—Scientific Note—

Diagnosis of Nutrient Deficiencies, Pest, and Disease Disorders on Citrus Leaves using Deep Learning Machine Vision

PERSEVERANÇA MUNGOFA*, ARNOLD SCHUMANN, LAURA WALDO, AND TIMOTHY EBERT

Citrus Research and Education Center, University of Florida, IFAS, Lake Alfred, FL

Additional index words. citrus leaf, CNNs, diagnosis, machine vision, symptoms

Late and imprecise diagnosis of crop biotic and abiotic stresses cause significant decreases in productivity and subsequent economic losses as a result of delays in implementing corrective actions. Therefore, supplemental tools are needed to complement the laborious conventional analytical methods employed to identify these stresses. Visual diagnosis of citrus leaf symptoms is challenging in the presence of confounding factors from plant-pathogen-environment interactions. The advances in artificial intelligence (AI) and machine vision have made it possible to develop accurate and inexpensive diagnostic tools to analyze plant properties from digital images. In this study, machine vision in the form of convolutional neural networks (CNNs) was applied to develop image classification models for rapid, non-destructive, and accurate analysis of citrus leaves from digital images captured by a smartphone camera.

The research was conducted at the Citrus Research and Education Center, Soil and Precision Agriculture Laboratory <https://crec.ifas.ufl.edu/>. A targeted leaf sampling was carried out to create a database of leaf images with common citrus disorders encountered in HLB-endemic Florida groves: nutrient deficiencies (N, Mg, Zn, Mn, and Fe), disease symptoms (citrus scab, HLB, citrus canker, phytophthora chlorosis, and greasy spot), pest damage (spider mites) and asymptomatic leaves (i.e., healthy leaves). The top and the underside of each leaf were photographed using a Samsung Galaxy S8 smartphone camera. Leaf tissue analysis along with the Diagnosis and Recommendation Integrated System (DRIS) web tool were used to identify the most limiting essential nutrient in samples. A total of 14,400 images divided into 24 classes (including the top and the underside of the leaf) were used for model calibration. Also, 1400 images from an external dataset were used for independent validation. Each class was trained with 600 images of representative samples of citrus cultivars and distinct degrees of symptoms initial/early, moderate/intermediate, and severe/late.

Transfer learning was implemented to calibrate two pre-trained CNN models, EfficientNet-B4 and VGG-16 to develop the citrus leaf diagnosis models (CLD-Models). The models were then fine-tuned on the database of digital images, with an 80%-20% training and validation ratio. Calibration was implemented in a Jupyter Notebook developed by Pérez and Granger in 2018, using the Keras API, developed by François Chollet in 2015, written in Python 3, running on the TensorFlow framework version 2.4, an open source platform developed by the Google Brain team. A subset of 240 images from the external dataset was used to develop a web survey <https://s.surveyplanet.com/YC17pXmhH>, to assess human classification performance in comparison to the models. The variables used to evaluate model performance were: accuracy, precision, recall, and F1 scores.

Five CLD-Models were developed: four models based on EfficientNet-B4, and one model based on VGG-16. The EfficientNet-B4 models had a better performance than the VGG-16 model during calibration and testing. The results showed a validation accuracy of 98% for the VGG-16 and 99% for the EfficientNet-B4 models. The CLD-Models achieved an independent validation accuracy ranging from 96 to 100% for most of the classes on the external dataset. The average model performance ranged between 97.99% and 98.26% for the EfficientNet-B4 models compared to 95.90% for the VGG-16 model. The model results were compared to the classification performance of two groups of individuals with distinct levels of expertise diagnosing citrus disorders, the Experts, and the Novices (Fig. 1). The Pearson’s Chi-square test showed that the models outperformed human Experts and Novices ($P < 0.001$). These results suggest that the CLD-Models are a reliable tool to supplement field and laboratory assessment of biotic and abiotic stress.

Fig. 1. (A) Confusion matrix with classification results from of group of novice scout and (B) the group experienced professionals.

*Corresponding author. Email: pmungofa@ufl.edu