

Received March 17, 2022, accepted April 12, 2022, date of publication April 21, 2022, date of current version May 2, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3169147

Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction—Blister Blight for Tea Plant

ZHIYAN LIU¹, RAB NAWAZ BASHIR^{1,2}, SALMAN IQBAL², MALIK MUHAMMAD ALI SHAHID^{1,2}, MUHAMMAD TAUSIF^{2,3}, AND QASIM UMER^{1,2}

¹Anxi Tea College, Fujian Agriculture and Forestry University, Cangshan, Fuzhou, Fujian 350002, China

²Department of Computer Science, COMSATS University Islamabad, Vehari 61100, Pakistan

³Department of Computer Science, Superior University, Lahore 54000, Pakistan

Corresponding author: Rab Nawaz Bashir (rabnawaz@ciitvehari.edu.pk)

ABSTRACT Crop plant diseases are a significant threat to productivity and sustainable development in agriculture. Early prediction of disease attacks is useful for the effective control of the disease by taking proactive actions against their attacks. Modern Information and Communication Technologies (ICTs) have a predominant role in Precision Agriculture (PA) applications to support sustainable developments. There is an immense need for solutions for the early prediction of the disease attack for proactive control against the plant disease attack. The present solution of disease detection using the computer vision approach can only detect the existence of the disease once the disease has already appeared. This study aims to propose a Machine Learning (ML) approach for the early prediction of the probability of disease attack based on Internet of Things (IoT) directly sensed crop field environmental conditions. Plant disease life cycles are strongly correlated with environmental conditions. The crop field environmental conditions are used to predict the occurrence of plant diseases. The Multiple Linear Regression (MLR) is applied as the ML model due to the existence of a linear relationship between disease attack and environmental conditions. Internet of Things (IoT) based crop field environmental conditions help to accurately predict the occurrence of plant diseases using the ML approach. The proposed model is implemented for the prediction of blister blight (*Exobasidium vexans*) for tea (*Camellia sinensis*) plant to check the effectiveness of the proposed solution. The implementation of the proposed model from 2015 to 2019 reveals that the accuracy of prediction of occurrence of the disease reached up to 91% in 2019.

INDEX TERMS Machine learning (ML), precision agriculture (PA), multiple linear regression (MLR), tea (*Camellia sinensis*), blister blight (*Exobasidium vexans*).

I. INTRODUCTION

Global climatic changes have adverse impacts on agriculture productivity. An increase in disease attacks is one of the major impacts of climatic change. Climatic changes have significant impacts on crop production, and tea plants are more affected by climatic changes [1]. Crop disease attacks cause severe loss to agriculture productivity [2]. The use of chemicals in agriculture has further diversified the issue. Early prediction of disease attacks is very important to take proactive, efficient,

and effective control against the disease [3] and to support sustainable developments.

Tea is grown in hot and humid regions as a cash crop that is used to make the world's most used beverage [4]. Fungal diseases of tea plants cause severe losses to the tea production and quality of the produce [5]. Blister blight (*Exobasidium vexans*) is the most common and destructive disease of the tea plant that seriously affects tea production across the world. Blister blight (*Exobasidium vexans*) can reduce the yield of tea by up to 40% [4].

Disease attack on crops is a serious hazard to sustainable developments and productivity in agriculture [6]. Early prediction and detection of disease attacks are very important to

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues¹.

preserve crop yield [7]. Crop diseases cause serious economic losses to farmers. Many computer vision approaches have been developed over the year for the automatic detection and identification of disease attacks. One major problem of these approaches is that they detect disease based on apparent characteristics of the disease attack. At this stage, the disease is usually incurable or has already caused serious damage to the crops. The losses due to disease at this stage are untreatable. There is a need for solutions that can predict the occurrence of the disease attack before it causes losses.

There is a need for a solution that can predict the occurrence of the disease before it damages the plants to take effective control against the disease attack. There is a strong correlation between the disease attack and environmental conditions. Therefore, the environmental conditions can effectively be used to predict the occurrence of crop plant diseases.

The study proposed a model of plant disease predictions based on directly sensed crop field environmental conditions. The Internet of Things (IoT), based on directly sensed crop field conditions, can help to accurately predict the occurrence of the plant disease based on real-time crop field environmental conditions. The most important environmental conditions in this regard are temperature, humidity, and rainfall. The directly sensed environmental conditions are fed into the ML model to determine the probability of occurrence of disease according to crop field directly sensed environmental conditions. The proposed solution is implemented for the prediction of blister blight of tea plants.

Early prediction of disease attacks is also useful for sustainable developments in agriculture by judicious use of the pesticide's applications in agriculture. Early predictions of disease attacks are helpful for the effective control strategy of the disease. The study proposed the prediction of blister blight (*Exobasidium vexans*) disease attacks on tea plants based on temperature, humidity, and rainfall conditions.

The rest of the study is organized into a literature review, follow by the contribution of the study. In the next section, the material and method for the proposed solution are described. In the end, the evaluation of the proposed solution is given.

A. BACKGROUNDS

Tea is the most used drink after water across the world [8], [9]. Tea is the most used beverage across the world [5]. According to one estimate, more than three billion cups of tea are used in one day. The tea production growth rate has been 4.4% over the last decade [9]. Tea is the major source of income for tea-producing countries, with thirteen million people, are involved in tea production [9]

Tea (*Camellia sinensis*) is the plantation crop of Africa, Asia, and South America regions, grown for commercial purposes. Tea was discovered around 2732 BCE in China. The cultivation of tea begins in the first century BCE[9]. Nowadays, tea is produced in more than 60 countries, covering around 5 million hectares[9].

The tea crop required specific conditions for its successful growth. Tea crops grow well at 10-30C, with annual precipitation of 1250 mm[10]. China, Sri Lanka, India, Kenya, and Vietnam are the major tea-producing countries. China is the leading tea producing country with a production of 2.5 Metric Tons of tea annually [9], [11], [12]. China, India, Sri Lanka, and Kenya are the major tea-producing countries, with 86% of the total world tea production [8] Fig. 1 shows a typical tea garden, and Fig. 2 shows the level of tea production in different regions of the world.



FIGURE 1. Tea plantation.



FIGURE 2. Global tea production [10].

Fresh leaves from tea plants are harvested and dried for brewing the tea. In Fig. 3, the shoot of the tea plant is shown that it is harvested for commercial purposes.

Tea (*Camellia sinensis*) production is affected by different pests and disease attacks. Fungal diseases seriously affect the yield and quality of the tea. Blister blight (*Exobasidium vexans*) is the major fungal disease of the tea that affects the leaves, shoots, and buds, as shown in Fig. 4. Blister blight not only affects the production of the tea but also damages the plants to such an extent that survival of the teal plant becomes impossible [4], [5].

