Farmer Helping System to Diagnose Plant Leaf Diseases

Prajakta Mahendra Sonawane[1] Pratiksha Amrut Kale[2] Nivedita Harishchandra Dahima[3] Gauri G.Bilaye[4]

[1,2,3]Student [4]Assistant Professor

[1,2,3,4]Department of Information Technology

[1,2,3,4]SKNCOE Pune, Maharashtra, India

Abstract— In a research of the diagnosing disease, the pattern of disease is an important part. Many diseases are occurring on the leaf, so its color for different diseases is also different. There are various other features related to the shape of an image such as holes are present on the leaf are of an elliptical shape. Farmers need information about fertilizers, pesticides, information of atmosphere, etc. many things. Today this information is available in a scattered manner, and it is not providing proper diagnosis about various diseases. We can detect it by Image analysis. Image analysis by using Content- based Image Retrieval(CBIR) is an important method that helps segment image into objects and background, the key steps in image analysis is feature detection. Our Android application can provide the exact differentiation between the variation of color and texture present on these leaves. Depending upon that variation the further comparison with database stored image features is based on the color and texture.

Key words: Texture, Content-Based Image Retrieval

I. INTRODUCTION

India is the agricultural country where most of the population depend on it. The main focus is on increased productivity and food quality at reduced expenditure. Vegetables and fruits a most important agricultural product. The agricultural study shows that due to plant diseases quality of the product is reduced. These diseases are caused by fungi, viruses, bacteria and adverse environmental conditions. Therefore, at the initial stage plant disease diagnosis is an important task[1]. For that continuous monitoring of experts is required which might be time-consuming. Therefore fast, less expensive and accurate method to automatically detect the diseases from the symptoms that appear on the plant leaf is a necessity. This paper mainly focuses on cotton plant and grapes plant leaf disease detection using CBIR (Content Based Image Retrieval) technique with texture and color of the leaf. CBIR is the retrieval of images based on visual features such as color, texture, and shape. To reduce time consuming and insufficient indexing by using traditional methods for image indexing, the CBIR is developed. In this, each image stored in the database has its features extracted and compared to the features of the query image[2].

II. RELATED WORK

There are many users attracted towards databases of satellite and medical imagery in various professional fields — for example, geography, medicine, architecture, and publishing. Effectively accessing desired images from large and varied image databases is now a necessity. CBIR or Content-Based Image Retrieval is the retrieval of images based on visual features such as color, texture, and shape. In this paper, features such as color and texture are used. Traditional methods of image indexing have proven to be insufficient, and extremely time-consuming. In CBIR, the features of the query image are compared with the extracted features of the database stored image[2].

It involves two steps:

A. Feature Extraction:

The first step in the process is feature extraction. Global features are extracted to make the system more efficient. Feature extraction for the image description is performed. We classify the various features as follows-

1) Color Features

a) Color:

In CBIR System, Color information of an image is represented by color histograms. A color histogram is one of a type of bar graph which is constructed by counting some pixels of each color in the RGB or HSV color space. But one disadvantage of RGB color model is its behavior when illumination in an image changes. On the X-axis, bins are represented. Bars in a color histogram are referred as bins. The number of bins is related to the number of colors in an image. The number of pixels in each bin is denoted by Y-axis. It gives the count of pixels in an image representing a particular color. Quantization is the process used in the reduction of color histogram .The process reduces the number of bins by taking colors that are similar to each other and putting them in the same bin. Quantization does not consider any actual distributions in a given image databases. In Mat Lab, using histogram function, by default the maximum number of bins one can obtain is 256.The color histogram gives the intensity of each color of the image.

The used color descriptor is composed of the following attributes

$$\text{Color Expectancy Ei} = \frac{1}{N}\sum_{j=1}^{N} Pij \qquad (2.1)$$

$$\text{Color Variance } \delta i = \left(\frac{1}{N}\sum_{j=1}^{N}(PijEi)^2\right)^{1/2} \qquad (2.2)$$

$$\text{Skewness } \sigma i = \left(\frac{1}{N}\sum_{j=1}^{N}(PijEi)^3\right)^{1/3} \qquad (2.3)$$

Where Pij is the (i; j) pixel color; N is the total number of pixels in the image.

