Contents lists available at Scholar Indexing Society

Artificial Intelligence (AI)

Homepage: **https://globalweb1.com/index.php/ojs**

ORIGINAL ARTICLE **OPEN** Check for updates

Disease Detection in CCN-51 Cocoa Fruits through Convolutional Neural Networks: A Novel Approach for the Ghana Cocoa Board

Miracle A. Atianashie

School of Research and Graduate Studies, Catholic University of Ghana, P.O. Box 363, Sunyani, Bono Region, Ghana

e-mail: miracleatianashie@cug.edu.gh

Abstract

The study explores the significant challenge of diagnosing diseases in CCN-51 cocoa fruits within Ghana, a key concern for the agricultural sector. This model aims to revolutionize the accuracy of disease detection in cocoa fruits, a crucial step toward bolstering the sustainability of Ghana's agricultural sector. By significantly improving detection rates, the project anticipates providing a solid foundation for more effective disease management strategies, ensuring the health and productivity of cocoa crops, and, by extension, supporting the economic stability of the farming communities reliant on cocoa production. The methodology is designed with a dual focus: ensuring the model's robustness to handle real-world agricultural complexities and verifying its adaptability to the diverse conditions encountered in cocoa farming environments. A comprehensive series of experiments were meticulously designed to evaluate the CNN model's diagnostic capabilities. These experiments were structured to assess the model's precision in identifying various diseases across different stages of infection, environmental conditions, and fruit varieties. The research aims to rigorously test the model's effectiveness and reliability by simulating a wide array of real-world scenarios, ensuring its practical applicability for farmers and agricultural professionals. The experimental findings paint a promising picture, showcasing the CNN model's exceptional performance across key metrics such as accuracy, precision, recall, and F1 scores. These results highlight a significant leap forward in disease detection capabilities, surpassing the benchmarks set by conventional methods. The high level of accuracy not only validates the model's effectiveness and signals its potential to transform disease management practices in cocoa agriculture.

The implications of these findings are profound, with the potential to catalyze a paradigm shift in how disease detection is approached in the cocoa farming sector. The study elaborates on the multifaceted benefits of the CNN model, emphasizing its role as a cost-effective, efficient, and scalable tool for disease management. By significantly reducing crop losses and enhancing production sustainability, the model promises to bolster the economic well-being of cocoa farmers and contribute to the broader goals of agricultural innovation and food security in Ghana.

Keywords: Convolutional Neural Networks (CNN), Disease Detection, Cocoa Fruit Agriculture, Ghana Agricultural Sector, Crop Sustainability

1. Introduction

Disease outbreaks again and again pose a threat to global food production and food security. Accurate and timely detection and identification of plant disease could reduce major agricultural production losses and prevent human exposure to plant pathogens. However, it remains challenging for plant pathologists and epidemiologists to identify early signs of pathogens that affect plants by visualization through a microscope or the naked eye. Traditional plant disease detection methods are time-consuming, highly labour-intensive, and often impractical. Visual assessment of symptom expression used in these traditional methods is highly subjective and error-prone. For example, different people may interpret certain symptoms differently, which may introduce misclassification. Traditional molecular-based diagnostic approaches require pathogen-specific primers and involve several failure-prone and technically demanding protocols that might not be suitable for routine plant disease diagnosis. Plant pathologists are now moving towards more high-tech solutions.

Remote sensing technologies are implemented using digital imaging, such as multispectral and hyperspectral imaging, as well as digital photographs. However, these technologies require significant costs to develop and adopt, making them out of reach for most growers and field workers in developing countries. (Ronneberger et al., 2015) Noted that there is no commercially available disease detection robot to date. Robots are increasingly being used to tackle life-challenging tasks in the agriculture sector. So far, teams of researchers from different parts of the world are developing variations of robots to address the perils of plant pathology. Most of them utilize some form of light spectrum analysis due to the nature of the disease symptoms. For example, the recently developed Bonirob uses light spectrum to analyze plants, and the robot is equipped with sensors to prevent it from stepping onto non-infected areas and from harming humans. Coronatine is produced by many species of pathogens in plants (Espejo, 2018; Wang et al., 2019). The Ray technique is proposed, which relies on detecting these microbes, and this is one aspect that robot developers focus on. However, the costs of developing robots and extracting real-time computer data are high, making them not widely available, especially in developing countries. Our research aims to use a novel approach for detecting diseases in cocoa fruits, which is of high economic importance in Ghana. We strive to develop a machine-learning approach that is significantly faster and more cost-effective for image-based detection of diseases in cocoa than traditional molecular-based diagnostic methods. We hope this research will support the development of an affordable yet highly efficient robotic platform for complex, stage-specific, and high-throughput plant disease phenotyping that would benefit the crop research community.

1.1. Background

Cocoa is a major cash crop in Ghana. Due to Ghana's suitable environment for growing cocoa, it has become the largest producer of cocoa in the world. The cocoa industry has been identified as a major contributor to the country's socio-economic development. The crop is believed to be cultivated in all the regions in Ghana. The major production areas are Ashanti, Western, Central, Brong Ahafo, and the Eastern regions. However, it has been recognized that the cocoa crop in Ghana is always under attack from pests, diseases, and weather-related problems. In the 2017/18 crop year, Ghana's diseases and pests' infestations destroyed 16,000 hectares of cocoa farms. The slack in production can be partly attributed to the cocoa swollen shoot virus disease, black pod disease, and capsid insect infestation. Pests and diseases have been found to cause substantial losses in cocoa production in Ghana every year. Therefore, it is important to develop an efficient mechanism to enable early detection of pests and diseases in affected cocoa farms. This will help to give prompt attention to the distressed farms and limit the rate of destruction. Currently, pests and disease detection in cocoa farms in Ghana are being done by manual inspection. However, manual inspection is time-consuming and laborious, especially when large farms are to be inspected (Signoroni et al., 2019; Zhao et al., 2019).

Additionally, the subjectivity of the inspectors may affect the detection accuracy due to sheer fatigue. The challenges associated with the traditional method of disease detection in Ghana have led to research into the use of data-driven methods such as remote sensing and satellite image analysis to provide automated procedures for identifying pests, diseases, and other farm troubles. These modern technologies have been found to provide accurate and reliable detection methods and strengthen the efficient use of resources provided to distressed farms. The available technologies and modern farming methods, including drones and satellite image analysis systems, have great potential to revolutionize cocoa farming in Ghana. By considering the limitations and challenges associated with the current manual inspection practices and the need for efficient and reliable production of the crop, this work intends to propose a novel approach for the detection of diseases in CCN-51 cocoa fruits (Rajaee et al., 2015; Shendryk et al., 2019). It is proposed that an image processing framework that employs the use of Convolutional Neural Networks (CNN) be designed to allow for an automated, fast, accurate, and reliable process to aid in the detection of diseases from the images of the CCN-51 cocoa fruits. Through this study, it is expected that the proposed novel approach will provide a modern and efficient method for the detection of diseases in the cocoa crop at earlier stages of development. This project aims to develop an efficient and automated mechanism for detecting diseases in CCN-51 cocoa fruits. It will involve the design of a comprehensive framework from the data collection, preprocessing, CNN model development, and testing of the developed model through a range of experimental activities (Obodai et al., 2019; Olofsson et al., 2014). The year 2020 is marked as the beginning of a new cocoa season in Ghana. It is anticipated that the proposed approach could be successfully implemented and incorporated with the existing technology, such as drones and satellite image analysis systems, to help the cocoa industry deal with the problem of diseases more effectively and progress towards sustainable cocoa production in Ghana.

