Fruit Disease Classification and Detection Using Machine Learning

Ashish Patel, Department of Computer Science & Engineering, RDEC, Ghaziabad. patel.ashish02@gmail.com

Abstract

Due of visual impairments, researchers and regulators of fruit diseases used to rely on human eyesight. The minor shifts in fruit colour represent the movements. Sometimes, little differences in colour and pattern might be a sign of genuineness. Physical monitoring and detection of microorganisms, the subsequent stage in infection diagnosis, is a laborious, costly, and inaccurate procedure. Therefore, to conduct a quick and accurate diagnosis, it is best to employ certain MATLAB approaches that are more dependable than other, more antiquated ways. The presence of lesions, diseased fruit, and leaf spots indicates that the plant is suffering from an infection or sickness. Accurately identifying the illness from the given picture is the objective of this assignment. It is necessary to do image segmentation, preprocessing, feature extraction, and labelling. Infectious disorders such as the flu, strep throat, and staph may spread via the environment, which can include bugs and environmental factors. Here, we'll need to look at spoiled fruit to figure out what went wrong. In order to classify the illness, we will extract image characteristics such the major and minor axes of the fruit.

Keywords:K-Means Clustering, Local Binary Pattern, Multi-class Support Vector Machine, TextureClassification

I. Introduction

Computer vision research primarily aims to build a reliable recognition system that can perform as well as, or better than, humans. As far as agricultural scientific data collection and interpretation is concerned, a picture really is worth a thousand words. Up until recently, photography was the only method for reliably recording and reproducing such data. Mathematical processing and quantification of visual input is challenging. Thanks to the growth of computers and microelectronics alongside conventional photography, new technologies for digital image processing and analysis have emerged, successfully overcoming these issues. Images captured with any lens, from the most minute to the most telescopic, may benefit from this application.

Monitoring the health of crops, such as fruit trees and vegetables, on a regular basis is essential for sustainable farming. Regrettably, there aren't any commercially accessible sensors that can continuously monitor the health of trees at this time. Scouting is the most common way to assess a tree's health, but it may be a time-consuming, costly, and physically demanding process. Diagnosing fruit illnesses requires molecular methods like polymerase chain reaction, which is labor-intensive and takes a long time and requires a lot of samples.

A wide variety of fruit diseases have the potential to drastically cut harvests. In addition to decreasing yields, fruit diseases may cause once-popular cultivars to decrease or even become extinct. For the purpose of controlling disease vectors, increasing production, and making effective use of fungicides, disease-specific pesticides, and insecticides, early disease and crop health identification is essential. Experts have always relied on naked eye examinations as the primary method for identifying and diagnosing fruit illnesses. Accessing a clinic or hospital with trained professionals may be a daunting and expensive task in certain underdeveloped nations.

Fruit infections may drastically lower yields and quality if they show up during harvest. The fungus known as soybean rust has wreaked havoc on the global economy. But if the disease can be completely eliminated, farmers may get back more than $11 million.

If a disease develops in the fruit, it might spread to the tree's branches and leaves as well. Reducing such losses and halting the spread of disease might be possible with the development of early detection technologies for fruit concerns.

There has been a lot of effort to automate the use of machine vision for visually inspecting fruit for size and colour. Nevertheless, flaw identification in images is still challenging due to the wide variety of flaw kinds, the existence of stems and calyxes, and the inherent skin colour variation across fruit species. Analysing the data is vital for taking the required safeguards the next year to prevent the same losses.

This study lays out a plan for creating autonomous systems for farming using photographs captured from far away. The development of several image processing programmes has greatly benefited agricultural endeavours. To import photos into these apps, you'll need either a camera-based system or a colour scanner. We have made an effort to apply cutting-edge methods for processing and analysing images to various agricultural issues.

The advancement of computing technology is closely followed by the evolution and improvement of computer-based image processing. Unfortunately, the specialised imaging systems that are now on the market are quite costly and lack flexibility, requiring the user to click a few buttons before the data is shown. Furthermore, the means by which these outcomes are achieved are not readily apparent. Fruit diseases that cause spots may be disastrous if not addressed promptly.Because agricultural goods are more likely to have harmful residue levels on them, pesticide usage for fruit disease treatment is a major source of ground water pollution. Careful application is required when using pesticides because of their high manufacturing costs. Hence, we've made an effort to develop a method for early disease detection in fruit trees, which is crucial for effective treatment.

