MAIZE CROP DISEASE IMAGE CLASSIFICATION

1.0 Business Understanding

1.1 Overview

Maize is a cereal plant that yields large grains (corn or sweetcorn) set in rows on a cob. The leafy stalk of the plant produces pollen inflorescences and separate ovuliferous inflorescences called ears that yield kernels or seeds, which are fruits. Maize matures after 3-4 months especially for the short season varieties with others going up to 10 months or more. Harvesting can be done while the maize is green or when dry. Cobs are harvested by hand or by use of mechanized harvesters e.g. combine harvesters.

In Kenya, Maize provides a basic diet to millions of people. The total land area under maize production is about 1.5 million hectares, with an annual average production estimated at 3.0 million metric tons, giving a national mean yield of 2 metric tons per hectare. Typically, yields range from 4 to 8 T/Ha in the high potential highlands of Kenya, representing only 50% (or less) of the genetic potentials of the hybrids. Highland maize varieties are grown on some 40-50% of the total maize area, representing 600,000 - 800,000 Ha.

Maize production has not kept pace with the population increase, although breeders and agronomists have exploited its genetic potential for yield. Constraints for maize production include drought, low soil fertility, pests, and diseases. Foliar (leaf) and stalk/ear rot diseases and stem-borers cause great losses in maize production in the humid transitional and high tropics of Kenya. Crop protectionists have put tremendous effort into identifying the disease and pest problems. Some of the most common types of maize diseases affecting Kenyan farmers include Downy mildew, Northern Corn (Turcicum), Southern leaf blight, gray leaf spot (GLS), Common rust, Common smut, Head Smut, Maize lethal necrosis disease, and Maize streak virus.

1.2 Problem Statement

Maize production has not kept pace with the population increase, although breeders and agronomists have exploited its genetic potential for yield.Constraints for maize production include drought, low soil fertility, pests, and diseases. Foliar (leaf) and stalk/ear rot diseases and stem-borers cause great losses in maize production in the humid transitional and high tropics of Kenya. Developing techniques to accurately classify crop leaf disease categories is critical for disease prevention. Computer vision systems can be useful by building a model that can identify the various types of maize crop diseases and classify them accurately.

1.3 Objectives

Our main objective for this study will be to provide a system that can be used by farmers to identify the kind of disease affecting their maize crop and offer possible solutions and remedies to fight the disease and prevent it from spreading.

Specific objectives:

1. To determine the model's overall effectiveness in classifying diseases using both validation and test dataset.
2. To identify the specific features that differentiate the different kinds of maize crop diseases.
3. Deploy the trained model to create an easy-to-use web application.

To achieve the objective of this study, the analysis will seek to answer the following questions:

1. What type of disease does an image belong to?
2. What are the specific features that differentiate the different kinds of maize diseases?
3. How will the model be designed, validated and deployed?

1.4 Impact

This research seeks to reduce the losses in yield suffered by maize farmers due to crop diseases and support them in increasing their crop yield by correctly identifying the type of disease affecting their maize and offering possible solutions to the problem.

1.5 Success Criteria

A successful analysis will correctly recognize and classify maize crop diseases according to discriminant image features. The model will assess the confusion matrix which depicts the inter-class variability in disease classification.

Precision and Recall will be key metrics in model evaluation to ensure that diseases are correctly identified so that the appropriate disease management measures can be undertaken by the farmer. These metrics will also be compared for models tested with various image sizes and augmentation settings.

1.6 Assessing the situation

1.6.1 Resource inventory

1. Datasets: Scraped data from the internet.
2. Softwares:
    a. Github - Collaboration
    b. Google Colab
    c. WandB - Metrics Tracking
    d. Tensorflow Library
    e. Streamlit - Deployment
    f. Fatkun Extension - Scrape Images
    g. Trello - Project Tracking
    h. Captum - model interpretability.
3. Project Implementation links

    i) Wandb

    ii) Github

    iii) Trello

    iv) Project presentation

1.6.2 Constraints

Insufficient image data for our classes.

1.7 Implementation plan

Phase	Time-Frame
Business Understanding	1 Day

Data scraping	3 Days
Data Understanding	1 Day
Data Preprocessing and Cleaning	3 Days
Modeling	5 Days
Evaluation	5 Hours
Deployment	2 Days

2.0 Data Understanding

2.1 Data Collection

Image Data

We used the Fatkun extension to develop web scraping scripts to collect images from the internet for this project. We annotated the data according to the various classes chosen for the study.

