A CNN-based Decision Support System for Pests and Disease Control in Cucumber Plant

[1]Aborisade, Dada O., [1]Adejimi, Alaba O., [1]Falana, Olorunjube. J., [1]Olaleye, Taiwo O., and [1]Daniel, Dauda W.

1. Department of Computer Science, Federal University of Agriculture, Abeokuta, Ogun State.

Corresponding Author: aborisadeda@funaab.edu.ng

ABSTRACT

Cucumber is a well-known vegetable because of its benefits and nutritional composition such as low-calorie, high water content, essential vitamins and minerals. The relatively short cultivation period and low requirement for specialized care also make cucumber farming an accessible option for small-scale farmers, contributing to improved livelihoods and food security. Cucumber cultivation contributes to sustainable agricultural practices; in that they serve as effective cover crops, they prevent soil erosion and suppress weed growth. However, pests such as insects, mites, and rodents, as well as diseases caused by fungi, bacteria, and viruses rapidly spread through cucumber, causing widespread infections, reduced quality, and market value depreciation. It has been found that pests reduce the average yield of vegetables (i.e. cucumber) by up to 50%. Cucumber is easily infested by pests and diseases which in turn affect their growth and consequently reduce crop yields. Aphids, spider mites, and cucumber beetles are among the pests that feed on the cucumber. Several efforts have been made to control pests and disease in cucumber cultivation, but because of their apparent challenges such as problem of efficiency, accuracy, and real-time analysis capabilities, with resulting inadequacy in management, delay in identifying and addressing pest and disease outbreaks, there is a need for a more accurate and effective system. In this paper, a smart system for the control of pest and disease in cucumber cultivation was presented. Dataset from standard repository was used for the model training. Opencv was used for feature extraction image data. The trained model was evaluated, and the evaluation results showed that precision of 92.7%, recall of 91.4%, and accuracy of 98%.

Keywords: CNN, Decision Support system, Pest Disease, Pest Control, and Plant

1.0 Introduction

Cucumbers are creeping plants with various species and cultivars, the most common being *Cucumis sativus*, commonly known as the garden cucumber. Garden cucumber is believed to have originated from India, where it has been grown for thousands of years (Manavalan, 2021). Multiple variations have also been grown in various regions over the centuries. Cucumber cultivation has versatile applications and economic benefits because of its low-calorie, high water content, and its nutritional composition. Essential vitamins and minerals like vitamins (K,

ISSN NUMBER: 1116-249X

and C) including dietary fiber are found in cucumbers, making it a vital component of a balanced diet. The relatively short cultivation period and low requirement for specialized care also make cucumber farming an accessible option for small-scale farmers, contributing to improved livelihoods and food security (Sultana**,** *et al.*, 2022). Cultivation of cucumbers also contributes to sustainable agricultural practices, as these plants serve as effective cover crops, preventing soil erosion and suppressing weed growth.

Cucumber plants are easily infested by different types of pests and diseases that can significantly hamper their growth and reduce crop yields, posing challenges for farmers and agricultural productivity. Sultana *et al.*(2022) found that pests reduce the average yield of vegetables by up to 50%. Among the most prevalent issues are pests such as aphids, spider mites, and cucumber beetles, which feed on the plant's leaves and stems, to cause to stunted growth, chlorosis, and sometimes death of the plant (Skoneczny *et al*., 2020). Fungal diseases like powdery mildew and downy mildew can spread rapidly, causing leaf discoloration, deformation, and fruit rot, thereby diminishing the overall quality and market value of the harvested cucumbers. Bacterial infections, including bacterial wilt and angular leaf spot, can result in wilting, yellowing, and necrosis of plant tissues, with attendant loss to farmers.

Efforts to manage pest and disease in cucumber plants have been faced a lot of challenge, owing exposure of the crop to a diverse range of pests and diseases. The complexities arising from the distinct life stages of pests, combined with the rapid spread of diseases, has been impeding the overall yield and quality of the cucumber crop, thereby jeopardizing food security and economic sustainability.

These challenges have shown the limitations of traditional pest and disease observation methods. Conventional approaches often lack the necessary efficiency, accuracy, and real-time analysis

ISSN NUMBER: 1116-249X

capabilities required for timely and effective intervention. Consequently, delays in identifying and addressing pest and disease outbreaks can lead to inadequate or untimely management interventions, exacerbating the negative impact on crop yield and quality. There is therefore the need for adoption of more intelligent technology and innovative agricultural practices to improve the control and management of pests and cucumber disease in its cultivation for sustainable production and safeguarding global food supplies.