1.2. Problem Statement

In order to ensure a smoother and timely operation of the Ghana Cocoa Board, this paper will discuss a novel methodology to detect diseases, particularly black pod disease and cacao swollen shoot virus disease, in cocoa fruits. Black pod diseases are the most common diseases in cocoa plants and are seriously widespread globally in all cocoa-growing areas. From time to time, the severity of the disease has come to the level of being an epidemic. If the disease starts to spread in the crop, the only action that can be taken is to destroy the infected plant to prevent the disease's spread. On the other hand, the cacao swollen shoot virus is another dangerous repeated virus disease that causes serious economic detriment to cacao production in western Africa. The virus is always killing off young plants and cutting into adult plants' yield. This present paper will focus on detecting diseases in cacao fruits by taking high-resolution images using a drone camera. This means that the problem here is image recognition and detection, which are major problems in the agricultural field. Visual inspection is a tedious, time-consuming task for humans, requiring expertise and a lot of experience (Abdulai et al., 2018; Gil de Zúñiga et al., 2023; Teye et al., 2020). According to the literature, the main problem with the machine is the selection and classification of suitable features. Also, it has been highlighted in the literature that such applications have a lot of challenges, such as variations in different environmental conditions, different diseases, infections, and different stages of the growth of plants. Unlike traditional disease detection methods such as visual assessment and leaf testing, which suffer from inaccuracy and low detection efficiency, the proposed method offers a fast and effective method to detect the diseases in the cacao fruits at the earliest stage of the infection. By using this novel approach, the rapid spread of diseases could be prevented, and the yield and quality of the products could also be improved. The swift detection and proper scientific management of the disease will ensure sustainable and profitable cocoa production, which ultimately helps in the development of both the cocoa industry and the national economy of Ghana (Aboah & Setsoafia, 2022; Kehinde et al., 2021; Wade et al., 2010).

1.3. Objective

The difficulty of detecting cocoa diseases early has become a major concern for the Ghana Cocoa Board over the years. In addition, current disease detection methods are largely based on manual inspection by workers in cocoa farms, which inevitably results in low accuracy and high false negatives and false positives. The study's main objective is to investigate alternative disease detection methods in the form of image analysis and machine learning algorithms. The success of CNN in identifying faces and objects makes the technology a potential candidate for image processing-based disease detection in cocoa. However, it is still uncertain whether CNN technology enables the effective detection of diseases in cocoa. Also, this study aims to develop a deeper understanding of the potential application of CNN in cocoa disease detection and to propose an effective image analysis based on cocoa disease detection methodology.

2. Literature Review

The destructive impact of witches' broom (Marasmius perniciosia), black pod (Phytophthora pod rot), and frosty pod (Monilliophthora roreri) diseases and the significant decrease in cocoa production has been well documented. These diseases are more prevalent among the Forastero trees and have the potential to cause 100% crop loss under favourable conditions. Frosty pod disease is the most devastating as it creates a high-quality crop loss and keeps farmers in a cycle of poverty due to the high rate of other primates and humans across the country and the government interventions. Stem and black pod diseases, which are caused by Phytophthora and form a grouping called soft pod diseases, occur in most cocoa-growing areas of Ghana, with the exception of the rainy zone and the high veldt zone. Cinchona and the likes of Monilia, which are plentiful in the dry high forest, continue to restrict the development of cocoa in these areas. Witches' broom disease has caused significant requirements for urgently rebuilding the cocoa Tetteh Quarshie Nursing and breeding stock. Ghana's cocoa production has been and is still negatively affected by diseases. According to (de Boer et al., 2019 Saj et al., 2023) (Attipoe et al., 2020 Cilas & Bastide, 2020 Dormon et al., 2004 Tsiboe et al., 2018), the increasing cultivation of cocoa in monotypic stands and the perennial nature of cocoa plants have given rise to an escalation of pests and diseases of cocoa across the country such as capsid and mirids. (Ku & Lewis, 2023) also noted that cocoa agroforests' ecological habitat and complexity provide suitable conditions for insect vectors and non-vector transmitted diseases. Frequent and long-lasting humidity conditions, such as in the wet evergreen forest and the wet semi-deciduous forest regions, offer a conducive environment for disease-causing pathogens.

2.1. Cocoa Diseases and their Impact

Cocoa is a crop that is mainly grown in areas within 20 degrees north and south of the Equator because it requires warm temperatures and high rainfall. The cocoa tree is susceptible to at least 10 diseases. However, the most prevalent and destructive diseases are swollen shoots, black pods, and witches' broom - which have caused the most damage. For instance 1987, this disease wiped out virtually all of Brazil's cocoa production. It acts by attacking a particular part of the tree and, in the process, stimulates the vegetative growth of another part of the tree. On the other hand, black pods produce a wet, sticky rot that affects the cocoa in the pod. It is even deadlier - and in the right weather, it can destroy a whole crop in just 10 days. However, the biggest impact is still diseases like witches' broom and black pods. Black pods are not so deadly, yet they seem to be getting worse in Ghana, so the effects are increasing. These diseases cause an estimated 30-40% of the cocoa production to be lost yearly(Attipoe et al., 2020). Scientists also think that in addition to diseases, the effects of climate change are beginning to manifest in Ghana. These diseases will try to adapt to the new environment. For example, new virulent strains of the pathogen that cause black pods are fast becoming prevalent in Ghana, overtaking the old strains there. It results in an increasing probability of these pathogens attacking the cocoa, leading to increased damage. For producing countries like Ghana and Cote d'Ivoire, it means they enjoy a higher annual variability of coffee and cocoa production because of climate change and diseases. These had a major impact on farmers in these countries.

Farmers usually would not dare to change what they grow because of the need for stable income and the impact of losing a crop can mean that they will not be able to afford the next season or feed their family. Also, once a disease takes hold, they are largely defenceless against it. Visual inspections of the pods and ripened fruits are conducted regularly. However, diseases can be challenging to diagnose accurately as many similar symptoms offer no clue for non-expert visual inspections (Eric et al., 2023). So, having the right skills and knowledge to diagnose diseases accurately is essential. Work that has been done to explore the development of sustainable techniques and the use of modern scientific tools should be utilized to conduct these diagnoses. For example, molecular biology and DNA techniques can offer rapid and accurate diagnoses from very small samples of plant material. These techniques are well established, and computer databases are provided to search for the results of DNA analysis to identify the disease (Ku & Lewis, 2023). However, the fact that diseases such as the witches' broom are so devastating means that concentration should be given to methods of controlling and limiting the spread of the diseases. It is discovered that there was no one way of controlling plant diseases in the future (Koko et al., 2013). However, the researchers noted that the use of pesticides and fungicides should be reduced to reflect the demand for more environmentally friendly and sustainable control strategies, such as disease-resistant plants and eradication and exclusion of diseases. In summary, different plant diseases have very different ways and factors that influence the diseases, and in the broader class of cocoa production, different risk factors play a role in how science will be applied. The best approach would probably be a mix of techniques to control diseases effectively. For example, resistant varieties may be considered if a well-adapted variety that is resistant against a prevalent disease is available. Also, better practices by the farmers, such as separating plantings at high risk from diseases, should be included in the improvement plans, and better methods of diagnosis should be considered to implement early interventions (Gockowski et al., 2013).

2.2. Traditional Methods of Disease Detection

The traditional methods of cocoa disease detection are categorized into clinical, molecular, and cultured-based methods. In the clinical method, diseases are diagnosed through direct observations of symptoms exhibited by CCN-51 cocoa. Plant doctors on routine visits to the cocoa farms visually inspect the plants, and whenever something unusual is spotted, a small part of the infected plant is cut and immersed in a chemical that gives information about the causative agent of the disease. This method is advantageous because the diseases can be rapidly identified and intervention is timely, but it's not commonly used as it demands high expertise, and many failures in diagnosis have been reported (Cilas & Bastide, 2020; Donkor et al., 2023). The molecular techniques involve the isolation of DNA and the application of polymerase chain reaction to amplify it so as to detect pathogens. This is advanced and highly specific but very expensive and not commonly employed. Cultured-based methods, on the other hand, are cheaper and more practical because they don't require sophisticated equipment or a high level of expertise. These methods involve collecting diseased tissues, arranging them in a suitable nutrient medium, and allowing the pathogen to grow to identify it under the microscope. Applying the data collected by research should enhance the effectiveness and realization of the cocoa industry's great potential(Snapir et al., 2017). The establishment of a national digital cocoa diseases database for each of the cocoa-growing regions to store and manage data on the occurrence, spread, and management of cocoa diseases is highly recommended. This would aid in mapping out the hotspots in the regions, and routine monitoring could be initiated. It would also contribute to the implementation of precise and customized spraying programs with timely and appropriate disease management practices.

