Detection and Classification of Plant Disease Using Deep Learning

Sumathi R[1], Jane Marie Harshitha E[2], Nandini Devi M[3], Prathema Rachil A[4], Ramya P[5]

[1]*Head of the Department, Information Technology, Saranathan College of Engineering, Trichy*

[2,3,4,5] *Information Technology, Saranathan College of Engineering, Trichy*

***Abstract*- The purpose of Agriculture is not only to feed ever growing population but it's an important source of energy and a solution to solve the problem of global warming. Plant diseases are extremely significant, as that can adversely affect both quality and quantity of crops in agriculture production. Plant disease diagnosis is very essential in earlier stage in order to cure and control them. Generally the naked eye method is used to identify the diseases. In this method experts are involved who have the ability to detect the changes in leaf color. This method involves lots of efforts, takes long time and also not practical for the large fields. Many times different experts identify the same disease as the different disease. This method is expensive as it requires continuous monitoring of experts. Plant diseases can increase the cost of agricultural production and may extend to total economic disaster of a producer if not cured appropriately at early stages. The producers need to monitor their crops and detect the first symptoms in order to prevent the spread of a plant disease, with low cost and save the major part of the production. Hiring professional agriculturists may not be affordable especially in remote isolated geographic regions. Machine learning algorithm in image can offer an alternative solution in plant monitoring and such an approach may anyway be controlled by a professional to offer his services with lower cost. It includes features extraction and classification which includes local binary method and discrete wavelet transform algorithm and also image classification approach which includes neural network algorithm to predict various types of diseases. And also extend the approach to recommend the fertilizers based on severity analysis with measurements.**

***Index terms*- Plant disease prediction, Features extraction, Classification, Fertilizer Recommendation, Neural network approach**

I.INTRODUCTION

India is an agricultural country. Farmers have wide range of diversity to select suitable fruit and vegetable crop. Research work develops the advance computing system to identify the diseases using infected images of various leaf spots. Images are captured by digital camera mobile and processed using image growing, then the part of the leaf sport has been used for the classification purpose of the train and test. The technique evolved into the system is both Image processing techniques and advance computing techniques. Agriculture is the mother of all nations. Research in agriculture domain is aimed towards increase the quality and quantity of the product at less expenditure with more profit. The quality of the agricultural product may be degraded due to plant diseases. These diseases are caused by pathogens viz.., fungi, bacteria and viruses. Therefore, to detect and classify the plant disease in early stage is a significant task. Farmers require constant monitoring of experts which might be prohibitively expensive and time consuming. Depending on the applications, many systems have been proposed to solve or at least to reduce the problems, by making use of image processing and some automatic classification tools.

Image Analysis Can Be Applied For The Following Purposes:

- To detect diseased leaf, stem, fruit.
- To quantify affected area by disease.
- To find the boundaries of the affected area.
- To determine the color of the affected area.
- To determine size & shape of leaf.
- To identify the Object correctly. Etc.

Disease management is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Precise quantification of these visually observed diseases, pests, traits has not studied yet because of the complexity of visual patterns. Hence there has been increasing demand for more specific and sophisticated image pattern understanding.

Various Types of Leaf Spot Diseases:
- Bacterial
- Fungal
- Viral

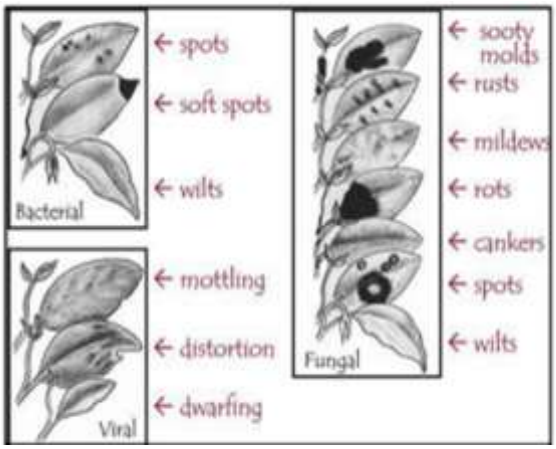

Fig 1: Various types of diseases

Most leaf diseases are caused by fungi, bacteria and viruses. Fungi are identified primarily from their morphology, with emphasis placed on their reproductive structures. Bacteria are considered more primitive than fungi and generally have simpler life cycles. With few exceptions, bacteria exist as single cells and increase in numbers by dividing into two cells during a process called binary fission viruses are extremely tiny particles consisting of protein and genetic material with no associated protein. In biological science, sometimes thousands of images are generated in a single experiment. There images can be required for further studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc. Almost all of these tasks are processed manually or with distinct software packages. It is not only tremendous amount of work but also suffers from two major issues: excessive processing time and subjectiveness rising from different individuals. Hence to conduct high throughput experiments, plant biologist need efficient computer software to automatically extract and analyze significant content. Here image processing plays important role. This paper provides image processing techniques used for studding leaf diseases.

