CROP DISEASE DETECTION USING COMPUTER VISION

Kaustubh Shah
BTech, Department Of Electronics And Communication Engineering
Motilal Nehru National Institute Of Technology Allahabad, Prayagraj

Abstract
Agriculture is a rich source to meet growing food demand of increasing population. To protect the food for the overall population, it is important to predict crop diseases at the onset. This study is aimed to inform farmers about emerging technologies to reduce crop diseases. Approaches related to computer vision can detect those diseases in most common vegetables like tomatoes. This study proposes the most accurate algorithm to detect crop diseases as per early symptoms.

Keywords*: agriculture, crop diseases, computer vision, algorithm, food demand*

1. Introduction

Advanced technologies have enabled production of sufficient food to meet the growing demand. Crop safety and security is still a major concern. Factors like crop diseases, changing climate, reduced pollinators, etc. are troubling the farmers. It becomes important to address those challenges (Krithika & Veni, 2017; Prakash et al., 2017).
Analysis and detection with the latest technology can help farmers to avoid those problems. In this day and age, several latest technologies have helped in preventing diseases to the crops (Prakash et al., 2017; Mishra et al., 2017; Pooja et al., 2017). Plant diseases are known to be the major threat to public health and lives as they may cause famines, droughts, and epidemics. These diseases can cause significant losses to farmers. This way, technologies like Machine Learning can help prevent crop diseases (Yu et al., 2020). Some researchers and scientists have observed key challenges to analyze crop diseases –

- High quality image is required for processing
- Dataset must be publicly available
- Leaf samples may be affected by noisy data
- Leaves’ color changes in different environments
- Samples should go through proper testing to identify diseases in segmentation
- Different crops may have different diseases. So, it will not be easy to detect diseases.
- Another major challenge is classification of leaf disease.

To address those challenges, the model proposed in this study would combine machine learning with image processing (IP) for better accuracy.

2. Literature Reviews

Krithika & Veni. (2017) detected leaf disease on cucumber leaves by image resizing, color-space conversion, and contrast enhancement using multiclass SVM. They performed K-Means clustering for feature extraction and segmentation with GLCM. Narmadha & Arulvadivu (2017) used digital cameras to capture images and enhanced images using median filters. They used SVM for classification and segmentation is done using K-Mean clustering.
Segmentation is performed by Pooja & Kanchana. (2017) to detect the infected region, i.e., area of interest. It is performed with the K-Means clustering model. RGB is converted into HIS with Otsu’s detection. Later, spot and boundary detection are performed using segmentation. Gupta et al. (2017) used normalization and contrast adjustment for image pre-processing. They converted and transformed the color into YCBCR and performed Bi-level thresholding. The HMM and GLCM are used for feature classification and extraction (Yu et al., 2021).
Dhaware & Wanjale (2017) segmented image of crop for backdrop subtraction. They conducted the classification approach with ANN, SVM, and KNN approaches. KNN classified the samples with closest distance among testing and trained subjects (Esmaeel, 2018). Singh et al. (2017) took samples of beans/rose (to detect bacterial infection), banana leaves (to detect early scorch), lemon leaves (to detect sunburn), and beans (to detect fungal infection) captured with a digital camera. They got segmented imaging with a genetic model. They adapted color co-occurrence for extracting features from segmented pictures. They also used the SVM classifier and “Minimum Distance Criterion” for classification. They have achieved 97.6% accuracy.

2.1 Research Gap

A lot of studies have been conducted to detect leaf diseases using various approaches over the years and they have constantly attracted attention of researchers to do further studies. This study uses computer vision to detect crop disease infection with the combination of image processing and machine learning.

2.2 Research Question

- How to detect crop disease using computer vision?

