Enhancing Coffee Crop Management with IoT and Machine Learning: Automated Monitoring and Disease Control

Chris Chettissery, P.S. Rajakumar and S. Geetha

March 26, 2023

Enhancing Coffee Crop Management with IoT and Machine Learning: Automated Monitoring and Disease Control

Chris Chettissery
Dr.M.G.R. Educational and Research Institute
Chennai-95
chrisphilip08@gmail.com

Dr.P.S. Rajakumar
Dr.M.G.R Educational and Research Institute
Chennai-95
rajakumar.subramanian@drmgrdu.ac.in

Dr. S. Geetha
Dr.M.G.R Educational and Research Institute
Chennai-95

***Abstract*- A rise in food production is necessary to keep pace with the rapid growth of the human population. Diseases with a high rate of spreading can severely reduce plant yields and even wipe out the entire plantation. One cannot overstate the value of early disease detection and prevention. Due to the increasing use of cell phones, even in the most remote areas, researchers have recently turned to automatic feature analytics as a technique for diagnosing crop disease. The convolutional, activation, pooling, and fully connected layers of the CNN have therefore been used in this work to create a disease identification approach. Predictions of soil factors including pH levels and water contents, illnesses, weed identification in crops, and species recognition are the sectors that have received the most attention. The micro-controller system keeps track of meteorological and atmospheric changes and uses sensors to estimate how much water should circulate in accordance. If a pesticide sprayer is attached to the hardware, the technique can also treat plant diseases. Data from the system is tracked and documented using a mobile application. Future farmers will benefit intelligently from the proposed methodology.**

Keywords—Machine Learning, Convolutional Neural Network (CNN), Automatic Coffee Disease Prediction, Image Processing.

I. INTRODUCTION

Crops are often thought of as the agriculture industry's heartbeat, but occasionally different illnesses can appear on plants, posing problems because of the damage they cause. When farmers are not able to recognize illnesses, they need to be educated, which will raise the price of farming. This makes manually tracking plant health and identifying plant infections a difficult and time-consuming activity. Agriculture and crop yields have a significant role in any country's economy. Diseases affecting plants and agriculture have a big impact on output. The bulk of manual inspection methods makes it difficult to detect plant diseases, which has a negative impact on crop output or quality. It is difficult for farmers or cultivators to keep track of plants and crops that are dispersed over a broad area. In rare circumstances, the farmer might not be aware of the sickness.

The technology uses a common smartphone and a machine learning methodology to forecast plant illnesses. The suggested system gathers information, such as photos of plant diseases, and uses that dataset to identify the different diseases that affect plants and crops. The ability to quickly diagnose illnesses with little to no human involvement could be advantageous to farmers. The suggested method also aids in protecting the produce by spotting illnesses in their early stages. A neural network-based model is developed to recognize different crop varieties and plant illnesses. Additionally, the algorithm recommends which insecticides to use for each group of diseases.

The growth of the human population and food production are both been greatly impacted by plant disease. When epidemics strike, food shortages might easily develop, causing havoc on human civilization due to low harvests and a lack of large storing facilities. Visual disease detection in plants could only be accomplished in certain locations and takes more time and precision. Automated disease detection is simpler and less expensive when symptoms are observed on plant leaves. In order to provide image-based automated controls, inspections, and robot guiding, machine vision is also enabled. To avoid losses in the quantity and quality of agricultural goods, it is essential to identify plant diseases. For agriculture to be sustained, the plant's health must be monitored and diseases must be found. The most popular method for studying plant disease involves patterns that are readily visible on plant leaves. Physically diagnosing plant illnesses is ineffective, difficult, and time-consuming. It also calls for continuous inspection, which can be expensive on large farms and requires an understanding of plant diseases. It is essential to identify plant diseases quickly, automatically, and accurately. Consequently, image processing technology is used to identify plant illnesses.

II. RELATED WORKS

Keeping agriculture effective and simple for farmers is the major goal of the advancements in smart farming methods. The technologies behind intelligent agriculture systems primarily focus on plant disease identification technologies and automatic smart drip irrigation facilities. In illness diagnosis solutions, the no. of diseases identified and the accuracy attained depend on the quantity of the dataset used.

