

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION (IRJMT)

http://www.mapletreejournals.com/index.php/IRJMT

Received 20 September 2019ISSN 2582-1040Accepted 23 September 20192019; 1(1);8-16Published online 25 September 2019

Fruit and Leaves Disease Prediction Using Deep Learning Algorithm

V.Sathya^{1*}, H.Rafidha², G.Sumitha Rani²

¹Assistant Professor, Department of CSE, AVC College of Engineering, Mayiladuthurai, TN, India ²UG Student, Department of CSE, AVC College of Engineering, Mayiladuthurai, TN, India

*Corresponding author E-Mail ID: <u>ranscreation1819@gmail.com</u>

DOI: https://doi.org/10.34256/irjmt1952

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. Tree leaves and fruit 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 image segmentation and image classification approach to predict various types of diseases using Otsu thresholding method and convolutional neural network method.

Keywords: Disease prediction, Features Extraction, Segmentation, Classification, Neural network

1. 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 project provides image processing techniques used for studding leaf diseases.

Fruit diseases have turned into a dilemma as it can cause significant reduction in both quality and quantity of agricultural products. Recognition Systems of normal and infected apple operations use computer vision strategies that consider features like color, shape and texture for

recognition. In naked eye observation of experts is the main approach adopted in practice for detection and identification of normal and infected apple, but it is a very expensive and expert oriented and time consuming approach, which is not practical Always. Machine learning based on detection and recognition of apple and its diseases can provide clues to identify and treat the diseases in its early stages. Fruit diseases are one of the dangerous causes that reduce quantity and degrade quality of the undeveloped products. There are frequent characteristics and behaviors of such apple diseases in which many of them are noticeable. Hence an intelligent decision supports system for Prevention process for human knowledge is required. Features extraction represents the roundness, shape, color values help for the identification of apple and its structure. Pattern recognition system is a combination of different features together, including entropy, color, shape, and boundary size to perform sequential pattern classification. This method can be applied as a useful image processing tool for other object classification and recognition of normal and infected fruit problems various fields, such as education, food packing image retrieval and plant science research. The fruit diseases are shown in fig 2.



Fig 2. Fruit diseases a) apple scab b) apple rot c) apple blotch

2.RELATED WORK

J. Wäldchen, et.al.,[1] focused on the most discriminating characteristics and narrowing down the set of candidate species. This series of answered questions leads eventually to the desired species. However, the determination of plant species from field observation requires a substantial botanical expertise, which puts it beyond the reach of most nature enthusiasts. Traditional plant species identification is almost impossible for the general public and challenging even for professionals that deal with botanical problems daily, such as, conservationists, farmers, foresters, and landscape architects. Even for botanists themselves species identification is often a difficult task. The situation is further exacerbated by the increasing shortage of skilled taxonomists. The still existing, but rapidly declining high biodiversityand a limited numbers of taxonomists represent significant challenges to the future of biological study and conservation. Recently, taxonomists started searching for more efficient methods to meet species identification requirements, such as developing digital image processing and pattern recognition techniques. The rich development and ubiquity of relevant information technologies, such as digital cameras and portable devices, has brought these ideas closer to reality. Digital image processing refers to the useof algorithms and procedures for operations such as image enhancement, image compression, image analysis, mapping, and geo-referencing. The influence and impact of digital images on the modern society is tremendous and is considered a critical component in a variety of application areas including pattern recognition, computer vision, industrial automation, and healthcare industries.

S. H. Lee, et.al,...[2] implemented deep learning which is a class of techniques in machine learning technology, consisting of multiple processing layers that allow representation learning of multiple level data abstraction. This paper begins with an introduction to deep learning. Next,

proceed to a critical and comprehensive review of existing methods and a description of the context of plant identification - i.e. how species are delimited by botanists using morphology. Then, introduce the idea of deep learning for automatic processing and classification in order to learn and discover useful features for leaf data. And describe how computational methods can be adapted and learnt using visual attention. The universal occurrence of variability in natural object kinds, including species, will be described, showing first how it can confound the classification task, but also how it can be exploited to provide better solutions by using deep learning. In the plant identification domain, numerous studies have focused on procedures or algorithms that maximize the use of leaf databases, and this always leads to a norm that leaf features are liable to change with different leaf data and feature extraction techniques. Heretofore, we have been engaged with ambiguity surrounding the subset of features that best represent the leaf data. Hence, in the present study, instead of delving into the creation of feature representation as in previous approaches, reverse engineer the process by asking DL to interpret and elicit the particular features that best represent the leaf data. By means of these interpretation results, are able to perceive the cognitive complexities of vision for leaves as such, reflecting the trivial knowledge researchers intuitively deploy in their imaginative vision from the outset. And quantify the characteristics of features in each CNN layer and find that the network exhibits layer-by-layer transition from general to specific types of leaf feature. And find that this effect emulates the botanists' character definitions used for plant species classification.

