Using Artificial Intelligence (AI) for Potato Disease Detection in Kenya

Strathmore University

Centre for Intellectual Property and Information Technology Law

Acknowledgements

CIPIT sincerely thanks the team whose collaborative efforts made this report possible. We would like to express our deep appreciation to the authors, Dr. Angeline Wairegi, Joseph Theuri Gitonga, and Josephine Kaaniru, for their invaluable contributions and expertise that shaped the content of this report. We would also like to acknowledge Florence Ogonjo for her exceptional editorial services, which significantly enhanced the clarity and quality of the report. We are grateful to Gloria Muthoka for her creative design of the report. Lastly, we wish to extend our gratitude to the administrative team for their unwavering guidance and support throughout the research and writing process.

Table of Contents

Table of Figures

Table of Tables

Abstract

Crop diseases threaten yield losses, food security, and livelihoods, especially in Kenya, where agriculture is a significant industry. Early detection and treatment of these illnesses can prevent crop losses, yet traditional disease detection methods are often inefficient, reducing agricultural yields and costing farmers money. This study explores the use of artificial intelligence (AI) to detect potato plant diseases through image classification to prevent losses.

The study evaluates the accuracy of an AI tool that uses Convolutional Neural Networks (CNN) for disease detection and classification. It also outlines the current state of potato disease detection and the integration of AI, considering challenges such as limited access to technology, insufficient training data, low technical knowledge among farmers, and potential disregard for indigenous agricultural knowledge. Recommendations to address these challenges include expanding training and capacity-building for farmers, facilitating access to technology and infrastructure in rural areas, availing quality training data, integrating indigenous knowledge into agricultural technology design, and establishing AI regulations.

The study also examines legal and policy considerations, including privacy, security, accountability, and AI regulatory frameworks, emphasizing the importance of protecting personal data and ensuring transparent and accountable technology use. By addressing these challenges, potato farmers can effectively use AI applications for early disease detection, preventing plant losses and improving yields and food security in Kenya.

Keywords: Convolutional Neural Networks, Potato disease detection, privacy, accountability

Glossary

Artificial Intelligence (AI) – these are computer programs and systems that can perform tasks that require human intelligence, such as making decisions, learning from data, understanding natural language and solving problems.[1]

Blight – a plant disease, typically one caused by fungi.[2]

Convolutional Neural Network (CNN) – this is a type of deep learning model that is created to process and analyse visual data such as videos and images; it utilises convolutional layers to automatically extract features from input data.[3]

Dataset - a collection of structured and organized data typically used for analysis and research.[4]

Deep Learning – this is a machine learning subfield that uses artificial neural networks with multiple layers, to learn and make predictions from data. It excels in tasks like image and speech recognition.[5]

Machine Learning (ML) - a subfield of artificial intelligence (AI) that focuses on designing algorithms and models capable of autonomously improving their performance by learning from data through pattern recognition, prediction, and adaptation.[6]

Neural Networks – this is a machine learning model composed of interconnected nodes that can learn from data to make predictions or decisions inspired by the human brain.[7]

Potato crop diseases - these are various plant diseases that can affect potatoes, such as early blight, late blight, common scab, pink rot, black scurf, fusarium dry root, silver scurf, and black dot.[8]

[1]IBM, 'What Is Artificial Intelligence (AI)?' (IBM2024) <https://www.ibm.com/topics/artificial-intelligence>.
[2]NPCK, 'What Every Potato Farmer Should Know about Late Blight and Early Blight – NPCK' (Npck.org2019) <https://npck.org/what-every-potato-farmer-should-know-about-late-blight-and-early-blight/> accessed 14 August 2024.
[3]Mayank Mishra, 'Convolutional Neural Networks, Explained' (Medium27 August 2020) <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>.
[4]IBM, 'What Is a Data Set?' (www.ibm.com) <https://www.ibm.com/docs/en/zos-basic-skills?topic=more-what-is-data-set>.
[5]IBM, 'What Is Deep Learning?' (www.ibm.com2023) <https://www.ibm.com/topics/deep-learning>.
[6]IBM, 'What Is Machine Learning?' (IBM2023) <https://www.ibm.com/topics/machine-learning>.
[7]AWS, 'What Is a Neural Network? AI and ML Guide - AWS' (Amazon Web Services, Inc.2023) <https://aws.amazon.com/what-is/neural-network/>.
[8]Hugo Campos and Oscar Ortiz, The Potato Crop (Hugo Campos and Oscar Ortiz eds, Springer International Publishing 2020) <https://link.springer.com/book/10.1007/978-3-030-28683-5>.

1. Introduction

1.1 Background

In Kenya, potatoes are the second most important staple food crop after maize, valued at 30 million USD annually.[9] Around 800,000 Kenyans directly profit from potato production, while 2.5 million people earn money from the potato industry.[10] Kenya is Africa's fourth-largest potato producer after Algeria, Egypt, and South Africa, but it has the lowest yield per acre of any of these nations.[11] The country also has the fourth lowest yield in East Africa at 8.6 tonnes per hectare, slightly better than Uganda's 4.3 tonnes per hectare.[12] However, production has consistently reduced over the last decade from 22 tonnes per hectare in 2008 to 8.6 tonnes per hectare in 2018.[13] This has been attributed to several issues, the most significant of which is the high prevalence of pests and diseases, some of which are seed and soil-borne.[14] Pests and diseases contribute to an estimated 80% reduction in production, threatening the availability of seeds and food security.[15] Small-scale farmers' harvests and incomes can be improved by assisting them in detecting and controlling crop pests and diseases. Most Sub-Saharan African countries have a fraction of the extension agents required to reach all their farmers, ranging from one agent for every 1,000 farmers to one for every 5,000-10,000 farmers, depending on the country.[16] This lack of adequate expertise to assist farmers further complicates disease detection and management among potato farmers. To help overcome this challenge, researchers are using digital technologies backed by Artificial Intelligence (AI) to bridge the extension gap.

Artificial Intelligence (AI) is the simulation of human intelligence in computers intended to think and behave like humans.[17] It also refers to any system that exhibits human-like characteristics such as learning and problem-solving.[18] AI is based on the idea that human intelligence should be described so that a computer can easily mimic it and perform tasks ranging from the most basic to the most complex.[19] Machine learning (ML) and computer vision have already enabled game-changing precision agriculture capabilities by allowing for the optimization of farm returns, preservation of natural resources, reduction of unnecessary fertilizer use, and identification of disease in crops and animals from remotely sensed imagery.[20] Thus, farmers can use smart mobile-based systems to accurately identify the different types of plant diseases. For instance, free smartphone apps such as PlantVillage Nuru are helping farmers to produce more food.[21] Nuru's cassava disease

[9]David Dudenhoefer, 'Smartphone App Helps Farmers Control Potato and Sweetpotato Diseases' (International Potato Center4 April 2022) <https://cipotato.org/blog/smartphone-app-help-farmers-control-potato-sweetpotato-diseases/>.

[10]Patrick Andati and others, 'Effect of Climate Smart Agriculture Technologies on Crop Yields: Evidence from Potato Production in Kenya' (2023) 41 Climate Risk Management 100539 <https://www.sciencedirect.com/science/article/pii/S2212096323000657>.

[11]CABI, 'Low Potato Yields in Kenya Not Down to Management Practices Alone, CABI Study Suggests - CABI.org' (CABI.org2022) <https://www.cabi.org/news-article/low-potato-yields-in-kenya-not-down-to-management-practices-alone-cabi-study-suggests/> accessed 14 August 2024.

[12]CABI, 'Surveillance of Potato Diseases in Kenya - CABI.org' (CABI.org10 March 2022) <https://www.cabi.org/projects/surveillance-of-potato-diseases-in-kenya/> accessed 14 August 2024.

[13]Agatha Ngotho, 'Scientists Launch Research into Potato Diseases to Boost Yields' (The Star18 July 2020) <https://www.the-star.co.ke/news/2020-07-18-scientists-launch-research-into-potato-diseases-to-boost-yields/> accessed 14 August 2024.

[14]CABI, 'Surveillance of Potato Diseases in Kenya

[15]Jean B Ristaino and others, 'The Persistent Threat of Emerging Plant Disease Pandemics to Global Food Security' (2021) 118 Proceedings of the National Academy of Sciences <https://www.pnas.org/doi/10.1073/pnas.2022239118>.

[16]David Dudenhoefer, 'Smartphone App Helps Farmers Control Potato and Sweetpotato Diseases' (International Potato Center4 April 2022) <https://cipotato.org/blog/smartphone-app-help-farmers-control-potato-sweetpotato-diseases/>.

[17]IBM, 'What Is Artificial Intelligence (AI)?'

