

# Image Segmentation K-Means Clustering Algorithm for Fruit Disease Detection Image Processing

**P. Kanjana Devi**

Research Scholar

Dept of master of computer applications,  
N.G.M College, Pollachi, Tamil Nadu, India.

**Dr. Rathamani**

Assistant professor

Dept of master of computer applications,  
N.G.M College, Pollachi, Tamil Nadu, India.

rathababu@yahoo.com

**Abstract:** Fruit diseases are always considered as a remarkable issue in the cultivating business carried out across the globe. This arises the need for manual checking framework. In this way, agriculturists require the manual analysis of fruits. Nevertheless, the continually manual watching does not provide adequate results and they generally require a heading from an expert. The world economy is primarily depending on the agri-business as its development is diminishing when it has been appeared differently in relation to the expansion in intrigue and this ratio of intrigue versus creation is foreseen to be high in the upcoming years. Recently, clustering and fruit image segmentation algorithms are implemented for identifying the fruit diseases. To exhibit its importance, an algorithm plot is surveyed by utilizing various estimations. For instance, intensity ratio, specificity ratio, and probability ratio.

**Keywords-** Image processing, K-means clustering algorithm, intensity ratio, specificity ratio, probability ratio, fruit disease.

## I INTRODUCTION

The farming field is something past being taking care of source in the present world. Regardless, in view of climatic and various changes consistently, crop yields and farming yield have gotten slanted to certain critical issues which are a subject of certifiable concern [6]. India is the green land by producing 44.04 million tons of fruit and it is a second-greatest producer of fruits. India contributes 10% to the world's fruit creation. Indian ranchers build up a variety of fruits those are apple, banana, citrus, grape, mango, guava, papaya, and watermelon. For fruit disaster, it will consider water-soaked wounds superficially, it converts the low resulted images to a high resolute image. This was identified in the entire fruit image or partially in a part of an image. There's squeezing found a workable pace disaster at the primary stage. The fruit business contributes around 20% of the country's progression. The production of significant worth fruits has been reduced taking into account the disgraceful improvement of fruit, absence of maintenance, and manual examination. The amount and nature of the cultivating things are decreased by the disease of fruits [12].

The fundamental for fruit diseases are contaminations and minuscule creatures [4]. The diseases are also realized by awful common conditions. The assurance of fruit disease is noteworthy. The meeting charges of capable expert's square measure high and it's conjointly unachievable to provoke it on time in a remote territory. Hereafter, there's a need for customized fruit undesirable location framework inside the starting time of the disaster. Ranchers as a rule will watch visual side effects of distress on fruit.

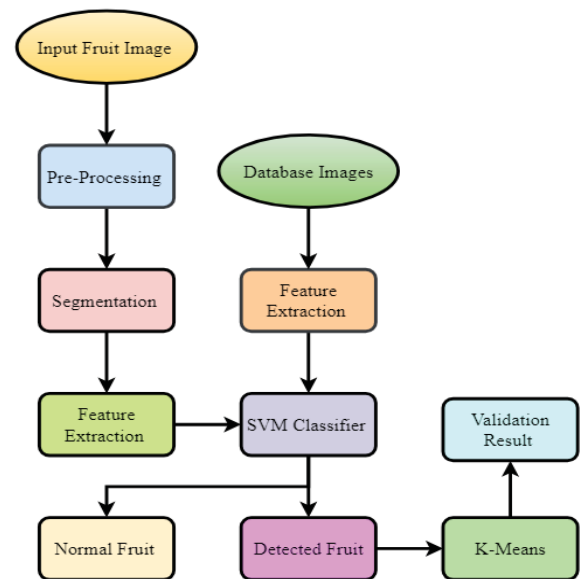


Fig 1: Block Diagram

Authorities could watch visual appearances of disaster on fruit. Specialists could simply break down the disaster or it could place trust in lab investigations. Most of the eventual practices for the fruit disaster location framework in the Republic of India square will measure eye recognition by space skilled. There are different characteristics and practices of such fruit diseases in which a significant parcel of them are less discernable. Agri-business contrasts from farming all

things considered in the size zone of improvement and the number of harvests to be created. The purpose of this research paper is to look out the diseases in the pomegranate fruit. This procedure acknowledges the contribution as an image of a fruit and decides it as either tainted or non-contaminated. The purpose chase a framework that makes the ranchers spot unfortunate by prescribing the significant pictures to address the image from data.