Blister blight is caused by *Exobasidium vexans* pathogen. This disease is present in almost every tea-growing area except in America and Africa. Tea is the sole host for the *E. vexans* pathogen. Moderate temperature, from 15 to 25°C,

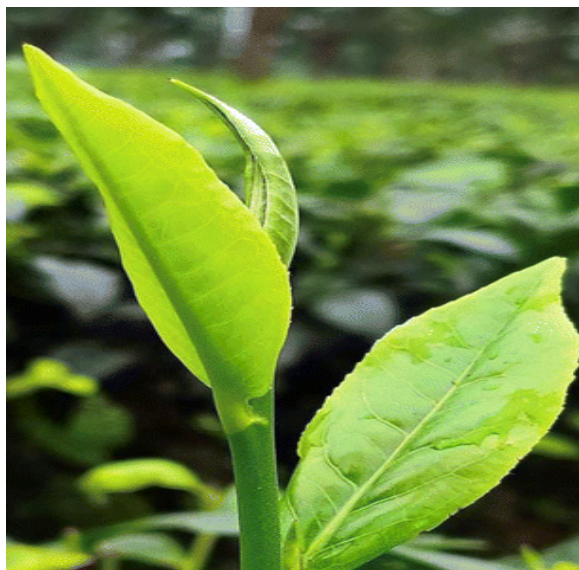


FIGURE 3. Tea producing countries [9].



FIGURE 4. Tea (*Camellia sinensis*) disease affected cotton plant A. Front Side of Leaves B. The backside of leaves C. Twigs of tea affected with blister blight.

and high humidity above 80% are the favorable environmental condition for the development of blister blight by *Exobasidium vexans* [1], [10]. Wet and cool weather favors the development of blister blight disease on tea plants [1], [4]. The *E. vexans* affect the tea plant at every stage and almost every part of the plant. Lemon color spots at leaves are the first symptoms of the disease in the tea plant. Blister blight (*E. vexans*) can affect 20-50% of the yield of tea [4]. Apart from crop production, blister blight also affects the quality of the tea in terms of caffeine production [5].

China is the major tea-producing country, and the production of tea has increased tremendously during the last decades. The increase in demand for tea across the world is due to a tremendous increase in population across the world.

For better control of the disease, early prediction can be very helpful in effectively managing the disease and taking proactive control measures. For better disease management, farmers must be aware of the probability of disease attack in advance. There is a strong correlation between environmental

conditions and plant disease attacks. High humidity coupled with moderate temperature favors the development of blister blight disease on tea plants. Humid conditions in summer are ideal for *E. Vexans* to grow and reproduce. Moderate temperature with high humidity favors the development of the disease life cycle [13]. The correlation between the environmental conditions and disease life cycle can be used to effectively predict the occurrence of the disease in advance before the occurrence of the disease. Early warning of the disease attack can help the farmers effectively deal with the disease. Many ML-based approaches are used to detect and predict the occurrence of the diseases that are reviewed in the next section.

II. LITERATURE REVIEW

Farber *et al.* review different spectroscopic techniques like infrared, reflectance, and Ramen for plant disease identifications. Each of these techniques is discussed with its advantages and disadvantages [7]. Jawade *et al.* proposed weather-based mango plant disease prediction using the Random Forest Machine learning algorithm. The proposed technique is very accurate in mango disease predictions [14]. Chen *et al.* proposed the Internet of Things (IoT) and machine learning-assisted rice blast disease detection [15]. The rice field imagery data is converted to hyperspectral data for a machine learning model of rice blast disease prediction.

Vishnoi *et al.* reviewed different techniques of disease detection from leaf images by using computer vision approaches. Different feature extraction techniques and their performance is compared [2]. Devaraj *et al.* proposed an image processing technique for pre-processing, segmentation, classification, and segmentation for plant disease classification [16]. Khamparia *et al.* [17] proposed a hybrid approach by autoencoders and a Convolution Neural Network (CNN) model for the identification of leaf diseases of crop plants. The proposed techniques gain 97% accuracy in disease identification.

Orozco-Fuentes *et al.* proposed a mathematical and computational model of disease forecasting by observing the pathogen and host relationship to the environment [18]. Sambasivam and Geoffrey Duncan Opiyo proposed Cassava (*Manihot esculenta* Crantz) disease detection and proposed a technique to deal with the imbalance in class identification in Cassava Mosaic Disease (CMD) and CBB (Cassava Brown Streak Virus Disease) classes [19].

Truong *et al.* proposed fungal disease prediction based on prevailing environmental conditions [20]. Raghavendra *et al.* identified the relationship between the disease and weather conditions on tea plants [3]. Liu *et al.* proposed disease predictions based on environmental conditions using a neural network. [21]. Liu and Wei Wang proposed tomato disease detection based on the environment using deep learning techniques from images of the tomato plants [22]. Shang *et al.* proposed an artificial neural network and Genetic Algorithm (GA) to predict insect attacks [23].

Yakkundimath *et al.* proposed a plant health detection system using real-time monitoring of the plant's characteristics [24]. Ramesh and Vydeki recommended IoT-based plant disease detection by real-time monitoring [25]. Arsenovic *et al.* discussed the limitations of the existing models of disease detection. The study proposed two-stage neural network architecture for plant disease detections in a real-time environment, with 93.67% accuracy [26]. Nagaraju and Chawla reviewed different CNNs techniques of plant disease detection with their performance comparison [27]. Yang and Guo reviewed the machine learning approaches for plant disease classification and plant resistance genes discovery [28]. Liu and Wang reviewed and analyzed different techniques of deep learning for plant disease detection [29]. Francis *et al.* [30] proposed a CNN-based technique for disease detection of apple and tomato leaf images.

Kim *et al.* proposed strawberry disease prediction by FAAS (Farm as a Service) approach for managing the data, devices, and models in an integrated manner [31]. The proposed integrated agriculture specialized FaaS system ensures the implementation of the system in a smooth manner.

The Materne and Inoue recommended IoT-based environment monitoring for predictions of disease populations to deal with the issue of climate change impacts on the crop [32].

Araby *et al.* recommended an IoT-based disease warning system by direct observations of the crop field and the use of machine learning to generate an early warning system [33]. Truong *et al.*, proposed fungal disease detection using IoT and machine learning capabilities. IoT is used to directly observe the crop field's environmental conditions. The directly observed environmental conditions are used to predict the occurrence of fungal disease [20]. Syarif *et al.* find the regression model and correlation between the disease population and weather conditions in corn crops. The proposed correlation identification helps to identify the occurrence of disease at a particular time [34]. Cohen proposed MedCila named Decision Support System (DSS) for the control of medflies in citrus orchids [35]. The biological, mathematical, and statistical models were developed for the identification of medflies. The proposed approach was developed on a rule-based decision tree. The performance of the proposed approach was high in medflies attack detection in citrus orchids when compared to the expert judgment for medflies attack detection.

Shashank *et al.*, discuss the way to improve precision agriculture practices by effective detection of plant disease, with a major focus on leaf and plant diseases by using image processing. The study analyzes the Otsu's and K-means clustering for segmentation and feature extraction of images for disease classification [36]. Pragya *et al.* proposed tomato leaf disease and defects identification by a proposed solution named SENet, with CNNs. SENet, with CNN's based hybrid approach for tomato leaf disease detection, is proposed. The major objective of the proposed approach is to efficiently utilize the computing resources for the identification of plant diseases [37].