It is becoming more and more important to use the Agriculture Cyber-Physical System (A-CPS) to increase crop growth and quantity while using the least area of property possible. In order to build the ACPS, this paper explains automated plant disease identification and introduces the Internet of Agro-Things (IoAT). Microbial infections affect the majority of the crop throughout agricultural output. In addition, farmers are unable to identify the constantly changing illnesses, prompting the creation of a disease predictive model. The study [1] prevents this by using a training CNN model to evaluate the cropped photos obtained by a health maintenance program. The solar sensor node is responsible for image capture, continual sensing, and smart automation. In comparison to competitors, the sensor node has a long life expectancy and is home to a recently founded soil moisture sensor. By using a solar sensor node with a camera module, a microcontroller, and a smartphone app, the now suggested system can be used in real time and enable a farmer to monitor the fields. After three months of testing, the prototype was found to be dependable, maintaining its rust-free condition while withstanding various weather patterns. A 99.24% accuracy rate is reached by the proposed plant disease predictive model.

The production of maize could be significantly hampered by a disease epidemic, costing millions of rupees. The danger of crop loss brought on by disease outbreaks can be reduced using machine learning techniques. The main symptoms of plant diseases include colour changes, the development of spots or decaying areas in the leaves, or both. Numerous image processing-based criteria are used in this work to identify corn infections based on these findings. To identify colour, objects, and other information in pictures, key points on the picture include “RGB, Scale-Invariant Feature Transform (SIFT), Sped-Up Robust Features (SURF), Oriented FAST and rotated BRIEF (ORB), and the Histogram Of Oriented Gradients (HOG)”. Several machine learning approaches are used in the article [2] to examine the efficacy of these features. The four techniques are “Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF), and Naive Bayes (NB)”. According to our tests, colour is the component of this project that is the most revealing. RGB is the attribute that has the best accuracy for most of the classifiers we test.

Any crop that farmers grow is said to be vulnerable to one or several illnesses[3]. It is difficult to manually evaluate plant health and find diseases. Image segmentation can therefore be a helpful and efficient method for identifying plant diseases. Diseases are categorized using colour attributes and edge enhancement. The technology enables both preventive measures and an illness percentage. Pre-processed, segmented pictures captured with a mobile camera are used to feature extraction and identify illnesses. Python sickness detection methods will be developed using the Open CV platform. In this study, the methods for diagnosing pomegranate diseases were discussed.

Plant disease poses a serious threat to both plant diversity and food supply, so monitoring tools are essential for sustainable maintenance. [4] Rice occupies a large part of Bangladesh's farmland, making it one of the most important food crops in the country. Since rice quality and quantity are so important to economic and food production, it is crucial to diagnose rice infections as soon as possible. In order to identify and categorize diseases in rice plants, this research uses a fundamental photo-analyzing technique, wavelet transform. This technology divides the incoming picture into horizontal, vertical, and diagonal subbands in order to use the Discrete Wavelet Transform (DWT) for multi-resolution research. The horizontal, vertical, and diagonal sub-bands up to two levels of the input image are used to rebuild various textures and colour qualities, which are then translated into features. The data were categorised using an ensemble of linear classifiers and Random Subspace Methodology (RSM). With a classifiers performance level of 95%, the proposed methodology has successfully classified the four main forms of diseases that affect rice plants: rice bacterial blighted, rice brown spots, rice bacterial sheath brown rot, and rice blasting.

Numerous diseases[5], such as bacterial, viral, and fungal disorders, can affect plants. The five primary tomato diseases—tomatillo late blight, septoria spot, bacterial spot, bacterial canker, tomato leaf curl—as well as healthier tomato plants' leaves and stems—are described and discussed. Pictures of healthy and diseased tomato plants were classified using colour, shape, and texture information. The method of feature extraction starts after the classification stage. Segmentation process images are used for the features extracted that are then submitted to the classification model. Ultimately, disorders were classified using these six different categories. In the categorization of six major tomato photo types, an overall accuracy of 97.3 percentage was discovered.

III. METHODOLOGY

A. Proposed System

The suggested program's primary goal is to autonomously monitor the crop. Diseases are found using ML approaches, and they are treated using an automated pesticide sprayer. For automated prediction and diagnosis in coffee plants, the CNN model is employed. The equipment keeps track of several aspects of the soil and environment. A mobile application can be used to access the acquired data from the cloud where it was stored. The system keeps an eye on the coffee plantation as a whole. The tracking uses a variety of sensors to track rain, temperature, soil fertility, and soil richness. Additionally, ML techniques were used to identify the plant's ailment, and automatic pesticide sprayers were used to treat it. A single piece of hardware handled every operation.

