Research Article

Application of Image Processing and Transfer Learning for the Detection and Classification of Blemishes and Canker Diseases of Citrus

Aqsa Batool[1], Hina Safdar[1]* and Furqan-ur-Rehman[2]

[1]Department of Plant Pathology, College of Agriculture, University of Sargodha, Sargodha, Pakistan; [2]College of Agriculture, University of Sargodha, Sargodha, Pakistan.

Abstract | Citrus blemishes and canker, have significant threats to agriculture worldwide, leading to substantial economic losses. Conventional methods for detection of diseases like visual inspection is time-consuming and have a chance of errors. Therefore, this study is planned to utilize the google teachable machine to develop an automated system for identifying and classifying citrus diseases using an image processing based method i.e. data augmentation and transfer learning approach. A properly curated dataset of citrus fruits infected with blemishes and cankers was compiled from databases of agriculture and personal photographs captured using smartphones. Preprocessing of the dataset included resizing, background reduction and brightness normalization. Data augmentation such as rotation. Zooming and flipping were applied to enhance model robustness. The analysis based on confusion matrix and performance metrics, it was observed that the model achieved 100% accuracy, precision, F1-score and recall with no misclassification errors. Furthermore, the model demonstrated fast convergence and stability during training, reaching 50 epochs with low loss and without overfitting. These results represent the potential of artificial intelligence for accurate disease detection. However, one of its limitations is that it is sensitive to variations in the real-world and constraints intrinsic to using the GTA platform. This study offering cost effective and scalable solutions with the use of artificial intelligence into agriculture for better disease management.

Received | September 03, 2025; **Accepted** | December 8, 2025; **Published** | April 03, 2026
***Correspondence** | Hina Safdar, Department of Plant Pathology, College of Agriculture, University of Sargodha; **Email:** hina.safdar@uos.edu.pk
Citation | Batool, A., H. Safdar and F.U. Rehman. 2026. Application of Image processing and transfer learning for the detection and classification of blemishes and canker diseases of citrus. *Sarhad Journal of Agriculture,* 42(2): 601-608.
DOI | https://dx.doi.org/10.17582/journal.sja/2026/42.2.601.608
Keywords | Deep learning, Machine learning, Google teachable machine, Transfer learning, Image processing

Introduction

Citrus is a tropical and subtropical fruit crop, cultivated in more than 140 countries with 158.49 million metric tons annual output (FAOSTAT, 2024). Citrus hold an important role in agriculture and are generally consumed by everyone (Zhang *et al*., 2021). In Punjab (Pakistan) total fruit production is 30 %. Lemon, sweet orange, grapefruit and mandarin face various diseases like black spot and scab. These diseases not only lower fruit quality but also its yield. Traditional methods for diagnosis of disease is slow, less accurate and expensive but also depends upon expert intervention (Butt *et al*., 2025). The automated systems provide cost effective, consistent and efficient methods to mitigate the challenges of manual diagnosis. The suggested framework is compared to modern approaches, certifying its performance (Butt *et al*., 2024). The industry of agriculture has designate many deep learning (DP) models to tackle different challenges like fruit detection, plant leaf categorization, leaf and fruit disease detection and insect identification (Mamat *et al*., 2022). The approaches associated with DP, computer vision, and machine learning (ML) provide efficient solutions for classification of diseases and their detection. These digital technologies are integrated with precision agriculture for better management and diagnosis of disease through IoT based devising, drones and remote sensing which collect huge amount of information about plant health (Ayaz *et al*., 2019).

The application of DP, ML and different techniques of image processing has become critical for automated detection and classification of plant disease to minimize the need of manual inspection. These systems require minimum physical work, time and labor (Das *et al*., 2024). Supervised ML create models that categorize new data found on datasets. DL enables convolutional neural networks and hierarchical feature extraction highly effective for image analysis (Anwar *et al*., 2022). The computer techniques have made it uncomplicated to detect deviations in crops in real-time (Wang *et al*., 2017). These methods had significant success in recognizing and diagnosing diseases of plants. However, they have limitation in different task of image processing and image segmentation (Iqbal *et al*., 2018 and Singh *et al*., 2019), pattern recognition (Asghar *et al*., 2019), K-nearest neighbor method (Wang *et al*., 2020), feature extraction (Luaibi *et al*., 2021) and artificial neural network (Ali *et al*., 2021). Transfer learning (TL) is an effective approach for disease classification, uilizing pre-trained models such as ResNet and VGG16, which were developed to recognize large datasets images like ImageNet (Zhu *et al*., 2024). TL allows the application of previously trained models, including ResNet, GTM and VGG16 to categorize and identify citrus diseases as precisely as possible. It decreases the demands of huge data since the trained models already have a massive collection of attributes that are trained. The models do not require a fine tuning, making the training process faster (Hossen *et al*., 2025).