II. RELATED WORK

Anil a. Bharate, et.al,…[1] implemented various techniques such as spectroscopic and imaging technology which can be used to obtain patterns. Farmers can make use of automation techniques and tools to integrate knowledge, product and services to improve productivity, grade and yield with the help of smart farming. It helps farmers to identify the plant disease in primary stage and take time to time decisions which in turn saves time and reduces loss of plant due to diseases. Farmers will also be able to identify different grades of fruit before marketing. The purpose of this paper is to survey on how to monitor diseases on plants and suggest better solution for healthy yield and productivity. Many existing systems use two databases of images, one for query images and other for training images. Diseases of three fruits namely apple, grapes and pomegranate have been identified by many systems. Only grapes and other few related species are affected by this fungus. It is most frequently occurring disease on grapes. Primary symptom of powdery mildew is whitish or greenish powdery patches appearing on the underneath of basal leaves. It also causes leaf curling, withering along with blotched or deformation of badly infected leaves.

Jyotismitachaki, et.al,[2]prepared digital plant cataloging systems for recognizing plant species in efficient ways. From this perspective, the current work proposes an innovative scheme of a plant recognition system based on digital images of plant leaves. People recognize a specific plant type by prominent characteristics of its leaf like shape and texture. Data modeling techniques have been employed here to represent these characteristics using a set of computer recognizable features. Shape of the leaf is represented by Curvelet transform coefficients together with Invariant Moments, while texture is modeled with Gabor filter outputs and metrics derived from Gray Level Co-occurrence Matrices. Features are subsequently fed to two neural based classifiers to discriminate them into a number of predefined classes. Experimentations are done using features individually as well as in various combinations to study optimal conditions. The approach proposed by author a combination of texture and shape modeling techniques, since these are thought to be significant parameters for discrimination. Texture of a plant leaf is captured using complex Gabor filter (GF) and gray level co-occurrence matrix (GLCM) while shape of the leaf is captured using curvelet transforms (CT) and invariant moments (IM). The feature values generated are however sensitive to the size and orientation of the leaf image.

Neerajkumar, et.al,[3]developed to greatly speed up the manual process of plant species identification, collection, and monitoring. Without visual recognition tools such as Leafsnap, a dichotomous key (decision tree) must be manually navigated to search the many branches and seemingly endless nodes of the taxonomic tree. Identifying a single species using this process by answering dozens of often-ambiguous questions, such as, are the leaves at and thin?" may take several minutes or even hours. This is difficult for experts, and exceedingly so (or even impossible) for amateurs. Untrained users initially try to take photos of leaves in-situ, with multiple leaves present amid clutter, often with severe lighting and blur artifacts, resulting in images that cannot handle (usually due to segmentation failures). In addition, many users also take photos of objects that are not leaves. And address both of these issues by first running a binary leaf/non-leaf classifier on all input images. If this classifier detects that an input image is not valid of a single leaf, placed on a light, un-textured background with no other clutter and informs the user of this fact and instructs them on how to take an appropriate image. Then found this simple procedure very helpful for training users without the need for long tutorials or help pages, which often go unread. It also greatly reduces the computational load on our server, as images that fail this classification are discarded from further processing.

Naresh.Y.G, et.al,[4] implementedthe digitization for useful species of plants and their information is necessary. This digitization is primarily based on the images, which comprise of the plant leaves and other parts of the plants such as fruits, flowers, pollen grains etc. This leads to enormous collection of digital data, which initiates the need of classification and retrieval. In order to provide the information to laymen about the medicinal plants, it is very important to have a computer vision system to identify the plant species with the help of their digitized databases. In this work, a symbolic approach for classification of plant leaves based on texture features is proposed. Modified Local binary patterns (MLBP) is proposed to extract texture features from plant leaves. Texture of plant leaves belonging to same plant species may vary due to maturity levels, acquisition and environmental conditions. Hence, the concept of clustering is used to choose multiple class representatives and the intra-cluster variations are captured using interval valued type symbolic features. The classification is facilitated using a simple nearest neighbor classifier. Extensive experiments have been carried out on newly created UoM medicinal plant dataset as well as publically available Flavia, Foliage and Swedish plant leaf datasets. Results obtained by proposed methodology are compared with the contemporary methodologies. The Outex dataset is also considered for experiments and the results are promising even on this synthetic dataset.

