Identification and Classification of Rice Plant Diseases Using Cluster Based Thresholding Algorithm-A Review

Amanpreet Kaur [1], Vijay Bhardwaj[2]

Research scholar [1], Associate professor[2]

Guru Kashi University, Talwandi Sabo, Punjab, India

Abstract: In leaves recognition analysis, plenty has been done regarding general options extraction or recognition between completely different categories of objects. Just in case of specific domain recognition, taking under consideration the distinctive characteristics that belong to the current class, improves the performance of the system. Despite the high technical side of this work, addressing leaves provides a biological connotation. Plant diseases have become a quandary because it will cause vital reduction in each quality and amount of agricultural product. In Republic of India seventieth of the population rely upon agriculture. Farmers have wide selection of diversity to pick out appropriate Fruit and Vegetable crops. During this analysis work I actually have reviewed completely different analysis papers of various authors. Each have baby-faced completely different issues, out of that some issues Measure baby-faced by me i.e. Plant leaf classification, the overlapping of the plant diseases and another drawback to spot the ripening stage of the rice crop and fruits (like guava and bananas etc.).

Keywords: Plant, Leaf, Diseases, Image, Agriculture etc.

I.   INTRODUCTION

Plant diseases have became a quandary because it will cause important reduction in each quality and amount of agricultural product. In India seventieth of the population rely upon agriculture. Farmers have big selection of diversity to pick appropriate Fruit and Vegetable crops. However, the cultivation of those crops for optimum yield and quality manufacture is very technical. [1] It is improved by the help of technological support. The management of perennial fruit crops needs shut observance particularly for the management of diseases that may have an effect on production considerably and afterwards the postharvest life. In [2] the authors have worked on the event of ways for the automated classification of leaf diseases supported high resolution multispectral and stereo pictures. Leaves of sugar beet are used for evaluating their approach. Sugar beet leaves could be infected by many diseases, like rusts (Uromyces betae), mildew (Erysiphe betae). [2] Unwellness is caused by infectious agent that is any agent inflicting unwellness. In most of the cases pests or diseases are seen on the leaves or stems of the plant. Thus identification of plants, leaves, stems and looking for the tormenter or diseases, share of the tormenter or unwellness

incidence, symptoms of the tormenter or unwellness attack, plays a key role in triple-crown cultivation of crops. It's found that diseases cause significant crop losses amounting to many billion bucks annually. It is not solely tremendous quantity of labor however additionally suffers from 2 major issues: excessive time interval and subjectivity rising from completely different people .Hence to conduct high output experiments, plant man of science would like economical laptop software package to mechanically extract and analyze important content. Here image process plays necessary role. [3]

II.   BACKGROUND

Since recent decades, digital image process, image analysis and machine vision are sharply developed, and that they became a really vital a part of computer science and the interface between human and machine grounded theory and applied technology. These technologies are applied wide in trade and drugs, however seldom in realm associated with agriculture or natural habitats.[5]
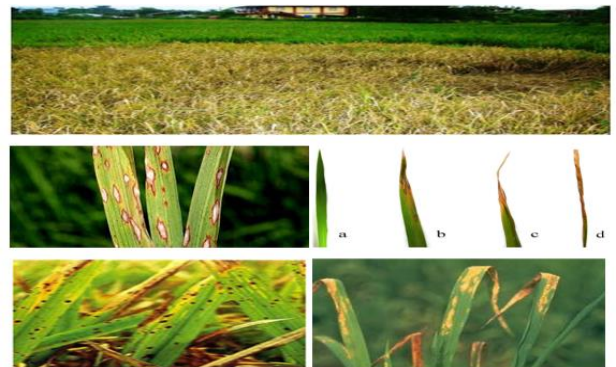

Figure.1. various images of infected rice leave

Despite the importance of the topic of characteristic plant diseases victimisation digital image process, and though this has been studied for a minimum of thirty years, the advances achieved appear to be somewhat timid. Some facts result in this conclusion: strategies square measure too specific. The best technique would be able to establish any quite plant. Evidently, this can be impossible given the present technological level. However, several of the strategies that square measure being projected not solely square measure able to subsume only 1 species of plant, however those plants got to be at an explicit growth stage so as to the algorithmic

