

Black Sigatoka Classification Using Convolutional Neural Networks

Cristian A. Escudero, Andrés F. Calvo, and Arley Bejarano

Abstract—In this paper we present a methodology for the automatic recognition of black Sigatoka in commercial banana crops. This method uses a LeNet convolutional neural network to detect the progress of infection by the disease in different regions of a leaf image; using this information, we trained a decision tree in order to classify the level of infection severity. The methodology was validated with an annotated database, which was built in the process of this work and which can be compared with other state-of-the-art alternatives. The results show that the method is robust against atypical values and photometric variations.

Index Terms—Black sigatoka, convolutional neural network, decision tree, plant disease detection.

I. INTRODUCTION

Farming is one of the most important activities around the world and is the economic foundation of many developing countries [1]. In Colombia, 26% of the entire banana crop is produced by small farmers [2], who are challenged by the presence of diseases that cause 16% of the crop to be lost. Currently, there are no low-cost tools that allow the diagnosis of the disease in its early stages.

One pathogen that has a significant impact on banana crops is black Sigatoka. This disease destroys leaf tissue, inducing visual patterns [3], which can be classified using computer vision techniques that represent new methods of disease detection. Visual representation systems such as HSV, TSL, LAB, and YCbCr commonly use Gaussian filters that soften images; then methods including histogram analysis and OTSU, among others, can be used to threshold images [4]. However, these developments are validated by laboratory testing, under controlled conditions, ignoring lighting conditions growing crops where these algorithms cannot perform properly. Other sturdier methodologies use classic machine learning techniques to perform feature extraction processes on and classification of the objects of study. The authors of [5] performed threshold operations and color space dilatation using YcbCr, which is necessary for color feature extraction, and then used ANFIS models and support vector machines to classify images of leaves infected with black Sigatoka. These models have a detection rate of 100%; however, the database only has a small number of samples, which generates biased classifiers.

Other studies defined different types of illnesses [6], for example, followed the same preprocessing, feature extraction, and classification steps to train a k-nearest neighbor algorithm to classify bacterial blight, *Alternaria alternata*, antracnose, and cercospora. In general, the use of wavelet transform, Fourier transform, HOG, SIFT, and SURF techniques in order to train support vector machines, k-neighborhoods, and neural networks is common [7], [8]. However, these methodologies are limited to establishing the presence or absence of the disease; processes in the field are required to detect the progress of infection by the disease and to perform effective control in its early stages. Convolutional neural networks (CNNs) are powerful tools that enable precision agriculture applications such as leaf and stem counts, leaf size measurements, root localization, and plant recognition. Due to the importance of the farming process, the most frequently used application is disease detection [9]. One study [10] trained an Inception-v3 network to classify infections in yucca plants, accomplished 93% of efficiency. The authors of [9] used a public domain dataset to train the AlexNet and GoogleNet architecture to identify 26 diseases that impact 14 types of corps, accurately identifying disease in 99.35% of cases. Even though the reported accuracy rates for these studies are high, the tests were made under laboratory conditions where there were no disturbances caused by image noise. These classification models must satisfy all the disturbances which it would generate sturdy systems with high implementation possibilities in lands.

In the field of Sigatoka detection, some studies have shown that the use of LeNet architecture trained with RGB and GRAY model samples can detect disease 98% of the time [3]. A further study introduced the application of pre-training networks such as ResNet50 and Inception-v3, which generated efficiencies of 99.9% [11]. Even though those efficacies have high accuracy, these methodologies only detect the disease while skipping metrics that define different levels of severity like the Four éscale [12] and each stage must carry a different action control. Disease detection approaches using CNNs classify diseases using images with a resolution of 64×64 to 150×150 [13], whereas current mobile devices have cameras with average resolutions of 4160×3120 . Therefore, it is necessary to develop a suitable methodology for deep learning methods. Nevertheless, these developments show the viability of implementing methods of disease detection using digital images at low computational cost, allowing the potential advancement in mobile devices of easy access for small farmers.

This article describes a method of detecting different levels of severity of black Sigatoka infection using a methodological variation of the LeNet CNN architecture. The training of the network parameters was undertaken by

Manuscript received October 25, 2019; revised October 11, 2020.

