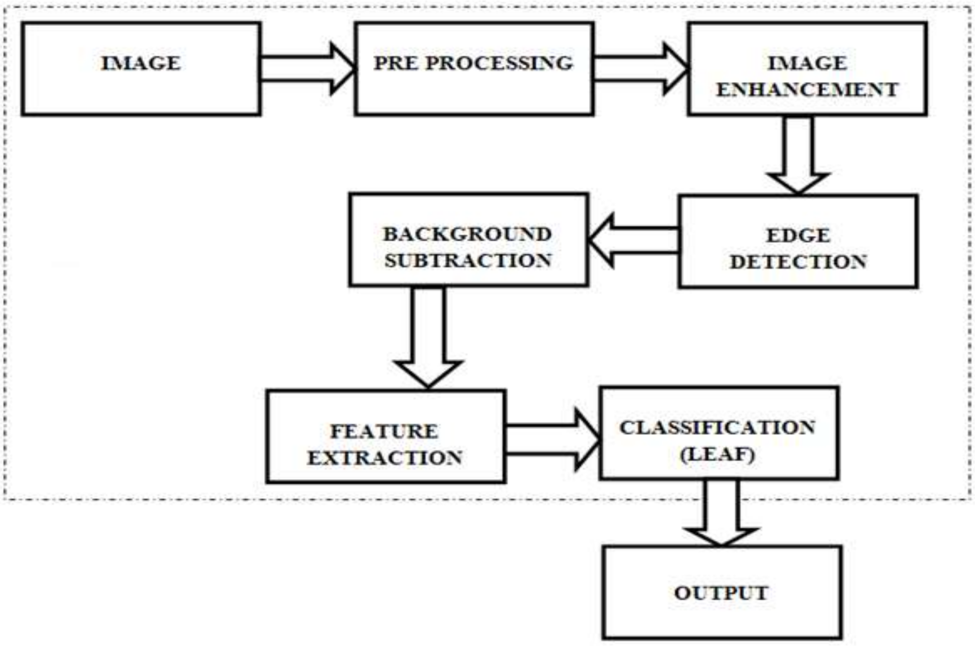

Fig 1- system architecture

5.1 INPUT REQUIREMENTS

Image of affected leaf part of the Paddy crop.

5.2 OUTPUT REQUIREMENTS

To show the name of the disease and the pesticides for that paddy crop.

5.3 FUNCTIONAL REQUIREMENTS

Classification of the type of disease that occurred to the paddy crop and to suggest the corresponding pesticide.

6. MODULE DESCRIPTION

6.1 PREPROCESSING

Data pre-processing is a crucial step within the data processing process. The phrase "garbage in, garbage out" is especially applicable to data processing and machine learning projects. Data-gathering methods are often loosely controlled, leading to out-of-range values impossible data combinations, missing values, etc. Analyzing the data which has not been carefully screened for such problems can produce the misleading results. Thus, the representation and quality of knowledge is first and foremost before running an analysis. Often, data pre-processing is that the most vital phase of a machine learning project, especially in computational biology.

The presence of more number of irrelevant and redundant information or noisy and unreliable data makes the knowledge discovery during the training phase harder to process. Data preparation and filtering steps can take considerable amount of time interval. Data pre-processing is a process which includes the following: cleaning, Instance selection, normalization, transformation, feature extraction and selection, etc. the merchandise of knowledge pre-processing is that the final training set.

6.2 IMAGE PRE-PROCESSING

Pre-processing is one of the general operation with images at its lowest level of abstraction. The iconic images produced after pre-processing are of very similar kind as the original data which is captured by the sensor. It is generally with an intensity image represented by a matrix of image function values (brightnesses). The mail purpose of pre-processing is improvement of image data that suppresses the unwilling distortions or it enhances some of the image features which are important for further processing. Although geometric transformations of the images (e.g. rotation, scaling, translation) are classified among several pre-processing methods here since similar techniques are used.

6.3 IMAGE ENHANCEMENT

In photography, colorimetric, and computing, a grayscale image carries only intensity information. Images of this type is also referred to as black-and-white image or monochrome image and are composed exclusively of reminder gray. These images are varying from black at its weakest intensity and to the white at its strongest.