[18]IBM, 'What Is Artificial Intelligence (AI)?

[19]Natnael Tilahun and Beakal Gizachew, 'Artificial Intelligence Assisted Early Blight and Late Blight Potato Disease Detection Using Convolutional Neural Networks' (2020) 8 Ethiop. J. Crop Sci 15 <https://www.ajol.info/index.php/ejcs/article/view/237248/224191> accessed 14 August 2024.

[20]Ravesa Akhter and Shabir Ahmad Sofi, 'Precision Agriculture Using IoT Data Analytics and Machine Learning' (2021) 34 Journal of King Saud University - Computer and Information Sciences <https://www.sciencedirect.com/science/article/pii/S1319157821001282>.

[21]AI in Agriculture: How Kenyan Farmers Benefit from PlantVillage Nuru App - Scholar Media Africa' (Scholar Media Africa17 October 2023) <https://scholarmedia.africa/agribusiness/ai-in-agriculture-how-kenyan-farmers-benefit-from-plantvillage-nuru-app/> accessed 14 August 2024.

diagnosis accuracy has been found to be higher (65%) than that of farmers (18%-31%) and that of experts (40%-58%).[22] Therefore, these technological detection methods could be applied on a large scale to aid in quick detection and management of potato diseases while providing expertise to farmers on their devices.

1.2 Problem Statement

Potatoes are one of the leading food crops in Kenya, but their production is hampered by the prevalence of pests and diseases, which results in significant yield losses. In most cases, especially in developing countries and small farms, farmers identify crop diseases with the naked eye based on visual symptoms. This and other traditional disease detection methods are time-consuming, labour-intensive, and often ineffective, leading to reduced crop yields and economic losses for farmers.[23] There is a need for a more efficient and practical approach to potato disease detection that can help farmers identify diseases early and take appropriate measures to manage them.[24]

The use of artificial intelligence (AI) has emerged as a promising approach for disease detection in agriculture, including potato crops.[25] However, with limited existing research on the use of AI for potato disease detection in Kenya, this study aims to examine the accuracy of a bespoke Machine Learning tool, providing valuable insights into the accuracy benefits of using AI for potato disease detection. The study also examines barriers to the adoption of AI in potato disease detection, identifying regulatory and policy interventions to protect farmers' privacy, promote fairness and accountability, and address the digital divide challenges faced by farmers when using AI in disease detection.

1.3 Objectives of the Study

To examine the adoption of AI in potato disease detection, including assessing a custom Machine Learning tool, analyzing adoption barriers, and considering policy and regulatory factors for deployment.

1.3.1 Specific Objectives

a. To evaluate the accuracy of a bespoke potato disease detection application.

b. To explore the use of AI in potato disease detection, taking into consideration the current state of potato disease detection in Kenya.

c. To identify the barriers to AI use in potato disease detection in Kenya.

d. To explore the legal and policy considerations in using AI for potato disease detection in Kenya.

1.4 Research Questions

a. How accurate is a bespoke AI potato disease detection application in potato plant disease detection?

b. How is AI currently being applied to potato disease detection within the broader context of potato disease detection in Kenya?

c. What barriers affect AI use in potato disease detection in Kenya?

d. What are the legal and policy considerations for using AI for potato disease detection in Kenya?

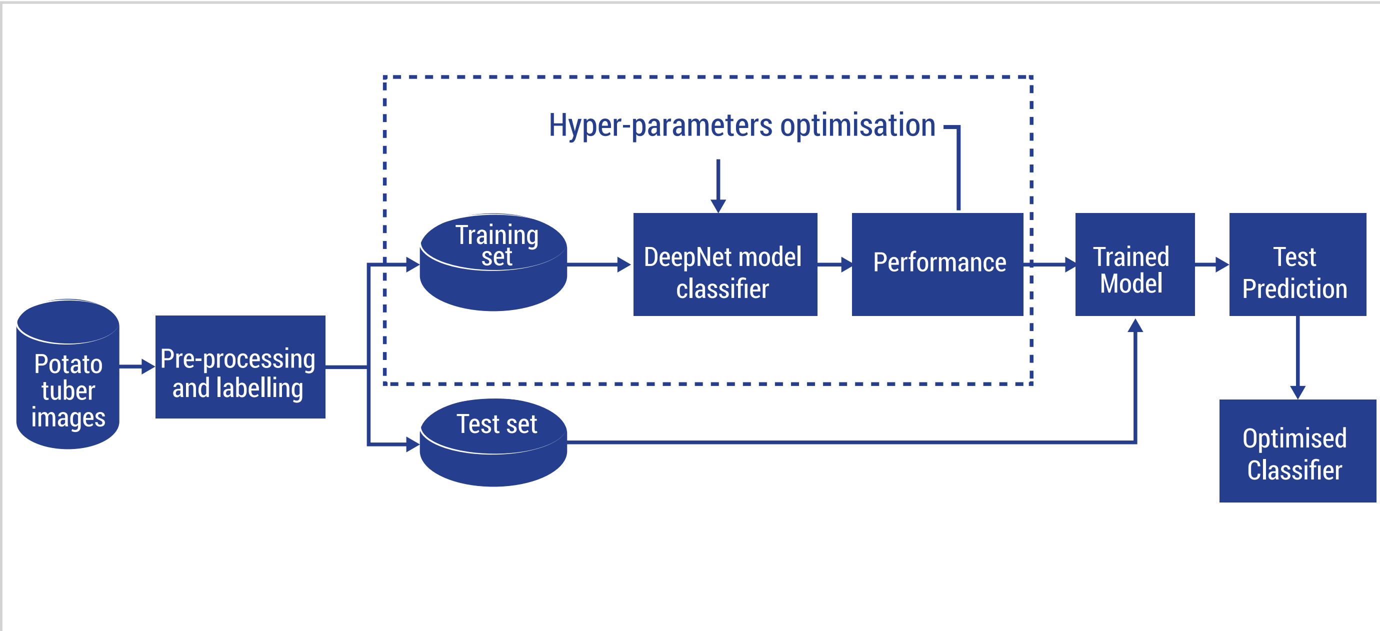
1.5 Conceptual Framework

The conceptual model for developing the Machine Learning tool is inspired by the deep learning process used for disease detection in potato tubers. Figure 2.1 illustrates the detailed steps involved in constructing the disease classifier. Initially, images of potato tubers are pre-processed and labelled. The entire dataset is then divided into a training set and a test set. The training set is used to train the DeepNet model (classifier), followed by performance evaluation. Hyper-parameters are

[22] Latifa M Mrisho and others, 'Accuracy of a Smartphone-Based Object Detection Model, PlantVillage Nuru, in Identifying the Foliar Symptoms of the Viral Diseases of Cassava-CMD and CBSD' (2020) 11 Frontiers in Plant Science 590889 <https://pubmed.ncbi.nlm.nih.gov/33391304/>.
[23] CABI, 'Surveillance of Potato Diseases in Kenya
[24] Houda Orchi, Mohamed Sadik and Mohammed Khaldoun, 'On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey' (2021) 12 Agriculture 9 <https://www.mdpi.com/2077-0472/12/1/9>.
[25] Simon Y Liu, 'Artificial Intelligence (AI) in Agriculture' (2020) 22 IT Professional 14 <https://ieeexplore.ieee.org/abstract/document/9098011>.

optimized during this process. The trained model is subsequently tested on the test set, leading to the development of an optimized classifier.

Figure 2.1 Conceptual framework



1.6 Research Methodology

In addition to developing a bespoke application, the research study takes a qualitative approach to comprehensively examine the research objectives. Additionally, secondary data containing training images for the AI platform was sourced from various online sources and documented accordingly. These images were used to develop the ML algorithm.

2. Developing the Machine Learning Application

2.1 Machine Learning

Machine Learning (ML) is a subfield of artificial intelligence (AI) that focuses on developing algorithms and models capable of autonomously improving their performance by learning from data through pattern recognition, prediction, and adaptation.[26] Machine Learning (ML) offers a promising approach to improving potato disease detection, especially for farmers in Kenya, where early diagnosis is crucial. Convolutional Neural Networks (CNNs), a key ML tool, excel at analyzing images to identify patterns, making them ideal for recognizing disease symptoms in crops.[27] CNNs such as ResNet18, DenseNet169, VGG16 and AlexNet can quickly and accurately detect diseases, providing a more reliable alternative to traditional methods, and ultimately helping farmers prevent yield losses.

2.1.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a type of deep learning model specifically designed to process data with a grid-like structure, such as images.[28] CNNs consist of multiple layers of neurons, each applying a different transformation to the input data. One of the critical features of CNNs is their use of convolutional layers, which apply filters to the input data to extract spatial hierarchies of features.[29] Mathematically, the convolutional layer can be expressed as:

$$y\_i=\sigma(\sum(x\_j*w\_ij)+b\_i)$$

In this equation, represents the output activation of the i-th feature map, σ denotes the activation function, represents the input activation of the j-th feature map, represents the corresponding filter weight, and is the bias term.

