Analyzing the Effectiveness of Digital Image Processing in Plant Disease Detection

---

A Research Proposal
Presented to the Faculty of the Basic Education Department
of the Holy Cross College of Calinan

In Partial Fulfillment of the Requirements
In Practical Research 1

BY:

Navales, Christian Kyle

Alpes, Aldren

Ponpon, Kaya Xenia

Salarde, Myra Jessa

Pontino, Olive Gracelle

Mangubat, John Mark

August, 2020

TABLE OF CONTENTS

Appendices

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Background of the Study

Agriculture is a science of soil planting, harvesting and rearing of livestock, as well as a science or art of growing plants and animals that are useful to humans and, to various degrees, the preparation of such items for human use and disposal (Bareja, 2019). It is the major and crucial source of income for a variety of countries, including in the Philippines. On the other hand, agricultural production is the use of cultivated plants or animals to produce goods to support or improve human life (Chait, 2020). It leads to increased volumes that can stabilize and boost domestic and export markets, as well as lower food costs for customers. In addition, with access to more varied seeds, farmers can grow new crops that allow customers to make more choices on local markets. However, plant crop disease is a significant cause of declining the quantity and quality of produce, resulting in economic losses. Symptoms of plant disease are visible in different areas of plants. However, the most common application of plant leaves is to diagnose infection (Kumar & Vishonoi, 2020).

A large variety of plant diseases from a few hundred nucleotide viroid to higher plants cause disease in the crops. Their effects vary from minor symptoms to disasters in which vast fields cultivated with food crops are devastated. Catastrophic plant disease aggravates the global food production shortage in which at least 800 million people are inadequately fed. Plant pathogens are hard to observe because their species differ in time, space and genotype (Scott & Strange, n.d).

Subsequently, Peronosclerospora Philippinensis, a triggering pathogen of mildew maize, is one of the major diseases identified in some maize-growing countries especially the Philippines. High incidence of disease has been documented in many regions of the country, specifically in Northern Luzon and many parts of Mindanao, despite breakthroughs in controlling or alleviating the disease through cultural and chemical regulation (CABI, 2020). It is known to be the most pernicious mildew pathogen in maize. In certain areas, the yield loss is 40% to 60%, but in the best case it can be as high as 80% to 100%. In the 1974-5 season, national yield losses in the Philippines were estimated at 8 per cent, with a dollar value of US$23 million. In general, the yield loss is related to the percentage of contaminated plants, but the loss may be even greater due to secondary attacks from stem bore and other secondary parasites and pathogens (Murray, 2009).

In 2013, at least nine villagers in the Davao del Norte town of Sto. Tomas was infected with Panama disease, also known as fusarium wilt, causing the local government to provide technical assistance to affected farmers. The province of Davao del Norte had the highest Fusarium incidence in Davao Region with 13,743 hectares of contaminated banana farms (Palicte, 2019). The soil-borne fungus Fusarium oxysporum f. sp. cubense causes Fusarium wilt of banana, also known as Panama disease. The fungus reaches the plant through the roots and colonizes the xylem vessels, obstructing water and nutrient supply (Vezina, 2020).

Identification of plant diseases is the solution in avoiding productivity losses and improving the quality of the agricultural commodity. Modern approaches are accurate but it requires an available human resource for visual examination of plant leaf patterns and for the diagnosis of disease. Traditional approach consumes more effort and is timeconsuming work. The problem with this conventional approach is that it is necessary to decide which features are relevant in each image. As the number of groups to classify grows, function extraction becomes more difficult (Mahony et al., 2019).

In huge farmland, early diagnosis of plant disease using automated techniques which is a vision-based automatic detection of plant disease using the Image Processing Technique that would reduce productivity losses. Image processing algorithms are developed for detecting plant infection or disease by identifying and recognizing the color of the leaf area. K mean algorithm is used for color segmentation and GLCM or the Gray Level Co-occurrence Matrix is used for disease classification (Chandramouleeswaran & Meyyappan, 2018). Its functions interpret the texture of an image by calculating how often pixel pairs with particular values and in a defined spatial relationship occur in an image, creating a GLCM, and then extracting statistical measurements from this matrix (Messaouda, 2020).

Though image processing is ideally expected to carry automatic processing, control and analysis of such, visual knowledge plays an increasingly important role in many areas of wide range of disciplines and fields of science and technology. Therefore, the goals of this study is to assess the effectiveness of the image processing method to detect various of plant diseases on the basis of the data and inputs given and are available.

Statement of the Problem

This study focused on assessing the Effectiveness of Image Processing in Plant Disease Detection to a variety of plants by analyzing former studies, the researchers sought to answer the following questions:

1.) Are there any significant factors to be considered which affect the disease detection?

2.) How many percent does these factors affect the disease detection?

3.) How effective is image processing in detecting diseases present in plants in terms of:

a.) Accuracy of the disease detection

b.) Consumed time

Review of related literature

This part of the paper represents the related literature and studies that have contributed and used as a reference or guide for the researchers. This would also present the theoretical framework and conceptual framework to fully elaborate the study to be done.

Technological advancement is relevant in today's world. Furthermore, it can be used in the different fields of society such as in agriculture. The diseases have caused economic, social, and ecological loss to farmers. Diseases on plants can lead to a significant reduction in both the quality and quantity of agricultural products.

The studies of plant disease refer to the studies of visually visible patterns on the plants. Before globalization, the monitoring and analysis of plant diseases were done manually by the expertise person in that field. This requires enormous amount of work and also requires excessive processing time.

Technology is advancing and with the help of innovated digital methods the process of identifying disease on plants would be much easier and precise.

I.) VARIOUS FRUITS TESTED FOR DISEASE DETECTION THROUGH IMAGE PROCESSING

Disease detection on plants happen during the observation of its specific patterns. Furthermore, there are methods used to identify the patterns on plants such as spectroscopic and imaging technology. Through certain methods, farmers were able to easily identify various plants diseases during an earlier stage which is a benefit to reduce the spreading of disease (Bharate & Shirdhon kar, 2017).

a.) *Passion Fruit Disease Detection Using Image Processing*

An image processing approach is proposed for identifying passion fruit diseases. This can be used to identify passion fruit diseases quickly and spontaneously. This system is composed of the following main steps; Image Acquisition, Image Preprocessing, Image Segmentation, Feature Extraction, Dataset Preparation, Training & Testing. Healthy and two types of passion fruit diseases namely passion fruit scab and woodiness images, were used. This was tested according to passion fruit disease type and its stages, such as mild, moderate and severe. K-Means clustering was used for segmentation. Images were clustered according to k values, such as 2, 4, 6 and 8. Before the segmentation, images were converted to RGB, L*a*b, HSV and Grey color models to find out the most suitable color model for this approach. Local Binary Pattern was used for feature extraction and Support Vector Machine was used for creating the archetype. 70% of each dataset was used to train the SVM and the other 30% was used to test the archetype. Collected images included the leaves infected by Powdery Mildew and Downy Mildew. Features were extracted based on both color and texture for taking accurate disease data. The classification model was used to detect and identify the leaf disease. LSVM was used in this study for the classification of leaf diseases. This system could detect and identify the examined disease successfully. The accuracy of this system was 88.89% (Dharmasiri & Jayalal, 2019).