In this paper, a decision support system for pests and disease control in cucumber plant using Convolutional Neural Network (CNN) is presented. Implementation of the proposed support system would contribute to sustainable and efficient pest and disease control practices in cucumber cultivation. The remaining sections include section 2 where existing literatures were reviewed, section 3 where methods adopted were discussed, and section 4 which describes the implementation, results, evaluation, and conclusion.

2.0 Literature Review

In (Ferentinos, 2018), models of many diseases and many crops were trained. The study was based one crop and critical study of the leaves (Ferentinos, 2018). Pictures of leaves which serve as input data are captured using camera, while approach adopted included unbiqutous-based technology which for the purpose accurately identify. Liu *et al.* (2021) provided a synthesis of studies that have utilized CNN to classify pests and diseases in cucumber plants. Examining different feature extraction techniques was the main focus in order to improve the models' prediction abilities. It also included issues with the limitations of the CNN technique for classifying pests and diseases. It also included working with unbalanced datasets, adjusting hyper-parameters, and ensuring the model's predictions are accurate (Liu *et al.*, 2021). This comparative study underscored the unique advantages and potential trade-offs of the CNN model

ISSN NUMBER: 1116-249X

with respect to evaluation metrics like accuracy, interpretability, processing efficiency, and scalability.

K-Nearest Nieghbour (KNN) algorithm has been found beneficial in identifying similar patterns of pest and disease occurrences in cucumber plants, allowing farmers to anticipate potential outbreaks and implement preemptive control measures (Li *et al.*, 2023). By leveraging the proximity-based classification, KNN facilitates the identification of comparable historical instances, enabling timely and targeted interventions to prevent the escalation of pest and disease infestations.

Jeny *et al.* (2021) highlighted the fact that cucumber disease diagnosis by machine has the potential to considerably aid remote farmers in the agriculture sector. Given the difficulties in discriminating between distinct cucumber disease forms due to their comparable symptoms, this study introduced a cucumber disease identification system that is computer-based and has capability for evaluating photos obtained by mobile phones for ease of use by rural farmers. The study entailed extraction of discriminative features of leaf photos, feature segmentation using K-means (to different the infected part of the leaf from others), and five different classification methods were used to classify the disorders.

Agarwal *et al.* (2020) explored the significance of understanding the various diseases affecting cucumber crops and emphasized the pivotal role of leaf analysis in disease identification. This proposed model, using the plant village dataset, showed better performance compared with other models with 94.4% accuracy and adaptability accuracy of 93.7% across datasets beyond the confines of plant village, thereby underscoring the efficacy and versatility of this innovative CNN-based approach in addressing crop diseases. Ma *et al.* (2018) investigated the challenges associated with the timely and effective identification of cucumber diseases, without losing sight

ISSN NUMBER: 1116-249X

issues of subjective diagnosis and excessive pesticide use in agriculture. The study focused on the adoption of convolutional neural networks and other algorithms for disease recognition in greenhouse cucumbers, based on image processing and computer vision technology. The proposed image segmentation method, employing a novel combination of color features and an interactive region growing approach, demonstrates a robust disease symptom image segmentation capability, achieving an impressive overall accuracy of 94.29%.

Ngugi *et al.* (2021) presented experimental results that highlighted leaf disease identification capabilities of half-score convolutional neural network(CNN) architectures, also showing their performance with standard evaluation metrics, training duration, and storage requirements. The study also offered insights into the most suitable architectures for implementation environments. In (Omer *et al.,* 2022), an improved CNN algorithm was developed to identify five distinct cucumber diseases and healthy leaves. The study involved an integrated process of image enhancement, feature extraction, and classification, enabling automated and accurate disease identification. Study results showed the CNN algorithm could accurately identifying cucumber leaf diseases, outperforming alternative algorithms. This contributed significantly to the advancement of disease identification techniques in agriculture farming.

Zhang *et al*. (2021) investigated challenges involved in plant leaf diseases in complex real-field environments, particularly concerning the utilization of images from the agricultural Internet of Things (IoT) system. Their study proposed an innovative and integrated small sample size and deep convolutional neural networks for accurate recognition of cucumber leaf diseases (Zhang *et al*., 2021). Study results showed identification accuracy of 96.11% and 90.67% for lesion and disease leaf images, respectively. The study outcome highlighted its potential for practical

ISSN NUMBER: 1116-249X

implementation within the agricultural Internet of Things domain, outperforming existing approaches and showcasing promising prospects for field applications.

The integration of CNNs in plant disease recognition systems has led to enhanced accuracy and efficiency in disease identification and classification. In (Lu *et al.*, 2021) the inherent ability of CNNs to capture features and patterns from images to classify various disease symptoms and provide accurate and reliable diagnoses (Lu *et al.*, 2021).

