IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 1, March 2024, pp. 925~934
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp925-934

Extracting features of tomato viral leaf diseases using image processing techniques

Sanjeela Sagar[1], Jaswinder Singh[2]
[1]Department of Information Technology, Vidyalankar School of Information Technology, Mumbai, India
[2]Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, India

Article Info

Article history:

Received Aug 10, 2023
Revised Oct 24, 2023
Accepted Nov 15, 2023

Keywords:

Clustering in image processing
Color detection
Image processing techniques
Moments in image processing
Tomato disease detection

ABSTRACT

Agriculture is the main livelihood of Indians. More than 50% of Indian population Is dependent on it and it contributes about 18% of Indian gross domestic product (GDP). According to Inc42, the agricultural sector of India is predicted to increase to US$ 24 billion by 2025. With the increase in population, the demand for food also increases, but more than 30% of crops get affected due to crop diseases. Overall, India lost approximately five million hectares of crop area to flash floods, cyclonic storms, floods, cloudbursts, and landslides till 2021. In that case, there is a need to prevent crops from diseases to fulfil demand supply ratio. This paper presents the feature extraction of tomato viral leaf diseases using various image processing techniques. Most of the research uses Convolutional Neural networks to extract the features of these diseases, but these neural networks are not performing much accurately in real scenarios, so there is a need to extract the features using image processing methods. During the study, it is found that these diseases have different colours, shapes and textures and these features can be used with convolution neural networks to bring more accurate results in real scenarios.

Corresponding Author:

Sanjeela Sagar
Department of Information Technology, Vidyalankar School of Information Technology
B-401, Mahavir Vihar, Plot No-10, Sector-25, Nerul East, Navi Mumbai, Maharashtra, India
Email: sanjeela.sagar@vsit.edu.in

1. INTRODUCTION

Tomato plants are consumed by almost everyone in the world. India is the 2nd largest producer of tomato in the world, but the crop gets affected by various diseases due to which production per hectare is less as compared to other countries. The various solutions which exist to identify these diseases take lots of time and need expertise. With the advent of technology, we can get much better solutions. In India following are the major states producing tomatoes with their production in hectares and percentage share as given in Figure 1 and Figure 2. When the production per hectare is compared with other countries, it is seen that India is producing less tomatoes. There is a need to increase the production per hectare. It is found that majority of the times the tomato plants get affected due to various diseases which lead to 30% to sometimes more than 80% of crop loss. There is a great demand to detect and classify these diseases at the early stages to save the crop.

Several studies have done previously in which dataset from Plant Village, a publicly available dataset is taken [1]-[3]. The various image processing techniques like segmentation [1], color extraction, shape, contrast, annotation are the features used. Transfer learning was used to categorize the images [2]. To identify the shape and pattern, gray color was used along with 4-layers of color channels [3]. Image augmentation done on real data in the lab, then trained and got the accuracy of more than 99% [4]. The texture and other color characteristics were extracted to spot the diseases in tomato leaves [5]. In some studies, thresholding is used to differentiate the diseased and non-diseased leaves [6]. Other image processing techniques like annotation, blur, stretch, contrast, rotation were used and it is observed that convolution neural network (CNN) then has given promising results [7]. In another study, it is observed that image smoothening with Multi-class support vector machines (SVM) has given more than 93% accuracy [8]. Sabrol and Satish [9] various image processing features like colour, shape and texture were used to diagnose the diseased sections on the leaves. Image segmentation and clustering with deep learning achieved more than 98% accuracy. The paper [10] is a review of tomato diseases methods with the datasets of tomatoes available till date.

