IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

PLANT LEAF DISEASE IMAGE DETECTION AND CLASSIFICATION USING ANN

Mrs.A.Anitharani [1], Assistant Professor/EC

Department of Electronics and Communication Engineering

Excel Engineering College (Autonomous), Komarapalayam, Namakkal, India

A.Vasanth [2], C.Hariharan [3], V.Maheswaran[4], S.Pranesh[5]

UG Students, Department of Electronics and Communication Engineering

Excel Engineering College (Autonomous), Komarapalayam, Namakkal, India

ABSTRACT

In the emerging technology plant disease detection are important and the source of energy implies to human beings. Plant disease detection can influence the plant at convenience which leads to major loss on the production of crops and inexpensive value in the trade market. Therefore, plant disease detection plays a significant role in agricultural field. However, it requires huge manpower, more processing time and extensive knowledge about plant diseases. Since, machine learning technique under supervised learning like SVM, Gaussian Naive bayes, Random forest is consider detecting diseases in plant leaves as it analyzes the data from different aspects, and classifies it into the predetermined set of classes. The topological features are classified based on color, potency and measurements of the plant leaves are taken into consideration. This paper confer an overview on various types of plant disease detection and different grouping techniques in machine learning of CNN model that are used for recognizing diseases in different plant leaves using smart farming .

INTRODUCTION

Digital media archives are increasing to colossal proportions in the world today, which includes audio, video and images An Image refers as a picture produced on an electronic display .A digital image is a numeric representation of a two-dimensional image. Digital image processing refers to processing of digital images by using digital computers. Nowadays, most of the applications prefer digitalized version, to reduce memory space. Lot of application depends on digital images. One of the important application is medical image processing.

OBJECTIVE

a) To detect unhealthy region of plant leaves

b) Coding is used to analyze the leaf infection

c) Classification of plant leaf diseases using ANN

EXISTING SYSTEM

The Feature Discontinuity extracts the regions having different properties like intensity, color, texture etc. Similarity groups the image pixels into groups with some predefined criteria. PCA Based on pixel similarity with the neighboring pixel, the algorithm used is region based. In leaf disease identification, segmentation is used to identify the diseased area. From this, features of a region are computed; we have to extract the features corresponding to the disease in this system Not Clearly Recognition Leaf Disease Result.

LITERATURE SURVEY

Agriculture is one of the most important sources for human sustenance on Earth. Not only does it provides the much necessary food for human existence and consumption but also plays a major vital role in the economy of the country. Disease is caused by pathogen which is any agent causing disease. In general, there are two types of factors which can bring death and destruction to plants; living (biotic) and nonliving (a biotic) agents. Living agent's including insects, bacteria, fungi and viruses. Nonliving agents include extremes of temperature, excess moisture, poor light, insufficient nutrients, and poor soil pH and air pollutants.

In this project, we have used ANN classifier to identify the pest & type of disease in cotton plant. Image acquisition devices are used to acquire images of plantations at regular intervals. These images are then subjected to preprocessing. The pre-processed leaf images are then segmented using K-means clustering method. Then the color features(mean, skewness), texture features such as energy, entropy, correlation, contrast, edges are extracted from diseased leaf image using gray level co-occurrence matrix (GLCM) in the texture & then compared with normal cotton leaf image. The Artificial neural networks (ANN) classifier is used to classify the pest and Disease in cotton crop.

The diseases on the cotton leaves are classified as,

a) Bacterial disease: e.g. Bacterial Blight, Crown Gall, Lint Degradation
b) Fungal diseases: e.g. Anthracnose, Leaf Spot
c) Viral disease: e.g. Leaf Curl, Leaf Crumple, Leaf Roll

THRESHOLDING

Thresholding technique is effective for images based on objects of contrast background. In thresholding technique, each pixel in images are separates into foreground (binary "1") and background (binary "0") object of classes based upon their correspondence in gray level intensity. The drawback of thresholding is not always straight forward and assigned pixels to a single class.

EDGE DETECTION

Edge Detection is a general technique that operates on an image and results in a line drawing of the image. The lines represent changes in values such as intersections of planes, cross sections of planes, textures, colors, and lines as well as differences in shading and textures. The main purpose of edge detection is to identify areas of an image where a large change in intensity occurs. In edge detection, canny edge detector is mostly used to detect wide range of edges of images. Canny detection technique has three steps i.e. first, it clears the noise in image and second it obtains gradients of intensity and direction. Last, it determines strong edges of image and also finds begin and end of edges by threshold value. This detection technique is better because, it reduces noise. Edge detection technique is used to extract the image features (color, shape, texture, etc).