The utilization of CNN has significantly enhanced the precision and reliability of plant disease identification, allowing for more precise and timely interventions in agricultural practices. Additionally, the adaptability of CNNs to handle large and complex datasets has facilitated their widespread adoption in plant disease research and has led to the development of sophisticated decision support systems for farmers and agricultural experts (Tugrul et al., 2022).

3.0 Methodology

3.1 ***CNN Architecture*****:** Figure 1 shows how the architectural representation of convolutional neural network (CNN) would to automatically learns and extracts intricate features from raw visual data, in this case cucumber leaf image, making it an effective tool for image classification, object detection, and image segmentation. The architecture has feature learning component and classification component. The feature learning is made up of convolutional layer and pooling layer, while the classification layer is made up of the fully connected layers.

ISSN NUMBER: 1116-249X

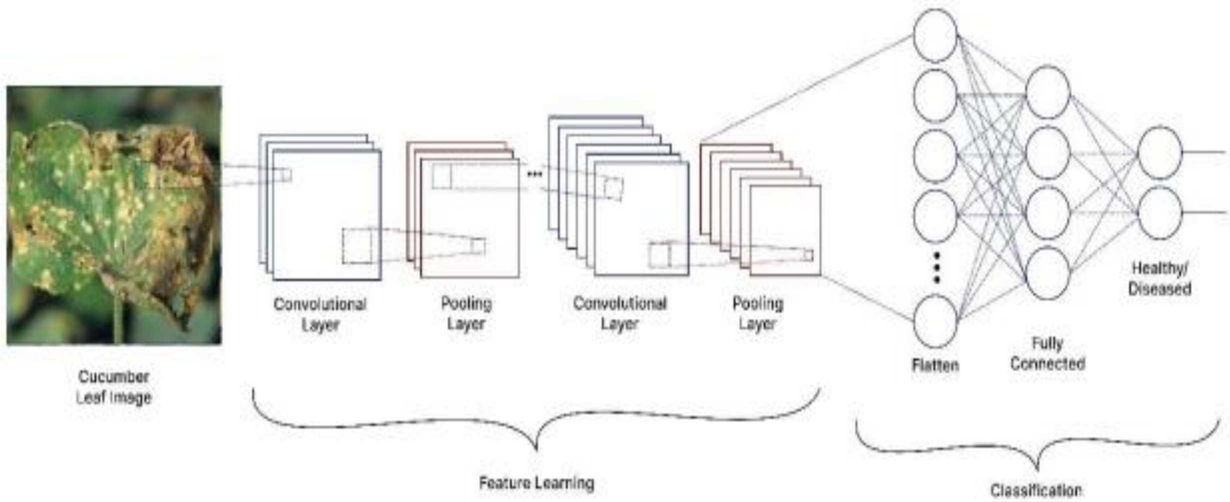

Figure 1: Convolutional Neural Network (CNN) Architecture

3.2 Proposed model framework

ISSN NUMBER: 1116-249X

Figure 2 describes the different modules in the proposed framework as shown below;

