Article in Press IJIMAI journal

IoT Detection System for Mildew Disease in Roses Using Neural Networks and Image Analysis

Laura Torres[1], Luis Romero[1], Edgar Aguirre[1]*, Roberto Ferro[2]

[1] Department of Informatic and Electronic, Faculty of Engineering, Minuto de Dios University Corporation - UNIMINUTO (Colombia)
[2] Francisco José de Caldas District University, Bogotá (Colombia)

Received 30 August 2021 | Accepted 9 May 2023 | Early Access 5 July 2023

Abstract

Artificial intelligence presents different approaches, one of these is the use of neural network algorithms, a particular context is the farming sector and these algorithms support the detection of diseases in flowers, this work presents a system to detect downy mildew disease in roses through the analysis of images through neural networks and the correlation of environmental variables through an experiment in a controlled environment, for which an IoT platform was developed that integrated an artificial intelligence module. For the verification of the model, three different models of neural networks in a controlled greenhouse were experimentally compared and a proposed model was obtained for the training and validation sets of two categories of healthy roses and diseased roses with 89% training and 11% recovery. validation and it was determined that the relative humidity variable can influence the development and appearance of Downy Mildew disease when its value is above 85% for a prolonged period.

Keywords

Classification, Convolution Neural Network, Images, Information System, Risk.

DOI: 10.9781/ijimai.2023.07.001

I. Introduction

New technological paradigms are impacting the agricultural area, such as artificial intelligence that is used to monitor relevant aspects of crops [1], generating new challenges towards the use of disease and pest detection systems, which are presented in correlation to other factors such as environmental ones, for which the use of different sources of information from different media allow enriching the analysis of certain problems that arise in the sector, two main elements are the analysis of images and the Internet of Things IoT, which they have become fundamental resources to understand the behavior of diseases and in the same way that of environmental variables, which, by correlating them, allow a better understanding of the appearance of diseases in crops.

In crops, different conditions can lead to the appearance of diseases and pathogens, especially in the crop of interest, which is roses. The appearance of this disease can affect productivity and in extreme cases can cause the total loss of a crop, due to the above, the need to detect the appearance of the downy mildew disease in time[2], is evident, to diagnose the state of a plant, helping the analysis and decision making through the analysis of images with neural networks, to rapid and accurate diagnosis of a culture.

Colombia is a Latin American country that is essentially agricultural where in recent years the gross domestic product of the agricultural sector has shown low growth, but during the 2012-2016 period it grew on average 2.8%, compared to 4.2%. of the national economy [3], and the increases that have occurred are due to the increase in production, the main flower exporters worldwide managed to produce US$ 8,852 million, among which are the Netherlands (37%), Colombia (15.2%), and Ecuador (9.6%). Also, among the main importers worldwide are the USA (16%), Germany (15%), and the United Kingdom (13%) [4].

Likewise, in 2016 the main destination markets for flowers produced in Colombia were the United States with 76.2%, the United Kingdom 4.7%, and Japan 3.2%. According to the Colombian flower association, US$1,312 million were exported. In this period, roses ranked first in exports with 20.5%, followed by carnations with 17.9%, chrysanthemums with 16.4%, alstroemeria with 8%, hydrangeas with 7.5%, and 29.8% from other species [4].

From the context of microclimatic variability, variables such as temperature and relative humidity that occur in greenhouses are related, therefore another challenge is to obtain those environmental variables that, if not known and controlled, can favor the appearance of downy mildew, especially in the rainy season and develops under certain environmental characteristics [5].

Downy mildew generates large economic losses year after year. Currently, the disease causes a 10% decrease in the total production of roses in the country [6], and losses reach up to 100% of flower stems [5]. Likewise, it significantly affects production if corrective measures are not taken in time, so it is necessary to anticipate the development of the disease due to its aggressiveness [5].

For this reason, in one way or another, flower growers carry out some type of monitoring, even if it is very simple, to detect the appearance of diseases and determine whether or not to apply phytosanitary treatments in time. However, they do not have the

* Corresponding author.

E-mail address: eaguirre@uniminuto.edu

Please cite this article in press as:
L. Torres, L. Romero, E. Aguirre, R. Ferro. IoT Detection System for Mildew Disease in Roses Using Neural Networks and Image Analysis, International Journal of Interactive Multimedia and Artificial Intelligence, (2023), http://dx.doi.org/10.9781/ijimai.2023.07.001

appropriate methodology for the correct decision-making on several occasions since they are based on their knowledge and experience of the patients.

Now, the analysis of images and neural networks for the recognition of plant pests and diseases [7], proposes a new method for the detection of pests [8], and diseases [9], and the automatic supply of pesticides in greenhouse crops.

The detection of diseases in plants has been developed by different authors, in general, different types of classifiers are used to extract the particular characteristics, in the case of downy mildew in the context of corn [10] they have designed an image processing system of symptoms for disease detection using an ANN classifier.

Performance evaluation [11] is done through the K-nearest neighbor, naive bays (NB), LDA, and random forest tree (RFT) classifiers to classify diseases such as melanosis, greasy spot, and scabies. Likewise, the identification of foreign bodies [12] is used with a neural network in bulk food grains, having color and texture characteristics extracted from the sample images.

In the case of the evaluation of the quality of potatoes [13] an SVM classifier was implemented to detect defects, likewise, the same detection classifier [14] of disease spots is used.

In the case of coffee [15], it is used to quantify the severity of the rust disease by segmenting the infected areas from multispectral images, checking spots on the fruit, also in the processes [16] segmentation and feature extraction where it allows rapid and accurate detection of plant leaf diseases [17], he use of multispectral imaging [18] to classify leaf diseases is implemented particularly in cercospora beticola, and uromyces betae that is found in sugar beets.

In addition, another use of multispectral images [18] is for a K-NN classifier and a Bayesian classifier similar to [19] for the detection of defective apples due to scab, rot, and apple spot diseases, the above is based on k-means, which is also used in the detection and recognition of plant pests by k-means clustering and correspondence filter.

Due to the above, there is a problem in the early detection of the disease, taking into account that there are certain microclimatic conditions that allow it to reproduce, so the question of this work is in relation to how to detect the downy mildew disease. in roses? and the hypothesis is how to detect the disease using a neural network to analyze images and determine the occurrence of downy mildew disease in roses.

This research is focused on the development of a system for the detection of downy mildew disease in roses through image analysis using neural networks and the correlation of environmental variables through an experiment in a controlled environment, with the development of an IoT platform that integrates an artificial intelligence module, for the verification of the model, three different models of neural networks were compared with another to know the behavior of the system.

II. Literary Review

A systematic analysis of the state of the art was carried out through five guiding questions that formed the research argument, the questions were: what was the objective of the investigations? what methodologies were used? in what context were the investigations carried out? investigations? what elements were used for the detection of diseases? And what kind of results did the authors obtain?

In the investigations found, different objectives were identified Fig. 1, which are described in this section, among which the advances in the different image processing techniques used for the study of plant diseases, traits, and pests are observed.

The recognition of diseases and pests was carried out through the implementation of recognition models [20] of diseases and pests, using the leaves of the plants based on images captured from different devices, and the automatic identification of plant diseases, based on visible range imaging [21].

The detection and classification [22], [23] of diseases are carried out through the automatic identification of diseases through image processing [24], [25]. The automatic identification [26] of the disease is also processed through artificial neural networks, from information crossed with a classifier [20] with PSO feature selection [26], it is also performed with disease detection (DDS) to detect and classify leaf diseases [27].

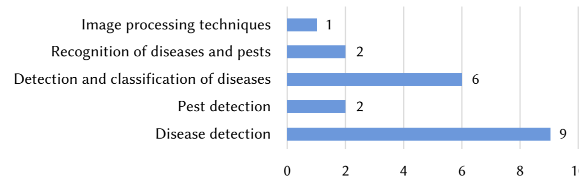


Fig. 1. Research objectives.

The methodologies and elements used are shown in Fig. 2, where one of the main elements is image acquisition [24]-[28], hence image preprocessing [29], [30], segmentation [29]-[31], comparison[32], [33], the analysis of histograms is performed and the use of detection techniques [26], [22] of edges that allow the training of the system [22], generate the knowledge base, perform the extraction of characteristics and classification, with Support Vector Machine, Matlab and neural networks.

