ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

Secure Crop Disease segmentation and Detection image search on Hadoop Environment

Ajit Gund, Sapna Jagtap, Omkar Sargar, Pratik Kumbhar, Prof. J. S Patil

UG Student, Department of Computer Engineering Jayawantrao Sawant College of Engineering, Pune, India

HOD, Department of Computer Engineering Jayawantrao Sawant College of Engineering, Pune, India

ABSTRACT: Sugarcane is one of the most important crops in India. Indian sugar industry is the second largest agro based industry, next only to the textiles. But, being a long durational crop, sugarcane is prone to the number of disease caused by pathogens viz. fungi, bacteria, viruses and phytoplasmas like organisms. Image processing techniques have been proved to be changing the scenario of agriculture in India with a number of research and applications like automatic disease detection, drone based pesticides and fertilizer dispensing, estimation of yield, vegetative growth, fruit sorting etc. This abstract is carried out to study the effectiveness of Image Processing and computer vision techniques for detection of disease in sugarcane plants using HDFS base MapReduce framework by observing the leaves. Few major diseases in sugarcane plant like red rot, mosaic and leaf scald have been studied and detection algorithm for the same has been implemented in this research work.

KEYWORDS: Mapreduce, computer vision, image processing, crop disease analysis

I. INTRODUCTION

Plant disease detection is basically a process which is used for identification and detection of plant diseases. The plant can be affected by a range of diseases which may be present on leaf, stem and root of plants. The diseases on any of the part of plant directly affect the quality and quantity of crop production. Indian farmers face a range of problems such as less crop production due to the diseases. So it is necessary to detect the diseases earlier as possible to reduce the loss in crop production. Plant disease detection methods using image processing tools are affective to use for identifying and classifying the disease. Automated disease detection systems are accurate and efficient for the purpose of disease detection. Mainly the plant disease are categorized into three categories: bacterial, fungal and viral diseases. The automated systems designed using image processing tools makes use of color properties and other features of leaf image to detect the disease at early stage. The quality and quantity of crops can be increased by detecting the diseases earlier with the help of automated disease detection system. The proposed work illustrates the sugarcane disease detection using HDFS base MapReduce framework.

II. LITERATURE SURVEY

Dr. E. L. Lydia, Dr. A. K. Mohan et. al, [1] presented in the Processing Image Files Using Sequence File In Hadoop, The proposed system presents MapReduce as a distributed data processing model utilizing open source Hadoop framework for work huge volume of data. The expansive volume of data in the advanced world, especially multimedia data, makes new requirement for processing and storage. As an open source distributed computational framework, Hadoop takes into consideration processing a lot of images on an unbounded arrangement of computing nodes by giving fundamental foundations. We have lots and lots of small images files and need to remove duplicate files from the available data. As most binary formats—particularly those that are compressed or encrypted—cannot be split and must be read as a single linear stream of data. Using such files as input to a MapReduce job means that a single mapper will be used to process the entire file, causing a potentially large performance hit. The proposed system proposes

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

splitable format such as SequenceFile and uses MD5 algorithm to improve the performance of image processing. This gives a survey of distributed processing procedures and the programming models. Moreover a couple works are considered which have been done of late using Hadoop open source system. Hadoop and its processing model are as of late surrounded and like some other new developments may have its own issues, for instance, nonappearance of nature of the lion's offer of IT society with it, nonattendance of enough ace forces, and undesirable flaws and issues on account of its peculiarity.

Vemula, Sridhar, and Christopher Crick. [2]Presented in the Hadoop Image Processing Framework, With the rapid growth of social media, the number of images being uploaded to the internet is exploding. Massive quantities of images are shared through multi-platform services such as Snapchat, Instagram, Facebook and WhatsApp; recent studies estimate that over 1.8 billion photos are uploaded every day. However, for the most part, applications that make use of this vast data have yet to emerge. Most current image processing applications, designed for small-scale, local computation, do not scale well to web-sized problems with their large requirements for computational resources and storage. The emergence of processing frameworks such as the Hadoop MapReduce platform addresses the problem of providing a system for computationally intensive data processing and distributed storage. However, to learn the technical complexities of developing useful applications using Hadoop requires a large investment of time and experience on the part of the developer. As such, the pool of researchers and programmers with the varied skills to develop applications that can use large sets of images has been limited. To address this we have developed the Hadoop Image Processing Framework, which provides a Hadoop-based library to support large-scale image processing. The main aim of the framework is to allow developers of image processing applications to leverage the Hadoop MapReduce framework without having to master its technical details and introduce an additional source of complexity and error into their programs.