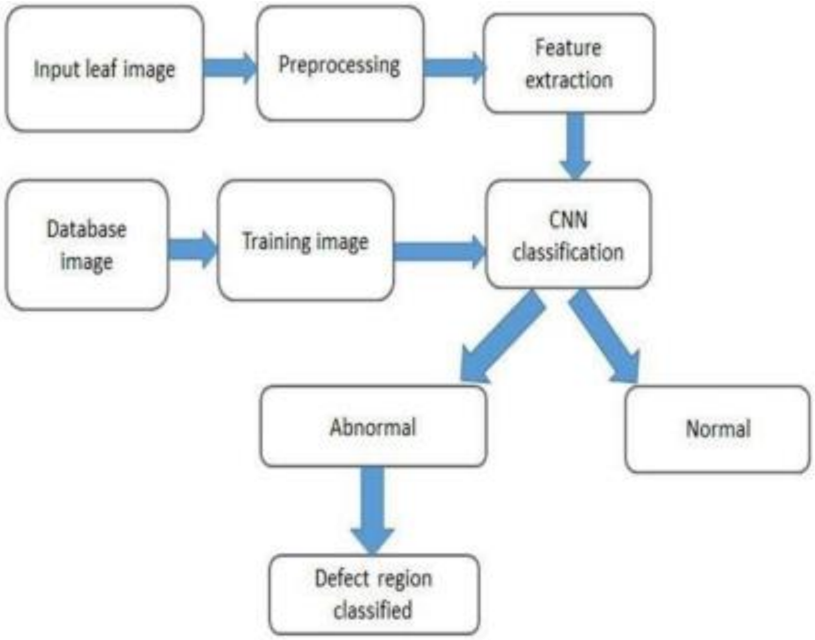

Figure 2: Proposed framework of the CNN Model

3.2.1 Collection and Description of Data

Mendeley Cucumber Disease Recognition Dataset was used in the classification of pests and diseases in cucumbers was downloaded. The dataset is separated into folders, and each image is annotated with accurate labels for the corresponding class, such as 'healthy,' 'powdery mildew,' 'downy mildew,' and other relevant pest and disease categories. The dataset contains eight types of cucumber classes, namely Anthracnose, Bacterial Wilt, Belly Rot, Downy Mildew, Pythium Fruit Rot, Gummy Stem Blight, healthy leaves, and healthy cucumber. The classification of cucumber diseases were done with the assistance of agricultural expert. 1280 images of cucumbers were also collected and augmented before use in model training.

ISSN NUMBER: 1116-249X

3.2.2 Pre-Processing of Data

In order to standardize the images for model training, preprocessing steps like resizing images, and adjust the pixel values. Data was also augmented to 6400 images using rotation, flipping, and zooming to increase the dataset's diversity and prevent over-fitting during training. Each folder is uploaded to Google Drive, where it can be accessed and used by the model before, during, and after training.

3.2.3 Data Splitting

The dataset was partitioned into distinct subsets of a training set, a validation set, and a test set. The data splitting was done to ensure that the model is trained, validated, and tested on a number of samples, to promote robust performance assessment and accurate predictions in its applications.

Model Training Set: 70% of the dataset is used to as the training set for the model's parameters to learn patterns and extract features.

Model Validation Set: 20% of the dataset was used to fine-tune the model's hyper-parameters and prevents over-fitting, ensuring better generalization to unseen data.

Model Evaluation Set: 10% of the dataset served as an independent benchmark to assess and provide insights into the model's capabilities and predictive accuracy.

3.2.4 Model Training

Training of model was carried out on Google Colab to leveraging on its GPU-accelerated computing capabilities to efficiently train deep learning models. This hardware acceleration significantly reduced training times, enabling efficient model development. The collaborative

ISSN NUMBER: 1116-249X

and cloud-based nature of Google Colab streamlines the training and validation processes, allowing for seamless collaboration and efficient experimentation in the development of deep learning models.

3.3 Algorithm

Algorithm 1: Algorithm for the proposed model

Input: i,n, x, $Data_{train}$, $Data_{test}$, $Data_{valid}$

Output: $\mathbf{Y_{class}}$, $\mathbf{Z_{class}}$

Step 1: *Load Import* $Data_{train}$, $Data_{test}$, $Data_{valid}$

Step 2: *import OpenCVin BGR format*

Step 3: *feature extraction* ← *img_rgb(cv2.cvtcolor(image,cv2.color_brgr2rgb)*

Step 4: *population initialization of* $P\{x_1, x_2, x_3,................x_n\}$

Step 5: *for* $i \leftarrow 1$ *to n* ***do***

Step 6: *Build the CNN model architecture*

Step 7: *Compile the model*

Step 8: ***end for***

Step 9: *Perform full training on the CNN architecture*

Step 10: *Determime the image class based on the class label* Y_{class} *or* Z_{class}

Step 11: $\boldsymbol{Y_{class}}$**|** $\boldsymbol{Z_{class}}$

4.0 Implementation, Results and Evaluation

TensorFlow, Keras, and PyTorch, were used to provide high-level abstractions, and making it easier to create and train complex neural networks. Python's compatibility with various data processing libraries, such as NumPy, Scikit-learn, Pandas, and Matplotlib, facilitates data preprocessing, manipulation, and visualization, made it possible to prepared datasets for CNN

ISSN NUMBER: 1116-249X

training. Figure 3 to 5 show the screenshots of interface section for the uploaded cucumber leaves images. Figure 6 represents the cucumber leaf image classification.