These filters can be considered pattern detectors looking for specific input data features. For this, the activation function introduces non-linearity to the network, enabling it to learn complex relationships in the data. Common activation functions include Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent (tanh). These activation functions can be represented as:

$$\textit{ReLU:}\ f(x)=\max(0,x)$$

$$\textit{Sigmoid:}\ f(x) = \frac{1}{1+e^{-x}}$$

$$\textit{Tanh:}\ f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$

Additionally, pooling layers are used to down-sample the data by summarising the information in small patches of the input, reducing the data's dimensionality and improving efficiency. Max pooling and average pooling are commonly used operations. Mathematically, they can be represented as:

$$\textit{Max Pooling:}\ y=\max(x)$$

$$\textit{Average Pooling:}\ y=\text{averagex}$$

The overall architecture and equations used in CNNs vary across different models, each designed to address specific challenges in computer vision tasks.[30] This

[26] IBM, 'What Is Machine Learning?'
[27] Burak Gülmez, 'Advancements in Rice Disease Detection through Convolutional Neural Networks: A Comprehensive Review' (2024) 10 Heliyon <https://www.sciencedirect.com/science/article/pii/S2405844024093599>.
[28] IBM, 'What Are Convolutional Neural Networks? | IBM' (www.ibm.com) <https://www.ibm.com/topics/convolutional-neural-networks>.
[29] Adrian Rosebrock, 'Convolutional Neural Networks (CNNs) and Layer Types' (PyImageSearch14 May 2021) <https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/>.
[30] Y Lecun and others, 'Gradient-Based Learning Applied to Document Recognition - IEEE Journals & Magazine' (Ieee.org2019) <https://ieeexplore.ieee.org/document/726791>.

study explored using five prominent CNN models, including Resnet18, Resnet50, Densenet169, Vgg16, and Alexnet.

2.1.1.1 ResNet18

ResNet18 is a convolutional neural network (CNN) model developed by Microsoft Research.[31] It is a variant of the more general ResNet model, known for its ability to train very deep networks (i.e., networks with many layers) without suffering from the problem of vanishing gradients.[32]

The "18" in "ResNet18" refers to the number of layers in the model. It has 18 layers in total, including 16 convolutional layers and 2 fully connected layers. This makes it a relatively shallow network compared to other ResNet models, such as ResNet50 or ResNet101, which have 50 and 101 layers, respectively.[33]

ResNet18 is often used as a general-purpose CNN model for image classification and object detection tasks. It is relatively simple, easy to train, and can achieve good performance on many tasks.

The equations for ResNet models are similar, with higher order ResNets incorporating bottleneck blocks for increased efficiency. The residual block equation (with identity mapping) is given by:

$$y=x+Fx,W$$

Fx,W represents the output of the residual function, which typically consists of convolutional layers with activation functions and parameters {W}. The input x is added element-wise to the output, facilitating the learning of residual mappings.

2.1.1.2 DenseNet169

DenseNet is known for its efficient use of computational resources and ability to learn complex features from data.[34] The "169" in "DenseNet169" refers to the number of layers in the model. It has 169 layers in total, including 168 convolutional layers and 1 fully connected layer. This makes it a very deep network compared to other DenseNet models, such as DenseNet121 or DenseNet201, which have 121 and 201 layers, respectively.[35]

The DenseNet is relatively efficient in using computational resources, making it a good choice for applications with limited computational resources. DenseNet leverages dense connections to enhance information flow and feature reuse across layers. In DenseNet169, dense blocks are connected with concatenation whose equation can be given by:

$$y=[x\_0,F(x\_0,W)]$$

represents the concatenation of the input feature maps and the output of the dense function. The dense function typically consists of convolutional layers with activation functions and parameters {W}.

2.1.1.3 VGG16

VGG16 is a convolutional neural network (CNN) model developed by the Visual Geometry Group at the University of Oxford.[36] It is a variant of the more general VGG model known for its simplicity and effectiveness.

The "16" in "VGG16" refers to the number of layers in the model. It has 16 layers in total, including 13 convolutional layers and 3 fully connected layers. This makes it a relatively shallow network compared to other VGG models, such as VGG19, which has 19 layers. It

[31] Kaiming He and others, 'Deep Residual Learning for Image Recognition' (2015) <https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf>.
[32] Muhammad Shafiq and Zhaoquan Gu, 'Deep Residual Learning for Image Recognition: A Survey' (2022) 12 Applied Sciences 8972 <https://www.mdpi.com/2076-3417/12/18/8972>.
[33] Mingyu Gao and others, 'A Novel Deep Convolutional Neural Network Based on ResNet-18 and Transfer Learning for Detection of Wood Knot Defects' (2021) 2021 Journal of Sensors 1 <https://onlinelibrary.wiley.com/doi/10.1155/2021/4428964>.
[34] Gao Huang and others, 'Densely Connected Convolutional Networks' [2017] 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2261 <https://ieeexplore.ieee.org/document/8099726>.
[35] Rabiul Hasan, Shah Muhammad Azmat Ullah and Sheikh Md Rabiul Islam, 'Recent Advancement of Deep Learning Techniques for Pneumonia Prediction from Chest X-Ray Image' [2024] Medical Reports 100106 <https://www.sciencedirect.com/science/article/pii/S2949918624000718> accessed 19 August 2024.
[36] Karen Simonyan and Andrew Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition' (arXiv.org10 April 2015) <https://arxiv.org/abs/1409.1556>.

uses 3x3 convolutional filters and 2x2 max pooling throughout the network, which can be represented as:

$$y=\sigma\sum x*w+b$$

In addition to the activation function σ, padding can be applied to ensure the output feature map has the same spatial dimensions as the input.

2.1.1.4 AlexNet

AlexNet is a convolutional neural network (CNN) model that was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton.[37] It is a significant milestone in deep learning and has played a crucial role in popularising CNNs for image classification tasks.

The AlexNet architecture consists of eight layers, five of which are convolutional and three fully connected. It utilizes the rectified linear unit (ReLU) activation function and employs dropout regularisation to prevent overfitting. One notable feature of AlexNet is the use of GPU acceleration, which greatly improved training speed and made it possible to train deeper networks efficiently.

Local Response Normalisation (LRN) equation:

$$b^{i}_{x,y}= \frac{a^{i}_{x,y}=}{(k+a^* \Sigma\, (a^{i}_{x',y'})^{2}\,)\ \beta}$$

represents the output of the LRN layer for neuron i at position (x, y). denotes the input activation of neuron i at position (x, y). LRN normalizes the neuron's response based on the sum of squared activations within a local region, controlled by parameters k, α, and β.

2.2 Dataset Description

To develop the initial potato disease detection model, an initial dataset containing images collected from the web was used to test and train the model. The initial dataset is crucial in understanding how the trained model behaves and knowing the types of images that would be encountered by farmers using the model in the field. The success of an image classification project is dependent on the data used to train the deep learning algorithms. The steps for generating the initial dataset are detailed below:

2.2.1 Data Collection Method

The most important tool for the data collection exercise is Google Images. This can be conducted through the steps highlighted below:

i. Enter a search query/term on Google Images and let the page load. This should ideally have an infinite scroll, where content keeps loading as you scroll down.

ii. The second tool in use is JavaScript. As the page loads with the different images, keep scrolling down to allow more images to load.

iii. Once an adequate number of images have loaded, a JavaScript script captures the URLs of all the images loaded on Google Images.

iv. In this instance, JavaScript is executed in REPL-like manner.

v. The JavaScript script is used to simulate a right-click on an image and extraction of the URL.

vi. All the captured URLs are then stored in a text file.

After extracting an adequate number of URLs, the next step is downloading the images. For this, we will be using Python:

i. First, ensure you have access to the URLs file containing the potato images' URLs.

ii. The main python package used for this exercise is requests. Requests allows you to send HTTP/1.1 requests extremely easily.

[37] Alex Krizhevsky, Ilya Sutskever and Geoffrey E Hinton, 'ImageNet Classification with Deep Convolutional Neural Networks' (2012) 60 Communications of the ACM 84 <https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.

iii. We would then use command-line parsing to point the script to the URL file and, later, to the output folder where all the downloaded images will be stored.

iv. As a rule of thumb, data verification must be done. For this we use OpenCV, where we try and see if all the images can be opened, if they cannot be opened, they are deleted.

v. After checking for valid images, the next step is to check for relevant images. This is a manual process of sifting through the images to remove irrelevant or wrongly placed images.