Another important use of such technology is the invention of mobile based technologies to detect and diagnose in real time. Farmers can receive real-time diagnosis by taking images of citrus plants with smartphones by integrating IP and TL algorithms into a mobile application (Nayak *et al*., 2023). A mobile application system has proposed for detection of disease in tomato. In this study scientist proposed a TL approach to fine tune convolutional neural network architectures like ResNet50, CGG19, SequeezeNet1.1 and DenseNet121. In this study highest accuracy achieved in Densenet121 model (99.85%) and this model was designed as an application in mobile phone for detection of disease in tomato. This program requires a high speed internet connection to perform diagnosis in 10 s (Nag *et al*., 2023). Another web based study was performed to diagnose disease with a validity of 93.91% (Asani *et al*., 2023). However, TL and IP for detection of disease in citrus have promising results, but also have several challenges. A great issue is the variability of disease symptoms and environmental factors (Qadri *et al*., 2025). By the creation of AI trained models leads to an increase in the farmers trust who gain a decision making process. The use of mobile based and drone based imaging will allow to track diseases in real time world.

Materials and Methods

Data collection

A dataset comprising healthy and infected images of citrus fruits (canker and blemishes) was compiled from reliable agricultural databases, peer-reviewed publications, and smartphone photographs under natural lighting conditions. This study ensured a dataset with variations in lighting, angle, infection severity and fruit size that support better generalization of model.

Data categorization

The collected images were manually labelled into two categories: citrus blemishes and citrus canker (Figure 1). Accurate labelling was necessary for reducing label noise and ensuring the model could learn visual differences between disease types.

Image processing

The preprocessing steps to enhance model learning and standardize input were:

Background reduction: The background elements were removed to focus on the region.
Resizing: The images were resized to meet the input requirements
Brightness normalization: The lighting variabilities were minimized to emphasized disease specific features.

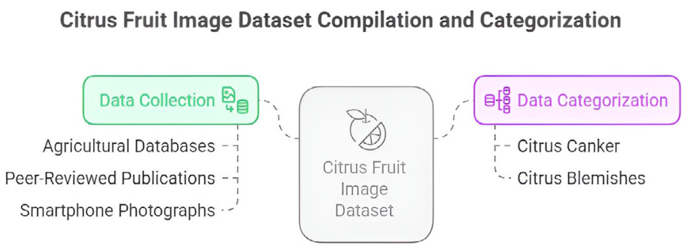

Figure 1: *Data collection and data categorization*

Data augmentation
The augmentation techniques (Figure 2) were used to enlarge the dataset size improve robustness were as following:

Rotation
Flipping
Zooming
Colour and contrast adjustments

Implementation of GTA
A web based tool developed by Google (Saini *et al*., 2024) was used for model development and training (Figure 3).

Uploading images
Images were uploaded via local disk or webcam that ensured class completeness and balance.

Model training
The model was trained in the browser and images were labeled (Figure 4).

Model evaluation
The model was assessed using a separate test dataset. The key metrics were:

Accuracy (correct predictions overall)
Precision (correct optimistic predictions out of total predicted positives)
Recall (correct optimistic predictions out of actual positives) (Malahina *et al*., 2024)

Model deployment and user accessibility
The trained model was utilized as a web-based link. The users can classify images of disease images by uploading a file or using a mobile camera. This makes the tool easily accessible to field officers, researchers and farmers, without software installation or technical expertise. It demonstrates a real-world application of AI in agriculture.