Ms. G.Shobana, Ms. M. Sugunaet. Al [3], presented in the Identification of Crop Disease by Predictive Analysis in Hadoop Environment, t-Agriculture is one of the important sources for survival of life. For the past few decades, new technologies are used to improve the better productivity of the crops. Many researchers have developed monitoring and automation system for different functionalities of farming. To yield high profit in agriculture, growth of crops has to be monitored by IOT enabled devices as result massive unstructured data are generated in regular interval of time. In the proposed system, predictive analyses in Hadoop environment are used to identify the disease of crop and growth of the crops. Based on prediction, type of pesticides and the nutrition required for fields are identified. The people can easily make decision from historical data with agricultural analysis to yield high productivity.provides a path in diagnosing the disease by the identification of presence of disease in plant by predictive analysis. The approach is well suitable for different crop types and diseases. The result will be expected to improve considerably with more training data. Early prediction disease in the agriculture field helps the farmer for identifying the severity of the detected disease as the result it will increase the productivity of the crops.

Izay A. et. al, [4]presented in the Digital Image Processing for Detecting and Classifying Plant Diseases, Agriculture is the backbone of human sustenance in this world. With growing population, there is need for increased productivity in agriculture to be able to meet the demands. Diseases can occur on any part of a plant, but in proposed system only the symptoms in the fruits of a plant is considered using segmentation algorithm and edge/ sizing detectors. We also looked at image processing using fuzzy logic controller. The system was designed using object oriented analysis and design methodology. It was implemented using MySQL for the database, and PHP programming language. This system will be of great benefit to farmers and will encourage them in investing their resources since crop diseases can be detected and eliminated early.Digital image processing technique helps to further increase the detection of diseases found in the fruit of plants using segmentation algorithm and edge/ sizing detectors. The system could be used for the easy analysis of infected fruits on a farm land since it is easier to understand an image than a comparable table of numbers.This system applies a comparison testing sample. It compares the analysis of the new image to the specification sample in the system.. The main involvement in this new system is the presence of a new architecture for processing images using Fuzzy logic controller, which enables the combination between image understanding and the detection of diseases in the fruits. The accuracy of classification of plant diseases varies dependingon the image acquisition. Thus we can

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

conclude that image processing is an effective tool that can be applied in the agriculture domain with great accuracy for analysis of agronomic parameters.

ArifaKhan,Manmohan Singh Yadav et. al,[5] presented in the Image Processing Based Disease Detection for Sugarcane Leaves, Sugarcane is one of the most important crops in India. Indian sugar industry is the second largest agro based industry, next only to the textiles. But, being a long durational crop, sugarcane is prone to the number of disease caused by pathogens viz. fungi, bacteria, viruses and phytoplasmas like organisms. Image processing techniques have been proved to be changing the scenario of agriculture in India with a number of research and applications like automatic disease detection, drone based pesticides and fertilizer dispensing, estimation of yield, vegetative growth, fruit sorting etc. This research is carried out to study the effectiveness of Image Processing and computer vision techniques for detection of disease in sugarcane plants by observing the leaves. Few major diseases in sugarcane plant like red rot, mosaic and leaf scald have been studied and detection algorithm for the same has been implemented in this research work. A computer vision based technique to detect leaf disease in sugarcane plant has been implemented. Image processing is nowadays for a number of applications in agriculture. The combination of feature extractions like color, size, and shape with different classifiers has added accuracy to these applications. Red Rot disease, leaf scald disease, and mosaic disease are three most common diseases occurring in sugarcane plant in India. A thorough study of the causes and symptoms of this diseases is presented in this research paper. Image processing techniques along with k-means classification have been used to classify the leaf dataset, according to the disease. The leaf image is compared with this database to detect the type of disease existing in the plant.

Zhao J., Zhao X. and Liu Y.[6],presented a study thatA Method for Detection and Classification of Glass Defects in Low Resolution Images,A quantity of diseases does not have any detectable symptoms associated and display very late. In today's modern era, agriculture has become one of the most important part of the population. Energy is the fundament part of the global warming and major source of plants also. Number of diseases is there which badly affected ecological, economical and society losses. To detect plant pathologies there are many ways. A quantity of diseases do not have any detectable symptoms associated and display very late. In those cases, powerful microscopes, is necessary for sophisticated analysis. In other cases, the signs can only be detected in parts of the electromagnetic spectrum that are not visible to humans.