Mónica g. Larese, et.al,[5] implemented the segmentation algorithm is performed using the unconstrained hit-or-miss transform and adaptive thresholding. Several morphological features are computed on the segmented venation, and classified using four alternative classifiers, namely support vector machines (linear and Gaussian kernels), penalized discriminant analysis and random forests. The performance is compared to the one obtained with cleared leaves images, which require a more expensive, time consuming and delicate procedure of acquisition. The results are encouraging, showing that the proposed approach is an effective and more economical alternative solution which outperforms the manual expert's recognition. The model’s saliency map is endowed with internal dynamics which generate attentional shifts. This model consequently represents a complete account of bottom-up saliency and does not require any top-down guidance to shift attention. This framework provides a massively parallel method for the fast selection of a small number of interesting image locations to be analyzed by more complex and time consuming object-recognition processes. Extending this approach in “guided-search,” feedback from higher cortical areas (e.g., knowledge about targets to be found) was used to weight the importance of different features, such that only those with high weights could reach higher processing levels

III. EXISTING METHODOLOGIES

Leaves are the most obvious and widespread choice for tree species recognition, even though the botanical classification was not built upon their properties. They can be found almost all year long, are easy to photograph, and their shapes present well studied specificities that make the identification, if not trivial, possible. With the aim of being an educational tool, it relies on high-level geometric criteria inspired by those used by botanists, that make a semantic interpretation possible, to classify a leaf into a list of species. Digital image processing will improve the quality of the image by removing noise & other unwanted pixels and obtain more information from image. Image segmentation is a mid-level processing technique used to analyze the image and can be used to classify or cluster an image into several disjoint parts by grouping the pixels to form a region of homogeneity based on the pixel characteristics like gray level, color, texture, intensity and other features. The main purpose of the segmentation process is to get more information about the image, the region we are interested in and to clearly differentiate the object and the background in an image. The criteria for segmenting the image is very hard to decide as it varies from image to image and also varies significantly on the modal quality of image. In some cases interactive methods can be laborious and time consuming and in some cases manual interaction to segment the image may be error-prone while the fully automated approach can give error output

IV. PROPOSED METHODOLOGIES

Even when considering trees only, leaves show an impressively wide variety in shapes. It is however necessary to come up with a representation of what a leaf is, that is accurate enough to be fitted to basically any kind of leaf. The general shape of a leaf is a key component of the process of identifying a leaf. Botanists have a whole set of terms describing either the shape of a simple leaf, of the lobes of a palmate leaf, or of the leaflets of a compound leaf. The problem being that the borders between the different terms are not well defined, since leaves can naturally have non-canonical, intermediate shapes. The margin of the leaf is also a very important feature to spot. Its shape can be determining when trying to discriminate two species that have more or less the same global shape. It may consist of teeth of various sizes and frequencies, regularly arranged or not, from large spiny points, to small regular saw-like teeth, or even to a smooth entire border. We present a study on segmentation of leaf images restricted to semi-controlled conditions, in which leaves are photographed against a solid light-colored background. Such images can be used in practice for plant species identification, by analyzing the distinctive shapes of the leaves. We restrict our attention to segmentation in this semi controlled condition, providing us with a more well-defined problem, which at the same time presents several challenges. The most important of these are: the variety of leaf shapes, inevitable presence of shadows and specularities, and the time constraints required by interactive species identification applications. We evaluate several popular segmentation algorithms on this task. In everyday more urbanized and artificial world, the knowledge of plants, that used to constitute our most immediate environment, has somehow been lost, except for a handful of specialists. What is allegedly seen as unquestionable progress also scattered away the names and uses of so many trees, flowers and herbs. But nowadays, with a certain resurgence of the idea that plant resources and diversity ought to be treasured, the will to regain some touch with nature feels more and more tangible. And making it possible, for whoever feels the need, to identify a plant species, to learn its history and properties, is as much a way to transmit a vanished knowledge, as to allow people to get a glance at nature's unfathomable richness. The identification of species is the first and essential key to understand the plant environment. Botanists traditionally rely on the aspect and composition of fruits, flowers and leaves to identify species. But in the context of a widespread non-specialist-oriented application, the predominant use of leaves, which are possible to find almost all year long, simple to photograph, and easier to analyze from two-dimensional images, is the most sensible and widely used approach in image processing. In the process of tree identification from pictures of leaves in a natural background, retrieving an accurate contour is a challenging and crucial issue. In this paper we introduce a method designed to deal with the obstacles raised by such complex images, for simple and lobed tree leaves. A first features extraction method based on Local binary pattern and

