

Detection of Healthy and Defected Diseased Leaf of Rice Crop using K-Means Clustering Technique

Prabira Kumar Sethy
SUIIT, Sambalpur University
Jyoti Vihar, Burla
Odisha, India

Baishalee Negi
SUIIT, Sambalpur University
Jyoti Vihar, Burla
Odisha, India

Nilamani Bhoi
Veer Surendra Sai University
of Technology
Siddhi Vihar, Burla
Odisha, India

ABSTRACT

Rice covers about 69 percent of the cultivated area and is the major crop covering about 63 percent of total area under the food grains [1]. It is a staple food of almost entire population of Odisha; therefore state economy has directly affected the production of rice in the state [1]. The common disorder found in the rice crop and usually appears at the tillering and panicle initiation stage which shows on the leaves of rice crop [2]. The disorder is due to mineral deficiency and infection caused by the pest. The disorders are visualized by discoloration and dead spots on leaves. It may be beneficial for detecting defected diseased leaf by the symptoms that found on the surface of leaves [3]. In this paper, a novel approach to identify defected diseased leaf by using K-Means clustering or 3-Means clustering method is proposed. The experimental outcomes demonstrate that the proposed method is an impressive technique for the detection of defected part present on the leaves of the rice crop.

Keywords

Defected leaf disease detection; K-Means clustering; rice crop; defected segmentation.

1. INTRODUCTION

Diseased on rice crop implies significance reduction in both quality and quantity of rice products. Monitoring of healthy and diseased area on leaves of rice crop plays a vital role for successful cultivation. In early days the monitoring and analysis were done manually by the expert person in the field. This traditional process takes excess time and workload. In most of the cases, disease symptoms are seen on the leaves of the rice crop. Now a day the defected diseased portion of the leaf can be detected by the new emerging technique like Image Processing. Clustering is the method of grouping a given set of patterns into disjoint clusters [4]. This is done such that patterns belonging to two different clusters are different and patterns in the same cluster are same [4]. The K-Means method has been shown to be practical efficient to generate good clustering results for many real-time applications. However, K-Means technique requires time proportional to the product of a number of clusters per iteration and number of patterns. This is computationally expensive for large datasets. A novel algorithm for practical implementation of K-Means technique is proposed. The approach algorithm generates the same or approximate (due to the round-off errors) clustering outputs to the direct K-Means method. It has significantly better performance compared to direct K-Means algorithm in most cases [4]. K-Means is one of the simplest unsupervised learning techniques that deals with the notable clustering problem. In this paper, K-Means clustering or 3-means clustering segmentation method (assuming the value of K as 3) to detect the defected diseased

leaf is approached. The main objective of this paper is to find out the defected diseased portion on a leaf of the rice crop.

This paper represents an effective image segmentation technique i.e. K-Means clustering algorithm depend on color features from the input image. Defect segmentation is carried out in two phases. At early phase, the pixels are clustered based on their spatial features and color, where the clustering technique is very effective. Then in next phase, the clustered blocks are merged to a specific number of regions. Using these two following step, it is possible to increase the estimation capability, if only excluding feature extraction for every pixel present in the input image of leaves. Although the color is not generally used for K-Means based segmentation, because it produces a confined power for different regions of the image.

2. LITERATURE SURVEY

Many researchers worked on detection of defected diseased leaf using different concepts.

Jimita Baghel, Prashant Jain [3], developed Automatic Leaf Disease Detection by using K-Means Segmentation Method. The proposed system uses Euclidean distance and K-Means clustering method for segmentation of an image to segment the leaf area, disease area and background area of input image leaf in order to calculate the severity of disease in a leaf.

Prof. Dr. Sanjay B. Patil, Prof. T.M. Dudhane [5], they discussed K-Means clustering algorithm and CBIR based system for detection of leaf disease. By using the K-Means clustering the disease severity can be measured which requires computation of infected area of the leaf. This new image processing technique is used for the segmentation of a diseased portion of the healthy portion of a leaf.

Shiv Ram Dubey et al. [6], they represent that K-Means clustering algorithm is usually used to determine the natural patterns of pixels present in an input image. For the defect segmentation of infected fruit, image can be achieved by the help of K-Means clustering which contains three or four clusters for a better result. After the experimental observation, it found that the proposed method is able to accurately segment the defected area of fruits present in the image.

Smita Naikwadi, Niket Amoda [7], proposed that K-Means clustering method considers labeling the objects based on a group of features into K number of objects. In this paper, K-Means clustering technique is used for image segmentation for which it can determine the infected part present in an image accurately. Here in this paper also discussed features like color co-occurrence texture methodology, gray-level co-occurrence matrix etc. can be extracted by using K-Means clustering technique for the classification purpose.