The study in this paper reveals that the majority of work to detect the diseases automatically can be done through deep learning methods. The study in [11] is review of various disease detection methods and focuses on the importance of image processing methods to be included for detecting and diagonising the disease type. Wang *et al.* [12] the focus is on various deep learning algorithms mainly Region based-CNN to segment the affected parts. It also shows R-CNN combined with object detection algorithm to not only find the type of diseases, but also to detect and mark the infected sections. The study done in Bangladesh [13] supports the use of CNN in identifying the tomato diseases. The study is mainly done for bacterial and leaf spot, tomato blight, mosaic virus and yellow leaf curl virus and in the image processing methods- the most used method image augmentation is used and found more than 90% accuracy for the above mentioned diseases. In Thapar University, [14] the study was done to detect the tomato diseases using image thresholding to segment the infected part and k-means clustering to detect the disease and then the percentage of infected portion in a leaf was detected. Four layer deep CNN [15], was used with different filters in every layer like 10 filters in first layer, 20 in second, 64 in third and 30 in fourth layer to detect the various plant diseases that is, tomato, pepper and potato and got an accuracy of more than 98%. In 2015, the study was done to extract the features of tomato diseases using gray-level-coocuurence matrix (GLCM) [16] and then training through SVM different kernel functions like linear, quadratic, Multi-Layer Perceptron (MLP), radial basis function (RBF) and this study proved that linear kernel shows the highest accuracy with 99.8%.

Gaussian filter was used for image smoothening, GLCM for feature extraction and then segmentation was done for similarity and disimilarity texture analysis with various CNN implementations like AlexNet, artificial neural network (ANN) and deep where deep CNN has shown remarkable detection of tomato diseases [17]. A recent study was done in Malaysia on pre-trained CNN's to detect the diseases in tomato plant in which different types of CNN's like Densenet-120, ResNet (18, 30, 50, 101), visual geometry group (VGGNet), SqueezeNet were compared and it was found that Densenet-120 was able to detect more accurately with more than 99% accuracy [18]. Bensaadi and Louchene [19], the four tomato diseases namely tomato leaf mould, septoria, tomato early blight and tomato powdery mildew were studied and trained using low cost three layered CNN and an accuracy of 99% approximately is achieved. There are methods to improve the quality of images as discussed in [20] like fuzzy noise filter reduction is used which is capable of preserving the edges as well. Baral *et al.* [21], to detect and classify roadsise vegetation, computer vision annotation tool is used to annotate the images into five different classes like Vegetation that requires repair, tree, vegetation, water accumulation and background and then trained using U-Net model and achieved an accuracy of more than 75%. The study of pest disease is done by Nazari, *et al.* [22] in which more than 170 infected leaves were studied against 100 healthy leaves and used adaptive network-based fuzzy inference system to diagonse the infected leaves. The study in [23] used stereo machine vision and deep learning to quantify stem diameter, branch angle, and branch diameter of loblolly pine trees with promising results. From the above literature view, following image processing methods are found to be used to detect the leaf diseases:

- Image segmentation: Image segmentation is the process of partitioning an image into different regions or objects. It can be used to differentiate healthy plant parts from diseased areas as done in. The various techniques that can be applied to segment the image are thresholding [24], clustering [25], or edge detection [26].
- Colour-based analysis: Generally, the diseases on the leaves can be identified by analysing the colour properties of the image, such as hue, saturation, and intensity as diseases occur in the plant issue and it affects the colour of the leaf, it is possible to detect abnormal colour regions that may indicate the presence of disease [27].
- Texture analysis: The texture of the plant leaves gets affected by the various diseases [28]. Texture analysis techniques, such as gray-level co-occurrence matrix (GLCM) or wavelet transform, can be applied to extract texture features that distinguish healthy and infected regions.
- Pattern recognition: Certain diseases exhibit distinctive patterns of lesions or spots, which can be identified using pattern recognition algorithms [29], [30].
- Machine learning: Image classification algorithms, such as transfer learning, CNNs, RCNN, deep CNN [31] can be trained on a dataset of labeled images to automatically recognize and classify plant diseases. This approach requires a substantial amount of labeled training data to achieve accurate results.

