

Detecting Healthiness of Leaves Using Texture Features



Srishti Shetty, Zulaikha Lateef, Sparsha Pole, Vidyadevi G. Biradar, S. Brunda and H. A. Sanjay

Abstract Agriculture is the dominant sector of our economy and contributes in various ways but the yield in the productivity leads to a significant reduction in the farmer's income. Monitoring crop health is important to increase the quality and quantity of the yield. But this requires manually monitoring the crops and also expertise in the field. Hence, automatic disease detection using image texture features is used for ease and to detect the disease at an early stage. The proposed methodology for the project is to design and implement the algorithm on two sets of databases: firstly, a locally generated leaf database which contains images of leaves and secondly, a standard database which is a common test database. The basic steps for crop disease detection include image acquisition, image preprocessing, image segmentation, feature extraction, and classification using image processing techniques. The acquired leaf images are preprocessed by removing undesired distortion and noise, and then, the processed image is further subjected to K-means-based segmentation. The segmented image is further analyzed using Haar wavelet transform and GLCM based on its texture by extracting feature vector. SVM is used for classification of image. Thus, the presence of diseases in leaf is identified along with all the features values of the leaf. It also calculates the accuracy rate of the prediction made by the system.

Keywords Image processing · Gray-level co-occurrence matrix · Support vector machine-nearest neighbor · Haar wavelet · K-means

S. Shetty · Z. Lateef · S. Pole · V. G. Biradar · S. Brunda (✉) · H. A. Sanjay
Department of ISE, NMIT, Bangalore 560064, India
e-mail: brunda.s@nmit.ac.in

© Springer Nature Singapore Pte Ltd. 2019
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Advances in Intelligent Systems and Computing 906,
https://doi.org/10.1007/978-981-13-6001-5_40

483

1 Introduction

The main aim of a development project is to change the current situation into a better one over time. Though industry has been playing an important role in Indian economic scenario, the contribution of agriculture still cannot be denied [1]. Agriculture is one of the biggest shares of economy in India, in terms of generating employment as well as provision of food for the ever-increasing population. The recognition and classification of crop diseases are of major economical and technical importance in agricultural industry. Maize popularly known as “corn” is one of the most versatile and major cash crops having wider adaptability under varied climatic conditions. Maize is one of the major cereal crops and is the third most contributing major crop in India after rice and wheat. Estimated losses due to major diseases of maize in India are about 13% of which fungal diseases cause major yield losses [2]. As per survey, bacterial blight, anthracnose, leaf spot, and rust are major commonly occurring diseases affecting maize crops in northern parts of Karnataka. They affect the photosynthesis with adverse reduction in the yield to an extent of 2891%. These diseases affect the quality as well as quantity of the agricultural products and lead economic losses. The leaves are the first one to be affected in case of any disease, generally before or after blooming which effects the growth plants. Our main objective is to concentrate on maize leaf disease detection based on the texture of the leaf.

2 Related Work

See Table 1.

3 Methodology

Image processing methods involve versatility, repeatability, and the preservation of original data precision. The various image processing techniques are:

1. Image acquisition
2. Image preprocessing
3. Image segmentation
4. Feature extraction
5. Image classification

(i) Image Acquisition

The images of the maize leaves are captured using a digital camera in JPEG format. The size of each image is 1500×1500 . The samples are collected from a field located in the University of Agricultural Sciences, Dharwad. Three