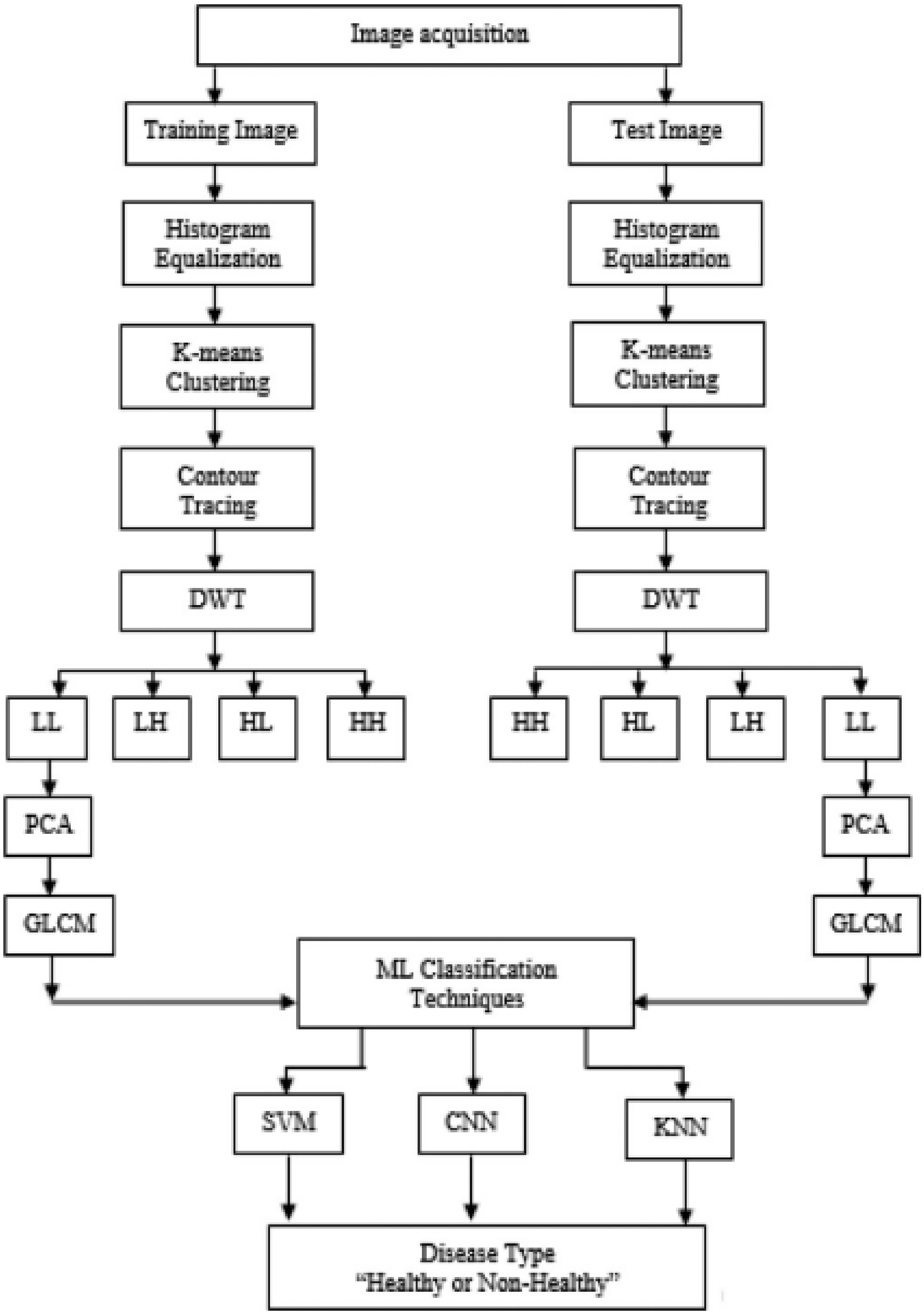
2.3 Research Objective

- To propose ML and IP based model to detect crop disease using computer vision

3. Research Methodology

To fulfill the research objectives, this study uses secondary data collected from relevant studies to propose the model based on ML and image processing methods to detect leaf diseases. In the model which combines CNN, GLCM, PCA, and DWT, tomato samples have been considered with 6 disorders to determine the accuracy and classify the same as Healthy/Unhealthy (Figure 1). Tomato samples are resized for image processing to 256x256 pixels to retain equal size.