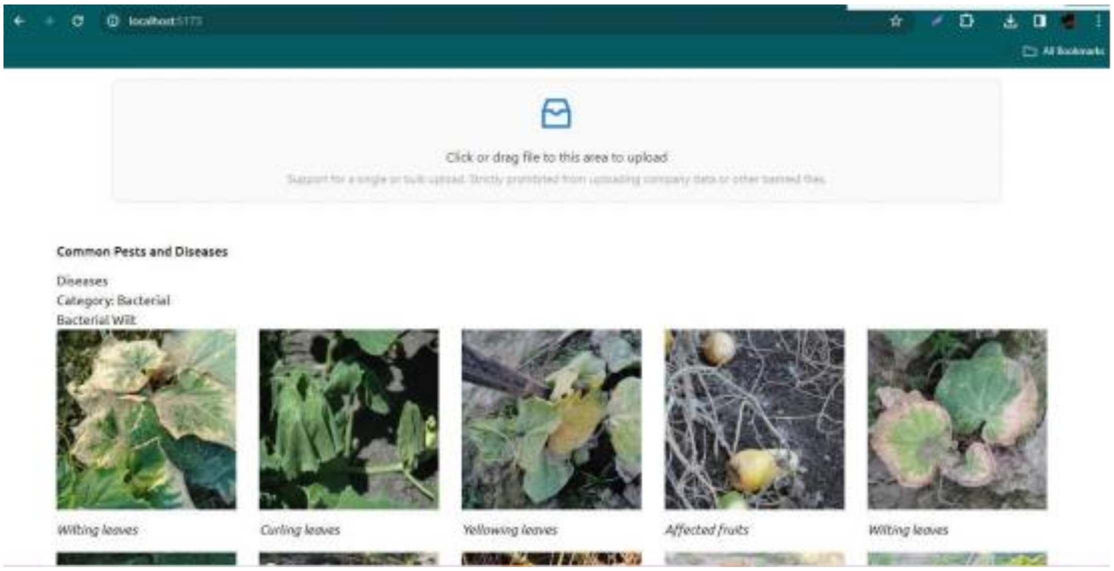

Figure 3: Interface Section for Uploading Images

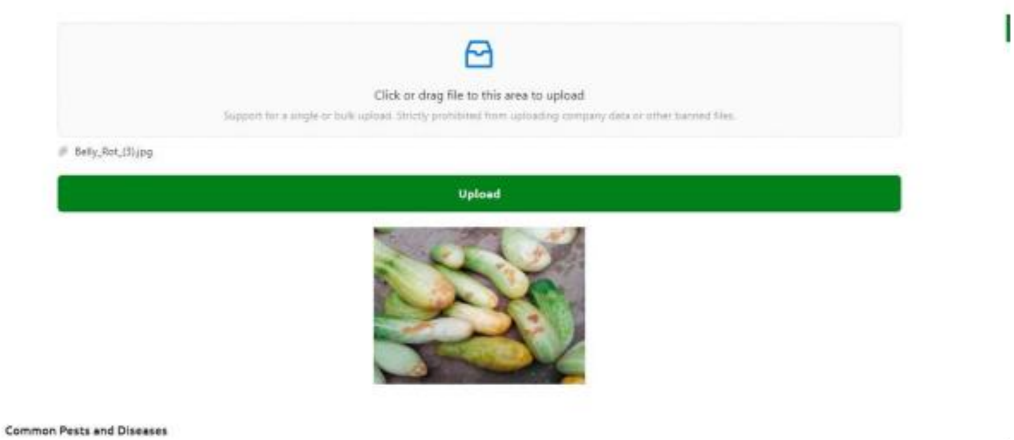

Figure 4: Image Uploaded by User

ISSN NUMBER: 1116-249X

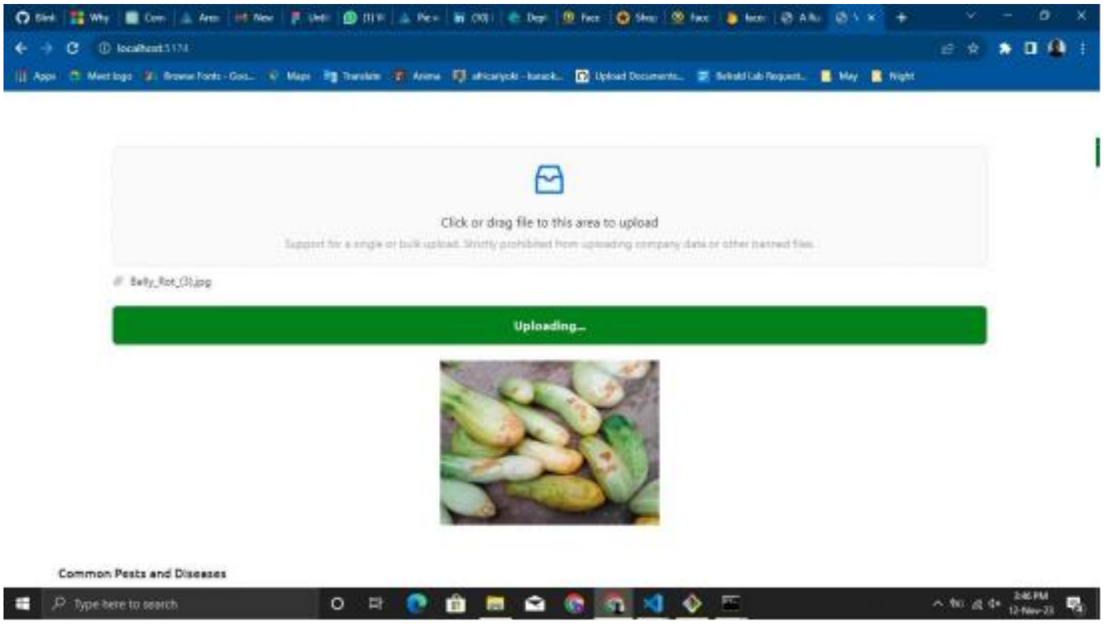

Figure 5: Image Uploading to System

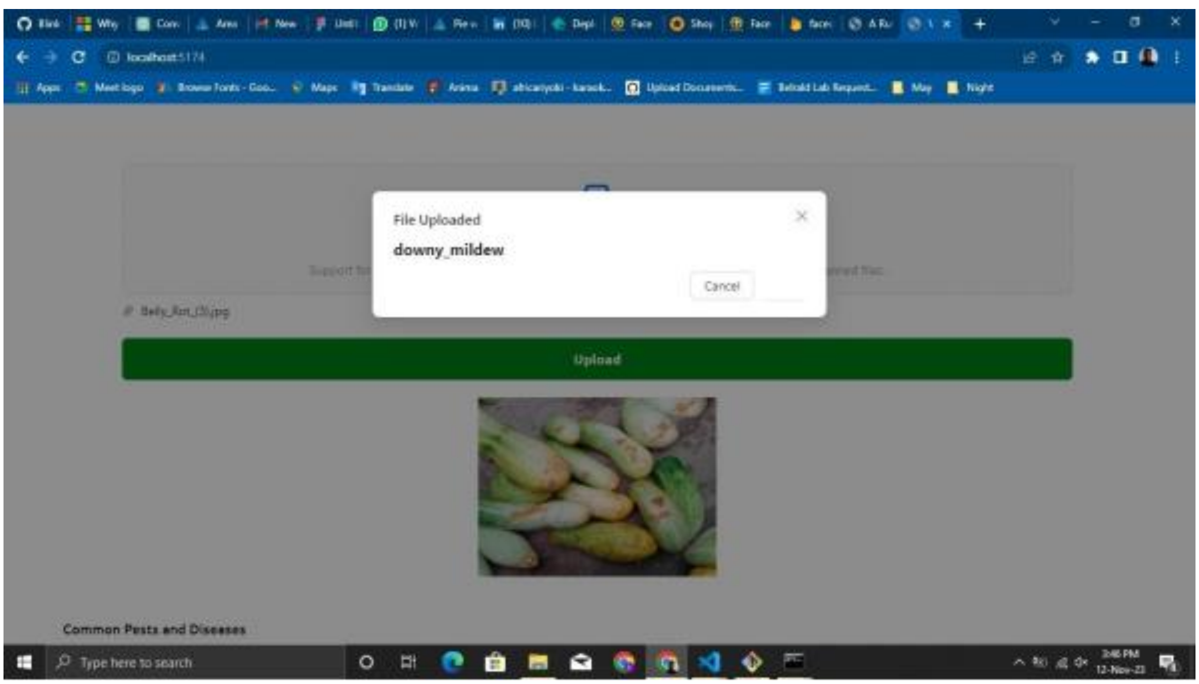

Figure 6: Screenshot of the system Image Classification

ISSN NUMBER: 1116-249X

4.1 Evaluation Metrics

Evaluation metrics are quantitative measures used in evaluating performance of a model during the training phase. Evaluation metrics such as accuracy, training loss, validation loss, learning curves, recall, and precision were used to determine how well the model is learning from the raining data. Accuracy was used to measure the model's capability to make correct predictions, loss indicated the error between predicted and actual values, helping to assess how well the model is minimizing errors during training.

Training loss: metric was used to determine the difference in error between the predicted output of a model and the actual target values in the training dataset, and to assess how well the model is learning during the training process. The intention of model's training is have a very low training loss, so as to gain very high accurate predictions rate. The training loss when measured over successive epochs could also in determining model's generalization to new, unseen data to improve its predictive capabilities. Figure 10 shows that the model learns with every epoch as the training loss comes down.

Validation loss: is the metric adopted in this study to assess the performance of the model during the validation phase of training. Validation loss was used to quantify the error between the model's predicted values and the actual target values on a validation dataset, to provide insights into the generalization capability of the model, and to identify over-fitting or under-fitting. As the value of validation loss reduces, the model performance improves.