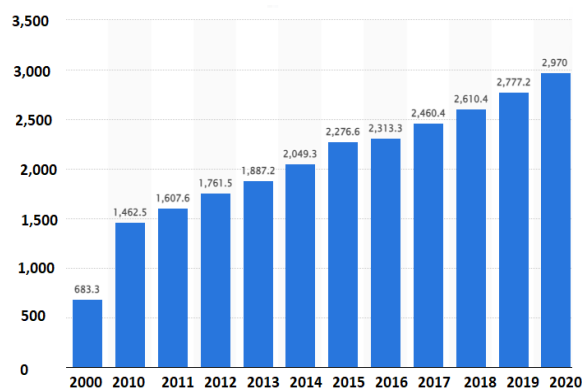


FIGURE 5. The volume of tea production in China (1,000 metric tons) [11].

Many solutions were proposed for disease detection using the computer vision approach from the apparent characteristics of the disease symptoms on plants. These approaches are applicable once the disease has already been set and causes significant losses to the crop. There is a need for a solution to predict the occurrence of disease to take proactive actions. There is a strong correlation between environmental conditions and the disease life cycle. The prediction of occurrence of crop disease based on temperature, humidity, and rainfall is proposed that has not previously been targeted. The unique feature of the proposed solution is that it can forecast the probability of occurrence of tea (*Camellia sinensis*) disease attack and extend to the prediction of any disease that has a strong correlation with environmental conditions.

A. CONTRIBUTION OF THE STUDY

The study proposed a solution for early prediction of the occurrence of plant disease with the help of directly sensed crop field environmental conditions. Machine Learning (ML) model is proposed and implemented to predict the occurrence of blister blight disease in the tea plants in China. The proposed solution is equally applicable for any disease predictions that have a strong correlation with environmental conditions.

III. MATERIAL AND METHODS

In this section, the method of environmental data collection and flow chart for crop disease prediction, the environmental data, and the machine learning model is described.

A. AREA OF STUDY

Major tea-producing is of the world is marked in Fig. 6. To implement the proposed solution to predict the occurrence of the disease Hainan Island is chosen.

B. PROTOTYPE FOR ENVIRONMENTAL DATA SENSING

The crop field environmental data is captured by developing an IoT-based hardware prototype. The Arduino platform is used with DHT-22 (Digital Temperature and Humidity) and a rain sensor to directly sense the temperature, humidity, and

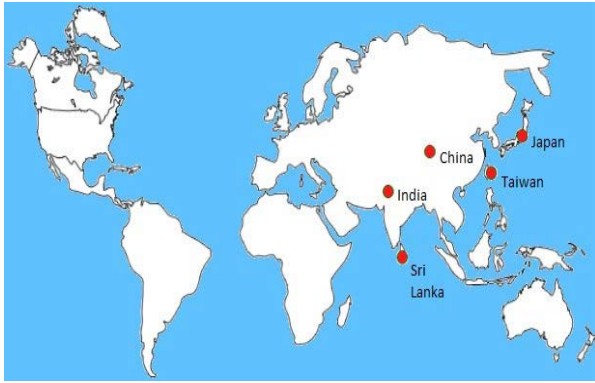


FIGURE 6. Tea production area [12].

rainfall from the crop field. The hardware prototype is shown in Fig. 7, with the DHT-22 sensor. The environmental data is sent to the server for further processing according to the proposed solution.



FIGURE 7. Hardware prototype.

C. FLOW CHART OF THE PROPOSED DISEASE PREDICTION

The proposed solution determines the probability of blister blight disease of tea plants from prevailing temperature, humidity, and rainfall conditions. Daily temperature and humidity are used to determine the average monthly temperature and average monthly humidity. The maximum rainfall of any day in the month is taken as the maximum rainfall of the month. The average monthly temperature, humidity, and maximum monthly rainfall are used to determine the probability of the occurrence of a blister blight disease attack on tea plants. The predictions made by the model are validated by field observations. If the field observations show the disease attack more than 15% of the plants, that is the Economic Threshold Level (ETL) of blister blight on tea plants. The predictions are compared against the field observations and validated. The validations of the predictions are provided as feedback to the machine learning model to improve the performance over time. The complete flow of information is shown in Fig. 8 by a flow chart.

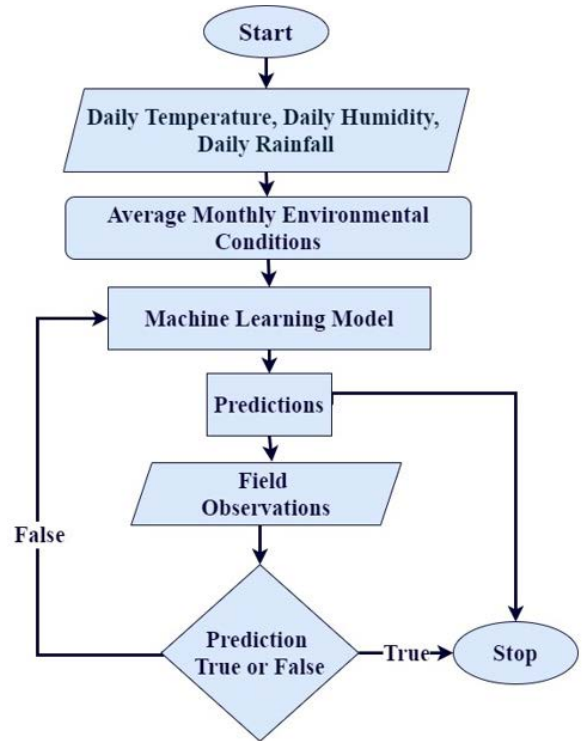


FIGURE 8. Disease prediction flow chart.

D. ENVIRONMENTAL CONDITIONS

This section presents the environmental conditions that are important for the development of blister blight disease attacks on tea plants. The blister blight disease attack on tea plants is affected by temperature coupled with humidity. The growth of the blister blight disease attack on tea plants is severe with moderate temperature and high humidity. The rainfall reduces the disease population; therefore, it is also used in the prediction model.

For monthly predictions, average monthly temperature (T_{avg}), average monthly humidity (H_{avg}), and maximum monthly rainfall (R_{max}) are used. The environmental data is directly observed from the crop field using the prototype described in the previous section. The environmental data is given in the next section.

