UTPDS-ML: Utility Techniques for Plant Disease Identification Using Machine Learning

N. Mahendran and T. Mekala

Abstract Agriculture plays a significant and critical role in the Indian economy, employing over half of all Indian workers. India is often regarded as the world's leading producer of rice, pulses, spices, wheat, and spice crops. Farmers routinely examine grasses to discover illnesses, even using specialized fertilizers. Recent machine learning algorithms illustrate the monitoring of crops database and assist botanists in disease diagnosis with high precision. The suggested work combines the cross-central filter (CCF) method for image noise removal with the k-means algorithm for predicting the suspicious zone from the original region.

Keywords Agriculture · Machine learning · CCF · K-means

1 Introduction

Horticulture is a branch of agriculture that focuses on the advanced art and science of plant gardening. Seed germination, reproductive phase, panicle starting, blooming, maturing, ripening phase, and senescence are the plant budding phases. Quality of service (QoS) is an important function of agricultural products and is evolving from the knowledge of providing healthy food. By employing the quality auditing method of nutrition foods, quality assurance is important to authorize the quality of the agricultural product that monitors the natural defects [1].

Agriculture is the backbone of Indian economy, wherein 17% of India's GDP is based on the agriculture. Unfortunate changes in the atmospheric conditions, insufficient irrigation methods, inadequate familiarity in current expertise and pesticides make a variance in agriculture, irregular harvest of plants and harmful crops thereby threatening the global food safety. [2] Based on the details of Food and Agriculture Organization (FAO), the humanity population reaches 9.1 billion in 2050. Accordingly, agricultural production requirements to be enhanced up to 75% to convince the food needs of a progressively growing people. According to [3], the humanity

N. Mahendran (✉) · T. Mekala
M. Kumarasamy College of Engineering, Karur, Tamilnadu, India

P. Karrupusamy et al. (eds.), *Sustainable Communication Networks and Application*, Lecture Notes on Data Engineering and Communications Technologies 93,
https://doi.org/10.1007/978-981-16-6605-6_57

Fig. 1 Example of plant diseases in various plants

population is anticipated to hit 9.1 billion in 2050. Consequently, farming manufacture needs to be enlarged up to 75% to complete the food necessities of a gradually increasing people.

Sometimes, the person is not competent to discover the plant infection exactly as shown in Fig. 1. The technology like advanced image classification and image processing to be helpful to identify the infections accurately. The premature disease recognition in plants reduces the hazard of maximum crop malfunction and enlarges the yields. Plant diseases prediction method is to formulate crops that are maximum healthful, and minimize the health-related problems for people. In this paper, a plan is prepared in a competent and accurate manner which might be used to find the infected crops and also to compute the strength of disease in the plant. The proposed work combines the CCF method to remove the noise in the image and k-means algorithm to identify the objects in the image to predict the suspicious region from the original region.

2 Related Work

Author [4] describes smart horticulture and techniques used for prediction and detection of diseases in crops. The digital image processing and sensor networks play a vital role, in which plant leaf image is captured for under consideration of analysis[5–7] through modern communication technology too [6, 13–21]. The author describes the general flow of the image processing technique shown in Fig. 2. The first phase describes noise removal process, in which the input images are transformed to the necessary layout. The second phase is to segment the preprocessed image samples to allocate the image into dissimilar partitions. The important features are obtained from the segmented image and classified accordingly. The author finds the infectious crops by image segmentation and soft computing techniques as described in [4].

Author [8] presents a method for image processing which can be used to find the plant diseases using mobile app. The scheme isolates the perceptible that at dissimilar parts of the crops. The scheme tests the grape vineyard diseases based on leaves' image. The method extracts the injury features such as the particular gray level area, different spots, and histogram image pixels that can indicate RGB colors like red, blue, and green. A fixed threshold value is considered for an individual leaf disease. The RGB range is less than threshold, the particular pixel range is blackened else whitened. In this manner, the output image consist of a pattern of black spots, which can be predicted by evaluating the input data set and calculating whether the crops endure from the different disease or not. The technique attained an accuracy of 80%. The step by step process shown in Fig. 3.

Author [9] proposed an SVM algorithm for the identification of crop infections. The SVM is a type of nonlinear classifier that is mostly utilized in texture partition and pattern identification. The RGB feature pixel-counting method is to detect the affected areas like plant stem, leaf, and fruit diseases. The SVM nonlinear input data are matched into linearly divide data in high-dimensional space. The impure part is computed concerning percentage and the affected area is noticed based on SVM classifier. The algorithm step is shown in Fig. 4.