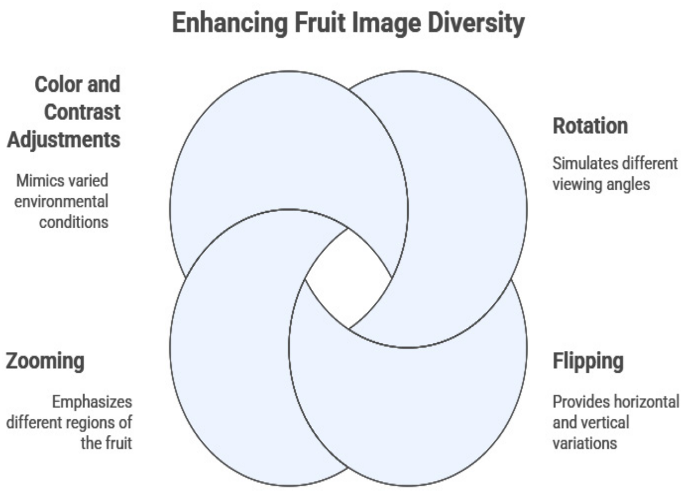

Figure 2: *Data augmentation techniques*

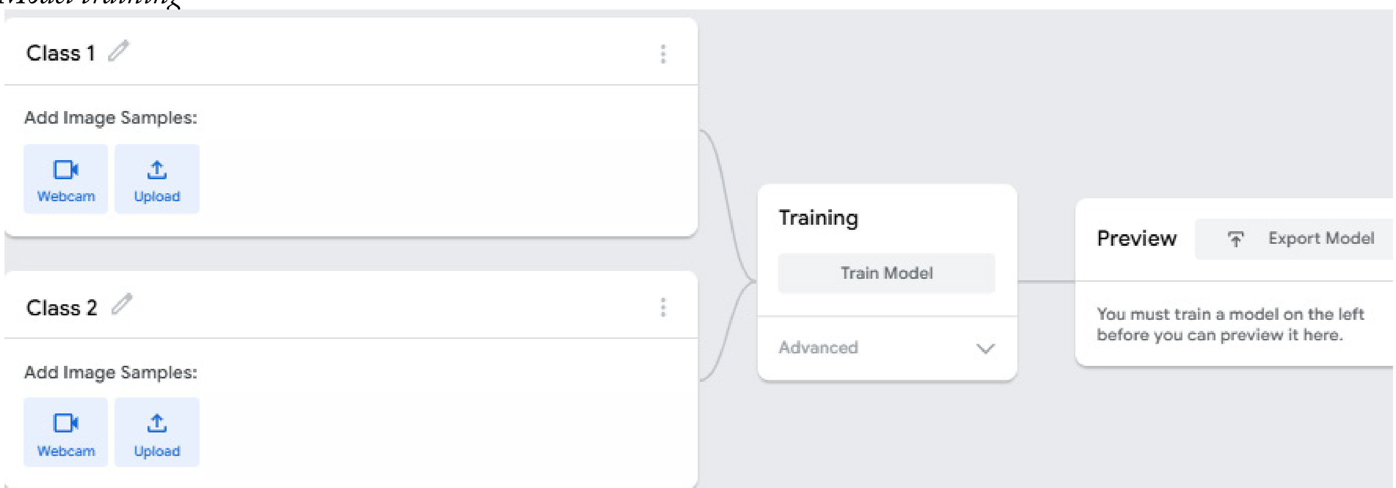

Figure 3: *General view of the Image Project*

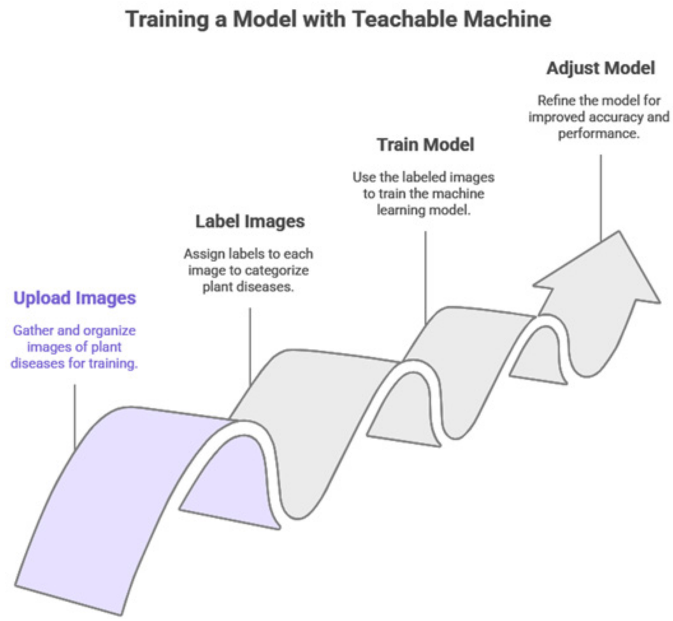

Figure 4: *GTM Training process*

Results and Discussion

Confusion matrix analysis

After training the model successfully, a confusion matrix (Figure 6) was generated to evaluate its performance in classifying two categories: citrus blemishes and citrus canker. Tit is used tool in classification tasks, detailing the differences between predicted and actual labels. It displays the number of instances in each category, helped to assess model misclassification and accuracy.