Cristian A. Escudero are with the Faculty of Engineering, Tecnológica University of Pereira, Pereira, Risaralda, Colombia (e-mail: cristian-escuder@utp.edu.co, afcalvo@utp.edu.co, abjarano@utp.edu.co).

using a database created under real conditions, generating all the cases for field implementation. This study is in two parts: in Section II, we describe the contents of each stage including the methodological development of machine learning models, the construction of the annotated database, and the statistical validation processes. In Section III, we describe our experiments and discuss the results. In Section IV, we present the conclusions of this study and its benefits.

II. MATERIALS AND METHODS

This section presents the proposed methodology for training a CNN to classify black Sigatoka infection according to different levels of severity. Fig. 1 shows the classification and training processes.

A. Classification Models

To perform the disease classification, we proposed three structured steps that allowed us to detect the disease at different levels of severity.

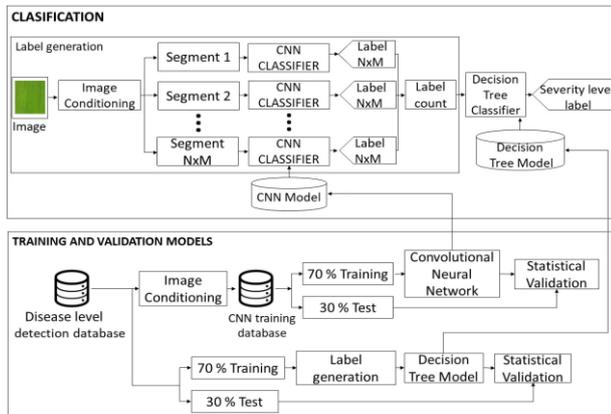


Fig. 1. Methodological diagram.

B. Image Conditioning for Classification

This step is necessary because classification in CNNs requires images with a resolution approximating 96×96 pixels [13]. Consequently, we segmented the signal into small windows that were classified using deep learning models. Even more when the captures made by the database construction are in high resolution. Fig. 2 shows the procedure for conditioning an image. This process segments the image into windows of 500×500 pixels and then resizes it to 96×96 pixels. This window size was chosen for segmentation since allows the information of the studied objective.

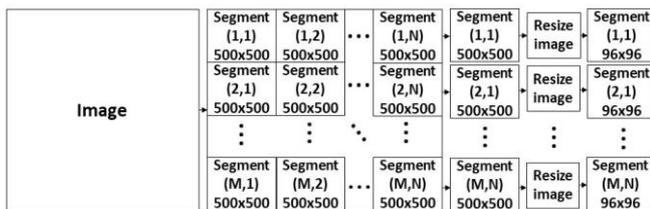


Fig. 2. Image conditioning block diagram for classification.

C. Convolutional Neural Network Model

The proposed classifier used the LeNet CNN topology [14], which consists of feature extraction and classification stages. Feature extraction is performed by a convolution

layer. This applies filters in a slider window form that computes the dot product between the kernel and the image, giving as result a bidimensional activation map that extracts morphological patterns such as curves and edges. This information is important for disease description. The first and third convolution layers used a kernel size of 5×5 and 3×3 . In addition, they had ReLU activation functions that introduced a non-linearity to the network, generating significant separability [15]. Equation (1) shows the convolution calculation.

$$Y_j = g\left(b_i + \sum_i K_{ij} * Y_i\right) \quad (1)$$

where Y_j is the neuron output and j is the matrix from the dot product between the image Y_i and the kernel convolution K_{ij} . The term b_i corresponds to the bias parameter of each neuron, and g is the activation function. The second and fourth layers perform a sub-sampling according to the max pooling process [16], which reduces the image size and generates invariance to potential rotations and input translation.