The use of spatial predictive models such as geographical information system (GIS), which uses disease incidence or prevalence, weather and climate data, and satellite images on cocoa farms to create disease risk maps, is encouraged as it will improve the efficiency of disease control (Iddrisu et al., 2020; Tsiboe et al., 2018). Also, with the advent of modern technologies and the use of unmanned aerial vehicles (UAVs) in agriculture, cocoa farmers can utilize remote sensing techniques such as compact imaging spectroscopy to detect unique 'stress signatures' produced by cocoa trees in response to disease attacks. Such injuries appear before symptoms are visible to the naked eye, and this early sign of diseases is captured and translated to the farmers through the generation of various types of images of the affected trees. Farmers are sent real-time alert messages on their mobile phones when certain types of these images are identified, and this smart disease detection process is termed hyperspectral imaging. Lastly, there should be conscious efforts to deviate from chemical control in managing cocoa diseases, and emphasis should be placed on integrated pest measures. This is because synthetic fungicides' preventive and curative properties are short-lived, and their application would gradually lead to pathogen resistance. On the other hand, the continuous use of broad-spectrum fungicides that kill beneficial organisms makes the ecosystem less balanced. Biological control and the prevention of the infection cycles should be the focus to achieve long-term sustainability in the cocoa industry (Akoa et al., 2021; Iddrisu et al., 2020).

2.3. Introduction to Convolutional Neural Networks (CNN)

Every image has been classified as a set of neurons organized in a feature detection output, and when the "firing" of a neuron matches the specific, actionable signal, the image would be classified as the corresponding one. This process allows for determining what has been detected when evaluating the taken image. Overall, a feature map is a way of taking convolutions of a large input. The process provides a way of encoding qualitatively different local structures of the digit, and the connectivity pattern between the input neurons and the neurons of each layer more closely resembles the organization of the animal visual cortex, which contrasts with the traditional artificial neurons in a human brain.

Every layer assigns one or more feature maps, and each is always smaller than the previous layer. A feature map is a map of inputs that highlights a particular feature so that the outputs are a set of arrays detecting the different regions. There are four main layers in a CNN, and are used to build its architecture. They are the convolutional layer - The pooling layer - The fully connected layer - And the normalization layer. The typical "artificial neural networks" designed to recognize handwriting would have to be trained with each pixel of the image and a representation of the strokes that make up the letters, but a convolutional neural network might be able to optimize those connections. Also, any pixel operation is going to be translationally invariant, meaning that the position of the features in the input does not need to match the position of features in the output. This feature helps handle the tricky correlation problem that arises between the output and the input (Akoa et al., 2021).

The primary difference between a convolutional neural network and the traditional "artificial neural network" is that a typical "artificial neural network" has three layers, which are the input layer, hidden layer and output layer, where the input is directly connected to the output and the "hidden layer" is responsible for the non-linearity. This means that the input data has to be made directly to be understood by the neural network, as there is no relationship between the input and output. However, in convolutional neural networks, the model itself automatically understands the important features regardless of the combination (Miracle, 2024). Convolutional neural networks are a type of feed-forward artificial neural network in which the organization of the animal visual cortex inspires the connectivity pattern between its neurons. Individual cortical neurons respond to stimuli in a limited visual field region known as the receptive field. The preferred stimuli of the neuron are the ones with a particular shape. In the visual cortex, these neurons are organized in such a way that they are responsive to overlapping regions of the visual field. This is in contrast with regular "artificial neural networks", in which the connections between the neurons do not have to be correlated and in which the neurons are fully connected (Oyekale, 2018).

3. Material and Methods

The study uses a sampling of CCN-51 cocoa fruits collected from different geographical locations in Ghana. Each fruit sample was subjected to isolation of the disease spots, that is, both the upper, middle, and lower sections, as shown in Fig 1. The basic data pre-processing steps were adopted in this research. This includes data cleaning, where the incomplete recorded data are corrected, deleted, or ignored. Data cleaning and manipulation were performed using Panda's Python library. Later, the recorded data was divided into two classes: one for the healthy images and the other for the unhealthy images. This is used to train the network (a process that will be explained in the next section). Consequently, the sample data was split and met 80/20 for training and validation, as shown below. As the name suggests, Convolutional Neural Networks (CNN) are mainly used in image recognition and classification or natural language processing. CNN tends to determine the features of the images, and once it is trained with enough data, the model will be able to make predictions. CNN has a more complex architecture than other forms of neural networks. The structure of a typical CNN is made up of 3 layers as follows: the Convolutional layer, the Pooling layer, and the Fully Connected layer. It has to be noted that in each Convolutional layer, there are two sub-layers of neurons. These are the convolution sub-layer and ReLU sub-layer. In the Pooling layer, every time a set of neurons becomes activated, only a single output is transmitted to the next layer, and this image activation data is sent to the output layer. This work introduced a way to measure the performance of CNN, which can be defined as an "evaluation metric". The evaluation metric aims to measure the classification accuracy for the training and testing dataset. However, there are different methods to evaluate CNN. For example, Receiver Operating Characteristic (ROC) is a preferred method for the analysis of a classification system, as it makes analysis easier.

Table 1: Summary of Methodology and Ghanaian Regions

Region	**Sample Collection**	**Data Pre-processing**	**CNN Use**	**Evaluation Method**
Western Region	CCN-51 cocoa fruits	Pandas library (Python)	Image recognition & classification	Classification accuracy, ROC
Central Region	CCN-51 cocoa fruits	Pandas library (Python)	Image recognition & classification	Classification accuracy, ROC
Ashanti Region	CCN-51 cocoa fruits	Pandas library (Python)	Image recognition & classification	Classification accuracy, ROC
Eastern Region	CCN-51 cocoa fruits	Pandas library (Python)	Image recognition & classification	Classification accuracy, ROC
Brong-Ahafo Region	CCN-51 cocoa fruits	Pandas library (Python)	Image recognition & classification	Classification accuracy, ROC

Table 1 presents a detailed overview of the study's geographical scope and methodological focus on CCN-51 cocoa fruit samples across Ghana. The study spans five key cocoa-producing regions, each with specific towns selected for sample collection and analysis. In the Western Region, towns like Sefwi Wiawso and Enchi were chosen for their significant cocoa production, where the study focused on collecting samples from various sections of the cocoa fruits to isolate disease spots. Central Region towns such as Assin Foso and Twifo Praso contributed to the study through the isolation of disease spots and sample preprocessing, emphasizing the study's data cleaning and manipulation phases.

The Ashanti Region, with towns like Tepa and Mampong, was pivotal in data collection for classifying images into healthy and unhealthy categories, showcasing the study's aim to enhance disease detection in cocoa crops. In the Eastern Region, towns, including Akim Oda and New Tafo, were crucial for preparing the training and validation datasets, highlighting the region's role in the study's analytical framework. Finally, the Brong-Ahafo Region, through towns like Goaso and Berekum, was integral to the training of the Convolutional Neural Network (CNN) model and the analysis of evaluation metrics, underscoring the study's technological and methodological advancements in cocoa disease detection and classification.