Dheeb Al Bashish, Malik Braik, Sulieman Bani-Ahmad [8], approached that by using K-Means clustering methodology works efficiently for the detection of infected area present on a leaf. Also in this paper, the value of K is assigned as four. So, one or more clusters contain the disease area in case of when the leaf is infected by more than one disease. Here in this paper K-Means uses squared Euclidean distances. Also by using the back propagation neural network classification of diseases can be possible.

3. PROPOSED METHODOLOGY

In this section, the flow chart (Figure 1) represents basic steps are followed for plant disease detection.

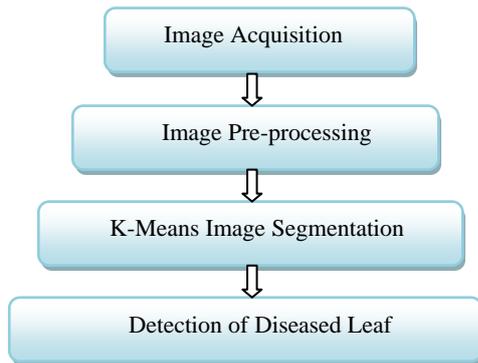


Figure 1: Flow chart of Proposed Methodology

According to the flow chart that mentions in figure 1, it is discussed below briefly.

3.1 Image Acquisition

Image Acquisition considers as the first step for the proposed methodology. First take the input leaf image of the rice crop has captured by the camera. The input image is in RGB (Red, Green and Blue) form. Then, the RGB leaf image is converted into suitable color space as per the requirements (Figure 3).

3.2 Pre-Processing of the Image

In this process, the query image converted to suitable color space i.e. L^*a^*b color space on which the algorithm can be worked. To extract the required information from the image more efficiently (Figure 4-a) by using image resizing and contrast enhancement image pre-processing techniques.

3.3 K-Means Based Image Segmentation

This step includes the segmentation of an image using K-Means algorithm. It is quite a helpful method for detection of the object which is based on a set of features into K number of classes. By minimizing the sum of the squares of the distance between the corresponding cluster and the object it can able to detect the interesting part of the input image. In K-Means clustering techniques, the clusters are determined by the groupings of pixels having the same value present in an image. Practically, the computational speed of this new image processing technique is very fast as well as gives more accurate output. The input dataset is partitioned into K number of clusters and each cluster is considered by a cluster center which is adaptive by nature. Initially considered values are known as seed-points and inputs are also known as data points. Estimation of the distances between the centers, inputs, and allocate inputs to the nearest center is only possible by using K-Means clustering technique [10].

Following steps represents K-Means clustering algorithm:

- In the first step, choose the value of K of all points as a partition centers randomly.
- Next, the distance between each data point on the set and those centers computed successfully, save that information for the further need.
- In final step calculations, assign each point to the nearest cluster centers, so that the minimum distance calculated for index each point. Then, add that point of the specific partition set.
- The distance between the cluster center and the pixels can be minimized by assigning each pixel in the image to the cluster as mention in the equation (Eq. 1).

$$c^{(i)} = \underset{j}{\operatorname{argmin}} \|x^{(i)} - \mu_j\| \quad (1)$$

- The cluster centers are computed by averaging all of the pixels present in the cluster according to the equation (Eq. 2). The convergence is attained by repeating the steps 3 and 4.

$$\mu_i = \frac{\lim_{i \rightarrow m} \sum_{j=1}^m 1\{c_{(i)} = j\} x^{(i)}}{\lim_{i \rightarrow m} \sum_{j=1}^m 1\{c_{(i)} = j\}} \quad (2)$$

3.4 Detection of Defected Diseased Leaf

After the successful implementation K-Means cluster based segmentation it is easily recognize the healthy and infected area present on the leaves of the rice crop.

4. DEFECTED AREA SEGMENTATION

K-Means algorithm is a quite useful method for image analysis purpose. By using this algorithm it is able to get the interest of the region. Aim of this work is to separate the object and background exactly by blurring the boundary of the input image.

Step1. First, read the input image of infected leaves.

Step2. Then, the input image should have to convert from RGB to L^*a^*b color space. L^*a^*b color space consists of two chromaticity layer in ‘*a’ and ‘*b’ channels and a luminosity layer in ‘L*’, channel. All of the color information is present in the ‘*a’ and ‘*b’ layers only, so by using L^*a^*b color space computational speed can increase. In MATLAB software, the RGB image can convert to L^*a^*b color space by using ‘makecform’ and ‘applycform’ command directly.

Step3. Colors are classified by using K-Means clustering in ‘*a*b’ space and simultaneously the difference between two colors can be evaluated with the help of Euclidean distance metric which is considered here.

Step4. Each pixel of the image will be labeled with its assigned cluster index (Figure 4-b & Figure 5-b).