The above-discussed schemes are undemanding and do not have any systematic processes. Though, in the recent technology where systematic plays a vital role, there is a need to concern and to maximize the accuracy. Authors [10, 11] made up with a KNN and GLCM procedure for plant disease identification. After loading the input

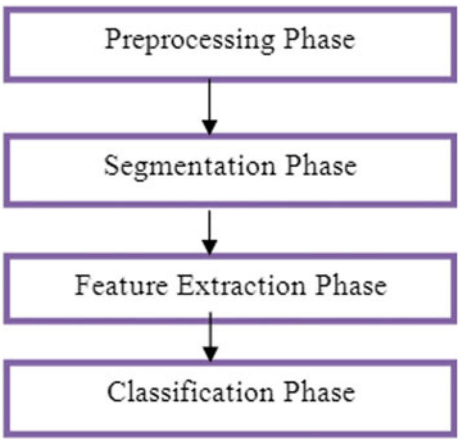


Fig. 2 Phase of İmage Processing

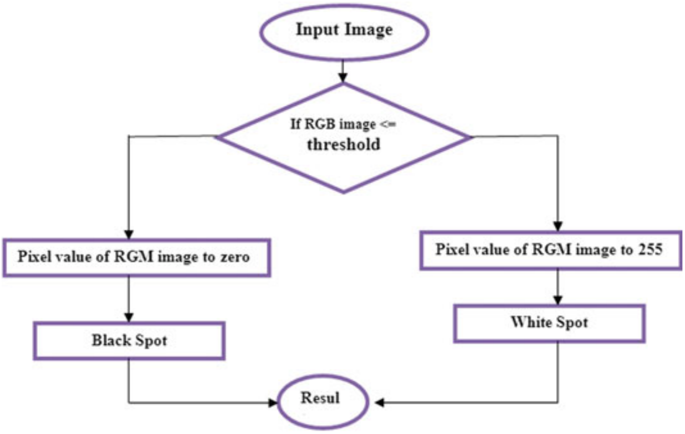
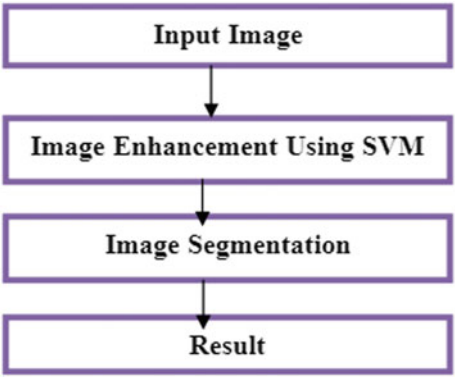


Fig. 3 Steps of threshold algorithm

Fig. 4 Steps in SVM classifier



plant leaf dataset, preprocessing of the image from data set observed by segmentation using k-means algorithm. GLCM helps to extract the feature and clustering is performed through KNN.

3 UTPDS-ML System Design

The scheme applies image processing method for crop disease identification; image segmentation and feature extraction have been observed through k-means algorithm as well as SVM-classification to categorize the crop disease as shown in Fig. 5.

The noise removal from the input images and transforming to the necessary layout is the first step called preprocessing, and it provides an insight to the process that has been observed by the enhancement of crop leaves. The next phase handles the

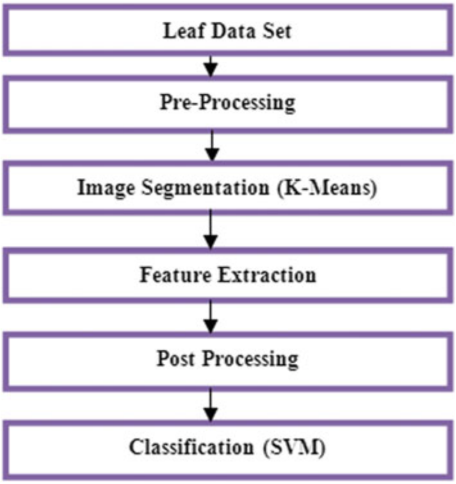


Fig. 5 Steps in proposed system

feature extraction followed by post-processing; false features are removed from the collection of obtained image.

