Detection Leaf Blight Using Guava Leaves Imaging

Subjects: Computer Science, Artificial Intelligence

Contributor: Sidrah Mumtaz , Saeed Ur Rehman , Mudassar Raza , Ofonime Dominic Okon , Adham E. Ragab , Hafiz Tayyab Rauf

Fruit is an essential element of human life and a significant gain for the agriculture sector. Guava is a common fruit found in different countries. It is considered the fourth primary fruit in Pakistan. Several bacterial and fungal diseases found in guava fruit decrease production daily. Leaf Blight is a common disease found in guava fruit that affects the growth and production of fruit. Automatic detection of leaf blight disease in guava fruit can help avoid decreases in its production.

CNN DarkNet-53 deep learning entropy

1. Introduction

Food is the fundamental requirement for the existence of human beings, and it is the notable outcome of agricultural activities. Agriculture is assumed to be the backbone of economic development, as it exhibits the cultivation of multiple crops, fruits, and vegetables. There is a large difference between the cultivation and annual production of fruits because of inappropriate advancements in technology, lack of knowledge, and diseases that negatively affect the production [1]. Disease detection in plants is a challenging task and is essential to diagnose at early stages. Diseases are mostly diagnosed through leaves because they tend to highlight contaminated parts immediately. Guava is an important fruit in agriculture; therefore, its leaves are selected for the detection and recognition of diseases [2]. Guava is nutritionally beneficial, serving calcium and iron to the human body. It is cultivated in America, especially in Mexico, Thailand, South Africa, and many other countries. Many laboratories such as the Central Institute of Subtropical Horticulture (CISH) and different institutes are continuing to work on guava production in different areas of the world [3]. Several diseases, such as bacterial and fungal diseases, attack the guava fruit, which badly affects its production [4]. There are different techniques of ML applied for disease detection. Almost 177 types of diseases are found that damage leaves, causing leaf blight and leaf spots. Known diseases include brown roots, twig drying, bacterial wilt, anthracnose, ring rots, and many others [5].

Many researchers aim for innovations in disease detection. Disease detection relies on five major steps. Usually, the first step in image processing is image acquisition. After obtaining images, preprocessing incorporates multiple steps that result in better accuracy. After preprocessing, feature extraction is performed, where the features of the images are boosted for further computation and selection. The final stage is classification. A variety of models are presented using diverse methodologies such as convolution neural network (CNN), gradient descent (GD), and many others for classification purposes [6]. Convolution neural networks play an essential part in the extraction of features through hidden layers, as manual extraction is costly and time-consuming [7]. Plant pathologists need an automatic detection system to diagnose leaf blight in plants.

2. Detection Leaf Blight Using Guava Leaves Imaging

Diseases in fruit plants and leaves are a major cause of destruction and economic loss. Automated systems help greatly with the detection of diseases at early stages. While considering the field of detection of disease in plants, deep neural networks work perfectly to identify and classify diseases. These networks are mentored to conduct high-value results in detecting and classifying diseases, and to fulfill the demands of food deformation prevention.

There are different methods for collecting images under certain conditions. Images are captured by multiple appliances, such as cameras, sensors, mobile phones, and other devices. In this era, more datasets containing guava are publicly available on multiple forums, such as Kaggle, Mendeley, and many others [8]. Pre-processing of images is an important phase in image processing. The pre-processing phase entails multiple steps which help highlight the focused parts and remove irrelevant information from guava leaf images. In the real world, label noise on images is a matter of concern. Multiple techniques have attained the best results in denoising images, especially mixed noise, speckle noise, and salt and pepper noise. Low contrast and color distortion in guava leaf images make them blur. Scattering and light absorption also affect clear image visualization [9]. Images are preprocessed by using rotational filters such as horizontal and vertical flipping. Data augmentation techniques are used, such as applying rotations and zooming into images [10]. Color spacing techniques are extensively applied in image processing. RGB, CIELAB, and CMYK models are mostly used as color spacing techniques to give the best results.

Feature extraction is a process in which features are reduced from the raw dataset and new features for manageable processing are created. Texture analysis has a wide range of applications [11]. Pattern recognition requires feature extraction to solve problems in prediction, cluster discrimination, and representation of data in the best way [12][13]. Content-Based Image Retrieval (CBIR) converts high-level image visuals into feature vectors that contain some properties [14]. There are multiple techniques to extract the features from guava leaf images, such as handcrafted-based features, region-based features, deep CNN-based features, texture-based features, color-based features, morphological-based features, etc. Extraction of features is categorized into hand-crafted-based features and deep-based features.

