Investigation of early symptoms of tomato leaf disorder by using analysing image and deep learning models

Surendra Reddy Vinta[1, *] Ashok Kumar Koshariya[2], Sampath Kumar S[3], Aditya[4], Annantharao Gottimukkala[5]

[1] School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India
[2] Department of Plant Pathology, School of Agriculture, Lovely Professional University Jalandhar, Punjab India
[3] Department of Computer Science and Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamilnadu, India
[4] National Institute of Food Technology Entrepreneurship and Management, Kundli (NIFTEM-K) 131028, Haryana – MoFPI, Government of India
[5] Department of Computer Science and Engineering, Koneru Lakshmaiah Educationa Foundation, Green Fields, Vaddeswaram-522302, Guntur Dirstrict, Andhra Pradesh, India

Abstract

Despite rapid population growth, agriculture feeds everyone. To feed the people, agriculture must detect plant illnesses early. Predicting crop diseases early is unfortunate. The publication educates farmers about cutting-edge plant leaf disease-reduction strategies. Since tomato is a readily accessible vegetable, machine learning and image processing with an accurate algorithm are used to identify tomato leaf illnesses. This study examines disordered tomato leaf samples. Based on early signs, farmers may quickly identify tomato leaf problem samples. Histogram Equalization improves tomato leaf samples after re sizing them to 256 × 256 pixels. K-means clustering divides data space into Voronoi cells. Contour tracing extracts leaf sample boundaries. Discrete Wavelet Transform, Principal Component Analysis, and Grey Level Co-occurrence Matrix retrieve leaf sample information.

*Corresponding author. Email: vsurendra.cse@gmail.com

1. Introduction

The domestication of primary food crops and animals today can be attributed to the advancement of agriculture several millennia ago. Food insecurity is a significant global challenge confronting humanity, with plant diseases representing a primary contributing factor. As per a scholarly estimate, plant diseases are responsible for causing a reduction in crop yield of approximately 16% on a global scale [3]. According to research, the estimated global loss potential for wheat due to pests is about 50%, while for soybean, it is estimated to be between 26-29%. This information is supported by reference [3]. Plant pathogens may be broken down into many distinct categories, such as fungi, fungal-like organisms, bacteria, viruses, viroid, virus-like organisms, nematodes, protozoa, algae, and parasitic plants. Applications in the fields of renewable energy power prediction [4, 5], biomedicine [6, 7], and computer vision [8,

9] have all benefited greatly from the use of AI, ML, and CV. As a consequence of the COVID-19 epidemic, there has been an increase in the use of AI for the diagnosis of lung-related diseases [8, 9, 10, 11] and other predictive applications [12]. The implementation of comparable cutting-edge technology may alleviate plant ailments' unfavourable impacts through prompt identification and assessment during their initial phases. Currently, a significant amount of research is being conducted on the utilisation of artificial intelligence and computer vision for the automated detection and diagnosis of plant diseases. This is because manual plant disease monitoring is a laborious, time-consuming, and arduous task.

Plants play a crucial role in human existence as they are responsible for the production of sustenance and protection against hazardous radiation. The presence of plants is essential for maintaining all terrestrial life and the safety of the ozone layer, which serves as a filter for ultraviolet radiation. The tomato plant is a nutrient-dense crop commonly grown and consumed as a vegetable [1]. Globally, an estimated 160 million tonnes of tomatoes are consumed annually [2]. The tomato is

regarded as a source of income for farm households and is recognized as a significant contributor to poverty reduction [3]. The cultivation and production of tomatoes notably impact the agricultural economy due to their high nutrient density, making them one of the most valuable crops globally. The tomato is considered a nutrient-dense food item, and it has been found to exhibit pharmacological properties that can safeguard against various ailments, including hypertension, hepatitis, and gingival bleeding [1]. The growing utilization of tomatoes has led to a surge in demand for this commodity. As per statistical data, it has been observed that small-scale farmers are responsible for generating over 80% of the agricultural yield [2]. However, they face a significant challenge in crop loss, which amounts to approximately 50% due to the impact of diseases and pests. Identifying field crop diseases and parasitic insects is crucial in determining their effects on tomato growth. Therefore, research on field crop disease diagnosis is necessary.

2. Proposed Methodology

This section presents a model for detecting leaf disease developed using the IP and ML approaches. The proposed model for detecting leaf diseases and utilizes a combination of computer vision and machine learning techniques, specifically DWT, PCA, GLCM, and CNN.