4 Image Enhancement

The plant image enrichment is predictable to progress the dissimilarity between affected images and shrink noises in the leaf images. Maximum eminence leaf image is most significant for infection confirmation or classification to work accurately. In actual life, the feature of the plant image is unnatural by noise.

4.1 *Binarization Process (B_P)*

B_P is the procedure that decodes a grayscale input image into binary (1's and 0's) image. This improves the dissimilarity among the precious places in a given input image and formulates the potential feature extraction. Hence, B_P involves evaluating grayscale values of every pixel, if values is larger than the inclusive threshold set binary value as 1, else 0.

4.2 ROI Extraction (R_E)

In the image dataset, the region of interest (ROI) is the part of an input image, which is significant for the extraction of feature points. R_E has two different morphological operations like OPEN (OP) and CLOSE (CL). The OP operation can enlarge images and eliminate peaks initiated by background noise. The CL operation can minimize and reduce small differences.

4.3 Post-processing (P_P)

The segmented images may have numerous specious points. This might happen due to the presence of lines could not be improved even after enhancement results in negative points. The redundant points are eliminated in the P_p stage[12].

5 Classification and Clustering

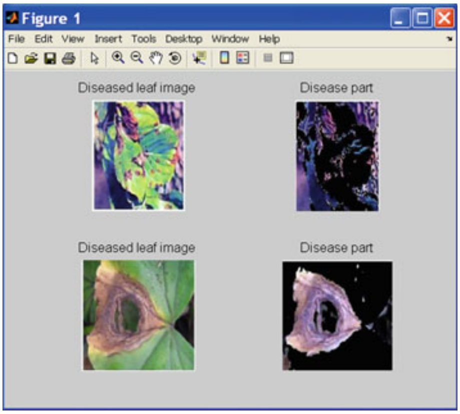
Clustering with k-means scheme divides the input image into different partitions with more similarity along with the substance of awareness. Structures of the clustered image support in a peaceful identification of dynamic or diseased illustrations.

Algorithm: K-means clustering

Input: set of leaf images $\{S_{L1}, S_{L2}, \ldots S_{Ln}\}$
Output: cluster the similar objects into k groups

do
    Arbitrarily choose k objects from S_{Ln} as the initial cluster centers;
Repeat
    Divide the input dataset into similar groups
    Calculate the Euclidean distance (d)
    Every pixel is referred to its closest K-centroid
while (no change);

Fig. 6 Image segmentation



Algorithm: SVM Classification

Input: cluster data C_d = {group of similar data}
Output: classification according to their relevance

do (violating point)
    Find disease name
    cluster data C_d =U C_d
Repeat
    violator
while (until pruned);

6 Results and Discussion

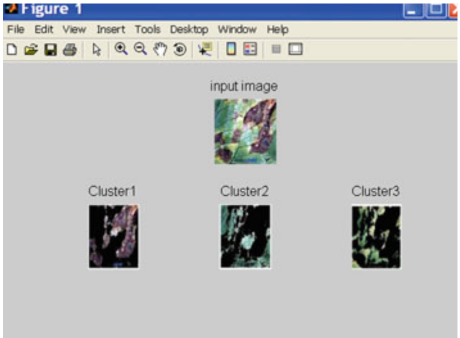
Figure 6 shows the feature extraction-based segmentation process using K-means algorithm. The algorithm computes the centric point through Euclidian distance and images are grouped through similarity. After clustering, the classification processed by SVM method.

Clusters formation completed based on a k-means clustering. This formation is transformed into a different feature vector (F_V) and transformed to the support vector machine for identification of the crop disease shown in Fig. 7.

7 Conclusion

The proposed work mainly focuses on identifying the plant disease and to construct an effective method. Based on the SVM classification, the corrective measures are reported. The proposed process entirely removes the necessity to check with

Fig. 7 Formation of clusters



a specialist to resolve whether a crop is disease affected or not. The humanity is changing near an age based on knowledge. Farmers frequently look on the grasses to identify diseases, even using exclusive fertilizers. The proposed machine learning methods demonstrate the crop monitoring database and support botanists in disease diagnosis with a large deal of precision.