Step5. The pixels present in the input image have been separated by color using pixel labels which will result in the different image based on the number of clusters.

The flowchart (Figure 2) shows in below is the framework for the defect segmentation of an infected leaf.

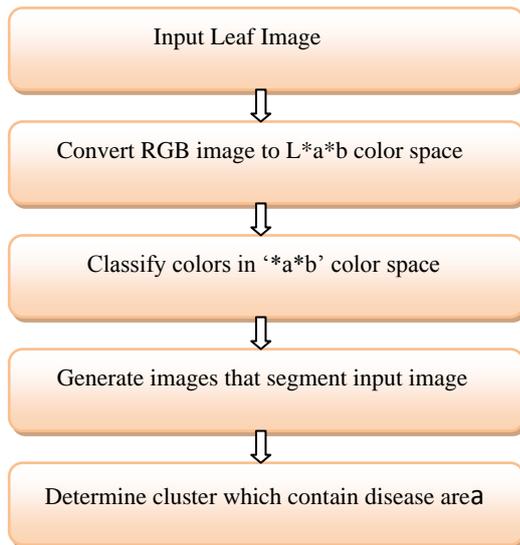


Figure 2: Flow chart of Defected Area Segmentation

The performance of the proposed algorithm can be evaluated by taking leaves of a rice crop as a case study fig. 3. The introduced method is evaluated on some diseased leaf which is infected by brown spots and leaf scald disease. The presence of a lot of variations in the data set makes it more realistic.

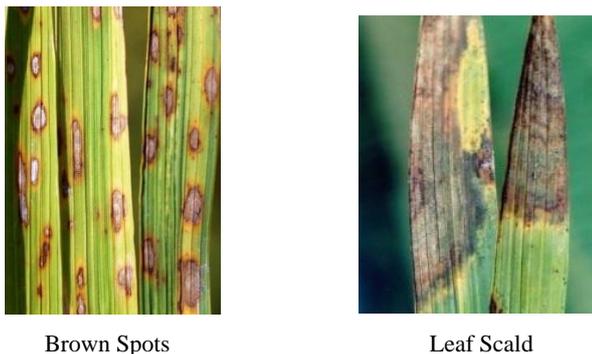


Figure 3: Input Images of Various Defected Diseased Leaf

Figure 4 and 5 shows the defect segmentation results of the leaf which is infected with brown spots, leaf scald diseases using K-Means clustering technique. The input image has been segmented into three clusters and it is very clear that the second cluster correctly segments the infected portion of the image and contains diseased affected area (Figure 4-d & Figure 5-d). From the above observations it is found that the 3-Means clustering technique gives good segmentation results. So in this experiment, the input image is partition into 3 clusters as per the requirements.

The figures 4 and 5 as mention in below shows the detection results of diseased leaf image which is infected by brown spots and leaf scald diseases simultaneously. If the cluster value is considered as 2, one cluster contains the leaf part while the second one contains the infected part and background portion also. When the number of the cluster is increase to 3, then one cluster contains the background area (Figure 4-c & Figure 5-c), the second one contains the infected leaf area (Figure 4-d & Figure 5-d) and the third one contains healthy leaf area (Figure 4-e & Figure 5-e). By using 3-Means clustering the outputs are visualized more accurately. If the number of the cluster is increase to 4 or 5, then it is not able to segment the entire defected portion into a single

cluster. But for some cases, 3 clusters are not sufficient to recognize the infected part because if the color of the infected part is having quite a similar color with input leaf image. On that case, it is bound to consider more number of the cluster for the better result.

Here in this paper, it is easily detect the healthy and infected area by using 3-Means clustering. So according to the study it is say that if defected area is smaller, more clusters will be required while if defected area is larger, the fewer cluster will be needed. It means a number of clusters required for the disease detection of the infected image are inversely proportional to the infected area.

The experimental results suggest that the introduced method for the cluster based segmentation in this paper is a robust technique because it can accurately segment the healthy part, infected part, and background also.

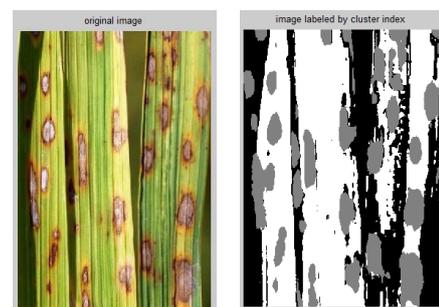


fig. (a)

fig. (b)



fig. (c)

fig. (d)

fig. (e)

Figure 4: 3-Means clustering for defected leaf that is infected with Brown spots disease (a) The infected leaf image, (b) The input image label with its cluster index, (c) 1st cluster, (d) 2nd cluster, (e) 3rd cluster.