Tomatoes with six different diseases are used to see how well they can detect and categorize leaf diseases. To ensure that the tomato samples are the same size throughout the experiment, they are scaled to 256 256 pixels as part of the picture processing. HE and K-mean are clustering to improve the quality and divide the leaf samples. The presence of illness in a leaf may be anticipated early in the process using K-means clustering. Using contour tracing, we can identify the limits of the leaf samples.

(a)Healthy tomato leaf   (b)Late bright tomato leaf

(c) Spector spot- tomato leaf   (D) mosaic- tomato leaf

(e)Yellow curved tomato leaf (f) bacterial spot tomato leaf

Figure 1: tomato leaf with different diseases

Samples' functional regions and characteristics are extracted using the DWT, PCA, and GLCM. The second step, also part of machine learning methodologies, involves classifying the features and keeping track of the model's performance using SVM, KNN, and CNN.

2.1. Development of Datasets

Tomato plants in the village database that many illnesses have hit are considered. To conduct the tests for leaf disease identification we capture photos of a tomato leaf affected by six different diseases. Figure 1 displays several representative database leaf pictures.Images of leaves clustered by using the K-means algorithm. K-means clustering determines how far each cluster's centre is from its neighbours, applied to each image. Figure 2 displays leaf samples clustered using the k-mean technique.

The process of contour tracing and utilized on digital leaf specimens to extract their overall shape characteristics. Upon completion of the contour extraction process, its distinct features are meticulously examined and subsequently used for pattern classification. Frequently aids the assessment of feature extraction process efficiency. Figure 3 displays the images that were generated after executing contour tracing.
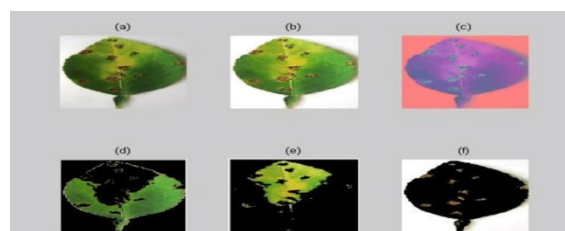

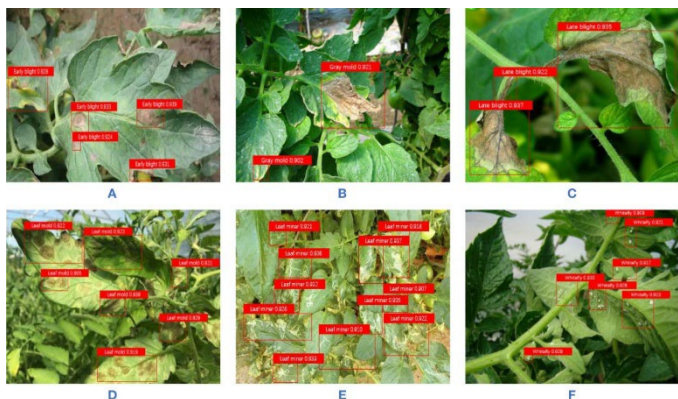**Figure 2:** K mean cluster- Tomato leaf

Figure 3: Tracing of the disease on tomato leaf

2.2. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) technique extracts significant features from improved tomato specimens. The Discrete Wavelet Transform (DWT) partitions the signal into sub-bands that consist of components with lower frequencies (LL, LH, HL) and higher frequencies (HH). According to Figure 4, the low-low part of the discrete wavelet transform (DWT) exhibits a greater degree of information availability in comparison to its higher frequency counterparts.

Principal Component Analysis extracts the most valuable features from a wavelet decomposition of a grey-level co-occurrence matrix. The GLCM uses higher-order grey values determined by nearby criteria in its distribution. The GLCM method is used to extract characteristics from leaves, and this is where the various attributes come from.