b.) Mango Disease Detection Using Image Processing

Mango is one of the most commonly consumed fruits around the world, and so dealing with diseases and pests that affects the crop is quite a challenging matter for most of the agriculturists. This problem motivated the researchers to find and develop a solution for the uprising problem. The study elaborated new techniques to identify and diagnose the diseases that affect the mango plant. The process was a built-in function in MATLAB. They undergo through the process called "contrast enhancement", done by scaling pixel values and its features such as contrast, correlation, energy, entropy, homogeneity, cluster, cluster shade, variance, and dissimilarity. The system proposed an accuracy of 90% in testing on 92 samples. An image is being captured on the infected part of the mango plant and it would be sent to the system via the internet. Once the image would be received by the internet it would be input to the system for some testing and image processing. A technical process would be used which includes steps like resizing of images, image acquisition, image pre-processing, image segmentation, feature extraction and classification of images. Image segmentation follows after previous procedures on which the images would be divided into parts for better results. They had discovered three different diseases that infect the mango plant namely, Anthracnose, Powdery Mildew and Red Rust (Veling, Kalelkar, Ajgaonkar, Mestry & Gawade, 2019).

c.) Banana Plant Disease Detection and Grading Using Image Processing

Banana is the most vigorous fruit that was high in demand mostly in Asia. Large productions in industries of bananas managed to provide a large number of products to be exported. They described the information about the diseases that can be acquired by the banana plants; symptoms can be found on the plant leaves. In solution to this problem, they proposed a software solution for detecting banana plant diseases, which also includes the percentage infection using the image processing methods. The system involves various steps, that include dataset detection, pre-processing image, feature creation, and artificial neural network-based training. They developed a new system which could extract features and information about the plant diseases and could classify the diseases using the artificial neural network (ANN) Algorithm. The first step is the image acquisition, in this stage the image for various plant diseases would be captured on its leaf by a digital camera with a resolution of 16 megapixels. All the images would be filed to a jpg format on a disk which would be created as the database. The database creation serves as the most important part for it is responsible for the accuracy of the classifications of the diseases detected. The next process includes image pre-processing, through this process, the image would be enhanced as it passes through image cropping, resizing, and color conversion. Since the images differ in terms of dimensions, they decided to have a measurement of 256 by 256 dimensions for efficient processing. After the image pre-processing, it would proceed to feature extraction. This part would present the features and the major attributes contained by the image which would be used for the classification of the diseases being detected. There would be two kinds of features extracted from the image: the histogram of template features and color features. The feature file that was created would be given as input to the neural network toolbox for automatic detection of the diseases (Tigandi & Sharma, 2016).

d.) Apple Fruit Disease Detection using Image Segmentation Algorithm

It takes a long time to manually identify defective fruit. However, only a few segmentation algorithms can detect fruit diseases. The researchers developed a method for detecting disease on fruit by using the K - clustering segmentation algorithm and color images of fruits for defect segmentation. To begin, the RGB color image is converted to Lab color space. Taking the absolute difference between each pixel and the clustering center in Lab color space does clustering. The researchers used apples as a case study and tested the suggested method on defective apples. The results show more than 95% segmentation accuracy for three common diseases on apple.

The proposed approach is tested on defective apples. For the segmentation of diseased parts of the apple, researchers used three common diseases: apple scab, apple rot, and apple blotch. The study's data visualization includes several photographs of the data collection contaminated with different diseases. The fact that the data set contains a large number of variations makes it more practical. The research presents an effective image segmentation method based on color features from the images using the K-means clustering technique. There are two steps to defect segmentation. The clustering process begins with the pixels being clustered based on their color and spatial properties. We looked at three different forms of apple diseases to support the proposed method: apple scab, apple rot, and apple blotch.

For segmenting the diseased portion of the apple with three clusters, the proposed method used the K-means clustering technique. Experiments on images of apple fruit with three common diseases demonstrated that the system performs satisfactorily in terms of accuracy, performance, and automation. The proposed method appears to be capable of accurately segmenting the defected areas of apple fruits in image databases, according to simulation results (Abhijeet & Patil, 2017).

II.) AN APPLICATION FOR FRUIT DISEASE DETECTION USING IMAGE PROCESSING

There are many methods being used to identify plant pathologies. Though, most of the disease have no visible symptoms and have developed quickly were a caused to the rapid loss of crops. Some of those case, sophisticated analysis, laboratory testing, and other processes were necessary. Methods of image processing were applied in order to attain the results.

This starts off with image segmentation. It is the process of partitioning a digital image into multiple segments such as sets of pixels. The image segmentation performs a major role within the field of image processing thanks to its wide range of applications in the agricultural fields to spot plants diseases by classifying various diseases. With the use of classification, it helps classifies the plants diseases on different morphological characteristics (Masood & Khan, 2016).

a.) Matlab application for disease detection

Maturity is a key factor that determines the storage life and ripening quality of fruits and in order to provide marketing adaptability and to bargain for acceptable eating quality to the buyer, it is very crucial to determine the right maturity stage. The proposed system is based on implementation of image processing techniques on the Joint Photographic Experts Group or JPEG images of different maturity stages of the plum variety 'Satluj Purple' grown under sub-tropical conditions. A uniform thresholding operator available in MATLAB Image Processing Toolbox was used for locating the plum fruit in an image. In uniform thresholding, pixels above the specified level were set to white while those below the specified level were set to black. In this manner, the pixels pertaining to the fruit were separated from the background pixels and the region of interest (ROI) was segmented. The output of this phase was a binary image which had 1's in the plum fruit region and 0's in the background region. Multi-Attribute Decision Making theory was used in the process for taking final decisions. The developed system accurately determined the maturity level of the fruit. The color was found to be the most dominant factor for classifying the plums according to maturity level. The error percentage was less than 2.4%, when the length and width were calculated from application were compared with the manual readings (Kaur, Sawhney & Jawandha, 2018).