B. Components

1) Temperature sensor

The sensor determines the atmospheric temperature by converting the physical feature into an electrical voltage. A temperature controller's output voltage is directly proportional to the temperature being sensed (in degrees Celsius).

2) Soil moisture sensor

Soil moisture sensors have been used to measure the volumetric water content of soil [6]. Since the straight gravimetric dimensions of soil moisture need to be eliminated, dryness and sample weighing is crucial. These sensors make an indirect estimation of the volumetric moisture content based on other soil characteristics.

3) Rain sensor

A rainfall sensor is used on the ground to gauge rainfall. The sensor has modules for controlling and measuring rain. Analog and digital outcomes from the control scheme may be generated. The digital output is used to track rainfall, whereas the analog output is used to both track and identify rainfall. The rain-detecting board has two separate PCB tracks, each 50 mm × 40 mm in size.

4) Solenoid pump

Positive displacement compressors, often known as solenoid pumps, operate by transferring fluid to the point of discharge using a diaphragm and solenoid assembly. The solenoid circuit is composed of an electromagnet and a spring. Upon accepting power, the electromagnetic core of a solenoid forces a diaphragm into the discharging condition.

5) Camera module

It connects straight to the Raspberry Pi and has a 5 MP sensor with a camera module that can capture still photographs and 1080p video. The Raspbian system software, which is suitable for applications like time-lapse photographs, motion sensors, security, and many other things, has recently undergone an update.

C. System Architecture

The model's components included a camera module, pumps, a raspberry pi 4 computer, the cloud, and a smartphone app. Sensors are used to gather a variety of information, including temperature, soil moisture, and rainfall. The gathered information is kept in the cloud and then accessed using a mobile app. The system's main controller is the Raspberry Pi. It analyzes the leaf image captured by the camera module and determines whether or not the coffee plant is afflicted. The CNN model is used to predict coffee sickness. “Miner [8], Rust [9], Phoma [10], and Cercospora [11]” are the diseases that have been identified. The Pesticide Sprayer automatically recognizes the ailment and sprays the proper pesticide. Through a mobile application, the farmer can manage the equipment.

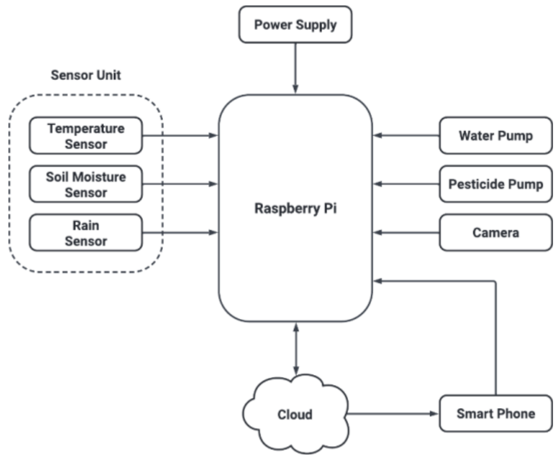

Fig. 1. System Architecture

1) Power supply: The Raspberry Pi must be powered with a power source of at least 5V.

2) Sensor Unit: The sensor unit includes a rain sensor, a soil moisture sensor, and a temperature sensor.

3) Water pumps: All along the system, pumps are employed to supply water to the plants. These are connected to the Raspberry Pi so that the motors operate following Raspberry Pi instructions.

4) *Cloud*: The cloud stores all the information gathered by Raspberry Pi. Using the IP address, one can retrieve the information stored.

5) Mobile Application: A smartphone app created to keep track of the system's performance.

A. Image processing techniques

Convolutional Neural Networks (CNN) are used to develop a computer model that takes unorganized image inputs and converts them into labels that are compatible with matching classification outputs. Multi-layer neural networks, which are the category where these technologies fall, can be designed to predict the traits necessary for categorization. They do various important functions with less pre-processing than standard approaches, which leads to better results.

1) Image preprocessing

The phrase "image preprocessing" is frequently used to describe the initial phases of modifying a raw image. In order to do this, the image must primarily be adjusted, undesired noise must be suppressed, and certain characteristics must be enhanced for further processing. The probing image taken by the camera module linked to the hardware will be used in agricultural image processing to create photographs of the crops.