These are the most common texture-based features utilized today. A sum of squares, Mean, Standard Deviation, Autocorrelation, Dissimilarity, Entropy, and Homogeneity. A feature vector composed of the parts collected using DWT, GLCM, and PCA is concatenated to recognize and categories pictures.
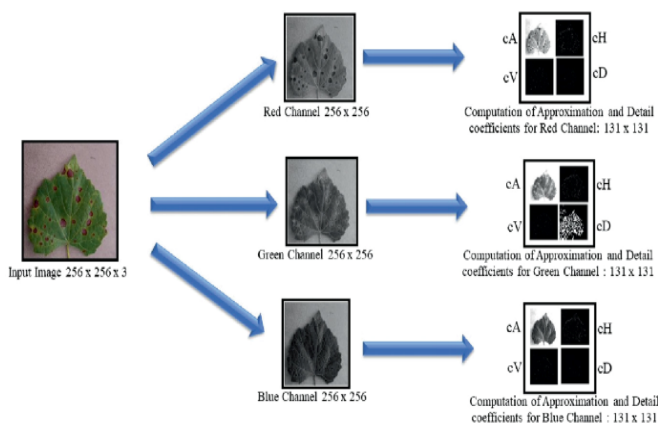

Figure 4: Discrete Wavelet Transform- Tomato leaf

2.3. Machine learning Models

Samples are classified using several methods including SVM, KNN, and CNN. The Convolutional Neural Network is an example of an ANN created for data analysis. Input (IL),

output (OL), and hidden (HL) layers make up CNN's complex architecture. Convolutional layers, a RELU layer that executes the activation function, pooling layers, a fully connected layer, and a normalization layer are all part of the HL. Its design is cross correlation rather than convolution, as shown by the fact that the matrix indices have mathematical importance. In Fig.5, we see a typical -3-layer network.
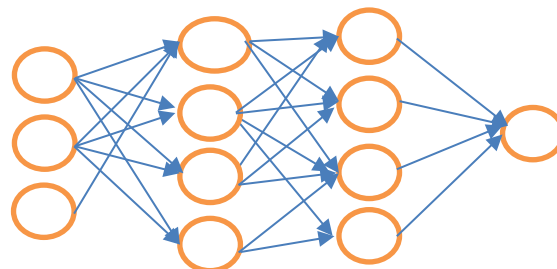

Figure 5: ANN Model – Input, hidden and output layer

Fig. 6 shows the CNN confusion matrix for output and target classes. Fig. 7 shows CNN-classified leaf feature training samples and their accuracy and mistakes.

Figure 6: CNN- Performance

5 13.3%	0 0%	0 0%	2 2.8%	0 0%	72.4% 27.6%	
1 3.9%	7 21.1%	0 0%	0 0%	0 0%	88.1% 11.1%	
0 0%	0 0%	7 21.1%	0 0%	0 0%	99.9% 0.1%	
1 2.8%	0 0%	0 0%	0 0%	7 21.0%	99.9% 0.1%	4
72.4 27.6	100% 0%	100% 0%	100% 0%	72.4% 27.6	88.3% 11.7	

Figure 7: Confusion Matrix

2.4. Performance evaluation

IP-enabled Parameters for the proposed model are computed and presented in, including Precession, Recall, and the F-measure.

$$Precision\ in\ \% = \frac{TP}{TP+FP} * 100$$

$$Recall\ in\ \% = \frac{TP}{TP+FP} X 100$$

$$F - Measure = \frac{2*precesion*Recall}{Precesion*Recall}$$

3. Result

The suggested model will be evaluated using data collected from tomato leaf samples in a village dataset. One hundred leaf specimens in good health are used to put the proposed model through its paces. Using this model, we could correctly classify 99 samples at the 99% level. The model correctly identifies 100 out of 100 tomato samples infected with the Mosaic virus.

The model achieves perfect accuracy in the field of leaf moulds. The programme accurately predicted 99 out of 100 instances of yellow curl. Spotted spider mite and Target Spot provide identical success rates of 99% and 100%, respectively. The suggested model is evaluated by testing 685 samples from the tomato village dataset, with an impressive accuracy of 99.1%.