Figure 1 – Machine Learning model for image processing using Computer Vision



Source – Harakannanavar et al. (2022)

4. Analysis of Study

Yu et al. (2021) and Vadivel & Suguna (2022) considered tomato leaves for the village database. Those plants were infected with different diseases. They took pictures of tomato leaves with 6 disorders to detect leaf disease. Figure 2 illustrates the leaf samples in the database.

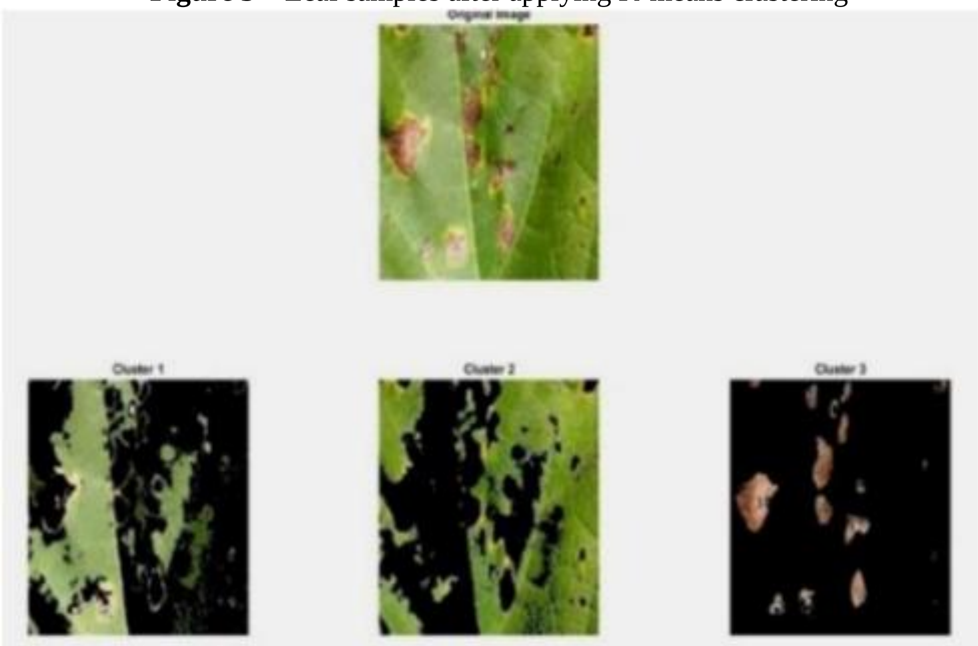
Figure 2 – Diseases detected in tomato leaves

Source - Yu et al., 2021; Vadivel & Suguna, 2022

4.1. Preprocessing

K-means clustering was applied on leaf pictures to detect the infected area (Esmaeel, 2018; Nagashetti et al., 2021; Arya et al., 2018). This model can get the data center of the picture, make the clusters, and calculate the distance from other clusters. Figure 3 illustrates leaf samples after using K-means clustering (Chandrasekaran et al., 2021).