Al-Bashish, D., M. Braik, and S.Bani-Ahmad [7], Detection and Classification of Leaf Diseases Using KMeans based Segmentation and Neural Networks Based Classification.A common approach in this case is the utilize of remote sensing techniques that explore multi and hyper spectral image captures. The methods that adopt this approach often employ digital image processing tools to achieve their goals. The naked eye observation method is generally used to decide diseases severity in the production practice but results are subjective and it is not possible to measure the disease extent precisely. Image processing technology in the agricultural research has made significant development.

Nasreen N. et al. [8], presented a study that focus on theapplication of texture analysis in microscopic images of plantstem cross sections. The texture features are extracted usingGLCM method in this project. The accuracy and precisionwere obtained as 97%. The SVM classifier Performance maybe improved by giving a large training set. In conclusion, this automatic classification system has efficiency of 97% and isreliable and fast.

Dhaware C.G. et al. [9], proposed a method that focus onautomatic leaf unhealthiness classification which establish onleaf image processing. The system approach will give adviceto the farmer with minimum efforts. The farmer most effectiverequire to seize the image of the plant leaf the usage of mobilecamera and forward it to the DSS, without any additionalinputs.

Francis J. et al. [10], presented an algorithm that producesbetter results and healthy and unhealthy plants can bedifferentiated with the help of this algorithm. This algorithmhelps in identifying the presence of diseases by observing thevisual symptoms seen on the leaves of the plant. With the aidof imaging technology the plant disease detection systemsautomatically detect the symptoms that appear on the leavesand stem of a plant and helps in cultivating healthy plants in afarm. These systems monitor the plant such as leaves and stemand any variation observed

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

from its characteristic features,variation will be automatically identified and also will beinformed to the user. This paper provides an evaluative studyon the existing disease detection systems in plants.

III. PROPOSED SYSTEM DESIGN

The proposed system follows the complete image processing strategy in HDFS framework. Working flow of proposed approach includes following steps:
1. Take RGB image
2. Image Color transformation of RGB to Gray scale
3. Image segmentation
4. Feature extractions
5. Statistical analyses
6. According to analysis matching of feature extraction are done.

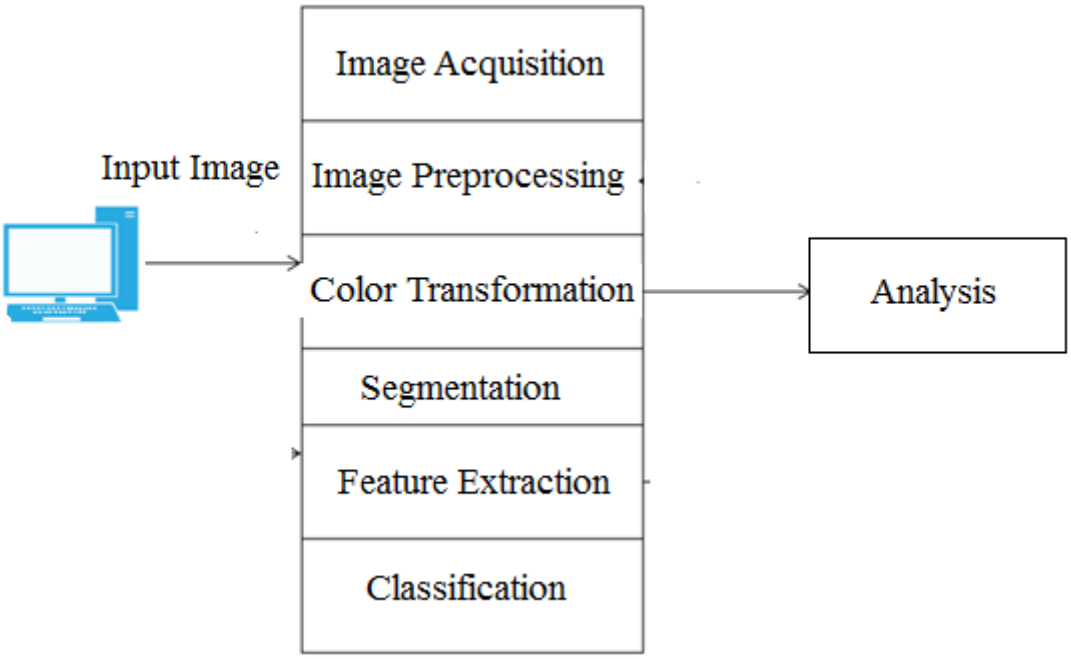

Figure 1 : Proposed System Architecture

IV. SYSTEM MODULES

Image Acquisition
Sugarcane diseased leaf images are taken in controlled environment and are stored in the JPEG format Disease detection start with taking input image from digital camera, those are in RGB format. Better quality resolutions are used for image-analysis that images are in the format such as TIF, JPEG, PNG, BMP etc, the same time system use automatic image resize according to training as well as testing feature policy like [height*width] etc.