Discrete Wavelet transform are performed, and later used to guide the evolution of leaf boundaries. And combining global shape descriptors given by the polygonal model with local curvature-based features, the leaves are then classified over leaf datasets. And implement classification algorithm to classify the diseases and recommend the fertilizers to affected leaves. The classification algorithm includes the convolutional neural network algorithm to improve the accuracy in disease classification and recommend the leaf disease based on multi-class classification. The proposed layout is shown in fig 2.

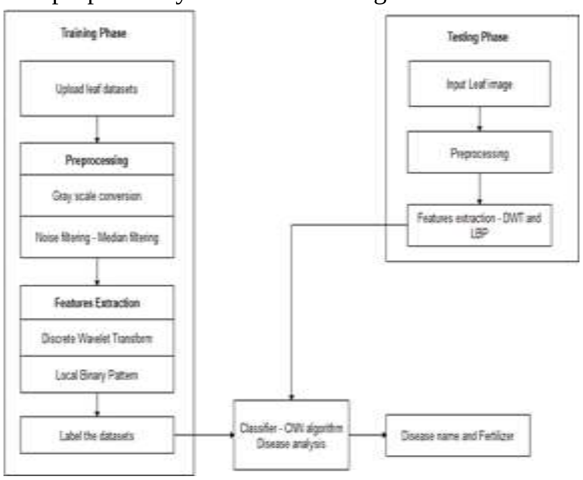

Fig 2: Proposed framework

4.1 Image acquisition and Preprocessing

Leaves are structures specialized for photosynthesis and are arranged on the tree in such a way as to maximize their exposure to light without shading each other. In this module, we can upload the leaf images from the datasets. This database called LEAF was originally created for experiments with recognition of wood species based on a leaf shape. It contains leaves of species growing in the Czech Republic, both trees and bushes; native, invasive and imported (only those imported species which are common in parks are included). The number of samples (leaves) of one species varies from 2 to 25; their total number in the database is 795. The leaves were scanned with 300 dpi, threshold (binarized), preprocessed (denoising and cleaning) and saved in PNG format. In this module convert the RGB image into gray scale image. The colors of leaves are always green shades and the variety of changes in atmosphere cause the color feature having low reliability. Therefore, to recognize various plants using their leaves, the obtained leaf image in RGB format will be converted to gray scale before pre-processing. The formula used for converting the RGB pixel value to its grey scale counterpart is given in Equation.

Gray = 0.2989 * R + 0.5870 * G + 0.1140 * B

where R, G, B correspond to the color of the pixel, respectively.

Then remove the noises from images by using filter techniques. The goal of the filter is to filter out noise that has corrupted image. It is based on a statistical approach. Typical filters are designed for a desired frequency response. Filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges. And implement image binarization tasks.

4.2 Image Segmentation

In this module, we can implement local binary pattern and discrete wavelet transform with automatic descriptors. Unconstrained active contours applied to the complex natural images we aim at dealing with would produce unsatisfying contours, that would try and make their way through every possible gap and aw in the border of the leaf. The solution we propose is to use the polygonal model obtained after the first step not only as an initial leaf contour but also as a shape prior that will guide its evolution towards the real leaf boundary.

Use the resulting polygon as a shape prior to drive the evolution of an active contour

- Set the initial contour on a contracted version of the polygon
- Constraint the contour to remain close to the polygon

Energy Formulation

- For a contour τ delineating a region $\Omega(\tau)$:
- $E(\tau) = \alpha ELeaf(\tau) + \beta EShape(\tau) + \gamma EGradient(\tau) + \delta ESmooth(\tau) - \delta EBalloon(\tau)$

Instead of having an external energy term based on color consistency, or distance to a mean, we decided to reuse the dissimilarity map from the previous step, considering we have already an efficient measure of how well a pixel should fit in the leaf, in terms of color.