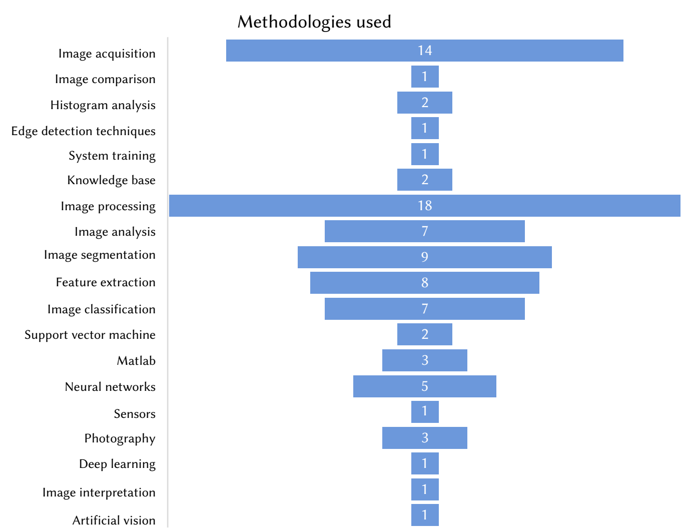


Fig. 2. Methodologies used.

The context in which the investigations are framed according to Fig. 3, is specifically divided into three parts, the first corresponds to the detection of diseases in different types of plants [21]- [34], where the rose is mentioned [35], [36]. The second is aimed at the detection of pests in different crops [37], [38]-[39], and finally, the improvement of yield in agricultural production is mentioned [40], [41]. Likewise, the investigations were carried out in different countries such as India, Mexico, Peru, and Ethiopia, and some of these in crops with controlled greenhouse environments [35] - [39].

The components and technologies used in the detection of pests and diseases Fig. 4, were mainly achieved with the implementation of image analysis, followed by image segmentation, neural networks, and support vector machines (SVM). Without neglecting tools such as OpenCV, Image Bank, CANNY Borders, and Matlab.

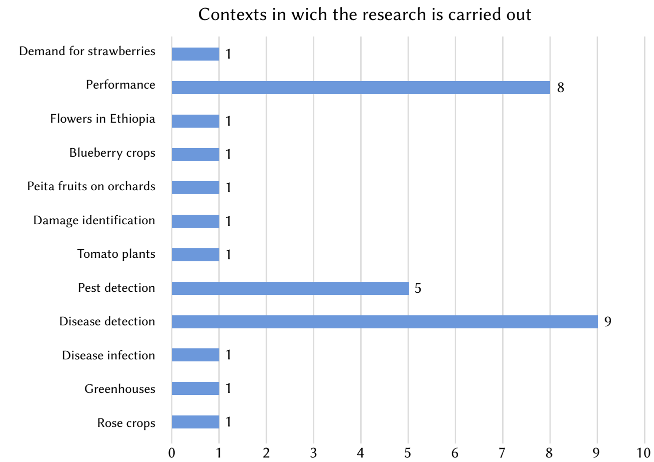


Fig. 3. Contexts in which the research is carried out.

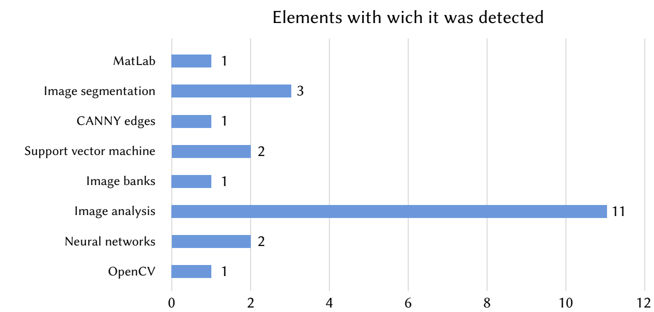


Fig. 4. Elements with which it was detected.

Of equal importance is to mention the devices and sensors that were used in the research work, where the common denominator was a simple digital camera, a cellular device, or the Everio JVC GZ-HD30 camera [35]. Additionally, a light sensor [35], was used to determine the intensity and obtain better quality images [2] as well as a scanner [39] allowing to improve the quality of the image for the detection of the target pest.

The results obtained by the investigations Fig. 5 are segmented into eight groups, within which the main one is oriented to the adequate detection of the disease [24], [42], [35], in second place is the precision in the disease identification [41], [36], [43], followed by: adequate detection of pests [20], [43], erroneous detection of diseases [28], images correctly classified with SVM [28], imágenes clasificadas correctamente con SVM [31], [32], [38], construction of hybrid systems between software and experts, reduction of human errors [37] and better performance of simple networks.

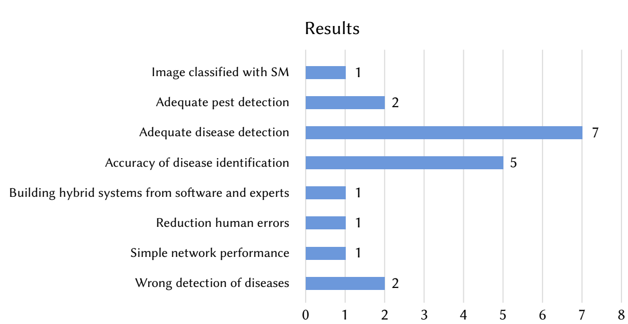


Fig. 5. Results obtained in the investigations.

III. Proposed Method

The developed system was framed in three components that allowed the experimental verification, Fig. 6, which are the on-site system, which corresponds to all the parts in charge of capturing the data of the microclimatic variables in the physical place to be monitored, the data processing, analysis, and administration is responsible for the storage, management of information and analysis through neural networks, the presentation is the last part and allows the user to access the information through a web interface.

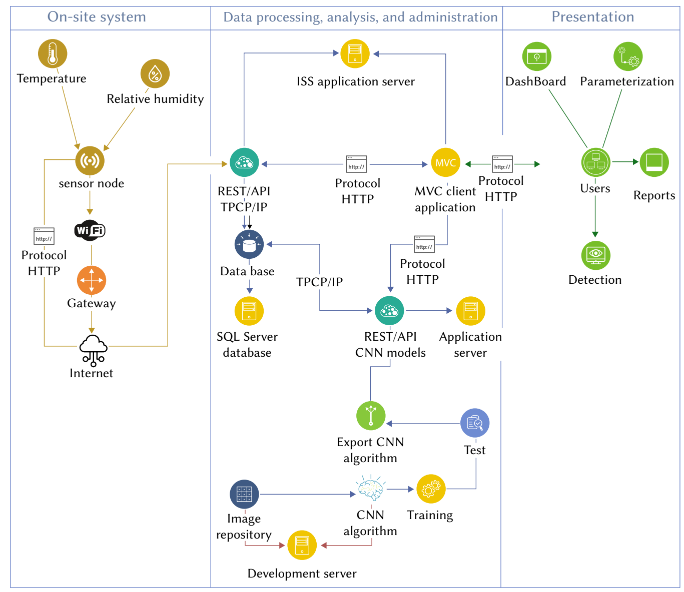


Fig. 6. Scheme of the system architecture.

The on-site system is made up of a microclimate station that integrates a temperature and relative humidity sensor, and has an internet connection, sending the captured data for storage is done through a REST/API, according to a period of time configured in the scheduled time. The installation of the station gives the user the possibility of obtaining data on the status of the variables automatically, to determine if at any time favorable conditions exist in which the rose bushes could be infected by the disease. downy mildew.

Data processing, analysis, and administration: it is the central component of the entire proposed solution since the network infrastructure and the application servers that support the operation of the system are located here; its parts are a database server in the SQL Server engine, an IIS application server that exposes all the REST/APIs that execute transactions with the database and additionally supports the client's web application, developed in ASP .Net MVC framework. On the other hand, in the detection of Downy Mildew disease based on the analysis of images or photographs, there is a machine that houses the development environment, where the training and testing of the convolutional neural network algorithms are carried out.