Once the image is acquired, it is transferred, following the requirements, for image preprocessing. Usually, analysis is done for agricultural data to assess the strength, length, and width of the fruits or crops. If the image obtained by the classifier is inappropriate for preprocessing, image enrichment is utilized to speed up this phase. The extra background image can be cropped for quicker processing of the image as the Disease Detection system evaluates the leaves. These cropped photos are fitted into a 60 X 60 graph, or into blocks that can be used to train the model. The tagged photos of the leaves with dimensions of 60 X 60 are the pre-processing images' final output. As input dimensions for the neural network, these dimensions are a perfect fit. The size of the images can be increased, though, for better results, which necessitates more processing power and training time for the model. Therefore, for this project, the model is trained and predictions are made using 60 X 60 dimensions. The photographs in the database were scaled down to a resolution of 60x60 in order to expedite model performance and make it technically possible. The training procedure often proceeds more rapidly when the inputs or target output are normalized. This is achieved by making the optimization problem's numerical factors stronger. Furthermore, it is made sure that the variable data values used for startup and termination are appropriate. To achieve our goals, we normalized the images using the mean and standard deviation such that all pixel values lie within the same region. It is referred to as the Z-score in the context of machine learning.

2) Image Segmentation

The division of a digital photo into divisions is referred to as image analysis. By simplifying the appearance of the image, image segmentation simplifies analyses. Finding boundaries and entities, like leaves, fruits, and vegetables in this case, in the picture is the primary goal of categorization. The dataset was run through the image analysis method developed by Otsu. RGB colour elements were utilized to construct colour information, the sector props function to generate feature descriptors, and the Gray Level Co-occurrence Matrix (GLCM) to construct texture characteristics. The extracted features were all combined to form a feature extraction module. A set of sections that fully enclose the picture will always result from this.

3) Feature Extraction and Classification

Feature extraction is crucial to the sciences of image processing, deep learning, and pattern classification. In order to start computing, values (features) are first derived from unspecified original information [12]. A database of samples that will be contrasted to the testing dataset generated for assessment serves as the fundamental building block of classifications in image recognition. Convolutional, activating, pooled, and fully connected information layers are present in the CNN model that we used in our approach. Every frame's construction elements are convolutional, activated, and maximum pooling layers. This technology's architecture contains three of these blocks, fully connected layers, and softmax triggering. Convolutional and pooling layers are used to extract features in classification, while fully connected layers are used for feature extraction. Activation layers are employed to introduce nonlinearity into the network. Convolutional operation is used by the convolutional layer to extract features. As depth increases, the complexity of the recovered characteristics also advances. The size of the filter is fixed at 5×5, even though the quantity of filtering progressively rises as we move from one block to the next. 20 filters are included in the first convolutional block; 50 and 80 are included in the second and third blocks, correspondingly. The use of pooling layers for each of the blocks has reduced the dimensions of the feature maps, which must be compensated for by adding more filters. Additionally, zero padding of the feature maps maintain the size of the image after the convolutional process is performed. The max pooling reduces the dimensionality of feature maps, speeds up learning, and reduces the model's sensitivity to minute input adjustments. A 2x2 kernel is needed for maximum pooling. ReLU activation layers are used for each of the frames to generate non-linearity. To avoid overfitting the train set, the Dropout regularized process has also been used with maintained probabilities of 0.5. Dropout regularization minimizes the model's variability and streamlines the networks, reducing overfitting. By periodically deleting synapses from the networks, it accomplishes this throughout each training cycle. With a categorized cross-entropy loss function, a learning rate of 0.0001, and 32 batches, the system is designed using the Adam optimizer. Furthermore, the classifier is constructed with 3 distinct max-pooling layer configurations, 3 distinct numbers of epochs, and three different train, validation, and test ratios. The classification block is composed of two sets of fully connected neural different networks with 500 and 10 neurons, respectively, in each layer. A softmax activation function is used after the second dense layer to calculate the probability scores for the ten classes.

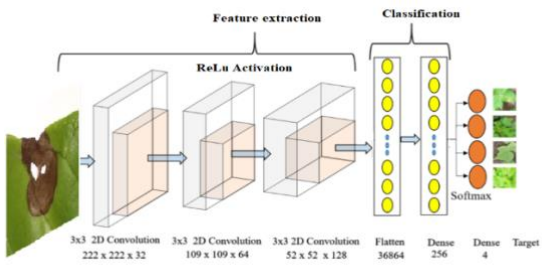

Fig. 2. Proposed CNN Architecture

IV. RESULT

Because of the drip irrigation system and the leaf disease detection technology, farmers can properly monitor their fields and reduce their effort. For changes in the climate, soil moisture, and plant diseases, precise tracking and forecasts are carried out. Additionally, the connected pesticide sprayer showed superior results against plant disease.