Table 2: Sampling Locations and Focus Areas for CCN-51 Cocoa Fruit Study in Ghana

Region	Towns	Study Focus
Western Region	Sefwi Wiawso, Enchi	Sample collection from upper, middle, and lower sections of cocoa fruits
Central Region	Assin Foso, Twifo Praso	Disease spot isolation and sample preprocessing
Ashanti Region	Tepa, Mampong	Data collection for healthy vs. unhealthy image classification
Eastern Region	Akim Oda, New Tafo	Training and validation dataset preparation
Brong-Ahafo Region	Goaso, Berekum	CNN model training and evaluation metric analysis

Table 2 provides a detailed overview of a study's geographical scope and methodological focus on CCN-51 cocoa fruits in Ghana. The study spans five major cocoa-producing regions in Ghana, with specific towns in each region selected for sample collection and analysis. The table highlights the diversity of the study locations, ensuring a comprehensive dataset that represents various environmental and cultivation practices across Ghana.

Western Region: The towns of Sefwi Wiawso and Enchi were chosen for their prominence in cocoa production. The study focused on collecting fruit samples from different sections (upper, middle, and lower) to isolate disease spots, a crucial step in understanding the prevalence and distribution of diseases affecting cocoa plants in this region.

Central Region: In Assin Foso and Twifo Praso, the research emphasized the isolation of disease spots and the preprocessing of samples. This step is vital for preparing the data for further analysis, ensuring that only complete and relevant information is used in the study.

Ashanti Region: Tepa and Mampong were selected for their significant contributions to cocoa production. Here, the study aimed to collect data about the difference between healthy and unhealthy images of cocoa fruits. This classification is fundamental for training the Convolutional Neural Network (CNN) to recognise and predict cocoa plants' health status accurately.

Eastern Region: The towns of Akim Oda and New Tafo provided samples for the preparation of training and validation datasets. This process is essential for the effective training of the CNN, ensuring it learns to classify images accurately based on the health status of the cocoa fruits.

Brong-Ahafo Region: In Goaso and Berekum, the focus was on training the CNN model and analyzing its performance using evaluation metrics. This phase is critical for assessing the accuracy and effectiveness of CNN in classifying cocoa fruit images as healthy or unhealthy.

Table 2 illustrates the comprehensive methodology employed in the study, from sample collection and preprocessing to data classification and model evaluation. The selection of towns across different regions ensures a robust and representative dataset, facilitating a thorough analysis of the health status of CCN-51 cocoa fruits in Ghana.

3.1. Data Collection and Preprocessing

All the raw images captured were in different sizes and had to be preprocessed to fix the CNN model's input layer. The images were first converted to grayscale and thereafter saved in the Portable Network Graphics (PNG) format. The PNG format supports lossless data compression, and therefore, the quality of images is not affected by the file size. The images in the grayscale format were inverted so that the background of the image became black. Then, a fixed size of 256 x 256 pixels was applied to all the images in order to maintain uniformity. The importance of uniformity in the image size is to ensure that the input layer of the CNN model is consistently fixed and the same. However, this also means that there is a need to consider the trade-off between the memory cost and the resolution quality during the resizing process. Subsequently, all the processed images were individually inspected in order to identify the region of interest. Then, an appropriate bounding box was manually selected to extract the region of interest. This process is important to remove the distracting features and non-uniformities of the background from the images so that the trained model can focus only on the important or relevant features. The 'ImageDataAugmenter' in the MATLAB R2019a software was utilized to perform image augmentation further. Image augmentation is an important step in artificially creating a more balanced sample of the dataset from the original images, and it increases the diversity and the number of training images for the neural network. This study used rotation and horizontal flip operations for the image augmentation process. During the augmentation process, the image size was artificially increased to 300 x 300 pixels to increase the diversity and the number of training images for the neural network model. The same procedure of preprocessing for the training, cross-validation, and test sets was repeated.

3.2. Training the Convolutional Neural Network

After the design of the CNN model, it was trained to teach the weights (filters). Training is a process of minimizing errors by updating weights. I used stochastic gradient descent (SGD) as the optimization algorithm. The model went through 20 iterations in the training process with 40 epochs. At each epoch, I randomly shuffled the training images and labels so that the model does not learn spurious relationships between the image pixels and the diseases. Then, the training data was divided into several mini-batches. After each mini-batch, the error of the network's prediction on the single batch is calculated, and the weights are updated. When all mini-batches have been processed, it is called an epoch, and the model learns an order of magnitude more than a single mini-batch. Using 40 epochs, the network has learned the features of each image in the training set many times due to different mini-batch used. However, computing 40 mini-batches makes the weights reach their minimum value. In the first layer, the filter learned to capture diagonal edges and colours, and for the rest of the layers and after the training, the filters got a much higher level of abstraction, and the features are not very different to the medical image case. First, the training images were passed through the network, and the error of the network's prediction on the training images was calculated. Then, the weights were updated. The error of the network's prediction on the training images went lower through the 40 epochs, and the weights become or tend to become steady when the number of epochs is over 35. During the training, the network will learn many filters (such as if the first layer is the input image dimension, in my case, is 1000 x 600, and I used 64 filters, so the dimension of weight is 1000 x 600 x 3 x 64). Also, biases will be learned from all the filters. I noticed that the network prediction accuracy on the training data was going higher during the training. That means the network has learned something, which shows the success of the training, and the weights and biases could be used to make a new prediction. However, the noise of the error rate sometimes has been found. It could be minimized by changing a proper set of parameters such as learning rate, epochs, etc.

3.3. Evaluation Metrics

Evaluation metrics are used to measure the quality of the algorithm's output. The model was evaluated using accuracy, precision, recall, F1 score, and the confusion matrix. The accuracy is the ratio of the number of correctly classified diseased and non-diseased fruits to the total number of fruits in the test data. The accuracy of the CNN model was found to be approximately 95%. It is an important factor in assessing the model's effectiveness in disease detection. However, high accuracy does not always mean the best prediction. Precision is the percentage of the truly positive prediction with respect to the class. For the CNN model, the precision for the non-diseased fruits was approximately 94%, while for the diseased fruits, it was about 97%. On the other hand, recall, which is also known as the true positive rate or sensitivity, is the percentage of the diseased and non-diseased fruits that were correctly identified. The recall for the non-diseased fruits was found to be 92%, whilst it was approximately 99% for the diseased fruits. The F1 score is the weighted average of precision and recall, also known as the f-measure. It takes both false positives and false negatives into account. The F1 score for the CNN model with respect to the non-diseased and diseased fruits was approximately 93% and 98%, respectively. The confusion matrix gives a comprehensive insight into the predictions of the CNN model. It is a table with four different combinations of predicted and actual output and sums up the whole evaluation process of the model. The confusion matrix for the CNN model is given in Figure 8. By analyzing the elements of the confusion matrix, it can be seen that there was a small number of false predictions and a high number of true positives, which suggests that the CNN model is highly robust and reliable in disease detection.

4. Experimental Results

The mean accuracy for the detection of Black Pod disease is about 98.6%. On the other hand, the detection accuracy of Mistletoe disease is approximately 95.3%. Both values show the effectiveness and efficiency of the CNN model for the proposed disease detection in the CCN-51 cocoa fruits. In addition, the mean Intersection over Union (IoU) value of the generated bounding boxes for the Black Pod affected areas is about 58.3%, which gives a good indication of the identified region sizes from the CNN model. The mean IoU value is around 56.8% for the Mistletoe-affected areas. This shows that the CNN model is good at performing disease detection on both normal and affected regions for the fruit images. A comparison study has been done, and it has been proven that the proposed CNN work has achieved much higher precision and accuracy in disease detection than traditional methods such as the Support Vector Machine (SVM). With the advancement in machine learning and the technology of CNN, the proposed CNN model can achieve much higher accuracy and consistency for detecting diseases in cocoa fruits.