Scabs, rots, and blotches are some of the most common illnesses that may harm apples. Garish or brown corky patches are what you see on apples when they have scabs. Infected apples may have crimson haloes around brown or black spots. Blotch is a fungal disease that affects apples in a way that looks like black spots with uneven or lobed borders.

Here, we provide an adaptive method for automated picture-based fruit disease identification and experimentally assess its efficacy. A Multi-class Support Vector Machine is used to detect illnesses in fruit after first segmenting the photos of fruit using the K-Means clustering approach. Subsequently, certain state-of-the-art characteristics are retrieved from the segmented images. We demonstrate that classifiers like Multi-class Support Vector Machine and clustering strategies for disease division may significantly enhance the efficacy of automated fruit pathogen diagnosis. Three distinct apple diseases—apple blotch, apple rot, and apple scab—have been taken into account to guarantee the practicability of the suggested method. Experimental results for automatically identifying and diagnosing fruit illnesses have shown the efficacy of the suggested method. Magnetic resonance imaging (MRI), x-ray imaging (x-ray), etc. are imaging techniques that scientists use to find fruit defects; however, these methods are prohibitively expensive for farmers to implement, necessitate an excessive amount of space, demand an advanced degree of scientific literacy from consumers, and have a detrimental effect on the research specimens. When it comes to diagnosing problems with fruit, doctors can only rely on their own eyes. The vast distances between cities in certain underdeveloped nations could make it difficult to get an appointment with a specialist.Because of this, their potential users and the population they cater to are both constrained. One disease may infect a tree all the way down to its twigs. Whether it's a darker spot on the fruit's surface or an inside mark, every fruit disease is easily identifiable. With these peaks, we might potentially find the first indicators of fruit destruction. Determining if fruit is diseased calls for extensive training and manual labour.

Methodology

Many potential causes could be at the heart of the decline in fruit yield. One important factor, however, is the possibility that infectious illnesses are present in the fruit itself. It is difficult to forecast how accurate the procedure will be for detecting the fruit disease. Consequently, state-of-the-art strategies and procedures are required for the accurate and quick classification of agricultural diseases. Failure to diagnose fruit illnesses in a timely manner might have a detrimental effect on fruit yield. It is possible for farmers to lose fruit output due to a misdiagnosis, especially when they physically examine fruits and seek pesticide guidance from local agricultural authorities. Manual inspection in agricultural contexts is labor-intensive and tedious. By applying machine learning algorithms to preprocessed photos, tasks such as identifying the source of a fruit disease and choosing the appropriate pesticide may be completed more quickly and with more precision.Converting images to grayscale, reducing noise, smoothing them out, and other enhancements are typical in the first step of image processing. The picture will get its own distinct characteristics and a more refined appearance after preprocessing.Secondly, the input window is notified of the most important changes in the image's characteristics—the pixel values. Compressing the picture or locating its boundaries may help in component extraction.Last but not least, segmentation isolates the target region from the original picture.Sorting images according to their attributes is the fourth stage.Deep Learning Architectural Methods: One kind of deep learning that has been widely used in computer vision applications is convolutional neural networks. Convolutional, pooling, and fully connected layers are some of the building blocks that may go into the model. Then, to learn more complex spatial feature systems, it might be trained using the back propagation approach. Neural network-based disease classification in citrus fruits:

It is necessary to break down an input image into its individual pixels before it can be understood. One way to describe the picture is as a three-by-three grid of the three primary colors—red, blue, and green.Following the first picture processing, the second CNN layer is responsible for extracting features from the input matrix.

In this step, convolutional techniques are used to both the input matrix and the filter grid.Maxpooling in the third layer: The resultant convolved feature map is sent to Maxpooling after applying the convolutional layer learned in the previous CNN model layers. As a whole, the pooling layer averages out the picture lattice components.We apply many filtering techniques to the feature map at this layer in order to retrieve the high-level features. As the CNN model progresses through its layers, it hones down on characteristics, eventually concentrating on the most important ones.

One of the primary goals of the second Max pooling layer is to reduce the dimensionality for easier feature retrieval. The second max pooling layer produced a long featured vector, which the sixth layer, the flattening layer, is tasked with simplifying.

Functions for Activation Module The images are shown in galleries on this page. Data obtained from straightening layers may be used to characterise citrus fruit diseases. To determine the likelihood of the most common diseases impacting citrus crops, we use a SoftMax enactment method. To categorise the supplied image, the most prevalent likelihood is employed.

