

Big Data Visualization for Black Sigatoka Disease of Bananas and Pathogen–Host Interactions (PHI) of Other Plants

Richard S. Segall

 <https://orcid.org/0000-0001-7627-2609>

Arkansas State University, USA

Soichiro Takahashi

 <https://orcid.org/0009-0002-7309-2797>

Arkansas State University, USA

Prasanna Rajbhandari

 <https://orcid.org/0009-0009-3394-2137>

Arkansas State University, USA

ABSTRACT

Black sigatoka is a leaf spot disease affecting banana plants that has caused significant yield reductions of up to 50% (Arman et al., 2023). This research presents data visualizations of 8,761 data points related to black sigatoka in banana plants, encompassing attributes such as time, canopy temperature, and relative humidity. The paper also reviews related work, including the application of data mining to plant studies, the use of deep learning neural networks for plant disease data, and the use of machine learning for predicting crop yields and detecting disease. Additionally, it presents big data visualizations for 20,952 values obtained from the web-accessible Pathogen–Host Interactions database (PHI-base), covering various plant disease categories and providing an epidemiological analysis of prevalent causative agents in specific plant diseases.

KEYWORDS

Big Data, Black Sigatoka (Banana), Canopy Temperature (CT) Data Visualization, Pathogen–Host Interaction (PHI), Plant Pathogens, Relative Humidity (RH)

INTRODUCTION

Many researchers have investigated applications of data mining, neural networks for deep learning, and machine learning (ML) for detection of plant pathogens. However, this paper presents the applications of techniques for big data visualization to selected the significant leaf spot disease of black sigatoka for banana and also interactions between pathogens and host plants. Some related work is discussed below and, in the tables, presented in the next section.

Big Data for Plant Omics

Ohyanagi et al. (2023) edited a book on plant omics focusing on advances in big data biology. The book discusses deep learning approaches in plant omics research, including deep learning on plant images and genetic sequences for classifications and regressions, object detection, and image

DOI: 10.4018/IJARB.361940

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

segmentation. It also provides an online resource guide for plant omics databases. Zheng and Terauchi (2023) explored plant–pathogen interactions from the new perspective of “omics.”

Machine Learning (ML) for Plant Disease Detection

Ghosh et al. (2021) provided a survey on plant disease detection using ML approaches. Mehta et al. (2023) researched artificial intelligence (AI)-based plant disease detection and classification using pretrained models. Tripathi et al. (2021) presented a survey of deep learning methods for plant disease detection. Verma et al. (2019) created a deep learning-based mobile application for plant disease diagnosis with a case study on tomato plants. Additionally, Dekka (2022) presented a study on plant DNA barcoding.

Bowers et al. (2022) conducted an in-depth study on surveillance for the early detection of high-consequence pests and pathogens, including plant pathogens, featuring a case study on soybean rust. Yakimovich (2021) used ML and AI for the prediction of host–pathogen interactions (PHI). Sperschneider (2020) used ML in plant–pathogen interactions for empowering biological predictions.

Deep Learning for Plant Disease Detection

Sangeeta et al. (2023) investigated plant leaf disease detection using deep learning. Kour et al. (2023) studied the implementation of ML techniques such as artificial neural networks, linear discriminant analysis, and logistic regression for wheat yield prediction. Kaundal et al. (2022) created a comprehensive deep learning platform named deepHPI for accurate prediction and visualization of host–pathogen protein–protein interactions.

Previous Related Work by Authors

Previous work on data mining related to the environmental stress tolerances of plants has been presented by Segall et al. (2004, 2008), Segall and Nonis (2004), and Segall (2005a, 2005b). In addition, Segall (2006) and Segall and Zhang (2006) have presented studies on microarray databases in bioinformatics. The authors Segall and Rajbhandari (2024) discussed the applications of image processing for plant diseases of four different plant categories.

BACKGROUND AND RELATED WORK

What Is Data Visualization?

Data visualization involves presenting sets of primarily quantitative raw data in a schematic form. It is an essential strategy for many business organizations. Common tools for data visualization include charts (e.g., bar, line, and pie), plots (e.g., scatter, bubble, and box), and maps (e.g., dot distribution maps), some of which are used in the research presented in this paper. Tableau and Power BI by Microsoft are two of the most popular data visualization tools.

Related Work

A comparative study of performances of plant disease detection and classification techniques for big data was presented by Demilie (2024).

Hasan et al. (2022) provided a comprehensive review of plant leaf disease detection using image processing, discussing the algorithms for image processing and AI. Other investigators who used image processing for detection of plant diseases include Bharate and Shirdhonker (2017).

Nagila and Mishra (2023) and Kumar (2022) discussed the effectiveness of ML and image processing in detecting leaf disease. Sahu et al. (2021) discussed challenges and issues in plant disease detection using deep learning. Aldakheel et al. (2024) studied the detection and identification of plant leaf disease using an algorithm named YOLOv4 (meaning, “you only look once”) that compared its performance with established target identification methods.

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/big-data-visualization-for-black-sigatoka-disease-of-bananas-and-pathogenhost-interactions-phi-of-other-plants/361940

Related Content

e-OpenDay: Open Virtual Environment for Biomedical Related Research, Business and Public Resources

Vasileios G. Stamatopoulos, George E. Karagiannis, Michael A. Gatzoulis and Anastasia N. Kastania (2010). *Biocomputation and Biomedical Informatics: Case Studies and Applications* (pp. 215-227).

www.irma-international.org/chapter/openday-open-virtual-environment-biomedical/39615

Transcriptome-To-Metabolome™ Biosimulation Reveals Human Hippocampal Hypometabolism with Age and Alzheimer's Disease

Clyde F. Phelix, Richard G. LeBaron, Dawnlee J. Roberson, Rosa E. Villanueva, Greg Villareal, Omid B. Rahimi, Sandra Siedlak, Xiongwei Zhu and George Perry (2011). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 1-18).

www.irma-international.org/article/transcriptome-metabolome-biosimulation-reveals-human/62298

Towards Optimal Microarray Universal Reference Sample Designs: An In-Silico Optimization Approach

George Potamias, Sofia Kaforou and Dimitris Kafetzopoulos (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 1676-1687).

www.irma-international.org/chapter/towards-optimal-microarray-universal-reference/76141

Parameterless Clustering Techniques for Gene Expression Analysis

Vincent S. Tseng and Ching-Pin Kao (2006). *Advanced Data Mining Technologies in Bioinformatics* (pp. 155-173).

www.irma-international.org/chapter/parameterless-clustering-techniques-gene-expression/4251

Genetic Code: Emergence of Symmetrical Pattern, Beginnings of Matrix Genetics

Sergey Petoukhov and Matthew He (2010). *Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications* (pp. 1-30).

www.irma-international.org/chapter/genetic-code-emergence-symmetrical-pattern/37894