ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

DISEASE DETECTION ON COTTON LEAVES AND ITS POSSIBLE SOLUTIONS

Ms Swarupa Kamble*, Atul V. Kondekar, Swapnil Mane, Mukul Wanjare, Vinay Kalambate
Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

ABSTRACT

The application takes image as input and gives the related information about the disease and its possible diagnosis. Image analysis is one important method that helps segment image into objects and background. Content Based Image Retrieval (CBIR) helps in the retrieval of image data from a huge collection. Image representation based on certain features which helps in retrieval process. There are three important visual features of an image include Color, Texture and Histogram. The input image can be either uploaded on the website or the android application.

KEYWORDS: Image processing, Color feature extraction, Histogram based extraction, Texture feature extraction, image database.

INTRODUCTION

Today most of the farmers know how to use an android phone or access the web, even though they are not so well educated. The farmers who plant cotton plants in their farms may not know about its diseases and precautions to take, so this project helps the farmers to easily detect and diagnosis it. In a research of identifying and diagnosing cotton disease, the pattern of disease is crucial part in that, features of the images are extracted, the image should be the actual infected image, there are many diseases occurred on the cotton leaves so the color of leaf for different diseases is also different and also there are various other features related to shape of image, there are different shape of holes and spots are present on the leaf of the image, generally the leaf of infected image have elliptical shape of holes or spots.

About 80 to 90 % of disease on the cotton plant is on its leaves. For that study of interest is the leaves of the cotton tree not the whole cotton plant, the cotton leaf is mainly suffered from diseases like foliar leaf spot of cotton, fungus, Alternaria leaf spot of cotton.

Content Based Image Retrieval (CBIR) is approach that helps in the retrieval of image data from the Database. Image representation based on features helps in retrieval process. There are three important visual features of an image include Histogram, Color and Texture. The input image can be either uploaded on the website or the android application.

RELATED WORK

Some of the image retrieval databases are maintained below:

A) BDLP (Berkeley digital library project)

Text strings are used as storage medium for image feature. For instance, an image of a sky with clouds might have little large white section and a large amount of blue, so it would have a text string attached to the image as mostly blue large white few. Matching is done by using the substring matching for example in the above query string -large white%. The retrieved photos are presented unordered, photographer, id-number, and collection keys.

B) FIDS (Flexible image database system)

In this the L1 distances are the distances between the histograms. The weighted difference is the distance between wavelet coefficients of the images. By taking the weighted sum, minimum or maximum of the individual feature distances, which have metric properties, an overall distance can be composed. It also depends on the user to select how many of those the true distance must be calculated.

C) AMORE (Advanced multimedia oriental retrieval engine)

C Research Laboratories NEC USA had developed it. Then the exact work initially an association among regions in the target and query image is found. The regions associated with the same regions in the other image are fused together. The histogram resemblance among those two regions is actually based on the number of pixels which are overlapping, a type of pattern matching. Without an explicit order, the retrieved images are revealed as thumbnails. The obtained result images were displayed as a scatter plot, along with their histogram and color similarity values at the axes.

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

PROPOSED WORK

Disadvantages of related work

In the above systems the image was either attached with a string to it or just compared with their color histogram. When a string is attached with the image, there would not be any content checking of that retrieved image, so the accuracy would be very low and most of the times the generated output would not be the expected output. And when the image is compared using only the color histogram, it would drop its accuracy textures can be different.

System Introduction

Only simple features of image information cannot give the exact description of an image. We are using the three retrieval features color, texture and histograms to get images that are similar to query images. The working is based for dominant color, texture and a histogram features.

Color Histogram:

In imaging, a color histogram is a representation of the distribution of colors like RGB of that image. For digital images, a color histogram is nothing but the representation of the number of pixels that have colors in each of a fixed list of color ranges of the color space. The color histogram can be built for any kind of color space, but the term used for three-dimensional spaces is like HSV or RGB. In color histogram x-axis is RGB and y-axis is the frequency.