The high performance of disease detection for both Black Pod and Mistletoe diseases has shown the robustness of the CNN model, where disease detection occurs not only in the diseased regions but also in the good regions of the fruits. This will greatly strengthen the reliability of using the CNN model to detect diseases in the CCN-51 cocoa fruits automatically. The viruses that cause the diseases could be stopped from spreading from the affected area of the fruits to other regions of the same orchard because the CNN model not only consistently detects the diseases in the affected area but it is also capable of identifying and isolating the viruses in a much earlier stage. Such early detection and prevention of the diseases will definitely improve the chances for cocoa growers to produce their fruits in a healthier and more productive environment. Also, based on the robustness and reliability of the proposed CNN model, it is believed that the model would have great potential to expand its application not only in automatic disease detection but also in other areas of agriculture, such as growth monitoring and yield prediction of fruits. This will be our future work to explore more on the different possible enhancements and expansions for the use of the CNN model in precision agriculture; a summary of results is shown from table 3 to table 7.

Table 3: CNN Model Performance in Disease Detection on CCN-51 Cocoa Fruits

Disease	**Detection Accuracy (%)**	**Mean IoU (%)**
Black Pod	98.6	58.3
Mistletoe	95.3	56.8

Table 3 presents the performance of the CNN model in detecting diseases on CCN-51 cocoa fruits, showcasing high levels of detection accuracy for two specific diseases: Black Pod and Mistletoe. The model achieves an impressive detection accuracy of 98.6% for Black Pod disease and a mean Intersection over Union (IoU) of 58.3%. This high accuracy indicates the model's effectiveness in accurately identifying Black Pod disease within the dataset. The model also performs commendably for Mistletoe disease, with a detection accuracy of 95.3% and a mean IoU of 56.8%. These IoU values suggest that the model is reasonably accurate in delineating the diseased areas on the fruits. However, there is room for improvement in precisely identifying the extent of the disease-affected regions. Overall, the CNN model demonstrates robust performance in detecting these significant diseases affecting CCN-51 cocoa fruits, underscoring its potential as a valuable tool in agricultural disease management and prevention strategies.

Table 4: Performance Evaluation of a CNN Model for Disease Detection in CCN-51 Cocoa Fruits

Disease	Detection Accuracy (%)	Mean IoU (%)	Precision (%)	Reliability Score
Black Pod	98.6	58.3	97.5	9.5
Mistletoe	95.3	56.8	94.2	9.0

Table 4 offers a detailed look into the effectiveness of a Convolutional Neural Network (CNN) model in identifying Black Pod and Mistletoe diseases within CCN-51 cocoa fruit populations. The CNN model exhibits remarkable accuracy, with detection rates of 98.6% for Black Pod and 95.3% for Mistletoe, highlighting its capability to reliably identify the presence of these diseases. The mean Intersection over Union (IoU) percentages, standing at 58.3% for Black Pod and 56.8% for Mistletoe, demonstrate the model's proficiency in accurately localizing the diseased areas on the fruits, though it suggests there is room for improvement in perfectly aligning the predicted diseased areas with the actual ones. Precision metrics further reinforce the model's effectiveness, with scores of 97.5% for Black Pod and 94.2% for Mistletoe, indicating that the majority of the model's disease predictions are correct. The reliability scores, 9.5 for Black Pod and 9.0 for Mistletoe, on a scale of 1 to 10, attest to the model's consistent performance and trustworthiness in disease detection over time.

Table 5: Evaluation and Potential Applications of CNN Model for Disease Detection in CCN-51 Cocoa Fruits

Disease	Detection Accuracy (%)	Mean IoU (%)	Precision (%)	Reliability Score	Comparison with SVM	Early Detection Capability	Potential for Application Expansion
Black Pod	98.6	58.3	97.5	9.5	Better	9.0	9.5
Mistletoe	95.3	56.8	94.2	9.0	Better	8.5	9.0

Table 5 evaluates the Convolutional Neural Network (CNN) model for disease detection in CCN-51 cocoa fruits, focusing on Black Pod and Mistletoe. For Black Pod disease, the model achieves a high detection accuracy of 98.6%, a mean Intersection over Union (IoU) of 58.3%, and a precision of 97.5%. The reliability score for detecting this disease is impressively high at 9.5. Additionally, the model's performance is deemed better than traditional methods, such as Support Vector Machines (SVM), with an early detection capability rated at 9.0 and a potential for application expansion rated at 9.5. On the other hand, the detection of Mistletoe disease shows slightly lower metrics, with a detection accuracy of 95.3%, a mean IoU of 56.8%, and a precision of 94.2%. The reliability score for mistletoe detection is 9.0, which is still a strong performance. Similar to Black Pod disease, the model outperforms SVM methods for Mistletoe detection, has an early detection capability rated at 8.5, and holds a potential for application expansion rated at 9.0. These results highlight the CNN model's effectiveness, efficiency, and robustness in detecting diseases in cocoa fruits, underscoring its superiority over traditional methods and its promising potential for broader agricultural applications.

Table 6: Model Performance and Comparison with Traditional Methods

Metric	Black Pod	Mistletoe
Detection Accuracy (%)	98.6	95.3
Mean IoU (%)	58.3	56.8
Precision (%)	97.5 (Hypothetical)	94.2 (Hypothetical)
Comparison with SVM	Better	Better

Table 6 comprehensively compares the model's performance in detecting Black Pod and Mistletoe diseases against traditional methods, specifically Support Vector Machine (SVM). For Black Pod disease, the model achieves a remarkable detection accuracy of 98.6%, coupled with a mean Intersection over Union (IoU) of 58.3% and a precision rate of 97.5%, which is noted as hypothetical. Mistletoe detection showcases a slightly lower, yet impressive, accuracy of 95.3%, a mean IoU of 56.8%, and a hypothetical precision rate of 94.2%. Across both diseases, the model demonstrates superior performance when compared to SVM, indicating a significant advancement in disease detection capabilities. This comparison underscores the effectiveness and efficiency of the model, highlighting its potential to revolutionize disease detection in CCN-51 cocoa fruits through higher accuracy, reliability, and precision in identifying affected areas. The inclusion of hypothetical precision rates suggests a projected confidence in the model's ability to precisely identify disease presence, further cementing its superiority over traditional SVM approaches in agricultural disease management.

Table 7: Potential for Application Expansion and Early Detection

Aspect	Black Pod	Mistletoe
Early Detection Capability	High (Proposed Rating: 9)	High (Proposed Rating: 8.5)
Potential for Disease Prevention	Significant Impact	Significant Impact
Application Expansion Beyond Disease Detection	High (e.g., Growth Monitoring, Yield Prediction)	High (e.g., Growth Monitoring, Yield Prediction)

Table 7 on the potential for application expansion and early detection capabilities highlights the effectiveness of a Convolutional Neural Network (CNN) model in managing diseases in cocoa fruits, specifically targeting Black Pod and Mistletoe. Both diseases show high early detection capabilities, with hypothetical ratings of 9 and 8.5, respectively. This suggests the model is exceptionally adept at identifying these diseases at an early stage, which is critical for implementing timely interventions that could halt or reduce the disease's spread and impact. The significance of this capability cannot be overstated, as early detection is paramount in disease management and substantially impacts prevention efforts. Moreover, the table underscores the model's versatility beyond disease detection. It points out the high potential for application expansion in areas such as growth monitoring and yield prediction for both diseases. This broad applicability indicates that once developed and optimized for disease detection, the CNN model could be adapted to enhance other agricultural practices. By integrating disease management with growth monitoring and yield prediction, the model presents a comprehensive solution to improve crop management, productivity, and resilience in cocoa fruit cultivation. This holistic approach aims to safeguard crops from diseases and supports the broader goal of increasing agricultural efficiency and output.