III.) IMAGE PROCESSING TECHNIQUE IN DETECTING PLANT DISEASES

The development in certain disciplines of computer science like Artificial Intelligence, Pattern Recognition, Image Processing, Neural Networks, and etc., promise the required technological support to tackle the various issues in computer vision. Applications considered in these studies are concerned with processing of images of agriculture and horticulture crops affected by different kinds of diseases.

a.*) Plant Disease Detection using Digital Image Processing and GSM*

There are two methods for obtaining photographs of the plant leaf. The first method is to capture an image with an external camera (in this case, a web camera), and the second method is to retrieve an image from an email. The color space of the input image is transformed. The methods are preceded by various pre-processing techniques such as image clipping for obtaining the concerned region by cropping; then, image segmentation, partitioning of images into various parts of the same feature or having some similarity; and finally, image segmentation, partitioning of images into various parts of the same feature or having some similarity. They discovered that the morphological result surpasses the other features. The SVM then creates a hyper plane or a group of hyper planes in a high or infinite dimensional space that can be used for classification, regression, or other tasks, while the mat file stores the various features of the affected leaf. Finally, the transmission of files or photographs in which the name of the pesticide would be sent to the farmer's mobile phone via GSM after identification of the disease and its corresponding pesticide. The accurate identification and classification of plant diseases is critical for effective crop production, and this can be accomplished with digital image processing. (Marathe, 2017)

The research reviews aid readers and future researchers in comprehending the various aspects of research on disease detection on plants using image processing. Different techniques are proposed by different researchers with the help of digital image processing for rapid and accurate plant disease identification in agricultural fields, for greater quality and quantity of crops in agricultural fields. The two most important factors to remember are time and accuracy. So if compared to manual processes, the accuracy of various approaches using image processing techniques can be enhanced. These modern methods also help save time. The best techniques will be studied among all of these various techniques to see which ones have the most benefits. Regardless of its maximum benefits, every technique has its own set of limitations. While symptoms were evident on the appearances of plants, symptoms that were not visible could also be considered in some cases. As a result, image analysis can be used to detect symptoms quicker. To classify the morphological features of the plants, classification of various diseases will be needed.

Different image processing techniques have been established in the studies in order to achieve a higher level of representation in identifying disease on a specific variety of plants.

Theoretical Framework

This study is anchored in the Fuzzy Set Theory that was introduced by Lotfi A. Zadeh, professor for computer science at the University of California in Berkeley in 1965. It is commonly used in optimal control models, decision-making under uncertainty, demographic data, behavioral experiments, and artificial intelligence techniques. Fuzzy set theory is a technique for explaining circumstances where the information is inaccurate or ambiguous other forms of imperfection in image processing include geometric fuzziness and ambiguous information of image characteristics. (Zadeh 1965)

The Fuzzy Set Theory is applicable for this research because it has content and a background about image processing, which is relative to the study. Fuzzy image processing is the selection of all approaches that recognize, represent and process images, segments and characteristics as fuzzy sets. The representation and processing are focused on the chosen fuzzy approach and the problem to be solved. The Researcher proponents attempt to represent a concept or characteristic of the image related to perception in some instances to be able to cope with information uncertainty.

Conceptual Framework

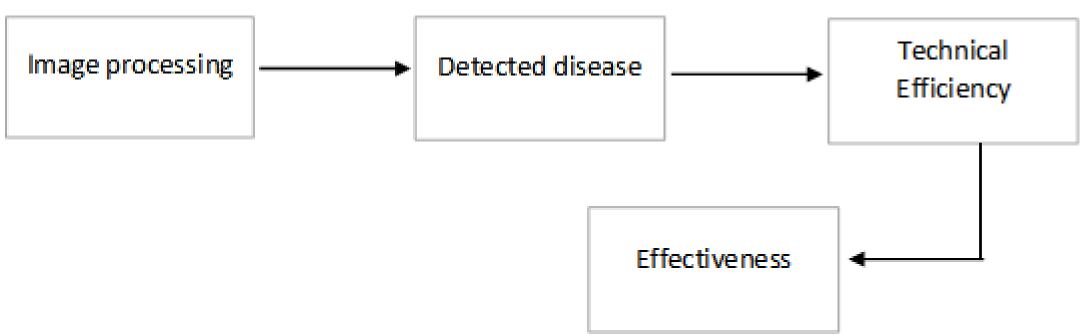

Figure 1: Conceptual Framework

Figure 1 shows the conceptual framework of this study that tackle the variables involved. The process in the framework shows, that image processing is the method used to detect the diseases, and these detected diseases would then identify the technical efficiency of image processing by giving precise and accurate output. The data that is acquired would then be used in order to provide information that would help identify the effectiveness of image processing in plant disease detection, which would give people information on the effectiveness of image processing in plant disease detection.

Significance of the study

The purpose of this study is to measure the effectiveness of image processing to identify particular diseases on variety of plants. In which it also includes the mission to understand the characteristics or significant factors that are related to the identification of disease on the plant which makes the image processing be easier to detect diseases. The types and ages of plants would be followed and the occurrence of health problems in relation to these diseases.

The research problem indicates the area of concern, which is the effectiveness of image processing and identification of diseases present on plants. Furthermore, the purpose of this qualitative study is to generate specific knowledge about the how effective image processing is in terms of disease detection on plants. The results of this study would be of great benefit to the future researchers, local laboratory experts, consumers and businessmen.

The ideas presented may be used as reference by the future researchers in conducting the new research or in testing the validity of other related findings. This study would also serve as their cross-reference that would give them a background or an overview of Plant Disease Detection. Along with it, Department of Agriculture and local laboratory experts may use the study as a back-up reference to prove if the image processing is an effective and reliable method in terms of plant disease detection for farmers. Moreover, the result of this study would be beneficial to consumers and businessmen, thus, enables the farmers to discern whether the crops that are produce is infected with diseases and can also give them an advantage to effectively manage their time harvesting and separating fruits.

Scopes and Limitations of the Study

This study focuses mainly on evaluating the effectiveness of image processing in the detection of different plant diseases. It would not include its other functions such as human recognition, object recognition, video processing, and more. The study would be limited in rating the effectiveness of image processing in the field or sector of agriculture and would not include other sectors like the medical field.

The study would only apply digital image processing to recognize the diseases present on plants and would not cover conventional image processing. More methods on distinguishing plant diseases are beyond the range of the scope of this study. The study would be using the data gathered from several past pieces of research from year 2013-2020, articles and journals about image processing as a tool for plant disease detection, which is the basis to prove and check its effectiveness in disease detection.

Definition of terms

Digital Image Processing	-The method used to perform some operation to the uploaded plant image, in order to extract some useful information to determine the diseases present on the plant.
Technical Efficiency	– This step identifies the efficiency of image processing in terms of disease detection if it is able to acquire accurate and precise output with minimum effort.
Plant Diseases	**-** These are the diseases found on the plants which are identified with the use of image processing.
Coding	**-** It is the process of categorizing and analysis of statements/ codes gathered from past researches to a set of meaningful, cohesive categories.
Matlab	**-** It is a programming platform used in the study entitled matlab application for disease detection to create a digital image processing application

CHAPTER 2

METHODS

This chapter provides and discusses the various methods in gathering data, the research design, ethical considerations, and instruments that would be used in conducting this study.

Research Design

The method used for this study is content analysis of image processing, disease detection, and plants. Content analysis assures not only that all units of analysis receive equal treatment, whether they are entered at the beginning or at the end of an analysis but also that the process is objective in that it does not matter who performs the analysis or where and when (Krippendorff, 1989).

In this study, Content Analysis would be utilized systematically to collect data from a set of texts and would be narrowed down into codes. The researchers would focus on counting and measuring the effectiveness of detecting plant disease detection with the use of digital image processing.