All tests are done on this platform, which has the following technical specifications: "a Google Colab Pro with Python 3.6.9 as the coding environment, an Intel (R) core i3-4005 CPU operating at 1.70 GHz, 8.00 GB RAM and 64-bit Windows 10".

1) Performance Parameters

We evaluated the effectiveness of the model using metrics including accuracy, precision, recall, and F1 score. These measurements are scaled on a scale of 0 to 1, with 0 being the worst and 1 being the best.

a) Accuracy: The percentage of accurately predicted images (TP+TN) to all available predictions (TP+TN+FP+FN) is how accurate a prediction is. Following is the equation for the evaluation of accuracy.

$$\text{Accuracy}=(TP+TN)(TP+TN+FP+FN) \quad (1)$$

b) Precision: The ratio of occurrences that were accurately predicted (TP) to all positive predictions (TP+FP) is used to determine how accurate a forecast is. The following equation can be used to calculate an estimation of it (2).

$$\text{Precision}=(TP)(TP+FP) \quad (2)$$

c) Recall: The proportion of occurrences that were correctly predicted (TP) to all actual cases (TP+FN) is known as recall, also known as sensitivity. Equation (3) can be used to calculate Recall.

$$\text{Recall}=(TP)(TP+FN) \quad (3)$$

d) F1 Score: A assessment of recall and precision is combined into one score by the F1-score. The Harmonic mean of recall and precision is a mathematical model that is determined using Equ (4).

$$\text{Recall}=(TP)(TP+12(FP+FN)) \quad (4)$$

Where TP: True Positives FP: False Positives TN: True Negatives FN: False Negatives.

Using 30 training epochs, the sickness recognition results show a strong validation accuracy of 94.8% and a high training accuracy of 99.3%. It has been possible to attain an average validation accuracy of 94%. This provides an effective measure of how accurately the deep learning model categorized the items. The charts of training and testing set correctness and loss against the epochs in Fig. offer a mechanism for visualizing and providing a gauge of the rate of model converging. As can be seen, the model has stabilized after approximately 20 epochs, and the metrics do not suggest a marked enhancement in the last 10 epochs. According to the results, the model can detect leaf diseases with the least environmental impact based on the datasets.