Likewise, an application server is integrated that exposes the REST/APIs with the Flask-Python application that gives access to the classification algorithms, stores the images in the database and returns the result of the classification thrown by each of the algorithms implemented towards the client application.

The presentation corresponds to the web interface that allows the user to view the information handled by the system and its configuration, the above, through its four main modules that are the DashBoard, where the data on changes in temperature and humidity relative to the on-site system captures them, and the alerts generated based on configured thresholds, the parameterization, here the user makes the system configurations according to his personalization data taking into account the entity to which he belongs, monitoring devices which have, among others, the following module is the reports module that allows historical queries of the behavior data of the variables and alerts generated, the detection module enables a web interface with a form to upload an image in JPEG format, which is passed to the CNN algorithms and the answer is returned with two possible options, the first that the plant is sick or that is healthy.

IV. Experiment and Results

The data architecture was approached through the design of the component diagram Fig. 7, which contains five groups of components, starting first with the connection of monitoring devices, where there are three components corresponding to the on-site monitoring device, which communicates via WiFi to send data to the Internet, the second component is the business logic, where there is the REST APIs that allow transactions between the database and the presentation MVC application and finally the processing of Images where part of the example image repository that uses the classification model for training and validation, which is ultimately exported to be consumed from the REST API by the MVC application.

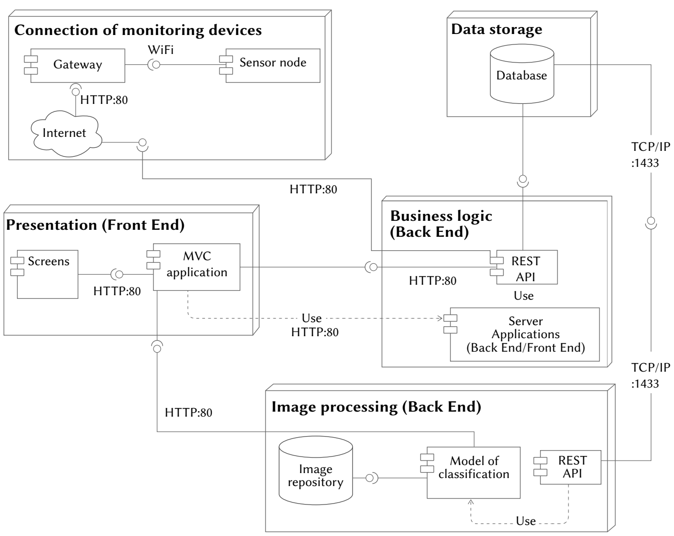


Fig. 7. Component diagram.

Second, the design of the Fig. 8 deployment diagram was made, which contains five groups of devices and a package corresponding to the IoT Hardware. The correct order for the implementation of the system begins with the implementation of the SQL Server, followed by the deployment of the IoT hardware, third the CNN model server is implemented, fourth the implements the back-end server, fifth the server is deployed to finally deploy the client.

Thirdly, the design of the entity-relationship model Fig. 9 was carried out, where eleven entities corresponding to the functionalities of the system with their respective attributes were established, the main entity is Image, which in turn is related to the Monitoring Device entity, which is related to the entities of Concentrator, User, Location and Variable, the latter is related to Captured Data, and Threshold that is related to Alert and Disease.

Finally, the database was developed in the SQL SERVER database manager, considering the entity-relationship model, the database was composed of twelve tables, within which is the Image table, where each one was stored. of the images analyzed with the date and the result obtained by each of the four classification algorithms, in addition, the image with a Device Monitoring table that in turn is related to User, Location, Concentrator, Threshold, VariableXDevice and Captured Data, these last two stores the variable assignment for each device and the data sent by the monitoring device respectively. In addition, they are related to the Variable table where the variables to be captured by the monitoring device are stored, each variable has many records in the Alert and Threshold table, which in turn are related to the Disease table.

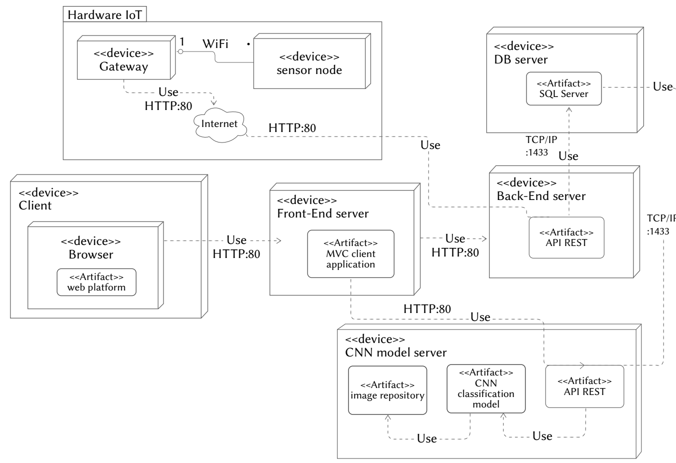


Fig. 8. Deployment diagram.

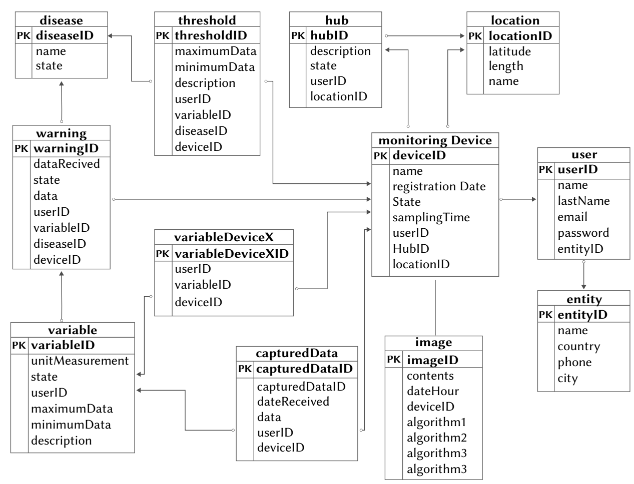


Fig. 9. Relational database diagram.

The development and coding of the IoT platform were carried out by implementing Microsoft technology in the ASP.Net framework with the C# programming language for the Front-End and Back-End, together with HTML, CSS, and JavaScript. Four modules were established based on the defined user stories.

The dashboard performs the real-time display of the data captured by the monitoring device, for the graphical representation meters from the Highcharts library were used, whose data was loaded from the JSON file returned by the REST API.

The parameterization module contains the general configurations of the system corresponding to the entity, company, or institution, location with geographic coordinates, hub, WiFi gateway or router, a monitoring device, microclimate variable, disease, and a maximum and minimum value that represents a characteristic. of a variable.

In the report module, there are the graphs corresponding to the reports generated by the system, among which are history, line graph, basic statistics, bar graph, heat map, behavior graph, alert history, table with the information of generated alerts, location of stations with a map implementing OpenStreetMap.

In the detection module, there is the upload section where the image is uploaded in .jpeg format, when you click query the image is passed to the REST API via an HTTP POST request as a file, then the REST API decodes and applies the pre-processing of the image, then passes it to the different algorithms and these return the sick or healthy response, which is encoded in JSON format so that the REST API returns it to the corresponding client.

The monitoring stations were two as in Fig. 10, both are the same in their construction, they receive the data from the sensors and transmit them to the internet using WiFi technology, each station used a pair of MCU esp32 nodes, based on the ESP32 LX6 microcontroller dual-core, 40 MHz and 520 KB SRAM with 802.11 bgN WiFi connectivity, HT40 transceiver and Bluetooth Low Energy (BLE), lithium battery charging circuit and CP2102 USB interface for connection to PC.

The function of the station was to capture the data from the sensors, the first is a temperature and relative humidity sensor, the Dht22 module was used, which has a relative humidity range between 0-99%, a temperature range between -40- 80°C, and a resolution of 8 bits, the second sensor was the TSL2561 which is a sensor that digitally measures the intensity of light ranging from 0.1 to 40,000 Lux, it operates between temperature ranges from -30 to 80 lux: from 0.1 to 40,000, operating voltage: from 2.7 to 3.6V, and manages an interface with I2C protocol, the use of this station and the transmission for data storage, which allows comparing the values of sensors against image analysis detection.