Color expectancy allows estimating the average color; Color variance shows the dispersion of color values from the average color and skewness is related to the symmetry of their distribution on the whole image.

2) Texture Features

a) Texture characterization based on GLCM:

For the computation of the GLCM not only the displacement but also the orientation between neighbor pixels must be established. The orientations can be horizontal (0º), vertical (90º)[3].Commonly used GLCM texture features are shown in the following:

Matrix elements are computed by the equation showed as follow GLCM expresses the texture feature according to the correlation of the couple pixels gray-level

All rights reserved by *www.ijsrd.com*

at different positions. It quantification all describes the texture feature. In this paper, four features are selected, include energy, contrast, entropy, inverse difference

Energy E= $p(x, y)$ ^ 2 (2.4)

Energy is a measure of homogeneity. It is opposite to the entropy. This feature will tell us texture uniformity. Higher Energy value shows the bigger homogeneity of texture. The range of it is [0, 1], where Energy is 1 for a constant image.

Contrast I = ΣΣ(x − y) 2 p(x, y) (2.5)

Contrast is the local grey level variation in grey level co-occurrence matrix. The contrast of the image is low when the neighboring pixels are similar in their grey level values.

Low contrast values are for smooth texture while high values for heavy texture. The range of Contrast is [0, (size (GLCM,1)-1)^2] where Contrast is 0 for constant image.

Entropy S= ΣΣ p(x, y) log p(x, y) (2.6)

Entropy in any system represents disorder. It measures image texture randomness. A completely random distribution would have very high entropy because it represents disarray. This feature will tell us for heavy textures entropy is bigger or for the smooth textures giving us information about which type of texture can be considered statistically more disarray.

B. Matching:

The second step involves matching. The most common method for comparison of two images in CBIR is by using image distance measure. In CBIR while building an image database, features from are to be extracted and then store in another database for future use. When given a query image its features are computed. An image distance measure compares the similarity between the query image and the database stored an image in the form of color, texture, region, shape. If the distance between feature vectors of the query image and image in the database is small enough, the corresponding image in the database is matched to the query. Though there is not an exact match but the search is usually based on similarity the retrieval results are then ranked accordingly to a similarity index[3].

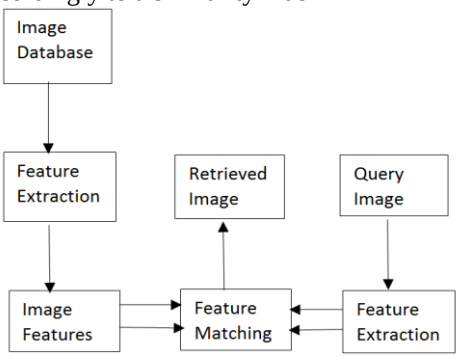

Fig. 1: Block diagram of CBIR system

III. PROPOSED METHOD

The proposed system is divided into following modules:

A. Web-based GUI:

The server will be the web-based application, and this module will be responsible for taking inputs from admin. The GUI will be developed in HTML and JavaScript.

B. Android Client:

Android application will be needed for the farmer to provide input and to see the output on his device. Farmer will send the image of the plant leaf to a server.

This module will take care of camera capturing activity that needs to be performed to take an image.

C. Color Feature Extraction:

Color feature extraction represents the color information of the image with a histogram. The intensity of each color is given by color histogram. A three-dimensional RGB (8*8*8) histogram contains total 512 bins. At the time of image indexing, the color of each pixel is found out, and its corresponding bin's count incremented by one[3]. As we already discussed some of the attributes such as color expectancy, color variance, skewness are calculated to find the color intensity of an image.

D. Texture Feature Extraction:

GLCM is used for texture feature extraction. It contains information about the frequency of occurrence of two neighboring pixel combination in an image. Although 22 features can be derived from GLCM, usually only some of them are considered as parameters of importance: Contrast, Energy, and Entropy. By calculating this texture features it is possible to see how they behave for different textures.

E. Matching Logic:

Euclidean Distance method is used to find out distance measures. These distance measurements indicated the similarities. The low value of distance measurement represent the close (good) similarity relation otherwise, the high value of distance measurement represent the open (bad) similarity relation between two images. The metric used in the Euclidean distance is called the Euclidean metric to find out the distance measurements.