DRAWBACKS

a. Noise High
b. Output Not Clearly Recognition Leaf Disease
c. To identify the disease in plant is less accurate

PROPOSED SYSTEM

In the proposed work, we have concentrated on identification of Leaf Spot disease and Leaf Miner from the photographic signs and classify them using image processing techniques. The proposed framework has been implemented in three steps. First, image segmentation is performed using K means clustering to identify the infected area. In the next step leaf features are extracted from segmented regions using feature extraction techniques such as GLCM. These features are then used for classification into infected or non-infected leaf type. As third step these features given to the classifier to classify the disease in the cotton crop. We used ANN classifier to obtain efficient results.

DESCRIPTION

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analog image processing; it allows a much wider range of algorithms to be applied to the input data, and can avoid problems such as the build-up of noise and signal distortion during processing.

PLANT DISEASE DETECTION: A PRAGMATIC REVIEW

India is an agriculture based country. In this context agriculture plays vital role in Indian economy. 58 percent of rural people depend on agriculture as their principle means of livelihood. It is an important source of raw materials for many agro-based industries. There are many causes for plant diseases which affect the yield and hence the economical condition of both farmers and in turn the entire nation as whole. Plant diseases can be infectious or noninfectious. Noninfectious diseases are usually referred to as disorder caused by various causes such as nutrient deficiency, by waterlogged or polluted soil, and by polluted air from industry or excessive use of herbicides and pesticides. Infectious plant diseases (usually not visible to the naked eye) are caused by pathogens, living microorganisms that infect a plant and deprive it of nutrients.

The common approach used by the farmers to identify the plant diseases is by consulting the experts which is a very tedious, consumes more time and money. In addition there are some diseases which cannot be identified by the naked eyes. Hence in this context, a fast, reliable and automatic method is required to accurately identify the plant diseases. This paper provides the review of image processing techniques applied for identifying plant diseases detection. The figure represents the general image processing steps adopted for plant disease detection using image processing.

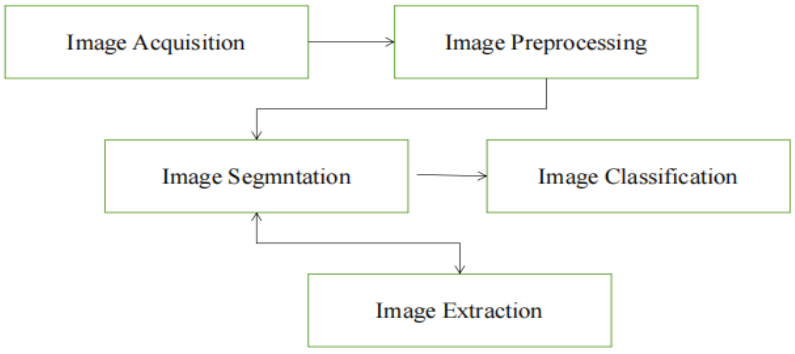


Fig. General steps in Digital Image Processing

Image Acquisition: The digital images are acquired using a digital mobile camera or digital camera and given as input to the identification system. This is the image in which the leaf disease has to be identified by the system

Image Preprocessing: It is the technique for improving picture quality prior to computational processing and also used to remove the low frequency noise, reflections and masking portions of the images.

Image Segmentation: segmentation is the process of partitioning the digital image into multiple parts/segments. It helps in simplifying the image into more meaningful and easier to analyze