The selection of features from plant leaf images is carried out after the extraction of hand engineered and deep-based features. The set of features is chosen while noisy, poor, and extra features are eliminated from the original set of features [15]. There are five main types of feature selection, which are (1) Linear Method, (2) Non-Linear Method, (3) Filter-Based Method, (4) Wrapper Method, (5) Embedded Method. Linear methods include PCA and LDA. PCA stands for Principal Component Analysis, which is used for data reduction [16]. LDA stands for Linear Discriminant Analysis and is used for the conversion of high dimension features into lower dimension features [17]. Non-linear methods include Entropy [18], Genetic Algorithm (GA) [19], Binary Gray Wolf [20], Slap Swarm [21], Atom Search [22] and many others. Filter methods includes mRMR [23], Missing Value Ratio [24], and many others.

Wrapper methods include Jackstraw [25] and Boruta [26]. Finally, embedded methods include LASSO [27], Ridge [28], Elastic [29], and many others. Image fusion [30] helps greatly in improving classifier accuracy with less computational cost [31]. Different algorithms are proposed that use image fusion to get the best accuracy results.

Image classification is the last step in image processing [32]. Classification tends to dominate the feature vector to determine which object belongs to which class [33]. There are different types of techniques used for the classification of healthy and diseased images of plant leaves. Image classification is divided into three main categories, which are (1) Supervised Learning, (2) Unsupervised learning, and (3) Object-based image analysis. Supervised Learning is used to detect the new category of the object from training data [34]. Unsupervised Learning is a process in which an image is identified in an image collection without using labeled training data [35]. Object-based analysis involves the grouping of pixels on the basis of some similarities such as shape and neighborhood [36]. To get the most accurate results, the Plant Village dataset is used for testing and training purposes. A total of 80% of guava leaf images are used for testing while 20% of them are used for training purposes. The achievable accuracy is 97.22% using Alex-Net and Squeeze-Net after segmentation and classification [37]. Atila et al. [38] designed the Efficient-Net architecture, which is designed for classification purposes. Different architectures are applied using CNN for model training to get highly accurate results. The model training is performed on the dataset of 87,848 images. Images are preprocessed using different techniques such as downscaling and squaring methods, they are then classified, and an accuracy of 99.53% is achieved using AlexNet, VGG-16, and GoogLeNet [39]. Several algorithms are used for image classification. These are SVM [40], K-Nearest Neighbor (KNN) [41], Naïve Bayes [42], Shadow algorithms [43], Minimum Mean Distance (MMD) [44], Decision Trees [45], K-Means Clustering [46] and many others. The datasets are frequently classified by SVM. This involves supervised learning and comprises points that are in the sample space and different regions [47]. Segmentation is performed on the preprocessed data in three stages. In the first stage, the deep CNN is trained to learn the mapping from the space map. In the second stage, prediction-based labels are acquired. At the last stage, these acquired labeled images are sent to SVM for classification and achieve an accuracy of 86% [48]. In machine learning, KNN is a statistical classification algorithm. It gathers the objects selected by neighbors having the highest number of votes [49]. KNN is inspected for the detection of weeds from UAV images of the chili crop of Australia. In comparison with KNN, SVM and Random Forest (RF) are used. The achievable accuracies across RF, SVM, and KNN are 96%, 94%, and 63%, respectively [50]. KNN is also used for classifying facial expressions [51]. Additionally, KNN is used for the classification of grape leaves into healthy and unhealthy leaves. Texture-based and color-based features are extracted from grape leaf images and are classified by the KNN classifier, and an accuracy of 96.66% is achieved [52]. **Table 1** depicts an overview of recent works related to plants diseases analysis

Table 1. An overview of recent literature regarding plants diseases analysis.

Ref.	Year	Techniques	Dataset	Diseases	Results %
[39]	2018	Downscaling and squaring method, AlexNet, VGG-16, AlexNetOWTBn.	87,848 58 classes	Apple Scab, Black Rot, Early Blight,	99.53