Table 1 Literature review

Author name	Preprocessing	Feature extraction	Result	Name of the data set used
Dheeb Al Bashish, Malik Braik, Sulieman Bani-Ahmad [10], 2010		CCM SDGM Neural networks	93%	Early scorch Cottony mold, ashen mold, late scorch, tiny whiteness
Sanjay B. Dhaygude, Nitin P. Kumbhar [8], 2013		CCM	Satisfied	Self
Kiran R. Gavhale, Ujwala Gawande, Kamal O. Hajari [3], 2014	DCT L^*a^*b YCbCr CES	GLCM	96%— SVMRBF 95%— SVMPOLY	Created 300
Aakanksha Rastogi, Ritika Arora, Shanu Sharma [4], 2015			Satisfied	Hydrangea leaf Maple leaf
Sachin D. Khirade, A. B. Patil [9], 2015	Histogram equalization	CCM	Satisfied	Wheat
Shivaputra S. Panchal, Rutuja Sonar [6], 2016	Image enhancement	GLCM	Satisfied	Pomegranate leaf, image database consortium
Anand R., Veni S., Aravinth J. [5], 2016	Histogram equalization, L^*a^*b	CCM SDGM	Satisfied	
R. Meena Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari, T. Kaviya [2], 2017	L^*a^*b	GLCM	0.9–1.0	Self-created database 60 citrus leaves using camera
Amruta Ambatkar, Ashwini Bhandekar, Avanti Tawale, Chetna Vairagade, Ketaki Kotamkar [7], 2017	Histogram equalization	CCM GLCM SDGM	Satisfied	Self
Ahmad Nor Ikhwan Masazhar, Mahanijah Md Kamal [1], 2017	L^*a^*b	GLCM	Chimaera 97% Anthracnose 95%	Self

hundred images of each category: healthy, northern blight, southern blight, and rust, are collected by keeping constant distance of 1 foot. One hundred and fifty images of each category are used for training, and remaining 150 images are used for testing. Thus, we have 600 images for training and 600 images for testing.

(ii) **Image preprocessing**

1. The images are of varying contrast. Preprocessing is necessary in order to correct nonuniform illumination. First, RGB image is converted into L^*a^*b [1–3] color space and the luminance component (L) is extracted. Then, contrast limited adaptive histogram equalization (CLAHE) is applied to L component and concatenated with ‘ a ’ and ‘ b ’ components, and finally, the image is converted from $L^*a^*b^*$ color space to RGB space (Fig. 1).

(iii) **Image segmentation**

Image segmentation is the way toward partitioning a propelled image into various bits (sets of pixels, generally called super-pixels). The goal of division is to contemplate the picture in detail and also change the features of an image into something that is more vital and less complex to analyze. Image segmentation is consistently used to discover objects and edges (lines, curves, etc.) in images. More definitively, image segmentation is the path toward consigning a label to every pixel in an image to such a degree, to the point that pixels with a comparative feature share certain characteristics. The yield of this division is a course of action of portions that cover the entire image. Each pixel in a fragmented zone is similar with respect to some basic element, for instance, color, texture, intensity, and contrast. Neighboring locales are diverse relating to similar characteristics.

(iv) **Feature Extraction using Haar wavelet and GLCM**

Texture analysis plays an important role in image analysis, such as disease detection, medical imaging, machine vision, and content indexing of image databases. Texture analysis can be done by wavelet transform. The Haar transform has been used as a necessary tool in the wavelet transform for feature extraction. The method is based on the application of Haar wavelet on RGB image of the preprocessed image to obtain horizontal, vertical, and diagonal coefficients. Then, gray-level co-occurrence matrix [1, 3, 6, 7] (GLCM) is constructed into two directions (0° and 90°) for each of the coefficients. GLCM matrix produces four statistical features like contrast, correlation, energy, and homogeneity.

(v) **Classification using SVM classifiers**

The concept of support vector machine (SVM) was presented by Vapnik and collaborators. It gets predominance as it offers alluring highlights and competent equipment to handle the issue of order. The SVM depends on measurable learning hypothesis. SVM’s better speculation execution depends on the standard of structural risk minimization (SRM). The idea of SRM is to expand the edge of class partition. The SVM was characterized for two-class issue and it searched for ideal hyperplane, which amplified the separation, the edge, between the closest cases of the two classes, named SVM. At exhibit, SVM is