Color feature extraction:

For color feature extraction is obtained by color moments. Color moments are mainly used for color indexing purposes and extracting features in image for retrieval applications so we can compare how similar two images are based on color. Usually one uploaded image is compared to a database of images with pre-computed features in order to find and retrieve an output image which is most accurate. Each comparison between the two images results in their similarity score, and the lower this score is the more identical the two images are supposed to be. Since color moment encodes both color and shape information of the image, so they are a better features to use under different lighting conditions. There are three color moments are computed per channel. There are 9 moments for the color model is RGB.

Following are the color moments:

Color Expectancy

It is also called as mean. It can be interpreted as the average color in the image and can be calculated by using the formula given below

$$E_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij}$$

Where N is the number of pixels in the image and P_{ij} the value of the j^{th} pixel of the image at the i^{th} color channel.

Color Variance

It is also called as Standard Deviation. It is nothing but square root of the variance of color distribution in the image.

$$\sigma_i = \sqrt{(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^2)}$$

Skewness

It measures how asymmetric the color distribution is and it gives information about the shape of the color distribution. The formula is given below

$$s_i = \sqrt[3]{(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^3)}$$

Texture Extraction:

The texture extraction is usually by GLCM/CCM (gray-level co-occurrence matrix/color co-occurrence). The image is first is converted into the normalized GLCM matrix. Using this normalized matrix following features are extracted

Energy

Energy is a texture measure of image which represents repitition occurrence of similar changes in images. Gray-scale uniformity of weight and texture.

$$\text{Energy } E = \sum_x \sum_y P(x, y)^2$$

Contrast

Contrast this feature shows the clarity of image and how the values of matrix are dispersed in the given images. As the value of contrast goes increased the clearer image looks.

$$\text{Contrast } I = \sum_x \sum_y (x-y)^2 P(x,y)$$

Entropy

Entropy measures unpredictability in the image texture. If the value of the GCM is equal then the entropy is minimum and if value GCM is not equal then its value is greater. Therefore, entropy is always maximum given by GCM.

$$\text{Entropy } S = \sum_x \sum_y P(x, y) \log P(x, y)$$

ARCHITECTURE

Figure:

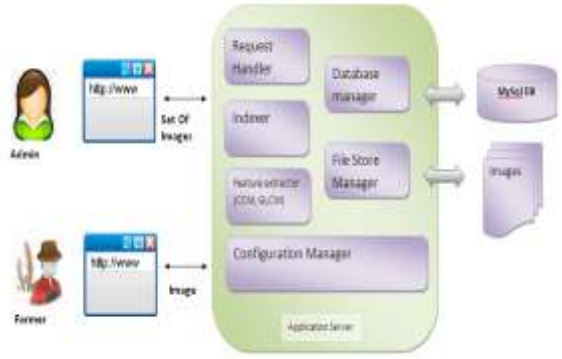

Fig1.Architecture of the proposed system

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

ALGORITHM

The algorithm used here is of GLCM
Input:Image File
Output:Energy ,Contrast, Entropy
Steps

1. Resize the image as per database images size.
2. Converts RGB values to grayscale values by forming a weighted sum of the R, G, and B components:
   0.2989 * R + 0.5870 * G + 0.1140 * B
3. Create an empty GLCM matrix.
4. For every pixel:
   4.1 Calculate how often a pixel with the intensity (gray-level) value i occurs in a specific spatial relationship to a pixel with the value j.
   4.2. Sum the number of times that the pixel with value i occurred in the specified spatial relationship to a pixel with value j in the input image.
   4.3. Update GLCM matrix [i,j] with above sum value
   4.4. Any more pixels, yes continue, no break.
5. After GLCM created, calculate:
   5.1. Contrast - Measures the local variations in the gray-level co-occurrence matrix.
   5.2. Entropy - Measures the joint probability occurrence of the specified pixel pairs.
   5.3. Energy - Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.