References

1. A.B. Mankar, Datamining—an evolutionary view of agriculture. Int. J. Appl. Innovation Eng. Manage. **3**(3), 102–105 (2014)
2. Food and Agriculture Organization (FAO). Declaration of the World Summit on Food Security. Rome: Food and Agriculture Organization; 2009
3. K. Golhani, S.K. Balasundram, G. Vadamalai, B. Pradhan, A review of neural networks in plant disease detection using hyperspectral data. Inform. Process. Agric. **5**(3), 354–371 (2018). https://doi.org/10.1016/j.inpa.2018.05.002
4. V. Singh, A.K. Mishra, Detection of plant leaf diseases using image segmentation and soft computing procedia computer science. Inform. Process. Agric. **65** (2016)
5. N. Mahendran, T. Mekala, An efficient optimization technique for scheduling in wireless sensor networks: a survey. Stud. Comput. Intell. 223–232. https://doi.org/10.1007/978-981-10-8797-4_24
6. N. Mahendran, T. Mekala, A survey: sentiment analysis using machine learning techniques for social media analytics. Int. J. Pure Appl. Math. **118**(8), 419–423 (2018)
7. M. Natarajan, S. Subramanian, A cross-layer design: energy efficient multilevel dynamic feedback scheduling in wireless sensor networks using deadline aware active time quantum for environmental monitoring. Int. J. Electron. **106**(1), 87–108 (2019).https://doi.org/10.1080/00207217.2018.1501615
8. N. Petrellis, A smart phone image processing application for plant disease diagnosis, in *2017 6th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, Thessaloniki, pp. 1–4. https://doi.org/10.1109/MOCAST.2017.7937683
9. S. Varshney, T. Dalal, A novel approach for the detection of plant diseases, Int. J. Comput. Sci. Mobile Comput. **5**(7), 44–54 (2016). ISSN 2320-088X
10. Ch.R. Babu, D. Srinivasa Rao, V.S. Kiran, N. Rajasekhar, Assessment of plant disease identification using GLCM and KNN algorithms. Int. J. Recent Technol. Eng. (IJRTE) **8**(5) (2020). ISSN 2277-3878

11. B.J. Sowmya, C. Shetty, S. Seema, K.G. Srinivasa, Utility system for premature plant disease detection using machine learning. Hybrid Comput. Intell. 149–172 (2020).https://doi.org/10.1016/b978-0-12-818699-2.0
12. T. Mekala, M. Natarajan, in *Proceedings of IEEE International Conference on Computer Communication and Systems ICCCS14* (2014), pp. 151–155
13. T. Mekala, N. Mahendran, Improved security in adaptive steganography using game theory. Int. J. Pure Appl. Math. **118**(8), 111–116 (2018)
14. N. Mahendran, S. Shankar, T. Mekala, LSAPSP: load distribution-based slot allocation and path establishment using optimized substance particle selection in sensor networks. Int. J. Commun. Syst. (2020). https://doi.org/10.1002/dac.4343
15. N. Mahendran, S. Shankar, T. Mekala, EMA-PRBDS: efficient multi-attribute packet rank based data scheduling in wireless sensor networks for real time monitoring systems. Int. J. Electron **106** (2019). Taylor and Francis.https://doi.org/10.1080/00207217.2019.1692244
16. N. Mahendran, T. Mekala, Energy-distance and degree: multi-objective based optimized clustering and route discovery in wireless sensor network. IOP Conf. Ser. Mater. Sci. Eng. **872**(1) (2020)
17. R. Gomathi, N. Mahendran, An efficient data packet scheduling schemes in wireless sensor networks, in *Proceeding 2015 IEEE international Conference on Electronics and Communication Systems (ICECS'15)* (2015), pp. 542–547. ISBN 978-1-4799-7225-8
18. R. Gomathi, N. Mahendran, S. Shankar, A survey on real-time data scheduling schemes in wireless sensor networks. Int. J. Appl. Eng. Res. **10**(9), 8445–8451 (2015)
19. N. Mahendran, T. Mekala, Improving energy efficiency of virtual resource allocation in cloud datacenter. Indian J. Sci. Technol. **11**(19), 1–8 (2018)
20. S. Vanithamani, N. Mahendran, Performance analysis of queue based scheduling schemes in wireless sensor networks, in *Proceeding 2014 IEEE international Conference on Electronics and Communication Systems (ICECS'14)* (2014), pp. 1–6. ISBN 978-1-4799-2320-5
21. N. Mahendran, Collaborative location based sleep scheduling with load balancing in sensor-cloud. Int. J. Comput. Sci. Inform. Sec. **14**, 20–27 (2016)