A. Joly, et.al,...[3] builded accurate knowledge of the identity, geographic distribution and uses of plants is essential for a sustainable development of agriculture as well as for biodiversity conservation. Unfortunately, such basic information is often only partially available for professional stakeholders, teachers, scientists and citizens, and often incomplete for ecosystems that possess the highest plant diversity. A noticeable cause and consequence of this sparse knowledge, expressed as the taxonomic gap, is that identifying plant species is usually impossible for the general public, and often a difficult task for professionals, such as farmers or foresters and even for the botanists themselves. Speeding up the collection and integration of raw botanical observation data is a crucial step towards a sustainable development of agriculture and the conservation of biodiversity. Initiated in the context of a citizen sciences project, the main contribution of this paper is an innovative collaborative workflow focused on image-based plant identification as a mean to enlist new contributors and facilitate access to botanical data. Since 2010, hundreds of thousands of geo-tagged and dated plant photographs were collected and revised by hundreds of novice, amateur and expert botanists of a specialized social network. An image-based identification tool - available as both a web and a mobile application - is synchronized with that growing data and allows any user to query or enrich the system with new observations. An important originality is that it works with up to five different organs contrarily to previous approaches that mainly relied on the leaf. This allows querying the system at any period of the year and with complementary images composing a plant observation. Extensive experiments of the visual search engine as well as system-oriented and user-oriented evaluations of the application show that it is already very helpful to determine a plant among hundreds or thousands of species.

A. R. Sfar, et.al,...[4] studied fine-grained categorization, the task of distinguishing among sub-categories of a more basic category, such as an object or shape class, focusing on identifying botanical species from leaf images. Whereas people can usually immediately recognize instances from basic categories(trees, dogs, etc.), fine-grained categories (e.g., species of plants, breeds of dogs) are usually recognized only by experts. The difficulty arises because taxonomic categories often have very fine differences which are hard to notice for the common eye. Generally, the situation is the same in other domains of fine-grained categorization, and raises the question of what extent of semi-automation is required to provide useful results. To this end, we employ the

user, with the goal of achieving something sensible between the two extremes of an inaccurate but fully automated identification and a very accurate but fully-manual identification. The baseline scenario is the standard one with no human intervention: given an image of a leaf, usually scanned against a flat background, the system automatically provides a single estimate of the true species. Even with scanned leaves, the utility of this approach is questionable due to relatively high error rates on large databases which contain very similar species and display high variability within the same species. This motivates the design of semi-automated systems. One natural possibility is to envision human participation at the end of the process in the sense of final disambiguation. The implementation exploits domain-specific knowledge about landmarks and taxonomy to automatically build the hierarchical representation of species based on purely foliar characteristics. As indicated earlier, we also introduce different identification scenarios and tackle the problem of the cluttered leaf images without using segmentation algorithms. State-of-the art results are obtained in all cases in which comparisons with previous work are possible.

J. Chaki, et.al,...[5] implemented the system for Plants play a crucial role in Earth's ecology by providing sustenance, shelter and maintaining a healthy breathable atmosphere. Plants also have important medicinal properties and are used for alternative energy sources like bio-fuel. Building a plant database for quick and efficient classification and recognition is an important step toward their conservation and preservation. This is especially significant as many plant species are at the brink of extinction due to incessant de-forestation to pave the way for modernization. In recent year's computer vision and pattern recognition techniques have been utilized to prepare 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 here uses 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.

3. 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. Our goal with the Folia application is then to build a system for leaf shape analysis that processes, unlike what has been done to date, pictures in a natural environment. 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.

4. PROPOSED WORK

Even when considering trees only, leaves and fruits show an impressively wide variety in shapes. It is however necessary to come up with a representation, 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 or fruit. Botanists have a whole set of terms describing either the shape of a simple leaf or fruit, of the lobes of a palmate leaf or fruit, or of the leaflets of a compound leaf. The problem being that the borders between the different terms are not well defined, since leaves or fruits can naturally have non-canonical, intermediate shapes. The margin of the leaf and texture of fruit are 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 or fruits 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 or fruits. 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 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 or fruits, 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 project we introduce a method designed to deal with the obstacles raised by such complex images, for simple and lobed tree leaves and fruits. A first segmentation step based on a light polygonal leaf or fruit model is first performed, and later used to guide the evolution of an Otsu there sholding. Combining global shape descriptors given by the threshold model with local curvature-based features, the leaves and fruits are then classified over datasets. In this project we introduce a method designed to deal with the obstacles raised by such complex images, for simple. A first segmentation step based on Otsu threshold method is first performed, and later used to guide the evolution of affected boundaries. And implement convolutional neural network classification algorithm to classify the fruit and leaf diseases. The proposed framework is shown in fig 4.