fig. (a)

fig.(b)

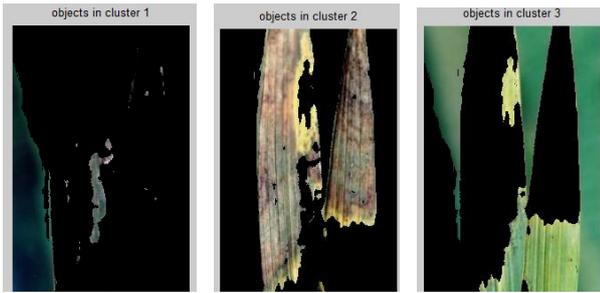


fig. (c)

fig. (d)

fig. (e)

Figure 5: 3-Means clustering for defected leaf that is infected with Leaf scald disease (a) The infected leaf image, (b) The input image label with its cluster index, (c) 1st cluster, (d) 2nd cluster, (e) 3rd cluster.

5. INFECTED AREA CALCULATION

After the successful execution of K-Means algorithm, the total leaf area (A_t) as well as diseased area (A_d) of the leaf is calculated. The percentage of infection (P) is calculated by the following equation (Eq. 3).

$$P = \frac{A_d}{A_t} * 100 \quad (3)$$

6. EXPERIMENTAL OBSERVATION

After the successful computation of the algorithm in MATLAB software the defected region is observed as (Table. 1);

Table 1. Observation Table

Name	1 st Cluster (Background Area)	2 nd Cluster (Infected Area)	3 rd Cluster (Healthy Area)
Brown Spots	22.0268%	15.430%	29.699%
Leaf Scald	28.3185%	15.013%	24.952%

For Figure 4-a, the input sample is infected by the brown spots disease. The infected area is covered by the second cluster (Figure 4-d), so according to the observation, the input leaf image contains 15.4302% of affected area.

For Figure 5-a, the input sample is infected by the leaf scald disease. The infected area is covered by the second cluster (Figure 5-d), so according to the observation, the input leaf image contains 15.013% of affected area.

7. CONCLUSION

A framework for the detection of defected diseased leaf using K-Means clustering based segmentation is proposed and evaluated in this paper. The proposed K-Means clustering technique segmentation for segmenting defect area with three

clusters and found the outputs accurately. Here the leaves of rice crop are considered as a case study and according to the experimental observation. Suggest that the proposed approach is able to detect the healthy leaf area and defected diseased area accurately.

8. FUTURE WORK

The future work includes that, if the leaf is infected by more than one disease, then by taking 4 or 5 Means clustering algorithm the diseases are detected accurately. Also, a researcher can focus on automatic determination of a number of clusters required to segment the defected area more accurately.

9. REFERENCES

- [1] S.R. Das, Department of Plant Breeding and Genetics, "Current status of the rice crop in Odisha", University of Odisha Agriculture and Technology, BBSR.
- [2] Karen Moldenhauer et al., "Development and Growth of rice crop", the University Of Arkansas Division Of Agriculture.
- [3] Prashant Jain, Jimita Baghel, "Leaf Disease Detection by using K-Means Based Segmentation". Int. Journal of Engineering Research and Application, ISSN: 2248-9622, Issue 3, (Part -5), Vol. 6, March 2016.
- [4] Khaled Alsabti, Sanjay Ranka, Vineet Singh, "An Effective K-Means Clustering Method", IITL, Hitachi America, Ltd.
- [5] Prof. T.M. Dudhane, Prof. Dr. Sanjay B. Patil, "Leaf disease detection by the help of K-Means Segmentation Method and CBIR system", IRJET, e-ISSN: 2395 -0056, p-ISSN: 23950072, Dec-2015.
- [6] Pushkar Dixit et al, "Diseased Fruit Area Detected by using K- Means Clustering Segmentation Method", ijimai.2013.229, Vol. 2, NO. 2.
- [7] Niket Amoda, Smita Naikwadi, "DETECTION OF PLANT DISEASES USING IMAGE PROCESSING TECHNIQUE", IJAIEM, Volume 2, Issue 11, ISSN 2319 – 4847.
- [8] Malik Braik, Dheeb Al Bashish, Sulieman Bani-Ahmad, "Leaf Disease Detection and Classification by using K-means algorithm and Neural-networks classification", Information Technology Journal, ISSN 1812-5638.
- [9] A. B. Patil, Sachin D. Khirade, "Detection of Plant Disease Using Advance Image Processing Technique", International Conference on Computing Communication Control and Automation, IEEE, 978-1-4799-6892-3/15, 2015.
- [10] Tapas Kanungo and David M. Mount, "Analysis andImplementation of An Efficient K-Means Clustering Method", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No.7.