program to be effective. That is acceptable if the plant is in this specific stage, however it's terribly limiting otherwise. several of the researchers don't state this sort of data expressly, however if their coaching and take a look at sets embrace solely pictures of an explicit growth stage, that is commonly the case, the validity of the results can not be extended to alternative stages.[6] Operation conditions square measure too strict. Several pictures accustomed develop new strategies square measure collected underneath terribly strict conditions of lighting, angle of capture, distance between object and capture device, among others. This can be a typical apply and is absolutely acceptable within the early stages of analysis. However, in most globe applications, those conditions square measure nearly not possible to be enforced, particularly if the analysis is predicted to be dole out in a very non-destructive method. Thus, it's a tangle that a lot of studies ne'er get to the purpose of testing and upgrading the strategy to subsume additional realistic conditions, as a result of this limits their scope greatly. Lack of technical data regarding additional refined technical tools. The best resolution for a tangle is typically the desirable one. Within the case of image process, some issues are often solved by victimisation solely morphological mathematical operations, that square measure simple to implement and perceive. [7] But, additional advanced issues typically demand additional refined approaches. Techniques like neural networks, genetic algorithms and support vector machines are often terribly powerful if properly applied. Sadly, that's typically not the case. In several cases, it looks that the utilization of these techniques is in additional demand three within the scientific community than in their technical appropriateness with regard to the matter at hand. As a result, issues like over fitting, overtraining, small sample sets, sample sets with low representativeness, bias, among others, appear to be a widespread plague. Those issues, though simply recognisable by a knowledgeable individual on the subject, appear to travel wide unnoted by the authors, in all probability owing to the shortage of data regarding the tools they're using. The result's a full cluster of technically imperfect solutions. In recent times, laptop vision methodologies and pattern recognition techniques are applied towards machine-driven procedures of plant recognition. Digital image process is that the use of the algorithms and procedures for operations like image sweetening, compression, image analysis, mapping, geo-referencing, etc [8]. The influence and impact of digital pictures on fashionable society is tremendous and is taken into account as a vital part in style of application areas as well as pattern recognition, laptop vision, industrial automation and tending industries. One amongst the foremost common strategies in leaf feature extraction is predicated on morphological options of leaf. Some straightforward geometrical options square measure ratio, angularity, convexity, sphericity, type issue etc. One will simply transfer the leaf image to a laptop and a laptop will extract options mechanically in image process techniques. Some systems use descriptions employed by botanists. However it's difficult to extract and transfer those options to a laptop mechanically [9].

The main purpose of this program is to use MATLAB resources. Indeed, there square measure many benefits of mixing MATLAB with the leaf recognition program. The result proves this methodology to be an easy associate degreed

an economical try. Future sections can discuss a lot of on image preprocessing and acquisition which incorporates the image preprocessing and sweetening, bar chart exploit, edge detection. any on sections introduces texture analysis and high frequency feature extraction of a leaf pictures to classify leaf pictures i.e. constant quantity calculations and so followed by results[5][8].

III.   USAGE

The progressive of leaf/plant/tree recognition is today employed by phytologist. However, the program developed in this case is restricted for one task and is of no use in additional general applications. Actually, these straightforward techniques focus solely on a number of options (like color), and aren't economical during a lot of general purpose. Consequently a lot of general image classification strategies square measure used, as a result of it's a widespread topic, and there square measure lots of well-known options (such as color bar chart, SIFT (Scale invariant feature transform), HOG (Histogram of homeward gradient), form descriptors, OCR (Optical Character Recognition)[9]. so as to specialize in the most structure of the program, the MATLAB implementation, the information retrieving and specific feature creation, we are going to take good thing about the designed functions out there in MATLAB for Digital Image process. the ultimate program as seen in Figure one provides a segmentation rule. additionally, like in most of the image recognition programs, a information of plant or leaf image has got to be done, also as a learning methodology to extract the options for the information, and an identical methodology to retrieve the most effective match from the information. many further little programs are enforced to collect data for results. a singular set of options square measure extracted from the leaves by slicing across the key axis and parallel to the minor axis[10].

```
┌─────────────────────────┐
│   Image Preprocessing   │
└─────────────────────────┘
              │
              ▼                    Leaf Image
┌─────────────────────────┐
│   Feature Extraction    │
└─────────────────────────┘
              │
              ▼                    7 Base Features
┌─────────────────────────┐
│      Classification     │
└─────────────────────────┘
              │
              ▼
         Plant Type
```
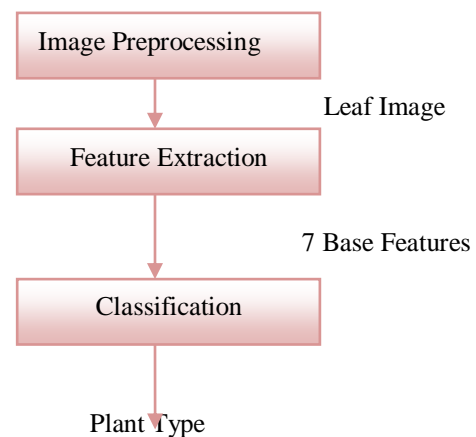
Figure 2. Main stages of the system.