The station processes are described in algorithm I, where the "WiFi.h" and "SPI.h" libraries were imported, the access credentials to the platform were declared by the get post method and the network credentials SSID and password, the variables were defined as String of temperature, humidity and light, later the variables were verified and the communication through WiFi was initialized, the program verifies the transmission of the package, being the data previously stored temporarily in variables, which are declared in the variable fields of the HTTP hypertext transfer protocol, to build and send the message over the network.

Algorithm I. Algorithm monitoring system1 import libraries

Precondition:

```
declare network variables
declare temperature variable
declare humidity variable
define port temperature &humidity sensor
define type sensor
create an instance of temperature & humidity sensor
Connect to Wi-Fi network with SSID and password
function: connection
    set serial connection
    call begin to start sensor
end function
function: read sensors
    Read humidity sensor
    Read temperature as Celsius
    Check if any reads failed and exit early
    Print Failed to read from sensor
return
send packets POST HTTP
end function
```

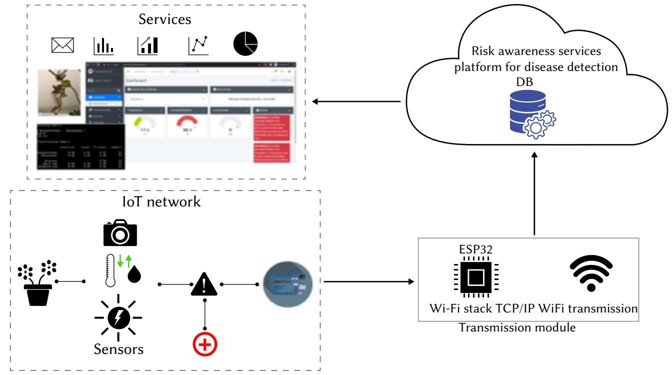
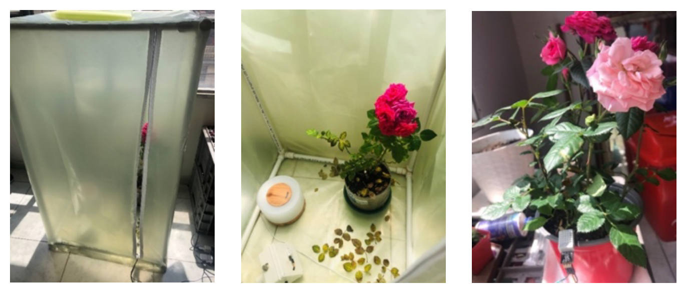


Fig. 10. Systems and subsystems of the hardware system implementation of the monitoring station.

In the context of the experiment, two scenarios were considered. The first consisted of a scale greenhouse Fig. 11a where a variety of healthy mini rose was planted, a humidifier, and a monitoring device that captured the microclimatic variables of temperature and relative humidity Fig. 11b. The second included a healthy mini rose variety in an outdoor environment and a monitoring device Fig. 11c.

The objective of the humidifier inside the greenhouse was to increase the percentage of relative humidity in the environment, to make the rose sick according to the values established in the state of the art of research for the appearance of Downy Mildew, and the monitoring device was included to be able to record the data in both scenarios.



(a) (b) (c)

Fig. 11. (a) scale greenhouse (b) interior scale greenhouse (c) healthy rose exterior. (Author, 2021)

Within the experiment, photographs were taken at different times of the day for a week Fig. 12, both greenhouse rosebush and the outdoor rosebush, obtaining a total of 100 images with 50 sick and 50 healthy ones to upload them to the rose detection module system and also make the comparison concerning the temperature and relative humidity data captured by both monitoring devices.

Fig. 12. Sample of images taken.

The main element of objective three was to implement the digital image processing algorithm shown in Fig. 13.

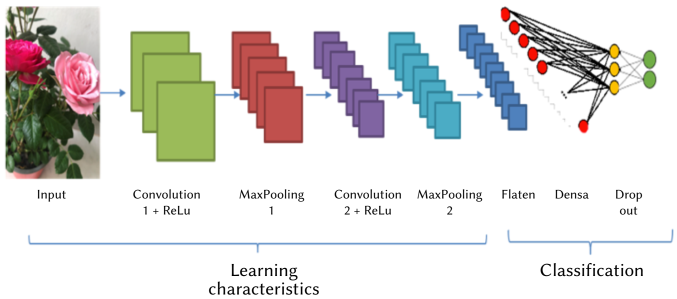


Fig. 13. Digital image processing algorithm.

Where the first thing that was carried out was the obtaining of images that allowed forming the set of images through algorithm II, for the training and testing of the four models to be implemented, a total of 200 images of different sizes were collected, one hundred are healthy rose bushes and one hundred are rose bushes that show visible symptoms of the disease with spots on the leaf.

Algorithm II. Digital image processing

Precondition:

- **set** epochs
- **set** height, length
- **set** batch size
- **set** steps
- **set** steps for validation
- **set** filters Conv
- **set** size filter
- **set** size pool
- **set** classes

1 **function:** Model
2 **set** cnn TO Sequential
3 **add** cnn ⟵ filtersConv, size filter, activation
4 **add** cnn ⟵MaxPooling, size pool
5 **add** cnn ⟵ Convolution, filters, size, activation
7 **add** cnn ⟵ compile, loss, optimizers, accuracy
8 **end function**
9 **function: Training**
10 **set** fit cnn train, steps, epochs, validation
11 **end function**

The second is the pre-processing of the images, which mainly consisted of resizing them to a size of 32 x 32 pixels or 100 x 100 pixels in RGB format. The categories or classes where zero represents those that belong to diseased roses and one those that belong to healthy roses were also coded. In the third point, the Fig. 14 repository was segmented into images for training with 89% of the samples equivalent to 178 and the remaining 11% of the test equivalent to twenty-two. In this way, the feature matrix containing the pixels of each image that is passed to the algorithm for training and testing.

In the training process, the images are passed as an array of pixels through each of the layers of the CNN sequentially and in this way it learns the weights that allowed it to classify the new images. At the heart of the CNN are the convolution layers, which using multiple k kernels allow to obtain significant characteristics of each image, such as the recognition of edges, points, objects, among others. An example is shown below in Fig. 15.

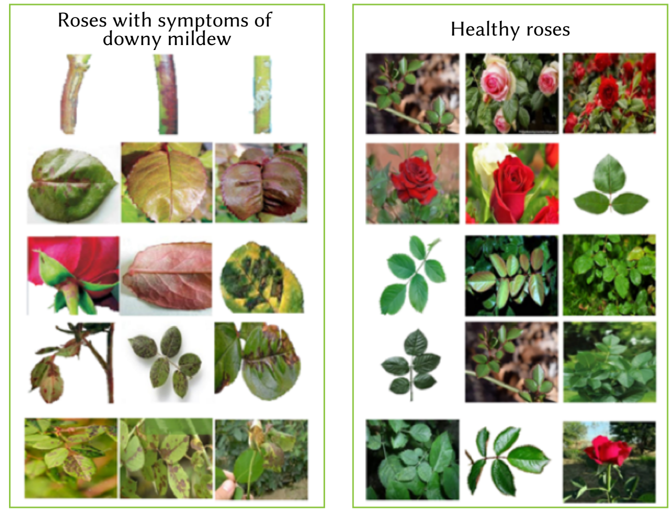


Fig. 14. Image repository sample.

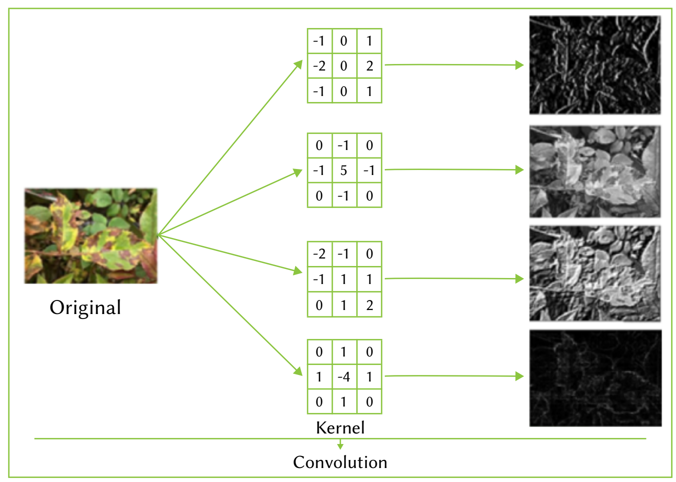


Fig. 15. Images after applying a convolution.