Training and testing the model on subsamples of the dataset ensures its accuracy. The following are the recommended system requirements for running the programme. In other words, Windows 10 operating system, GPU-NVIDIA core, and Python programming language. OpenCV, Keras, tensor flow, NumPy, Matplot, MatPlot, and the image data generator,Tomato (Six Disorder) samples from the Plant Village collection. Precision, Recall, and F1 score measure the effectiveness of the suggested model.
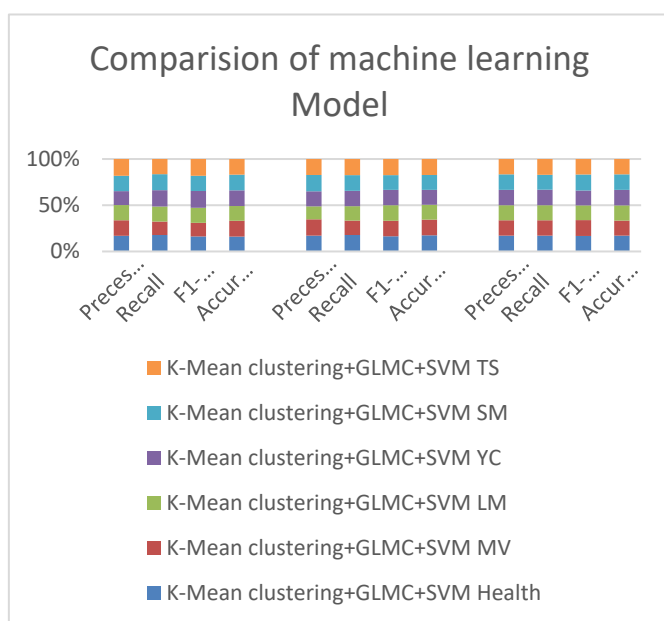

Figure 8: Comparison of proposed Deep learning

model

The suggested model is evaluated using a 600-sample dataset of tomato leaf diseases. Fig. 8 contrasts the new model with other commonly used ones. Table 1 compares SVM, KNN, and CNN based on their precession, recall, and F1 scores.

Table 1: Comparison of different Deep learning model combinations

		Precession %	Recall%	F1-score %	Accuracy%
K-Mean clustering+GLMC+SVM	Health	0.921	0.89	0.89	82%
	MV	0.92	0.73	0.81	87.20%
	LM	0.9	0.82	0.89	81.10%
	YC	0.82	0.89	0.99	86.30%
	SM	0.91	0.87	0.91	86.20%
	TS	0.99	0.82	0.99	86.10%
	Overall	0.91	0.84	0.91	0.85%
LBP+PCA+K MEAN CLUSTER	Health	0.86	0.98	0.91	96.1%
	MV	0.89	0.86	0.92	93.7%
	LM	0.71	0.85	0.93	88.5%
	YC	0.82	0.92	0.92	88.2%
	SM	0.89	0.93	0.88	89.1%
	TS	0.87	0.96	0.96	95.1%
	Overall	0.84	0.92	0.92	91.78333%
DWT+PCA+GLCM+CNN	Health	0.99	0.99	0.96	99.1%
	MV	0.95	0.95	0.98	95.2%
	LM	0.94	0.94	0.91	95.6%
	YC	0.96	0.97	0.92	97.2%
	SM	0.96	0.92	0.99	98.3%
	TS	0.96	0.99	0.96	96.2%
	Overall	0.96	0.96	0.95	96.93%

The outcomes of the proposed model are contrasted with those of pre-existing models. The study indicates that the proposed model, which combines DWT, PCA, GLCM, and CNN, exhibits a higher level of accuracy at 99.09% in comparison to other pre-existing models.

3. Conclusion:

In the suggested model, computer vision methods are used for preprocessing, including converting from RGB to greyscale,

Histogram Equalisation (HE), K-means clustering, and contour tracing. Feature extraction methods that are renowned for their capacity to extract useful features are applied to the leaf samples. These methods include the Discrete Wavelet Transform, Principal Component Analysis, and the Generalised Linear Model. To distinguish between healthy and diseased leaves, researchers use machine learning methods including Support Vector Machines (SVM), K-Nearest Neighbours (K-NN), and Convolutional Neural Networks (CNN). The proposed model has been shown to be suitable for the CNN machine learning classification methodology, with its recommended degree of accuracy when compared to other modern methods. Using fusion approaches to extract significant characteristics and testing the model on more leaf sample datasets might improve future research.

References

[1] M. E.H. Chowdhury et al., 'Tomato Leaf Diseases Detection Using Deep Learning Technique', Technology in Agriculture. IntechOpen, Oct. 13, 2021. doi: 10.5772/intechopen.97319.

[2] Bhandari, M.; Shahi, T.B.; Neupane, A.; Walsh, K.B. BotanicX-AI: Identification of Tomato Leaf Diseases Using an Explanation-Driven Deep-Learning Model. J. Imaging 2023, 9, 53. https://doi.org/10.3390/jimaging9020053

[3] Trivedi, Naresh & Gautam, Vinay & Anand, Abhineet & Aljahdali, Hani & Gracia Villar, Santos & Anand, Divya & Goyal, Nitin & Kadry, Seifedine. (2021). Early Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network. Sensors. 21. 7987. 10.3390/s21237987.