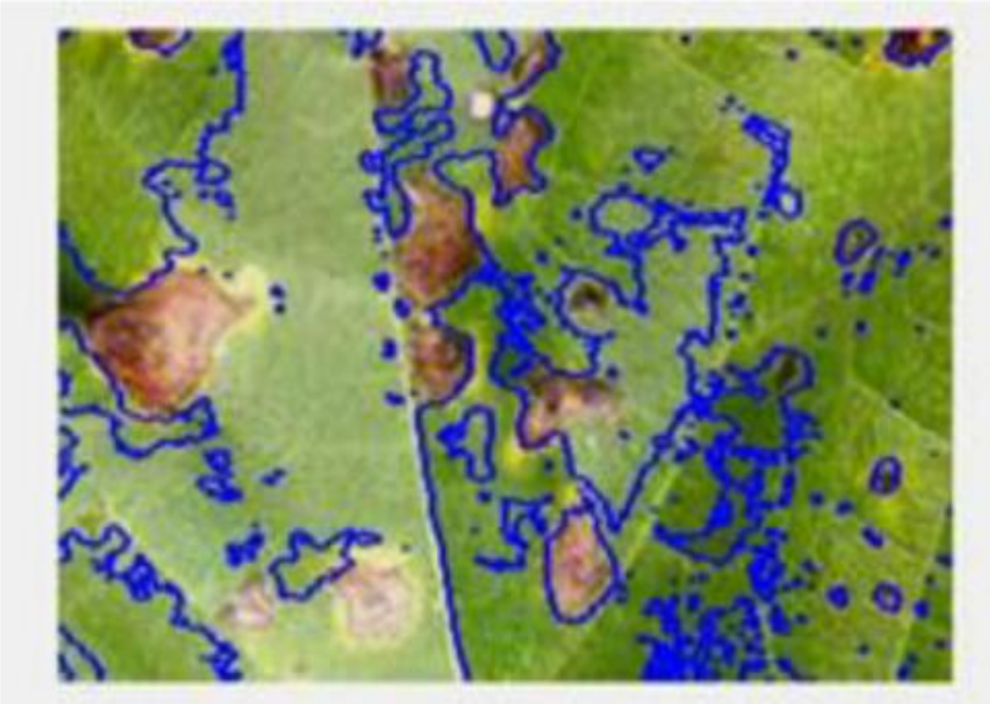
Figure 3 – Leaf samples after applying K-means clustering



Source – Chandrasekaran et al. (2021)

On digital leaf samples, contour tracing is done to extract the overall data of their shape (21,22). The characteristics of the contour are analyzed and used to perform pattern classification. It is usually helpful to determine the efficiency of the process of feature extraction (Chandrasekaran et al., 2021). Figure 4 illustrates the images of contour tracing.

Figure 4 – Sample after performing contour tracing

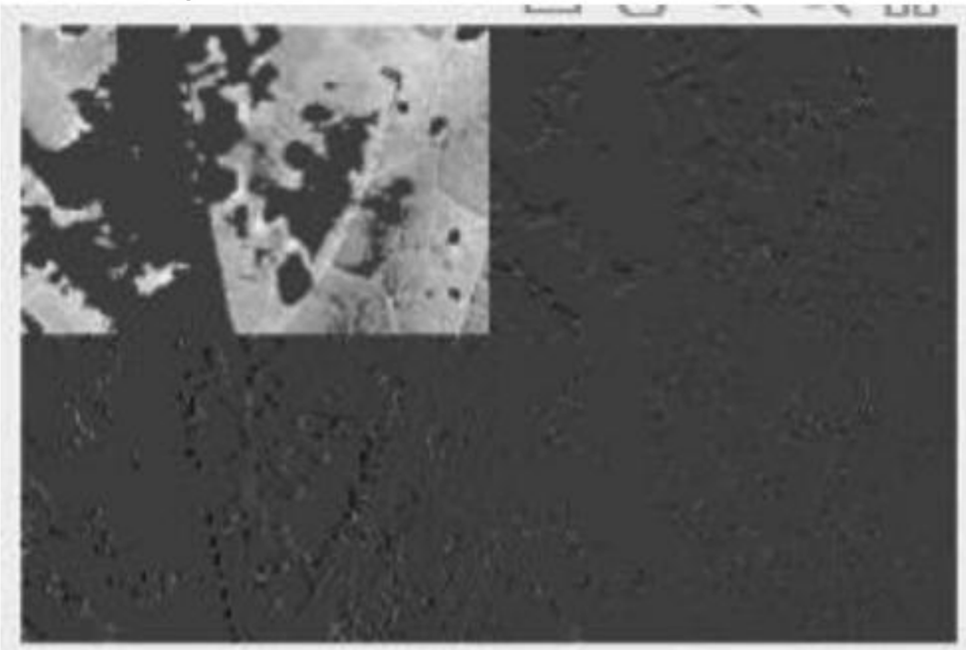


Source – Hossain et al., 2018; Zhang et al., 2021

4.2. Feature extraction

Nema & Dixit (2018) performed "Discrete Wavelet Transform (DWT)" on tomato samples to gather more important features. It decomposes into higher frequency (HH) and lower frequency components and their sub-bands. Maximum data is available for the LL part of DWT as compared to the HH part of DWT (Figure 5).