B. RGB to Gray scale Conversion
In this method while converting the true color images RGB to gray scale by eliminating the hue and saturation information.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

C. Image Preprocessing

By using image pre-processing reject unwanted part of data from the image such as filter the noise, image processing feature include the color, size and texture of image.

D. Clustering

The proposed system generate clusters using C-Means clustering Algorithm (CMA). Suppose a 3 dimensional dataset of points x = (xR , xG , xB) contains the red, green and blue component of a color pixel x. This space contains the color histogram of an image and is called the color space. After a linear transformation, another color space can be obtained, which can be, e.g., more adapted to the human visual system. In this paper, RGB space is used. The result of input image segmentation for a plant disease detection system is to preserve only the infected area in the output image for detection purpose

D. Image Segmentation

Image segmentation is the important step to separate the different regions with special significance in the image, these regions do not intersect each other and each region should meet consistency conditions in specific regions. The division of an image into its constituent objects or
regions is called Segmentation. The level to which the subdivision is carried out depends on the problem being solved. That is, segmentation should stop when the objects of interest or the regions of interest in an application have been isolated. Several techniques have to be applied to
achieve a desirable level of segmentation sufficient to carry out recognition.
The system used threshold base segmentation to generate the different segment of image. Thresholding is probably the most frequently used technique to segment an image. The thresholding operation is a grey value remapping operation G defined by:

$$G = P[x] < Th \rightarrow 0 \text{ else } 1$$

where v represents a grey value, and t is the threshold value. Thresholding maps a grey-valued image to a binary image. After the thresholding operation, the image has been segmented into two segments, identified by the pixel values 0 and 1 respectively.
The result of input image segmentation for a plant disease detection system is to preserve only the infected area in the output image for detection purpose. However, due to the diversity of disease types, the plants are growth in different environmental conditions. To predicate the final result it may be hard to detect and problem of accuracy. There is information need for prediction of disease is physical location of plant in environment and its infected area of disease according to this information prediction is done. Process for image segmentation based on color of image.

V. RESULTS AND DISCUSSION

The proposed system has describes image disease detection using c-menas classification approach. Basically system first executes initily image processing paradigm and then evaluate the performance of system. Below figure 2 shows the time required for single image processing with various function in stand alone system . Figure 3 illustrates time required for single image processing with various function in stand alone system, and figure 4 shows disease classification detection accuracy using c means algorithm.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