4.3 Disease prediction

Leaves are affected by bacteria, fungi, virus and other insects. In this module implement convolutional neural network algorithm to classify the leaf image as normal or affected. Vectors are constructed based leaf features such as color, shape, textures. Then layers can be constructed with conditions to categorize the preprocessed leaves. And also implement multiclass classifier, we can predict diseases in leaf images with improved accuracy.

Steps in CNN algorithms:

Step 1: Randomly initialize the weights and biases.

Step 2: feed the training sample.

Step 3: Propagate the inputs forward; compute the net input and output of each unit in the hidden and output layers.

Step 4: back propagate the error to the hidden layer.

Step 5: update weights and biases to reflect the propagated errors.

Training and learning functions are mathematical procedures used to automatically adjust the network's weights and biases.

Step 6: terminating condition

4.4 Fertilizer Recommendation

In this module recommend the fertilizer for affected leaves based on severity level. Fertilizers may be organic or inorganic. Admin can store the fertilizers based on disease categorization with severity levels. The measurements of fertilizers can be extracted based on disease severity.

EXPERIMENTAL RESULTS

In this system implemented for disease prediction using MATLAB. The proposed framework is shown in following figures.

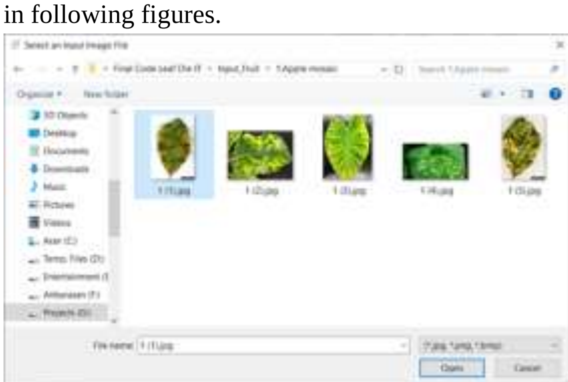

Fig 3: Image Upload

The preprocessing results are shown in fig 4 and 5

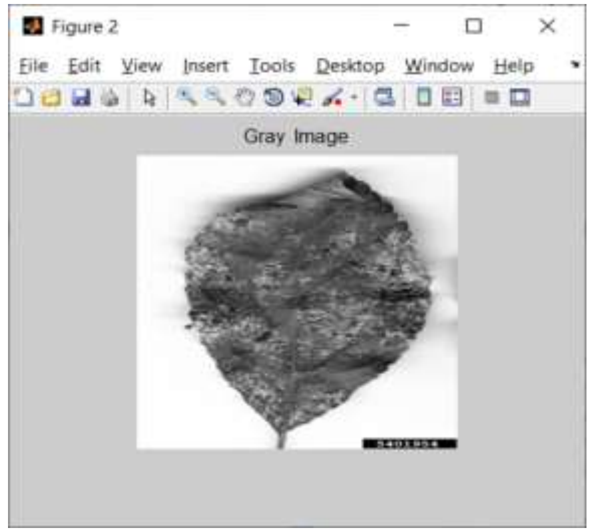

Fig 4: Gray Scale conversion

Fig 5: Noise filtering with Contrast enhanced

Based on these above figures, noises are removed and perform HIS conversion and shown in fig 6.

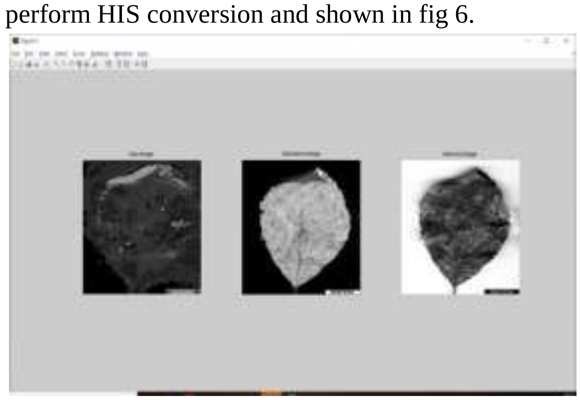

Fig 6: Hue, saturation and intensity image

The leaf features are extracted and segmented to extract the leaf boundaries and shown in fig 7