Ethical Consideration

Ethics refers to the set of moral principles or values that generally governs the conduct of an individual or group. Hence, in this study the researchers ensure professionalism, confidentiality of data, anonymity of the authors name, stability, accuracy, and reliability.

There are three types of reliability identified by Krippendorff (2004): stability, reproducibility, and accuracy. Stability is concerned with possible changes in the coding outcomes of repeated testing. A coder re-analyzing the same manifesto after a period of time has this type of reliability to highlight any intra-coder disagreement. Reproducibility is a stronger measure of reliability, also called inter-coder reliability. This test measures the degree of replication of coding outcomes by two different coders that function independently. Intra-coder variations and inter-coder gaps in the understanding and execution of the coding system are covered. Accuracy tests the adherence to any canonical norm of the coding process and data generation technique, and is considered to be the highest reliability measure. It can be used efficiently at the training stage because it is possible to equate the output of the coder to any real results.

Data Analysis

Sampling Frame

Jewel Ward (2012) stated that, if a researcher applies the quantitative method, then systematic sampling is used to provide for the generalization of the results. Systematic sampling is an extended implementation of probability sampling in which each unit of the group is selected at regular periods to form a sample. In this study, a systematic sampling method would be used wherein the researchers would zero down on the desired population they want to research. The researchers would calculate the sampling interval by dividing the entire population size by the desired sample size. The samples used in this study would be former studies, research articles, and literatures conducted 4 years in time.

Unit of Analysis

According to Trochim (2020), unit of analysis is the major entity that you are analyzing in your study. The unit itself may be physical, temporal, or conceptual (Spurgin & Wildemuth, 2009). For this study, the physical unit of analysis would be used to identify the population, measure variables, or report the analysis. The units to be observed for the data analysis includes words, phrases, and statements that are found in former studies, research articles and literatures. The study had analyzed and identified the frequency of the words, phrases, and statements about the accuracy of the image processing on disease detection to acquire the data needed to answer the research questions. Thus, the information found in research articles, and literatures on the percentage of accuracy and the time consumed during the disease detection had been further evaluated to prove the effectiveness of image processing in plant disease detection.

Coding Procedure and Scheme

A content analysis would be conducted where former studies involving plant disease detection using image processing as the unit of analysis. These former studies would be referred to as 'unit' in this study. Thematic coding is a type of qualitative data analysis that finds themes in text by analyzing the meaning of words and sentence structure. It can be deductive or inductive. For this study, inductive thematic coding would be used in this process where codebook wouldn't be necessarily used but rather building on from scratch based on the data collected. Five research studies would be coded. Coding categories would be identified as: (1) unit, (2) statements/ codes, (3) themes, and (4) frequency.

Unit	Statements/ Codes	Themes	Frequency
Unit 1			
Unit 2			
Unit 3			
Unit 4			
Unit 5			

Figure 2: Sample table to be used in coding

Themes	**Frequency**	**Percentage**
Theme 1		
Theme 2		
Theme 3		
Theme 4		
Total		

Figure 3: Sample table to be used in getting the frequency and percentage

CHAPTER 3

RESULTS AND DISCUSSION

This chapter presents the findings of the researchers' data collection efforts. Discussions about the analysis conducted in this study, as well as visual representations and tables, are included here for better comprehension. Studies from the Review of Related Literature are also discussed to demonstrate whether or not the study's findings are agreed upon.

Problem #1: Are there any significant factors to be considered which affects the disease detection?

Factors considered which affects the disease detection
COLOR
APPEARANCE
IMAGE QUALITY
PLANT AGE

Figure 4: Factors considered which affects the disease detection

The figure above shows the different factors which affects the disease detection in plants. These were identified by the researchers within the selected analyzed units. With the following factors, this makes it difficult to detect plant disease.

Problem #2: How many percent does these factors affect disease detection?

Percentages of factors affecting disease detection	
COLOR	38.46 %
APPEARANCE	30.77 %
IMAGE QUALITY	23.08 %

PLANT AGE	7.69 %

Figure 5: Percentages of factors affecting disease detection

The figure above shows the different percentages of the factors that affects disease detection. This depicts that Color, such as color variations and color segmentations greatly affect disease detection with the percentage of 38.46%. On the other hand, factors such as appearance, image quality, and plant age have also affected the disease detection.

Problem #3: How effective is image processing in detecting diseases present in plants in terms of:

a.) Accuracy of the Disease Detection

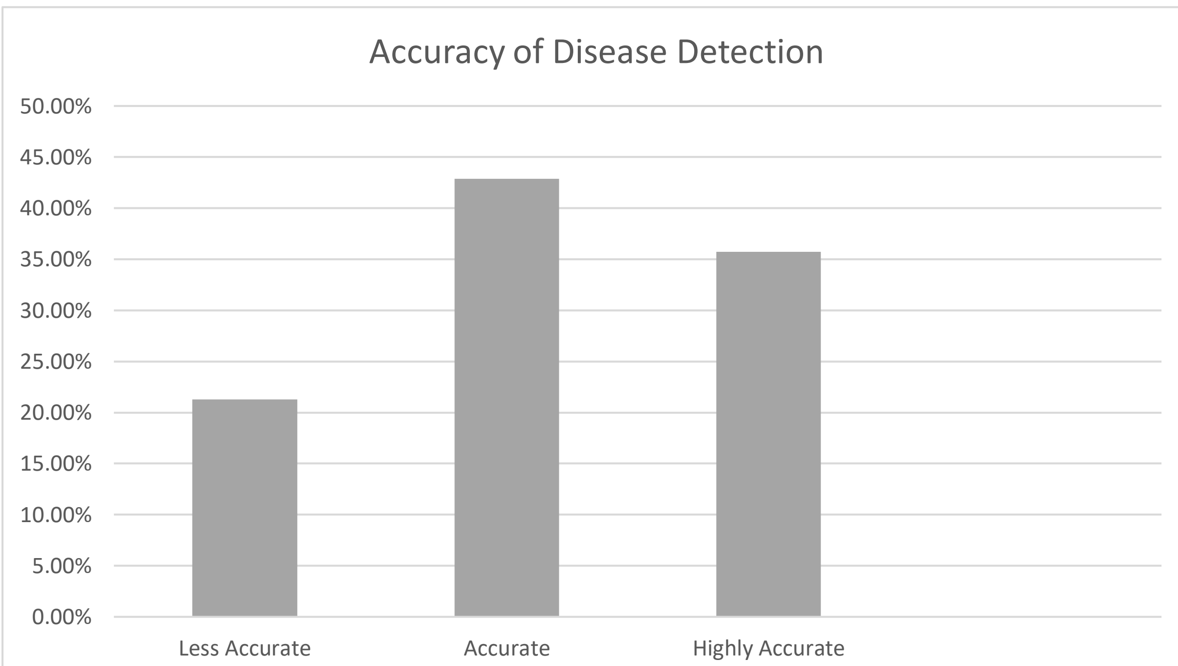

Figure 6: Accuracy of Disease Detection

The figure above summarizes the level of accuracy in disease detection taken from the different units analyzed by the researchers. Apparently, 21.43% were considered as less accurate, 42.86% were considered as accurate, and 35.71% were considered as highly accurate.