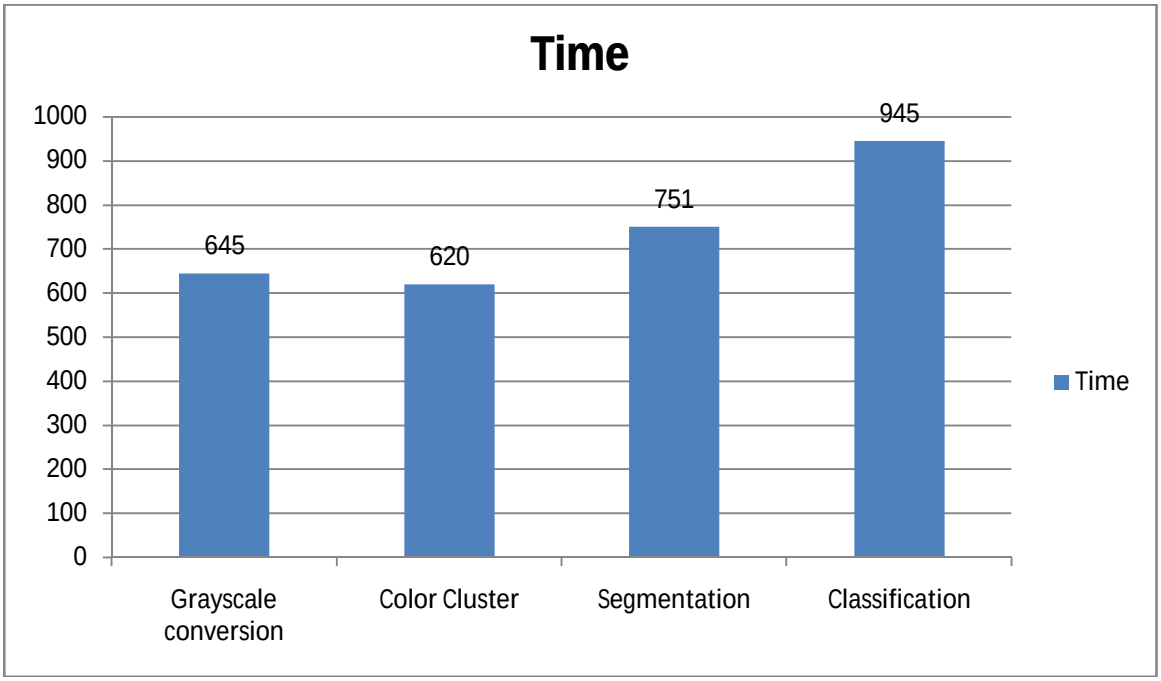

Figure 2 : Time required for single image processing with various function in stand alone system

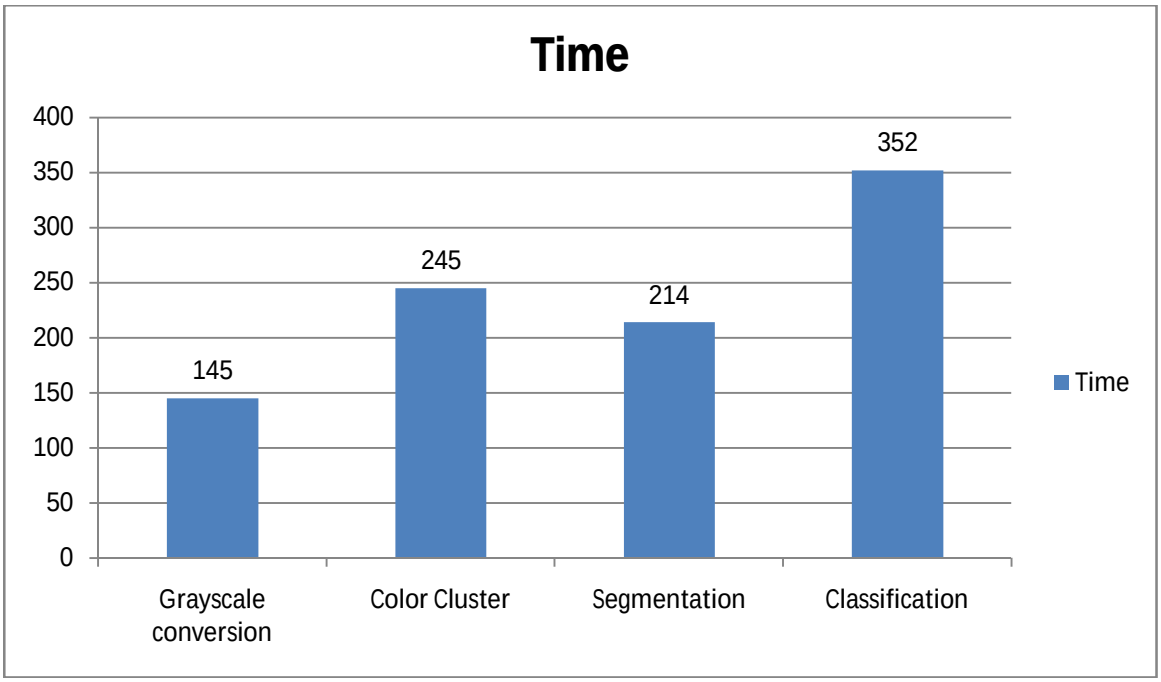

Figure 3 : Time required for single image processing with various function in single node HDFS