Then the feature points square measure normalized by taking the quantitative relation of the slice lengths and leaf lengths (major axis). These options square measure used as inputs to the SVM. The SVM Classifier was trained with few straightforward leaves from a unique plant species. Input file preparation: Once the feature extraction was complete, 2 files were obtained. They were: (1) coaching texture feature

---

knowledge and (2) take a look at texture feature knowledge Classification exploitation Support Vector Machine supported Hyper plane classifier: A code routine was written in MATLAB that might soak up .mat files representing the coaching and take a look at knowledge, train the classifier exploitation the train files so use the take a look at file to perform the classification task on the take a look at knowledge. Consequently, a MATLAB routine would load all the knowledge the info the information files (training and take a look at data files) and build modifications to the information in line with the projected model chosen[11][12].

IV.   K-NEAREST-NEIGHBOR CLASSIFICATION

The well-known k-NN approach to classification has verified victorious in several applications. During this methodology, we have a tendency to live the gap from a check set item to every of the coaching set things, noting the k coaching set things that are nearest. We have a tendency to then categoryify the check set item by whichever class is most typical among that k "nearest neighbors," holding every neighbor "vote." (In case of ties, we've chosen to incorporate all coaching set things no farther away than the kth nearest neighbor; therefore during this case there'll be over k voters.) Varieties of investigators have thought-about the question of however best to live distance: approaches have enclosed world metrics [5], native metrics [6], metrics that are specific to the matter [6] so on. out and away the foremost common metric, though, has been Euclidian distance, beneath that the gap between 2 points x_r and x_s, say, is given by the root of the (possibly weighted) total of the square distances over every co-ordinate. Though generalizations are potential, we have a tendency to use the easy form:

$$d(\mathbf{x}_r, \mathbf{x}_s) = \left[\sum_{i=1}^{p} c_i (x_{ri} - x_{si})^2 \right]^{1/2} .$$

In normal Euclidian distance, the weights c_i; $i = one;::::;p$ area unit all adequate 1. However, expertise suggests that 2 connected steps will improve classification accuracy. First, we would expect a number of the measurements to be orthogonal to the matter. Naturally we have a tendency to hope to be able to offer weights of zero to those orthogonal columns. This echoes, of course, the variable choice downside that seems in virtually each applied mathematics model. A second, relevant variable is also measurement similar quantities on quite totally different scales. Beneath these circumstances, it appears obvious that reducing every of the variables to a typical scale could facilitate k-NN classification by preventing one in every of the measurements from dominating all the others. a 3rd downside that has to be tackled is that of choosing the most effective price of k, the quantity of neighbors to be considered[15]. In our approach, we have a tendency to attack the alternatives of k and of that variable to incorporate by employing a stepwise approach. Our implementation permits either forward or backward selection; for reasons of speed and parsimony we have a tendency to typically use the previous. during this theme, we have a tendency to begin with each variable out of the model, and a vector of attainable values of k. Currently, this set isn't chosen by relevancy the data; we

have a tendency to just use the set 1; 3; 5;::::; thirty one since this includes a "reasonably large" vary. we have a tendency to use leave-one-out cross-validation to estimate the misclassification rate of the classifier for every selection of k. That is, every part of the coaching set is classed by all the others, victimization the present set of variables within the model and also the entire vector of k's. Of course, at the terribly starting of this method once each variable is "out" of the model, each coaching set item is equal from each check set item, and notwithstanding k, each coaching item gets to vote[14]. this offers North American country the so-called" name classifier," within which each check set item is solely given the foremost frequent coaching set classification. Then one in every of the variables is additional to the model and a brand new set of misclassification rates, one for every k, is computed. This is often in dire straits every variable successively. At the top of this method we elect the mixture of k and additional variable that produces all-time low misclassification rate. If no addition produces associate improvement then the method is diminished; the present set of variables and also the best k area unit used. If the addition of a variable produces a misclassification rate strictly higher than that of the present set, then that variable is additional to the present set and also the method continues[16][17]. Our approach, like different "greedy" algorithms, is cheap however not sure to turn out associate best set of variables. Since we have a tendency to need strict improvement at every stage we have a tendency to expect our routine to be proof against the presence of "noise" variables, and that we have seen this in associate example (see "noise resistance" below). Finally, we have a tendency to perform this whole stepwise routine twice: once with the information in its original kind and once with every column scaled to possess variance adequate one. (We have conjointly checked out scaling by median absolute deviation however in our examples we've found it makes very little distinction.) in fact we elect whichever of the scaled and un-scaled results has the lower misclassification rate. So once we point out k-NN classification, we have a tendency to area unit pertaining to a method that will all 3 of (i) variable selection; (ii) selection of k; and(iii) selection of victimization scaled or un-scaled information.