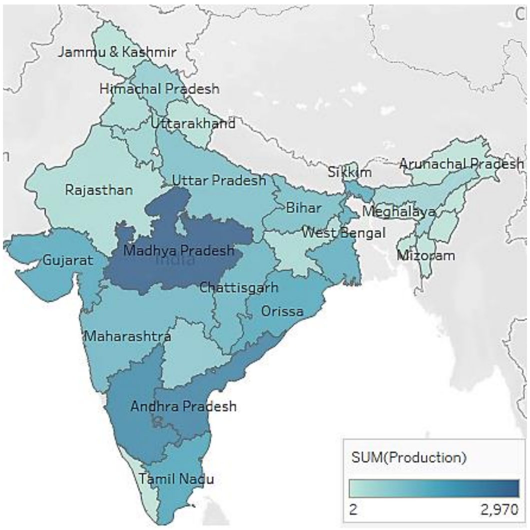


Figure 1. Indian production of tomato in 2021-22

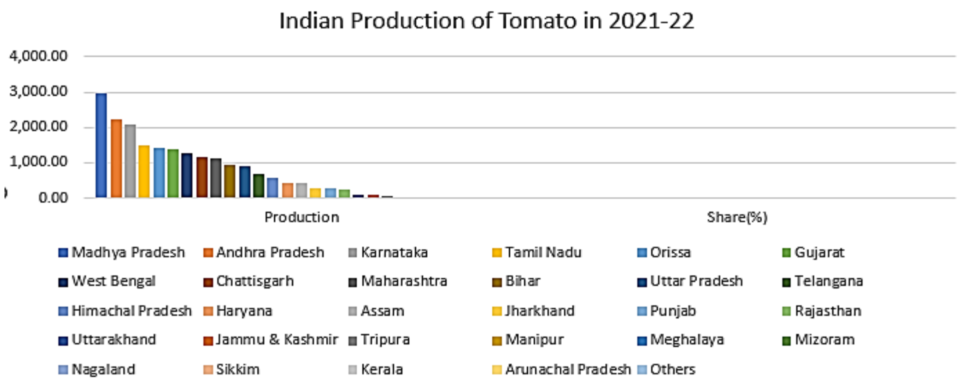


Figure 2. Indian production of tomato in 2021-22 (NHB, India)

2. METHOD

2.1. Data collection

Data of tomato healthy and diseased leaf images was collected from Plant village, a publicly available repository. 1000 images each of tomato mosaic virus, tomato target spot, tomato septoria leaf spot, tomato yellow leaf curl virus is taken. A sample of these diseases is shown in Figure 3. Images were annotated and then used for extracting features. This study focuses on the extracting features of diseases detection and classification of the following tomato viral diseases using image processing methods:

Figure 3. Various tomato viral diseases studied in this study

2.2. Extraction of features

The following image processing methods are used to extract the features:

a) Background removal: The background was removed from all the images as shown in Figure 4(a) 'image with background' to 'image after background removal in Figure 4(b). Algorithm for background removal:
- Step 1: Two clusters were created, namely, first, for leaf pixels and another for unwanted background using color space method.
- Step 2: Otsu's thresholding method was used to create binary image.
- Step 3: Binary image was used as a mask.
- Step 4: Background removed.

(a) (b)

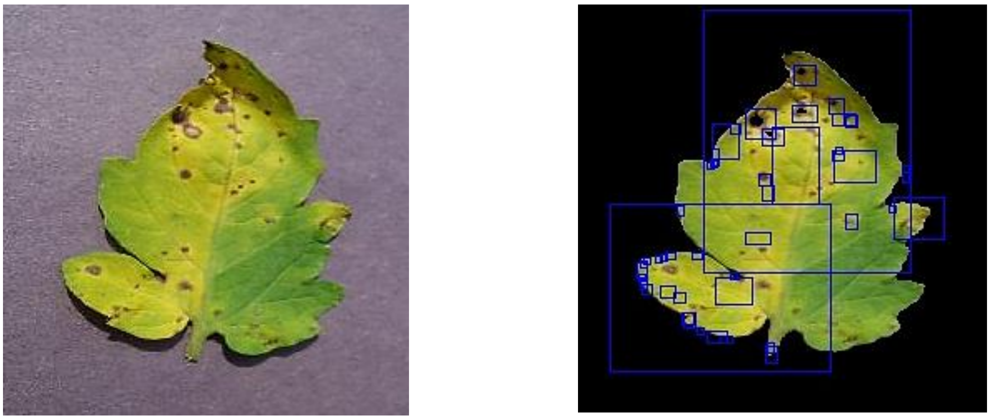
Figure 4. Tomato septoria leaf spot (TSLS) with (a) background and (b) after background removal