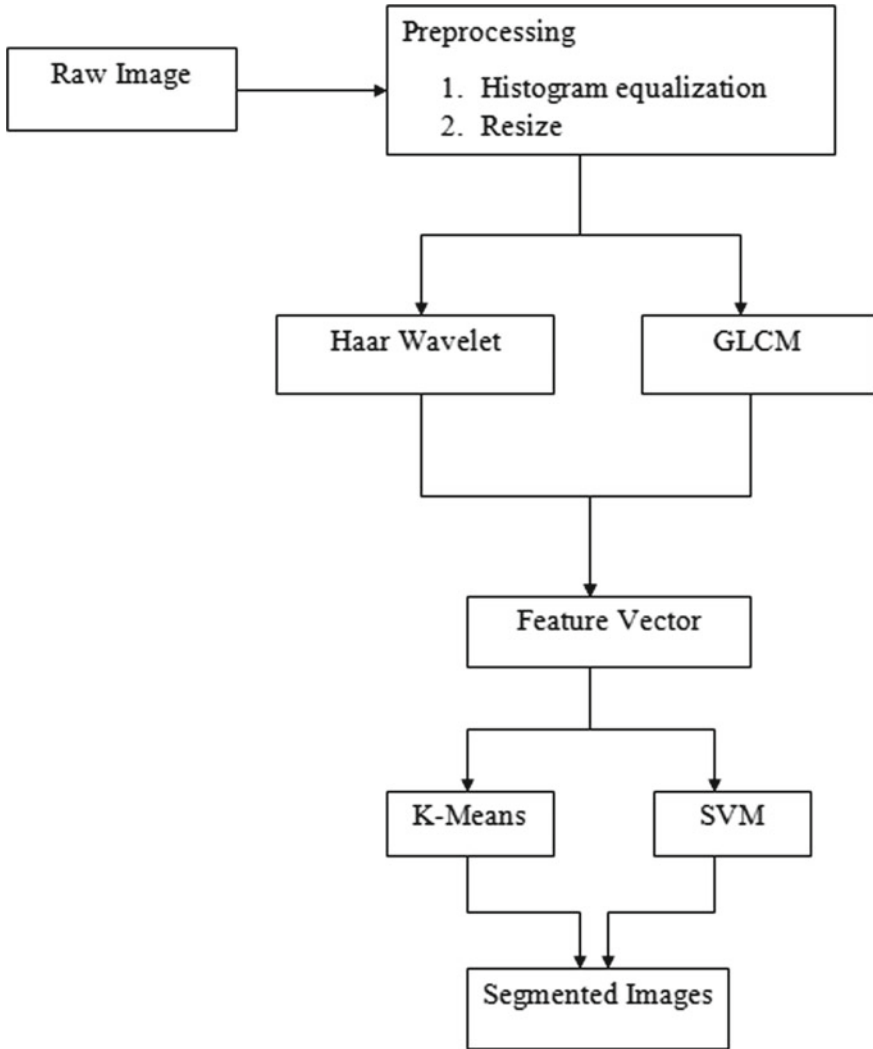


Fig. 1 Steps in leaf segmentation

mainstream grouping instrument utilized for design acknowledgment and other characterization purposes. The standard SVM classifier takes the arrangement of information and predicts to characterize them in one of the main two unmistakable classes. For multiclass order issue, we decay multiclass issue into various paired class issues, and we plan reasonable joined numerous parallel SVM classifiers.

4 Data set

For the purpose of this project, focus is laid mainly on maize leaves. The samples are gathered from a field situated in University of Agricultural Sciences, Dharwad. A total of 1200 deals were gathered with 600 images utilized for preparing and 600 for testing. In any case, non-dominant part of this maize is exhausted clearly by individuals. A bit of the maize production is used for corn ethanol, animal support, and other maize substances, for instance, corn starch and corn syrup. Some of the normal ailments found in maize leaves are northern blight, southern blight, rust, anthracnose, etc. These sicknesses are essentially found in Southern Indian districts.

5 Software Requirements

- The system will identify whether the leaf image is healthy or diseased.
- The system will detect different types of diseases that commonly occur in maize.
- The system is able to detect the percentage of the disease-affected area.
- The system is able to suggest the right treatment plan for the identified disease in maize leaf.

6 Nonfunctional Requirements

1. **Performance:** Fast and accurate detection of disease. (3–5 s).
2. **Usability:** The system will be useful for farmers and agricultural analysts.
3. **Reliability:** The system will be able to detect the three commonly occurring diseases in maize leaves and suggest the remedy.
 - a. **SVM:** 93–95% reliable with training data set of 400 images.
 - b. **KNN:** 86–90% reliable with training data set of 400 images.
4. **Availability:** The proposed system is readily available on any platform and provides on-demand service.
5. **Simplicity:** The proposed system is user friendly.