Table 1: *Values of the confusion matrix*

	Predicted citrus canker	Predicted citrus blemishes
Actual citrus canker	225 (TP)	0 (FN)
Actual citrus blemishes	0 (FP)	225 (TN)

Explanation of terms:

TP (True positive): The model correctly predicted citrus canker when the sample belonged to the citrus canker class.

TN (True Negative): The model correctly predicted citrus blemishes when the sample belonged to the citrus blemishes class.

FP (False Positive): The model incorrectly predicted citrus canker when the actual class was Citrus blemishes.

FN (False Negative): The model incorrectly predicted citrus blemishes when the actual class was citrus canker.

The confusion metrics revealed zero misclassification for FN and FP values. The model achieved perfect accuracy in the test dataset.

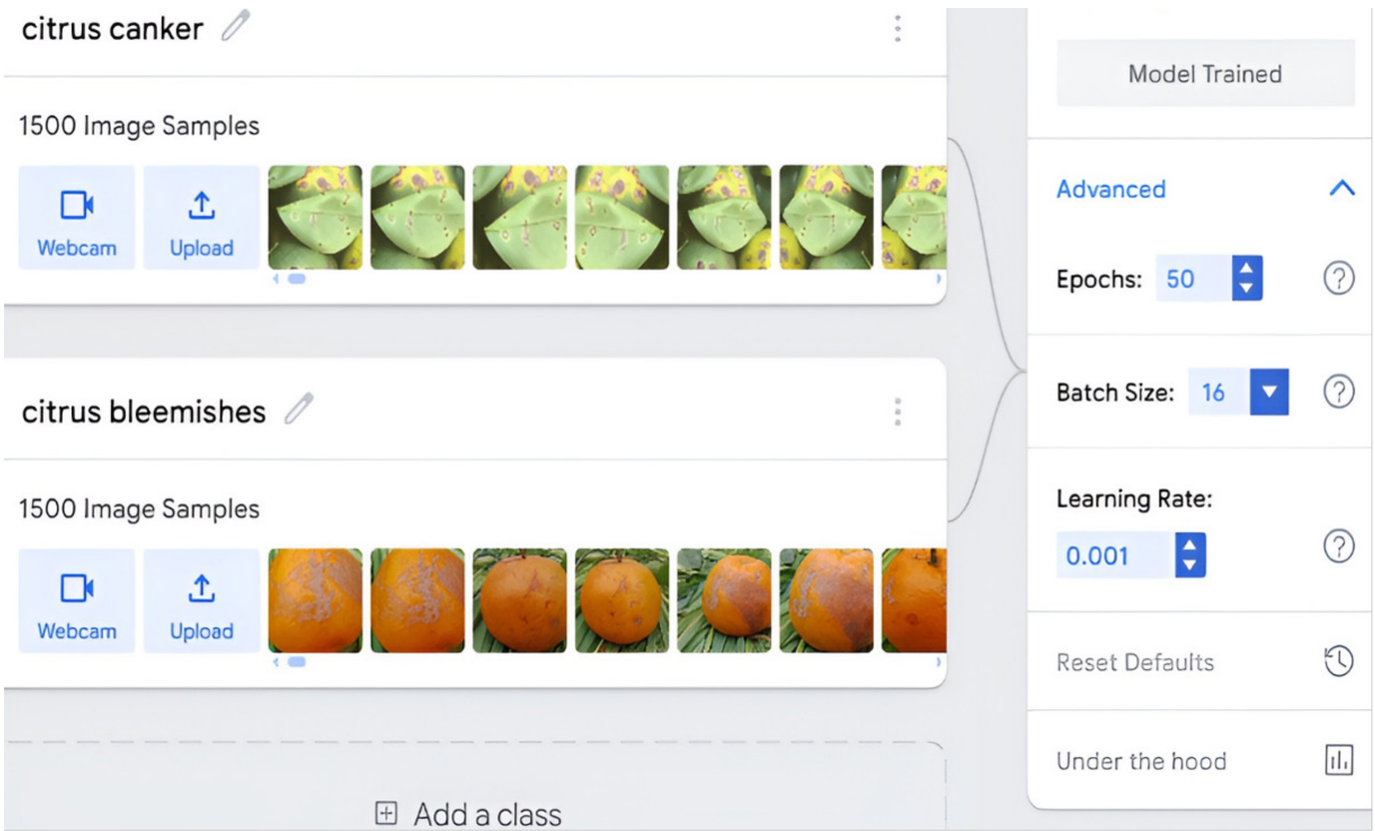

Figure 5: *Trained model*

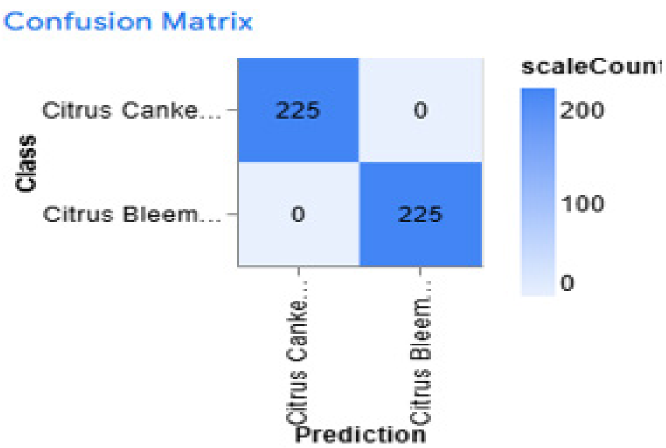

Figure 6: *Confusion matrix*

The performance metrics calculated were:

Accuracy

Definition: Total correct predictions proportion.

Formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Where

TP=225

TN=225

FP=0

FN=0 (Table 1)

Calculations:

$$\frac{225+225}{225+225+0+0} = 1.0 = 100\%$$

Precision (Positive predictive value)

Definition: How many were predicted as Citrus Canker?

Formula

$$\text{Precision} = \frac{TP+FP}{TP}$$

Calculation

$$\frac{225+0}{225} = 1.0 = 100\%$$

Recall (Sensitivity or true positive rate)

Definition: How many were correctly identified?

Formula:

$$\text{Recall} = \frac{TP}{TP+FN}$$

Calculation

$$\frac{225}{225+0} = 1.0 = 100\%$$

F1 Score

Definition: Harmonic mean of recall and prcision.

Formula:

$$\text{F1 Score} = \frac{2 \times precision \times Recall}{precision + Recall}$$

Calculation:

$$\frac{2 \times 1.0 \times 1.0}{1.0+1.0} = 1.0 = 100\%$$

Accuracy per epoch

Figure 7 demonstrated the accuracy trend over 50 training epochs, representing both training accuracy (acc) and testing accuracy (test acc). The model quickly achieved a high accuracy of nearly 1.0 (100%), in the early stage and maintained its performance throughout. The minimum gap between training and test accuracy reflects better generalization on the test data. The consistent closeness of the curves suggests a stable learning process, free from underfitting or overfitting. It indicates that the model effectively learned to classify images and sustained its performance on unseen data.

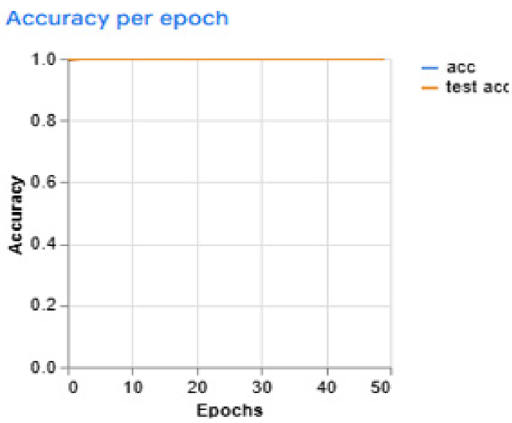

Figure 7: *Accuracy graph*

Loss per epoch

Figure 8 demonstrated the loss values over 50 epochs, plotting both training and testing loss. Both losses were high and drop sharply within few epochs. Then they converged near zero (around 0.001) and remained stable throughout the training process. This decline indicated effective minimization of the error function. The minimum gap between the training and test loss curves suggests the model did not over fit. This convergence confirmed that the model was efficiently learned relevant patterns.