b.) Consumed Time

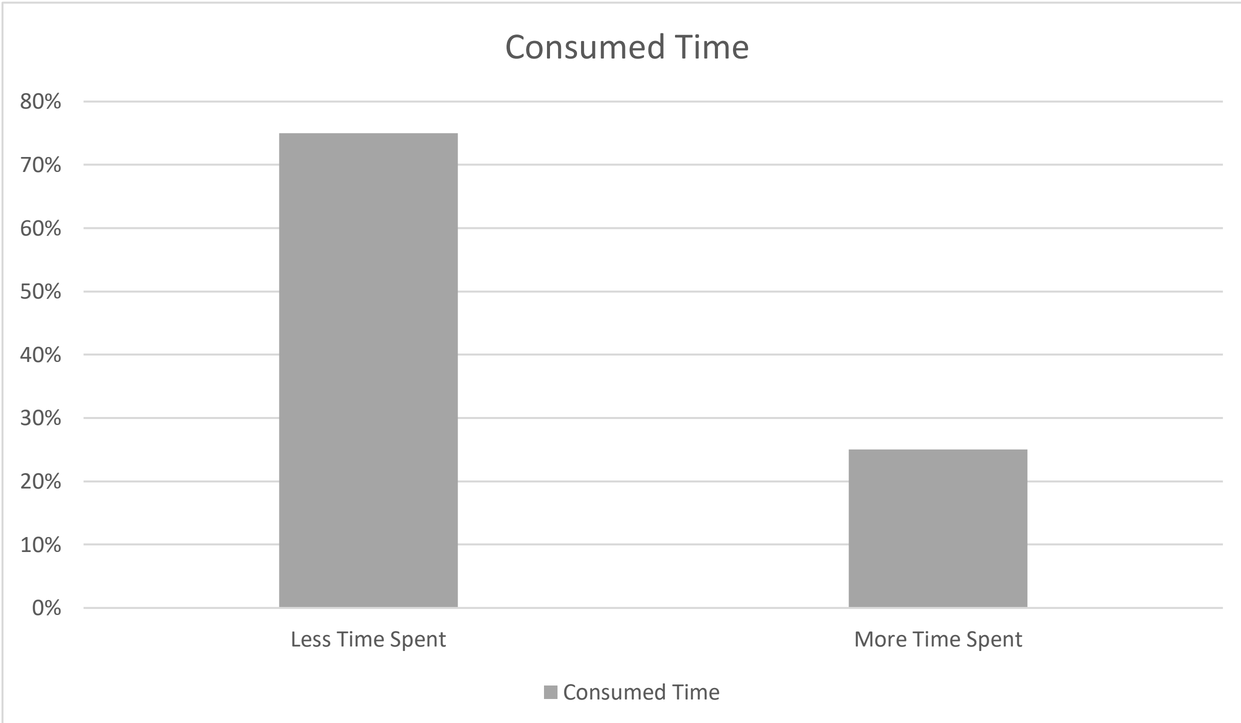

Figure 7: Consumed time in detecting diseases

The figure above summarizes the consumed time in the detection of diseases taken from the different units analyzed by the researchers. Apparently, there are less time spent in the process of detecting diseases.

CHAPTER 4

CONCLUSION AND RECOMMENDATION

This chapter summarizes the results garnered in the past chapter along with the recommendations that were given.

Conclusion

Based on the findings about the effectiveness of image processing in plant disease detection, the following conclusions were made:

There were four (4) identified factors which affects the disease detection namely; color, appearance, image quality, and plant age. The following factors makes it difficult to detect plant diseases. These factors have different percentages which depicts the severeness of their impact to the process of disease detection. Color, such as color variations and color segmentations with 38.46% greatly affect the process of disease detection. On the other hand, appearance with 30.77%, image quality with 23.08 %, and plant age with 7.69% have also affected the disease detection. Then, the level of accuracy in disease detection that were taken from the different units analyzed apparently had 21.43% considered as less accurate, 42.86% considered as accurate, and 35.71% as highly accurate having a total accuracy of 78.57%. The consumed time in the detection of diseases have less time spent in the process of detecting diseases. Therefore, having a total accuracy of 78.57% and a less time spent in the process of detecting diseases, we conclude that the digital image processing is effective.

Recommendation

Based on the acquired results, the following recommendations were made:

Firstly, a special program must be conducted to all of those who are in the agricultural sector. This is to present and educate them more about image processing. Knowing more about this image processing would greatly help farmers in monitoring and preventing the spread of plant diseases.

Secondly, technological advancement has such great possibilities. One's innovation and intervention in creating new programs, methods, or applications which can be used in plant identification, plant classification, and etc. could be the future.

Thirdly, image processing has a broad spectrum, thus discovering new uses for image processing such as image sharpening and restoration, remote sensing, transmission and encoding would be greatly beneficial to the society.

This study only focused on assessing the effectiveness of digital image processing in plant disease detection. Thus, this does not cover other methods on distinguishing plant diseases. Hence, future researchers can study more methods and in different fields or sectors such as the medical field.

Reference

Auti, N.A., Kadam, S., Bhapka, P., and Dhumal, S. (2016). Image Processing based Banana Leaf Recognition and Recommendation using machine learning. *International Journal for Research in Applied Science & Engineering Technology*, 6(6), 2349-7300 retrieved from https://www.ijirmps.org/papers/2018/6/429.pdf.

Barbedo, J.G.A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. SpringerPlus 2, 660 (2013). https://doi.org/10.1186/2193-1801-2-660

Bareja, B. (2019). What is Agriculture, definition of Agriculture?. Retrieved from https://www.cropsreview.com/what-is-agriculture.html

Bharate, A. A. & Shirdhonkar, M. S. (2017). A review on plant disease detection using image processing. International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 103-109, doi: 10.1109/ISS1.2017.8389326

CABI. (2020). Peronosclerospora philippinensis (Philippine downey mildew of maize). Retrieved from https://www.cabi.org/isc/datasheet/44646

Chait, J. (2020). What is agricultural production? Retrieved from https://www.thebalancesmb.com/what-is-an-agricultural-product-2538211

Crop Life International. (2014). Economic development boosted by plant science. Retrieved from https://croplife.org/news/economic-development-boosted-by-plant-science/#:~:text=Better%20crop%20productivity%20leads%20to,more%20choices%20in%20local%20markets.

Dharmasiri, D., & Jayalal, S. (2019, June 15). *Passion Fruit Disease Using Image Processing.* Retrieved from https://www.researchgate.net/publication/333803210_Passion_Fruit_Disease_Detection_using_Image_Processing

International Center for Tropical Africulture. (2019). The world needs a global system to detect and halt the spread of emerging crop diseases. Retrieved from https://croplife.org/news/economic-development-boosted-by-plant-science/#:~:text=Better%20crop%20productivity%20leads%20to,more%20choices%20in%20local%20markets.