7 Algorithms Used

1. Histogram equalization

Histogram shows the frequency of pixel intensity values. It is used for analysis of image and thresholding. Histogram equalization is used for contrast adjustment.

2. K-Means Clustering

The K-means clustering is utilized for division and characterization of leaves in view of an arrangement of features, into K number of groups. The classification is done by minimizing clusters.

The algorithm for K-means clustering:

- Choose the center of K cluster, randomly or based on some logic.
- Relocate each pixel in the leaf image to the cluster that minimizes the distance between the pixels and the cluster center.
- Recompute the cluster centers by averaging all of the pixels in the cluster.
- Repeat steps 2 and 3 until effective clustering is achieved.

3. Haar Wavelet

The Haar wavelet is used for texture analysis. In this method, the preprocessed image is used to obtain the horizontal, vertical, and diagonal coefficients. The obtained coefficients are used for texture analysis.

4. Gray-Level Co-occurrence Matrix (GLCM)

It is a statistical method of examining texture that considers the spatial relationship of pixels. GLCM is used to calculate how often a pixel with gray-level value ‘i’ occurs either horizontally, vertically, or diagonally to an adjacent pixel with gray-level value ‘j’.

Statistics derived from the GLCM include contrast, correlation, energy, homogeneity, etc., of the image.

5. Support Vector Machine (SVM)

Linear support vector machine is used for classification of leaf diseases. SVM is a binary classifier which uses a hyperplane called the decision boundary between two classes. It maximizes the margin around the separating hyperplane.

1. Mean: It measures the mean value of all pixels in the relationships that contributed to the GLCM. It can be measured using the following formula:
where

P_{ij} Element i, j of the image.

N Number of gray levels

$$\mu = \sum_{i,j=0}^{N-1} i(P_{ij}) \tag{1}$$

2. Variance: The variance of the intensities of all the reference pixels is calculated using the following formula:

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{ij}(i - \mu)^2 \quad (2)$$

3. Contrast: Contrast measures the quantity of local change in an image. It reflects the sensitivity of the textures in relation to the changes in the intensity. It returns the measure of intensity contrast between a pixel and its neighborhood. Contrast is 0 for a constant image. It can be measured using the following formula:

$$\sum_{i,j=0}^{N-1} P_{ij}(i - j)^2 \quad (3)$$

4. Energy: Energy also means uniformity. It returns the sum of squared elements in the GLCM. The more homogenous the image is, the larger the value. When energy equals to 1, the image is believed to be a constant image. It can be measured using the following formula:

$$\sum_{i,j=0}^{N-1} (P_{ij})^2 \quad (4)$$

5. Correlation: This feature is used to measure how correlated a pixel is to its neighborhood. Correlation is 1 or -1 for a perfectly positive or negative correlated image. It can be measured using the following formula:

$$\sum_{i,j=0}^{N-1} P_{ij} \frac{(i - \mu)(j - \mu)}{\sigma^2} \quad (5)$$

6. Homogeneity: Homogeneity measures the similarity of pixels. A diagonal gray-level co-occurrence matrix gives homogeneity of 1. It becomes large if local textures only have minimum changes. It can be measured using the following formula:

$$\sum_{i,j=0}^{N-1} \frac{P_{ij}}{1 + (i - j)^2} \quad (6)$$

8 Result

Sample 1 (Fig. 2).

- Sample 2** (Fig. 3).
- Sample 3** (Fig. 4).
- Sample 4** (Fig. 5).
- Sample 5** (Fig. 6).
- Sample 6** (Fig. 7).
- Sample 7** (Fig. 8).
- Sample 8** (Fig. 9 and Table 2).