V.   LITERATURE SURVEY

In this work totally different completely different authors studied different analysis algorithms and strategies, some authors like Anand R et. al. [2016] have studied associate degree approach for careful detection of diseases. The goal of projected work was to diagnose the illness of aubergine leaf victimization image process and artificial neural techniques. The diseases on the aubergine were crucial issue that makes the sharp decrease within the production of aubergine. The study of interest was the leaf instead of whole aubergine plant as a result of regarding 85-95 maximize diseases occurred on the aubergine leaf like, microorganism Wilt, Cercospora Leaf Spot, mosaic virus (TMV). The methodology to find aubergine plant disease during this work includes Kmeans cluster rule for segmentation and Neural-network for classification. The projected detection model based mostly artificial neural networks were terribly effective in recognizing leaf diseases.[1] Azeil Louisse Codizar et. al. [2016] have

studied the quantity of famous and unknown plant species will increase as time goes by. analysis on plant species may be more advanced if there was a fast and correct system which will establish plants and hasten the classification method. This method wasn't solely facilitate in fast plant classification, however was additionally enable folks that wasn't morphological specialists to conduct their own studies. Leaves were associate degree application designed to classify completely different plant species supported the leaf's form and venation. This method uses completely different image process and machine learning techniques as well as centroid-radii, moment invariableness, clever edge detection, morphological operations, image distinction and artificial neural networks. [2] Chitra Anil Dhawale et.al.[2016] have studied Citrus trees was the nutrition food for humans moreover as animals. However, owing to the unsure weather conditions, it'll liable to completely different pathological disorders thanks to the organic process deficiency. In Vidarbha regions, citrus suffer from bound deficiencies of essential parts, within which plants gain from the soil. The segmentation of illness symptoms in citrus leaf pictures may be a valuable aid for the detection of organic process deficiencies and disorders. During this analysis, completely different digital image segmentation techniques had been utilized that analysis the regions of the citrus leaf caused by some diseases like spots and wavy structure. This paper investigates the abnormalities in citrus leaf caused by the diseases by the segmentation methodologies. The organic process deficiency of the citrus was directly mirrored on its plants. If any temporal a part of the symptom is disconnected then, it may be metameric to its original half victimization the cluster technique. The illness spots were known by the cluster wherever the wavy disorders square measure metameric by the brandy operator. This projected system analyzes the disorder of the citrus by analyzing its leaf victimization the segmentation technique with the integrated use of cluster feature and therefore the brandy operator. [3] Davoud Ashourloo et.al. [2016] have studied the spectra of the infected and non infected leaves in numerous illness symptoms were measured employing a non imaging spectro radiometer within the magnetic force region of 350 to 2500 nm. So as to provide a ground truth dataset, we tend to utilized photos of a photographic camera to calculate the illness severity and illness symptoms fractions. Then, completely different sample sizes of collected datasets were utilised to coach every methodology. PLSR showed constant of determination (R2) values of zero.98 (root mean sq. error (RMSE) = zero.6) and 0.92 (RMSE = zero.11) at leaf and cover, severally.[4] Harshal Waghmare et.al.[2016] have studied associate degree identification of disease through the leaf texture analysis and pattern recognition. during this work they centered on Grapes plant disease detection system. The system takes one leaf of a plant as associate degree input and segmentation was performed once background removal. The metameric leaf image was then analyzed through high pass filter to find the morbid a part of the leaf. The metameric leaf texture was retrieved victimization distinctive form based mostly texture feature. form based mostly options was regionally invariant in nature and so provides an honest texture model. the feel of each freelance illness are going to be completely different. The extracted texture pattern was then classified victimization

multiclass SVM. The work classifies specialise in major diseases usually discovered in Grapes plant that square measure mildew & plant disease. The projected approach avails recommendation of agricultural specialists simply to farmers with the accuracy of ninety six.6%.[5] John Barajas et.al.[2016] have studied Untreated surface water and groundwater was the first sources of potable in several elements of the planet, significantly in low-income regions. He investigated the surface assimilation of Moringa proteins (MOP) onto rice husk ash (RHA) and assessed whether or not the proteins maintained their medication properties. Their results recommend a possible sensible answer to the unsought co-release of soluble organic matter in Moringa seeds. [6] John Raymond Barajas et.al.[2015] have studied a multi-level factorial style to analyze the factors poignant the surface assimilation and medical aid processes. Results show that the MO macromolecule binds powerfully to any or all adsorbents, which sure proteins wasn't free into the answer. The MO surface assimilation capability was highest in carbon and lowest in sand. The functionalized adsorbents were ready to deactivate E. coli with the very best coliform removal discovered in rice husk and carbon. Results of unidirectional ANOVA indicate that the kind of sorbent material is a vital considers E. coli medical aid victimization MO functionalized adsorbents. However, there's no comfortable proof to conclude that carbon is superior to rice husk. Overall, these results recommend the likelihood of planning a low-priced biofilter that uses MO immobilized adsorbents as wadding.[7] Jyotismita Chaki et.al.[2016] have projected the present work proposes associate degree approach for the popularity of plants from their digital leaf pictures victimization multiple visual options to handle heterogeneous plant sorts. Recognizing the actual fact that plant leaves will had a range of recognizable options like color (green and non-green) and form (simple and compound) and texture (vein structure patterns), one set of options might not be economical enough for complete recognition of heterogeneous plant sorts. Consequently a superimposed design was projected wherever every layer handles a selected style of visual characteristics victimization its own set of options to form a made-to-order information model. Options from numerous layers were later fed to associate degree array of custom classifiers for a additional sturdy recognition. During this work we tend to enumerate on the colour and form layers solely. A dataset involving 600 leaf pictures divided over thirty categories and as well as inexperienced, non-green, straightforward and compound leaves, is employed to check the performance and effectiveness of the approach.[8] Neha Pandey et.al.[2015] have studied pattern recognition applications, limiting the quantity of options was a awfully vital demand owing to high dimensional value moreover because the risk of "over fitting" obligatory by the high-dimensional feature vectors. Feature set choice addresses the spatial property reduction downside by determinant a set of accessible options that is most essential for classification. He studied a Genetic rule. The GA is employed for feature set choice and neural network for classification. They calculated the accuracy of those 2 classifiers.[9] Phan Thi Thu Hong et.al.[2015] have studied a system for machine-driven classification of rice selection for rice seed production victimization laptop vision and image process techniques. He same that Rice seeds of various