## II LITERATURE SURVEY

**Milos Ilic, Sinisa Ilic, Srdjan Jovic, Stefan Panic** (2018) Proposed the cherry fruit [1] pathogen sickness location subject to data mining desire. The utilization of various methods finished in MATLAB for conceivable fruit tribulation contamination conjecture. Close by the figure of potential diseases [13] MATLAB systems were utilized for data processing, portrayal, exceptional case location, and amendment. By anticipating the perfect open door for contamination occasions, the manufactured confirmation is ceaselessly profitable so ranchers could set aside cash and in actuality, side, which is surely progressively basic, it gives continuously productive sustenance taking into account the diminished number of substance medications. **Md. Rasel Howlader, Umme Habiba, Rahat Hossain Faisal and Md. Mostafijur Rahman** (2019) proposed the modified acknowledgment of fruit leaf diseases by using an enhanced convolution neural framework. The proposed model applies to arrange basic diseases of guava leaf [2], and all the other methodology of the D-CNN framework detached into a few basic strides for conspicuous verification of guava leaf disease identification. The distinctive confirmation accuracy demonstrates that the proposed D-CNN model is progressively convincing and gives a superior plan to view and control guava leaf diseases. To remember more classes for the dataset and will plan and test various sorts of leaf diseases [14] using this proposed model. Furthermore, it considers interfacing the model on bleeding edge cells to make a quick and competent judgment which can help the ranchers instantly identifying and turning away diseases. **Youssef Es-saady, Ismail El Massi, Mostafa El Yassa, Driss Mammass, Abdeslam Benazoun** (2016) Proposed to the customized acknowledgment of plant leaves diseases subject to a consecutive blend of two SVM classifiers [3]. The proposed structure depends upon the successive blend strategy of two SVM classifiers. The fundamental classifier uses the concealing to arrange the photos; it considers, at this stage, the diseases with a family member or closest concealing having a spot with the same class. By that point, the following classifier is used to disengage between the classes with relative concealing as per the shape and surface highlights. Subsequently, the estimation is attempted six classes of plant leaves diseases [7] [8] including three sorts of aggravation creepy crawlies harms and three kinds of pathogens signs (Early criticize, Late scurge, and Powdery shape), which is a general sense attack the vegetable yields. Consequently, it will provide a choice to the steady framework that draws in the ranchers in assessment to sift through. **Md. Rasel Mia, Sujit Roy,**

**Subrata Kumar Das, Md. Atikur Rahman** (2019) proposed the mango leaf disease [5] acknowledgment using the neural framework and bolster vector machine. To perceive the indications of plant diseases feasibly with AI than a manual observing framework. Here, Trained data is conveyed by the grouping strategy gathering pictures of leaves that were various sickness affected. The proposed framework could successfully recognize and bunch the assessed illness with a commonplace accuracy of 80%. To perceive four sorts of contamination and one ordinary among through and through five social occasions. Each includes four exceptional sorts of pictures.

## III PROPOSED METHOD

### E-K-Means clustering Algorithm

The Enhanced K-means algorithm endeavours to divide the dataset into K pre-characterized particular non-covering subgroups (bunches) where each datum point has a spot with only one social occasion. K-means is one of the least unpredictable solo learning algorithms that take care of the striking clustering issue. The system follows a clear and straightforward way to deal with a bunch of a given instructive assortment through a specific number of groups (expect k groups) fixed apriori. The principle thought is to characterize k centers, one for each group. It requests a group of data that contains the aggregate of its models which is closer to its center and not to the point of convergence of some other bunch. Along these lines, when tests having a place with a group are closer to the point of convergence of a substitute bunch, the k-means algorithm moves the comparing data tests from their unique group to the new group. These centers should be set in a smart way by taking different territory into account with assorted result. K-Means clustering algorithm is an unaided algorithm and it is used to an area from the establishment. Bunches or fragments of the given data are converted into K-groups or parts reliant on the K-centroids. The point of convergence of the group can be considered as the operator of the bunch, because the center is close to all models in the bunch.

### Steps in K-Means clustering algorithm:

- Step 1:** Initially select and choose the quantity of clusters K.
- Step 2:** Also select the absolute arbitrary K points, the centroids (not essentially from your dataset).
- Step 3:** Assign each and every datum point to the nearest centroid → that structures K clusters.
- Step 4:** Finally, compute and place the new centroid of each cluster.
- Step 5:** Reassign every datum point to the new nearest centroid. On the off chance that any reassignment. Occurred, go to step 4, in any case, the model is prepared.

