

## Development of a digital image classification system to support technical assistance for Black Sigatoka detection

Cristian Andrés Escudero<sup>1</sup>, Andrés Felipe Calvo<sup>2</sup>, Arley Bejarano Martinez<sup>3</sup>, Ana María López<sup>4</sup>, Alexander Molina<sup>5</sup>

**Abstract-** A large percentage of Colombia's economic activity corresponds to the agricultural sector. In this sector, plantains rank second in production and planted area. However this crop is affected by different diseases, among which The Black Sigatoka stands out, caused by the fungus *Mycosphaerella fijiensis*. The disease highly reduces the production level of the crop and although there are prevention measures that allow reducing the incidence of the disease, there's a lack of support for small producers in Colombia, who do not have technological tools to support the disease detection processes. This article outlines the development of a support system for the detection of black sigatoka using digital images. For this, a characterization process of the agricultural user is carried out, then, a machine learning methodology is implemented to classify the disease on a mobile device. The support system is validated through laboratory tests, field tests and the feedback from the agricultural user.

**Index terms:** Black sigatoka, banana, agricultural sector, machine learning, mobile application.

## Desenvolvimento de um sistema de classificação de imagem digital para suporte a assistência técnica na detecção de Sigatoka Negra

**Resumo** – Uma grande porcentagem da atividade econômica na Colômbia corresponde ao setor agrícola. Neste, a banana verde ocupa o segundo lugar em produção e área plantada. No entanto, a plantação é afetada por diversas doenças, dentre as quais se destaca a Sigatoka negra, causada pelo fungo *Mycosphaerella fijiensis*. A doença reduz muito o nível de produção agrícola e embora existam medidas de prevenção que permitam reduzir sua incidência, falta apoio aos pequenos produtores, que não possuem ferramentas tecnológicas para apoiar os processos de detecção da doença.

Este artigo apresenta o desenvolvimento de um sistema de suporte à detecção de Sigatoka negra por meio de imagens digitais. Para isso, é realizado um processo de caracterização do usuário agrícola e, em seguida, implementa-se uma metodologia de aprendizado de máquina para classificar a doença em um dispositivo móvel. O sistema de apoio é validado a partir de testes laboratoriais e de campo e socialização com o usuário agrícola.

**Termos de indexação:** Sigatoka negra, banana, setor agrícola, aprendizado de máquina, aplicativo móvel.

**Corresponding author:**  
cristian-escuder@utp.edu.co

**Received:** July 09, 2020  
**Accepted:** October 05, 2020

**Copyright:** All the contents of this journal, except where otherwise noted, is licensed under a Creative Commons Attribution License.



<sup>1</sup>Msc. Tecnológica University of Pereira. Faculty of EEPC Engineering Research group on Electronic Engineering. E-mail: cristian-escuder@utp.edu.co (ORCID 0000-0003-3703-7520)

<sup>2</sup>Doctoral Student on Engineering. Tecnológica University of Pereira. Faculty of EEPC Engineering Research group on Electronic Engineering. E-mail: afcervo@utp.edu.co (ORCID 0000-0001-9409-8982)

<sup>3</sup>Msc. Tecnológica University of Pereira. Faculty of EEPC Engineering Research group on Electronic Engineering. E-mail: abejarano@utp.edu.co (ORCID 0000-0002-0470-537X)

<sup>4</sup>Doctoral Student of agricultural Sciences (University of Caldas). Tecnológica University of Pereira. Vegetal Biotechnology Laboratory. E-mail: alopez@utp.edu.co (ORCID 0000-0002-5138-1806)

<sup>5</sup>PhD. Tecnológica University of Pereira. Faculty of EEPC Engineering Research Group on Electromagnetic Fields and energetic phenomena E-mail: almo@utp.edu.co (ORCID 0000-0003-2448-2174)

## Introduction

In Colombia, plantain is one of the most important food products of the market basket; consumption per capita in the country is approximately 155 kg/person/year (RUÍZ; URUEÑA, 2009). Creating 286.000 permanent direct jobs a year and generating the livelihood of 57.000 families (ESPINAL et al. 2005).

Over the last decade, the plantain production sector's production has duplicated, going from 2.604.550 tons in 2008 to 4.316.726 tons in 2018 (SIOC, 2018); however, this is conceived by an increase in the area planted, not by the improvement in the production of crops, since generally there are low yields, being the national average of 8 Ton/ha/year.

Colombia produces 1.800.000 tons of high quality plantain, however, only 24% is produced by small farmers (JIMENEZ, 2018), who are faced with the challenge of generating production with few technological tools.

Another limiting factor is the incidence of diseases like the black sigatoka (*Mycosphaerella fijiensis*) (ALVAREZ et al. 2013), which reduces bunch weight by up to 50% and if control measures are not applied, it can cause losses of 100% of production, due to the deterioration of the quality of the fruit. (STOVER, 1980; MARÍN; ROMERO, 1992, MARÍN et al. 2003; RODRÍGUEZ; GAVIRIA; CAYÓN, 2008). In addition, the control of the disease reaches 27% of the total cost of production. Black sigatoka produces small pinches that become greyish spots covering the entire leaf, decreasing the photosynthetic area and the plant's production capacity. Therefore, it is essential to detect the disease and generate an adequate control from the early stages of occurrence. For this, farmers in Colombia perform leaf removal processes seeking to reduce the incidence of the fungus; however, most decisions are made based on empirical visual reviews which results are not standardized with a measurement pattern.

An adequate prediction requires trained personnel, generating a technical criteria from a standardized metric, such as the Fouré scale; that divides the disease in 6 stages in accordance to the visual symptoms presented by the leaf affected. This standardization allows taking adequate control measures, however, the technical assistance service in Colombia has serious problems regarding coverage, quality, financing, institutional disarticulation, absence of information and little technological innovation (AGRONET, 2012; CASTAÑO, 2016; BONILLA, 2020).

In order to solve these problems, the National Agricultural Innovation System (SNIA) is created, under the law 1876 of 2017, which role is to generate support services to manage the development of capacities of agricultural processes and their articulation with the environment; proposing the access to information technologies as an approach (BONILLA, 2020).