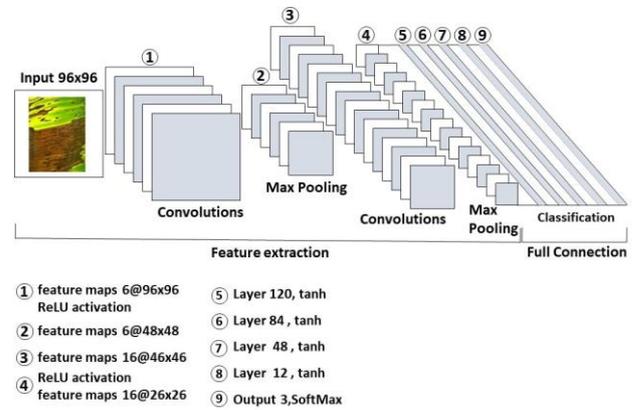


Fig. 3. LeNet topology.

Finally, we flattened the image, getting a feature vector that is classified through connected layers as shown in Fig. 3. These connection layers are numbered from 5 to 8 and have tanh activation functions to raise the gap of the neuron weights. The last layer uses a softmax activation function [17] that estimates the class membership likelihood (“high,” “middle,” or “low”).

D. Generation, Count, and Classification of Labels

This stage classified each image segment using the CNN, generating a set of $N \times M$ labels that correspond to “high,” “middle,” and “low” severity levels. The set represents a separate coding for each illness state that is characterized through a class count. These quantities are organized in a $X_{1 \times 3}$ vector, which generates a classifiable pattern from a decision tree. Fig. 4 shows the algorithm used to obtain the infection severity level of the banana leaf base in the classification count given by the CNN.

E. Training and Validation Models

As explained in the methodology section, it is necessary to use two learning models to detect the disease. Therefore, a CNN model and a decision tree must be trained. The CNN training was performed using the reduced-images database that is described in the next section. The proposed LeNet network learned the weights and biases by using the Adadelta algorithm; we then used the categorical cross-entropy loss

function, and finally, we evaluated the model by following the accuracy metric. To train the decision tree, we used the database presented in the next section. The model was trained with four nodes by using the Gini diversity index optimization model without pruning.

We evaluated our models using a cross-validation strategy that divided the database in k -folds. Each fold was divided into two parts: 70% of the data were used for training and 30% for testing. Nevertheless, generate the confusion matrix for each classifier. This process applied a Monte Carlo analysis, where the stop criterion was defined by

$$\|diag(\mathbf{M}_k) - diag(\mathbf{M}_{k-1})\|_2 < th \quad (2)$$

where M_k is the confusion matrix at iteration k and th is the error threshold.

```

// Label Counting and Classification
Input: NxM image segments
Output: Severity level
Start
  For i = 1 until i ≤ NxM
    Etiqueta(i) = CNN(Segment(i))
    If Label(i) == "High"
      CounterHigh = CounterHigh + 1
    If Label(i) == "Medium"
      CounterMedium = CounterMedium + 1
    If Label(i) == "Low"
      CounterLow = CounterLow + 1
  Enf for
  x = [CounterHigh, CounterMedium, CounterLow]
  Severity level = DT(x)
End
// CNN() = Convolutional neural network classifier
// DT() = Decision Tree classifier
    
```

Fig. 4. Labels count algorithm.

F. Database for Severity Level Detection

In the literature, there are databases with many images of Sigatoka infection. Unfortunately, these databases were not created according to a capture and sampling protocol that allowed the classification of leaves into different levels of infection severity. In addition, these databases have a reduced number of samples, avoiding the statistical relevance needed for the application of machine learning models. In some cases, the authors were not allowed access to databases. This made it difficult to apply a structured methodology like the one proposed in this article with the purpose of replicating findings and comparing methods. We sampled images from banana harvests in the Risaralda Department (Colombia); these were labeled with three severity levels according to the Fouré scale. The samples were taken using cellphones with camera resolutions of 9.6 MP (4128 × 2322) and 13 MP (3120 × 4160). Fig. 5 shows some examples of the base and the population per class.

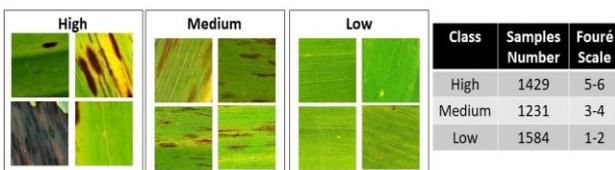


Fig. 5. Database.