b) Tomato mosaic virus (TMV) detection: In TMV, leaves become yellow or light green, so the major features extracted were to detect these colours in the leaf portion as given in Figure 5(a) and then create a bounded box around these spots as given in Figure 5(b). Algorithm for TMV detection:
- Step 1: Detection of yellow colour: To detect yellow colour, a mask is created with yellow colour range.
- Step 2: Find the contours in the mask.
- Step 3: If the contours were detected, then a rounded box created around that spot.
- Step 4: Detection of green colour: To detect light green colour, a mask is created with the colour range.
- Step 5: Find the contours in the mask.
- Step 6: If the contours were detected, then a rounded box will highlight that spot as shown in Figure 5(b).

(a) (b)

Figure 5. TMV (a) leaf and (b) spots highlighted after the algorithm executed

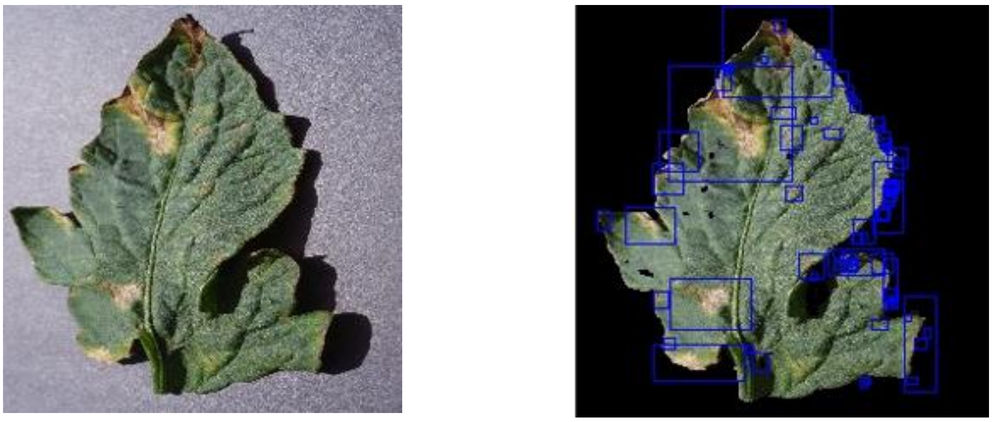
c) Tomato septoria leaf spot (TSLS) detection: In TSLS, circular spots of dark-brown margin with grey at the centres is spotted. These brown spots are surrounded by greenish yellow colour spots as seen in Figure 6(a), so the major features extracted were to detect these colours in the leaf portion and then create a bounded box around these spots as shown in Figure 6(b). Algorithm for TSLS detection:
   - Step 1: Detect brown colored circular spots with grey at the centre of the circle.
   - Step 2: Detect the greenish yellow colour spots around these spots.
   - Step 3: Create masks of brown and greenish yellow colours. Detect the contours.
   - Step 4: If the contours are detected, then a rounded box is created around that spot.



(a) (b)

Figure 6. TSLS (a) leaf and (b) spots highlighted after the algorithm executed

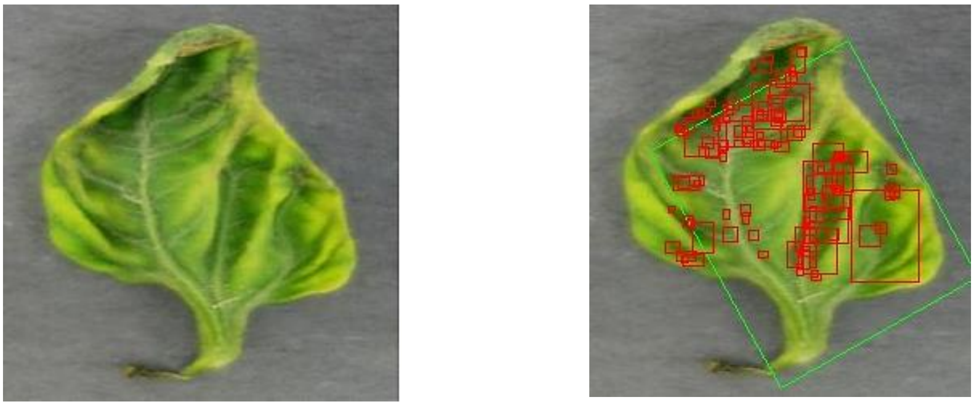
d) Tomato target spot (TTS) detection: In TTS, small dark to large light brown colour lesions appears on leaves and fruits. As seen in Figure 7(a), so the major features extracted were to detec these colours in the leaf portion and then create a bounded box around these spots as shown in Figure 7(b). Algorithm for TTS detection:
   - Step 1: Detect light brown and dark brown colored circular spots in the leaves. For this, the two-colour ranges namely, light brown and dark brown are defined and then both the ranges are combined to detect these colours.
   - Step 2: Find the contours in the mask.
   - Step 3: Check if the contours are found, then check if the brown circular spot is found. If it is found, then a rounded box is created around the spot.