Various means required during the time spent identification and characterization of diseases of fruit. At first, the establishment of the info image is ousted and the contaminated part is gotten by applying an image division system to the information image. In the accompanying stage, the extraction of highlights is done utilizing highlight extraction methodologies. A while later, utilizing arrangement procedure diseases are requested.

**Image Acquisition**

The development of re-establishing an image from some beginning is called Image obtaining. The procurement of an image is reliably the fundamental condition for the work methodology course of action of processing of an image. The image acquired is all-around ordinary and is the effect of any equipment which was figured out how to make it.

**Image Segmentation**

Before the component is removed from an image, the disease of the fruit must be partitioned. The division of an image is the way toward confining the modernized image into its constituent parts, which changes the depiction of the image and bringing about extricating the noteworthy part and which is less complex to analyze. When the fascinating part with regards to an application is disengaged the division should terminate. In the proposed framework, the clarification behind the division is to discover the districts in the image, which are diseased areas. Different methodologies are used for image division. Fluffy c-means clustering procedure and K-means clustering method are used in the proposed framework to perform division. K-means clustering methodology is a system for group assessment that relies upon apportioning and focuses on discernments into k absolutely irrelevant bunches and every observation will have a spot with one of the bunches with the nearest mean.

**Feature Extraction**

Highlight extraction is one of the critical walks in fruit disease grouping. GLCM framework is utilized for Feature extraction. The dark level co-event cross-section highlights the extraction strategy, depicts the outside of an image by figuring progressive repetition of sets of the pixel with explicit qualities, and in a predefined spatial relationship happen in an image, achieves making a GLCM, and later isolating credible measures from the structure.

**Classification**

Extricated highlights from GLCM are empowered into the classifier. KNN classifier is used. In KNN depiction object is portrayed by a greater bit of its neighbours, and the thing is given out to the class closest among its k closest neighbours.

IV EXPERIMENT RESULTS

Table 1: Comparison table of Intensity Ratio

Support Vector Machine (SVM) algorithm	Existing-K-Means Neighbour algorithm	S Clustering algorithm
50	60	78
55.4	65	85
59.6	70	92
60.9	75	95
70.3	81	99

The examination table of the intensity ratio of support vector machine (SVM) algorithm, existing K-Means neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. While contrasting the intensity ratio of the support vector machine algorithm, existing K-Means neighbour algorithm, E-K-Means clustering algorithm the E-K-Means Clustering algorithm is superior to the next two algorithms. The support vector machine (SVM) algorithm esteem begins from 50 to 70.3, the existing- K-Means neighbour algorithm esteems begin from 60 to 81 and E-K-Means clustering algorithm esteems begin from 78 to 99. Each time the E-K-Means clustering algorithm gives incredible outcomes.

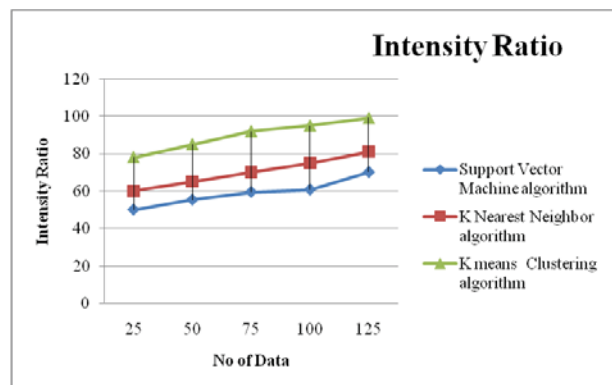


Fig 2: Comparison Chart of Intensity Ratio

Fig 2 shows the correlation chart of intensity ratio of support vector machine algorithm, existing-K-Means neighbour algorithm, and E-K-Means clustering algorithm with the various qualities. The number of information in the x pivot and intensity ratio in the y hub is measured. The E-K-Means clustering algorithm is superior to the next two algorithms. The support vector machine algorithm esteem begins from 50 to 70.3, the existing-K-means neighbour algorithm esteems begin from 60 to 81 and the E-K-Means clustering algorithm esteems begin from 78 to 99. Each time

the E-K-Means Clustering algorithm gives incredible outcomes.