In this sense, the technological tools for the diagnosis play a key role, in the state of the art,

classification methodologies have been proposed that could be part of tools for the diagnosis of diseases (CALVO, et al. 2014). Some techniques propose image processing to detect sick pixels (SINGH; MISRA, 2017), this also improves classification processes (CAMARGO; SMITH, 2009) where texture features are extracted as histogram of oriented gradients (HOG), speeded-up robust features (SURF), discrete transform Wavelet or Fourier (BEJARANO, et al. 2018) on color spaces HSV (hue, saturation, value), HSI (hue, saturation, intensity), TSL (tint, saturation, lightness), (VIPINADAS; THAMIZHARASI, 2017; SINGH; MISRA, 2017) to train classical machine learning techniques such as support vector machines (SVM)(ARISTISABAL, 2017), Bayes classifiers, decision trees, Kmeans, K-nearest neighbors (KNN)(MODENESE, 2008; PETRELLIS, 2015; POOJA et al. 2017; BALAJI, 2017).

Convolutional neural networks (CNN) have also been widely implemented in the disease detection processes (MOHANTY, et al. 2016). The Inception V3 network has been used to classify diseases in yucca (RAMCHARAN, et al. 2017). It is also common to use architectures such as AlexNet and Google that identify 26 diseases in 14 types of crops with a 99.35% efficiency (PRASANNA, et al. 2018) and the LeNet network reporting efficiencies of 98.67% classifying sigatoka.

However, all the techniques listed have two big problems. First, studies only validate their operation based on statistical tests in laboratory controlled conditions, do not contemplate validating the models in productive crops with real situations, the use and perception by the user is not verified either; if it is intended to generate an agricultural impact, it is necessary to carry out these validations. Second, there's no report of the construction of a database taken under a capture protocol that allows sampling all stages of the disease in real conditions (VARGAS, 2012; MODENESE, 2008). Nor is a labeling process reported with an agricultural criterion under a standardized metric such as the Fouré scale (OWOMUGISHA, 2019). For this reason, the classifiers listed only determine between healthy and diseased leaves, omitting that the disease presents different levels of severity, which requires a differentiated management.

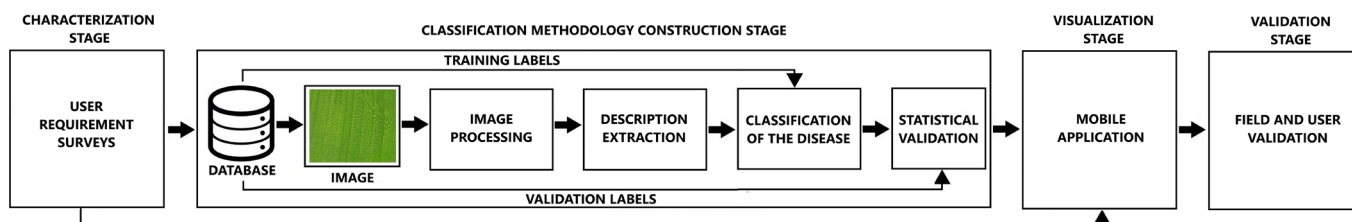
In this paper, a support system for monitoring black sigatoka disease is presented, introducing a robust system that is validated from the design of the methodology to the final stage with the user. The main contributions of this article are:

- Creation of a database with a capture and labeling protocol under expert criteria on four levels of severity.
- Design and validation of a methodology classifying black sigatoka on 4 levels of severity.
- Validation of the diagnosis system in real conditions with producers in the region.

## Materials and methods

In this section the methodology designed for the construction of the tool to diagnose black sigatoka is described in detail. First, a collection of requirements by the farmers of the department of Risaralda (Colombia's Midwest) was performed. Then, a classification system was designed where the model that generated the highest

efficiency of a set of methods proposed by the state of the art was used. The classification model was implemented in a mobile application that was socialized and validated with plantain producers from usability surveys. On Figure 1, the methodological diagram that will be explained in detail in this section is observed.



**Figure 1.** Explanatory diagram of the methodology proposed.

### Collection of user requirements

The support system is intended to be a tool for the farmer with a correct validation on field; therefore, it was necessary to characterize the needs of the agricultural environment. For this, a requirement collection process was performed with the plantain producers associations of the department of Risaralda registered in the Secretaries of Agriculture of Apia, Balboa, Belén de Umbría, la Celia, Marsella, Mistrató, Pereira, Santa Rosa de Cabal, Santuario, surveying 220 producers. Questions related to the access level, management and perception of information technologies were asked.

### Database

In the literature there are databases with images of the black sigatoka disease, however, they weren't built under an adequate capture and sampling protocol that allows to classify the leaves on different levels of severity. On the other hand, these have a reduced number of samples, ignoring the statistical relevance necessary to train machine learning models, or they were captured in laboratory settings, omitting all the noise that occurs when capturing a sample in real conditions. (VARGAS, 2012). Furthermore, the databases were not labeled with expert criteria from a standardized scale of visual patterns. Which makes it impossible to apply a structured methodology like the one proposed in this article, in order to replicate and compare similar methods. Due to this, an annotated database was built that fulfilled the required characteristics, for this, productive plantain crops were visited which characteristics are mentioned in Table 1.

**Table 1.** Description of the plantain farming systems visited to create the database.

Geolocation	Altitude (m)	Production system
4°52'09.7"N 75°45'24.4"W	1375	Plantain monoculture
4°53'59.5"N 75°43'59.1"W	1611	In association with coffee
4°58'56.5"N 75°39'39.9"W	1335	In association with coffee
5°13'24.0"N 75°52'36.6"W	1574	Plantain monoculture

The proposed study is intended to take samples in productive crops under real conditions, complying with user requirements. Therefore, a protocol was defined to allow capturing the patterns generated by black sigatoka and to control the noise produced by elements unrelated to the disease. Below, the components of the protocol are listed:

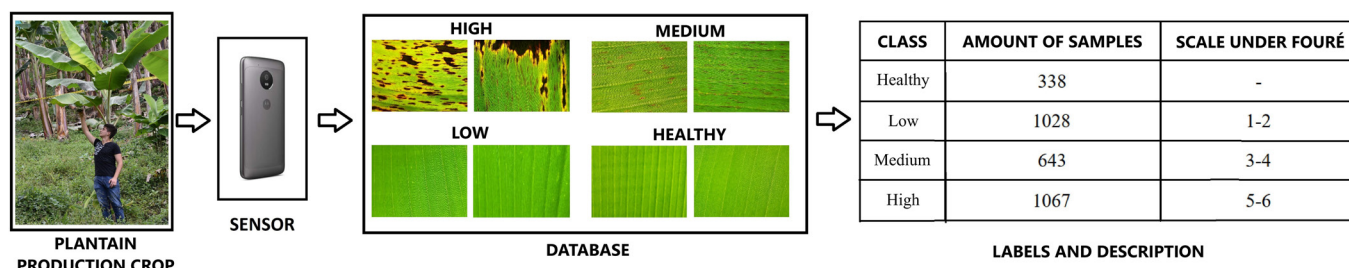
- Cell phone camera in automatic mode.
- Photographs of the plantain leaf taken against the light.
- Capture distance approximately 30 cm from the leaf.
- Sunny day from 10:00 am to 15:00 pm
- Avoid any aspect that is not the plantain leaf, such as mud, shadows, drops and background.

To guarantee a correct understanding of this procedure, an illustrative video is performed where the operation of the application is explained in detail and which link is presented in the following link: <https://www.youtube.com/watch?v=gW2YdBYwGpU&t=77s>.

Sampling was carried out with medium to low-end cell phones, which cameras have a 9.6MP (4128x2322) and 13Mp (3120x4160) resolution. This image size is adequate as it allows sampling all possible stages of black sigatoka.

The images were then labeled following the Fouré scale as a technical criterion. From this metric, 4 classes are generated according to the control actions to avoid economic damage. The early levels of infection demand cultural management (partial cuts of the leaf), without affecting economic performance. Here the learning system should give higher priority, as patterns are not easily detected.

The “Healthy” class corresponds to the absence of the disease. In this stage the farmer must carry out constant monitoring to warn in time the appearance of the disease. The “Low” class corresponds to stages 1 and 2 under Fouré. In this stage the farmer must perform low-cost control practices, with the purpose of avoiding a greater advance of the disease, at this point no economic losses are generated for the producer. The “Medium” class corresponds to stages 3 and 4 under Fouré. In this stage the farmer must remove the leaves of the affected parts more regularly and carry out chemical control. Finally, the high class corresponds to stages 5 and 6 under Fouré, at this point the leaf must be cut completely, generating economic losses. This class is easily detectable by the farmer; therefore, priority should be given to the appearance of the disease in the early stages (ITT2, 2016). Figure 2 presents a general diagram of the capture and labeling of the database.



**Figure 2.** Procedure performed for the creation of the database.

### Image processing

This methodological instrument processes the images with the purpose of highlighting the patterns produced by black sigatoka, for this an image-processing algorithm was implemented that increases the contrast of the pixels where the disease is found. In this, a Gaussian filter with sigma 5 is applied. Then a conversion from RGB to HSV is carried out and the V channel is selected, dividing the image into windows of 500x500 pixels. For each window, every pixel outside the  $\mu \pm 3\sigma$  range is considered sick and a gain of 3 is applied saturating its value. Finally, the windows are joined, obtaining the image in HSV, and they are also transformed to gray scale. In Figure 3 the result of applying the proposed algorithm is observed. The reader may consult (OPENCV, 2014) to see in detail the calculation and conversion process of color spaces.

### Descriptor extraction

In this stage, the main descriptors reported by the state of the art are computed to identify patterns in an image that has lighting variations as a result of an uncontrolled environment such as a crop (PRABUKUMAR et al. 2010; OWOMUGISHA, 2014). This process not only extracts the color properties in the image but also transforms space frequency, generating a robust process to lighting changes. This section is presented in Figure 4 and is explained below.

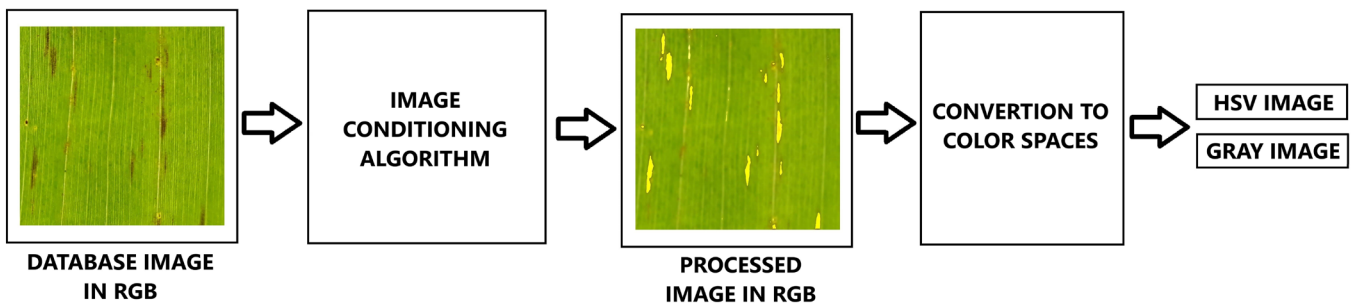


Figure 3. Process performed on the image processing stage.

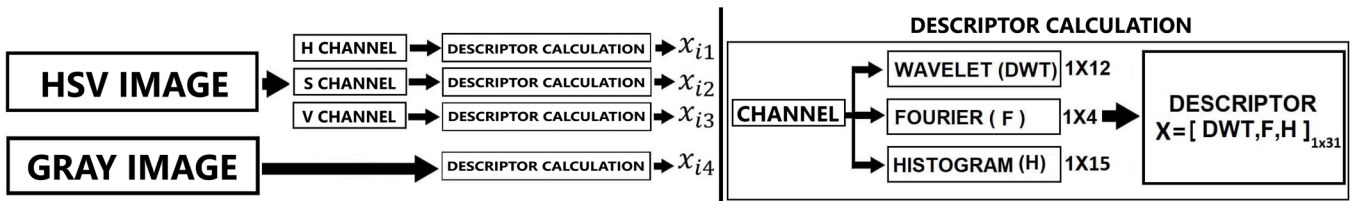


Figure 4. Process performed in the descriptor extraction stage.