2.2.2 Data Description

The data collected resulted in 1,198 images distributed into 10 classes, illustrated as follows:

Table 2.1 Data Distribution

Class	Composition	Class	Composition
Late Blight	18%	Fusarium Dryroot	9%
Common Scab	14%	Silver Scurf	9%
Pink Rot	13%	BlackDot	7%
Healthy Potatoes	10%	Early Blight	6%
Black Scurf	9%	Potato Virus Y	5%

Each class label is a disease category, and we try to predict the disease category given just the image of the potato tuber. The images below show samples photographs of the dataset:

Figure 2.2 Sample photographs of the dataset

2.2.3 Data Pre-processing and Preparation

Using random sampling, the entire dataset was initially divided into two samples: the training set and the testing set. In preparation for the data modelling process, we re-sampled our dataset into an 80-10-10 split for the train-validation-test composition to incorporate the validation set.

In addition, in all the approaches described in this study, we resized the images to 224 × 224 pixels and performed both the model optimization and predictions on these downscaled images.

All the collected images were then further subdivided into patches covering the entire tuber to capture the diseased area of the tuber completely. The patches appear as shown in Figure 2.3 (Appendix A). The data summary appears in Figure 2.4 (Appendix A) when segmentation is done on the entire dataset.

The dataset initially used had 1,198 images, which were later segmented to create a dataset with 16,283 images/patches.

Table 2.2 Data pre- and post-segmentation

Initial Dataset		After Segmentation	
Late Blight	211	Common Scab	2,624
Common Scab	164	Pink Rot	2,448
Pink Rot	153	Healthy Potatoes	1,824
Healthy Potatoes	114	Black Scurf	1,808
Black Scurf	113	Fusarium Dryroot	1,792
Fusarium Dryroot	112	Silver Scurf	1,792
Silver Scurf	112	BlackDot	1,408
BlackDot	88	Early Blight	1,184
Early Blight	74	Potato Virus Y	912
Potato Virus Y	57	Late Blight	491

2.2.4 Data Modelling Before Patching

We applied these five DNN models: Resnet18, Resnet50, Densnet169, VGG16, and Alexnet. The experiment aimed to assess the effectiveness of a convolutional neural network (CNN) in classifying nine different potato diseases and healthy potatoes. The goal was to evaluate the performance of the 5 CNN's learning algorithm in this task.

The results were as shown below; the best performance was from the Resnet50 model with a 43% error rate, which had an above 70% accuracy in predicting Healthy potatoes, Late Blight, Silver Scurf and Black Scurf. The most misclassified class was Black Dot, where most images were wrongly classified as Silver Scurf. The model performance was attributed to the low number of images, considering that CNN requires adequate data to train a model.

Table 2.3 Model performance

Models	Accuracy
Resnet50	56.5%
Vgg16	53.2%
Densenet169	52.4%
Resnet18	50.0%
Alexnet	49.1%

Even though this was the best performance among the base models, it still fell short of the desired threshold for effective classification. Hence, the researchers implemented a data pre-processing technique by splitting each image into patches that focused on the diseased areas of the tubers, resulting in a 13 times increase in the dataset size.

2.2.5 Data Modelling After Patching

Due to the poor performance of the base model, data pre-processing was needed to understand how it might affect the experiment. In this study, we applied these 5 DNN models: Resnet18, Resnet50, Densnet169, VGG16, and Alexnet. These models and their training and testing processes were implemented using Fast AI. The data used is a pre-processed version of the original dataset where every single image is split down into different patches, each focusing on the diseased part of the tuber and then resized to 224x224. This resulted in a 13x increase in the dataset.

All the different models presented were trained using the same training and tuning parameters. The overall model performance is highlighted in Appendix A.

Patching allowed for a more comprehensive representation of the disease characteristics, which is crucial for effectively training deep learning models. After retraining the models with the enhanced dataset, significant improvements in accuracy were observed. Densenet169 emerged as the top performer, achieving an accuracy of 89.7% and a low error rate of 10.3%.

During training, we compared the mean precision, mean recall, and mean F1 score over the whole period of training at regular intervals (at the end of every epoch). The performance of each model is outlined.

Table 2.4 Model accuracy after patching

Models	Accuracy
Densenet169	89.7%
Resnet50	86.1%
Vgg16	77.9%
Resnet18	73.4%
Alexnet	62.6%

Its confusion matrix based on the out-of-sample test set is shown below.

Actual \ Predicted	Early Blight	Healthy_Potatoes	Late_blight	Pink_Rot	Potato_virus_Y	Silver_Scurf	black_scurf	blackdot	ommon_scab	fusarium Dryroot
Early Blight	0.81	0.00	0.00	0.04	0.02	0.03	0.00	0.04	0.04	0.03
Healthy_Potatoes	0.00	0.99	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00
Late_blight	0.00	0.02	0.90	0.01	0.00	0.00	0.02	0.00	0.00	0.06
Pink_Rot	0.01	0.00	0.00	0.90	0.01	0.00	0.00	0.00	0.00	0.06
Potato_virus_Y	0.00	0.00	0.00	0.01	0.97	0.01	0.00	0.00	0.00	0.01
Silver_Scurf	0.01	0.01	0.00	0.01	0.00	0.93	0.01	0.02	0.02	0.00
black_scurf	0.01	0.00	0.00	0.02	0.01	0.03	0.87	0.02	0.02	0.03
blackdot	0.02	0.00	0.00	0.00	0.00	0.11	0.05	0.79	0.03	0.00
common_scab	0.02	0.00	0.00	0.00	0.01	0.00	0.03	0.02	0.89	0.03
fusarium Dryroot	0.01	0.01	0.00	0.05	0.01	0.00	0.01	0.01	0.00	0.92

Figure 2.5 Desnet169 confusion matrix

2.3 Summary of Findings

Initially, the study evaluated five models, including Resnet18, Resnet50, Densnet169, VGG16, and Alexnet. Initially, these models demonstrated varying levels of accuracy, with Resnet50 achieving the highest accuracy at 56.5%. Despite this being the best performance among the base models, it still fell short of the desired threshold for effective classification. The Resnet50 model showed particular strength in identifying Healthy potatoes, Late Blight, Silver Scurf, and Black Scurf, although it also had a 43% error rate.

Recognizing the limitations posed by the low number of training images, the researchers implemented a data pre-processing technique. They split each image into patches that focused on the diseased areas of the tubers, resulting in a 13 times increase in the dataset size. This strategy allowed for a more comprehensive representation of the disease characteristics, which is crucial for effectively training deep learning models. After retraining the models with the enhanced dataset, significant improvements in accuracy were observed. Densenet169 emerged as the top performer, achieving an accuracy of 89.7% and a low error rate of 10.3%.

This evaluation of models indicates the necessity for large amounts of data when training deep learning models. The application of data pre-processing techniques, such as patching, significantly enhanced model accuracy. Furthermore, the high accuracy of these models presents an advantage over traditional methods of disease detection in potato plants in Kenya, which often rely on slow visual inspections and expert knowledge, which is less reliable. Hence, using Machine Learning techniques to develop disease detection applications enables more timely and precise interventions to reduce crop losses.

3. The Use of AI in Potato Disease Detection in Kenya

3.1 Potato Disease Detection in Kenya

Monitoring potato diseases in Kenya is becoming more important because various pathogens threaten potato production.[38] As one of the country's most widely cultivated crops, potatoes are crucial in the agricultural economy, providing food security and livelihoods for millions of people.[39] However, the increasing incidence of diseases such as late blight and blackdot poses significant challenges to potato farmers.[40] These diseases reduce crop yields and compromise the quality of the produce, leading to economic losses and food scarcity.

Unfortunately, traditional methods of disease identification, which often rely on visual inspection and laboratory testing, can be time-consuming and may not provide timely information for farmers to take necessary actions.[41] Ethiopian researchers identified the similarity of potato crop diseases as a major challenge of crop disease detection through naked-eye observation.[42] As such, manually detecting and classifying such diseases is error-prone and tedious.[43] In response to the growing threat of pathogens, various initiatives have focused on comprehensive surveys to better understand pathogen prevalence and distribution across key potato-producing regions in Kenya.[44] In a survey by the Centre for Agriculture and Bioscience International (CABI) and the Kenya Plant Health Inspectorate Service (KEPHIS), researchers conducted extensive sampling from nearly 3,000 potato tubers, plant materials, and soil across six counties across Kenya.[45] The results revealed that one in ten samples tested positive for a potato disease known as blackleg, underscoring the widespread nature of potato diseases and highlighting the critical need for improved disease monitoring systems.[46]

As such, there is an urgent need to adopt innovative approaches that leverage technological advancements, such as artificial intelligence, to enhance the accuracy and efficiency of disease detection.[47] Integrating modern technologies in detection has enabled the monitoring of these and other plant diseases by generating timely and context-specific information, which enables farmers to improve their detection capacity and respond to potato diseases, safeguarding Kenya's potato production levels and food security for its population.