Finally, the results obtained from the training and testing processes of each model were evaluated to determine their quality in the classification of images representing a diseased plant. The metrics used are true positives (TP), which is the number of images that the model classified as rose plants with Downy Mildew and was correct, false positives (FP), which is the number of images that the model classified as rose plants with Downy Mildew. roses with Downy Mildew but that belong to images of healthy plants., true Negatives (VN) which is the number of images that the model classified as healthy rose plants and was correct, and false negatives (FN) which is the number of images that the model classified as healthy rose plants but belonging to images of rose plants with Downy Mildew.

Some of the metrics analyzed are found in Table 1 where precision, sensitivity and f1-score were obtained, where precision is the ratio between the number of true positives and the number of false positives. Precision is intuitively the classifier's ability to not label a negative sample as positive[44].

Sensitivity is the ratio of the number of true positives to the number of false negatives. Sensitivity is intuitively the classifier's ability to find all positive samples [44].

The F1-score can be interpreted as a weighted harmonic mean of accuracy and sensitivity, where an F-score reaches its best value of one and its worst score at zero [44].

TABLE I. Classification Metrics

Metric	Formula
Accuracy	*(positives correctly classified + negatives correctly classified)/(positives correctly classified+ false negatives + false positives + negatives correctly classified)*
Precision positives	*(positives correctly classified)/(positives correctly classified + false positives)*
Precision negatives	*(negatives correctly classified)/(positives correctly classified + negatives correctly classified*
Sensitivity	*(positives correctly classified)/(positives correctly classified + false positives)*
Specificity	*(negatives correctly classified)/(positives correctly classified + false positives)*

The results obtained from the IoT platform are framed in the final modules, the reports generated from the data captured by each monitoring device, and the analysis produced by some of the uploaded images. Fig. 16 shows the real-time values of a controlled greenhouse.

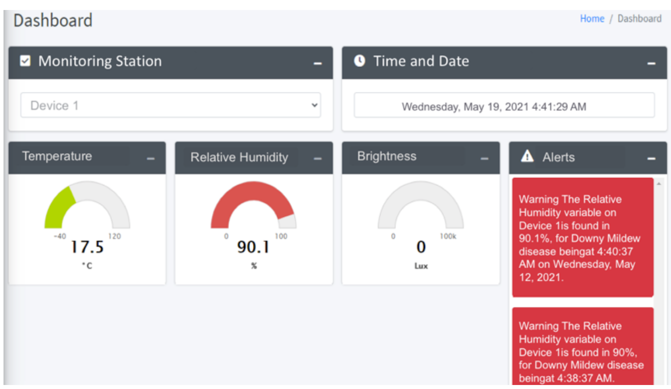


Fig. 16. Dashboard module.

From the environmental data obtained, an analysis is carried out through the stories to observe the behavior of the variables Fig. 17.

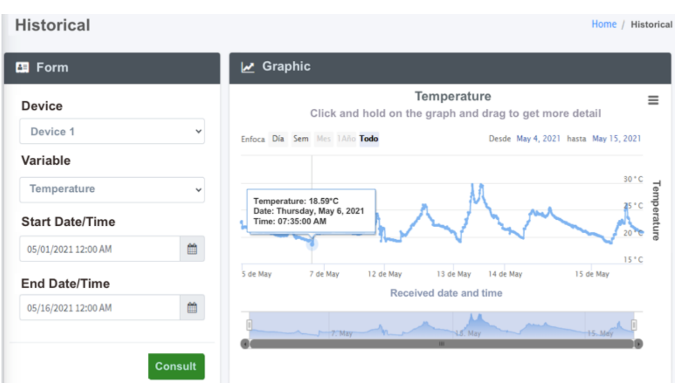


Fig. 17. History of environmental variables.

Likewise, the maximum, minimum, mode, mean, and mean values are determined in Fig. 18, which allows understanding the data set, in Fig. 19 a heat map is presented, and in Fig. 20 the loading of an image and the detection result.

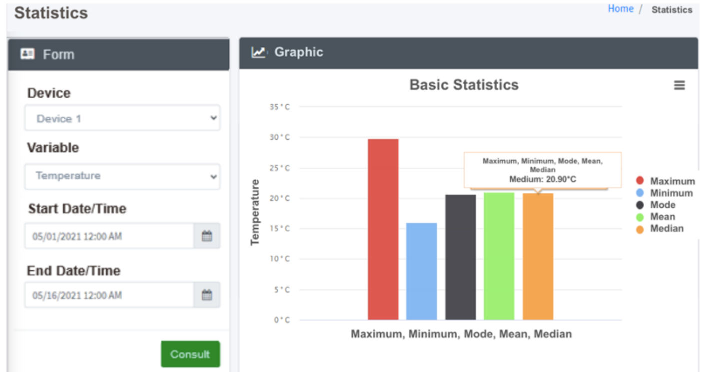


Fig. 18. Basic statistics.

Fig. 19. Heat map of the temperature variable.

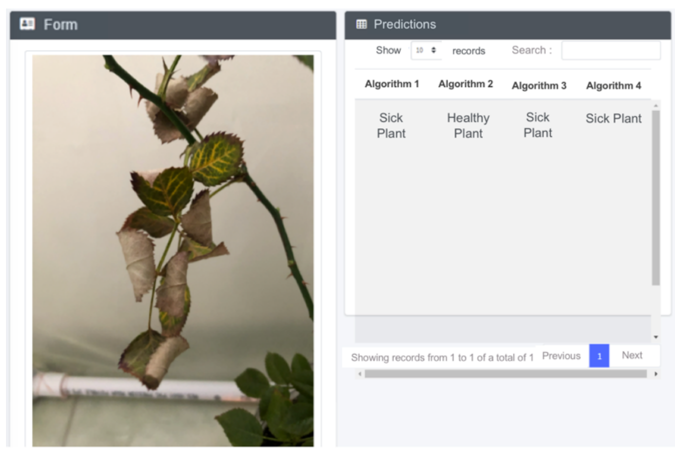


Fig. 20. Load image detection module.

To demonstrate the behavior of the microclimatic variables of temperature and relative humidity concerning the state of each rosebush, rose one corresponds to the one arranged in the greenhouse and rose two to the one located outside, to determine the effect of these on the state of the rose, a comparison was made of the data obtained by each monitoring device in a range of one week, which in this case are segmented as follows: device one corresponds to the sensor node located inside the greenhouse and device one device two corresponds to the sensor node located outside.

Fig. 21 represent the behavior and the temperature value that is repeated the most, respectively, for device one in a time range of one week. The most persistent data during the test was 22.50 °C, this being normal for the temperature required by the pink one.

Fig. 21. Device one temperature graph.

Fig. 22 represent the behavior and the most repeated humidity value, respectively, for device one in a time range of one week, the most persistent data during the test was 99.90% relative humidity. in the environment, this being very high and conducive to the appearance of Downy Mildew disease in the pink one.

Fig. 22. Humidity graph device one

Fig. 23 represent the behavior and the temperature value that is repeated the most, respectively, for device two in a time range of one week, the most persistent data during the test was 19.10°C, being this normal for the temperature required by pink two.