The first descriptor is the Discrete Wavelet Transformed (DWT) in two dimensions using the fourth order Symlets family, from this the horizontal (j), vertical (m) and diagonal (u) coefficient matrixes are extracted (PYWAVELETS, 2019). The mean statistical moments are calculated for each matrix (u), standard deviation ( $\sigma$ ), bias ( $\delta$ ) and kurtosis ( $\kappa$ ) obtaining the descriptor  $DWT_{(j,m,u)}$  given by:

$$DWT_{(j,m,u)} = [u_j, u_m, u_u, \sigma_j, \sigma_m, \sigma_u, \delta_j, \delta_m, \delta_u, \kappa_j, \kappa_m, \kappa_u]$$

The second descriptor is the Fast Fourier Transformed in two dimensions (SCIPY, 2019), that compute the matrix (y). The same statistical moments of DWT are calculated to this matrix, thus obtaining the descriptor  $F_y$  given by:

$$F_y = [u_j, \sigma_j, \delta_u, \kappa_u]$$

The last calculation performed at this stage is a histogram with 15 regular intervals (bins) (SCIPY, 2019), from which the descriptor H is obtained. Finally, the  $x_i$  data vector is obtained, resulting from concatenating the three proposed descriptors for each proposed representation system, in this manner:

$$x_i = [DWT_{(j,m,u)}, F_y, H]_{1 \times 31}$$

Where i is the color channel to compute.

### Classification of the disease

In this stage, the aim is to establish a classification method with the best relationship between efficiency and computational cost that allows it to be implemented in a user interface. For this the *Ensemble Boosted Trees* machine learning is proposed using the *adaptive boosting* methodology, with 30 learning cycles altogether (SCIKITLEARN, 2019). This classifier is chosen because the assembled models have obtained high efficiencies in classifying diseases. Another aspect to be taken into account is the imbalance possessed by the database due to the exponential growth of the disease. In these cases the state of the art suggests using weighted classification methods, therefore, a hierarchical classifier is used. For this, a bi-class model is trained to classify samples in “High” and “Not High” status. Then another model with a multiclass approach is trained using a number of balanced samples to generate labels “Medium,” “Low,” “Healthy.” This avoids generating classifiers biased towards the classes with the highest number of samples.

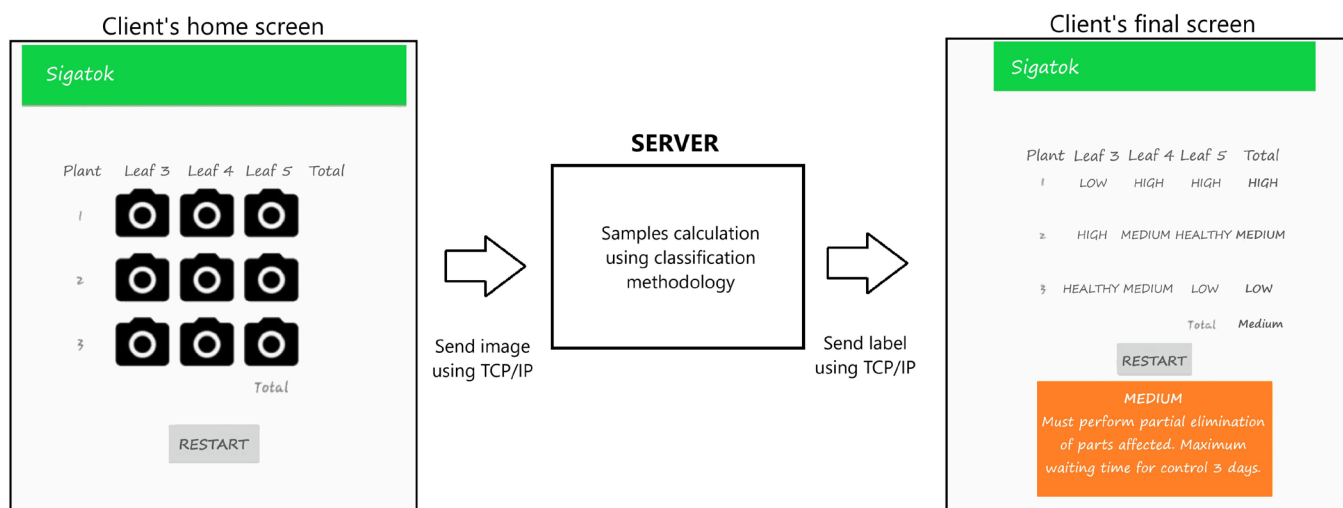
### Mobile application

A mobile application for Android is designed, with client-server architecture, using Kotlin language. The design was carried out following the requirement surveys; therefore, the application presents an intuitive design, simple to use and low computational cost.

The client (mobile device) has a visual interface for interaction with the users, where sampled images are entered with the stipulated capture protocol. The application, takes as reference the Stover sampling process (CUEVAS, et al. 2015) to select the images to capture, for this reason it requests to sample the leaves 3,4 and 5.

Every image is sent to a server using the TCP/IP protocol through the 8080 port, where they're processed using the classification methodology introduced in this article. The predictions obtained are sent to the client, presenting the status of the analyzed region and an agronomic management recommendation is made to the farmer. In Figure 5 the application's general diagram is observed.

This architecture allows running the application on low-end devices, since the client is only responsible for the sampling and presentation of results, while the processing or the information is performed in the server.



**Figure 5.** Graphic and functional description of the mobile Application.

## Validation

The support system for the detection of Sigatoka needs to be validated from three aspects: efficiency of the classification methodology, operation of the tool in the field and perception by agricultural users. Each section will be explained below.

### Statistical validation

In this stage the choice of the classification model is corroborated, comparing our classifier to machine learning methods widely used in the state of the art in disease detection, such as SVM, KNN, tree decision. In addition, the machine learning method is compared against a deep learning methodology that contemplates the use of convolutional neural networks. This modern method, divides the image in small segments of 96x96, to be classified by a convolutional neural network LeNet (AMARA, et al. 2017). The previous allows having a map of severity of the leaf, later a decision tree is used that weights the frequency of appearance to obtain the overall severity of the image. The Lenet architecture was chosen because of the good results generated in black sigatoka classification on state of the art works (JIHEN, et al. 2017).