3.2 The Use of AI in Potato Disease Detection in Kenya

The agricultural sector in Kenya is witnessing a shift with the integration of artificial AI, particularly in disease detection for staple crops like potatoes.[48]

[38] CABI, 'Surveillance of Potato Diseases in Kenya'
[39] CABI, 'Surveillance of Potato Diseases in Kenya'
[40] CABI, 'Surveillance of Potato Diseases in Kenya'
[41] CABI, 'Surveillance of Potato Diseases in Kenya'
[42] Natnael Tilahun and Beakal Gizachew, 'Artificial Intelligence Assisted Early Blight and Late Blight Potato Disease Detection Using Convolutional Neural Networks|Ethiopian Journal of Crop Science' (Ajol.info2020) <https://www.ajol.info/index.php/ejcs/article/view/237248>.
[43] Dor Oppenheim and others, 'Using Deep Learning for Image-Based Potato Tuber Disease Detection' (2019) 109 Phytopathology 1083 <https://pubmed.ncbi.nlm.nih.gov/30543489/>.
[44] Seed Potatoes Kenya, 'Potato Bacterial Disease Surveillance in Kenya -Explanatory Note #1' (2019) <https://www.wur.nl/en/show/Explanatory-note-1-Potato-Diseases-Surveillance-in-Kenya.htm> accessed 17 August 2024.
[45] Seed Potatoes Kenya, 'Potato Bacterial Disease Surveillance in Kenya -Explanatory Note
[46] Seed Potatoes Kenya, 'Potato Bacterial Disease Surveillance in Kenya -Explanatory Note
[47] CABI, 'Surveillance of Potato Diseases in Kenya'
[48] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers' (Consultative Group on International Agricultural Research: Research Program on Roots, Tubers and Bananas11 August 2021) <https://www.rtb.cgiar.org/news/ai-for-diagnosing-potato-diseases-will-benefit-millions-of-rural-farmers/#:~:text=New%20advances%20in%20artificial%20intelligence> accessed 17 August 2024.

Potatoes are susceptible to various diseases, such as early and late blight, which can lead to devastating losses if not identified and managed promptly.[49] The advent of AI-powered applications, such as the Plant Village Nuru app in Kenya, has presented promising benefits for Kenyan potato farmers in disease detection.[50] By leveraging machine learning algorithms and extensive databases of plant images, these applications enable users to diagnose diseases in real-time using their smartphones, empowering them with immediate access to critical information and faster response to potato plant diseases.[51] Furthermore, the Nuru app is designed to be user-friendly and accessible, providing support in multiple languages, including Swahili, French, and English, which broadens its reach among diverse farming communities.[52] The app fosters a collaborative environment where knowledge is shared, and best practices are disseminated by connecting farmers directly to expert advice from government agencies, universities, and research organizations.[53] This integration of technology contributes to increased agricultural productivity. It sets the background for a skilled class of farmers and farm workers across Kenya, where over two million farmers depend on potato cultivation for their livelihoods.[54]

Using AI in potato disease detection also addresses the challenges posed by limited access to expert knowledge and the need for timely interventions by allowing in-field diagnosis and direct links to expert advice through the application.[55] As Kenya continues to embrace these technological advancements, the potential for increased productivity, improved food security, and enhanced livelihoods for rural farmers increases.[56] However, some barriers to the adoption of these AI tools remain, which are discussed in the next section.

The advent of AI-powered applications, such as the Plant Village Nuru app in Kenya, has presented promising benefits for Kenyan potato farmers in disease detection.

[49] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers'
[50] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers'
[51] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers'
[52] NPCK, 'Artificial Intelligence for Diagnosing Potato Diseases – Npck' (Npck.org 2022) <https://npck.org/artificial-inteligence-for-diagnosing-potato-diseases/> accessed 15 August 2024.
[53] NPCK, 'Artificial Intelligence for Diagnosing Potato Diseases – Npck'
[54] NPCK, 'Artificial Intelligence for Diagnosing Potato Diseases – Npck'
[55] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers'
[56] CGIAR, 'AI for Diagnosing Potato Diseases Will Benefit Millions of Rural Farmers'

4. Barriers to AI Use in Potato Disease Detection in Kenya

Integrating AI in potato disease detection in Kenya presents a great opportunity to improve agricultural productivity and food security. However, several barriers hinder its effective implementation, including limited access to technology, insufficient data to train applications, low technical knowledge among farmers, and the potential for disregarding indigenous knowledge.

4.1 Limited Access to Technology

Lack of access to technology is a critical barrier to adopting AI in potato disease detection. Many smallholder farmers in Kenya, representing a significant portion of the agricultural workforce, often lack the necessary technological tools, such as smartphones, tablets, and computers.[57] This digital divide means that even if AI tools are available, a significant portion of the farming community cannot benefit from them since they cannot scan crops, access relevant information, or communicate with experts in real time.[58] Moreover, rural infrastructure in Kenya can be inadequate, with many areas lacking reliable internet connectivity. This hinders the use of cloud-based AI solutions that require constant internet access for data processing and analysis.[59] Without the necessary technological infrastructure, farmers cannot fully leverage AI tools that could enhance their ability to detect potato diseases. Addressing this barrier requires investments in rural technology infrastructure and subsidizing the costs of devices such as smartphones to increase access.

[57]Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective' (2023) 2023 Advances in Agriculture 1 <https://onlinelibrary.wiley.com/doi/epdf/10.1155/2023/1530629>.
[58]Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective'
[59]Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective'

4.2 Insufficient Training Data

AI systems rely heavily on data; they require large volumes of high-quality data for training and validation to function effectively.[60] Algorithms created using limited datasets typically do not perform well in field contexts. This is due to the vast differences in the images taken in the field, as well as the different focus and resolution results that occur in real-world agricultural settings.[61] The inadequacy of large quantities of high-quality training data is a significant challenge in Kenya's adoption of AI technologies.[62] The Kenyan government does not currently provide a public repository for crop and livestock disease data, which limits the availability of high-quality datasets necessary for training AI models.[63] There is also a lack of comprehensive datasets with historical information on diseases, environmental conditions, and agricultural practices.[64] This scarcity hinders the development of robust AI applications for potato disease detection.

4.3 Low Technical Knowledge among Farmers

The successful adoption of AI technologies in agriculture relies upon the ability of farmers and farm workers to understand and effectively use these technologies.[65]

[60]Joe Johnson and others, 'Enhanced Field-Based Detection of Potato Blight in Complex Backgrounds Using Deep Learning' (2021) 2021 Plant Phenomics 1 <https://spj.science.org/doi/10.34133/2021/9835724>.
[61]Joe Johnson and others, 'Enhanced Field-Based Detection of Potato Blight in Complex Backgrounds Using Deep Learning'
[62]Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective'
[63]Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective'
[64]Joe Johnson and others, 'Enhanced Field-Based Detection of Potato Blight in Complex Backgrounds Using Deep Learning'
[65]Mohd Javaid and others, 'Understanding the Potential Applications of Artificial Intelligence in Agriculture Sector' (2022) 2 Advanced Agrochem 15 <https://www.sciencedirect.com/science/article/pii/S277323712200020X>.

In Kenya, there is a general lack of knowledge regarding the potential applications of AI in potato disease detection, and many farmers may not be familiar with how AI can enhance their farming practices in rural communities. While organizations like PlantVillage train select farmers to become lead farmers who can then educate their peers, the broader farming community remains largely unaware.[66] Not all farmers receive adequate training on AI tools or interpreting the data these applications provide.[67] This lack of understanding can lead to underutilizing the technology, as farmers may be unsure how to apply the insights gained from AI tools to their farming practices.

4.4 Potential Disregard for Indigenous Knowledge

Indigenous knowledge encompasses the practices and information that have been developed over generations within local communities. These practices and information are often tailored to specific environmental conditions and cultural contexts.[68] Research indicates that a major challenge of adopting computer vision technology among Kenyan farmers is the potential for this technology to disregard the indigenous farming practices that the farmers have gained from generations of practice.[69] When adopted, data-driven approaches often take precedence over contextual practices such as agricultural Indigenous knowledge, potentially undermining the effectiveness of past successful agricultural practices.[70] Also, farmers who have relied on indigenous knowledge may resist adopting new technologies that they perceive as foreign or incompatible with their traditional practices, which creates a divide between modern technological solutions and established farming methods, making it challenging to achieve widespread acceptance of technologies in farming. Thus, the adoption of AI in potato plant disease detection must integrate indigenous knowledge through collaborative approaches that respect and incorporate traditional knowledge in designing data-driven products for agricultural applications.