To train the CNN model, it was necessary to create a second database. This base was constructed with the image segments obtained through the conditioning process, explained in Section II.A. We then obtained 96 × 96 pixel

sub-images. Next, these images were deparated and labeled according to the Fouré scale. As a result, we secured 4244 samples. Fig. 6 shows some examples from the database and the population per class. The databases described can be found here: <https://sites.google.com/a/utp.edu.co/black-sigatoka-disease-database/>.

III. RESULTS

The presentation of the results of this study is divided into two parts. First, we show the performance of the classification methods in detecting the state of infection by the disease in different segments of the leaf. Second, we present the results obtained by validating the decision tree to classify the stage of infection in the whole banana leaf using the classifiers proposed in the first stage. We used the validation approach described in this section to obtain the confusion matrices of our method and other studies.

Detection of Infection Levels in Different Segments of the Leaf

To understand the scope of the proposed method, we conducted a comparative analysis of our model and other state-of-the-art methodologies. We used the CNN Inception-v3 [11] and SVM [7] methods, which have been successfully applied in disease detection. In addition, the SVM is a methodology widely used in classification problems and has demonstrated high performance in classifying patterns. The SVM configuration for training is a polynomial kernel and a feature descriptor with wavelet transform in the RGB and HSV color spaces.

To detect disease with the CNN Inception-v3 approach, we used this network to extract the features vector and then, it is connected to a final layer to classify. In Table I, we show the traces of the recognition confusion matrices for the detection of black Sigatoka in a banana leaf. The CNN LeNet method provides the best disease detection results with an average detection rate of approximately 90.03% ± 1.73%, which is substantially higher than those achieved using the other methods. On the other hand, the CNN Inception-v3 method alone performs relatively poorly; however, it can still obtain a detection rate of 78.79% ± 1.6%, whereas the SVM presented with a performance of approximately 86.16% ± 2.1%.

TABLE I: A COMPARISON OF CLASSIFIERS

	CNN LeNet	CNN Inception-v3	Support Vector Machine
High	87% ± 2.1%	92.16% ± 2.4%	80% ± 3.1%
Medium	80% ± 1.9%	13.29% ± 1.4%	83% ± 2.4%
Low	99% ± 1.7%	70.77% ± 2.0%	94% ± 2.5%
Total	90.03% ± 1.7%	74.51% ± 1.9%	86.16% ± 2%

The CNN LeNet method shows an accuracy rate of 87% ± 2.1% and 99% ± 1.7% in classifying images from the high and low classes, respectively. Although the performances are suitable for the application, the middle class has more overlap.

Detection of Infection Levels in the Whole Banana Leaf

In this part, we observed the performance of the algorithm in detecting disease in banana leaves. It is important to

highlight that this classification represents the diagnosis that the system would deliver to a person interested in analyzing the disease. In Table II, we show the results obtained by validating the decision tree. The results correspond to 89 Monte Carlo iterations.

TABLE II: CONFUSION MATRIX FOR THE DECISION TREE CLASSIFIER

	High	Medium	Low
High	91% \pm 1.9%	8% \pm 1.0%	2% \pm 0.7%
Medium	11% \pm 1%	71% \pm 2.3%	19% \pm 1.3%
Low	<1% \pm 0.5%	4% \pm 0.9%	96% \pm 2.3%

Our results demonstrate that our method can accurately classify high and low classes. However, it was less effective with respect to medium classes. This is due to the high level of overlap in the data, which makes it difficult to find a classification method that can separate classes properly.

IV. CONCLUSIONS

We developed an automatic method for detecting black Sigatoka infection in banana leaves in its early stages. This method will allow users to perform corrective practices to avoid economic losses. In comparison, most studies have detected the infection in its final stages when farmers can no longer act. The proposed approach was tested in an annotated dataset that was created specifically for this work because there was no publicly available database with the different levels of severity labeled by an expert or labeled according to a standardized method such as the Four é scale. We made the dataset publicly available to facilitate comparisons and accelerate research in this area. In the future, the database must be expanded to validate our approach in a wider set of activities.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Cristian Escudero conducted initial literature review, proposed and implemented the model, collected and preprocessed datasets, run experiments and drafted the first version of the paper under supervision of Andrés Calvo and Arley Bejarano. Andrés Calvo and Arley Bejarano edited and extended the manuscript.