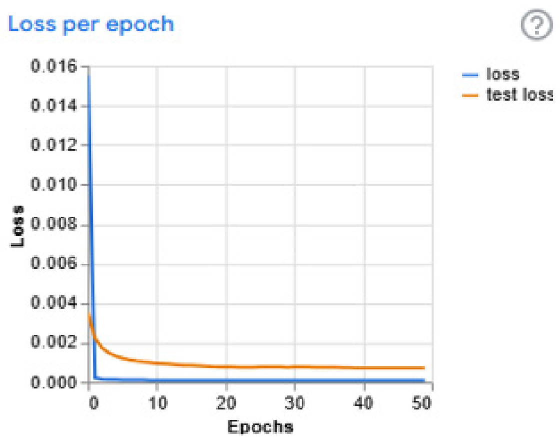

Figure 8: *Loss graph*

Results comparison

Our model has achieved a 100% recall, accuracy and F1-score in classifying citrus canker, with no FN or FP and TN and TP were 225. Similar results were observed by Odabas *et al.* (2024) thataccuracies ranging from 95% to 98% for *Aphis* spp. and *Venturia inaequalis* with minor misclassification in *Monilinia laxa* and *Venturia inaequalis*. Our model achieved perfect results and avoid overfitting. Malahina *et al.* (2024) observed trained class accuracies between 98 and their performance dropped to 35-96% on external dataset and highlighting generalization challenges. Ancheta *et al.* (2023) gained 98.17% accuracy for banana leaf diseases but they observed misclassifications between visually similar diseases, such as the black and yellow Sigatoka. Parate *et al.*, (2023) observed 100% accuracy for leaves of guava, orange, and mango; however, strict metrics such as recall and precision were not reported, as shown in Figure 9. The evaluation was performed on a relatively small dataset of 300 images. In contrast of these studies, our observation not only achieved perfect accuracy but also integrate robust validation metrics. It included the F1 score, ensuring that the model avoids over fit and showed reliable generalization within tested dataset.

Previous studies showed the importance of accurate disease detetion methods. Kala *et al.* (2023) highlighted the crucial role of advanced technology in decreasing the effects of plant diseases on yields. Similarly, another study (Storey *et al.*, 2022) highlight the significance of ML techniques for precise diagnosis of disease. These findings correlate with the outcomes of the present study, further validating the efficacy of utilizing a teachable machine for detection of disease.

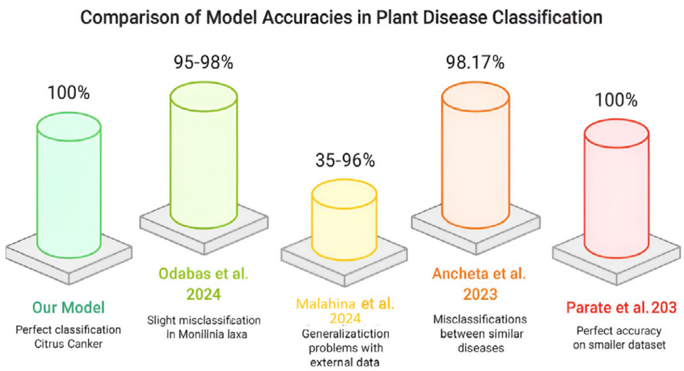

Figure 9: *Result comparison*

Limitations and drawbacks

Limited dataset size & complexity

- GTM performs optimally with small datasets (a few hundred images per class), whereas plant disease detection ideally requires thousands of high-quality images.
- With limited data, there is a risks of overfitting, where the model memorizes images rather than learning generalizable features, potentially reducing real-world accuracy.

Image quality & background sensitivity

- Complex backgrounds (e.g., soil, shadows) may mislead the model if not properly managed during training.
- Model performance can decline when training and test images differ in lighting, angles, or resolution.

Conclusions and Recommendations

The study reflects the efficacy of GTM for the detection and classification of citrus diseases, specifically blemishes and canker. Techniques such as image processing, data augmentation, and TL, the model achieved perfect performance metrics (100%). This indicates that the model has strong capability to distinguish between the two diseases with zero misclassifications and showed its potential as reliable alternatives to visual inspection methods. The model performance was affected by real-world variability in the training data. The dependency on GTM imposes constraints on size of dataset and model complexity, which could impact scalability for broader applications in agriculture.