All classification tests are validated using cross-validation with 100 folds where the samples are fractioned in a random manner following a Gaussian distribution. In this article 70% of the data is used for training and the remainder for validation. All the tests are carried out in a computer with an Intel Core i7 5500U processor and 8Gb of RAM memory using Python language.

### Field Validation

Once the classification methodology is verified and implemented in the application, a socialization process is followed, where the operation of the system is validated on field, for this, workshops were arranged with the producer associations in the abovementioned municipalities and surveys were conducted with 56 small plantain producers. Attendees were explained the capture protocol with the system generated, the Stover sampling procedure and they were proposed to validate the status of a crop sector using the diagnostic tool, in total farmers captured 192 samples, in the results section the confusion matrix resulting from this test is presented.

### User validation

After performing the use and image capture tests, TAM (technology acceptance model) surveys were applied allowing to evaluate the utility, effectiveness, productivity and performance of the support system.

## Results and discussion

In this section, the results obtained in the collection of user requirements, the design of the classification methodology and the validation of the system with farmers in the region are presented.

Characterization revealed that 91% of the surveyed or some member of their family, owns a device with a camera and Internet access. Also, 86% uses mobile software, however, only 20% has used mobile applications related to agriculture.

96% of the surveyed considers easy to have an application that selects the images and to have a button to open the camera. While 97% consider useful a mobile application that indicates the level of disease and helps making management decisions.

They were also asked what processes should the system support through the new information technologies, to which 33% considered useful to indicate the status of the disease and 36% generate recommendations for taking measures in accordance to the diagnosis. These low results reveal the high ignorance about technologies and corroborate the need to include agricultural tools that improve this perception. These technological elements should be developed for use in mobile devices already owned by the farmers, which facilitates the use and avoids incurring in high technology acquisition costs. The graphs summarizing the results obtained are presented on Figure 6.

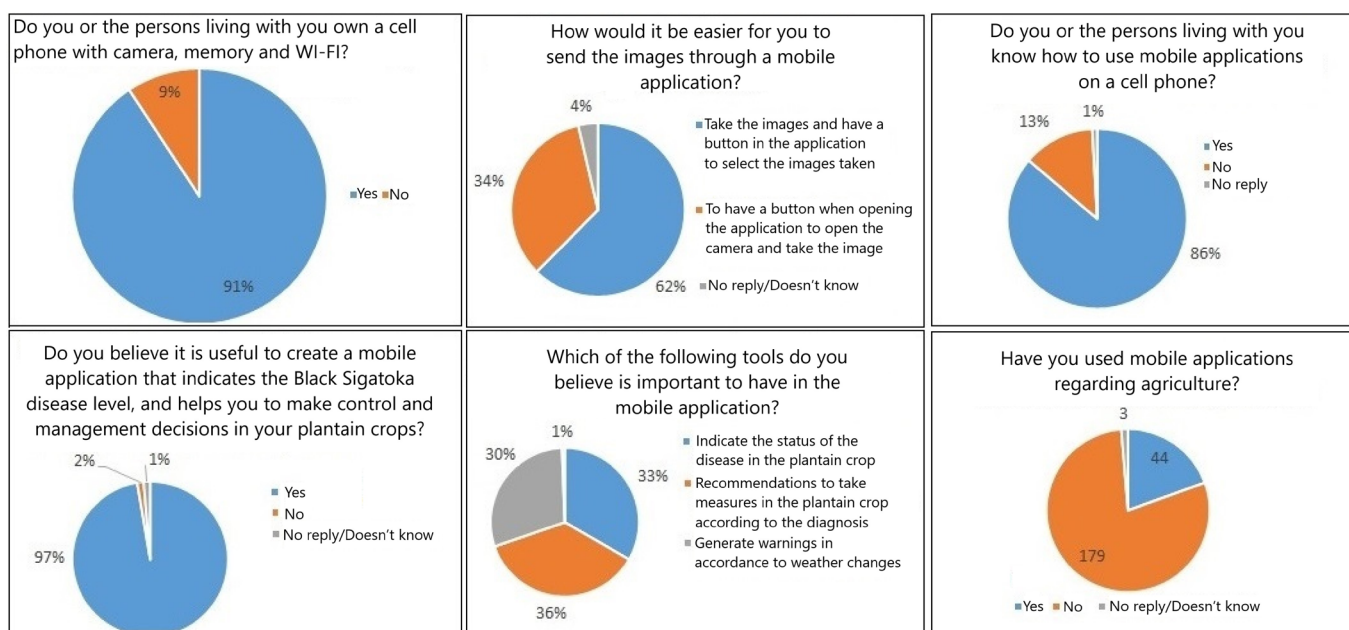


Figure 6. User requirements survey results.

The characterization surveys confirm that the farmers have basic technological equipment and Internet access; they also have a great disposition and know the need to include technologies to support the country. However, ignorance of technological tools is common. Support systems need to be made from agricultural needs, easy access, simple use, ease of installation on mobile devices.

Regarding the classification methodology, on Table 2, the main diagonals of each classifier and also the computing time needed to carry out the crossed validation of each model are presented.

The results show that the proposed model presents a greater efficiency on the High and Healthy classes, but the SVM and the CNN LeNet surpass it on the Medium class. Despite the foregoing, the Ensemble Boosted Tree classifier generates the most consistency in all the classes and generates the best separation between the Low and Healthy class.

Separability in the first classes is fundamental because it's the point of the disease where the farmer can apply all the adequate management tools and avoid economic losses from happening. If the disease is detected in an advanced stage like the medium class, control actions are less effective.

The disease advances rapidly in tropical rainy weathers like Colombia, therefore, few leaves are in a totally healthy state in the region's crops. This leads to the database presenting imbalance between its classes and requires applying techniques to avoid training biased models. The use of a hierarchical classification model was configured as an effective method to treat imbalance and allowed obtaining the best results in the detection of the disease.