Kaur, H., Sawhney, B.K. & Jawandha, S.K. (2018). Evaluation of plum fruit maturity by image processing techniques. *J Food Sci Technol*, 55, 3008–3015. Retrieved from https://doi.org/10.1007/s13197-018-3220-0

Khirade, S. & Patil, A.B. (2015). *Plant Disease Detection Using Image Processing.* Retrieved from: https://www.xilirprojects.com/wp-content/uploads/2019/11/khirade2015.pdf

Krippendorf, K. (1989). Content Analysis. Retrieved from: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1232&context=asc_papers

Krippendorf, K. (2004). Content analysis an introduction to its methodology (2nd ed.). Thousand Oaks, CA: Sage Publications.

Kumar, S. & Kaur, R. (2015). Plant Disease Detection using Image Processing-A Review. International Journal of Computer Applications (0975 – 8887). 124 (16). Retrieved from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.736.7605&rep=rep1&type=pdf

Masood R. & Khan, S.A. (2016). *Plant Disease Segmentation using Image Processing,* 1, 24-32. Retrieved from: https://pdfs.semanticscholar.org/48e5/582a2e9dbe993fa4a9d60e21c2c2e681c5f8.pdf

Medelayan, A. (2020). Coding qualitative data: How to code qualitative research. Retrieved from: https://getthematic.com/insights/coding-qualitative-data/

Messaouda, L. (2020). Re: What is significance of GLCM features?. Retrieved from https://www.mathworks.com/help/images/texture-analysis-using-the-gray-level-co-occurrence-matrix-glcm.html#:~:text=The%20GLCM%20functions%20characterize%20the,statistical%20measures%20from%20this%20matrix

Meyyappan, S. & Chandramouleeswaran, S. (2018). Plant Infection Detection Using Image Processing. International Journal of Modern Engineering Research (IJMER). 8. 2249-6645.

Murray, G. (2009). Philippine downy mildew of maize (Perenosclerospora philippensis) and Downy mildew of sorghum (P. sorghi). Plan Health Australia. https://www.planthealthaustralia.com.au/wp-content/uploads/2013/03/Downy-mildew-of-maize-and-sorghum-CP-2009.pdf

Pujari, J., Yakkundimath, R., & Byadgi, A. (2015). *Image Processing Based Detection of Fungal Diseases in Plants*, 4 (2015), 1802 – 1808. Retrieved from: https://pdf.sciencedirectassets.com

Selvaraj, M.G., Vergara, A., Ruiz, H. *et al.* (2019). AI-powered banana diseases and pest detection. *Plant Methods*, 15(92). Retrieved from https://doi.org/10.1186/s13007-019-0475-z

Spurgin, K.M. & Wildemuth, B.M. (2009). Content Analysis. In Applications of Social Research Method to Questions in Information and Library Science (pp. 297-307). Westport, CT: Libraries Unlimited.

Strange, R. & Scott, P. (n.d). Plant disease: a threat to global food security. Annu Rev Phytopathol. 2005; 43:83-116. doi: 10.1146/annurev.phyto.43.113004.133839. PMID: 16078878.

Tigadi, B., Sharma, B. (2016). Banana Plant Disease Detection and Grading Using Image Processing. *Department of Electronics and Communication Maratha Mandal Engineering College,* 6(6) 6512-6514, doi:10.4010/2016.1565 ISSN 2321 3361

Trochim, M.K. (2020). Unit of analysis. Retrieved from: https://conjointly.com/kb/unit-of-analysis/#:~:text=The%20unit%20of%20analysis%20is,groups

Veling, S., Kalelkar, R., Ajgaonkar, L., Mestry, N., Gawade N. (2019). Mango disease detection by using Image Processing. *International Journal for Research in Applied Science & Engineering Technology*, 7(4) 3717-3718.

Vishnoi, V.K., Kumar, K. & Kumar, B. (2020). Plant disease detection using computational intelligence and image processing. Journal of Plant Diseases and Protection. https://doi.org/10.1007/s41348-020-00368-0

Vision systems design. (2007, October 1). Retrieved from: https://www.vision-systems.com/boards-software/article/16738883/using-fuzzy-logic-in-image-processing

Ward, J. (2019). Content Analysis Methodology. Retrieved from: https://www.impactzone.co/portfolio/content-analysis-methodology/#population-sampling-unit-analysis

Zadeh, L. A. (1965, June 1). Fuzzy sets. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S001999586590241X

APPENDIX 1: Coding Procedures

Unit	Statement/Code	Themes	Frequency
Unit 1	- Increases the overall accuracy from 0.75 up to 0.81.	HIGHLY ACCURATE	1
	- Average processing time was 5.2 s	LESS TIME SPENT	1
	- Direct or indirect color variation detection - The AuC decreased to 0.7 when no color constancy algorithm is applied - The use of shades of gray color constancy algorithm assures AuC higher values	COLOR	3
	- Characteristic symptoms those are visible on some elements of an infected plant.	APPEARANCE	1
Unit 2	- The overall accuracy of the algorithm for common beans was 50% - The overall accuracy of the algorithm for cassava was 46%, - The overall accuracy of the algorithm for corn was 40%,	LESS ACCURATE	3
	- The overall accuracy of the algorithm for citrus was 56% - The overall accuracy of the algorithm for grapevines was 58% - The overall accuracy of the algorithm for passion fruit was 56% - The overall accuracy of the algorithm for soybean was 58% - The overall accuracy of the algorithm for sugarcane was 59% - The overall accuracy of the algorithm for wheat was 70%,	ACCURATE	6
	The overall accuracy of the algorithm for cotton was 76%	HIGHLY ACCURATE	1

	- Image with strong specular reflections and several light/shadow transitions - Specular lighting and shadowed and illuminated areas present simultaneously	IMAGE QUALITY	2
	- The images in the database contained leaves at various stages of maturity	PLANT AGE	1
	greenness variation among the samples the leaf's hue tends towards yellow	COLOR	1
Unit 3	- The overall system accuracy is measured to be 82%	HIGHLY ACCURATE	1
	- All features together take more time.	MORE TIME SPENT	1
	- Having small dot like structures on the fruit surface. - Spots on thefruit surface are darken, grow in size and a small crack is developed within. - Fruit surface darkens furthermore, and the cracks widen at such a scale that, the fruit splits almost entirely into two parts.	APPEARANCE	3
Unit 4	The final test accuracy has reached 89 %	HIGHLY ACCURATE	1
	Retrained - 0.547s Optimized - 0.422s Quantized - 0.406s	LESS TIME SPENT	3
Unit 5	Giving accuracy of 98.46%	HIGHLY ACCURATE	1
	Color-based segmentation is done	COLOR	1
	texture of images	IMAGE QUALITY	1