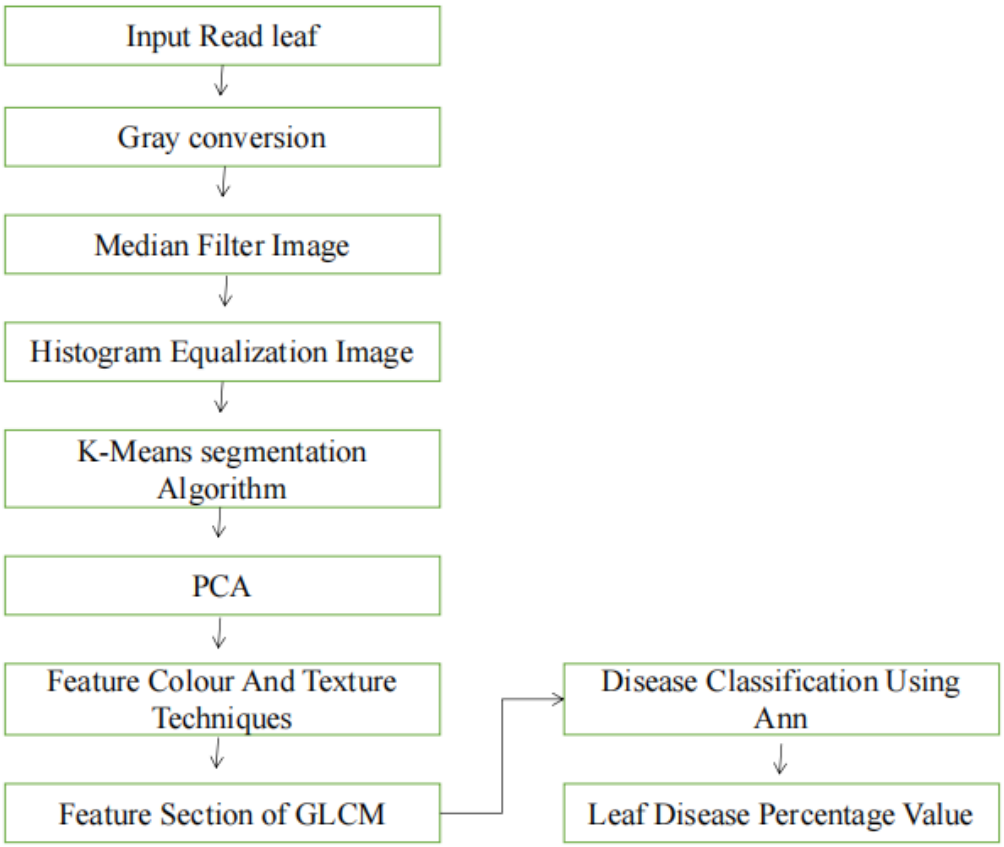
Feature extraction: After the segmentation process, various features are extracted from the infected region .The features which can be used in plant diseases detection are color, texture and shape etc.

Image Classification: It is most important part in digital analysis. Classification can be executed on spectral features like density, texture etc and then divides the features space into many classes using different machine learning algorithms.

Image Acquisition of Diseased Plants

The RGB color images of most frequently encountered Phyto-pathological problems affecting Cotton leaves were captured using camera. Images were stored in.JPG format

IMPLEMENTATION PROCESS



INPUT – RGB IMAGE

RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue. The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors.

MEDIAN FILTER

Median filtering is used as a noise removal in order to obtain a noise free image. After segmentation is done, the segmented image may still present some unwanted regions or noise. So to make the image a good and better quality, the median filter is applied to the segmented image. We can use different neighborhood of n × n. But generally neighborhood of n = 7 is used because large neighborhoods produce more severe smoothing.

BOUNDARY PROCESSING

The first value is repeated, as with the last value, to obtain enough entries to fill the window. This is one way of handling missing window entries at the boundaries of the signal, but there are other schemes that have different properties that might be preferred in particular circumstances:

- Avoid processing the boundaries, with or without cropping the signal or image boundary afterwards
- Fetching entries from other places in the signal. With images for example, entries from the far horizontal or vertical boundary might be selected
- Shrinking the window near the boundaries, so that every window is full

IMAGE PRE-PROCESSINGAND SEGMENTATION

The pre-processing involved the procedures to prepare the images for subsequent analysis. The affected lea images were converted from RGB color format t gray scale images. Segmentation refers to the process of clustering the pixels with certain properties into salient regions and these regions correspond to different faces, things or natural parts of the things. We proposed k-means segmentation technique to fragment goal areas. Target regions are those areas in the image that represented visual symptoms of a fungal disease.

K-MEANS SEGMENTATION

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

K-MEANS CLUSTERING ALGORITHM

Clustering is a method to divide a set of data into a specific number of groups. It's one of the popular method is k-means clustering. In k-means clustering, it partitions a collection of data into a k number group of data. It classifies a given set of data into k number of disjoint cluster. K-means algorithm consists of two separate phases.

In the first phase it calculates the k centroid and in the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance. Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new Euclidean distance is calculated between each center and each data point and assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid.

The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters.