Table 2: Comparison table of Specificity Ratio

**Specificity Ratio**

Support Vector Machine (SVM) algorithm	Existing-K-Means Neighbour algorithm	E-K-Means Clustering algorithm
30.6	45.6	60
36.7	50	65
39.2	54.8	70
45.6	60.9	75
49.5	68.3	80

Table 2 shows the specificity ratio of the Support Vector Machine (SVM) algorithm, Existing-K-Means Neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. While comparing the different algorithms like the intensity ratio of the support vector machine algorithm, the existing K-Means neighbour algorithm, and the E-K-Means clustering algorithm, the E-K-Means clustering algorithm is superior to the next two algorithms. The Support Vector Machine algorithm esteem begins from 30.6 to 49.5, the Existing-K-Means neighbour algorithm esteems begin from 45.6 to 68.3 and the E-K-Means Clustering algorithm esteems begin from 60 to 80. Each time the E-K-Means Clustering algorithm gives incredible outcomes.

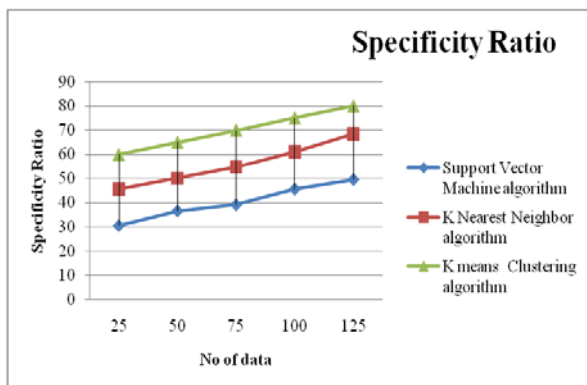


Fig 3: Comparison Chart of Specificity Ratio

The correlation chart of the specificity ratio of the Support Vector Machine (SVM) algorithm, Existing-K-Means Neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. The number of information in the x hub and the specificity ratio in the y hub are measured. The E-K-Means Clustering algorithm is superior to the next two algorithms. The Support Vector

Machine (SVM) algorithm esteem begins from 30.6 to 49.5, the Existing-K-Means Neighbour algorithm esteems begin from 45.6 to 68.3 and the E-K-Means Clustering algorithm esteems begin from 60 to 80. Each time the E-K-Means Clustering algorithm gives incredible outcomes.

Table 3: Comparison table of Probability Ratio

**Probability Ratio**

Support Vector Machine (SVM) algorithm	Existing-K-Means Neighbour algorithm	E-K-Means Clustering algorithm
40.6	49.6	55
47.3	56.2	62
55.5	62.1	69
59.3	68.3	77
66.9	77.5	86

The correlation table of the probability ratio of the Support Vector Machine algorithm, Existing-K-Means Neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. While gazing at the intensity ratio of the Support Vector Machine algorithm, Existing-K-Means Neighbour algorithm, E-K-Means Clustering algorithm the E-K-Means Clustering algorithm is superior to the next two algorithms. The Support Vector Machine algorithm esteem begins from 40.6 to 66.9, the Existing-K-Means Neighbour algorithm esteems begin from 49.6 to 77.5 and the E-K-Means Clustering algorithm esteems begin from 55 to 86. Each time the E-K-Means Clustering algorithm gives extraordinary outcomes.

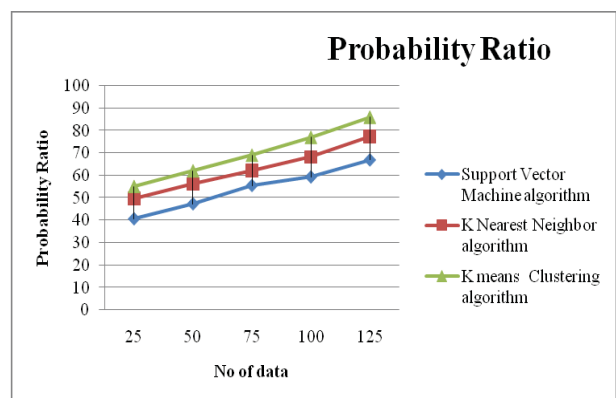


Fig 4: Comparison Chart of Probability Ratio

The correlation chart of probability ratio of Support Vector Machine (SVM) algorithm, existing-K-Means neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. The number of information in the x hub and the probability ratio in the y hub are measured. The E-K-Means Clustering algorithm is superior to the next two

algorithms. The Support Vector Machine (SVM) algorithm esteem begins from 40.6 to 66.9, the Existing-K-Means Neighbour algorithm esteems begin from 49.6 to 77.5 and E-K-Means Clustering algorithm esteems begin from 55 to 86. Each time the E-K-Means Clustering algorithm gives incredible outcomes.