The grayscale images are distinct from black-and-white images, which contains only two colors, black and white (also called bi-level or binary images). These grayscale images have many reminder gray in between. Grayscale images are often the results of measuring the intensity of sunshine at each pixel consistent with a specific weighted combination of frequencies (or wavelengths), and in those cases they are said as monochromatic proper when only one frequency (in practice, a narrow band of frequencies) is captured. The frequencies can within the ory be from anywhere in the spectrum (e.g. infrared, light, ultraviolet, etc.).

A colorimetric (or more specifically photometric) grayscale image is a picture that features a defined grayscale colorspace, which maps the stored numeric sample values to the achromatic channel of a typical colorspace, which itself is predicated on measured properties of human vision.

If the first color image has no defined colorspace, or if the grayscale image isn't intended to possess an equivalent human-perceived achromatic intensity because the color image, then there is no proper and unique mapping from that color image to a grayscale image.

6.4 EDGE DETECTION

Edge detection may be a well-developed field on its own within image processing. The region boundaries and edges are closely related to each other, since there often occurs a pointy adjustment in intensity at the region boundaries.

The edges identified by the edge detection techniques are often disconnected. To segment an object from a picture however, one needs closed region boundaries. The desired edges are the boundaries found between objects or spatial-taxons.

Spatial-taxons are the ones which provides data such as information granules which consists of a crisp pixel region.That is stationed at abstraction levels within hierarchical nested scene architecture. They are almost like the Gestalt psychological designation of figure-ground, but are extended to incorporate foreground, object groups, objects and salient object parts. Edge detection methods are often applied to the spatial-taxon region, within the same manner they might be applied to a silhouette. This method is especially useful when the disconnected edge is a component of an illusory contour

Segmentation methods also can be applied to edges obtained from edge detectors. Edge detection is a well-developed field on its own within image processing. The edges and region boundaries are closely related to each other. Since there is also a sharp adjustment in intensity at the region boundaries.

The edges identified by these edge detection techniques are often disconnected. Inorder to segment an object from a particular image closed region boundaries are needed. The desired edges are known to be the boundaries between such objects or spatial-taxons.

Spatial-taxons are information granules, consisting of a crisp pixel region, stationed at abstraction levels within hierarchical nested scene architecture. Edge detection methods can also be applied to the spatial-taxon region, in the same manner which they would be applied to a silhouette. This method is considered as a necessary method because the disconnected edge is part of an illusory contour.

6.5 BACKGROUND SUBTRACTION

Background subtraction may be a technique for separating out foreground elements from the background and is completed by generating a foreground mask. This system is employed for detecting dynamically moving objects from static cameras. Background subtraction technique is vital for object tracking. There are several techniques for background subtraction

6.6 FEATURE EXTRACTION

Through feature extraction an initial set of data is minimized to manageable groups for processing. A characteristic of those large data sets may be a sizable amount of variables that need tons of computing resources to process.

The method of feature extraction is beneficial once you got to reduce the amount of resources needed for processing without losing important or relevant information. Feature extraction also can reduce the quantity of redundant data for a given analysis. Also, the reduction of the info and therefore the machine's efforts in building variable combinations (features) facilitate the speed of learning and generalization steps within the machine learning process.

6.7 CLASSIFICATION

K-means clustering is used for classifying the infected crop.

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells.

The problem is computationally difficult (NP-hard); however, efficient heuristic algorithms converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both k-means and Gaussian mixture modeling. They both use cluster centers to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes.