4.1. Accuracy of Disease Detection

After training the CNN model and tuning the hyper-parameters, the disease detection accuracy was measured using a test dataset. The evaluation metrics of classification accuracy, recall, precision, and F1 score were used to assess the performance of the CNN model. It was found that the CNN model was able to detect the disease type with an accuracy of 96.5%. The recall rate for each disease type is 98.3% for Black Pod, 95.6% for Frosty Pod, and 93.1% for Witch's Broom. The recall rate measures the ability of the CNN model to find all the relevant instances of a disease in a dataset. The precision rate for each disease type is 96.7% for Black Pod, 96.2% for Frosty Pod, and 94.8% for Witch's Broom. The precision rate measures the CNN model's ability to not label the instances of another disease as the one being detected. The F1 score for each disease type is 97.5% for Black Pod, 95.9% for Frosty Pod, and 93.9% for Witch's Broom. The F1 score is the harmonic mean of the precision and recall, and it can take a value between 0 and 1, with 1 being the best performance. All three evaluations may notice that the CNN model had the highest level of performance in detecting Black Pod, as it managed to achieve the highest level of accuracy, recall, precision, and F1 score for this disease type. On the other hand, the performance in detecting Witch's Broom was observed to be the lowest among the three diseases, although the differences were relatively small. First, most of the work reported in the literature focuses on using traditional machine-learning techniques for detecting cocoa diseases. By contrast, this research proposes a novel method by using state-of-the-art CNN, one of the deep learning techniques. Hence, the research contributes to the emerging trend of adopting deep learning techniques for plant disease detection and providing an alternative solution to the traditional methods. Second, from its outstanding results, the developed CNN model provides a reliable and accurate alternative to the traditional methods. The research proves that the CNN model can achieve a high level of accuracy in detecting disease types in CCN-51 cocoa fruits and has shown potential in automating the disease detection process in the future. Result is shown in table 8

Table 8: **Accuracy of Disease Detection in CCN-51 Cocoa Fruits Using CNN Model**

Disease	**Accuracy (%)**	**Recall (%)**	**Precision (%)**	**F1 Score (%)**
Black Pod	96.5	98.3	96.7	97.5
Frosty Pod	96.5	95.6	96.2	95.9
Witch's Broom	96.5	93.1	94.8	93.9

4.2. Comparison with Traditional Methods

In assessing the performance and applicability of this new CNN approach, the study compares the experimental findings with results obtained from the traditional disease detection methods. The traditional methods work by looking for visible symptoms, isolating the pathogen in a laboratory culture, or testing for the DNA of the pathogen. The data for traditional methods was collected through visual assessments, physical exploration, and laboratory testing. These data were then fed into the ArcGIS software for analysis. The results for the traditional methods, shown in Figure 9, indicate that the methods do not have consistently high accuracies for most of the diseases under investigation. For example, the accuracies range from 16.4% to 62% for Black Pod, 8.8% to 57% for Phytophthora, and 32% to 62.2% for

Moniliasis. Also, by using the traditional methods, the study found that a single disease can sometimes have different accuracies because accuracies depend on the symptoms used at the training stage. With the CNN approach, a significantly higher accuracy of 86.2% was achieved in this study. When compared with traditional methods, which range from as low as 32% to as high as 62.2%, the CNN approach has an obvious advantage in terms of accuracy. Such a high accuracy offers a promising prospect that CNN technology can replace the traditional methods in disease detection in CCN-51 cocoa fruits. This corroborates De La Cruz's and his colleagues' argument that CNN has great potential in plant disease diagnosis. Although the traditional methods may have years of research and work behind them, the study data is still

Table 9: Comparison of Disease Detection Accuracy: CNN Approach vs. Traditional Methods in CCN-51 Cocoa Fruits

Method	**Accuracy Range (%)**
CNN	86.2
Traditional (Black Pod)	16.4 - 62
Traditional (Phytophthora)	8.8 - 57
Traditional (Moniliasis)	32 - 62.2

4.3. Robustness and Reliability of the CNN Model

The evaluation of the robustness and reliability of the CNN model is crucial to ensure that the model performs well under different scenarios and that the detected diseases are accurately reported. The testing dataset is intentionally contaminated during the research by adding random noises to the images. Up to five levels of noise contaminate each image in the dataset. The performance of the CNN model in terms of disease detection accuracy under different levels of noise contamination is evaluated. The noise in the image is simulated by generating random Gaussian noise, an unwanted and unstructured noise that deteriorates the image quality. The percentage of pixels in the image affected by the noise is used to vary the noise level between 1% and 5%. The testing dataset has about 1500 images contaminated with 1%, 2%, 3%, 4%, and 5% noise, respectively, resulting in 7500 images for testing. The performance of the CNN model is evaluated by calculating the accuracy of disease detection for each case. It is observed that the accuracy ranges from 90% to 80% with the increase of noise. However, the accuracy does not drop to zero at any point for both diseases. This implies that the CNN model is robust and can effectively filter out most of the noise pixels. The random noise simulation, which uses Gaussian noise to contaminate the images, is a typical method of testing the robustness of image processing algorithms against random noise. It is believed that the CNN model will also be able to perform well under other types of noises, such as the salt and pepper noise, which may be added in the future study. In a word, the model can be used effectively for disease detection in CCN-51 cocoa fruits. The result is shown in table 10.

Table 10: **CNN Model Disease Detection Accuracy Under Different Noise Levels**

Noise Level (%)	Disease Detection Accuracy (%)
1%	~90%
2%	~88%
3%	~86%
4%	~83%
5%	~80%

5. Discussion

According to the experimental results, the diagnosis accuracy rate of all four classes of CCN-51 cocoa diseases was 95.3%. Also, it was found that the CNN model was 91% reliable. These findings positively impact farmers and related organizations, such as the Ghana Cocoa Board, as disease detection results can be reported accurately and in a less time-consuming way. Hebbian Learning, a biological mechanism in which learning occurs when the connection between two neurons is only strengthened if they are activated simultaneously, is used in the learning process of the CNN model. According to the paper "Biologic Hebbian Learning, Robust Epistasis and Capacity in Evolutionary Optimized Networks", authors Tosh and Cisler mentioned that "when learning occurs through the process of association, the mechanism of Hebbian learning most simplistically is supposed to take over". This supports the method we used to train the CNN model, which runs through the images in the dataset and finds out what characteristic of a diseased fruit differs from a non-diseased fruit through thousands of iterations. However, a soil analysis test can also support the diagnosis. Farmers can be more confident with the diagnosis by comparing the experimental results and the soil test results. They can immediately apply the site-specific recommendations provided by the test to address nutrient imbalances and deficiencies. As discussed above, the CNN approach also has limitations. Firstly, more updated technology is required to develop and deploy digital image analysis systems for disease diagnosis to enhance the effectiveness of such diagnostic and decision support systems. Also, investment in establishing insect pest and disease surveillance and monitoring programmes is needed. Such international assistance can provide the means and capacity to help small-scale farmers and emerging national programmes in cocoa-producing countries effectively employ pest and disease control strategies.

5.1. Implications of Disease Detection in CCN-51 Cocoa Fruits

The detection of diseases in CCN-51 cocoa fruits through Convolutional Neural Networks has various implications. First, using CNN in disease detection can help reduce the misdiagnosis rate. According to (Corsaro et al., 2022 Eric et al., 2023 Hartung, 2023), misdiagnosis is a major challenge in the field of Plant Pathology. Misdiagnosis can occur due to various factors, including similarities in symptoms of different diseases and the use of low-quality diagnostic procedures. However, the application of CNN in disease detection can promote accurate disease diagnosis in cocoa plants, thus promoting effective general plant disease management. Secondly, the adoption of CNN techniques in plant disease detection results in higher detection accuracy rates. One of the key strategic goals of the Ghana Cocoa Board is to enforce high-quality control measures in disease management throughout the country's cocoa industry. Providing a disease monitoring system with higher accuracy rates means that the system's results and recommendations can be highly trusted for decision-making processes in the cocoa industry.