1) TEMPERATURE

Temperature is highly related to the growth of diseases. The average monthly Temperature (T_{avg}) is obtained from the daily maximum temperature (T_{xi}) by Equation 1.

$$T_{avg} = \frac{\sum_{i=1}^n T_{xi}}{n} \tag{1}$$

where “ T_{avg} ” is the average monthly temperature and “ T_{xi} ” is the daily maximum temperature, and ‘ n ’ is the number of days in a month. The temperature in the selected area is observed from May to September for selected years. The average monthly temperature (T_{avg}) and daily maximum temperature (T_{xi}) for selected months from the

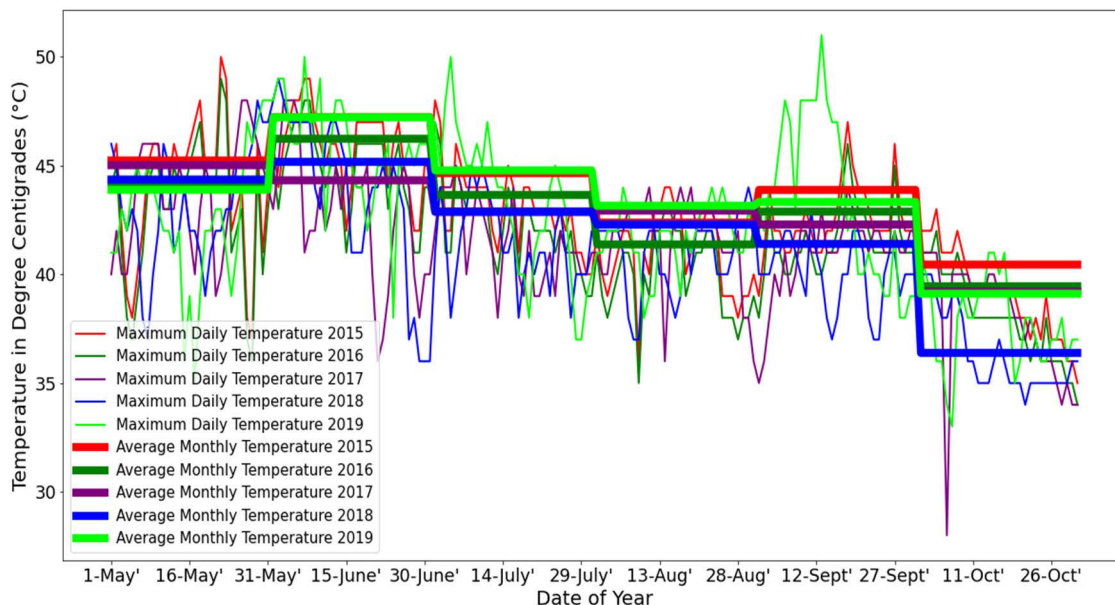


FIGURE 9. Daily maximum temperature and monthly average temperature from 2015 to 2019.

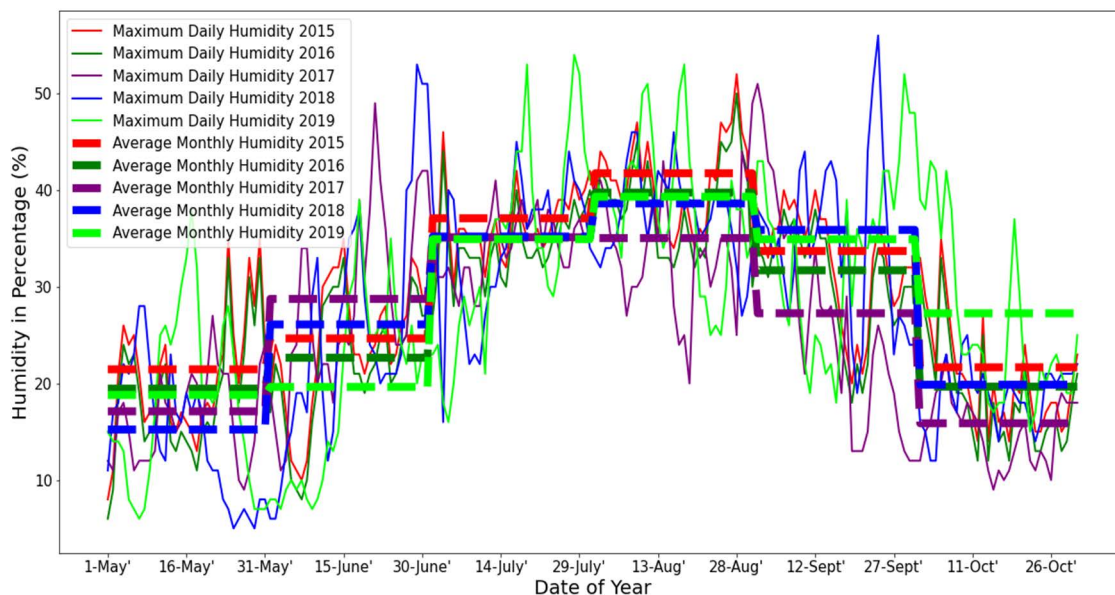


FIGURE 10. Daily maximum humidity and monthly average humidity from 2015 to 2019.

year 2015 to 2019 are shown in Fig. 9. The temperature in the selected area is mostly around 30°C during the tea plants season, which is favorable for the development of the disease.

2) HUMIDITY

Humidity is the percentage of moisture in the air. The daily maximum humidity (H_{xi}) and average monthly humidity (H_{avg}) for the years 2015 to 2019 are shown in Fig. 10. The humidity in July, August, and September are high as

compared to other selected months, which makes these more favorable for the development of blister blight attacks on tea plants. The humidity from July to September is 40 to 70% in the selected area. The average monthly humidity is determined by Equation 2.

$$H_{avg} = \frac{\sum_{i=1}^n H_{xi}}{n} \tag{2}$$

where “ H_{avg} ” is the average monthly humidity, and “ H_{xi} ” is the daily maximum humidity.

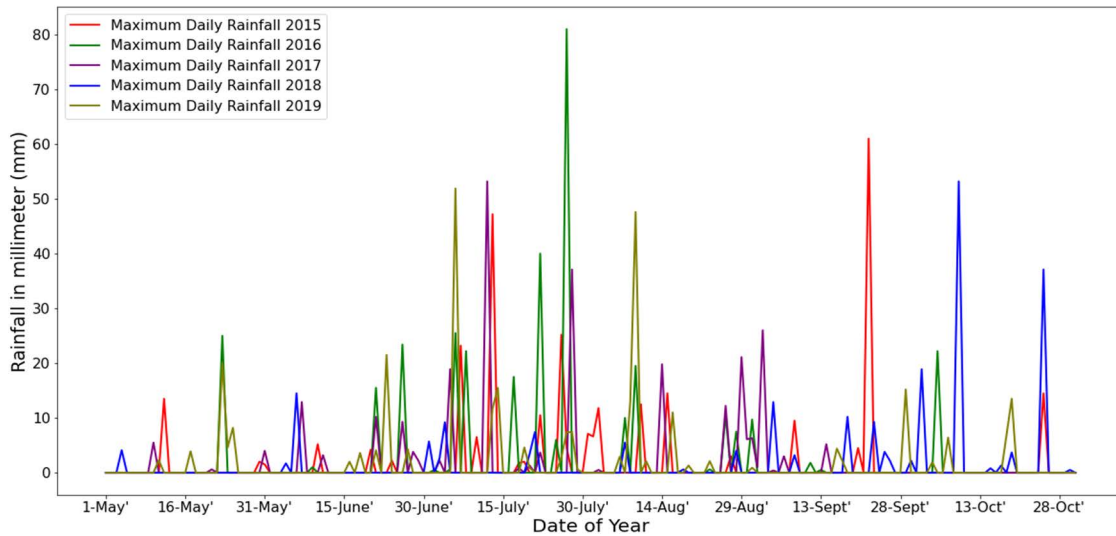


FIGURE 11. Daily maximum rainfall from 2015 to 2019.