FACTORS

Themes	Frequency	Percentage
Color	5	(5/13) x 100 = 38.46 %
Appearance	4	(4/13) x 100 = 30.77 %
Image quality	3	(3/13) x 100 = 23.08 %
Plant Age	1	(1/13) x 100 = 7.69 %
TOTAL	13	100%

ACCURACY

Themes	Frequency	Percentage
Highly Accurate	5	(5/14) x 100 = 35.71%
Accurate	6	(6/14) x 100 = 42.86 %
Less Accurate	3	(3/14) x 100 = 21.43
TOTAL	14	100%

CONSUMED TIME

Theme	Frequency	Percentage
Less time spent	3	(3/4) x 100 = 75%
More time spent	1	(1/4) x 100 = 25%
TOTAL	4	100%

APPENDIX 2: Units of analysis

Unit 1

ABSTRACT

A new image recognition system based on multiple linear regressions is proposed. Particularly, there are a number of innovations in image segmentation and recognition system. In image segmentation, an improved histogram segmentation method, which can calculate threshold automatically and accurately, is proposed. Meanwhile, the regional growth method and true color image processing are combined with this system to improve the accuracy and intelligence. While creating the recognition system, multiple linear regression and image feature extraction are utilized. After evaluating the results of different image training libraries, the system is proved to have effective image recognition ability, high precision, and reliability.

RESULTS AND DISCUSSION

The authentication system has the following: server-side configuration: 2.9 GHz Intel Core i7, Intel HD Graphics 630 and Radeon Pro 560, 16 GB memory, Mac iOS Version 10.13.3; tool platform: MATLAB R2017b.

After the establishment of multiple linear regression model, the images from training libraries are placed into the multiple linear regression model; then the disease recognition system is constructed by using the least squares method. Use the images inside and outside the training libraries to test the accuracy of the system.

As the disease situation becomes gradually complex, the number of errors gradually increased, because when the disease becomes more serious, the characteristic parameters would become more complex, which would make the results instable. Compare the multiple linear regression recognition system with the traditional minimum distance method with the same training data.

The results of multiple regression system can discriminate the severity of plant diseases better. Meanwhile, as the increase of the training images, the results are more accurate, which proves the accuracy and good potential of this system.

At the same time, due to the arbitrariness of the independent and dependent variables, user can also choose different parameters to establish new regression models. For example, the random variables can be only relevant to several specific diseases, which makes it more accurate to distinguish these diseases. Besides, the selection of the characteristic parameters depends on user, so it is easy to highlight the different lesion characteristics to meet different requirement. In conclusion, the recognition system in this paper has great applicability and modifiability.

CONCLUSION

This paper improves image segmentation and disease recognition system. An improved histogram segmentation method is proposed; this method can find appropriate threshold automatically rather than manually, which is more scientific, reliable, and efficient. Meanwhile, the linear regression model can be modified easily by changing the independent and dependent variables; it has accuracy, applicability, and greater potential.

Unit 2

ABSTRACT

Disease diagnosis based on the detection of early symptoms is a usual threshold taken into account for integrated pest management strategies. Early phytosanitary treatment minimizes yield losses and increases the efficacy and efficiency of the treatments. However, the appearance of new diseases associated to new resistant crop variants complicates their early identification delaying the application of the appropriate corrective actions. The use of image based automated identification systems can leverage early detection of diseases among farmers and technicians but they perform poorly under real field conditions using mobile devices. A novel image-processing algorithm based on candidate hot-spot detection in combination with statistical inference methods is proposed to tackle disease identification in wild conditions. This work analyses the performance of early identification of three European endemic wheat diseases – septoria, rust and tan spot. The analysis was done using 7 mobile devices and more than 3500 images captured in two pilot sites in Spain and Germany during 2014, 2015 and 2016. Obtained results reveal AuC

(Area under the Receiver Operating Characteristic –ROC– Curve) metrics higher than 0.80 for all the analyzed diseases on the pilot tests under real conditions.

RESULTS AND DISCUSSIONS

The presented algorithm was developed on Python programming language and deployed as a service on a Linux based processing server. The deployed service was connected to a middleware server that managed the connections from Android and Windows-phone applications. Average processing time of the algorithm was 5.2 seconds with a standard deviation of 2.6 seconds depending on the number suspicious hot spots found.

In order to validate the results of the proposed method, a database was created using the images acquired in 2014 and 2015 (named W-2014 and W-2015) with 987 images containing rust, 2505 containing sectorial and 657 containing tan spot on a total of 3637 images. This training database was divided into training and validation sets. In order to avoid over-fitting and biasing, the dataset was divided into ten folds where the picture acquisition date was used as set divider. This means that, at each fold, the pictures belonging to the acquisition dates selected for training would be selected for the training set whereas the rest were selected for validation. At each fold, 90% of the acquisition dates were set as train and the remaining 10% were set as validation. The Area under the Receiver Operating Characteristic (ROC) Curve (AuC) was selected as the most suitable algorithm performance metric, in order to account for the class imbalance present in the dataset (in such cases, the use of accuracy is discouraged). The AuC for a binary classification problem is constructed by first sorting all the samples by the disease presence probability predicted by the model for each of them. The classification threshold value is then moved all the way from 0 to 1, and the result at each threshold value is mapped into the plot representing False Positive (x-axis) vs. True Positive rates (y-axis), and measuring the resulting area (in the 0, 1 range, higher is better) under such curve. AuC values over 0.85 were obtained for all diseases on the k-fold validation sets. It is important to remark that even diseases at very early stages were identified. An additional test was performed in order to quantify the effects of color constancy normalization on the identification metrics. The use of color constancy normalization increases the overall accuracy from 0.75 up to 0.81.

In order to validate the results under real conditions, a pilot study was set in Germany and Spain in 2016. The pilot users employed a specific developed identification smartphone application as described in section 3. Captures leaves were divided into early stage diseased plants and medium-late stage diseased plants to validate early disease detection.

We observed a small decrease on the AuC from an average 0.88 to 0.81 when moving into real field conditions. Compared with other results that fail when moving into real conditions, the developed algorithm can cope better with the variability of real light illumination, different acquisition devices and multi-located users under real use conditions. Besides this, performance is not degraded too much when dealing with early diseases.

CONCLUSION

In this proposal a general use multi-disease identification algorithm has been presented. The algorithm was validated over three different kinds of diseases (rust, septoria and tan spot) on wheat images. The algorithm has been deployed on a real smartphone application and validated under real field conditions in a pilot study located in Spain and Germany over more than 36 wheat varieties. The results on real field tests obtained AuCs (Area under the Receiver Operating Characteristic –ROC– Curve) higher than 0.8 when assuring global color constancy on the image by means of the developed algorithm succeeding on real time conditions and being able to cope with different diseases simultaneously. The AuC decreased to 0.7 when no color constancy algorithm is applied on the processing workflow. However, the use of shades of gray color constancy algorithm assures AuC values higher than 0.8. We have also validated algorithm performance when dealing with early diseases. Although there is a small degradation on the performance of the algorithm when dealing with early diseases, the algorithm can obtain almost the similar performance on early and late diseases.