Preprocessing

To enhance the image data, either by amplifying certain picture elements necessary for further processing and analysis or by suppressing undesirable distortions, it is usual practice to execute operations on pictures at the lowest conceivable level of abstraction. Neither the images nor the information they carry are of higher quality. The techniques it employs make extensive use of visual redundancy. In most cases, you'll see very similar, if not identical, brightness levels among nearby pixels that depict the same physical item. If a distorted pixel is located independently in the picture, its initial value may be recovered by averaging it with its nearby pixels. The suggested technique first stores the acquired picture in an image database after passing it through a number of image pre-processing techniques.

The input picture shows maize fruits that have pest illnesses such as applefruitspot and foliar fruit spot, which are caused by bacteria. Various lighting settings are used to photograph the processed fruits after they are randomly picked from the maize field. Scaling the picture to 256x256 is the first step in getting it ready for segmentation.

Picture segmentation is a useful tool for separating out relevant parts of a picture for further study. We use the simple k-means clustering technique for picture segmentation. In order to separate and retrieve visual items, segmentation is necessary, regardless of how hazy their boundaries may be. In order to construct a high number of clusters, clustering depends on the capacity to precisely distinguish different picture objects from one another. Therefore, a chromatic-image-space transformation is what we're using. It makes complete sense to use a colour space with two chromaticity levels and a brightness layer as the components, since no extra colour components are needed. A matrix of geometric distances may be used to quantify the degree of similarity between colours. A sample is often placed into a category

SHRADHA EDUCATIONAL ACADEMY

using the K-means distance. Using the pixel coordinates, K-means gives each picture cluster a numerical value.

Segmenting Images using K-means clustering is common practice for the sake of analysis and comprehension. For all intents and purposes, a cluster is just a collection of related but different items. Clustering involves dividing a dataset into smaller, more manageable pieces based on shared characteristics using a predetermined distance measure. Collections of images may be defined by the forms or textures that they share, among other things. To divide the data into groups, K-means clustering uses a threshold that has already been specified. The clusters' starting centres are chosen at random. The next step is to find the centroid of the dataset and connect all of the data points to it. The geometric centre of each cluster is used to assign each pixel to that cluster using the Euclidean distance measure.

Block Diagram

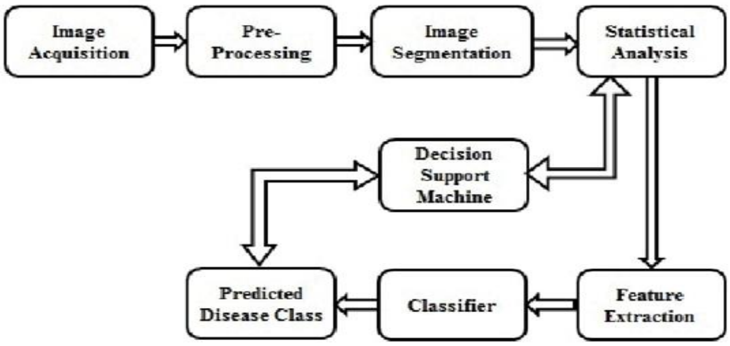

Fig1. Overall Diagram of System Architecture

Conclusion

Technology will play an increasingly important role in the future. On a daily basis, we get messages from farmers who are upset because they invested much in fertilisation only to have it wasted due to crop damage caused by viruses. Experts in this area are hard to come by. Getting advice from an expert is a good idea before doing something, as their view could differ from the average person's. It was shown that improving the accuracy of illness diagnoses is possible by increasing the quantity of training samples and modifying the parameters of the support vector machine. A framework for identifying and classifying several fruit-related illnesses is laid forth by this technique. K-Means segmentation is used to distinguish between healthy and infected tissue. Then, we use the GLCM and a support vector machine (SVM) to extract features from textures, and then we use those features to categorise the textures.

References

1.Tian Youwen and Li Chenghua, "Color Image Segmentation Method of Plant Disease Basing on Statistics Model Identification [J]", Journal of Jilin University (Gongxue Version), no. 02, pp. 291-293, 2004.

2.Shang Yijun, Zhang Shanwen and Zhang Yunlong, "Plant Disease Detection Method Basing on Image of the Plant Leaf [J]", Jiang su Agricultural Science, vol. 42, no. 04, pp. 340-342, 2014.

3.ChaoXiaofei, Research on Disease Identification and Disease Spot Segmentation Methods of Normal Apple Leaves Basing on Deep Learning [D], Northwest A&F University, 2021.