Ref.	Year	Techniques	Dataset	Diseases	Results %
				Brown Leaf Spot	
[53]	2018	ResNet-50, Deep Siamese convolutional network, TSNE method, KNN	PVD	Black Rot, Esca, Chlorosis	90
[54]	2018	Transfer learning, F-RCNN, classification.	4923	Phoma Rot, Leaf Miner, Target Spot	95.75
[55]	2019	F-CNN, S-CNN, Segmentation, annotation and labeling on region of interest (lesions), random transformation (stretch/rotation/brightness/contrast blur)	Independent dataset	Spider Mite, Target Spot	98.6
[56]	2019	VGG classification, resizing and transformation of images into grayscale,	2465	Black dot and scurf	96
[38]	2021	Efficient-Net (B5Ver), Alexnet, ResNet50, classification.	61,486	Late Blight, Bacterial Spot	99.97
[57]	2021	Resizing, normalizing and augmentation, Efficient-Net (B7Ver), Efficient-Net (B4Ver), U-net, and modified U-net segmentation model, Score-Cam visualization technique,	18,161	Target Spot	99.9

References

1. Rehman, A.; Jingdong, L.; Shahzad, B.; Chandio, A.A.; Hussain, I.; Nabi, G.; Iqbal, M.S. Economic perspectives of major field crops of Pakistan: An empirical study. Pac. Sci. Rev. B Humanit. Soc. Sci. 2015, 1, 145–158.
2. Rai, M.K.; Asthana, P.; Jaiswal, V.; Jaiswal, U. Biotechnological advances in guava (Psidium guajava L.): Recent developments and prospects for further research. Trees 2010, 24, 1–12.
3. Mitra, S.; Thingreingam Irenaeus, K. Guava cultivars of the world. In International Symposia on Tropical and Temperate Horticulture—ISTTH2016; CIRAD Publications: Cairns, Queensland, Australia, 2016; pp. 905–910.
4. Almadhor, A.; Rauf, H.T.; Lali, M.I.U.; Damaševičius, R.; Alouffi, B.; Alharbi, A. AI-Driven Framework for Recognition of Guava Plant Diseases through Machine Learning from DSLR Camera Sensor Based High Resolution Imagery. Sensors 2021, 21, 3830.
5. Misra, A. Guava diseases—Their symptoms, causes and management. In Diseases of Fruits and Vegetables; Springer: Berlin/Heidelberg, Germany, 2004; Volume II, pp. 81–119.
6. Dhiman, B.; Kumar, Y.; Hu, Y.-C. A general purpose multi-fruit system for assessing the quality of fruits with the application of recurrent neural network. Soft Comput. 2021, 25, 9255–9272.

7. Ashraf, R.; Habib, M.A.; Akram, M.; Latif, M.A.; Malik, M.S.A.; Awais, M.; Dar, S.H.; Mahmood, T.; Yasir, M.; Abbas, Z. Deep convolution neural network for big data medical image classification. IEEE Access 2020, 8, 105659–105670.
8. Bojer, C.S.; Meldgaard, J.P. Kaggle forecasting competitions: An overlooked learning opportunity. Int. J. Forecast. 2021, 37, 587–603.
9. Fu, X.; Cao, X. Underwater image enhancement with global–local networks and compressed-histogram equalization. Signal Process. Image Commun. 2020, 86, 115892.
10. Tang, Z.; Gao, Y.; Karlinsky, L.; Sattigeri, P.; Feris, R.; Metaxas, D. OnlineAugment: Online Data Augmentation with Less Domain Knowledge; Springer: Berlin/Heidelberg, Germany, 2020; pp. 313–329.
11. Liu, Z.; Lai, Z.; Ou, W.; Zhang, K.; Zheng, R. Structured optimal graph based sparse feature extraction for semi-supervised learning. Signal Process. 2020, 170, 107456.
12. Ghojogh, B.; Samad, M.N.; Mashhadi, S.A.; Kapoor, T.; Ali, W.; Karray, F.; Crowley, M. ‘Feature selection and feature extraction in pattern analysis: A literature review. arXiv 2019, arXiv:1905.02845.
13. Zhang, J.; Liu, B. A review on the recent developments of sequence-based protein feature extraction methods. Curr. Bioinform. 2019, 14, 190–199.
14. Latif, A.; Rasheed, A.; Sajid, U.; Ahmed, J.; Ali, N.; Ratyal, N.I.; Zafar, B.; Dar, S.H.; Sajid, M.; Khalil, T. Content-based image retrieval and feature extraction: A comprehensive review. Math. Probl. Eng. 2019, 2019, 9658350.
15. Raj, R.J.S.; Shobana, S.J.; Pustokhina, I.V.; Pustokhin, D.A.; Gupta, D.; Shankar, K. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. IEEE Access 2020, 8, 58006–58017.
16. Ali, L.; Wajahat, I.; Golilarz, N.A.; Keshtkar, F.; Bukhari, S.A.C. ‘LDA–GA–SVM: Improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine. Neural Comput. Appl. 2021, 33, 2783–2792.
17. Ozyurt, B.; Akcayol, M.A. A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA. Expert Syst. Appl. 2021, 168, 114231.
18. Zhao, J.; Liang, J.-M.; Dong, Z.-N.; Tang, D.-Y.; Liu, Z. Accelerating information entropy-based feature selection using rough set theory with classified nested equivalence classes. Pattern Recognit. 2020, 107, 107517.
19. Maleki, N.; Zeinali, Y.; Niaki, S.T.A. A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. Expert Syst. Appl. 2021, 164, 113981.