In Kenya, there is a general lack of knowledge regarding the potential applications of AI in potato disease detection, and many farmers may not be familiar with how AI can enhance their farming practices in rural communities.

[66] Juliet Akoth Ojwang, 'Kenyan Farmers Turn to WhatsApp & AI Tools to Combat Crop Diseases'
[67] Juliet Akoth Ojwang, 'Kenyan Farmers Turn to WhatsApp & AI Tools to Combat Crop Diseases'
[68] Astone Owino, 'Challenges of Computer Vision Adoption in the Kenyan Agricultural Sector and How to Solve Them: A General Perspective'
[69] Joseph N Kiplang'at and Daniel Chebutuk Rotich, 'Mapping and Auditing of Agricultural Indigenous Knowledge in Uasin Gishu and Keiyo Districts in Rift Valley Province, Kenya'

5. Key Regulatory Considerations in Using AI for Potato Disease Detection

Regulatory and policy concerns are a significant barrier to adopting AI in potato disease detection in Kenya. The regulatory environment may not fully support integrating AI technologies in agriculture, leading to uncertainty among stakeholders. Additionally, there may be a lack of clear policies and guidelines governing the use of AI, which can create challenges for both developers and users of AI solutions. Existing agricultural policies may not adequately address the unique challenges posed by AI technologies, such as data privacy, intellectual property rights, and the ethical use of AI in decision-making.[71]

To address these concerns, comprehensive policies that promote the responsible use of AI in agriculture, provide clear guidelines for stakeholders, and foster an environment conducive to innovation are essential.

Key areas of focus include data privacy and security, transparency and accountability, and the existing gap in AI regulation in Kenya. Addressing these concerns is important to ensure the effective integration of AI into agriculture while fostering a strong framework to enable innovation and safeguard the potato farming sector.

5.1 *Data Privacy and Security*

The first major consideration is data privacy and security when dealing with farmers' data. The principle of privacy underlies all data protection activities, and in Kenya, it is anchored in Article 31 of the Constitution. Article 31(c) clearly states that individuals have the right not to have their information unnecessarily revealed or required. Kenyan farmers using the Machine Learning model must ensure that their personal data is not unnecessarily required or revealed. This is in line with the principle of purpose limitation espoused by the Data Protection Act,[72] where data collection and processing is limited to only what is necessary for the processing in question.[73]

In this regard, farmers' personal data should not be collected or processed when deploying AI tools since it does not directly contribute to the effectiveness of the machine learning model in detecting potato diseases. Further, the act directs that personal data should only be kept as long as necessary for its intended purpose unless there are specific exceptions. Once the data is no longer needed, the data controller or processor is required to properly dispose of it by deleting, erasing, anonymizing, or pseudonymization.[74]

Additionally, data processors must implement robust security measures for storing and managing collected data to further safeguard against unauthorized access and potential breaches.[75]

This is especially important where farmers' personally identifiable information is stored in databases and datasets, which are then shared with AI tool developers.

5.2 *Transparency and Accountability*

Deployers of agricultural AI tools must communicate how the machine learning model operates by providing detailed explanations of its data processing to help farmers understand their functionality.[76]

[71]Rebecca Sarku, Ulfia A Clemen and Thomas Clemen, 'The Application of Artificial Intelligence Models for Food Security: A Review' (2023) 13 Agriculture 2037 <https://www.mdpi.com/2077-0472/13/10/2037>.

[72]Data Protection Act, Section 25 (c)
[73]Data Protection Act, Section 25 (d)
[74]Data Protection Act, Section 39
[75]Data Protection Act, Section 41
[76]Michael, 'Data Bias in AI Agriculture: Ensuring Fairness&Sustainability' (June 2024) <https://keymakr.com/blog/data-bias-in-ai-agriculture-ensuring-fairness-and-sustainability/#:~:text=Transparency%20and%20Accountability%20in%20AI%20Agriculture&text=This%20means%20clearly%20sharing%20info> accessed 21 August 2024.

Developers and operators of these tools must also take responsibility for their accuracy and performance by regularly conducting tests to improve reliability and precision in disease detection. In cases where the model produces incorrect results, clear protocols should address and correct these issues to ensure farmers receive the best quality of detection. This should also be accompanied by feedback mechanisms for farmers to report issues, allowing them to participate in improving the tools.

5.3 AI Regulation Gap

Currently, Kenya lacks a comprehensive regulatory framework for AI, raising concerns as AI tools become increasingly integrated into various sectors. Agriculture utilizes AI the most, accounting for nearly half of the total use cases identified across sectors in Kenya.

The absence of this regulatory framework creates uncertainties regarding the safety, efficacy, and ethical use of AI tools, as there are no established standards for algorithm accuracy and reliability. If deployed AI tools fail to perform as intended, resulting in incorrect disease diagnoses and adverse outcomes for farmers, no legal remedies are available. Therefore, as AI continues to be implemented in agriculture, developing a regulatory framework that specifically addresses its deployment across various sectors is essential. This framework should also consider smallholder farmers' unique challenges, who may lack access to precision agriculture technologies and encounter barriers related to data availability and digital literacy.

Developers and operators of these tools must also take responsibility for their accuracy and performance by regularly conducting tests to improve reliability and precision in disease detection.

6. Recommendations

6.1 Expand Training and Capacity-Building for Farmers

Targeted education efforts are necessary to increase the potential for widespread adoption of AI in agriculture.[79] These initiatives should focus on practical applications such as disease detection, data interpretation, and decision-making processes on disease. These trainings should also target female farmers, who often face tougher barriers to technology access due to gender disparities. The government should also invest in education, research, and development to cultivate skills in AI and other relevant technologies, emphasize practical applications in agriculture and empower farmers to leverage these tools effectively.

6.2 Facilitate Access to Technology and Infrastructure among Rural Communities

The government should take proactive measures to improve the affordability and accessibility of smartphones, tablets, and other digital devices for smallholder farmers. Additionally, it should enhance internet infrastructure in rural areas to ensure farmers can effectively utilize AI tools and technologies, thereby facilitating their engagement with digital agricultural solutions.[80]

6.3 Quality and Adequate Training Data

The government should improve the accessibility and quality of training data for AI applications in agriculture. Despite notable advancements in data collection in Kenya, issues persist regarding the completeness and quality of available datasets. The Kenyan government has made progress in developing open-source databases, such as the Census from the Kenya National Bureau of Statistics (KNBS) and various sector-specific datasets from the Kenya Open Data Initiative.[81] However, these datasets are more accessible to larger technology firms, which can disadvantage smallholder farmers and startups.[82]

Furthermore, gathering images and data from farmers will improve model accuracy and yield valuable insights for further improvements. A significant challenge in developing effective models has been obtaining curated data in controlled settings, leading to a reliance on Google Image data that necessitates extensive cleaning and transformation.[83]

Improving data availability and quality will empower startups to develop efficient AI tools that enable farmers to better identify diseases in their potato tubers and effectively manage them. This will support farmers in increasing their yields and improving food security in the country.

6.4 Integrate Indigenous Knowledge in Agricultural Technology Design

To effectively adopt AI for potato plant disease detection among Kenyan farmers, indigenous knowledge must be integrated through collaborative approaches that respect and incorporate traditional practices.[84] Engaging local farmers and knowledge holders in the design process ensures that technological solutions

[79] Juliet Akoth Ojwang, 'Kenyan Farmers Turn to WhatsApp & AI Tools to Combat Crop Diseases'
[80] ICJ, 'Leveraging Technology to Advance Access to Information in Kenya - ICJ Kenya' (29 September 2023) <https://icj-kenya.org/news/opinion-leveraging-technology-to-advance-access-to-information-in-kenya/>.

[81] Eugénie Humeau, 'AI for Africa: Use Cases Delivering Impact Kenya Deep Dive Author and Contributors'
[82] Eugénie Humeau, 'AI for Africa: Use Cases Delivering Impact Kenya Deep Dive Author and Contributors'
[83] Eugénie Humeau, 'AI for Africa: Use Cases Delivering Impact Kenya Deep Dive Author and Contributors'
[84] Joseph N Kiplang'at and Daniel Chebutuk Rotich, 'Mapping and Auditing of Agricultural Indigenous Knowledge in Uasin Gishu and Keiyo Districts in Rift Valley Province, Kenya'

are tailored to specific environmental and cultural contexts, enhancing their relevance and acceptance.[85] The adoption of AI in potato disease detection would be effective and improve farmers' livelihoods by fostering a partnership between modern data-driven methods and established Indigenous practices.