Let us consider an image with resolution of x×y and the image has to be cluster into k number of cluster. Let p(x, y) be an input pixels to be cluster and ck be the cluster centers. The algorithm for k-means13 clustering is following as:

1. Initialize number of cluster k and centre.
2. For each pixel of an image, calculate the Euclidean distance d, between
the center and each pixel of an image using the relation given below.
3. Assign all the pixels to the nearest centre based on distance d.
4. After all pixels have been assigned, recalculate new position of the centre using the relation given below.

$$Ck = \frac{1}{k} \sum_{y \in C_k} \sum_{y \in C_k} p(x, y) \qquad \text{(Eq.1.1)}$$

5. Repeat the process until it satisfies the tolerance or error value.
6. Reshape the cluster pixels into image. Although k-means has the great advantage of being easy to implement, it has some drawbacks. The quality of the final clustering results is depends on the arbitrary selection of initial centroid. So if the initial centroid is randomly chosen, it will get different result for different initial centers. So the initial center will be carefully chosen so that we get our desire segmentation. And also computational complexity is another term which we need to consider while designing the K-means clustering. It relies on the number of data elements, number of clusters and number of iteration.

PRINCIPLE COMPONENTANALYSIS (PCA)

To reduce the dimension of input vector of neural network, PCA is used to orthogonalize 12 features. The purpose of PCA is to present the information of original data as the linear combination of certain linear irrelevant variables. Mathematically, PCA transforms the data to a new coordinate system such that the greatest variance by any projection of the

data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on. Each coordinate is called a principal component.

FEATURE EXTRACTION

The symptoms associated with various Phyto-pathological problems of cotton leaves under investigation visible on the affected leaves were extracted from their respective images using K-mean. The image analysis was mainly focuses on the extraction of shape features and their color based segmentation. The image analysis technique is done using Gray-level co-occurrence matrix. The affected areas vary in color and texture and are dominant in classifying disease symptoms. So, we have considered both color and texture features for recognition and classification purpose. Picture texture, explained as a function of the spatial variation in pixel intensities (gray values). The use of color features in the noticeable light spectrum provided additional image characteristic features over traditional gray-scale representation. GLCM is a method in which both color and texture features are taken into account to arrive at unique features which represent that image.

STATISTICALANALYSIS

Statistical analysis tasks are completed to choose the best features that represent the given image, thus minimizing feature redundancy. We have found that only 13 features contribute as discriminating features as this is essential for better classification.

Magnitudes that are workable to guess via the co-occurrence matrix are:

- Energy
- Entropy
- Homogeneity
- Contrast
- Mean
- Standard Deviation
- RMS
- Variance
- Smoothness
- Kurtosis
- Skewness
- IDM and
- Correlation

ANN CLASSIFICATION

At present ANN is popular classification tool used for pattern recognition and other classification purposes. Artificial neural networks (ANN) are a group of supervised learning methods that can be applied to classification or regression. The normal ANN classifier takes the set of involvement data and calculates to classify them in one of the only two separate classes. ANN classifier is trained by a given set of training data and a model is willing to classify test data established upon this model. Most habitual classification models are established on the empirical risk minimization principle.

ANN implements the structural risk minimization principle which pursues to reduce the training error and a sureness interval term. A number of submission showed that ANN hold the superior classification capability in production with minor sample, nonlinearity and high dimensionality pattern identification. Support Vector Machines are based on the concept of decision planes that define decision boundaries. A decision plane is one that splits among a set of objects having different class association. Classifier that separate a set of objects into their corresponding classes with a line. Supreme classification tasks, however, are not that modest, and regularly more difficult structures are needed in order to make an optimal separation, i.e., correctly classify new objects (test cases) on the basis of the examples that are available (train cases). All the evidence from beyond processes is given to multiclass ANN .The Multiclass ANN were used for cotton disease classification.

ANN ALGORITHM

The presents a methodology for early and accurately plant diseases detection, using artificial neural network (ANN) and diverse image processing techniques. As the proposed approach is based on ANN classifier for classification and Gabor filter for feature extraction, it gives better results with a recognition rate of up to 91%. An ANN based classifier classifies different plant diseases and uses the combination of textures, color and features to recognize those diseases.

ACCURACY

In pattern recognition and information retrieval with binary classification, precision (also called positive predictive value) is the fraction of retrieved instances that are relevant, while recall (also known as sensitivity) is the fraction of relevant instances that are retrieved**.**

PRECISION

In the field of information retrieval, precision is the fraction of retrieved documents that are relevant to the query:

RECALL

Recall in information retrieval is the fraction of the documents that are relevant to the query that are successfully retrieved. Recall in this context is also referred to as the true positive rate or sensitivity, and precision is also referred to as positive predictive value (PPV); other related measures used in classification include true negative rate and accuracy.[6] True negative rate is also called specificity.