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

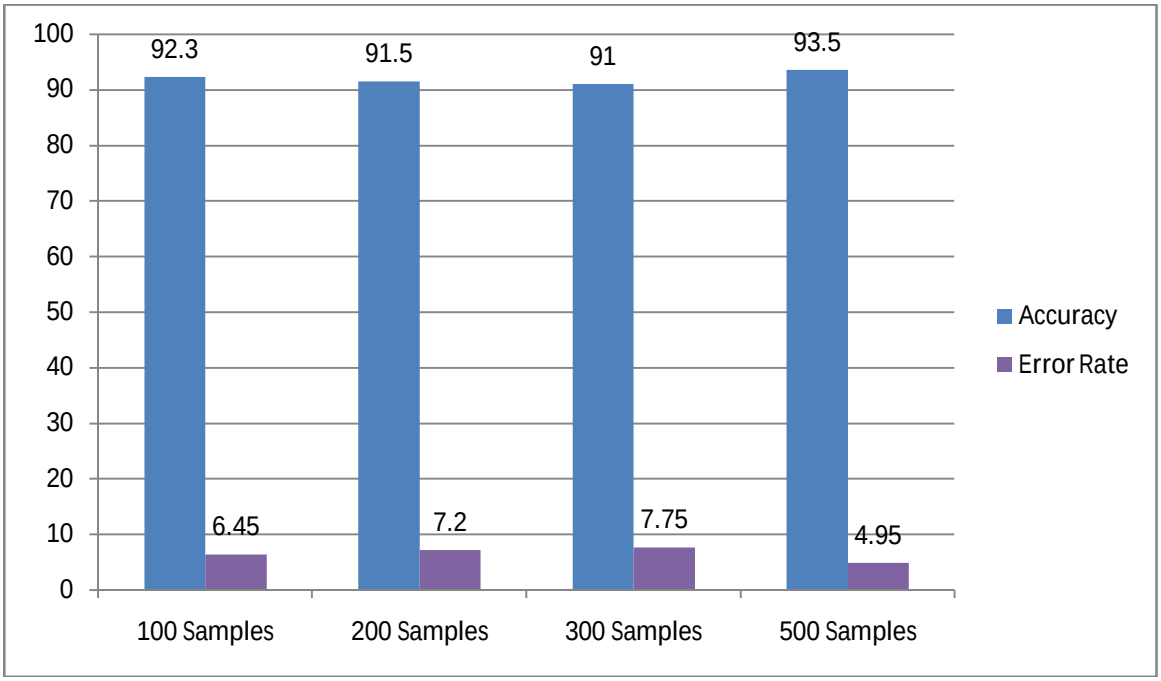

Figure 4 : disease classification detection accuracy using c means algorithm

VI. CONCLUSION

The early detection of diseases on plants is much required as a very small number of diseased crops can spread the infection to the whole crop in the field and thus affects further storage and sales of agriculture products. This system can be used for multiple purposes where agro based industries can launch their products and acquire feedbacks. Agricultural institutes can explore new inventions and technologies for farmers. This system will help to increase income from overall crop production. This system will be an interactive one which helps to overcome communication barrier through interaction between farmers and experts. The system reduces occurrences of diseases in crops and thus prevents future economical losses.

FUTURE WORK

To implement and enhance the system on android as well as some remaining mobile platform, to enhance usability to end user.

REFERENCES

1] Dr. E. L. Lydia, Dr. A. K. Mohan et. al, Processing Image Files Using Sequence File In Hadoop, IJESRT 2016
2] Vemula, Sridhar,et. al "Hadoop Image Processing Framework." BigData Congress. 2015.
3] Ms. G.Shobana, Ms. M.Sugunaet. al, Identification of Crop Disease by Predictive Analysis in Hadoop Environment, International Journal of Pure and Applied Mathematics 2018
4] Izay A. et. al, Digital Image Processing for Detecting and Classifying Plant Diseases,Circulation in Computer Science 2017
5] ArifaKhan,Manmohan Singh Yadav et. al, Image Processing Based Disease Detection for Sugarcane Leaves,IJARIIT 2017
[6] Zhao J., Zhao X. and Liu Y., (2011), "A Method for Detection and Classification of Glass Defects in Low Resolution Images," Sixth International Conference on Image and Graphics, , pp. 642-647.
[7] Al-Bashish, D., M. Braik, and S.Bani-Ahmad, (2011), "Detection and Classification of Leaf Diseases Using KMeans based Segmentation and Neural Networks Based Classification", Information Technology Journal, Vol. 10, Issue. 2, pp. 267-275.
[8] Nasreen N., Kumar C. and Nabeel A., (2016), "Automatic Classification of Monocots and Dicots Using SVM Classifier", International Conference on Information Science (ICIS), IEEE, pp. 28-31.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

[9] Dhaware C.G. and Wanjale K.H., (2017), "A Modern Approach for Plant Leaf Disease Classification which Depends on Leaf Image Processing", International Conference on Computer Communication and Informatics (ICCCI), IEEE, pp. 1-4.
[10] Francis J. and Anoop B.K., (2016), "Identification of Leaf Diseases in Pepper Plants Using Soft Computing Techniques", International Conference on Emerging Devices and Smart Systems (ICEDSS), IEEE, pp. 168-173.