(a) (b)

Figure 7. TTS (a) leaf and (b) spots highlighted after the algorithm executed

e) Tomato Yellow Leaf Curl Virus (TYLCV)*:* The symptoms of this disease are curled leaves with yellow colour as seen in Figure 8(a), so in this case curled leaves with yellow colours will be identified and a bounded box will be created at those spots as seen in Figure 8(b). Algorithm for Yellow Leaf Curl Virus detection:
   - Step 1. Read and pre-process the image.
   - Step 2. Use contour detection to find the contours of the leaf, Border of the leaves, finding the leaf in the image.
   - Step 3. Use image moments to calculate the centroid, area, and orientation of the leaf.
   - Step 4. Use the calculated properties to determine whether the leaf is curled up or not.
   - Step 5. Mark the curled part of the leaf by drawing a rectangle around it.
   - Step 6. Detect the yellow colour.
     6.1 Convert to HSV colour space.
     6.2 Define the lower and upper bounds of the yellow colour.
     6.3 Create a mask of the yellow areas.
     6.4 Find contours in the mask.
     6.5 Check if any contours are found.
       6.5.1 At least one yellow spot was detected.
       6.5.2 Create a bounded box around it.



(a) (b)

Figure 8. TYLCV (a) leaf and (b) spots highlighted after the algorithm executed

3. RESULTS AND DISCUSSION

The algorithms designed in the study are based on the colours these diseases have specifically and hence these algorithms can detect and diagnose these diseases appropriately and also able to mark them. These algorithm works in three major steps:
- Step 1: Removal of background: Background of the images is removed by using Otsu's thresholding.
- Step 2: To detect the diseases, identified their colours and then created the masks for those colour and used the moments and contours accordingly.
- Step 3: If the contours are found, create a bounded box around it.

4. CONCLUSION

The features of four tomato disease detection that is TMV, TTS, TYLCV, and TSLS are extracted using image processing methods. Many other techniques like image segmentation, colour analysis, texture analysis was observed. The algorithms in this study extract the features of these diseases mainly through colour analysis. Once the features specific to the disease are diagnosed, these are then highlighted using bounded boxes as shown in the figures. All the image processing techniques have some pros and cons when applied to classify diseases. The result of this experiment shows that features of these diseases are extracted nicely and marked more accurately as compared to other studies. In the future, these extracted features can then be trained using CNN to detect the diseases more accurately.

REFERENCES

[1] A. Kumar and M. Vani, "Image Based Tomato Leaf Disease Detection," in *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Jul. 2019, pp. 1–6, doi: 10.1109/icccnt45670.2019.8944692.

[2] A. Abbas, S. Jain, M. Gour, and S. Vankudothu, "Tomato plant disease detection using transfer learning with C-GAN synthetic images," *Computers and Electronics in Agriculture*, vol. 187, Aug. 2021, doi: 10.1016/j.compag.2021.106279.

[3] B. A. M. Ashqar and S. S. Abu-Naser, "Image-Based Tomato Leaves Diseases Detection Using Deep Learning," *International Journal of Academic Engineering Research*, vol. 2, no. 12, pp. 10–16, 2018, [Online]. Available: www.ijeais.org/ijaer

[4] I. Ahmad, M. Hamid, S. Yousaf, S. T. Shah, and M. O. Ahmad, "Optimizing pretrained convolutional neural networks for tomato leaf disease detection," *Complexity*, vol. 2020, pp. 1–6, Sep. 2020, doi: 10.1155/2020/8812019.