varieties was visually terribly similar in color, form and texture that build the classification of rice seed varieties at high accuracy difficult. They investigated numerous feature extraction techniques for economical rice seed image illustration. He had studied six completely different rice seed varieties in northern Vietnam were no heritable and analyzed. They calculated average accuracy of ninety.54% victimization Random Forest methodology with a straightforward feature extraction technique.[10] Pranjali B. Padol et.al.[2016] have studied Image process is one among the wide used technique is adopted for the plant leaf diseases detection and classification. They supposed to help within the detection and classification leaf diseases of grape victimization SVM classification technique. 1st the morbid region is found victimization segmentation by K-means cluster, then each color and texture options square measure extracted. Finally classification technique is employed to find the kind of plant disease. The projected system will with success detect and classify the examined illness with accuracy of eighty eight.89%. [11] Qicong Wang et.al.[2016] have studied to attain the effective plant leaf classification victimization manifold learning, the native pure mathematics structure of plant leaves was ready to be preserved effectively and a discriminant manifold-based projection ought to be learned to capture the dominant structure options higher. They foremost use physicist filter to model the feel of plant leaf pictures because the samples. Then for the high-dimensional options, we tend to construct the nearness data graph supported 2 constraints, i.e., low rank and scantness. There by, we tend to propose a completely unique conserving Discriminant Manifold topological space Learning (PDMSL) to imbed the knowledge graph and learn a typical topological space by introducing each graph Laplacian and distributed regularizers. The low-dimensional embedding and projection appreciate the learned manifold topological space had higher intra-class similarity and inter-class discriminant ability of physicist options of the leaf, and might additionally touch upon out-of-sample extension with efficiency. The experiments on Swedish leaf datasets demonstrate that the projected methodology was far more effective than different baseline strategies.[12] Santanu Phadikar et.al.[2016] have studied Automatic identification of plant diseases was the new difficult space for the researchers. One among the foremost vital steps in automatic identification of plant diseases is to extract the infected region from the traditional portion of the plant. He had studied the on the market VI Normalized distinction Vegetation Index (NDVI), inexperienced Normalized distinction Vegetation Index (GNDVI), increased Vegetation Indices (EVI), and Soil Adjusted Vegetation Index (SAVI). Then 5 completely different texture options specifically Homogeneity, Correlation, Contrast, Energy and Entropy of the infected regions square measure computed. These feature values were then used for classifying the diseases victimization fifteen completely different classifiers (such as: naïve Thomas Bayes, SVM, Part, J48 classifiers, etc.) on the market in rail tool. Among these VI, EVI provides the most effective average result.[13] author Marshall et.al.[2015] have studied Food quality analysis was a key space wherever reliable, nondestructive and correct measures was needed. He same Hyper spectral imaging was a technology that meets all of those necessities however on condition that acceptable

signal process techniques was enforced. During this paper he had given four completely different applications of hyper spectral imaging for food quality analysis: period of time estimation of baked sponges; beef quality prediction; classification of Chinese tea leaves; and classification of rice grains. The had studied the HIS model to spot the standard of the food. [14] Vijai Singh et.al.[2015] have studied associate degree rule for image segmentation technique used for automatic detection moreover as classification of plant leaf illness plant diseases and survey on completely different diseases classification techniques which will be used for plant leaf disease detection. Image segmentation, that is a vital facet for illness detection in plant disease, is completed by victimization genetic rule. [15] Yizhe Zhang et.al.[2015] have studied Glands square measure very important tissues found throughout the body and their structure and performance square measure full of several diseases. the power to section and find glands among different kinds of tissues is very important for the study of traditional and illness processes and was pronto envisioned by pathologists in microscopic detail. During this paper, we tend to develop a replacement approach for segmenting and police investigation internal organ glands in Hamper stained microscopic anatomy pictures, that utilizes a group of advanced image process techniques like graph search, ensemble, feature extraction and classification. The tactic computes quickly, and was ready to preserve secretary organ boundaries robustly and find glands accurately. we tend to tested the performance of secretary organ detection and segmentation by analyzing a dataset of 1723 glands from digitized high-resolution clinical microscopic anatomy pictures obtained in traditional and morbid intestines. The experimental results show that the tactic outperforms significantly the progressive strategies for secretary organ segmentation and detection tasks.[16]