Figure 5 – Discrete Wavelet Transform Decomposition

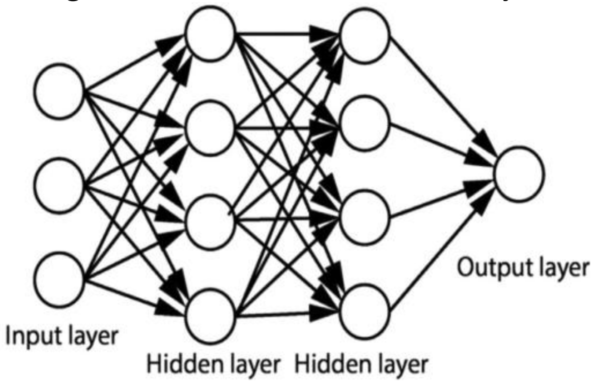


Source - Nema & Dixit (2018)

4.3. Classification of Samples

CNN, SVM and KNN are machine learning techniques used to classify the leaf samples. CNN is used for data processing and its architecture consists of "output layer (OL), input layer (IL) and hidden layer (HL)" (Gharpankar, 2020; Saleem et al., 2020; Pushpa et al., 2021). The HL consists of the RELU layer (for activation and pooling of data), convolutional layer, and normalization layer (Kaleem, 2021; Vadivel & Suguna, 2022). It has cross correlation architecture instead of convolution for indices in the matrix (Devi, 2021). Figure 6 illustrates a typical 3-layer neural network.

Figure 6 – A Neural Network with 3 layers



Source - Harakannanavar et al. (2022)