20. Hu, P.; Pan, J.-S.; Chu, S.-C. Improved binary grey wolf optimizer and its application for feature selection. Knowl. Based Syst. 2020, 195, 105746.

21. Tubishat, M.; Idris, N.; Shuib, L.; Abushariah, M.A.; Mirjalili, S. Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection. Expert Syst. Appl. 2020, 145, 113122.

22. Too, J.; Rahim Abdullah, A. Binary atom search optimisation approaches for feature selection. Connect. Sci. 2020, 32, 406–430.

23. Özyurt, F. A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine. Soft Comput. 2020, 24, 8163–8172.

24. De Silva, K.; Jönsson, D.; Demmer, R.T. A combined strategy of feature selection and machine learning to identify predictors of prediabetes. J. Am. Med. Inform. Assoc. 2020, 27, 396–406.

25. Germain, P.-L.; Sonrel, A.; Robinson, M.D. pipeComp, a general framework for the evaluation of computational pipelines, reveals performant single cell RNA-seq preprocessing tools. Genome Biol. 2020, 21, 1–28.

26. Kaur, P.; Singh, A.; Chana, I. Computational techniques and tools for omics data analysis: State-of-the-art, challenges, and future directions. Arch. Comput. Methods Eng. 2021, 28, 4595–4631.

27. Ghosh, P.; Azam, S.; Jonkman, M.; Karim, A.; Shamrat, F.J.M.; Ignatious, E.; Shultana, S.; Beeravolu, A.R.; De Boer, F. Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques. IEEE Access 2021, 9, 19304–19326.

28. Deepa, N.; Prabadevi, B.; Maddikunta, P.K.; Gadekallu, T.R.; Baker, T.; Khan, M.A.; Tariq, U. An AI-based intelligent system for healthcare analysis using Ridge-Adaline Stochastic Gradient Descent Classifier. J. Supercomput. 2021, 77, 1998–2017.

29. Amini, F.; Hu, G. A two-layer feature selection method using genetic algorithm and elastic net. Expert Syst. Appl. 2021, 166, 114072.

30. Zhang, H.; Xu, H.; Xiao, Y.; Guo, X.; Ma, J. Rethinking the Image Fusion: A Fast Unified Image Fusion Network Based on Proportional Maintenance of Gradient and Intensity. Proc. AAAI Conf. Artif. Intell. 2020, 34, 12797–12804.

31. Jung, H.; Kim, Y.; Jang, H.; Ha, N.; Sohn, K. Unsupervised deep image fusion with structure tensor representations. IEEE Trans. Image Process. 2020, 29, 3845–3858.

32. Shakya, S. Analysis of artificial intelligence based image classification techniques. J. Innov. Image Process. (JIIP) 2020, 2, 44–54.

33. Hong, D.; Wu, X.; Ghamisi, P.; Chanussot, J.; Yokoya, N.; Zhu, X.X. Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification. IEEE Trans.

Geosci. Remote Sens. 2020, 58, 3791–3808.