VI. JUSTIFICATION AND LIKELY BENEFITS

The plant leaf classification is the major problem in today's world. In this work I have studied the different plants real time artificial prototype model for detection of rice leaf diseases based on their color, shape and other features. Our method is more accurate then the exiting one. The previous work studied with the help of SVM and different methods but I will use different classifiers and morphological operators in this work.

VII. PROBLEM STATEMENT

Plant leaf classification is that the major downside and additionally their diseases are major downside. Most of plant leaf has attacks by snail, worm and fungi. Moreover, once the leaf had been infected or attacked, the others areas had been exposed to be infected. Thus, it'll decrease the lifetime of plant. there's the overlapping of the plant diseases downside. Plants play a crucial role in our lives, while not plants there'll not be the existence of the ecology of the world. The big quantity of leaf varieties currently makes the soul in a very front of some issues within the specification of the utilization of plants, the primary ought to grasp the utilization of a plant is that the identification of the plant leaf. The matter with variety of those techniques is that they need some manual intervention like properly positioning the image or distinctive the top

points of the leaf's main vein. These options extracted by vertical and horizontal cacophonous the leaf pictures. Another downside to spot the ripening stage of the rice crop and fruits (like guava and bananas etc.). as a result of lay men is unable to spot supported their color throughout the harvest system whether or not the crop and their fruits square measure ripped or not. There square measure completely different colours and forms of rice found in today's system and that we square measure unable to spot the assorted kinds of rice and their shape.

VIII. THE PROPOSED APPROACH – STEPBY-STEP DETAILS

The overall concept that is the framework for any vision related algorithm of image classification is almost the same.

Image Acquisition
Image Preprocessing
Image Segmentation
Feature Extraction
Statistical Analysis
Classification Based on Classifier

Figure 3: The basic procedure of the proposed image processing- based disease detection solution

First, the digital pictures are no heritable from the setting employing a camera. Then image-processing techniques are applied to the no heritable pictures to extract helpful options that are necessary for any analysis. After that, many analytical discriminating techniques are accustomed classify the pictures in keeping with the particular drawback at hand. Figure 3 depicts the essential procedure of the planned vision-based detection rule during this analysis. The planned approach step - by - step of the image segmentation and recognition processes is illustrated in rule one. Within the initial step, the RGB pictures of all the leaf samples were picked up. Some real samples of these diseases are shown in Figure two. It's obvious from Figure 3 that leaves happiness to early scorch, soft mildew; ashen mildew and late scorch have important variations kind greasy spot leaves in terms of color and texture. Also, Figure three shows 2 images; the left image is infected with small white illness, and also the right image may be a traditional image. However, the leaves associated with these six categories (early scorch, soft mildew, ashen mold, late scorch, small white and normal) had terribly little variations as discernible to the human eye, which can justify the misclassifications supported eye.

Different types of Diseases on Leaves of various Plants

The diseases on the cotton leaves are classified as,

A.1) microorganism disease: e.g. microorganism Blight, Crown Gall, Lint Degradation.

A.2) plant life diseases: e.g. Anthracnose, Leaf Spot.

A.3) infectious agent disease: e.g. Leaf Curl, Leaf Crumple, Leaf Roll.

There are totally different Phases to method the works that are given below:

PHASE 1 - K-MEANS bunch TECHNIQUE

There are two preprocessing steps that are required so as to implement the K-means bunch algorithm: The part starts initial by making device-independent color area transformation structure. During a device freelance color area, the coordinates accustomed specify the colour can turn out identical color despite the device accustomed draw it. Thus, we have a tendency to create the colour transformation structure that defines the colour area conversion. Then, we have a tendency to apply the device-independent color area transformation that converts the colour values within the image to the colour area per the colour transformation structure. the colour transformation structure specifies numerous parameters of the transformation. a tool dependent color area is that the one wherever the resultant color depends on the instrumentality accustomed turn out it. for instance the colour created victimisation pel with a given RGB values are altered because the brightness and distinction on the show device used. so the RGB system may be a color area that's dependent. The K-means bunch rule tries to classify objects (pixels in our case) supported a collection of options into K range of categories. The categoryification is finished by minimizing the add of squares of distances between the objects and also the corresponding cluster or class center of mass [3; 4]. However, K-means bunch is employed to partition the leaf image into four clusters during which one or additional clusters contain the illness just in case once the leaf is infected by quite one illness. In our experiments multiple values of range of clusters are tested. Best results were discovered once the quantity of clusters was three or four. A stem image infected with early scorch and its initial cluster (the infected object itself) is shown in Figure three.
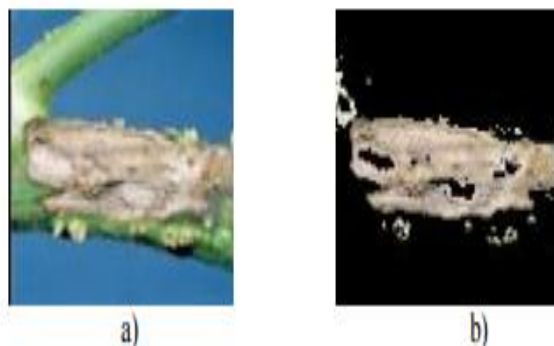

Figure 4: A stem image infected with early scorch; a) original image b) cluster 1 image

PHASE 2 – MASKING THE inexperienced PIXELS and also the PIXELS ON THE BOUNDARIES

This section consists of 2 steps: The largely inexperienced coloured pixels are known, so the worldwide image threshold exploitation Otsu's methodology [5; 6] has been applied so as to specify the varied threshold worth that chooses the brink to attenuate the interclass variance of the brink black and white pixels. Next, the inexperienced constituent's are cloaked as follows: if the inexperienced element of pixel intensities is a smaller amount than the computed threshold worth, then, the red, inexperienced and blue parts of the constituent are cleared. Successive step during this section is concentrated on deleting

each the pixels with zeros parts and also the pixels on the boundaries of the infected cluster(s).

PHASE 3– OPTIONS EXTRACTION

Within the planned approach, the methodology the tactic the strategy adopted for extracting the feature set is termed the colour Co-occurrence methodology or CCM method in brief. It's a technique, during which each the colour and texture of a picture are taken into consideration, to attain distinctive options, that represent that image.
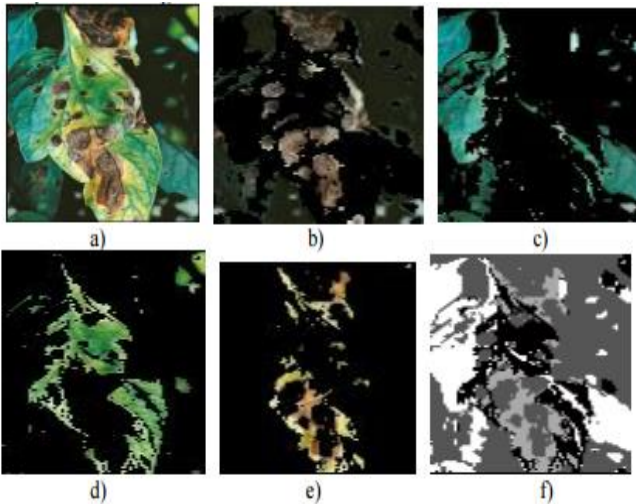
Figure 5: An example of the output of K-Means clustering for a leaf that is infected with early scorch disease. (a) The infected leaf picture. (b, c, d, e) the pixels of the first, second, the third and the fourth clusters. (e) a single gray-scale image with the pixels colored based on their cluster index.

The present work shows a comparison of classification of seven different representations of plant leaves using three features extracted from the images.
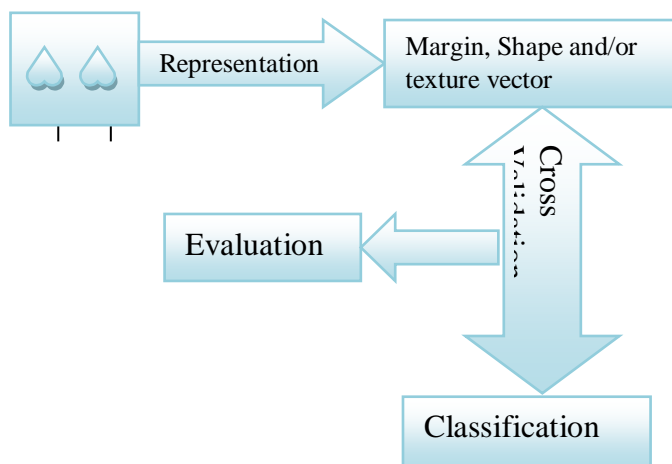
Figure 6: shows the architecture of proposed approaches

A.   Representation of samples

Beginning with representation of species by three features extracted from images: a fine-scale margin feature histogram,

then a Centroid Contour Distance Curve shape signature, and finally an interior texture feature histogram. We put values of each feature in 64 elements vector, then we tried to combine these vectors two by two in one 128 elements vector, and finally we presented species combining the three vectors together in one 192 elements vector

B.   Classification

In each case, we used three different approaches for classification: probabilistic approach using Naïve Bayes algorithm, hierarchical approach using Decision Tree C4.5 algorithm, and finally, approach based on distance calculation using K-nearest neighbors' (K-NN) algorithm with k = 3, 4, 5, 6, or 7 and using Euclidian distance.