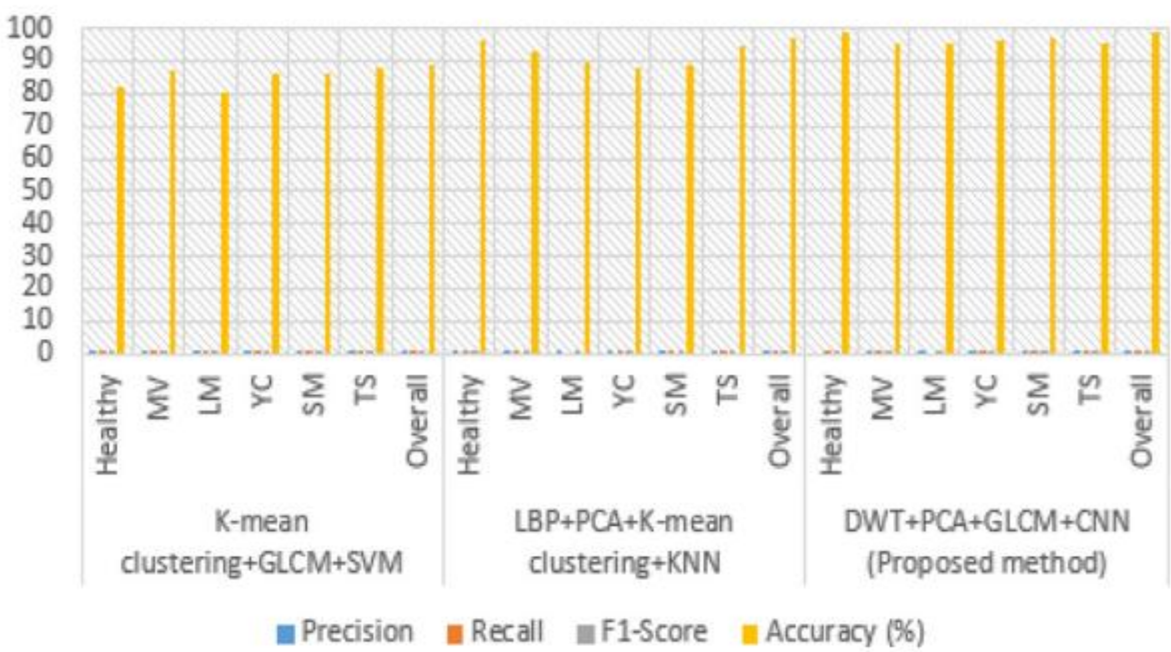
5. Results

In the village dataset, the tomato leaf samples are considered for evaluating the model proposed by Harakannanavar et al. (2022). They took 100 samples of heavy leaf for testing the model, which identified 99 samples with 99% accuracy. The model identified 100 samples with 100% accuracy in the 100 tomato samples infected with Mosaic virus. The model came up with 100% accuracy in leaf mold. The model achieved 99% accuracy in 100 yellow curl samples. In the same way, 99% and 100% accuracy has been achieved in "Spotted spider mite" and "Target Spot" samples, respectively. Overall, researchers have tested 600 tomato village dataset samples on the model and 99.5% accuracy has been achieved.

The researchers have tested and trained dataset samples for validating the model. They recommended the following software and hardware requirements for sample testing –
OS – Windows 10
Language – Python
Core – Nvidia GPU
Dataset – Tomato village dataset samples
Libraries – OpenCv, Image data generator, NumPy, Tensorflow, Mat plot

The overall performance of the model is evaluated using parameters like F1 score, Recall, and Precision. Overall, 600 samples have been used to test the proposed model on the dataset of tomato leaf disease. Figure 9 compares the same with other models.

Figure 9 – Comparing Proposed model with other existing models



Source - Harakannanavar et al. (2022)

Harakannanavar et al. (2022) compared the proposed model with other algorithms and found 99.09% accuracy of the proposed model, i.e., better than other existing models. Using ML classification and computer vision, the proposed model is compared with methodology suggested by Kanabur et al. (2020), Hossain et al. (2018), and Vadivel & Suguna (2021). The accuracy achieved is better in the proposed model in comparison to other models.

6. Conclusion

Computer vision techniques have been proposed in this study for RGB conversion to gray. In pre-processing, contour tracing and clustering are used. The researchers have also proposed "Principal Component Analysis, Discrete Wavelet Transform, and GLCM" to gather the data of leaf samples. Machine learning models like CNN, SVM, and K-NN can differentiate Healthy/Unhealthy leaves. The proposed model can be analyzed with CNN technique and researchers can achieve the desired accuracy. In future, fusion techniques can be employed to further improve the proposed model to extract more important features and examine other crop samples.