ISSN NUMBER: 1116-249X

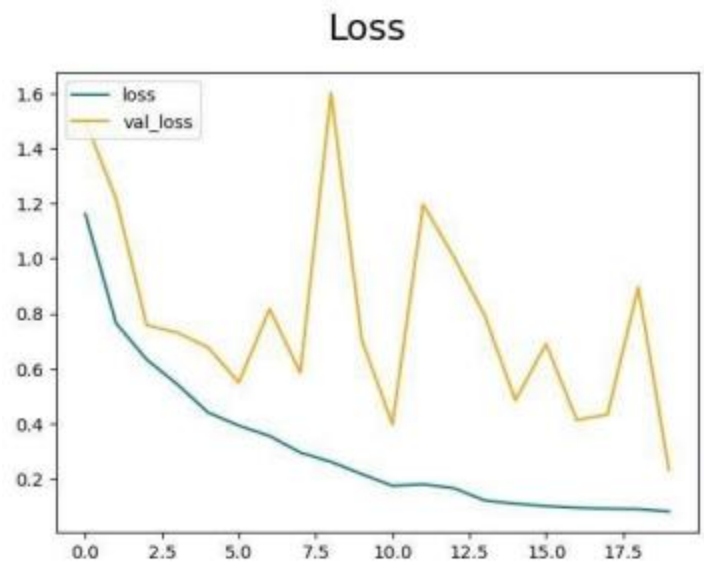

Figure 10: A Graph Depicting the Training and Validation Loss of the CNN Model

Training Accuracy: is a metric to assess the performance of a model while training. It gives the percentage of correctly classified instances within the training dataset, showing correctly the model is predicting the training data.

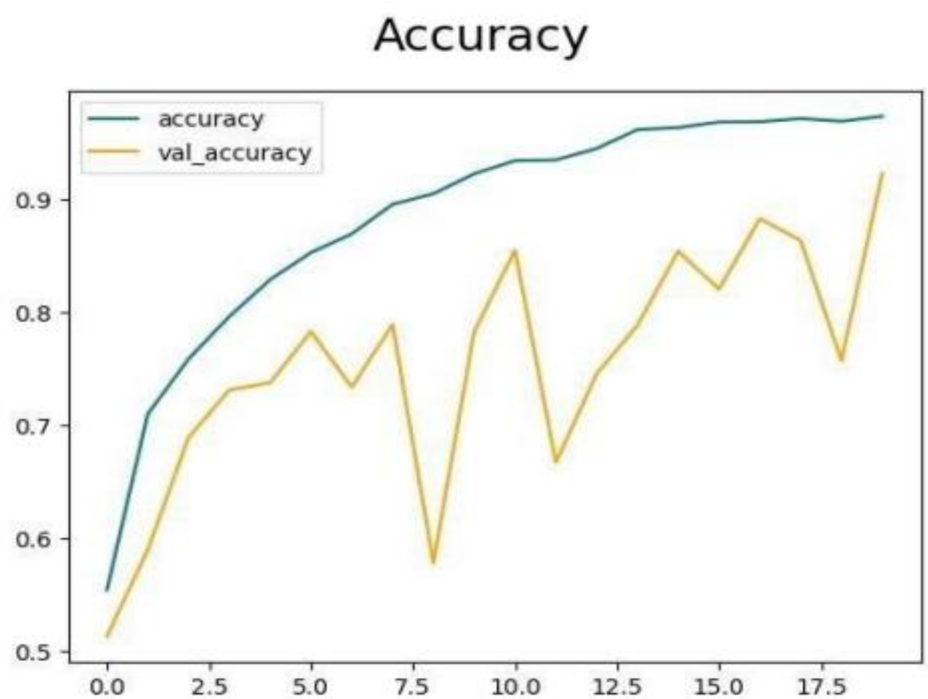

Figure 11: A Graph of the Training and Validation Accuracy of the CNN Model

ISSN NUMBER: 1116-249X

Precision as a metric was adopted to assess a classification model's effectiveness, particularly when it comes to classifying diseases and pests. In this study, the model precision detected a sizable percentage of the instances with a value of 92.7%.

Recall: metric was used to evaluate the model's capacity to accurately identify a sizable percentage of real positive cases in the dataset. It is also referred to as sensitivity. After evaluation and measurement, 91.4% was obtained as the recall for the model.