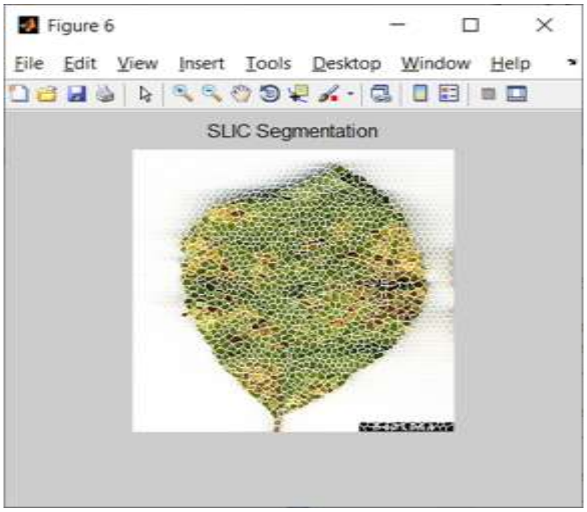

Fig 7: Segmented image

The disease prediction and fertilizer recommendation can be shown in fig 8

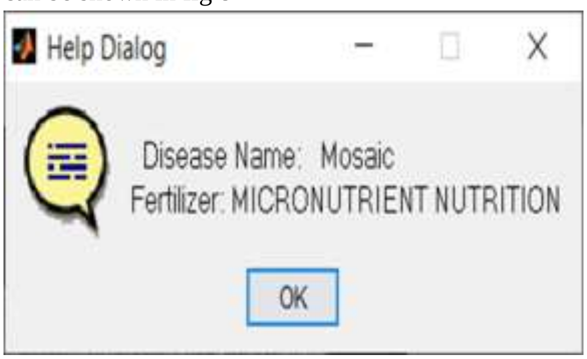

Fig 8: Disease prediction with fertilizer recommendation

In this chapter used real time datasets. This framework used the features extraction and classification techniques. Then can evaluate the performance using accuracy metrics. The accuracy metric is evaluated as

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100$$

The proposed algorithm provide improved accuracy rate than the machine learning algorithms.

Accuracy table shown in table 1

Algorithm	Accuracy (%)
Naivesbayes	20
SVM	50
CNN	90

Table (2) Accuracy table

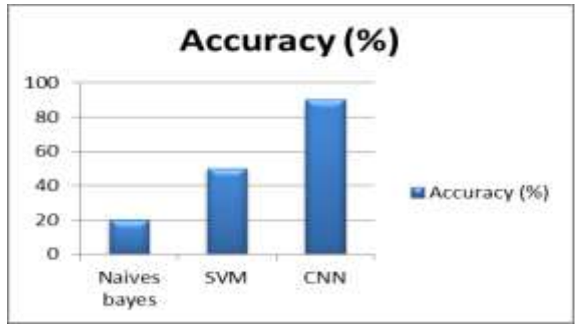

Fig 9: Performance report

From the performance chart, CNN provide high level accuracy than the existing machine learning algorithms.

VI. CONCLUSION

In this paper, we overview the various techniques and algorithms are proposed for segmentation and classification methods for improve the quality of segmentation. But the result shows that segmentation algorithms do not work properly and can't implement in large datasets rather than proposed graph cut model. We have presented a method designed to perform the segmentation of a leaf in a natural scene, based on the optimization of a polygonal leaf model used as a shape prior for an exact leaf boundary using Local binary pattern and discrete wavelet transform algorithm. It also provides a set of global geometric descriptors that, later combined with local curvature-based features extracted on the final contour, make the classification into tree species possible. The segmentation process is based on a color model that is robust to uncontrolled lighting conditions. But a global color model for a whole image may sometimes not be enough, for leaves that are not well defined by color only. Finally implement neural network classification algorithm to classify the leaf diseases as bacteria, fungi and virus. Then recommend the fertilizers to affected leaves based on measurements.

REFERENCES

[1] Bharate, Anil A., and M. S. Shirdhonkar. "A review on plant disease detection using image processing." 2017 International Conference on Intelligent Sustainable Systems (ICISS).IEEE, 2017.

[2] Chaki, Jyotismita, Ranjan Parekh, and Samar Bhattacharya. "Plant leaf recognition using texture and shape features with neural

classifiers." Pattern Recognition Letters 58 (2015): 61-68.

[3] Kumar, Neeraj, et al. "Leafsnap: A computer vision system for automatic plant species identification." European conference on computer vision.Springer, Berlin, Heidelberg, 2012.

[4] Naresh, Y. G., and H. S. Nagendraswamy. "Classification of medicinal plants: an approach using modified LBP with symbolic representation." Neurocomputing 173 (2016): 1789-1797.