IX.   CONCLUSION & FUTURE WORK

Disease management could be a difficult task. Largely diseases area unit seen on the leaves or stems of the plant. Precise quantification of those visually determined diseases, pests, traits has not studied nevertheless attributable to the complexness of visual patterns. Then there has been increasing demand for additional specific and complex image pattern understanding. In biology, generally thousands of pictures area unit generated during a single experiment. These pictures will be needed for additional studies like classifying lesion, rating quantitative traits, hard space consumed by insects, etc. most of those tasks area unit processed manually or with distinct package packages. The aim of the work is to develop a Leaf recognition program supported specific characteristics extracted from photography. This presents associate degree approach wherever the plant is known supported its leaf options like space, bar graph effort and edge detection and classification. of these work is enforced within the future, as a result of here solely review of labor is finished. within the future a man-made epitome model is developed for detection of leaf diseases supported their totally different options like color, form and different options mistreatment thresholding algorithmic rule.

REFERENCES

[1] Anand R et.al.(2016) "An Application of image processing techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method" IEEE Fifth International Conference On Recent Trends In Information Technology, pp. 27-34.

[2] Azeil Louisse Codizar et.al. (2016) "Plant leaf recognition by venation and shape using artificial neural networks", Information, Intelligence, Systems & Applications (IISA), IEEE 7th International Conference, pp. 225-233.

[3] Chitra Anil Dhawale et.al. (2016) "Analysis of nutritional deficiency in citrus species tree leaf using image processing" Advances in Computing, Communications and Informatics (ICACCI), IEEE International Conference, pp.-615-620.

[4] Davoud Ashourloo et.al.(2016) "An Investigation Into Machine Learning Regression Techniques for the Leaf

Rust Disease Detection Using Hyperspectral Measurement" IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, pp.202-206.

[5] Harshal Waghmare et.al. (2016) "Detection and Classification of Diseases of Grape Plant Using Opposite Colour Local Binary Pattern Feature and Machine Learning for Automated Decision Support System" IEEE 3rd International Conference on Signal Processing and Integrated Networks (SPIN), pp.200-204.

[6] John Barajas et.al. (2016) "Water Disinfection Using Moringa Protein Adsorbed On Rice Husk Ash" IEEE Systems and Information Engineering Design Conference (SIEDS '16), pp.195-200.

[7] John Raymond Barajas et.al. (2015) "Development of a low-cost water treatment technology using Moringa oleifera seeds" IEEE Systems and Information Engineering Design Symposium (SIEDS), pp.202-206.

[8] Jyotismita Chaki et.al.(2016)"Plant leaf recognition using a layered approach" Microelectronics, Computing and Communications (MicroCom), IEEE International Conference, pp.101-106.

[9] Neha Pandey et.al.(2015) "A Novel Feature Learning For Image Classification Using Wrapper Approach In GA", IEEE International Conference on Signal Processing and Integrated Networks (SPIN), pp. 1-8.

[10] Phan Thi Thu Hong et.al. (2015) "Comparative Study On Vision Based Rice Seed Varieties Identification", IEEE Seventh International Conference on Knowledge and Systems Engineering, pp.1-4.

[11] Pranjali B. Padol et.al.(2016) "SVM Classifier Based Grape Leaf Disease Detection", IEEE Conference on Advances in Signal Processing (CASP) Cummins College of Engineering for Women, Pune. Jun 9-11, pp.201-204.

[12] Qicong Wang et.al. (2016) "Preserving discriminant manifold subspace learning for plant leaf recognition", Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), IEEE 12th International Conference, pp.201-204.

[13] Santanu Phadikar et.al.(2016) " Vegetation Indices Based Segmentation for Automatic Classification of Brown Spot and Blast Diseases of Rice", IEEE 3rd International Conf. on Recent Advances in Information Technology(RAIT), pp.200-204.

[14] Stephen Marshall et.al.(2015) "Hyperspectral Imaging For Food Applications", IEEE 23rd European Signal Processing Conference (EUSIPCO), pp.206-210

[15] Vijai Singh et.al.(2015) " Detection of unhealthy region of plant leaves using Image Processing and Genetic Algorithm", IEEE International Conference on Advances in Computer Engineering and Applications (ICACEA) IMS Engineering College, Ghaziabad, India, pp.200-204.

[16] Yizhe Zhang et.al.(2015) " A seeding-searching-ensemble method for gland segmentation and detection" Bioinformatics and Biomedicine (BIBM), IEEE International Conference, pp.205-210.