Acknowledgments

The authors would like to thank the Department of Plant Pathology, College of Agriculture, University of Sargodha, for providing facilities and support necessary to conduct this research.

Novelty Statement

The present research introduced the utilization of GTM with IP and TL for the classification and detection of citrus blemishes and canker. The model scored 100% accuracy, precision, recall and F1-score without over fitting. These findings indicate that AI-based tools using smartphone can provide a low cost and straightforward solution for citrus disease detection and precision farming.

Author's Contribution

Aqsa Batool: Trained the model, analyzed results and wrote first draft.
Hina Safdar: Supervised and guided the methodology and reviewed the manuscript.
Furqan-ur-Rehman: Provided technical support in data handling, validation and editing.
All authors read and approved the final version.

Generative AI or AI assisted technology statement
The authors declare that no generative AI or AI-assisted technologies were used in the writing, analysis, or preparation of this manuscript.

Conflict of interest
The authors have no conflict of interest.

References

Ali, M., M.Z. Asghar and A. Baloch. 2021. An efficient approach for sub-image separation from large-scale multi-panel images using dynamic programming. Multimedia Tool Appl., 80(4): 5449-5471. https://doi.org/10.1007/s11042-020-09950-y

Ancheta, J.B., L.Q. Santos, L.C.M. De Jesus, J.V. David, A.M. Avelino and M.E.A. Grande. 2023. Development of a Classification Model for Banana Leaf Disease Using Google Teachable Machine. Inter. Conf. Arti. Intell. Inform. Commun., (ICAIIC): 0258-0263. https://doi.org/10.1109/ICAIIC64266.2025.10920850

Anwar, T. and H. Anwar. 2022. Citrus plant disease identification using deep learning with multiple transfer learning approaches. Pak. J. Eng. Technol., 3(2): 34-38. https://doi.org/10.51846/vol3iss2pp34-38

Asani, E.O., Y.P. Osadeyi, A.A. Adegun, S. Viriri, J.A. Ayoola and E.A. Kolawole. 2023. mPD-APP: A mobile-enabled plant diseases diagnosis application using convolutional neural network toward the attainment of a food secure world. Front. Art. Intell., 6. https://doi.org/10.3389/frai.2023.1227950

Asghar, M.Z., F. Subhan, M. Imran, F.M. Kundi, S. Shamshirband, A. Mosavi, P. Csiba and A.R. Varkonyi-Koczy. 2019. Performance evaluation of supervised machine learning techniques for efficient detection of emotions from online content. arXiv., 1908.01587. https://doi.org/10.20944/preprints201908.0019.v1

Ayaz, M., M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.H.M. Aggoune. 2019. Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. *IEEE Access.*, 7: 129551-129583. https://doi.org/10.1109/ACCESS.2019.2932609

Butt, N., M.M. Iqbal, I. Ahmad, H. Akbar and U. Khadam. 2024. Citrus Diseases Detection using Deep Learning. J. Comp. Biomed. Inform., 6(02): 23-33.

Butt, N., M.M. Iqbal, S. Ramzan, A. Raza, L. Abualigah, N.L. Fitriyani and M. Syafrudin. 2025. Citrus diseases detection using innovative deeplearning approach and Hybrid Meta Heuristic 2025 PLOS ONE., 20(1): e0316081 https://doi.org/10.1371/journal.pone.0316081

Das, P.K., S.S. Rupa, S. Pumrin, U.C. Das and M.K. Hossen. 2024. Deep Learning for Plant Disease Detection and Classification: A Systematic Analysis and Review. C. Appl. Sci Tech., 24(4): e0259016. https://doi.org/10.55003/cast.2024.259016

FAOSTAT. 2024. Food and agriculture data. Crop Statistics. Available Online: Http://Www. Fao. Org/Faostat.