The CNN LeNet presents the lowest efficiency in the healthy class, in this study this efficiency is to the fact that the differentiating traces between the low and healthy class are barely perceptible to the human eye and the resizing process produces loss of information,

which makes it impossible for the network to generate separability. Given the importance of separation between the low and healthy class the CNN LeNet isn't adequate for the help system proposed in this article. In addition, it is highlighted that using a set of images with higher resolution for training increases the number of weights and operations that the model must compute, making its use unfeasible in a low-cost device such as a smartphone (NUMERENTUR, 2018).

After verifying the laboratory results, field validation is performed. The captures made by the farmers which tested the system in productive crops, corroborate efficiency in the laboratory. The tool correctly classified 82.02% of the images processed. On Table 3 the confusion matrix is presented presenting in detail the test's result.

**Table 2.** Confusion matrixes main diagonal comparison of the classification models.

	<b>Ensemble Boosted Hierarchical Tree [%]</b>	<b>SVM [%]</b>	<b>Decision Tree [%]</b>	<b>KNN [%]</b>	<b>CNN LeNet [%]</b>
<b>High</b>	95.41±1.36	80.20±2.35	90,59±1.79	76.36±2.33	90.32±2.3
<b>Medium</b>	74.47±3.37	71.68±3.01	68,27±5.20	55.00±3.24	85.08±1.4
<b>Low</b>	72.47±2.83	78.61±2.36	84,95±3.92	86.92±1.76	80.54±2.31
<b>Healthy</b>	78.54±4.52	36,33±5,20	36,03±7.8	18.63±3.29	14.34±1.35
<b>Total</b>	<b>80.58±1.05</b>	<b>73.38±1.33</b>	<b>78.02±1.22</b>	<b>69.06±1.18</b>	<b>72.05±1.7</b>
<b>Time[s]</b>	245±3.29	48.09±1.21	15.71±0.64	16.34±0.53	421.90±2.1

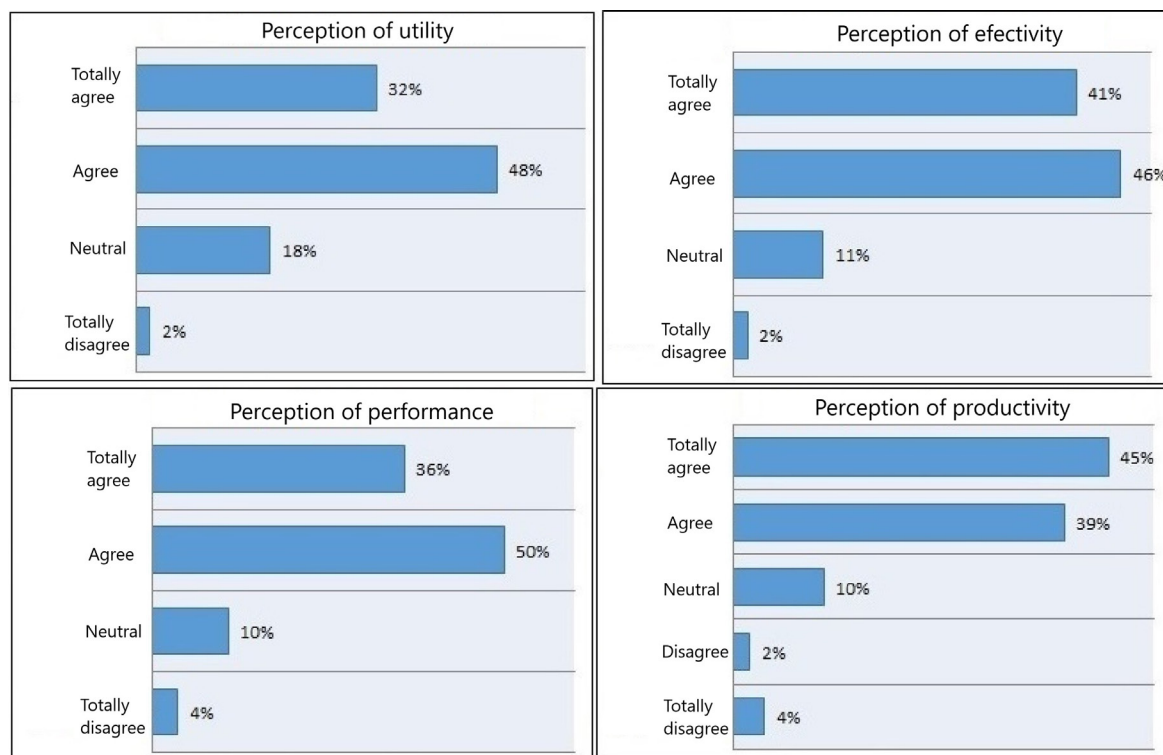
**Table 3.** Confusion matrix of field validation.

	<b>High</b>	<b>Medium</b>	<b>Low</b>	<b>Healthy</b>
<b>High</b>	76.19%	23.80%	0	0
<b>Medium</b>	0	80.76%	15.38%	3.84%
<b>Low</b>	0	3.33%	86.67%	10%
<b>Healthy</b>	0	0	23.07%	76.92%

The support system was designed bearing in mind all field conditions, this allowed achieving similar efficiencies between laboratory validation and field tests. It is necessary that the classification methodologies have verification stages in the field and from the user where the efficiency and operation are corroborated under all possible real conditions. The above allows this research to be useful for this sector, it becomes a product with

high potential of use and generate impact on the target population of the research. The TAM survey results reveal a high acceptability of the support system, 80% of the surveyed agree or highly agree with the utility perception, 87% has a high perception of affectivity, 87% consider a high perception in terms of performance, finally 84% positively graded the perception of productivity; as presented on Figure 7.





**Figure 7.** User validation survey results.

## Conclusions

Mass use of smart phones, including their penetration in agricultural contexts, and the vertiginous advance in its software capacities, have made them attractive for the government programs, such as agricultural technical assistance, to be supported by mobile applications that maximize their utility and make the resources destined for this purpose more efficient, primarily in countries like Colombia where access to certain rural areas is very difficult.