3) RAINFALL

The maximum monthly rainfall is the maximum rainfall on any day for a particular month and is given by Equation 3.

$$R_{max} = Max_{R_{xi}} \tag{3}$$

where “*Rmax*” is the maximum monthly rainfall, “*Rxi*” is the maximum daily rainfall. The daily rainfall for the selected month is shown in Fig. 11.

E. DISEASE INTENSITY OBSERVATIONS

The proposed model is implemented and validated by direct field observations to check the disease intensity. Ten acres of tea plants crop are selected, and if 15% of plants in each acre are affected, then the disease attack is considered above Economic Threshold Level (ETL).

The average of the blister blight disease attack from each of the ten-acre is taken. The population of the blister blight-affected tea plants from the year 2015 to the year 2019 is shown in Fig. 12. The intensity of the plant disease is strongly correlated with temperature and humidity.

F. MACHINE LEARNING MODEL

Due to the existence of a linear relationship, the multiple regression machine learning model is applied. The Multiple Linear Regression (MLR) model determines the dependent variable from one or more independent variables. In case of a problem on hand, the probability of occurrence of a tea (*Camellia sinensis*) disease attack is determined by the prevailing temperature, humidity, and rainfall. For the ‘n’ number of the dataset, the relationship between the dependent variable ‘y’ and the regressor vector is expressed by Equation 4.

$$y_i = \beta_0 + \beta_1 i + \dots + \beta_p x_{ip} + \epsilon_i \tag{4}$$

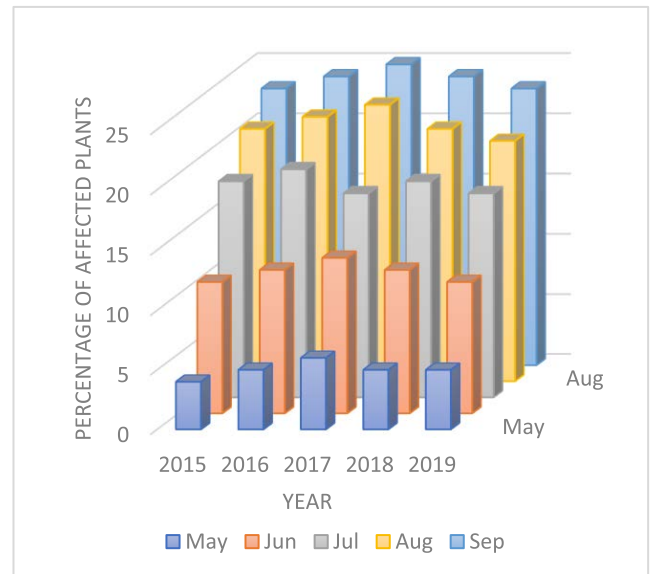


FIGURE 12. Distribution of blister blight disease on tea plants in experiment area from 2015 to 2019.

where $i = 1, \dots, n$. The Sciket Learn python library is used for implementation purposes. The dataset is partitioned into 80:20 for training and testing of the Machine Learning (ML) model.

IV. EVALUATIONS

Evaluations are performed based on.

1. Performance of the machine learning model
2. Accuracy of the predictions by field observations.

A. MACHINE LEARNING MODEL PERFORMANCE

In this section, different statistics are given to implement the machine learning model. In Table 1, the correlation between

TABLE 1. Correlation between environmental conditions and disease intensity.

	Temperature	Humidity	Rainfall
Disease Intensity	0.89	0.81	-0.50

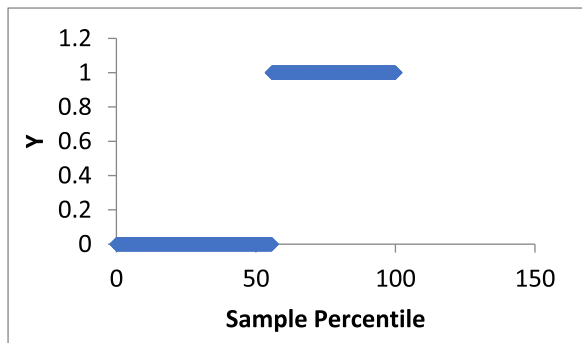


FIGURE 13. Probability distribution of predictions.

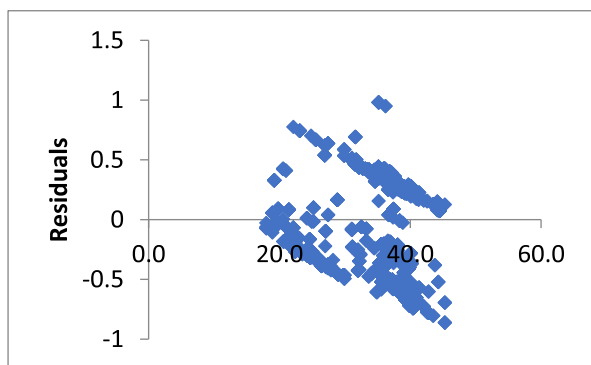


FIGURE 14. Residual plot for temperature.

environmental conditions and blister blight disease development is given. The development of blister blight disease is highly positively correlated with temperature and humidity while negatively correlated with rainfall. Due to the existence of a correlation between the predictor and response variable, the regression line model is used for implementation purposes.

Fig. 13 shows the distribution of “Yes” and “No” probabilities in the data set. The probabilities of prediction of “Yes” and “No” above the ETL are evenly distributed in the dataset. “Yes” is mapped with ‘1’ that is distributed to 50%, and “No” is mapped with “0” that is also distributed to 50% in the dataset. Fig. 14. shows the residuals values for temperature-based prediction of blister blight (*Exobasidium vexans*) disease of tea (*Camellia sinensis*), which are distributed evenly around the averages.

Fig. 15 shows the residuals values for humidity-based prediction of blister blight disease, which are also distributed evenly around the average.

FIGURE 16 shows the residuals values for rainfall-based prediction of blister blight (*Exobasidium vexans*) disease of

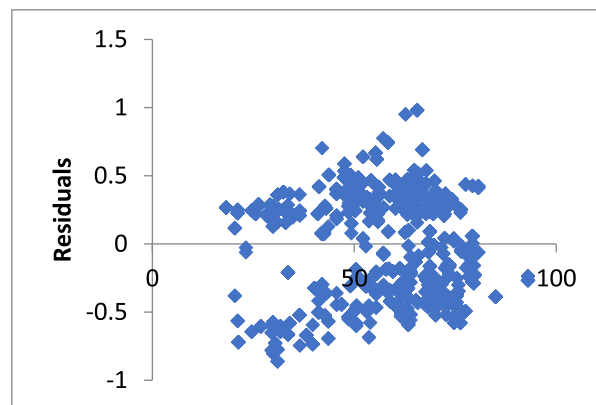


FIGURE 15. Residual plot for humidity

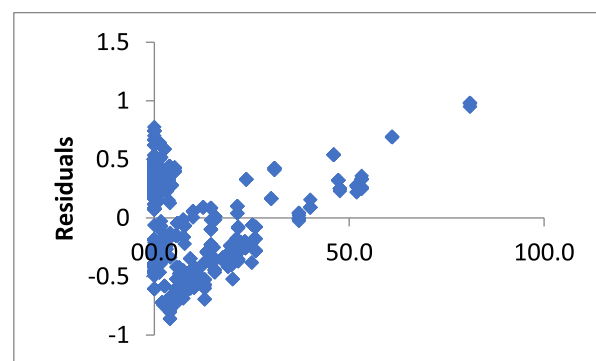


FIGURE 16. Residual plot for rainfall variable.