## V CONCLUSION

The correlation Chart of Probability Ratio of Support Vector Machine (SVM) algorithm, Existing-K-Means neighbour algorithm, E-K-Means Clustering algorithm shows the various qualities. No of information in x hub and the Probability Ratio in y hub. The E-K-Means Clustering algorithm is superior to the next two algorithms. The Support Vector Machine algorithm esteem begins from 40.6 to 66.9, the Existing-K-Means neighbour algorithm esteem begins from 49.6 to 77.5 and E-K-Means clustering algorithm esteem begins from 55 to 86. Each time the E-K-Means clustering algorithm gives the incredible outcomes.

## REFERENCES

- [1] Milos Ilic, Sinisa Ilic, Srdjan Jovic, Stefan Panic, "Early cherry fruit pathogen disease detection based on data mining prediction", *Computers and Electronics in Agriculture* © 2018 Elsevier.
- [2] Md. Rasel Howlader, Umme Habiba, Rahat Hossain Faisal and Md. Mostafijur Rahman, "Automatic Recognition of Guava Leaf Diseases using Deep Convolution Neural Network", 2019 *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 7-9 February, 2019
- [3] Youssef Es-saady, Ismail El Massi, Mostafa El Yassa, Driss Mammass, Abdeslam Benazoun, "Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers", *2nd International Conference on Electrical and Information Technologies ICEIT'2016*
- [4] Anne-Katrin Mahlein & Erich-Christian Oerke & Ulrike Steiner & Heinz-Wilhelm Dehn, "Recent advances in sensing plant diseases for precision crop protection", *Eur J Plant Pathol*, vol. 133, pp. 197-209, 2012.
- [5] Dr. N. Sasirekha, N. Swetha, "An Identification of Variety of Leaf Diseases Using Various Data Mining Techniques", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 10, October 2015
- [6] Samiksha Bhor, Shubha Kotian, Aishwarya Shetty, Prashant Sawant, "Developing an Agricultural Portal for Crop Disease Prediction", *IJRST-International Journal for Innovative Research in Science & Technology*, Vol 3, Issue 11, April 2017, ISSN (online): 2349-6010
- [7] Dr. Kamaljit Kaur, Manpreet Kaur, "Prediction of Plant Disease from Weather Forecasting using Data Mining", *International Journal on Future Revolution in Computer Science & Communication Engineering* April 2018
- [8] Saradhambal, Dhivya, Latha, R. Rajesh, "PLANT DISEASE DETECTION AND ITS SOLUTION USING IMAGE CLASSIFICATION", *International Journal of Pure and Applied Mathematics* Volume 119, No. 14, pp. 879-884, 2018.
- [9] A. Haldorai and A. Ramu, "Security and channel noise management in cognitive radio networks," *Computers & Electrical Engineering*, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784
- [10] R. Arulmurugan and H. Anandakumar, "Early Detection of Lung Cancer Using Wavelet Feature Descriptor and Feed Forward Back Propagation Neural Networks Classifier," *Lecture Notes in Computational Vision and Biomechanics*, pp. 103-110, 2018. doi:10.1007/978-3-319-71767-8\_9
- [11] Haldorai, A. Ramu, and S. Murugan, "Social Aware Cognitive Radio Networks," *Social Network Analytics for Contemporary Business Organizations*, pp. 188-202. doi:10.4018/978-1-5225-5097-6.ch010
- [12] Bindhu, V. "Biomedical Image Analysis using Semantic Segmentation." *Journal of Innovative Image Processing (JIIP)* 1, no. 02 (2019): 91-101.
- [13] Dhaya, R., and R. Kanthavel. "A Wireless Collision Detection on Transmission Poles through IoT Technology." *Journal of trends in Computer Science and Smart technology (TCSST)* 2, no. 03 (2020): 165-172.
- [14] Mr. Ketan D. Bodhe, Mr. Himanshu V. Taiwade, Mr. Virendra P. Yadav, Mr. Nikesh V. Aote, "Implementation of Prototype for Detection & Diagnosis of Cotton Leaf Diseases using Rule Based System for Farmers", *Proceedings of the International Conference on Communication and Electronics Systems (ICCES 2018) IEEE Xplore Part Number: CFP18AWO-ART; ISBN: 978-1-5386-4765-3*.