In this article, a support system is presented for the black sigatoka detection process that allows classifying stages of the disease from its earliest stages, which would allow a farmer to take corrective action, improving product quality and avoiding economic losses. This study contributes with the creation of a public database, labeled by an expert criterion on 4 levels of severity. This generates a new approach for classification studies where more information for the control of the disease is produced. It also allows researchers on machine learning, to have truthful information that allows research to be directed to useful products for the farmer.

The classification method proposed allows maintaining an adequate efficiency in the first stages of the disease, which is not achieved using other classification techniques proposed in the state of the art, as observed in Table 3. Detecting the disease in the first stages is of vital importance, since it allows the farmer to take preventive action, avoiding economic losses.

The Ensemble Boosted Tree model, following an adaptive boosting methodology, generated an  $80.58\% \pm 1.05\%$  efficiency. As future work, it is necessary to evaluate the performance of methodologies such as bagging, in addition to comparing it regarding methods composed by different types of classifiers.

Given the problems of technical assistance in the Colombian territory, technological elements become an important tool to improve the productive process. However, these tools do not substitute the evaluation and diagnostic process performed by an expert. The application presented generates a suggestion from a diagnosis an  $80.58\% \pm 1.05\%$  efficiency. This becomes an important help for low-income farmers who have no access to a technical assistance service.

## Acknowledgment

The authors express their gratitude to the Tecnológica University of Pereira (Colombia), to the Ministry of Science, Technology and Innovation for financing the project “Expert system for the automatic detection of black sigatoka produced by the fungus *Mycosphaerella fijiensis* in commercial plantain crops for the generation of early warnings using information technologies” contract 160-2018” code 511-3-024-15 and to the plantain producers, specially to mister German Pulgarin for providing the necessary support for the development of the project.

## References

- AGRONET. **Minagricultura pone en marcha programa de asistencia técnica para 450 mil pequeños y medianos productores**. Bogotá, 2012. Disponível em: <https://www.agronet.gov.co/Noticias/Paginas/Noticia512.aspx>>. Acesso em: 1 jul. 2020.
- ALVAREZ, E.; PANTOJA, A.; GAÑAN, L.; CEBALLOS. **La sigatoka negra en plátano y banano**. Roma: FAO, 2013. Disponível em: <http://www.fao.org/3/as089s/as089s.pdf>. Acesso em: 15 ago. 2019.
- AMARA, J.; BOUAZIZ, B.; ALGERGAWY, A. A deep learning-based approach for banana leaf diseases classification. *In: DATENBANKSYSTEME FÜR BUSINESS, TECHNOLOGIE UND WEB, 2017, Stuttgart. Workshopband [...]*. Bonn: Köllen Druck und Verlag GmbH, 2017. p.77-88.
- ARISTIZABAL, J. **Evaluación de un método de aprendizaje supervisado para la detección de las enfermedades, antracnosis y *Phytophthora infestans* en cultivos de fruta de Risaralda**. 2017. Tesis (Ingeniería Electrónica) - Universidad Tecnológica de Pereira, Pereira; 2017.
- BALAJI, N. A survey paper on crop disease identification and classification using pattern recognition and digital image processing techniques. **IOSR Journal of Computer Engineering**, India, p.14-17, 2017.
- BEJARANO, A.; CALVO, A.; HENAO, C. Descriptores espacio-frecuencia para identificación automática de patrones de textura en productos textiles utilizando aprendizaje supervisado. **Ciencia e Ingeniería Neogranadina**. Bogotá, v.28, n.2, p.63-82, 2018.
- BONILLA, S. **Los cambios del Acceso de la asistencia técnica agropecuaria**. Bogotá: Agronegocios, 2020. Disponível em: <https://www.agronegocios.co/analisis/sara-patricia-bonilla-2857810/los-cambios-del-acceso-a-la-asistencia-tecnica-agropecuaria-2962834>. Acesso em: 1 Jul. 2020.
- CALVO, A.; GUZMÁN, M.; GUZMÁN, J. Considerations about application of machine learning to the prediction of sigatoka disease. *In: WORLD CONFERENCE ON COMPUTERS IN AGRICULTURE AND NATURAL RESOURCES, 2014, San Jose de Costa Rica. Anais [...]*. Costa Rica: University of Costa Rica, 2014.
- CAMARGO, A.; SMITH, J. An image-processing based algorithm to automatically identify plant disease visual symptoms. **Biosystems Engineering**, London, v.102, p.9-21, 2009.
- CASTAÑO, G. **Calidad del servicio de asistencia técnica agropecuaria: análisis de dos casos en Boyacá**. Tesis [Maestría]. 2016. Universidad Nacional de Colombia, Bogotá. 2016.
- CUEVAS, J.; FLORES, C.; GUERRERO, J. Evaluación de un sistema de apoyo para el diagnóstico de Sigatoka Negra. **Research in Computing Science**, Ciudad de Mexico, v.108, n.1, p.73-80, 2015.
- DAVIS, F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS Quarterly**, Minneapolis, v.13, n.3, p.319-340, 1989.
- ESPINAL, C.; MARTÍNEZ, H.; Y PEÑA, Y. La cadena de plátano en Colombia: una mirada global de su estructura y dinámica 1991-2005. Bogotá: Ministerio de Agricultura y Desarrollo Rural, 2005. (Documento de trabajo, 102).
- ITT2 - Institut Technique Tropical. **Manejo de las sigatokas del banano**. 2016. Disponível em: [http://www.it2.fr/wp-content/uploads/2014/11/DOC\\_IT2\\_2014-Fiche-manuel-BGM-n1-Cercosporioses\\_ESP\\_BD1.pdf](http://www.it2.fr/wp-content/uploads/2014/11/DOC_IT2_2014-Fiche-manuel-BGM-n1-Cercosporioses_ESP_BD1.pdf). Acesso em: 14 jun. 2020.
- JIMÉNEZ, F. **Sector bananero crece en exportaciones, área cultivada y productividad**. Bogotá: El colombiano, 2018. Disponível em: <https://www.elcolombiano.com/negocios/bananeros-presentan-buen-balance-en-2017-FY8463412>. Acesso em: 15 Ago. 2019.
- MARÍN, D.; ROMERO, R. **El combate de la Sigatoka negra**. San José: C.R, Departamento de Investigaciones, 1992. p.21. (Boletín, 14)
- MARIN, D.; ROMERO, R.; GUZMÁN, M.; SUTTON, B. Black sigatoka: an increasing threat to banana cultivation. **Plant Disease**, New York, v.87, n.3, p.208-222, 2003.
- MODENESE, S. **Processamento digital de imagens para identificação da sigatoka negra em bananais utilizando análise de componentes principais e redes neurais artificiais**. 2008. Tese (Doutor em Agronomia) - Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, 2008.
- MOHANTY, P.; HUGUES, D.; SALATHÉ, M. Using deep learning for image-based plant disease detection. **Frontiers in Plant Science**, Lausanne, v.7, p.1419, 2016.
- NUMERENTUR. **CNN-RNA Convulucional**. 2018. Disponível em: <http://numerentur.org/convolucionales/>. Acesso em: 11 Jun. 2020.