[5] B. Luna-Benoso, J. C. Martinez-Perales, and J. Cortes-Galicia, "Tomato disease detection by means of pattern recognition," *International Journal of Computing and Optimization*, vol. 7, no. 1, pp. 35–45, 2020, doi: 10.12988/ijco.2020.9823.

[6] M. P Arakeri, M. Arun, and Padmini R K, "Analysis of Late Blight Disease in Tomato Leaf Using Image Processing Techniques," *International Journal of Engineering and Manufacturing*, vol. 5, no. 4, pp. 12–22, Nov. 2015, doi: 10.5815/ijem.2015.04.02.

[7] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Information Processing in Agriculture*, vol. 7, no. 4, pp. 566–574, Dec. 2020, doi: 10.1016/j.inpa.2019.11.001.

[8] S. Vetal and Khule R.S., "Tomato Plant Disease Detection using Image Processing," *IJARCCE*, vol. 6, no. 6, pp. 293–297, Jun. 2017, doi: 10.17148/ijarcce.2017.6651.

[9] H. Sabrol and K. Satish, "Tomato plant disease classification in digital images using classification tree," in *2016 International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2016, pp. 1242–1246, doi: 10.1109/iccsp.2016.7754351.

[10] R. Thangaraj, S. Anandamurugan, P. Pandiyan, and V. K. Kaliappan, "Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion," *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 469–488, Aug. 2021, doi: 10.1007/s41348-021-00500-8.

[11] S. S. Kumar and B. K. Raghavendra, "Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, Mar. 2019, pp. 313–316, doi: 10.1109/icaccs.2019.8728325.

[12] Q. Wang, F. Qi, M. Sun, J. Qu, and J. Xue, "Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and Object Detection Techniques," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1–15, Dec. 2019, doi: 10.1155/2019/9142753.

[13] T. T. Mim, M. H. Sheikh, R. A. Shampa, M. S. Reza, and M. S. Islam, "Leaves Diseases Detection of Tomato Using Image Processing," in *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*, Nov. 2019, pp. 244–249, doi: 10.1109/smart46866.2019.9117437.

[14] T. Mehra, V. Kumar, and P. Gupta, "Maturity and disease detection in tomato using computer vision," in *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2016, pp. 399–403, doi: 10.1109/pdgc.2016.7913228.

[15] M. A. Jasim and J. M. AL-Tuwaijari, "Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques," in *2020 International Conference on Computer Science and Software Engineering (CSASE)*, Apr. 2020, pp. 259–265, doi: 10.1109/csase48920.2020.9142097.

[16] U. Mokhtar *et al.*, "SVM-Based Detection of Tomato Leaves Diseases," in *Intelligent Systems'2014*, Switzerland: Springer International Publishing, 2015, pp. 641–652, doi: 10.1007/978-3-319-11310-4_55.

[17] S. Ashok, G. Kishore, V. Rajesh, S. Suchitra, S. G. G. Sophia, and B. Pavithra, "Tomato Leaf Disease Detection Using Deep Learning Techniques," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, Jun. 2020, pp. 979–983, doi: 10.1109/icces48766.2020.9137986.

[18] M. A. Ihsan Aquil and W. H. Wan Ishak, "Evaluation of scratch and pre-trained convolutional neural networks for the classification of Tomato plant diseases," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 467–475, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp467-475.

[19] S. Bensaadi and A. Louchene, "Low-cost convolutional neural network for tomato plant diseases classifiation," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 1, pp. 162–170, Mar. 2023, doi: 10.11591/ijai.v12.i1.pp162-170.

[20] S. N. Estiri, A. H. Jalilvand, S. Naderi, M. H. Najafi, and M. Fazeli, "A Low-Cost Stochastic Computing-based Fuzzy Filtering for Image Noise Reduction," in *2022 IEEE 13th International Green and Sustainable Computing Conference (IGSC)*, Oct. 2022, pp. 1–6, doi: 10.1109/igsc55832.2022.9969358.