6.5 AI Regulation

Kenya should establish a framework with clear standards for AI accuracy and ethical use, alongside legal protections for farmers against tool failures. Engaging stakeholders, especially smallholder farmers, in the regulatory process, is essential, as is enhancing their digital literacy through capacity-building initiatives. The draft Information Technology - Artificial Intelligence - Code of Practice for AI Applications published for public comment by the Kenya Bureau of Standards (KEBS) incorporates key standards for AI trustworthiness, which includes features such as AI robustness, reliability, explainability, and risk management framework.[86]

The risk assessment framework emphasizes identifying and quantifying AI risks and establishing controls relevant to the development and use of AI systems. It also advocates for continuous monitoring and review of risk management processes, ensuring that any emerging risks affecting users are promptly addressed and that users are actively involved in evaluating the trustworthiness of AI systems. When adopted, the standards set a strong background for compliance with responsible AI principles, ensuring users such as farmers are protected from potential AI risks.

[85] Joseph N Kiplang'at and Daniel Chebutuk Rotich, 'Mapping and Auditing of Agricultural Indigenous Knowledge in Uasin Gishu and Keiyo Districts in Rift Valley Province, Kenya'

[86] KEBS, 'Information Technology -Artificial Intelligence -Code of Practice for AI Applications' (2024) <https://www.dataguidance.com/sites/default/files/kebs-tc_094_n66_public_review_kenya_standard_dks_3007_ai_code_of_practice.pdf>.

7. Conclusion

The integration of AI for potato disease detection in Kenya offers a significant opportunity to enhance agricultural yields and food security. By using AI tools, particularly those based on CNNs for image classification, farmers can achieve early detection and effective management of potato diseases, preventing crop losses. However, adopting these technologies faces several challenges, including limited access to technology, insufficient training data, low technical knowledge among farmers, and the potential disregard for indigenous knowledge.

To overcome these barriers, it is essential to expand training and capacity-building initiatives for farmers, facilitate access to technology and infrastructure in rural communities, and ensure the availability of quality training data. Additionally, integrating indigenous knowledge into the design of agricultural technologies will enhance their relevance and effectiveness. Establishing and implementing regulations on data protection and AI would also address concerns about privacy, security, and accountability. By addressing these challenges and implementing these recommendations, the potential of AI in transforming potato disease detection can be fully realized, leading to improved agricultural productivity and food security in Kenya.