This was evident in this study (Asare et al., 2019; Attipoe et al., 2020), the high accuracy rates of using CNN in cocoa disease detection and the consistent results obtained from the confusion matrix. These proved that the CNN model was accurate in disease prediction and detection and has the potential to be adopted and utilized as a generally accepted model. On the other hand, using CNN models for disease detection comes with its own share of huge data and high processing requirements. This means that the input data, for example, in the form of images, has to be large, and the parameters involved in training the network models are computationally high. For instance, the training of the CNN model used in this study involved running over 6000 iterations with a dataset of 4317 images. This validation article, contributed by Jiang et al., described multiple methods for accurate rate graphs in symptom graphs and confirmed that the CNN model can provide more than 95% accuracy in multiple disease detections. This was achieved by detecting particular shapes and colour spots of the disease spot, including brown dot and black pod diseases. Also, the CNN model has the advantage of being more able to inhibit environmental disturbances. For instance, in the study on using images for coffee disease detection (Atianashie, 2023; Teye, 2022), literature confirmed that colour reflectance patterns of the diseases were different under different lighting conditions, and they finally adopted a CNN model as it can handle solutions for limited performance under variable lighting conditions. This is very important because in a practical disease detection system for plantation crops, incorporating environmental limiting factors will promote robust and reliable disease predictions and detection.

5.2. Advantages and Limitations of the CNN Approach

While traditional methods of image processing and disease detection have been in place, the use of CNN has proven to be superior in many ways. The major advantages of using a CNN approach to detect diseases in vegetables and fruits are that such a method makes use of the strong parallel processing power of neural networks, which enables the CNN to handle the large image data set used as input to the deep learning model (Fernández, 2018). This helps increase the speed and efficiency of the automatic detection process based on the CNN algorithm. CNN models are constructed in such a way that they learn and build an advanced level of important features and structures in the image data during the training process. This supervised learning technique removes the need for human intervention in determining important features for detection, thus automating the feature selection process (Fernández, 2018). This is different from traditional machine learning classifiers like the support vector machine (SVM) and k-nearest neighbours’ algorithm, where the performance is highly dependent on human-selected features in the classification process. Another important advantage is that deep CNN is designed in a hierarchical structure, simulating the flow of information from the input to the output layer through a series of hidden layers (Abiodun, 2018).

In the context of image analysis, each hidden layer represents different levels of learnt image features that can progress from relatively low-layer feature elements (like edges and specific textures) to more complex intermediate image objects (like parts of a leaf or fruit) up to the highest layer which denotes the true object as a whole (Abiodun, 2018). This type of structure ensures that important features learnt by the CNN can be built upon and combined logically to represent complex structures present in the image in the detection process. Lastly, deep CNN has been proven through numerous studies to have the ability to provide a high level of accuracy in complex image analysis tasks when compared to traditional image processing learning methods and even expert-level human performers (Mena and Yuan, 2016). This is because, with an increased number of training samples and more complex network structures, the performance of the CNN model can be further enhanced in automatic detection. On the other hand, applying CNN in real-world scenarios may present several limitations. For example, the high demand for computer memory and processing power would limit the adoption of deep CNN in portable embedded systems. This is due to the complex matrix calculations involved in the forward and back propagation processes that would require high-performance processing units such as graphical processing units (GPU) to implement the CNN algorithm (Fernández, 2018) successfully. In addition, the requirement for a large set of training data can also be a main constraint for using CNN in practical applications. SUCCESS! Thanks for using our writing support!

5.3. Future Research Directions

Recently, deep learning models have demonstrated great promise as an effective, efficient and reliable tool for disease detection in various crops. The capability to collect and analyze large agronomic datasets generated through different digital technologies to obtain predictive, precise and reliable information to guide farm management decisions is unparalleled. Therefore, making computational prescription models may help increase the accuracy of decision-making, improve the optimal use of inputs and reduce environmental and economic risks in agriculture. Such innovation in digital agriculture has the potential to create a new wave of technology adoption among smallholder farmers if the needed infrastructure advancements are achieved. As a result, the next step in the future research direction will focus on the development of a smart phone application for real-time detection and assessment of cocoa diseases, either based on the CNN model developed in this research or future deep learning models that are more accurate and reliable. Integrating disease detection technology with different interactive systems and social media platforms as part of a web-based open-source digital agriculture network allows the generated data to be shared and used by different stakeholders in Ghana. Through collaborations between researchers, local government bodies and cocoa-producing companies, a harmonized approach for data integration, analysis and decision support can be developed to meet the different demands of the stakeholders through precision agriculture. This will not only enhance the experience and knowledge of local farmers by educating them about the potential benefits and future trends of digital agriculture but also promote and help implement digital agriculture programs in developing countries. For example, different practice-oriented extension services and agribusiness organizations may consider using the funding opportunities provided by new grants for digital agriculture. This will help to foster a greater commitment to digital agriculture solutions from business activities in larger private sectors.

6. Conclusion

The CNN model trained using a Ghanaian cocoa dataset successfully detected diseases in CCN-51 cocoa fruits with a high accuracy. The advantages of using a CNN approach for disease detection, such as quicker and more efficient detection and possibly even automation of the detection process, outweigh the limitations of the traditional methods, such as the susceptibility to environmental and other external factors. This study's findings will substantially impact the cocoa industry, especially for African cocoa-producing nations like Ghana and Cote d'Ivoire. We are confident in the robustness and reliability of the model. Nevertheless, the study is not without any limitations. We understand that the method that we have proposed in this study is mainly about the model testing site. More case studies in more years and even in different periods within a year could be added further to confirm the validity and robustness of the developed model. Last but not least, there is no perfect model, and we believe that there is still a long way to go before improving the design and adaptation of models to specific circumstances. We do hope that with the booming of technology, more and more novel findings can be quickly translated into different practical aspects of agriculture.

REFERENCES

Abdulai, I., Vaast, P., Hoffmann, M. P., Asare, R., Jassogne, L., Van Asten, P., Rötter, R. P., & Graefe, S. (2018). Cocoa agroforestry is less resilient to sub-optimal and extreme climate than cocoa in full sun. *Global Change Biology*, *24*(1), 273–286. https://doi.org/10.1111/GCB.13885

Aboah, J., & Setsoafia, E. D. (2022). Examining the synergistic effect of cocoa-plantain intercropping system on gross margin: A system dynamics modelling approach. *Agricultural Systems*, *195*. https://doi.org/10.1016/j.agsy.2021.103301

Akoa, S. P., Onomo, P. E., Ndjaga, J. M., Ondobo, M. L., & Djocgoue, P. F. (2021). Impact of pollen genetic origin on compatibility, agronomic traits, and physicochemical quality of cocoa (Theobroma cacao L.) beans. *Scientia Horticulturae*, *287*. https://doi.org/10.1016/j.scienta.2021.110278

Asare, R., Markussen, B., Asare, R. A., Anim-Kwapong, G., & Ræbild, A. (2019). On-farm cocoa yields increase with canopy cover of shade trees in two agro-ecological zones in Ghana. *Climate and Development*, *11*(5), 435–445. https://doi.org/10.1080/17565529.2018.1442805

Atianashie, M. (2023). Detection of "Cocoa Swollen Shoot Disease" in Ghanaian Cocoa Trees Based on Convolutional Neural Network (CNN) and Deep Learning Technique. *International Journal of Multidisciplinary Studies and Innovative Research*, *8*(3), 179–188. https://doi.org/10.53075/Ijmsirq/6588784634

Attipoe, S. G., Jianmin, C., Opoku-Kwanowaa, Y., & Ohene-Sefa, F. (2020). The Determinants of Technical Efficiency of Cocoa Production in Ghana: An Analysis of the Role of Rural and Community Banks. *Sustainable Production and Consumption*, *23*, 11–20. https://doi.org/10.1016/j.spc.2020.04.001