Fig 4. Proposed work

4.1 IMAGE ACQUISITION

Leaves and fruits 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 or fruit images from the datasets. We can input any size of image and any type. This module we can input apple, grapes, pomegranate fruits and leaves images.

4.2 PREPROCESSING

In this module convert the RGB image into gray scale image. The colors of leaves and fruits are always in RGB shades and the variety of changes in atmosphere cause the color feature having low reliability. Therefore, to recognize various plants using their leaves and fruits, the obtained 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. The median filter process as follows:

4.3 IMAGE SEGMENTATION

In this module, we can implement Otsu threshold with automatic descriptors. Image segmentation is a convenient and effective method for detecting foreground objects in images with stationary background. Background subtraction is a commonly used class of techniques for segmenting objects of interest in a scene. This task has been widely studied in the literature. Specular reflections, background clutter, shading and shadows are the major factors that affect the efficiency of the system. Therefore, in order to reduce the scene complexity, it might be interesting to perform image segmentation focusing on the object's description only. In the first step, using the Otsu's method Segmentation operation has performed RGB channel by taking threshold value. The value of threshold has been decided on the basis of the extensive image analysis method. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal.

4.4 DISEASE PREDICTION

Leaves and fruits are affected by bacteria, fungi, virus and other insects. CNN has significant application in fruits and leaf quality evaluation. The purpose of ANN is to generate a network system with little errors but also yield good result from the testing data set. In this module implement artificial neural network algorithm to classify the leaf and fruit image as normal or affected. Vectors are constructed based leaf features such as color, shape, textures.

Convolutional neural networks (CNN) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains. The neural network itself isn't an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs. Such systems "learn" to perform tasks by considering examples, generally without being programmed with any task-specific rules. CNN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it. In common CNN implementations, the signal at a connection between artificial neurons are a real number, and the output of each artificial neuron is computed by some non-linear function of the sum of its inputs. The connections between artificial neurons are called 'edges'. Artificial neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that the signal is only sent if the aggregate signal crosses that threshold.

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

Based on these steps leaves and fruits are classified with disease names with improved accuracy rate.

5. CONCLUSION

In this paper, we overview the various techniques and algorithms are proposed for segmentation and classification methods for improve the quality of segmentation. 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 otsu threshold segmentation. 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. The use of an additional texture model or of an adaptive color model could lead to a good improvement. Finally implement convolutional neural network classification algorithm to classify the diseases in leaves and fruits such as apple, grapes and pomegranate. We can extend the framework to implement various deep learning classification algorithms and also classify the diseases not only leaves and fruits also in various vegetables with improved accuracy rate.

REFERENCES

[1]. J. Wäldchen and P. Mäder, "Plant species identification using computer vision techniques: A systematic literature review," Archives of Computational Methods in Engineering, pp. 1–37, 2017.

[2]. S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, "How deep learning extracts and learns leaf features for plant classification," Pattern Recognition, vol. 71, pp. 1–13, 2017.

[3]. A. Joly, H. Goëau, P. Bonnet, V. Baki'c, J. Barbe, S. Selmi, I. Yahiaoui, J. Carré, E. Mouysset, J.-F. Molino et al., "Interactive plant identification based on social image data," Ecological Informatics, vol. 23, pp. 22–34, 2014.

[4]. A. R. Sfar, N. Boujemaa, and D. Geman, "Confidence sets for finegrained categorization and plant species identification," International Journal of Computer Vision, vol. 111, no. 3, pp. 255–275, 2015.

[5]. J. Chaki, R. Parekh, and S. Bhattacharya, "Plant leaf recognition using texture and shape features with neural classifiers," Pattern Recognition Letters, vol. 58, pp. 61–68, 2015.

[6]. N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares, "Leafsnap: A computer vision system for automatic plant species identification," in Computer Vision–ECCV 2012. Springer, 2012, pp. 502–516.

[7]. Y. Naresh and H. Nagendraswamy, "Classification of medicinal plants: An approach using modified lbp with symbolic representation," Neurocomputing, vol. 173, pp. 1789–1797, 2016.

[8]. G. L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, "Deep learning for plant identification using vein morphological patterns," Computers and Electronics in Agriculture, vol. 127, pp. 418–424, 2016.

[9]. J. Charters, Z. Wang, Z. Chi, A. C. Tsoi, and D. D. Feng, "Eagle: A novel descriptor for identifying plant species using leaf lamina vascular features," in Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on. IEEE, 2014, pp. 1–6.

[10]. M. G. Larese, R. Namías, R. M. Craviotto, M. R. Arango, C. Gallo, and P. M. Granitto, "Automatic classification of legumes using leaf vein image features," Pattern Recognition, vol. 47, no. 1, pp. 158–168, 2014.

Conflict of Interest

None of the authors have any conflicts of interest to declare.

About the License

The text of this article is licensed under a Creative Commons Attribution 4.0 International License