4.Guo Xingang, Wang Jia, QuNuoxi and Cheng Chao, "Canny SLIC Image Segmentation Algorithm Basing on Gradient Direction[J]", Computer Simulation, vol. 38, no. 09, pp. 465-469+500, 2021.

5.LanLuo, Yalan Ye and ZehuiQu, "A Mini Fish tailed Lion: The Intelligent Fishbone Based on Golden Fish", the 2012 International Conference on Systems and Informatics, 5. 2012.

6.ChengYanyan, "Method Simulation of Self Adaptive Fast Removal of Unaligned Distributed Redundant Data[J]", Computer Simulation, vol. 36, no. 09, pp. 389-392, 2019.

7.WangXiaofang, Fang Dengjie, He Hairui and ZouQianying, "MSRCR Image Defogging Algorithm Basing on Multi-scale Detail Optimization[J]", Experiment Technology and Management, vol. 37, no. 09, pp. 92-97, 2020.

8.Retinex Enhancing Algorithm of Image Processing (SSR MSR MSRCR), [online] Available: https://www.bbsmax.com.
9.R Achanta, A Shaji, K Smith et al., Slic super-pixels[R], 2010.
10.S Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions[J]", Int J Unc Fuzz Knowl Based Syst, vol. 6, no. 2, pp. 107-116, 1998.
11.LiuHuasong, Research on Segmentation Method of Toxic Strain Embryo Egg Image Basing on In-depth Learning[D], Tianjin:Tiangong University, 2019.
12.Goyal P Linty, R Girshick et al., "Focal loss for dense object detection[J]", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. PP, no. 99, pp. 2999-3007, 2017.
13.Yang Qirui, "Deep Network Fire Disaster Image Identification Algorithm Basing on improved DenseNet[J]", Computer Application and Software, vol. 36, no. 2, pp. 258-263, 2019.
14.Wu Yunzhi, Liu Aoyu, Zhu Xiaoning, Liu Chenxi, Fan Guohua, Le Yi, et al., "FI-DenseNet: Convolution Network Applied to Plant Disease Image Identification[J/OL]", Journal of Anhui Agricultural University, pp. 1-7, 11 2021.
15.LanLuo, QiongHai Dai, ChunXiangXu and ShaoQuan Jiang, "An Applicaiton Study to the Ciphers Weigh in Faithful Transmission", Applied Mechanics and Materials Switzerland, 2012.
16.Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz and Ali Yeon Bin MdShakaffRohaniBinti S Mohamed Farook, "Feasibilit y Study on Plant Chili Disease Detection Using Image Processing Techniques", 2012 Third International Conference on Intelligent Systems Modelling and Simulation.
17.R. Dhaya, "Flawless Identification of FusariumOxysporum in Tomato Plant Leaves by Machine Learning Algorithm", Journal of Innovative Image Processing (JIIP), vol. 2, no. 04, pp. 194-201, 2020
18.J. Samuel Manoharan, "Flawless Detection of Herbal Plant Leaf by Machine Learning Classifier Through Two Stage Authentication Procedure", Journal of Artificial Intelligence and Capsule Networks, vol. 3, no. 2, pp. 125-139, 2021.
19.GodliverOwomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, "Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease", Preceding of the 1'st international conference on the use of mobile ICT in Africa, 2014.
20 M. Sharif, M. A. Khan, Z. Iqbal, M. F. Azam, M. I. U. Lali, and M. Y. Javed, ''Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection,'' Comput.Electron.Agricult., vol. 150, pp. 220–234, Jul. 2018.
21 Manavalan, ''Automatic identification of diseases in grains crops through computational approaches: A review,'' Comput. Electron.Agricult., vol. 178, Nov. 2020, Art. no. 105802.
22 W. Pan, J. Qin, X. Xiang, Y. Wu, Y. Tan, and L. Xiang, ''A smart mobile diagnosis system for citrus diseases based on densely connected convolutional networks,'' IEEE Access, vol. 7, pp. 87534–87542, 2019.
23 G. Wang, Y. Sun, and J. Wang, ''Automatic image-based plant disease severity estimation using deep learning,'' Comput.Intell.Neurosci., vol. 2017, Jul. 2017, Art. no. 2917536.
24 U. P. Singh, S. S. Chouhan, S. Jain, and S. Jain, ''Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease,'' IEEE Access, vol. 7, pp. 43721–43729, 2019..
25https://www.kaggle.com/c/fruit-classification