Fig. 2 **a** Input leaf and **b** Segmented leaf

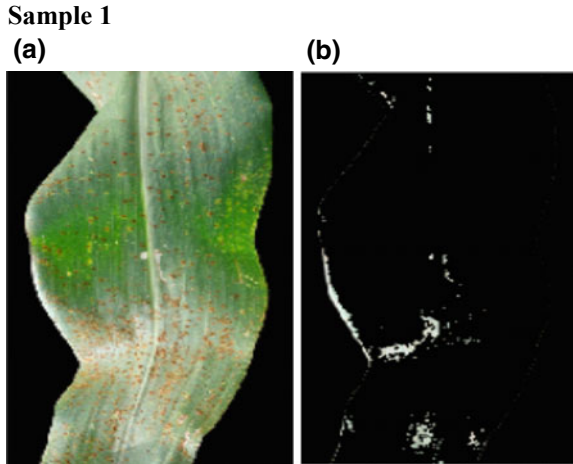


Fig. 3 **a** Input leaf and **b** Segmented leaf

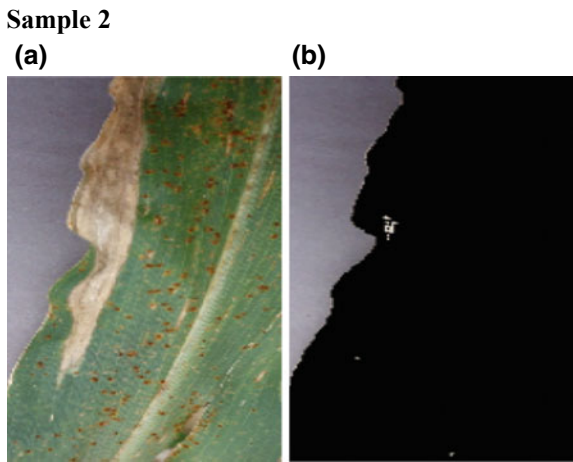


Fig. 4 a Input leaf and
b Segmented leaf

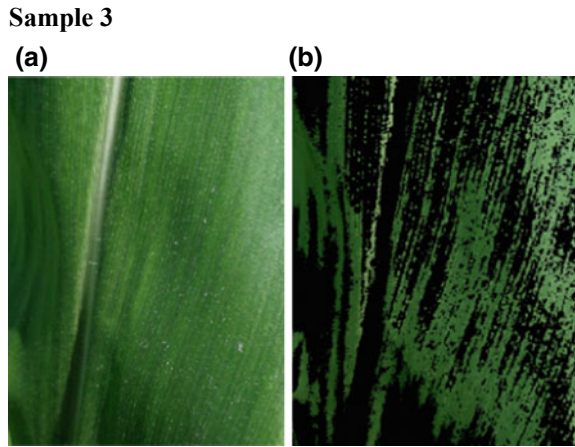


Fig. 5 a Input leaf and
b Segmented leaf

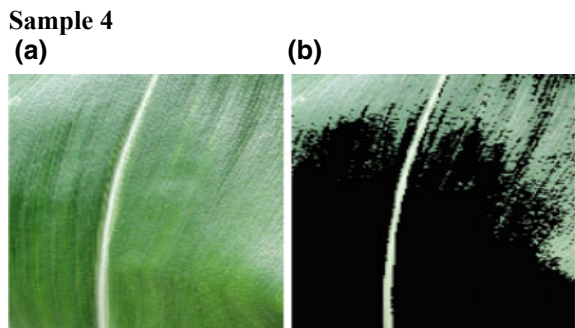


Fig. 6 a Input leaf and
b Segmented leaf

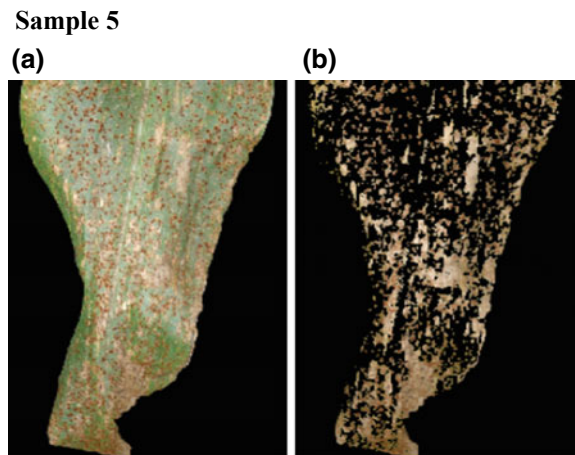


Fig. 7 a Input leaf and
b Segmented leaf

Sample 6
(a)