ACKNOWLEDGMENTS

Sincere thanks to the Universidad Tecnológica de Pereira, Colciencias agreement 160-2018 code 511-3-024-15 and Maestría en Ingeniería Eléctrica for the support provided during this research.

REFERENCES

- [1] La importancia de la agricultura en la actualidad, Organización de las Naciones Unidas para la Alimentación y la Agricultura.
- [2] Pláano, lo que más produce el campo colombiano, El Colombiano, January 4, 2016.
- [3] J. Amara, B. Bouaziz, and A. A. Algergawy, "Deep learning-based approach for banana leaf diseases classification," *Lecture Notes in Informatics (LNI)*, 2017.
- [4] V. Singh and A. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, 2017.

- [5] M. Vipinadas and A. Thamizharasi, "Detection and grading of diseases in banana leaves using machine learning," *International Journal of Scientific & Engineering Research*, 2016.
- [6] E. Hossain, M. Farhad, M. Anisur, "A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier," in *Proc. International Conference on Electrical, Computer and Communication Engineering*, 2019.
- [7] A. Camargo and J. Smith, "Image pattern classification for the identification of disease-causing agents in plant," *Computer and Electronics in Agriculture*, 2009.
- [8] L. Calvo, M. Guzmán, and A. Guzmán, "Considerations about Application of machine learning to the prediction of Sigatoka disease," in *Proc. World Congress on Computers in Agriculture*, 2014.
- [9] S. Prasanna, D. Peter, and M. Salathe, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, 2016.
- [10] A. Ramcharan, K. Baranowski, P. McCloskey *et al.*, 2017, "Deep learning for image-based cassava disease detection," *Frontiers in Plant Science*, 2017.
- [11] M. Gomez, A. Vergara, H. Ruiz *et al.*, "AI-powered banana diseases and pest detection," *Plant Methods*, 2019.
- [12] G. Cedeño, C. Suarez, V. Danilo, and C. Fadda, "Early detection of resistance to *Mycosphaerella Fijiensis* in local genotype of *Musa* in Ecuador," *Scientia Agropecuaria*, 2019.
- [13] A. Kumar, B. Ganapathysubramanian, S. Sarkar, and A. Singh, "Deep learning for plant stress phenotyping: Trends and future perspectives," *Trends in Plant Science*, 2018.
- [14] L. Rasdi, M. Fanany, and A. Murni, "Metaheuristic algorithms for convolution neural network," *Computational Intelligence and Neuroscience*, 2016.
- [15] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolution network," *Computer Science*, 2015.
- [16] J. Nagi, F. Ducatelle, G. D. Caro *et al.*, "Max-pooling convolutional neural networks for vision-based hand gesture recognition," in *Proc. IEEE International Conference on Signal and Image Processing Applications*.
- [17] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin SoftMax loss for convolutional neural networks," *International Conference on Machine Learning*, 2016.

Copyright © 2021 by the authors. This is an open-access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Cristian Andrés Escudero was born in April 1995 in Cartago, Colombia, and received his bachelor's degree in electronic engineering from Tecnológica University of Pereira, Colombia. He is currently a master's student at Tecnológica University of Pereira. He is a professor of electronic engineering. His research interests include computer vision, machine learning, and precision agriculture.



sectors.

Andrés Felipe Calvo was born in February 1988 in Pereira, Colombia. He received his bachelor's degree in electronic engineering and master's degree in electrical engineering from the Tecnológica University of Pereira in 2012 and 2015, respectively. He is a professor and the head of the Electronic Engineering Department at Tecnológica University of Pereira. His research focuses on the application of machine learning systems in industrial and agronomic



Arley Bejarano Martínez was born in April 1989 in Bogota, Colombia. He received his bachelor's degree in electronic engineering and master's degree in electrical engineering from the Tecnológica University of Pereira in 2012 and 2019, respectively. He is a professor of electronic engineering at the Universidad Tecnológica de Pereira. His research interests include machine learning, digital signal processing, embedded systems, and developing hardware systems to address problems in industry and precision agriculture.