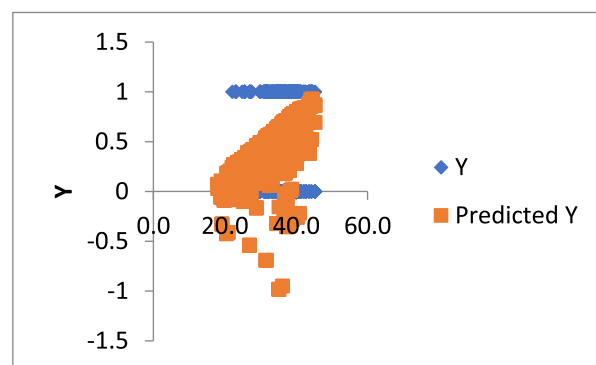


FIGURE 17. Temperature based Regression line fit for disease prediction.

tea (*Camellia sinensis*), which are also distributed evenly around the average.

Temperature, humidity, and rainfall are independent variables, and the occurrence of the disease is the dependent variable. FIGURE 17 shows the best-fitted regression line for temperature variables to make the temperature-based prediction. Figure 18 shows the best-fitted regression line for the humidity variable to make the humidity-based prediction. FIGURE 19, the best-fitted regression line for rainfall variable to make the rainfall-based prediction. These regression line models are used to predict the occurrence of disease.

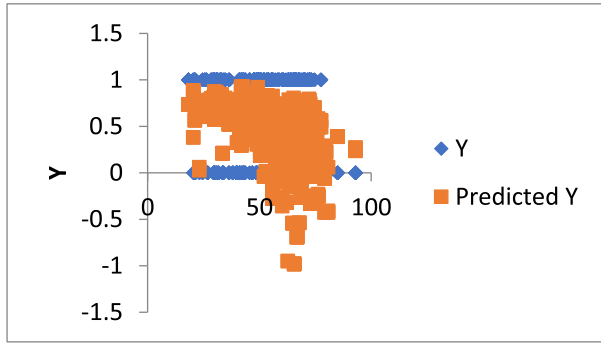


FIGURE 18. Humidity-based regression line fit for disease prediction.

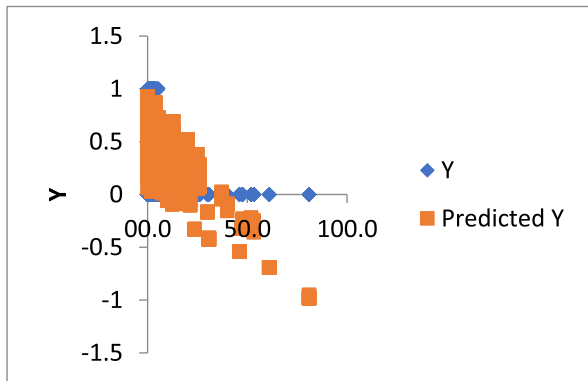


FIGURE 19. Rainfall regression line fit for prediction of the disease population.

TABLE 2. Regression analysis.

Regression Statistics	
Multiple R	0.64
R Square	0.40
Adjusted R Square	0.40
Standard Error	0.38
Observations	1540

TABLE 2 shows the regression statistics for all the variables. The values of multiple R, R Square, adjusted R Square, standard error, and the total number of observations are given in TABLE 2. The value of multiple regression is 0.63, R-Square is 0.4, adjusted R-square is 0.40, and the standard error is 0.38 with 1540 observations. These observations showed the accuracy of the ML model to predict the probability of the disease from a given set of environmental conditions.

Table 3 presents the test statistics for different errors. The errors in making predictions are very low. That makes the model suitable for the prediction.

B. FIELD VALIDATIONS

The predictions by the ML model are validated by field observations. TABLE 4 summarizes the accurate predictions from the Year 2015 to 2019. Accuracy of the prediction improves

TABLE 3. Standard errors.

	Co-efficient	Standard Error	P-value
Intercept	-0.530	0.092	8.81E-09
X Variable 1	0.032	0.002	5.27E-60
X Variable 2	0.001	0.001	0.245589
X Variable 3	-0.020	0.001	1.7E-112

TABLE 4. Year-wise prediction accuracy.

Year	Correct Prediction
2015	8
2016	9
2017	10
2018	10
2019	11

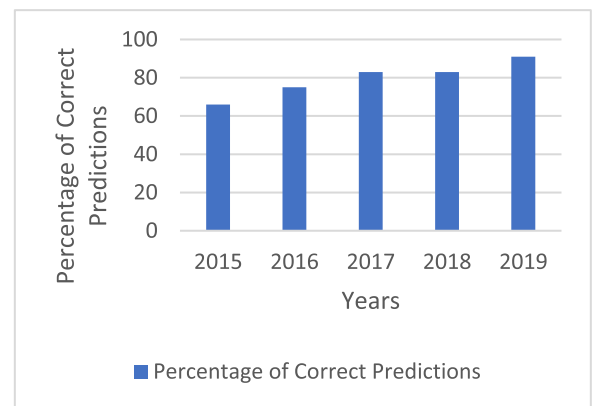


FIGURE 20. Year-wise accuracy of the predictions.

with years due to recursive feedback and training of machine learning with new data sets.

It reflects that accuracy improves from 2015 to 2019 gradually. In 2015, 66% predictions are correct, in 2016 75% predictions are correct, in 2017, 83% predictions are correct, in 2018, 83 % prediction are correct, and in 2019, 91% predictions are correct. Fig. 20 shows the accuracy of the predictions over the selected years, from where it is evident that the performance of the proposed solution increases over time.

The proposed solution accurately predicts the probability of occurrence of blister blight disease of tea (*Camellia sinensis*) before the occurrence of the disease attack. The technique is more useful against the image processing disease identification techniques that are only applicable after a disease attack has already occurred.

V. CONCLUSION

A machine learning model for the blister blight (*Exobasidium vexans*) disease prediction of tea (*Camellia sinensis*) plants is proposed by directly sensing environmental conditions from crop fields. Regression line models are developed to identify

the relationship between the environmental conditions and the development rate of the disease. Temperature, humidity, and rainfall are directly captured from the crop field using Internet of Things (IoT) capabilities. The crop field environmental data from 2015 to 2019 is used for training, testing, and validation of the proposed solution. The machine learning model shows high prediction accuracy when tested against the test data set. The predictions made by the proposed solution are also judged by direct observation from the field data. Each year the observation is also incorporated into the model as a training data set to improve the performance of the proposed model. It is observed that the prediction accuracy increased year by year. The proposed solution aims to support sustainable development in agriculture through judicious use of pesticides and effective control of the disease.