(b)

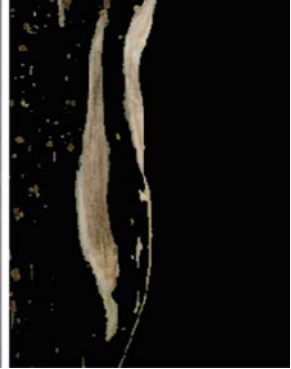


Fig. 8 a Input leaf and
b Segmented leaf

Sample 7
(a)



(b)



Fig. 9 a Input leaf and
b Segmented leaf

Sample 8
(a)



(b)



Table 2 GLCM values of different samples

	Mean	S.D	Entropy	Variance	Contrast	Correlation	Energy	Homogeneity
Sample 1	5.3614	26.3431	1.313	674.395	0.371148	0.953555	0.153258	0.867639
Sample 2	23.8789	49.098	1.89308	1136.33	0.256279	0.871701	0.206531	0.882174
Sample 3	31.4184	41.643	4.00298	1416.95	0.223564	0.884036	0.210835	0.898197
Sample 4	49.5676	74.519	2.97791	5088.52	0.402556	0.828097	0.128442	0.820587
Sample 5	29.887	57.3747	2.40529	2835.69	0.347677	0.960773	0.246748	0.879001
Sample 6	15.0888	41.6127	1.29395	1111.79	0.182057	0.908318	0.218444	0.914992
Sample 7	18.5059	48.7076	1.41064	1767.55	0.252881	0.825716	0.222579	0.886234
Sample 8	42.8826	79.4477	2.42895	5831.5	0.321507	0.80492	0.246217	0.858014

9 Conclusion

This work deals with crop yield detection based on the healthiness of leaves using image texture features. The texture features are Harlick features in addition to Haar wavelet features. The classification algorithm is a combinational one, SVM and K-means classification. The experiment is carried out on maize leaves database which is collected from Dharwad University. The results are found satisfactory. However, the accuracy may be improved by novel algorithms for texture features.

References

1. Masazhar, A. N. I., & Kamal, M. M. (2017). Digital image processing technique for palm oil leaf disease detection using multiclass SVM classifier. In *Proceedings of the 4th IEEE International Conference on Smart Instrumentation, Measurement and Applications*.
2. Meena Prakash, R., Saraswathy, G. P., Ramalakshmi, G., Mangaleswari, K. H., & Kaviya, T. (2017). Detection of leaf diseases and classification using digital image processing. In *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*.
3. Gavhale, K. R., Gawande, U., & Hajari, K. O. (2014). Unhealthy region of citrus leaf detection using image processing techniques. In *International Conference for Convergence of Technology*.
4. Rastogi, A., Arora, R., & Sharma, S. (2015). Leaf disease detection and grading using computer vision technology and fuzzy logic. In *2nd International Conference on Signal Processing and Integrated Networks*.
5. Anand, R., Veni, S., & Aravinth, J. (2016). An application of image processing techniques for detection of diseases on Brinjal leaves using K-means clustering method. In *Fifth International Conference on Recent Trends in Information Technology*.
6. Panchal, S. S., & Sonar, R. (2016). Pomegranate leaf disease detection using support vector machine. *International Journal of Engineering and Computer Science*, 5(6), 16815–16818.
7. Ambatkar, A., Bhandekar, A., Tawale, A., Vairagade, C., & Kotamkar, K. (2017). Leaf disease detection using image processing. In *International Conference on Recent Trends in Engineering Science and Technology (IJRITCC)* (Vol. 5, No. 1, pp. 333–336).
8. Dhaygude, S. B., & Kumbhar, N. P. (2013, January). Agriculture plant leaf disease detection using image processing. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(1), 599–602.

9. Khirade, S. D., & Patil, A. B. (2015). Plant disease detection using image processing. In *International Conference on Computing Communication Control and Automation* (pp. 768–771). New York: IEEE.
10. Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2010). A framework for detection and classification of plant leaf and stem diseases. In *International Conference on Signal and Image Processing* (pp. 113–118). New York: IEEE.