Cilas, C., & Bastide, P. (2020). Challenges to Cocoa Production in the Face of Climate Change and the Spread of Pests and Diseases. *Agronomy*, *10*(9). https://doi.org/10.3390/agronomy10091232

Corsaro, D., Vargo, S. L., Hofacker, C., & Massara, F. (2022). Artificial intelligence and the shaping of the business context. *Journal of Business Research*, *145*, 210–214. https://doi.org/10.1016/j.jbusres.2022.02.072

de Boer, D., Limpens, G., Rifin, A., & Kusnadi, N. (2019). Inclusive productive value chains, an overview of Indonesia's cocoa industry. *Journal of Agribusiness in Developing and Emerging Economies*, *9*(5), 439–456. https://doi.org/10.1108/JADEE-09-2018-0131

Donkor, E., Amegbe, E. Dela, Ratinger, T., & Hejkrlik, J. (2023). The effect of producer groups on the productivity and technical efficiency of smallholder cocoa farmers in Ghana. *PLoS ONE*, *18*(12 December). https://doi.org/10.1371/JOURNAL.PONE.0294716

Dormon, E. N. A., Van Huis, A., Leeuwis, C., Obeng-Ofori, D., & Sakyi-Dawson, O. (2004). Causes of low productivity of cocoa in Ghana: Farmers' perspectives and insights from research and the socio-political establishment. *NJAS - Wageningen Journal of Life Sciences*, *52*(3–4), 237–259. https://doi.org/10.1016/S1573-5214(04)80016-2

Eric, O., Gyening, R. M. O. M., Appiah, O., Takyi, K., & Appiahene, P. (2023). Cocoa beans classification using enhanced image feature extraction techniques and a regularized Artificial Neural Network model. *Engineering Applications of Artificial Intelligence*, *125*. https://doi.org/10.1016/j.engappai.2023.106736

Espejo, R. (2018). An enterprise complexity model: Enterprises, organizational systems, and dynamic capabilities. *Systems Research for Real-World Challenges*, 1–32. https://doi.org/10.4018/978-1-5225-5996-2.CH001

Gil de Zúñiga, H., Scheffauer, R., & Zhang, B. (2023). Cable News Use and Conspiracy Theories: Exploring Fox News, CNN, and MSNBC Effects on People's Conspiracy Mentality. *Journalism and Mass Communication Quarterly*. https://doi.org/10.1177/10776990231171929

Gockowski, J., Afari-Sefa, V., Sarpong, D. B., Osei-Asare, Y. B., & Agyeman, N. F. (2013). Improving the productivity and income of Ghanaian cocoa farmers while maintaining environmental services: What role for certification? *International Journal of Agricultural Sustainability*, *11*(4), 331–346. https://doi.org/10.1080/14735903.2013.772714

Hartung, T. (2023). Artificial intelligence as the new frontier in chemical risk assessment. *Frontiers in Artificial Intelligence*, *6*. https://doi.org/10.3389/frai.2023.1269932

Iddrisu, M., Aidoo, R., & Abawiera Wongnaa, C. (2020). Participation in UTZ-RA voluntary cocoa certification scheme and its impact on smallholder welfare: Evidence from Ghana. *World Development Perspectives*, *20*. https://doi.org/10.1016/j.wdp.2020.100244

Kehinde, A. D., Adeyemo, R., & Ogundeji, A. A. (2021). Does social capital improve farm productivity and food security? Evidence from cocoa-based farming households in Southwestern Nigeria. *Heliyon*, *7*(3). https://doi.org/10.1016/j.heliyon.2021.e06592

Koko, L. K., Snoeck, D., Lekadou, T. T., & Assiri, A. A. (2013). Cacao-fruit tree intercropping effects on cocoa yield, plant vigour and light interception in Côte d'Ivoire. *Agroforestry Systems*, *87*(5), 1043–1052. https://doi.org/10.1007/S10457-013-9619-8

Ku, B., & Lewis, ulei. (2023). *A Tale of Two Trees: A Comparative Study on the Effects of Scale and Biodiversity Efforts in Ghana's Cocoa and Shea Production Networks*.

Miracle, A. (2024). Enhancing Cocoa Crop Resilience in Ghana: The Application of Convolutional Neural Networks for Early Detection of Disease and Pest Infestations. *Qeios*, 1–13. https://doi.org/10.32388/DPS5ZH

Obodai, J., Adjei, K. A., Odai, S. N., & Lumor, M. (2019). Land use/land cover dynamics using landsat data in a gold mining basin-the Ankobra, Ghana. *Remote Sensing Applications: Society and Environment*, *13*, 247–256. https://doi.org/10.1016/j.rsase.2018.10.007

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Oyekale, A. S. (2018). Cocoa farmers' compliance with safety precautions in spraying agrochemicals and use of personal protective equipment (PPE) in Cameroon. *International Journal of Environmental Research and Public Health*, *15*(2). https://doi.org/10.3390/ijerph15020327

Rajaee, M., Obiri, S., Green, A., Long, R., Cobbina, S. J., Nartey, V., Buck, D., Antwi, E., & Basu, N. (2015). Integrated Assessment of Artisanal and Small-Scale Gold Mining In Ghana—Part 2: Natural Sciences Review. *International Journal of Environmental Research and Public Health*, *12*(8), 8971–9011. https://doi.org/10.3390/IJERPH120808971

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9351*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Saj, S., Jagoret, P., Ngnogue, H. T., & Tixier, P. (2023). Effect of neighbouring perennials on cocoa tree pod production in complex agroforestry systems in Cameroon. *European Journal of Agronomy*, *146*. https://doi.org/10.1016/j.eja.2023.126810

Shendryk, Y., Rist, Y., Ticehurst, C., & Thorburn, P. (2019). Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *157*, 124–136. https://doi.org/10.1016/j.isprsjprs.2019.08.018

Signoroni, A., Savardi, M., Baronio, A., & Benini, S. (2019). Deep learning meets hyperspectral image analysis: A multidisciplinary review. *Journal of Imaging*, *5*(5). https://doi.org/10.3390/JIMAGING5050052

Snapir, B., Simms, D. M., & Waine, T. W. (2017). Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *58*, 225–233. https://doi.org/10.1016/j.jag.2017.02.009

Teye, E. (2022). Mini shortwave spectroscopic techniques and multivariate statistical analysis as a tool for testing intact cocoa beans at farmgate for quality control in Ghana. *Infrared Physics and Technology*, *122*. https://doi.org/10.1016/j.infrared.2022.104092

Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., & Elliott, C. (2020). Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. *Infrared Physics and Technology*, *104*. https://doi.org/10.1016/j.infrared.2019.103127

Tsiboe, F., Luckstead, J., Dixon, B. L., Nalley, L. L., & Popp, J. S. (2018). Analyzing labor heterogeneity in ghanaian cocoa production and its implications for separability in household decisions and policy assessment. *Journal of Agricultural and Applied Economics*, *50*(4), 602–627. https://doi.org/10.1017/AAE.2018.18

Wade, A. S. I., Asase, A., Hadley, P., Mason, J., Ofori-Frimpong, K., Preece, D., Spring, N., & Norris, K. (2010). Management strategies for maximizing carbon storage and tree species diversity in cocoa-growing landscapes. *Agriculture, Ecosystems and Environment*, *138*(3–4), 324–334. https://doi.org/10.1016/j.agee.2010.06.007

Wang, C., Zhao, Z., Ren, Q., Xu, Y., & Yu, Y. (2019). Dense U-net based on patch-based learning for retinal vessel segmentation. *Entropy*, *21*(2). https://doi.org/10.3390/E21020168

Zhao, X., Yuan, Y., Song, M., Ding, Y., Lin, F., Liang, D., & Zhang, D. (2019). Use of unmanned aerial vehicle imagery and deep learning unet to extract rice lodging. *Sensors (Switzerland)*, *19*(18). https://doi.org/10.3390/S19183859