REFERENCES

- [1] S. Tibpromma, Y. Dong, S. Ranjitkar, D. A. Schaefer, S. C. Karunaratna, K. D. Hyde, R. S. Jayawardena, I. S. Manawasinghe, D. P. Bebbler, I. Promptutha, and J. Xu, "Climate-fungal pathogen modeling predict loss of up to one-third of tea growing areas," *Frontiers Cellular Infection Microbiol.*, vol. 11, p. 308, Apr. 2021, doi: [10.3389/fcimb.2021.610567](https://doi.org/10.3389/fcimb.2021.610567).
- [2] V. K. Vishnoi, K. Kumar, and B. Kumar, "Plant disease detection using computational intelligence and image processing," *J. Plant Diseases Protection*, vol. 128, no. 1, pp. 19–53, Aug. 2020, doi: [10.1007/S41348-020-00368-0](https://doi.org/10.1007/S41348-020-00368-0).
- [3] K. V. Raghavendra, D. S. B. Naik, S. Venkatramaphanikumar, S. D. Kumar, and S. V. Rama Krishna, "Weather based prediction of pests in cotton," in *Proc. Int. Conf. Comput. Intell. Commun. Netw.*, Nov. 2014, pp. 570–574, doi: [10.1109/CICN.2014.129](https://doi.org/10.1109/CICN.2014.129).
- [4] C. Chaliha and E. Kalita, "Blister blight disease of tea: An enigma," in *Diagnostics of Plant Diseases*, Dec. 2020, doi: [10.5772/INTECHOPEN.95362](https://doi.org/10.5772/INTECHOPEN.95362).
- [5] S. Sen, M. Rai, D. Das, S. Chandra, and K. Acharya, "Blister blight a threatened problem in tea industry: A review," *J. King Saud Univ. Sci.*, vol. 32, no. 8, pp. 3265–3272, Dec. 2020, doi: [10.1016/J.JKSUS.2020.09.008](https://doi.org/10.1016/J.JKSUS.2020.09.008).
- [6] X. Yang, J. Zhang, D. Guo, X. Xiong, L. Chang, Q. Niu, and D. Huang, "Measuring and evaluating anthocyanin in lettuce leaf based on color information," *IFAC-PapersOnLine*, vol. 49, no. 16, pp. 96–99, 2016, doi: [10.1016/j.ifacol.2016.10.018](https://doi.org/10.1016/j.ifacol.2016.10.018).
- [7] C. Farber, M. Mahnke, L. Sanchez, and D. Kourouski, "Advanced spectroscopic techniques for plant disease diagnostics. A review," *TrAC Trends Anal. Chem.*, vol. 118, pp. 43–49, Sep. 2019, doi: [10.1016/J.TRAC.2019.05.022](https://doi.org/10.1016/J.TRAC.2019.05.022).
- [8] H. M. Prathibhani, C. Kumarihami, and K. Jeong, "Parathibhai," *J. Korea Tea Soc.*, vol. 24, no. 3, pp. 79–87, 2018.
- [9] V. Voora, S. Bermúdez, and C. Larrea. (2019). *Global Market Report: Tea*[International Institute for Sustainable Development. Accessed: Mar. 16, 2022. [Online]. Available: <https://www.iisd.org/publications/global-market-report-tea>
- [10] A. K. Pandey, G. D. Sinniah, A. Babu, and A. Tanti, "How the global tea industry copes with fungal diseases—Challenges and opportunities," *Plant Disease*, vol. 105, no. 7, pp. 1868–1879, Jul. 2021, doi: [10.1094/PDIS-09-20-1945-FE](https://doi.org/10.1094/PDIS-09-20-1945-FE).
- [11] *China: Tea Production 2020* | Statista. Accessed: Mar. 16, 2022. [Online]. Available: <https://www.statista.com/statistics/275639/tea-production-in-china/>
- [12] *Tea Growing Regions*. Accessed: Mar. 16, 2022. [Online]. Available: <https://teapeople.co.U.K./tea-growing-regions>
- [13] S. Tibpromma, Y. Dong, S. Ranjitkar, D. A. Schaefer, S. C. Karunaratna, K. D. Hyde, R. S. Jayawardena, I. S. Manawasinghe, D. P. Bebbler, I. Promptutha, and J. Xu, "Climate-fungal pathogen modeling predicts loss of up to one-third of tea growing areas," *Frontiers Cellular Infection Microbiol.*, vol. 11, p. 1, Apr. 2021, doi: [10.3389/fcimb.2021.610567](https://doi.org/10.3389/fcimb.2021.610567).
- [14] P. B. Jawade, D. Chaugule, D. Patil, and H. Shinde, "Disease prediction of mango crop using machine learning and IoT," *Tech. Rep.*, 2020, pp. 254–260, doi: [10.1007/978-3-030-24322-7_33](https://doi.org/10.1007/978-3-030-24322-7_33).
- [15] W.-L. Chen, Y.-B. Lin, F.-L. Ng, C.-Y. Liu, and Y.-W. Lin, "RiceTalk: Rice blast detection using Internet of Things and artificial intelligence technologies," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1001–1010, Feb. 2020, doi: [10.1109/JIOT.2019.2947624](https://doi.org/10.1109/JIOT.2019.2947624).
- [16] A. Devaraj, K. Rathan, S. Jaahnavi, and K. Indira, "Identification of plant disease using image processing technique," in *Proc. Int. Conf. Commun. Signal Process. (ICCCSP)*, Apr. 2019, pp. 749–753, doi: [10.1109/ICCCSP.2019.8698056](https://doi.org/10.1109/ICCCSP.2019.8698056).
- [17] A. Khamparia, G. Saini, D. Gupta, A. Khanna, S. Tiwari, and V. H. C. de Albuquerque, "Seasonal crops disease prediction and classification using deep convolutional encoder network," *Circuits, Syst., Signal Process.*, vol. 39, no. 2, pp. 818–836, Jan. 2019, doi: [10.1007/S00034-019-01041-0](https://doi.org/10.1007/S00034-019-01041-0).
- [18] S. Orozco-Fuentes, G. Griffiths, M. J. Holmes, R. Ettelaie, J. Smith, A. W. Baggaley, and N. G. Parker, "Early warning signals in plant disease outbreaks," *Ecolog. Model.*, vol. 393, pp. 12–19, Feb. 2019, doi: [10.1016/J.ECOLMODEL.2018.11.003](https://doi.org/10.1016/J.ECOLMODEL.2018.11.003).
- [19] G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egyptian Inform. J.*, vol. 22, no. 1, pp. 27–34, Mar. 2021, doi: [10.1016/j.eij.2020.02.007](https://doi.org/10.1016/j.eij.2020.02.007).
- [20] T. Truong, A. Dinh, and K. Wahid, "An IoT environmental data collection system for fungal detection in crop fields," in *Proc. IEEE 30th Can. Electr. Comput. Eng. (CCECE)*, Apr. 2017, pp. 1–4, doi: [10.1109/CCECE.2017.7946787](https://doi.org/10.1109/CCECE.2017.7946787).
- [21] G. Liu, H. Shen, X. Yang, and Y. Ge, "Research on prediction about fruit tree diseases and insect pests based on neural network," in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov.*, vol. 187, 2005, pp. 731–740, doi: [10.1007/0-387-29295-0_79](https://doi.org/10.1007/0-387-29295-0_79).
- [22] J. Liu and X. Wang, "Tomato diseases and pests detection based on improved Yolo v3 convolutional neural network," *Frontiers Plant Sci.*, vol. 11, p. 898, Jun. 2020, doi: [10.3389/fpls.2020.00898](https://doi.org/10.3389/fpls.2020.00898).
- [23] Y. Shang and Y. Zhu, "Research on intelligent pest prediction of based on improved artificial neural network," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2018, pp. 3633–3638, doi: [10.1109/CAC.2018.8623592](https://doi.org/10.1109/CAC.2018.8623592).
- [24] R. Yakkundimath, G. Saunshi, and V. Kamatar, "Plant disease detection using IoT," *Int. J. Eng. Sci. Comput.*, vol. 8, no. 9, p. 18902, 2018.
- [25] D. Vydeki and S. Ramesh, *Ijppam.Eu*, vol. 118, no. 20, 2018, pp. 143–147.
- [26] M. Arsenovic, M. Karanovic, S. Sladojevic, and A. Anderla, "SS symmetry solving current limitations of deep learning based approaches for plant disease detection," *Tech. Rep.*, 2019.
- [27] M. Nagaraju and P. Chawla, "Systematic review of deep learning techniques in plant disease detection," *Int. J. Syst. Assurance Eng. Manage.*, vol. 11, no. 3, pp. 547–560, Jun. 2020, doi: [10.1007/S13198-020-00972-1](https://doi.org/10.1007/S13198-020-00972-1).
- [28] X. Yang and T. Guo, "Machine learning in plant disease research caused by pathogens lead significant crop yield loss globally," *Numerous Researchers Have Been Studying*, vol. 34, no. 8, pp. 1–5, 2019.
- [29] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: A review," *Plant Methods*, vol. 17, no. 1, pp. 1–18, Dec. 2021, doi: [10.1186/s13007-021-00722-9](https://doi.org/10.1186/s13007-021-00722-9).
- [30] M. Francis and C. Deisy, "Disease detection and classification in agricultural plants using convolutional neural networks—A visual understanding," in *Proc. 6th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Mar. 2019, pp. 1063–1068, doi: [10.1109/SPIN.2019.8711701](https://doi.org/10.1109/SPIN.2019.8711701).
- [31] S. Kim, M. Lee, and C. Shin, "IoT-based strawberry disease prediction system for smart farming," *Sensors*, vol. 18, no. 11, p. 4051, 2018, doi: [10.3390/s18114051](https://doi.org/10.3390/s18114051).
- [32] N. Materne and M. Inoue, "IoT monitoring system for early detection of agricultural pests and diseases," in *Proc. 12th South East Asian Tech. Univ. Consortium (SEATUC)*, Mar. 2018, Art. no. 8788860, doi: [10.1109/SEATUC.2018.8788860](https://doi.org/10.1109/SEATUC.2018.8788860).
- [33] A. A. Araby, M. M. Abd Elhameed, N. M. Magdy, L. A. Said, N. Abdelaal, Y. T. Abd Allah, M. S. Darweesh, M. A. Fahim, and H. Mostafa, "Smart IoT monitoring system for agriculture with predictive analysis," in *Proc. 8th Int. Conf. Modern Circuits Syst. Technol. (MOCASST)*, May 2019, pp. 1–4, doi: [10.1109/MOCASST.2019.8741794](https://doi.org/10.1109/MOCASST.2019.8741794).
- [34] I. Syarif, D. H. Indiarito, I. Prasetyaningrum, T. Badriyah, and E. Satriyanto, "Corn pests and diseases prediction using linear regression and natural spline methods," in *Proc. Int. Conf. Appl. Sci. Technol. (iCAST)*, Oct. 2018, pp. 383–387, doi: [10.1109/iCAST.2018.8751583](https://doi.org/10.1109/iCAST.2018.8751583).