The algorithm has a loose relationship to the k-nearest neighbor classifier, a popular machine learning technique for classification that is often confused with k-means due to the name. Applying the 1-nearest neighbor classifier to the cluster centers obtained by k-means classifies new data into the existing clusters. This is known as nearest centroid classifier or Rocchio algorithm

7. CONCLUSION

With the advancement of data innovation, information in the field of harvest infections and creepy crawly bothers have been aggregated. Building information diagram can improve the administration, the degree of sharing and utilization of harvest vermin and infections information. The improvement of insightful farming and accuracy horticulture is enormously significant and is likewise a future look into hotspot. The foundation, innovation, what's more, utilization of information guide of harvest ailments and creepy crawly vermin were condensed from the point of view of the application of information map in crop sicknesses and bug bugs. The primary difficulties looked by the information guide of yield illnesses also, creepy crawly irritations were condensed, and future research bearings were prospected. The mix of information map book and vermin information will advance the mechanization and intellectualization of yield vermin and ailment related frameworks. Numerous endeavours have been made to examine the information map of yield infections and creepy crawly bugs, yet they were not great also, inside and out, and require further research.

REFERENCES

[1] Kavitha Subramani and Murugavalli S , Recognizing Ancient Characters from Tamil Palm Leaf Manuscripts using Convolution Based Deep Learning, International Journal of Recent Technology and Engineering, Sep 2019, Volume 10, Issue 10, pp 6873-6880.

[2] Kavitha Subramani and Murugavalli S , Application of Machine Learning for Identification of Characters in Ancient Tamil Medicinal Palm Leaf Manuscripts, International Journal of Innovative Technology and Exploring Engineering, Jan 2020 , Volume 9, Issue 3, pp 541-548.

[3] H. Arnaout and S. Elbassuoni, ``Effective searching of RDF knowledge graphs,'' J. Web Semantics, vol. 48, pp. 66_84, Jan. 2018.

[4] C. B. Wang, X. Ma, J. Chen, and J. Chen, ``Information extraction and knowledge graph construction from geoscience literature,'' Comput. Geosci., vol. 112, pp. 112_120, Mar. 2018.

[5] P. K. Thornton et al., ``Site selection to test an integrated approach to agricultural research for development: Combining expert knowledge and participatory geographic information system methods,'' Int. J. Agricult. Sustainability, vol. 4, no. 1, pp. 39_60, 2006.

[6] M. M. Yusof, N. F. Rosli, M. Othman, R. Mohamed, and M. H. A. Abdullah, ``M-DCocoa: M-agriculture expert system for

diagnosing cocoa plant diseases,'' in Recent Advances on Soft Computing and Data Mining (Advances in Intelligent Systems and Computing), vol. 700, R. Ghazali, M. Deris, N. Nawi, and J. Abawajy, Eds. Cham, Switzerland: Springer, 2018.

[7] Y. Yao and W. Xu, ``Construction of ontology knowledge base for water and soil conservation measures suitability evaluation,'' in Journal of Fuzhou University (Natural Science Edition), 2016, pp. 188_195.

[8] W. D. van Eeden, J. P. de Villiers, R. J. Berndt,W. A. J. Nel, and E. Blasch, ``Micro-Doppler radar classi_cation of humans and animals in an operational environment,'' Expert Syst. Appl., vol. 102, pp. 1_11, Jul. 2018.

[9] B. Stewart-Kosteret al., ``Expert based model building to quantify risk factors in a combined aquaculture-agriculture system,'' Agricult. Syst., vol. 157, pp. 230_240, Oct. 2017.

[10] Z. Xu, Y. P. Sheng, L. R. He, Y. F. Wang, ``Review on knowledge graph techniques,'' J. Univ. Electron. Sci. Technol. China, vol. 45, no. 4, pp. 589_606, 2016.

[11] L. Qiao, L. Yang, D. Hong, L. Yao, and Q. Zhiguang, ``Knowledge graph construction techniques,'' J. Comput. Res. Develop., vol. 53, no. 3, pp. 582_600, 2016.

[12] T. Yu et al., ``Knowledge graph for TCM health preservation: Design, construction, and applications,'' Artif. Intell. Med., vol. 77, pp. 48_52, Mar. 2017.

[13] J. L. Martinez-Rodriguez, I. Lopez-Arevalo, and A. B. Rios-Alvarado, ``OpenIE-based approach for knowledge graph construction from text,'' Expert Syst. Appl., vol. 113, pp. 339_355, Dec. 2018.

[14] N. Guan, D. Song, and L. Liao, ``Knowledge graph embedding with concepts,'' Knowl.-Based Syst., vol. 164, pp. 38_44, Jan. 2019.