APPENDIX

Appendix A: Figures & Tables

Figure 2.3: Potato patches

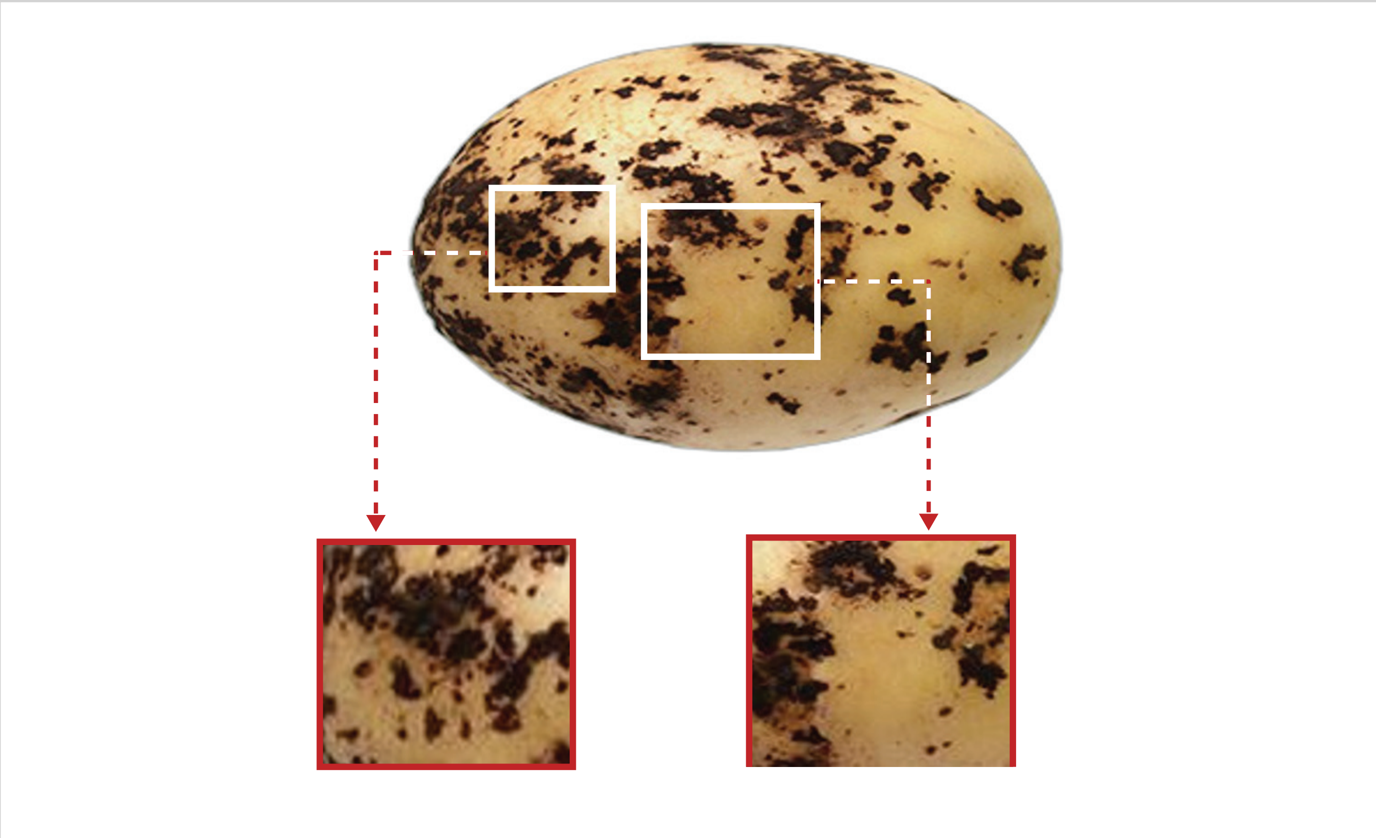


Figure 2.4 Data summary after segmentation

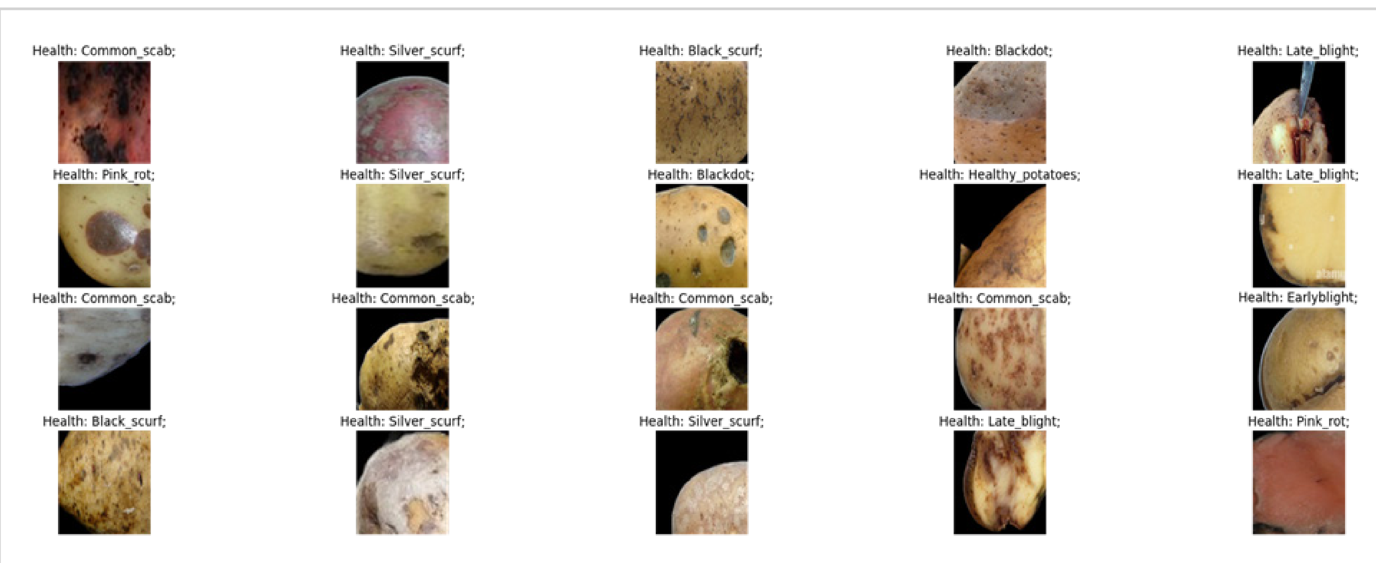


Figure 2.6 Confusion Matrix

Actual	Early Blight	Healthy_Potatoes	Late_blight	Pink_Rot	Potato_virus_Y	Silver_Scurf	black_scurf	blackdot	ommon_scab	fusarium Dryroot
Early Blight	0.20	0.00	0.10	0.10	0.00	0.00	0.20	0.20	0.10	0.10
Healthy_Potatoes	0.00	0.73	0.09	0.18	0.00	0.00	0.00	0.00	0.00	0.00
Late_blight	0.00	0.05	0.71	0.10	0.00	0.05	0.00	0.05	0.00	0.05
Pink_Rot	0.07	0.00	0.27	0.53	0.00	0.00	0.00	0.07	0.00	0.07
Potato_virus_Y	0.00	0.20	0.00	0.00	0.20	0.20	0.00	0.00	0.20	0.20
Silver_Scurf	0.00	0.00	0.00	0.09	0.09	0.82	0.00	0.00	0.00	0.00
black_scurf	0.08	0.00	0.00	0.00	0.00	0.00	0.77	0.08	0.08	0.00
blackdot	0.12	0.00	0.00	0.00	0.00	0.38	0.25	0.12	0.12	0.00
common_scab	0.19	0.06	0.06	0.06	0.06	0.00	0.00	0.06	0.44	0.06
fusarium Dryroot	0.00	0.08	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.57

Predicted

Resnet18

Table 2.5 Resnet18 performance

Class	Precision	Recall	F1-Score
Early Blight	0.76	0.47	0.58
Healthy Potatoes	0.88	0.88	0.88
Late blight	0.71	0.49	0.58
Pink Rot	0.75	0.78	0.76
Potato virus	0.73	0.66	0.69
Black scurf	0.76	0.81	0.78
Black dot	0.63	0.46	0.53
Common scab	0.73	0.84	0.78
Fusarium Dryroot	0.74	0.75	0.75

Resnet50

Table 2.6 Resnet50 performance

Class	Precision	Recall	F1-Score
Early Blight	0.81	0.78	0.8
Healthy Potatoes	0.95	0.96	0.95
Late blight	0.92	0.73	0.82
Pink Rot	0.85	0.90	0.87
Potato virus Y	0.94	0.87	0.90
Black scurf	0.84	0.89	0.86
Black dot	0.79	0.72	0.75
Common scab	0.87	0.88	0.87
Fusarium Dryroot	0.85	0.82	0.84

Densenet169

Table 2.7 Desnet169 performance

Class	Precision	Recall	F1-Score
Early Blight	0.88	0.81	0.84
Healthy Potatoes	0.98	0.99	0.98
Late blight	0.96	0.90	0.93
Pink Rot	0.90	0.90	0.90
Potato virus	0.90	0.97	0.93
Black scurf	0.89	0.87	0.88
Black dot	0.84	0.79	0.81
Common scab	0.93	0.89	0.91
Fusarium Dryroot	0.84	0.92	0.88

Vgg16

Table 2.8 Vegg16 performance

Class	Precision	Recall	F1-Score
Early Blight	0.73	0.58	0.65
Healthy Potatoes	0.86	0.88	0.87
Late blight	0.83	0.51	0.63
Pink Rot	0.79	0.83	0.81
Potato virus	0.82	0.74	0.77
Black scurf	0.84	0.81	0.82
Black dot	0.66	0.63	0.64
Common scab	0.82	0.83	0.82
Fusarium Dryroot	0.75	0.82	0.78

Alexnet

Table 2.9 Alexnet performance

Class	Precision	Recall	F1-Score
Early Blight	0.60	0.25	0.35
Healthy Potatoes	0.73	0.81	0.77
Late blight	0.48	0.33	0.39
Pink Rot	0.69	0.72	0.71
Potato virus	0.64	0.58	0.61
Black scurf	0.72	0.67	0.69
Black dot	0.41	0.37	0.39
Common scab	0.60	0.69	0.64
Fusarium Dryroot	0.65	0.63	0.64

Model Deployment

The model was deployed using Flask, a web application framework written in Python. It offers an easy-to-setup platform, and in this case, easy does not mean lacking in functionality since Flask is capable of scaling to complex systems.

In the projects, Flask was used to launch a web interface allowing users to upload their photographs and run the image through the model to get an output. The test frontend was developed using HTML, JS and CSS, as seen below:

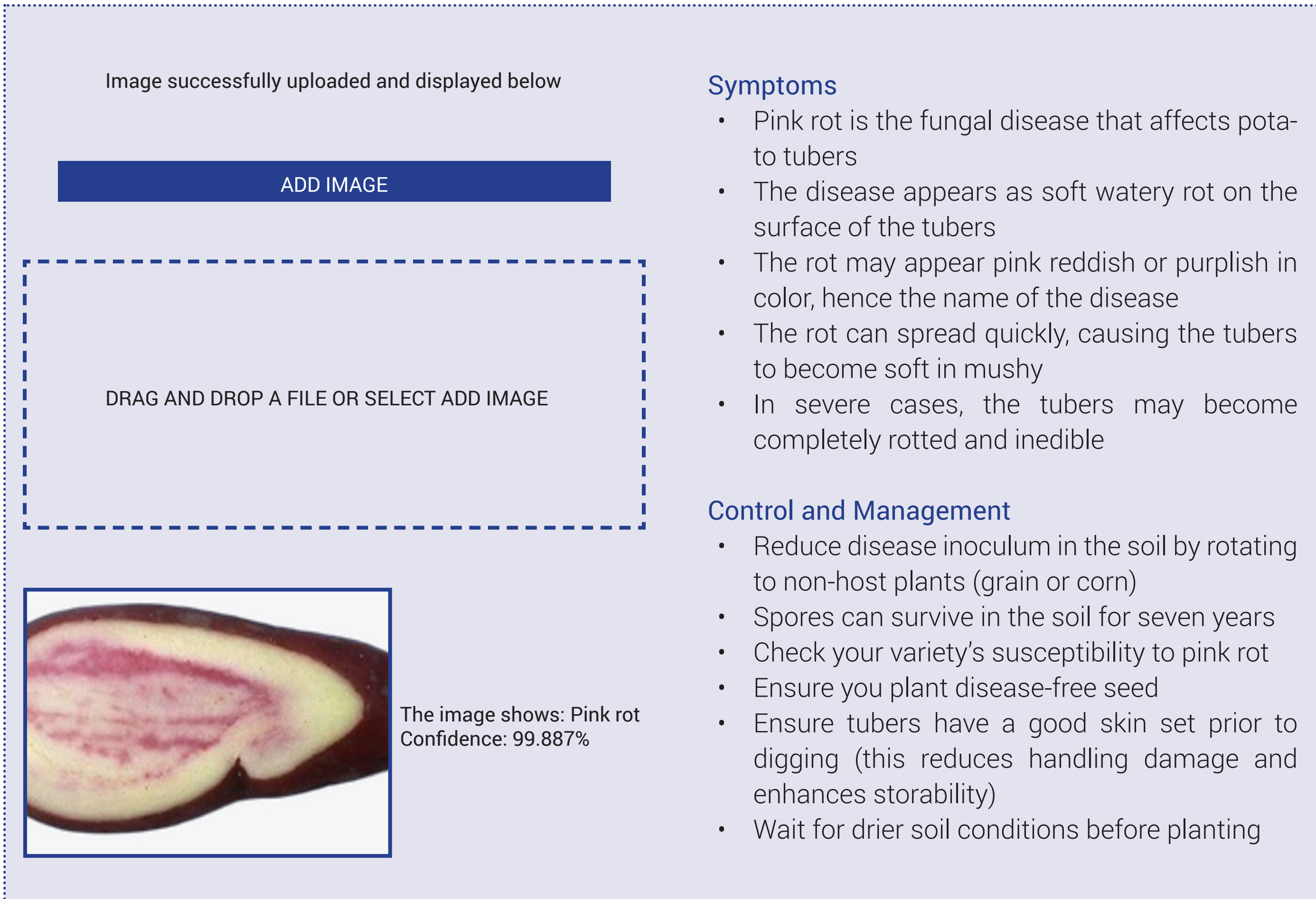


Figure 2.7 Image classification example

Flask was used for API calls between the client and the model (backend). The deployment was only possible because the deep learning model was loaded into a pickle file. The pickle file stores a serialized version of the model, and when we later want to use it, we deserialize it. It essentially stores an instance of the model without the data, and later, we run it with the data (an image).

This study was made possible by a grant provided by the International Development Research Center (IDRC). We thank the organization for their continued support.