Hossen, M.I., M. Awrangjeb, M.S. Pan and A.A. Mamun. 2025. Transfer learning in agriculture: A review. Artif. Intell. Rev., 58(4): 97. https://doi.org/10.1007/s10462-024-11081-x

Iqbal, Z., M.A. Khan, M. Sharif, J.H. Shah, M.H. Ur Rehman and K. Javed. 2018. An automa ted detection and classification of citrus plant diseases using image processing techniques: A review. *Comput. Electron. Agric.,* 153:12-32. https://doi.org/10.1016/j.compag.2018.07.032

Kala, K.U., M. Nandhini, M. Thangadarshini, M.K. Chakkravarthi and M. Verma. 2023. Leveraging Deep Learning for Effective Pest Management in Plantain Tree Cultivation. P425-434. In International Conference on Soft Computing and Signal Processing. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-8628-6_36

Luaibi, A.R., T.M. Salman and A.H. Miry. 2021. Detection of citrus leaf diseases using a deep learning technique. Int. J. Electr. Comput. Eng., 11(2): 1719. https://doi.org/10.11591/ijece.v11i2.pp1719-1727

Malahina, E.A.U., M. Saitakela, S.J. Bulan, M.I.J. Lamabelawa and Y.S. Belutowe. 2024. Teachable Machine: Optimization of Herbal Plant Image Classification Based on Epoch Value, Batch Size and Learning Rate. J. Appl. Data Sci., 5(2): 532-545. https://doi.org/10.47738/jads.v5i2.206

Mamat, N., M.F. Othman, R. Abdoulghafor, S.B. Belhaouari, N. Mamat and S.F.M. Hussein. 2022. Advanced technology in agriculture industry by implementing image annotation technique and deep learning approach: A review. Agric., 12(7): 1033. https://doi.org/10.3390/agriculture12071033

Nag, A., P.R. Chanda and S. Nandi. 2023. Mobile app-based tomato disease identification with fine-tuned convolutional neural networks. Comp. Electric. Eng., 112: e108995. https://doi.org/10.1016/j.compeleceng.2023.108995

Nayak, A., S. Chakraborty and D.K. Swain. 2023. Application of smartphone-image processing and transfer learning for rice disease and nutrient deficiency detection. Smart Agric. Tech., 4: 100195. https://doi.org/10.1016/j.atech.2023.100195

Odabas, M.S., N. Senyer and S.O. Kaya. 2024. Classification of Apple Diseases and Pests using The Google.com Powered Teachable Machine. J. Agric. F. Gazio., 41(2): 66-71. https://doi.org/10.55507/gopzfd.1287389

Parate, R.K., K.M. Dhole and S.J. Sharma. 2023. Classification of Leaf using Teachable Machine. Inter. J. Res. Appl. Sci. and Eng. Tech., 11(9): 307-311. https://doi.org/10.22214/ijraset.2023.55629

Qadri, S.A.A., N.F. Huang, T.M. Wani and S.A. Bhat. 2025. Advances and Challenges in Computer Vision for Image-Based Plant Disease Detection: A Comprehensive Survey of Machine and Deep Learning Approaches. IEEE Trans. Auto. Sci. Eng., 22: 2639-2670. https://doi.org/10.1109/TASE.2024.3382731

Saini, A. 2024. Identifying Plant Diseases with Image Recognition. Polygence. https://doi.org/10.58445/rars.1900

Singh, U.P., S.S. Chouhan and S. Jain. 2019. Multilayer convolutionneural network for the classification of mango leaves infected byanthracnose disease. IEEE Access., 7: 43721-43729. https://doi.org/10.1109/ACCESS.2019.2907383

Storey, G., Q. Meng and B. Li. 2022. Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture. Sustain., 14(3): 1458. https://doi.org/10.3390/su14031458

Wang, G., Y. Sun and J. Wang. 2017. Automatic image-based plant disease severity estimation using deep learning. Comput. Intell. Neurosci., e2917536. https://doi.org/10.1155/2017/2917536

Wang, Y., F. Subhan, S. Shamshirband, M. Zubair Asghar, I. Ullah, and A. Habib. 2020. Fuzzy-based sentiment analysis system for analyzing Student feedback and satisfaction. Comput. Mater. Continua., 62(2): 631-655. https://doi.org/10.32604/cmc.2020.07920

Zhang, H., Y. Chen, X. Liu, Y. Huang, B. Zhan, and W. Luo. 2021. Identification of Common Skin Defects and Classification of Early Decayed Citrus Using Hyperspectral Imaging Technique. Food Analytic. Methods., pp. 1-18 https://doi.org/10.1007/s12161-020-01960-8.

Zhu, H., D. Wang, Y. Wei, X. Zhang and L. Li. 2024. Combining Transfer Learning and Ensemble Algorithms for Improved Citrus Leaf Disease Classification. Agri., 14(9): 1549. https://doi.org/10.3390/agriculture14091549