- OPENCV - Open Source Computer Vision Library. **Image processing in OPENCV, changing colorspace**. 2014. Disponível em: [https://docs.opencv.org/3.0-beta/doc/py\\_tutorials/py\\_imgproc/py\\_colorspaces/py\\_colorspaces.html](https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_colorspaces/py_colorspaces.html). Acesso em: 21 Ago. 2019.
- OWOMUGISHA, G.; QUINN, J.; MWEBAZE, E.; LWASA, J. Automated vision-based diagnosis of banana bacterial wilt disease and black sigatoka disease. *In: INTERNATIONAL CONFERENCE ON THE USE OF MOBILE ICT IN AFRICA*, 1, 2014. Stellenbosch. **Proceedings [...]**. Disponível em: <https://www.semanticscholar.org/paper/Automated-Vision-Based-Diagnosis-of-Banana-Wilt-and-Owomugisha-Quinn/73d65fa4b006dc25515f2f026510291671e6f039>. Acesso em: 17 Ago. 2019.
- PETRELLIS, N. Plant disease diagnosis based on image processing, appropriate for mobile phone implementation. *In: INTERNATIONAL CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES IN AGRICULTURE*, 7.. 2015, Kavala. **Proceedings [...]**. Disponível em: [http://ceur-ws.org/Vol-1498/HAICTA\\_2015\\_paper28.pdf](http://ceur-ws.org/Vol-1498/HAICTA_2015_paper28.pdf). Acesso em: 16 Ago. 2019.
- POOJA, V.; RAHUL, D.; KANCHANA, V. **Identification of plant leaf diseases using image processing techniques**. Chennai: IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR), 2017. p.130-133. Disponível em: <https://ieeexplore.ieee.org/document/8273700>. Acesso em: 17 Ago. 2019.
- PRABUKUMAR, M.; BALAMURALI, K.; KAMALAKANNAN, J. Image processing and pattern classification technique in a machine vision system that identifies and classifies the plant diseases based on the visual symptoms. **International Journal of Advanced Research in Computer Science**, Mandsaur, v.1, n.4, p.302-307, 2010.
- PYWAVELETS. **Discrete wavelet transform. Api reference**. 2019. Disponível em: <https://pywavelets.readthedocs.io/en/latest/ref/dwt-discrete-wavelet-transform.html>. Acesso em: 25 Ago. 2019.
- RAMCHARAN, A.; BARANOWSKI, K.; MCCLOSKEY, P.; AHMED, B.; LEGG, J.; HUGHES, D. Deep learning for image-based cassava disease detection. **Frontiers in Plant Science**, Lausanne, v.8, p.1852, 2017.
- RODRÍGUEZ, P.; SALINAS, D. Physiological effect of *Mycosphaerella fijiensis* in banana leaves. **Agronomía Colombiana**, Bogotá, v.26, n.2, p.256-265, 2008.
- RUIZ, M.; URUEÑA, M. Situación actual y perspectivas del mercado del plátano. Bogotá: Economic Research Service Ers Programa MIDAS, 2009.
- SCIKIT-LEARN. Ensemble methods. *In: Scikit-learn: Machine Learning in Python*. 2019. Disponível em: <https://scikit-learn.org/stable/modules/ensemble.html>. Acesso em: 9 Set. 2019.
- SINGH, V.; MISRA, A. Detection of plant leaf diseases using image segmentation and soft computing techniques. **Information Processing in Agriculture**, Amsterdam, v.4, n.2, p.41-49, 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2214317316300154>. Acesso em: 16 Ago. 2019.
- SIOC -Sistema de Información de Gestión y Desempeño de Organizaciones de Cadenas. **Indicadores de plátano**. Bogotá: Ministerio de Agricultura y Desarrollo Rural, 2014.
- SIOC -Sistema de Información de Gestión y Desempeño de Organizaciones de Cadenas. **Cifras sectoriales**. Bogotá: Ministerio de Agricultura y Desarrollo Rural, 2018.
- STOVER, R. Sigatoka leaf spots of bananas and plantains. **Plant Disease**, Beltsville, v.64 n.8, p.751, 1980.
- THE SCIPY COMMUNITY. **SciPy.fft**. 2019. Disponível em: <https://docs.scipy.org/doc/numpy/reference/generated/numpy.fft.fft.html>. Acesso em: 27 Ago. 2019.
- THE SCIPY COMMUNITY. **Numpy.histogram**. Disponível em: <https://docs.scipy.org/doc/numpy/reference/generated/numpy.histogram.html>. Acesso em: 1 Set. 2019.
- VARGAS, E. **Monitoreo asistido de plantaciones agrícolas: Técnicas de procesamiento digital de imágenes e inteligencia artificial para diagnosticar y controlar la incidencia de la sigatoka negra (*Mycosphaerella fijiensis* var *difformis* Morelet) en plantaciones de banano y plátano curraré**. 2008. Dissertação (Mestrado) - Instituto Tecnológico de Costa Rica Sede Regional de San Carlos, San Carlos, 2008.
- VIPINADAS, M.; THAMIZHARASI, A. Detection and grading of diseases in banana leaves using machine learning. **International Journal Of Scientific & Engineering Research**, New Delhi, v.7, n.7, p.916-924, 2016.