[21] A. Baral, M. Sadegh Nasr, P. Darghiasi, B. Abediniangerabi, and M. Shahandashti, "Detection and Classification of Vegetation for Roadside Vegetation Inspection and Rehabilitation Using Deep Learning Techniques," *International Conference on Transportation and Development 2022*, Aug. 2022, doi: 10.1061/9780784484319.014.

[22] K. Nazari, M. J. Ebadi, and K. Berahmand, "Diagnosis of Alternaria disease and leafminer pest on tomato leaves using image processing techniques," *Journal of the Science of Food and Agriculture*, vol. 102, no. 15, pp. 6907–6920, Jun. 2022, doi: 10.1002/jsfa.12052.

[23] N. Niknejad, R. Bidese-Puhl, Y. Bao, K. G. Payn, and J. Zheng, "Phenotyping of architecture traits of loblolly pine trees using stereo machine vision and deep learning: Stem diameter, branch angle, and branch diameter," *Computers and Electronics in Agriculture*, vol. 211, Aug. 2023, doi: 10.1016/j.compag.2023.107999.

[24] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," *Environmental Monitoring and Assessment*, vol. 195, no. 1, Nov. 2022, doi: 10.1007/s10661-022-10656-x.

[25] S. M. Javidan, A. Banakar, K. A. Vakilian, and Y. Ampatzidis, "Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning," *Smart Agricultural Technology*, vol. 3, Feb. 2023, doi: 10.1016/j.atech.2022.100081.

[26] A. Dhiman and V. Saroha, "Detection of Severity of Disease in Paddy Leaf by Integrating Edge Detection to CNN-Based Model," in *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2022, pp. 470–475, doi: 10.23919/indiacom54597.2022.9763128.

[27] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, "A novel deep learning method for detection and classification of plant diseases," *Complex & Intelligent Systems*, vol. 8, no. 1, pp. 507–524, Sep. 2021, doi: 10.1007/s40747-021-00536-1.

[28] B. K. Mohammed Ali Al-windi, A. H. Abbas, and M. Shakir Mahmood, "Wheat stem rust leaf disease detection using image processing," *Materials Today: Proceedings*, vol. 80, pp. 3346–3350, 2023, doi: 10.1016/j.matpr.2021.07.251.

[29] A. V Panchal, S. C. Patel, K. Bagyalakshmi, P. Kumar, I. R. Khan, and M. Soni, "Image-based Plant Diseases Detection using Deep Learning," *Materials Today: Proceedings*, vol. 80, pp. 3500–3506, 2023, doi: 10.1016/j.matpr.2021.07.281.

[30] S. U. Rahman, F. Alam, N. Ahmad, and S. Arshad, "Image processing based system for the detection, identification and treatment of tomato leaf diseases.," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9431–9445, Sep. 2022, doi: 10.1007/s11042-022-13715-0.
[31] S. Sagar and J. Singh, "An experimental study of tomato viral leaf diseases detection using machine learning classification techniques," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 451–461, Feb. 2023, doi: 10.11591/eei.v12i1.4385.

BIOGRAPHIES OF AUTHORS

Sanjeela Sagar Assistant Professor, Vidyalankar School of Information Technology, Wadala East, Mumbai, India. She has 23 years of academic experience and is a Research Scholar of Guru Jambheshwar University of Science and Technology Hisar, Haryana. She completed her MCA in 2005. She has done few publications in National and International Conferences and authored 23 books with Vipul Publications, Mumbai. Her research area includes machine learning, data science, and data mining. She can be contacted at email: sanjeela.sagar@vsit.edu.in.

Jaswinder Singh is working as an Associate Professor in the Department of Computer Science and Engineering at Guru Jambheshwar University of Science and Technology, Hisar, Haryana. He has teaching experience of more than 20 years and he has published more than 25 research papers in international journals and conferences. He has completed his Ph.D. in Computer Science and Engineering from Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana and completed his M. Tech in Computer Science and Engineering from Kurukshetra University, Kurukshetra, Haryana. His areas of research include opinion mining, web information retrieval, search engine optimization, web mining, information system, and social network analysis. He can be contacted at email: jaswinder_singh@gjust.org.