[5] Larese, Mónica G., et al. "Automatic classification of legumes using leaf vein image features." Pattern Recognition 47.1 (2014): 158-168.

[6] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, ‘‘A comparative study of fine-tuning deep learning models for plant disease identification,’’ Comput. Electron. Agricult., to be published. doi: 10.1016/j.compag.2018.03.032.

[7] S. S. Chouhan, A. Kaul, and U. P. Singh, ‘‘Image segmentation using computational intelligence techniques: Review,’’ in Archives of Computational Methods in Engineering. Dotrich, the Netherlands: Springer, 2018, pp. 1–64. doi: 10.1007/s11831-018-9257-4.

[8] S. W. Chen et al., ‘‘Counting apples and oranges with deep learning: A data-driven approach,’’ IEEE Robot. Autom.Lett., vol. 2, no. 2, pp. 781–788, Apr. 2017. doi: 10.1109/LRA.2017.2651944.

[9] P. A. Dias, A. Tabb, and H. Medeiros, ‘‘Multispecies fruit flower detection using a refined semantic segmentation network,’’ IEEE Robot. Autom.Lett., vol. 3, no. 4, pp. 3003–3010, Oct. 2018. doi: 10.1109/LRA.2018.2849498.

[10] J.Ubbens, M.Cieslak, P.Prusinkiewicz, and I.Stavness,‘‘ Theuseofplant models in deep learning: An application to leaf counting in rosette plants,’’ Plant Methods, vol. 14, p. 6, Jan. 2018. doi: /10.1186/s13007-018-0273-z.

[11] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, ‘‘Fully convolutional networks with sequential information for robust crop and weed detection in precision farming,’’ IEEE Robot. Autom.Lett., vol. 3, no. 4, pp. 2870–2877, Oct. 2018. doi: 10.1109/LRA.2018.2846289.

[12] H. K. Suh, J. IJsselmuiden, J. W. Hofstee, and E. J. van Henten, ‘‘Transfer learning for the classification of sugar beet and volunteer potato under field conditions,’’ Biosyst. Eng., vol. 174, pp. 50–65, Oct. 2018. doi: 10.1016/j.biosystemseng.2018.06.017.

[13] J. G. A. Barbedo, ‘‘Factors influencing the use of deep learning for plant disease recognition,’’ Biosyst. Eng., vol. 172, pp. 84–91, Aug. 2018. doi: 10.1016/j.biosystemseng.2018.05.013.

[14] A. Kamilaris and F. X. Prenafeta-Boldú, ‘‘Deep learning in agriculture: A survey,’’ Comput. Electron. Agricult., vol. 147, pp. 70–90, Apr. 2018. doi: 10.1016/j.compag.2018.02.016.

[15] J.G.A.Barbedo, ‘‘Are view on the main challenges in automatic plant disease identification based on visible range images,’’ Biosyst. Eng., vol.144, pp. 52–60, Apr. 2016. doi: 10.1016/j.biosystemseng.2016.01.017.

[16] S. Kaur, S. Pandey, and S. Goel, ‘‘Plants disease identification and classification through leaf images: A survey,’’ Arch. Comput. Methods Eng., vol. 26, no. 2, pp. 507–530, 2019.doi: 10.1007/s11831-018-9255-6.

[17] A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, and A. Johannes, ‘‘Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild,’’ Comput. Electron. Agricult., to be published. doi: 10.1016/j.compag.2018.04.002.

[18] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, ‘‘Identification of maize leaf diseases using improved deep convolutional neural networks,’’ IEEE Access, vol. 6, pp. 30370–30377, 2018.doi: 10.1109/ACCESS.2018.2844405.

[19] Y.Lu, S.Yi, N.Zeng, Y.Liu, and Y.Zhang,‘‘ Identification of rice diseases using deep convolutional neural networks,’’ Neuro computing, vol. 267, pp. 378–384, Dec. 2017. doi: 10.1016/j.neucom.2017.06.023.

[20] R. Gandhi, S. Nimbalkar, N. Yelamanchili, and S. Ponkshe, ‘‘Plant disease detection using CNNs and GANs as an augmentative approach,’’ in Proc. IEEE Int. Conf. Innov. Res. Develop. (ICIRD), Bangkok, Thailand, May 2018, pp. 1–5. doi: 10.1109/ICIRD.2018.8376321.