F. Database Manager:

All database related activity are handled by this module. All the SQL queries will be taken care in this module. A database connection polling system will be present to avoid repeatedly opening and closing database connection.

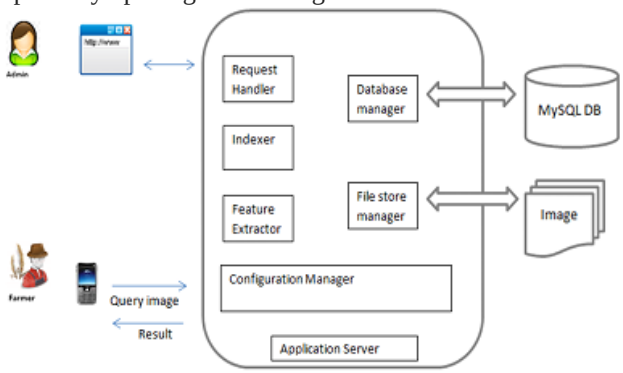

Fig 2: Proposed System Architecture

IV. ALGORITHM

A. Collection of Image Database

- We consider a database containing 100 images with any one of the formats .bmp, .jpg or .tiff.
- The images will be from RGB color model.

All rights reserved by *www.ijsrd.com*

B. Feature Extraction

- Feature Extraction is performed by using colors, using textures.
- The images are registered with their corresponding features such as color, texture.
- These extracted features will be forwarded to Feature Vector Module

C. Similarity Measures

The Direct Euclidian Distance between an image P and query image Q can be given as the
Equation below

$$ED=\Sigma (Vpi - Vqi). (Vpi - Vqi).$$

Where, Vpi and Vqi be the feature vectors of image P and Query image Q respectively with size 'n'

D. Comparison of results with other techniques

Distance measure is used for comparing the similarity between query image and image stored in database.

E. Finally, the image will be retrieved.

V. CONCLUSION

This paper proposes a farmer helping system to diagnose plant diseases using Content-based image retrieval technique. It is based on visual features such as color and texture. The system will provide a solution for the diseases and also find the factors responsible for causes of diseases. The prevention measures include various types of fertilizers, pesticides which we have to use for plants. The application will provide an efficient solution in less time so usability of the system will be more.

VI. FUTURE SCOPE

In future work application will be used for different plant diseases. This application will give information of various fields of farming. It also gives information about a type of soil, modern techniques of farming.

In future, by using the same technique, we will also implement the system for checking the quality of seed so the farmer will use a better quality of seed for farming.

REFERENCES

[1] Ms. Kiran R. Gavhale, Prof. Ujwalla Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. V (Jan. 2014), PP 10-16
[2] P. Revathi, M. Hemalatha, "Cotton Leaf Spot Diseases Detection Utilizing Feature Selection with Skew Divergence Method" International Journal of Scientific Engineering and Technology (ISSN: 2277-1581) Volume No.3, Issue No.1 pp: 22-30 1 Jan 2014
[3] Nitin Jain & Dr. S. S. Salankar," Color & Texture Feature Extraction for Content Based Image Retrieval" IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) e-ISSN: 2278-1676, p-ISSN: 2320-3331 PP 53-58,2014.
[4] Dharani, I. Laurence Aroquiaraj, "Content-Based Image Retrieval System Using Feature Classification with Modified KNN Algorithm", T. at International Journal of Computer Trends and Technology 2013
[5] Shafimirza, Dr.J.Apparao, "Retrieval of Digital Images Using Texture Feature with Advanced Genetic Algorithm" International Journal of Computer Trends and Technology- volume3Issue4- 2012
[6] P. Revathi, M. Hemalatha, "Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques", IEEE International Conference on Emerging Trends in Science, Engineering and Technology, pp-169-173, Tiruchirappalli, Tamilnadu, India, 2012
[7] Neetu Sharma**,** Paresh Rawat and Jai Karan Singh, "Efficient CBIR Using Color Histogram Processing", An International Journal(SIPIJ) Vol.2, No.1, March 2011
[8] Arti Khaparde, B. L. Deekshatulu, M.Madhavilath, Zakira Farheen, Sandhya Kumari V, "Content-Based Image Retrieval Using Independent Component Analysis", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.4, April 2008.

All rights reserved by *www.ijsrd.com*