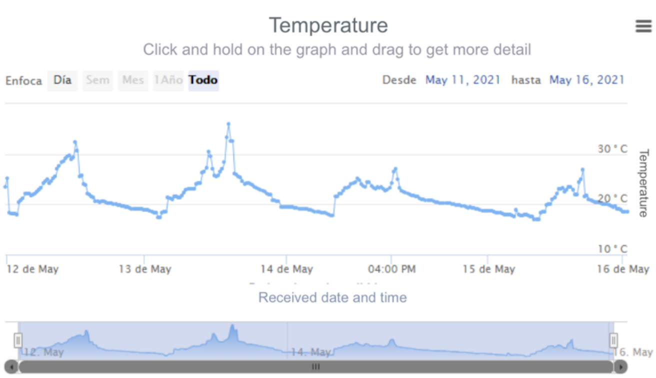


Fig. 23. Temperature behavior statistics

Fig. 24a shows the initial state of the rosebush before subjecting it to high relative humidity for a prolonged period of one week, and Fig. 24b shows the final state of the rosebush when observing the presence of Mildew Downy's disease.

Fig. 25a shows the initial state of rosebush two in an outdoor environment without alterations for one week and in Fig. 25b the final state of rosebush two is evidenced, observing that there are no changes in the state of the plant.

(a) (b)

Fig. 24. (a) Start of pink test 1. (b) End of pink test 1.

(a) (b)

Fig. 25. (a). Start of pink test 2. (b) End of pink test 2.

V. Analysis

About the data collected by the monitoring devices and the reports developed on the platform, the basic statistics graph was taken, which provides the mode information for each variable among other values, to observe what the value was of the data that was the most repeated during the execution time of the tests.

The mode value corresponding to the relative humidity for device one and rose one is outside the edaphoclimatic requirements of the rose, exposing it to an optimal level of humidity for the outbreak of Downy Mildew, which is triggered when the presence of the inoculum coincides. with high relative humidity above 85%, for a time greater than three hours according to [34] and when making the comparison between the beginning and the end of the state of the rose, it can be seen that indeed the alteration of humidity in levels as elevated for a long time, they can cause the appearance of Downy Mildew disease. Therefore, monitoring these microclimatic variables and generating alerts opens the possibility for the user to start monitoring with images for early detection of the disease.

The mode value corresponding to the relative humidity for device two and rosebush two is within the edaphoclimatic requirements of the rosebush, reducing the probability of the appearance of Downy Mildew disease as it is not exposed to microclimatic conditions that favor the appearance of the illness.

Regarding the results obtained in the different training and validation cycles, the CNN models were selected with the parameters of cycle four with input image size = 100 x 100 x 100 x 3 = 30,000, epochs = 100, and batch size = 32 since they were the ones that returned metrics closest to those expected, to obtain the greatest success in classifying new images in a real environment.

Table II presents the results classification report corresponding to the training stage of the final CNN models, to carry out a comparative analysis of them.

TABLE II. Classification Report

	Precision	Recall	F1-score	Support
Shallow Net				
Sick roses	0.67	0.55	0.6	11
Healthy roses	0.62	0.73	0.67	11
Accuaricy	no data	no data	0.64	22
Macro avg	0.64	0.64	0.63	22
weighted	0.64	0.64	0.63	0.63
LeNet				
Sick roses	0.75	0.27	0.4	11
Healthy roses	0.56	0.91	0.69	11
Accuaricy	no data	no data	0.59	22
Macro avg	0.65	0.59	0.54	22
weighted	0.65	0.59	0.54	22
MiniVGGNet				
Sick roses	0.75	0.55	0.63	11
Healthy roses	0.64	0.82	0.72	11
Accuaricy	no data	no data	0.68	22
Macro avg	0.7	0.68	0.68	22
weighted	0.7	0.68	0.68	22
Proposal				
Sick roses	0.67	0.36	0.47	11
Healthy roses	0.56	0.82	0.67	11
Accuaricy	no data	no data	0.59	22
Macro avg	0.61	0.59	0.57	22
weighted	0.61	0.59	0.57	22

In the first metric, training loss where the behavior is expected to decrease towards zero as times increase, which is an indicator that the model is learning from the example images, however, the results of all three. The former show pronounced variations in increase and decrease, since they show the described trend and reach a minimum value not less than 0.4.

This indicates that the learning process is carried out with errors, for the compensation of the error in the training set the images were repaired by the predictor and the ones that were detected were selected, to obtain the training set for which I review the result or output for each value $\hat{y}_i$ and compare this value with the target value y_i, having as reference the subtraction $\hat{y}_i$ - y_i, or the entire dataset, the total error committed was calculated as the sum of the individual errors Therefore, the best images and weights were chosen to minimize the total error, which was initially 50%, but as the input data was improved and the algorithm was improved, this was substantially reduced. On the other hand, the proposed model shows a more stable downward trend, reaching a value of at least 0.01, which indicates a better learning rate.

The second metric is the validation loss, which is applied to the validation images and deviated the most from ideal behavior in all training cycles with each of the algorithms. The expected trend should decrease towards zero as the epochs increase, however, the result indicates a growth with drastic variations in increase and decrease indicating possible overfitting of the model, that is, that the model only behaves well with the training images and when entering new ones it is not able to have an acceptable success rate in its classification as a sick rose or a healthy rose.

The third metric is the training accuracy where the behavior is expected to grow towards one as the epochs increase, which is an indicator that the model is learning from the example image features and classifying them correctly.

The fourth metric is the validation accuracy where the behavior is expected to grow towards one as the epochs increase, which would imply the correct classification of the images used in the validation set.

Table III presents the set of metrics of the confusion matrix and table IV the classification report of the results corresponding to the validation stage of the CNN models.

TABLE III. Confusion Matrix Table

Model name	True Positive (TP)	False Positive (FP)	False Negative (FN)	True Negative (TN)
ShallowNet	6	5	3	8
LeNet	3	8	1	10
VGGNet	6	5	2	9
Proposed	4	7	2	9

TABLE IV. Classification Report Confusion Matrix

Model name	Accuaricy	Precision positives	Precision negatives	Sensitivity	Specificity
Shallow Net	0.6	0.5	0.7	0.6	0.6
LeNet	0.5	0.2	0.9	0.7	0.5
VGGNet	0.6	0.5	0.8	0.7	0.6
Proposed	0.5	0.3	0.8	0.6	0.5

Taking the confusion matrix as a reference, it was possible to show that the LeNet and proposed algorithms succeed and fail in the classification, obtaining a total of thirteen successes and nine failures out of twenty-two images belonging to the validation repository. On the other hand, the ShallowNet and MiniVGGNet algorithms achieve a total of fourteen and fifteen hits, respectively, so in the analysis of the confusion matrix the best algorithm is MiniVGGNet.

On the other hand, the classification report shows how the LeNet and proposed algorithms reach low insensitivity scores for the diseased roses category, indicating that the algorithms have a low ability to classify images belonging to this category. Regarding the ShallowNet and MiniVGGNet algorithms, it can be seen that the scores obtained for the metrics improve slightly to those of the LeNet algorithms and proposed highlights the MiniVGGNet algorithm.

The set of images that allowed the test to be carried out in a laboratory environment is made up of a total of one hundred, of which fifty are photos of the part where the plant presented the disease and fifty of the healthy part.

To begin with, based on the confusion matrix, it is possible to determine that the ShallowNet classifier is not capable of delivering the correct category for the images that correspond to diseased roses, since there is a total of true positives (TP) equal to zero and false positives (FP) equal to nine. Regarding the classification of healthy roses, he managed to hit forty-one (VN) and failed a total of fifty (FN). The LeNet classifier managed to reach a total of thirty-five (VP) diseased roses and a total of forty-eight failed (FP), a total of two (VN) healthy roses, and a total of fifteen failed (FN). Therefore, LeNet is a better classifier than ShallowNet.

Continuing, the MiniVGGNet classifier returned a total hit of fifty (VP) as diseased roses, however, it missed a total of forty-nine (FP), and in the category of healthy roses it managed to hit a total of one (VN), its failure is equal to zero (FN). This provides an improvement over LeNet. Finally, the proposed classifier resulted in success in sick roses with a total of sixteen (VP) and failure in forty-one (FP), and in healthy roses, the classifier correctly returns nine (VN) but fails in thirty-four (FN). Thus, the proposed algorithm is below LeNet and MiniVGGNet in hits.