4.0 Conclusion

This study implemented the development of an intelligent model for controlling and managing pests and diseases in cucumber plant interfaced with a decision support system. In this paper, the trained model is connected to the web application (a decision support system application) which would serve as the front-end for the trained model. The implementation of a decision support system with Convolutional Neural Networks (CNN) to classify pests and diseases in cucumber plant images. This study is intended to mark a significant stride in the field of agricultural technology through adoption of artificial intelligence and image recognition to revolutionize the way farmers manage and safeguard their crop yield. Through the meticulous development and deployment of this technology, the agricultural sector is poised to benefit immensely, reaping rewards that extend beyond mere classification and detection. The integration of CNN technology into agricultural practices signifies a pivotal shift towards proactive and preventive measures in pest and disease control. Furthermore, the deployment of this decision support system streamlines and expedites the decision-making process for farmers, empowering them to make informed choices swiftly and effectively. The real-time analysis of Cucumber plant images

ISSN NUMBER: 1116-249X

enables prompt and accurate responses to potential threats, mitigating the potential economic losses that often arise from delayed or inadequate pest and disease management strategies.

References:

Agarwal, M., Gupta, S. K., & Biswas, K. K. (2020). Development of Efficient CNN Model for Tomato Crop Disease Identification. *Sustainable Computing: Informatics and Systems*, 28, 100407. ISSN 2210-5379. https://doi.org/10.1016/j.suscom.2020.100407.

Arogundade, O., Ajose, T., Olumide Matthew, J., & Adeyemi Osijo, I. (2021). Current and Emerging Pests and Diseases of Cucumber (Cucumis sativus, L.) in Africa. *IntechOpen*. https://doi:10.5772/intechopen.96692

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*. https://doi.org/10.1016/j.compag.2018.01.009

Jeny, A. A., Junayed, M. S., Islam, M. B., Imani, H., & Shah, A. F. M. S. (2021). Machine Vision-based Expert System for Automated Cucumber Diseases Recognition and Classification. *International Conference on Innovations in Intelligent Systems and Applications (INISTA)* (pp. 1-6). Kocaeli, Turkey. https://doi.org/10.1109/INISTA52262.2021.9548607

Li, M., Cheng, S., Cui, J., Li, C., Li, Z., Zhou, C., & Lv, C. (2023). High-Performance Plant Pest and Disease Detection Based on Model Ensemble with Inception Module and Cluster Algorithm. *Plants (Basel, Switzerland)*, 12(1), 200. https://doi.org/10.3390/plants12010200.

Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. *Plant Methods*, 17(1), 22. https://doi.org/10.1186/s13007-021-00722-9

Lu, J., Tan, L., & Jiang, H. (2021). Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. *Agriculture*. https://doi.org/10.3390/agriculture11080707

Manavalan, R. (2021). Cucumber Leaves Diseases Detection through Computational Approaches: A Review. *Asian Journal of Research in Biosciences*, 3(2), 143–153. Retrieved from https://globalpresshub.com/index.php/AJORIB/article/view/1302

Ma, J., Du, K., Zheng, F., et al. (2018). Disease Recognition System for Greenhouse Cucumbers

Based on Deep Convolutional Neural Network. *Transactions of the Chinese Society of Agricultural Engineering*, 34(12), 186–192. https://doi.org/10.1016/j.dib.2018.109320

Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. *Information Processing in Agriculture*, 8(1), 27-51. ISSN 2214-3173. https://doi.org/10.1016/j.inpa.2020.04.004.

Omer, S. M., Ghafoor, K. Z., & Askar, S. K. (2022). An Intelligent System for Cucumber Leaf Disease Diagnosis Based on the Tuned Convolutional Neural Network Algorithm. *Mobile Information Systems*, 16, 8909121. https://doi.org/10.1155/2022/8909121

Skoneczny, H., Kubiak, K., Spiralski, M., Kotlarz, J., Mikiciński, A., & Puławska, J. (2020). Fire blight disease detection for apple trees: hyperspectral analysis of healthy, infected and dry leaves. *Remote Sensing*, 12, 2101. https://doi.org/10.3390/rs12132101.

Sultana, N., Shorif, S. B., Akter, M., & Uddin, M. S. (2022). Cucumber Disease Recognition Dataset. *Mendeley Data*, V1. https://doi.org/10.17632/y6d3z6f8z9.1

Tugrul, B., Elfatimi, E., & Eryigit, R. (2022). Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review. *Agriculture*. https://doi.org/10.3390/agriculture12081192.

Zhang, J., Rao, Y., Man, C., Jiang, Z., & Li, S. (2021). Identification of Cucumber Leaf Diseases Using Deep Learning and Small Sample Size for Agricultural Internet of Things. *International Journal of Distributed Sensor Networks*. https://doi.org/10.4236/2021/31B0

ISSN NUMBER: 1116-249X