34. Ke, Z.; Qiu, D.; Li, K.; Yan, Q.; Lau, R.W. Guided Collaborative Training for Pixel-Wise Semi-Supervised Learning; Springer: Berlin/Heidelberg, Germany, 2020; pp. 429–445.
35. Wang, F.; Liu, H.; Guo, D.; Sun, F. Unsupervised Representation Learning by InvariancePropagation. arXiv 2020, preprint. arXiv:2010.11694.
36. Karantanellis, E.; Marinos, V.; Vassilakis, E.; Christaras, B. Object-based analysis using unmanned aerial vehicles (UAVs) for site-specific landslide assessment. Remote Sens. 2020, 12, 1711.
37. Durmuş, H.; Güneş, E.O.; Kırcı, M. Disease Detection on the Leaves of the Tomato Plants by Using Deep Learning; IEEE: Piscataway, NJ, USA, 2017; pp. 1–5.
38. Atila, Ü.; Uçar, M.; Akyol, K.; Uçar, E. Plant leaf disease classification using EfficientNet deep learning model. Ecol. Inform. 2021, 61, 101182.
39. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. 2018, 145, 311–318.
40. Azarmdel, H.; Jahanbakhshi, A.; Mohtasebi, S.S.; Muñoz, A.R. Evaluation of image processing technique as an expert system in mulberry fruit grading based on ripeness level using artificial neural networks (ANNs) and support vector machine (SVM). Postharvest Biol. Technol. 2020, 166, 111201.
41. Kasinathan, T.; Singaraju, D.; Uyyala, S.R. Insect classification and detection in field crops using modern machine learning techniques. Inf. Process. Agric. 2021, 8, 446–457.
42. Yao, J.; Ye, Y. The effect of image recognition traffic prediction method under deep learning and naive Bayes algorithm on freeway traffic safety. Image Vis. Comput. 2020, 103, 103971.
43. Atouf, I.; Al Okaishi, W.Y.; Zaaran, A.; Slimani, I.; Benrabh, M. A real-time system for vehicle detection with shadow removal and vehicle classification based on vehicle features at urban roads. Int. J. Power Electron. Drive Syst. 2020, 11, 2091.
44. Khan, M.A.; Khan, M.A.; Ahmed, F.; Mittal, M.; Goyal, L.M.; Hemanth, D.J.; Satapathy, S.C. Gastrointestinal diseases segmentation and classification based on duo-deep architectures. Pattern Recognit. Lett. 2020, 131, 193–204.
45. Chen, J.; Lian, Y.; Li, Y. ‘Real-time grain impurity sensing for rice combine harvesters using image processing and decision-tree algorithm. Comput. Electron. Agric. 2020, 175, 105591.
46. Moubayed, A.; Injadat, M.; Shami, A.; Lutfiyya, H. Student engagement level in an e-learning environment: Clustering using k-means. Am. J. Distance Educ. 2020, 34, 137–156.

47. Nanglia, P.; Kumar, S.; Mahajan, A.N.; Singh, P.; Rathee, D. A hybrid algorithm for lung cancer classification using SVM and Neural Networks. ICT Express 2021, 7, 335–341.

48. Wu, W.; Li, D.; Du, J.; Gao, X.; Gu, W.; Zhao, F.; Feng, X.; Yan, H. An intelligent diagnosis method of brain MRI tumor segmentation using deep convolutional neural network and SVM algorithm. Comput. Math. Methods Med. 2020, 2020, 6789306.

49. Basha, C.Z.; Rohini, G.; Jayasri, A.V.; Anuradha, S. Enhanced and Effective Computerized Classification of X-Ray Images; IEEE: Piscataway, NJ, USA, 2020; pp. 86–91.

50. Islam, N.; Rashid, M.M.; Wibowo, S.; Xu, C.-Y.; Morshed, A.; Wasimi, S.A.; Moore, S.; Rahman, S.M. Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm. Agriculture 2021, 11, 387.

51. Sasankar, P.; Kosarkar, U. A Study for Face Recognition Using Techniques PCA and KNN. EasyChair Print. 2021. Available online: https://wwww.easychair.org/publications/preprint_download/gS7Q (accessed on 30 January 2023).

52. Bharate, A.A.; Shirdhonkar, M. Classification of Grape Leaves Using KNN and SVM Classifiers; IEEE: Piscataway, NJ, USA, 2020; pp. 745–749.

53. Goncharov, P.; Ososkov, G.; Nechaevskiy, A.; Uzhinskiy, A.; Nestsiarenia, I. Disease Detection on the Plant Leaves by Deep Learning; Springer: Berlin/Heidelberg, Germany, 2018; pp. 151–159.

54. De Luna, R.G.; Dadios, E.P.; Bandala, A.A. Automated Image Capturing System for Deep Learning-Based Tomato Plant Leaf Disease Detection and Recognition; IEEE: Piscataway, NJ, USA, 2018; pp. 1414–1419.

55. Sharma, P.; Berwal, Y.P.S.; Ghai, W. Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. Inf. Process. Agric. 2020, 7, 566–574.

56. Oppenheim, D.; Shani, G.; Erlich, O.; Tsror, L. Using deep learning for image-based potato tuber disease detection. Phytopathology 2019, 109, 1083–1087.

57. Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Ayari, M.A.; Khan, A.U.; Khan, M.S.; Al-Emadi, N.; Reaz, M.B.I.; Islam, M.T.; Ali, S.H.M. Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques. AgriEngineering 2021, 3, 294–312.