Finally, additional metrics such as the precision that represents the percentage of correct answers at a general level for all the samples confirm that, despite the failures in the MiniVGGNet classification, it is the one that performs best. Furthermore, it is reflected in the metrics not mentioned so far: the precision per category, the sensitivity (correct true positives), and the represented specificity (true negatives) confirm that MiniVGGNet was the most successful.

VI. Conclusion

The theoretical foundation on the different methodologies, components, and technologies implemented in the detection, classification of pests and diseases in crops, allowed a better perception of the operation, development, training, validation, and deployment of four models of convolutional neural networks, so addressing the solution of the problem from the creation of the set of images to obtaining the result of the experiment.

Including the development of a web platform in conjunction with a monitoring system for microclimatic variables that create environments conducive to the appearance of mildew adds value to the project since it allows alerts when the temperature and/or relative humidity are out of range. and initiate follow-up by taking images that are passed to classification models. To detect the disease in its initial stage, avoiding major effects on the cultivation of roses.

When carrying out the implementation of the CNN, the segmentation process of the image repository was evidenced, the extraction of characteristics of an image by executing an example with a 2D convolution layer that applied several k kernels that resulted in transformed images highlighting the characteristics of the edges and spots on the leaf of a rose. This allows CNN to learn the filters (kernel group k) to then classify the new input images.

As can be seen, in the results of the CNN in its training and validation stage, it was possible to show that all the graphs that show the loss and validation precision metrics have behaviors with drastic variations (increase and decrease) and the trend is opposite. . than expected, indicating possible overfitting in the learning performed by the models. Identified due to the few image samples that could be obtained for the training and validation sets (200 in total, in 2 categories: healthy roses and diseased roses with 89% training and 11% validation) and consequently, the training of the algorithms was not performed. robust enough, making most results unacceptable.

According to the tests and the analysis of the results obtained with the microclimatic variables, it is observed that the relative humidity variable can influence the development and appearance of Downy Mildew disease when its value is above 85% during an extended period.

In the process of developing the research project, some future work fronts emerged related to neural networks and the implementation of an IoT platform for the storage of microclimatic data in real-time, since by merging these two modules it is possible to have a very robust system for the detection of any type of disease in plants. This is due to the detection module included in the CNNs and the possibility of real-time monitoring of the variables that affect a crop.

In the same way, the project can be improved by having a repository that has more examples of training and validation images and implementing other types of deeper CNNs that are not sequential, such as ResNet, ResNet50, among others, to compare them as well. with those implemented in this project.

Additionally, the monitoring device could be improved by including a camera as a sensor to take photos on-site and in real-time, which would allow a greater real knowledge of the current state of the crop.

References

[1] Y. Li, C. Xia, and J. Lee, "Vision-based pest detection and automatic spray of greenhouse plant," in *IEEE International Symposium on Industrial Electronics*, 2009, pp. 920–925. doi: 10.1109/ISIE.2009.5218251.

[2] A. Calderón and H. Hurtado, "Vista de Machine learning en la detección de enfermedades en plantas," *Tecnología, investigación y academia TIA*, vol. 7, no. 2, pp. 55–62, Dec. 2019, Accessed: May 30, 2023. [Online]. Available: https://revistas.udistrital.edu.co/index.php/tia/article/view/15685/15932

[3] F. Reyes, L. Cruz, N. Cáceres, and E. Valero, "Desempeño del sector floricultor," Bogotá D.C., 2017. Accessed: May 30, 2023. [Online]. Available: https://sioc.minagricultura.gov.co/Flores/Normatividad/2016-06-01%20Boletin%20desempeño%20sector%20floricultor.pdf

[4] S. Johnny, "Ventaja comparativa del sector floricultor colombiano que promueva su presencia y le permita fortalecerse en el marco del TLC con corea del sur," Bogotá D.C., 2018. Accessed: May 30, 2023. [Online]. Available: https://repositorio.uniagustiniana.edu.co/handle/123456789/370

[5] P. Israel *et al.*, "Current Status of Peronospora sparsa, Causal Agent of Downy Mildew on Rose (Rosa sp.) Estado Actual de Peronospora sparsa, Causante del Mildiu Velloso en Rosa (Rosa sp.)," *Rev. mex. fitopatol*, vol. 31, no. 2, pp. 113–115, 2013, Accessed: May 30, 2023. [Online]. Available: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-33092013000200004&lng=es.

[6] M. Ayala, I. A.; Luz, E. Argel-Roldan, S. Jaramillo-Villegas, M. Marín-Montoya, and M. M. Montoya, "Diversidad genética de peronospora sparsa (peronosporaceae) en cultivos de rosa de Colombia," *Acta biológica Colombiana*, vol. 13, no. 1, pp. 79–94, 2008, Accessed: May 30, 2023. [Online]. Available: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-548X2008000100005&lng=en.

[7] A. Flórez, O. Hurtado, and S. Ramos, "Procesamiento de imágenes para reconocimiento de daños causados por plagas en el cultivo de Begonia semperflorens (flor de azúcar)," *Acta Agronomica*, vol. 64, no. 3, pp. 273–278, 2014, doi: 10.15446/acag.v64n3.42657.

[8] C. Cáceres, D. Amaya, and O. Ramos, "Methodology for pest damage recognition in Begonia semperflorens link & Otto (sugar flower) crop through image processing," *Acta Agronomica*, vol. 64, no. 3, pp. 257–264, 2015, doi: 10.15446/acag.v64n3.42657.

[9] F. Qin, D. Liu, B. Sun, L. Ruan, Z. Ma, and H. Wang, "Identification of Alfalfa Leaf Diseases Using Image Recognition Technology," *PLoS One*, vol. 11, no. 12, p. e0168274, Dec. 2016, doi: 10.1371/journal.pone.0168274.

[10] S. A. Patil, D. S. Khot, O. D. Otari, and U. G. Malavkar, "Automated Disease Detection and Classification of Plants Using Image Processing Approaches: A Review," *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security*, vol. 203, pp. 641–651, 2021, doi: https://doi.org/10.1007/978-981-16-0733-2_45.

[11] S. Reddy Bandi, "Performance evaluation of various statistical classifiers in detecting the diseased citrus leaves," *International Journal of Engineering Science and Technology (IJEST)*, vol. 5, no. 2, pp. 98–307, 2013, Accessed: Feb. 28, 2022. [Online]. Available: https://www.idc-online.com/technical_references/pdfs/information_technology/PERFORMANCE.pdf

[12] U. Ansari, S. Shantaiya, and M. Ansari, "Identification Of Food Grains And Its Quality Using Pattern Classification," *International Journal of Computer and Communication Technology*, vol. 3, no. 1, 2012, doi: 10.47893/IJCCT.2012.1107.

[13] N. Razmjooy, B. S. Mousavi, and F. Soleymani, "A real-time mathematical computer method for potato inspection using machine vision," *Computers & Mathematics with Applications*, vol. 63, no. 1, pp. 268–279, Jan. 2012, doi: 10.1016/J.CAMWA.2011.11.019.

[14] S. Prasad, P. Kumar, R. Hazra, and A. Kumar, "Plant leaf disease detection using Gabor wavelet transform," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7677 LNCS, pp. 372–379, 2012, doi: 10.1007/978-3-642-35380-2_44.

[15] D. Cui, Q. Zhang, M. Li, G. L. Hartman, and Y. Zhao, "Image processing

methods for quantitatively detecting soybean rust from multispectral images," *Biosystems Engineering*, vol. 107, no. 3, pp. 186–193, Nov. 2010, doi: 10.1016/J.BIOSYSTEMSENG.2010.06.004.

[16] T. B. and S. V Rathod A N., "Image Processing Techniques for Detection of Leaf Disease," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 11, pp. 397–399, 2013.

[17] S. D. Bauer, F. Korč, and W. Förstner, "The potential of automatic methods of classification to identify leaf diseases from multispectral images," *Precision Agriculture*, vol. 12, no. 3, pp. 361–377, Jun. 2011, doi: 10.1007/S11119-011-9217-6/FIGURES/5.

[18] A. Pacheco, H. Bolivar-Baron, R. Gonzalez-Crespo, and J. Pascual-Espada, "Reconstruction of High Resolution 3D Objects from Incomplete Images and 3D Information," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 2, no. 6, p. 7, 2014, doi: 10.9781/IJIMAI.2014.261.