$$\text{True negative rate} = \frac{tn}{tn + fp}$$

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

RESULTAND DISCUSSION

Identified Phyto-pathological problems experiments modules are developed using MATLAB R2014a, which runs in the environment Windows 7,8,10. Two species of samples are taken for the experiment, whose digital images are obtained by a camera. FIG shows the species type and numbers of leaves images for these species for In pattern recognition and information retrieval with binary classification, precision (also called positive predictive value) is the fraction of retrieved instances that are relevant, while recall (also known as sensitivity) is the fraction of relevant instances that are retrieved**.**

SCHEME	**CA**	**PCA+GLCM+ ANN**
ACCURACY	1%	92%

OUTPUTFOR THE PADDY LEAF=92%

APPLICATION

A. To diseased leaf, stem, fruit

B. To determine affected area by disease

C. To find shape of affected area

D. To evaluate color of affected area

CONCLUSION

In this paper we describe our work concerned with the discrimination between healthy and diseased to cotton crops using an ANN. In this paper, respectively, the applications of K-means clustering have been formulated for clustering and classification of diseases that affect on plant leaves. Identifying the disease is generally the drive of the proposed method. By using segmentation technique it is easy for us to extract the features of disease leaf of the image. A new approach based on K-mean features extraction was proposed for cotton leaf recognition in this work. The whole process of leaf classification can be implemented In this study, extracted texture features was proposed and performed. The texture features have been extracted with using the Gray-Level Co-occurrence Matrix (GLCM) and the Principal Component Analysis (PCA) algorithms, using leaf detection, feature extraction and classification. For studying the proposed method, the composed dataset is used. The dataset contains diseased images. Images were preprocessed and cropped to a fixed standard size. Then, features are extracted from all the leaf images in the dataset using K-mean algorithm. For each image leaf more frequent K-mean key points are extracted to identify a unique feature. It permits finding related features for different image. Ultimately, the extracted K-mean and GLCM features are rendered to a ANN classifier for purpose of classification. In other words, differences between diseased and non-diseased leaves and the key points which are extracted from leaf are used for classifying Accuracy Using Precision and Recall Value Analysis Detection Result.

FUTURE SCOPE

For future study, different neural network architectures can be used for classification. We can extend this project to classify disease symptoms affected on fruits, vegetables, commercial crops etc.,we may work for better application like we develop a site where any person can upload their image they will find out there diseased and full detail about the disease. What they do for their fields and crops. What is the advantage and disadvantage of this disease and what should do to control it.

REFERENCE

[1] Sanjay B.Dhaygude,Mr.Nitin P.Kumbhar,(2013)“Agricultural plant leaf disease detection using image processing” ,vol.2,issue 1pg.no. 599-602.

[2] Ghulam Mustafa Choudhary,Vikrant Gulati,(2015)“Advance in image processing for detection of plant diseases”, vol 5.issues 7,pg.no.295-301.

[3] Y.Sanjana,AshwathiSivasamy and SriJayanth,(2015)“Plant disease detection using image processing techniques”,vol.4,special issue 6.

[4] Smita Naikwadi,Niket Amoda,(2013)“Advances in image processing for detection of plant diseases”,vol.2,issue11, pg.no.168-175.

[5] Jayamala K.patil,Raj Kumar,(2011)“image processing for detection of plant diseases”,vol 2,issue2,pg.no.136-141.

[6] Dae Gwan Kim, Thomas F. Burks, Jianwei Qin, Duke M. Bulanon,(2009) “Classification of grapefruit pee diseases using color texture feature analysis”, International Journal Agriculture and Biological Engineering, Vol:2,No:3, pp 41-50.

[7] J. G. Arnal Barbedo,(2013) “Digital image processing techniques for detecting, quantifying and classifying plant diseases,” SpringerPlus, vol. 2, article 660, pp. 1–12.

[8] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis,(2010) “A review of advanced techniques for detecting plant diseases,” Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1–13.

[9] P. R. Reddy, S. N. Divya, and R. Vijayalakshmi, (2015)“Plant disease detection techniquetool—a theoretical approach,” International Journal of Innovative Technology and Research, pp. 91–93.

[10] A.-K. Mahlein, T. Rumpf, P. Welke et al.,(2013) “Development of spectral indices for detecting and identifying plant diseases,” Remote Sensing of Environment, vol. 128, pp. 21–30.

[11] W. Xiuqing, W. Haiyan, and Y. Shifeng,(2014) “Plant disease detection based on near-field acoustic holography,” Transactions of the Chinese Society for Agricultural Machinery, vol. 2.