- [35] Y. Cohen, A. Cohen, A. Hetzroni, V. Alchanatis, D. Broday, Y. Gazit, and D. Timar, "Spatial decision support system for medfly control in citrus," *Comput. Electron. Agricult.*, vol. 62, no. 2, pp. 107–117, Jul. 2008, doi: 10.1016/j.compag.2007.12.005.
- [36] S. Chaudhary, U. Kumar, and A. Pandey, "Review: Crop plant disease detection using image processing," *Int. J. Innov. Technol. Exploring Eng.*, vol. 8, no. 7, pp. 472–477, 2019.
- [37] P. Pragya, V. Sharma, and V. Sharma, "Senet CNN based tomato leaf disease detection," *Int. J. Innov. Technol. Exploring Eng.*, vol. 8, no. 11, pp. 773–777, 2019, doi: 10.35940/ijitee.k1452.0981119.



ZHIYAN LIU was born in Lanzhou, Gansu, in 1998. She graduated from Fujian Agriculture and Forestry University, in 2020. She is currently pursuing the degree with Kyushu University. Her research interests include tea science, botany, and biochemistry.



RAB NAWAZ BASHIR received the M.S. and Ph.D. degrees in computer science from the Islamia University of Bahawalpur, Bahawalpur, Pakistan, in 2015 and 2021, respectively. Currently, he is working as a Lecturer at COMSATS University Islamabad, Vehari, Pakistan. His research interest includes the Internet of Things (IoT) applications in agriculture.



SALMAN IQBAL received the M.S. degree in CS from COMSATS University Islamabad, Lahore, Pakistan, in 2009, and the Ph.D. degree in network security from the University of Malaya, Malaysia, in 2017. Currently, he is serving as an Assistant Professor at COMSATS University Islamabad. He has published more than ten research articles in high-impact ISI index journals. His research interests include network security, the IoT, healthcare, and cybersecurity.



MALIK MUHAMMAD ALI SHAHID received the M.S. degree in computer science from UET Taxila, Pakistan, in 2008, and the Ph.D. degree in computing from the University of Technology Malaysia, Malaysia, in 2017. Currently, he is working as an Assistant Professor and the Head of the Computer Science Department, COMSATS University Islamabad, Vehari. His research interests include software reliability engineering, data encryption, and smart agriculture.



MUHAMMAD TAUSIF received the M.S. degree in computer science from COMSATS University Islamabad, Sahiwal, Pakistan, in 2016. He is currently pursuing the Ph.D. degree with Superior University, Lahore, Pakistan. Currently, he is working as a Lecturer with the Department of Computer Science, COMSATS University Islamabad, Vehari, Pakistan. His current research interests include machine learning, the Internet of Things (IoTs), and sensor networks.



QASIM UMER received the B.S. degree in computer science from Punjab University, Pakistan, in 2006, the M.S. degrees in net distributed system development and in computer science from the University of Hull, U.K., in 2009 and 2013, and the Ph.D. degree from the Beijing Institute of Technology, China. He is currently working as an Assistant Professor at the Department of Computer Sciences, COMSATS University Islamabad, Vehari Campus, Pakistan. His research interests include machine learning, the Internet of Things (IoTs), software reliability, and developing practical tools to assist software engineers.

• • •