[19] S. R. Dubey, P. Dixit, N. Singh, and J. P. Gupta, "Infected Fruit Part Detection using K-Means Clustering Segmentation Technique," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 2, no. 2, p. 65, 2013, doi: 10.9781/IJIMAI.2013.229.

[20] A. Devaraj, K. Rathan, S. Jaahnavi, and K. Indira, "Identification of plant disease using image processing technique," *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, pp. 749–753, 2019, doi: 10.1109/ICCSP.2019.8698056.

[21] J. G. A. Barbedo, "A review on the main challenges in automatic plant disease identification based on visible range images," *Biosystems Engineering*, vol. 144, pp. 52–60, 2016, doi: 10.1016/j.biosystemseng.2016.01.017.

[22] M. Khari, A. K. Garg, R. Gonzalez-Crespo, and E. Verdú, "Gesture Recognition of RGB and RGB-D Static Images Using Convolutional Neural Networks," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 7, p. 22, 2019, doi: 10.9781/IJIMAI.2019.09.002.

[23] Y. H. Robinson, S. Vimal, M. Khari, F. C. L. Hernández, and R. G. Crespo, "Tree-based convolutional neural networks for object classification in segmented satellite images," *International Journal of High Performance Computing Applications*, Jul. 2020, doi: 10.1177/1094342020945026.

[24] V. S. Bhong and P. B. V Pawar, "Study and Analysis of Cotton Leaf Disease Detection Using Image Processing," *International Journal of Advanced Research in Engineering and Technology*, vol. 3, no. 2, pp. 1447–1454, 2016, doi: 10.1088/1742-6596/2062/1/012009.

[25] R. Gupta, M. Khari, D. Gupta, and R. G. Crespo, "Fingerprint image enhancement and reconstruction using the orientation and phase reconstruction," *Information Sciences (N Y)*, vol. 530, pp. 201–218, Aug. 2020, doi: 10.1016/J.INS.2020.01.031.

[26] J. M. T. Ruiz, Jesús Gil and R. G. Crespo, "The Application of Artificial Intelligence in Project Management Research: A Review," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp. 54–66, 2021, doi: https://doi.org/10.9781/ijimai.2020.12.003.

[27] K. Indumathi, R. Hemalatha, S. A. Nandhini, and S. Radha, "Intelligent plant disease detection system using wireless multimedia sensor networks," *Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017*, vol. 2018-Janua, pp. 1607–1611, 2018, doi: 10.1109/WiSPNET.2017.8300032.

[28] G. Tigistu and Y. Assabie, "Automatic identification of flower diseases using artificial neural networks," *IEEE AFRICON Conference*, vol. 2015-Novem, 2015, doi: 10.1109/AFRCON.2015.7332020.

[29] V. Singh, Varsha, and A. K. Misra, "Detection of unhealthy region of plant leaves using image processing and genetic algorithm," *Conference Proceeding - 2015 International Conference on Advances in Computer Engineering and Applications, ICACEA 2015*, pp. 1028–1032, 2015, doi: 10.1109/ICACEA.2015.7164858.

[30] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," *Proceedings - 1st International Conference on Computing, Communication, Control and Automation, ICCUBEA 2015*, vol. 7677, pp. 768–771, 2012, doi: 10.1109/ICCUBEA.2015.153.

[31] J. N. Reddy, K. Vinod, and A. S. R. Ajai, "Analysis of Classification Algorithms for Plant Leaf Disease Detection," *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, pp. 1–6, 2019, doi: 10.1109/ICECCT.2019.8869090.

[32] R. Meena Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari, and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," *Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems, ICIIECS 2017*, vol. 2018-Janua, pp. 1–4, 2018, doi: 10.1109/ICIIECS.2017.8275915.

[33] F. Jobin, S. D. Anto, and B. K. Anoop, "Identification of leaf diseases in pepper plants using soft computing techniques," *Conference on emerging devices and smart systems (ICEDSS)*, pp. 168–173, 2016, doi: 10.1109/ICEDSS.2016.7587787.

[34] P. Revathi and M. Hemalatha, "Identification of cotton diseases based on cross information gain_deep forward neural network classifier with PSO feature selection," *International Journal of Engineering and Technology*, vol. 5, no. 6, pp. 4637–4642, 2013.

[35] N. Velázquez-López, Y. Sasaki, K. Nakano, J. M. Mejía-Muñoz, and E. R. Kriuchkova, "Detección de cenicilla en rosa usando procesamiento de imágenes por computadora," *Revista Chapingo Serie Horticultura*, vol. 17, no. 2, pp. 151–160, 2011, Accessed: May 30, 2023. [Online]. Available: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1027-152X2011000200008&lng=es&tlng=es.

[36] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors (Switzerland)*, vol. 17, no. 9, 2017, doi: 10.3390/s17092022.

[37] S. Baskaran, "Advances in Image Processing for Detection of Plant Disease," *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA)*, vol. 05, no. 02, pp. 08–10, 2017, doi: 10.9756/sijcsea/v5i2/05010140101.

[38] P. Rajan, B. Radhakrishnan, and L. Padma Suresh, "Detection and classification of pests from crop images using Support Vector Machine," *Proceedings of IEEE International Conference on Emerging Technological Trends in Computing, Communications and Electrical Engineering, ICETT 2016*, 2017, doi: 10.1109/ICETT.2016.7873750.

[39] P. Boissard, V. Martin, and S. Moisan, "A cognitive vision approach to early pest detection in greenhouse crops," *Computers and Electronics in Agriculture*, vol. 62, no. 2, pp. 81–93, 2008, doi: 10.1016/j.compag.2007.11.009.

[40] D. E. Kusumandari, M. Adzkia, S. P. Gultom, M. Turnip, and A. Turnip, "Detection of Strawberry Plant Disease Based on Leaf Spot Using Color Segmentation," *Journal of Physics: Conference Series*, vol. 1230, no. 1, 2019, doi: 10.1088/1742-6596/1230/1/012092.

[41] R. G. De Luna, E. P. Dadios, and A. A. Bandala, "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition," *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, vol. 2018-Octob, no. October, pp. 1414–1419, 2019, doi: 10.1109/TENCON.2018.8650088.

[42] L. Shanmugam, A. L. A. Adline, N. Aishwarya, and G. Krithika, "Disease detection in crops using remote sensing images," *Proceedings - 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development, TIAR 2017*, vol. 2018-Janua, no. Tiar, pp. 112–115, 2018, doi: 10.1109/TIAR.2017.8273696.

[43] C. Sullca, C. Molina, C. Rodríguez, and T. Fernández, "Detección de enfermedades y plagas en las hojas de arándanos utilizando técnicas de visión artificial," *Perspectivas*, vol. 15, no. 15, pp. 32–39, 2018.

[44] A. Navlani, "Naive Bayes Classification using Scikit-learn," *Datacamp*, 2020. https://www.datacamp.com/tutorial/naive-bayes-scikit-learn (accessed May 30, 2023).

Laura Estefania Torres Tovar

Systems engineer from the Corporación Universitaria Minuto de Dios-UNIMINUTO, works as a full-stack developer for private companies.

Luis Carlos Romero Cardenas

Systems engineer from the Corporación Universitaria Minuto de Dios-UNIMINUTO, works as a full-stack developer for private companies.

Roberto Ferro Escobar

Roberto Ferro Escobar is an Electronics Engineer from the Francisco José de Caldas District University, a master's in information and communication Sciences, and a Ph.D. in Computer Engineering from the Pontifical University of Salamanca. Full Professor at the same University. Researcher in HPC, Data Science, and Bioinformatics.

Edgar Alirio Aguirre Buenaventura

Edgar Aguirre He is a PhD in Engineering from the Francisco José de Caldas District University, master's in information and communication Sciences, Electronic Control Engineer, Electronics Technologist, works as director of the INDEC community development engineering institute, and is Professor in the Electronic Technology Program of the Corporación Universitaria Minuto de Dios-UNIMINUTO.