The preliminary hot-spot detection and its ulterior description by color and textural descriptors allow real time performance as only the suspicious regions are trained and described by the higher level classifiers and descriptors.
The presented image processing technology provides new possibilities for the detection of weeds and diseases in earlier stages. Next steps would be focused on measuring how this early stage detection can help the user to react in time and plan for some preventive

activities, e.g. crop protection application. Being able to react in an early stage could minimize the yield losses and therefore guarantee the food security in the upcoming years.

Unit 3

ABSTRACT

Crops are being affected by uneven climatic conditions leading to decreased agricultural yield. This affects global agricultural economy. Moreover, condition becomes even worst when any disease infects the crops. Also, increasing population burdens farmers to increase yield. This is where modern agricultural techniques and systems are needed to detect and prevent the crops from being effected by different diseases. In this paper, we propose a web-based tool that helps farmers for identifying fruit disease by uploading fruit image to the system. The system has an already trained dataset of images for the pomegranate fruit. Input image given by the user undergoes several processing steps to detect the severity of disease by comparing with the trained dataset images. First the image is resized and then its features are extracted on parameters such as color, morphology, and CCV and using k-means algorithm does clustering. Next, SVM is used for classification to classify the image as infected or non-infected. An intent search technique is also provided which is very useful to find the user intension. Out of three features extracted we got best results using morphology. Experimental evaluation of the proposed approach is effective and 82% accurate to identify pomegranate disease.

RESULTS AND DISCUSSIONS

5.1. Data set Preparation:

To demonstrate the proposed system we have created the dataset of pomegranate images with the help of domain expert (professor from agricultural college). Images are captured by digital camera of 10-mega pixels. Data set contains total 610 images of disease infected (440) and normal (170) pomegranate fruits. Under infected category three subcategories such as infected stage-I (180), infected stage-II (150) and infected stage-III (110). In infected stage-I of the fruit, the disease is at the preliminary level and it is still at surface of the fruit i.e. it affects by having small dot like structures on the fruit surface. In infected stage-II, the spots on the fruit surface are darken, grow in size and a small crack is developed within. In infected stage-III, the fruit surface darkens furthermore, and the cracks widen at such a scale that, the fruit splits almost entirely into two parts.

For training purpose, all images are captured by digital camera. But for testing purpose, we have considered im- ages those are captured by digital camera as well as mobile device camera. We considered here mobile device camera because farmers also use mobile camera to capture fruit images. The image from database closely matches to query image is shown first in the result. Images are shown according to descending order of rank we got from the algorithm, which is shown, in fig. 6. We got almost similar results in case of combining all features together and morphology. The results of histogram features are poor as compare to other features. The accuracy of intent search is calculated by the correct result shown for query image.

CONCLUSION

In this paper, a Web based Image Processing dependent approach for the Bacterial Blight ("Telya") disease for Pomegranate fruit is proposed. The input image is first pre-processed, then its features are extracted on three parameters namely- color, morphology, and CCV then, training and classification of the same are done. The proposed system provides two methods for the user to check the disease infection for the input pomegranate image as- with intent search and without intent search.

Experimental results display different accuracy levels of disease detection based on the input image quality and the stages of the disease. The overall system accuracy is measured to be 82%. Thus, this system takes one step towards promoting the farmers to do the smart farming and allowing them to take decisions for a better yield by making them capable to take the necessary preventive, corrective action on their pomegranate crop. In future, the system can be improved with the new features incorporated as- training the system to detect diseases for other fruits, increase dataset size to improve the overall system performance to detect diseases more accurately.

Unit 4

ABSTRACT

Plant diseases may cause many losses to agricultural crops around the world. Therefore methods for identification of disease found in any part of the plant play a critical role in disease management. Now a days the many aspects of the crop development process uses the Advance Computing Technology that has been developed to helps the farmer to take superior decision about the crop. Evaluating & diagnosing the crop diseases is critical in the field of agriculture to increase the crop productivity. The new technological strategies are used to express the captured symptoms of cotton leaf spot images and Algorithms are used to categorize the image. In this proposed work all the images are converted into standard resolution, pre- processed and stored in the database. The classifier is being trained to achieve intelligent farming, including early Identification of diseases. The mobile captured image would be pre-processed and Edge detection algorithm is applied to the pre-processed image. Then the segmentation technique such as k-means clustering would be applied and features like colour, shape and texture features are extracted. Finally support vector machine classifier is used to identify the Diseases comparing with the trained dataset.

CONCLUSION

The goal of this work is to develop an Advance Computing system that can identify the disease affected part of a cotton leaf spot by using the image analysis technique. The producers can amend the Yield and reduce the loss. Through this proposed system the farmers' burden has been reduced and saves their life. This work consist of mobile captured image would be preprocessed. Edge detection algorithm is applied to the pre-processed image. The segmentation technique such as k-means clustering would be applied and features like color, shape and texture features are extracted. Finally support vector machine classifier is used to identify the Diseases.

Unit 5

ABSTRACT

Due to the increasing population, agriculture sectors from around the globe are challenged to increase their yield per year. However, harvests suffer from defects due to plant diseases. The current methods to mitigate spreading plant diseases are entirely dependent on the detection and recognition of such. Detection and recognition systems of plants diseases are often required huge databases for reference and/or computationally expensive systems. In this paper, we present a computationally light neural network model for detection and recognition of plant disease and implement it to mobile platform. Here, a two-step training process is used: pre-training on image net data set of wide variety of objects and retrained on dataset of specific plant diseases. The model achieved a test accuracy of 89.0%.

RESULST AND DISCUSSIONS

A. Accuracy and Cross-empty

The generated bottleneck of each image is reused every training epoch. It shows that an increase-training epoch improves the training and validation accuracy of the classifier model. The final test accuracy has reached 89% during testing with very little loss of accuracy as describe by cross-entropy values.

B. Optimization, Quantization, and Compression

Deployment of classifier model to a mobile platform requires the mobile to be optimized to reduce multiplication-addition operations. Nodes that are not needed by Tensor flow libraries for mobile applications are removed during the optimization, which reduces complexity of calculations by merging batch norm operations into the architecture or structure of the classifier model.

C. Testing the Classifier

The classifier model which is Mobile net-based architecture developed on this research can classify five major diseases whose symptoms are appearing on the leaves of wide variety of plants. In terms of probabilities, the optimized, quantified, and compressed models are not significantly different. This means that doing quantization prior to compression reduces the size but does not degrade the performance of the classifier. We are also developing a system that uses the same model for processing video feeds of the farm for detection of the possible incidents of crop defects.

CONCLUSIONS

It was presented that even the relatively smaller data set, classifier based on Mobile net architecture can effectively classify plant disease. The classifier model achieved a relatively fair accuracy even when it is implemented on mobile platform only. All interferential processes can be autonomously done in mobile devices opening more opportunities for other applications. To improve the classifier more, expanding the data set to other diseases can be done. For future development, we are testing the model on smaller machines such as raspberry pi.