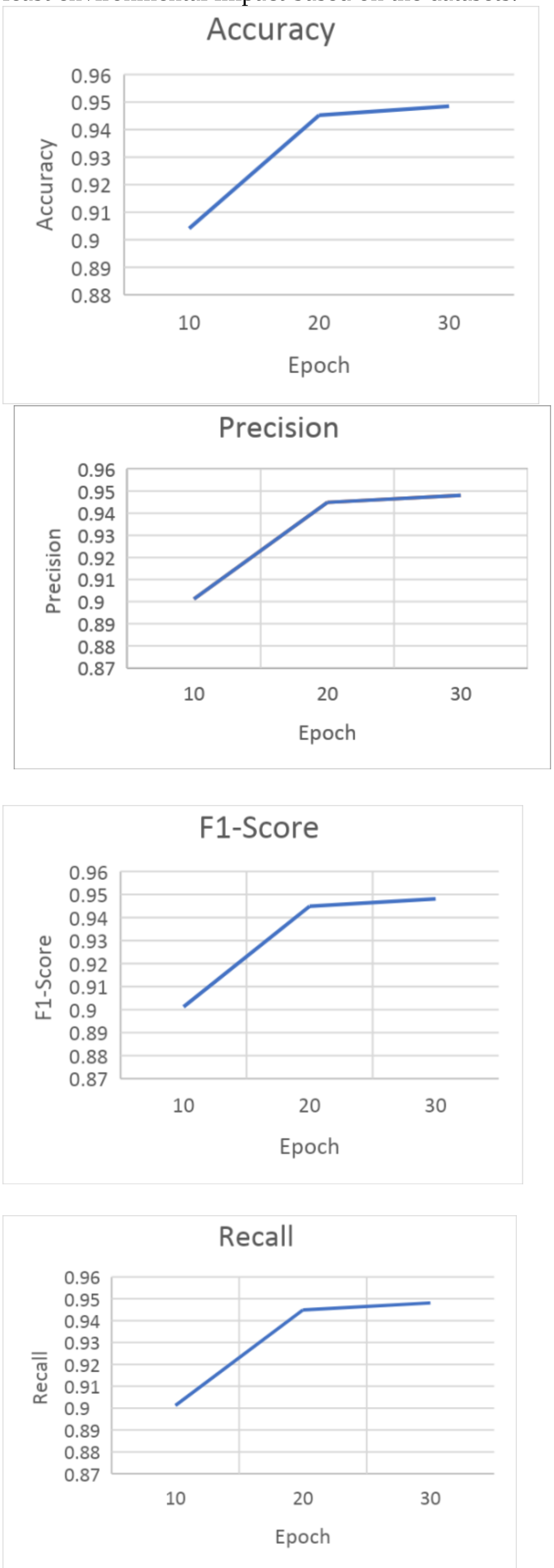

Fig. 3. Performance metrics of the proposed model

V. CONCLUSION

Crop disease forecasting is done to increase farming productivity and decrease labour intensiveness. The deployment of the suggested strategy is demonstrated in real-time, and it entails the uploading of crop-camera photos to the cloud for storage and later analysis, as well as the use of soil moisture values to automate the water pump used for irrigation. The tool can also be used to track irrigation operations and assist with inspecting crop photos for any signs of probable diseases. To the advantage of farmers worldwide, a CNN model is utilized for automatic plant disease recognition and characterization, crop assessment, and diagnostics at an early point. The disease detection task outcomes showed a maximum validation accuracy of 94.8% over 30 training epochs and a maximal training accuracy of 99.3%. There has been a 94% average accuracy for validation.

REFERENCES

1. V. Udutalapally, S. P. Mohanty, V. Pallagani and V. Khandelwal, "Crop: A Novel Device for Sustainable Automatic Disease Prediction, Crop Selection, and Irrigation in Internet-of-Agro-Things for Smart Agriculture," in IEEE Sensors Journal, vol. 21, no. 16, pp. 17525-17538, 15 Aug.15, 2021.
2. B. S. Kusumo, A. Heryana, O. Mahendra and H. F. Pardede, "Machine Learning-based for Automatic Detection of Corn-Plant Diseases Using Image Processing," 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA), 2018, pp. 93-97.
3. S. D.M., Akhilesh, R. M.G., S. A. Kumar and P. C., "Disease Detection in Pomegranate using Image Processing," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020, pp. 994-999.
4. S. M. Taohidul Islam and B. Mazumder, "Wavelet-Based Feature Extraction for Rice Plant Disease Detection and Classification," 2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 2019, pp. 53-56.
5. H. Sabrol and K. Satish, "Tomato plant disease classification in digital images using classification tree," 2016 International Conference on Communication and Signal Processing (ICCSP), 2016, pp. 1242-1246.
6. S. Puengsungwan, "IoT-based Soil Moisture Sensor for Smart Farming," 2020 International Conference on Power, Energy and Innovations (ICPEI), 2020, pp. 221-224.
7. M. Sankar, D. N. Mudgal, T. varsharani jagdish, N. w. Geetanjali Laxman and M. Mahesh Jalinder, "Green Leaf Disease Detection Using Raspberry pi," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019, pp. 1-6.
8. P, L. A. Pinto, L. Mary, M. P and S. Dass, "The Real-Time Mobile Application for Identification of Diseases in Coffee Leaves using the CNN Model," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp. 1694- 1700.
9. P. Marcos, N. L. Silva Rodovalho and A. R. Backes, "Coffee Leaf Rust Detection Using Convolutional Neural Network," 2019 XV Workshop de Visão Computacional (WVC), 2019, pp. 38-42.
10. L. Dutta and A. K. Rana, "Disease Detection Using Transfer Learning In Coffee Plants," 2021 2nd Global Conference for Advancement in Technology (GCAT), 2021, pp. 1-4.
11. M. Kumar, P. Gupta, P. Madhav and Sachin, "Disease Detection in Coffee Plants Using Convolutional Neural Network," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 755-760.
12. S. Donesh and U. Piumi Ishanka, "Plant Leaf Recognition: Comparing Contour-Based and Region-Based Feature Extraction," 2020 2nd International Conference on Advancements in Computing (ICAC), 2020, pp. 369-373.
13. L. Mohan, J. Pant, P. Suyal and A. Kumar, "Support Vector Machine Accuracy Improvement with Classification," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), 2020, pp. 477-481.
14. S. Gayathri, D. C. J. W. Wise, P. B. Shamini and N. Muthukumaran, "Image Analysis and Detection of Tea Leaf Disease using Deep Learning," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 398-403