[4] Brahimi, Mohammed & Kamel, Boukhalfa & Moussaoui, Abdelouahab. (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. Applied Artificial Intelligence. 31. 1-17. 10.1080/08839514.2017.1315516.

[5] Mohanty Sharada P., Hughes David P., Salathé Marcel, Using Deep Learning for Image-Based Plant Disease Detection, Frontiers in Plant Science, VOLUME 7, 2016, https://www.frontiersin.org/articles/10.3389/fpls.2016.01419, 10.3389/fpls.2016.01419

[6] Qimei Wang, Feng Qi, Minghe Sun, Jianhua Qu, Jie Xue, "Corrigendum to "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. 2021, Article ID 3751479, 1 pages, 2021. https://doi.org/10.1155/2021/3751479

[7] S. Khatoon, M. Maruf Hasan, A. Asif, M. Alshmari and Y. Yap, "Image-based automatic diagnostic system for tomato plants using deep learning," Computers, Materials & Continua, vol. 67, no.1, pp. 595–612, 2021.

[8] Jasim, Y. A. (2021). High-Performance Deep learning to Detection and Tracking Tomato Plant Leaf Predict Disease and Expert Systems. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 10(2). https://doi.org/10.14201/ADCAIJ202110297122

[9] Liu, J., Wang, X. Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model. Plant Methods 16, 83 (2020). https://doi.org/10.1186/s13007-020-00624-2

[10] Patnayakuni, S. P. . (2022). Tomato: Different Leaf Disease Detection Using Transfer Learning Based Network. Journal of Mobile Multimedia, 18(03), 743–756. https://doi.org/10.13052/jmm1550-4646.18313

[11] Bayram, H.Y., Bingol, H., Alatas, B. (2022). Hybrid deep model for automated detection of tomato leaf diseases. Traitement du Signal, Vol. 39, No. 5, pp. 1781-1787. https://doi.org/10.18280/ts.390537

[12] Tian, K., Zeng, J., Song, T., Li, Z., Evans, A. and Li, J. (2022) "Tomato leaf diseases recognition based on deep convolutional neural networks", Journal of Agricultural Engineering, 54(1). doi: 10.4081/jae.2022.1432.

[13] Singh, Ganesh Bahadur, et al. "Identification of Tomato Leaf Diseases Using Deep Convolutional Neural Networks." IJAEIS vol.12, no.4 2021: pp.1-22. http://doi.org/10.4018/IJAEIS.20211001.oa3

[14] Nawaz, M., Nazir, T., Javed, A. et al. A robust deep learning approach for tomato plant leaf disease localization and classification. Sci Rep 12, 18568 (2022). https://doi.org/10.1038/s41598-022-21498-5

[15] Keke Zhang, Qiufeng Wu, Anwang Liu, Xiangyan Meng, "Can Deep Learning Identify Tomato Leaf Disease?", Advances in Multimedia, vol. 2018, Article ID 6710865, 10 pages, 2018. https://doi.org/10.1155/2018/6710865

[16] Gayathri, S., et al. "Image analysis and detection of tea leaf disease using deep learning." 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, 2020.

[17] Sladojevic, Srdjan, et al. "Deep neural networks based recognition of plant diseases by leaf image classification." Computational intelligence and neuroscience 2016 (2016).

[18] Amara, Jihen, Bassem Bouaziz, and Alsayed Algergawy. "A deep learning-based approach for banana leaf diseases classification." Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband (2017).

[19] Amara, Jihen, Bassem Bouaziz, and Alsayed Algergawy. "A deep learning-based approach for banana leaf diseases classification." Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband (2017).

[20] Tiwari, Divyansh, et al. "Potato leaf diseases detection using deep learning." 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020.

[21] Panigrahi, Kshyanaprava Panda, et al. "Maize leaf disease detection and classification using machine learning algorithms." Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2019. Springer Singapore, 2020.

[22] Sujatha, Radhakrishnan, et al. "Performance of deep learning vs machine learning in plant leaf disease detection." Microprocessors and Microsystems 80 (2021): 103615.

[23] Atila, Ümit, et al. "Plant leaf disease classification using EfficientNet deep learning model." Ecological Informatics 61 (2021): 101182.

[24] Ramesh, Shima, et al. "Plant disease detection using machine learning." 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C). IEEE, 2018.