References

1. Krithika, P., & Veni, S. (2017, March). Leaf disease detection on cucumber leaves using multiclass support vector machines. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 1276-1281). IEEE.
2. Prakash, R. M., Saraswathy, G. P., Ramalakshmi, G., Mangaleswari, K. H., & Kaviya, T. (2017, March). Detection of leaf diseases and classification using digital image processing. In *2017 international conference on innovations in information, embedded and communication systems (ICIIECS)* (pp. 1-4). IEEE.
3. Mishra, B., Nema, S., Lambert, M., & Nema, S. (2017, March). Recent technologies of leaf disease detection using image processing approach—A review. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1-5). IEEE.
4. Pooja, V., Das, R., & Kanchana, V. (2017, April). Identification of plant leaf diseases using image processing techniques. In *2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)* (pp. 130-133). IEEE.
5. Yu, K., Lin, L., Alazab, M., Tan, L., & Gu, B. (2020). Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system. *IEEE transactions on intelligent transportation systems*, *22*(7), 4337-4347.
6. Narmadha, R. P., & Arulvadivu, G. (2017, January). Detection and measurement of paddy leaf disease symptoms using image processing. In *2017 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-4). IEEE.
7. Gupta, V., Sengar, N., Dutta, M. K., Travieso, C. M., & Alonso, J. B. (2017, July). Automated segmentation of powdery mildew disease from cherry leaves using image processing. In *2017 International Conference and Workshop on Bioinspired Intelligence (IWOBI)* (pp. 1-4). IEEE.
8. Yu, K., Tan, L., Lin, L., Cheng, X., Yi, Z., & Sato, T. (2021). Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote E-health. *IEEE Wireless Communications*, *28*(3), 54-61.
9. Dhaware, C. G., & Wanjale, K. H. (2017, January). A modern approach for plant leaf disease classification which depends on leaf image processing. In *2017 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-4). IEEE.
10. Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information processing in Agriculture*, *4*(1), 41-49.
11. Esmaeel, A. A. (2018, April). A novel approach to classify and detect bean diseases based on image processing. In *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)* (pp. 297-302). IEEE.
12. Vadivel, T., & Suguna, R. (2022). Automatic recognition of tomato leaf disease using fast enhanced learning with image processing. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, *72*(1), 312-324.
13. Nagashetti, S. M., Biradar, S., Dambal, S. D., Raghavendra, C. G., & Parameshachari, B. D. (2021, October). Detection of disease in Bombyx mori silkworm by using image analysis approach. In *2021 IEEE Mysore Sub Section International Conference (MysuruCon)* (pp. 440-444). IEEE.
14. Arya, M. S., Anjali, K., & Unni, D. (2018, January). Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino. In *2018 International Conference on Power, Signals, Control and Computation (EPSCICON)* (pp. 1-5). IEEE.

15. Chandrasekaran, G., Nguyen, T. N., & Hemanth D, J. (2021). Multimodal sentimental analysis for social media applications: A comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(5), e1415.
16. Hossain, S., Mou, R. M., Hasan, M. M., Chakraborty, S., & Razzak, M. A. (2018, March). Recognition and detection of tea leaf's diseases using support vector machine. In *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 150-154). IEEE.
17. Zhang, J., Yu, K., Wen, Z., Qi, X., & Paul, A. K. (2021). 3D reconstruction for motion blurred images using deep learning-based intelligent systems. *Computers, Materials & Continua*, *66*(2), 2087-2104.
18. Nema, S., & Dixit, A. (2018, December). Wheat leaf detection and prevention using support vector machine. In *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)* (pp. 1-5). IEEE.
19. Gharpankar, S. B. (2020). Plant Leaf Disease Detection using Convolutional Neural Networks. *Int. J. Res. Appl. Sci. Eng. Technol*, *8*(11), 494-499.
20. Saleem, M. H., Khanchi, S., Potgieter, J., & Arif, K. M. (2020). Image-based plant disease identification by deep learning meta-architectures. *Plants*, *9*(11), 1451.
21. Pushpa, B. R., AV, S. H., & Ashok, A. (2021, July). Diseased leaf segmentation from complex background using indices based histogram. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1502-1507). IEEE.
22. Kaleem, M. K. (2021). A modern approach for detection of leaf diseases using image processing and ml based svm classifier. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(13), 3340-3347.
23. Devi, P. R. (2021, August). Leaf Disease Detection Using Deep Learning. In *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1797-1804). IEEE.
24. Kanabur, V., Harakannanavar, S. S., Purnikmath, V. I., Hullole, P., & Torse, D. (2020). Detection of leaf disease using hybrid feature extraction techniques and CNN classifier. In *Computational Vision and Bio-Inspired Computing: ICCVBIC 2019* (pp. 1213-1220). Springer International Publishing.
25. Hossain, S., Mou, R. M., Hasan, M. M., Chakraborty, S., & Razzak, M. A. (2018, March). Recognition and detection of tea leaf's diseases using support vector machine. In *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 150-154). IEEE.