MATHEMATICAL MODEL

Figure:

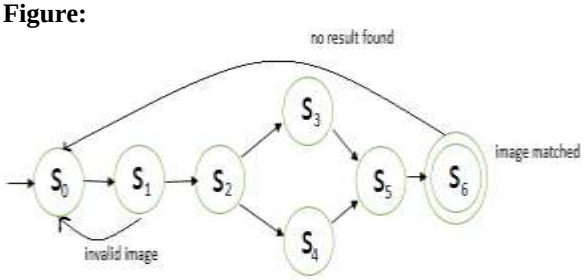

Fig2.Mathematical Model

S_0: Upload the image.
S_1: Image preprocessing (checks whether the image is in valid format) else return S_0.
S_2: Read the Image.
S_3: Color feature extraction and go to S_5.
S_4: Texture feature extraction and go to S_5.
S_5: Compare the image.
S_6: If image found generate the details else go to S_0.

ADVANTAGES

1. The main advantage of system that it provides the information of the disease by simply uploading the image.
2. The System helps the farmers to understand about the diseases about their cotton plant.
3. It also gives the precautions to take and avoid that kind of disease in the future.

LIMITATIONS

The System accepts any photo upload if it's valid format and process and finds no match.

CONCLUSION

The Project helps the farmers by giving the information about the diseases and its possible precautions to take by just clicking the image of their cotton leaves with most accuracy.

REFERENCES

[1] M.Rao, Dr.B.Rao, Dr.Govardhan, "Content based image retrieval using Dominant color and Texture features", International Journal of Computer science and information security, Vol.9 issue No: 2, February 2011.pp:41.
[2] X-Y Wang et al., "Effective image retrieval scheme using color, texture and histogram 0.features", Computer. Stand. Interfaces (2010), doi:10.1016/j.csi.2010.03.004.
[3] Chia-Hung Wei, Yue Li, Wing-Yin Chau, Chang-Tsun Li, "Trademark image retrieval using synthetic features for describing global histogram and interior structure", Pattern Recognition, 42 (3) (2009) 386–394.
[4] FAN-HUI KONG, "Image Retrieval using both color and texture features", proceedings of the 8th international conference on Machine learning and Cybernetics, Baoding, 12-15 July 2009.
[5] JI-QUAN MA, "Content-Based Image Retrieval with HSV Color Space and Texture Features", proceedings of the 2009 International Conference on Web Information Systems and Mining.
[6] Texture Feature Extraction of Video Frames Using GLCM http//www.ijettjournal.org/volume-4/issue-6/IJETT-V4I6P194.pdf
[7] Ashwini Vinayak Bhad, Komal Ramteke," Content Based Image Retrieval a Comparative Based Analysis for Feature Extraction Approach", 2015 International Conference on Advances in Computer Engineering and Applications (ICACEA)and IEEE 2015.

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

AUTHOR BIBLIOGRAPHY

	**Ms. Swarupa Kamble** is an Assitance Professor under Department Of Computer Engineering. She is having 5 years' experience in the field of teaching as well as research. Her Research interests include domains like Image Processing and Data Mining. Email : swar.kamble@gmail.com
	**Atul V. Kondekar** is a student pursuing his B.E Degree under Department Of Computer Engineering from University Of Pune. He is actively working on languages such as Java and Xml for making the scripts for the android application. He is also working in the development of frontend GUI of the application and web page. Email: atulkondekar@gmail.com
	**Mukul Wanajare** is a student pursuing his B.E Degree under Department Of Computer Engineering from University Of Pune. He is presently working on Mathematical Module of the application. Also, he is involved in the development of mathematical implementation of the sub algorithms involved in the System.
	**Swapnil Mane** is a student pursuing his B.E Degree under Department Of Computer Engineering from University Of Pune. He is currently working on the Mysql Database and CBIR library of the system. He is also involved in the cotton related diseases and it's solutions.
	**Vinay Kalambate** is a student pursuing his B.E Degree under Department Of Computer Engineering from University Of Pune. He is currenctly working on the Testing of the System.