Scope

Our research lays its focus on five classes of maize crop diseases that greatly affect maize yield which in turn results in food insecurity and production loss.
The following are the different types of diseases that this project focused on:

a) **Northern Leaf Blight**

Northern maize leaf blight is one of the major diseases that endanger the health of maize in especially in important maize growing areas of Kenya. Yield loss is caused predominantly
through the loss of photosynthetic leaf area due to blighting. If NCLB establishes before silking and spreads to upper leaves during grain filling, severe yield losses can occur.

Symptoms

In the beginning we will notice elliptical gray-green lesions on leaves. As the disease progresses these lesions become pale gray to tan color. Later stage the lesions look dirty due to dark gray spores particularly under the lower leaf surface. The disease can be easily identified in the field due to its long, narrow lesions which are unrestricted by veins.

Image illustration of Northern Leaf Blight

Close up of leaves demonstrating typical tan-colored, cigar-shaped lesions

b) Gray Leaf Spot (GLS)

Grey leaf spot (GLS) is a foliar fungal disease that affects maize, also known as corn. There are two fungal pathogens that cause GLS, which are Cercospora zeae-maydis and Cercospora zeina.

Symptoms

Gray leaf spot lesions begin as small necrotic pinpoints with chlorotic halos, these are more visible when leaves are backlit.

symptoms of GLS are rectangular, brown to gray necrotic lesions that run parallel to the leaf, spanning the spaces between the secondary leaf veins.Lesion expansion is limited by parallel leaf veins, resulting in the blocky shaped “spots”.
Because early lesions are ambiguous, they are easily confused with other foliar diseases such as anthracnose leaf blight, eyespot, or common rust.

Image illustration Gray Leaf Spot Image

c) Common Rust

Common rust is caused by the fungi, Puccina sorghi.

Symptoms

Oval or elongated cinnamon brown pustules on upper and lower surfaces of leaves; pustules rupture and release powdery red spores; pustules turn dark brown-black as they mature and release dark brown powdery spores; if infection is severe, pustules may appear on tassels and ears and leaves may begin to yellow; in partially resistant corn hybrids, symptoms appear as chlorotic or necrotic flecks on the leaves which release little or no spore.

d) Smut diseases

Smut is characterized by fungal spores that accumulate in sootlike masses called sori, which are formed within blisters in seeds, leaves, stems, flower parts, and bulbs. The sori usually breaks up into a black powder that is readily dispersed by the wind. Many smut fungi enter embryos or seedling plants, develop systemically, and appear externally only when the plants are near maturity. Other smuts are localized, infecting actively growing tissues. Examples of smut diseases that attack maize crops are head smut and common smut.

Visual representation of head smut

Visual representation of Common smut

3.0 Data Preparation

3.1 Loading the datasets

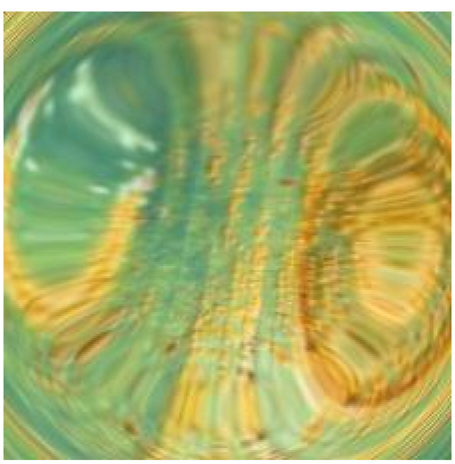
The scraped image folders were assigned to their respective classes; Northern Blight, Smut, Common Rust, Maize streak, Gray Leaf Spot and Healthy. The healthy class was introduced to act as a control target value. These were randomly split into the six class categories with 500 images per class and stored in the train, test and validation folders.
When loading the data into our coding environment, the train, test and validation sets were assigned paths and categories created.

3.2 Image Preprocessing

In order to prepare the data for training our data was resized into image sizes of 224 by 224. The features were then scaled so that the range of feature distribution would be equal.

3.3 Image Augmentation

Our original image dataset possessed high class imbalance which would in turn lead to overfitting and wrong prediction of our models. Augmentation transforms the image to increase the ability of the model to recognize different versions of an image. The breadth of the model increases and it becomes better suited to recognize target objects in images of varied contrast, size and from changed angles.



Example of an augmented common rust image from the train set.

4.0 Modelling

4.1 Computer Vision Systems

Computer vision combines cameras, edge - or cloud based computing, software, and artificial intelligence (AI) to enable systems to “see” and identify objects. The system uses deep learning to form neural networks that guide systems in their image processing and analysis. Once fully trained, computer vision models can accurately identify and detect objects, analyze and make meaningful interpretations out of a sequence of images.

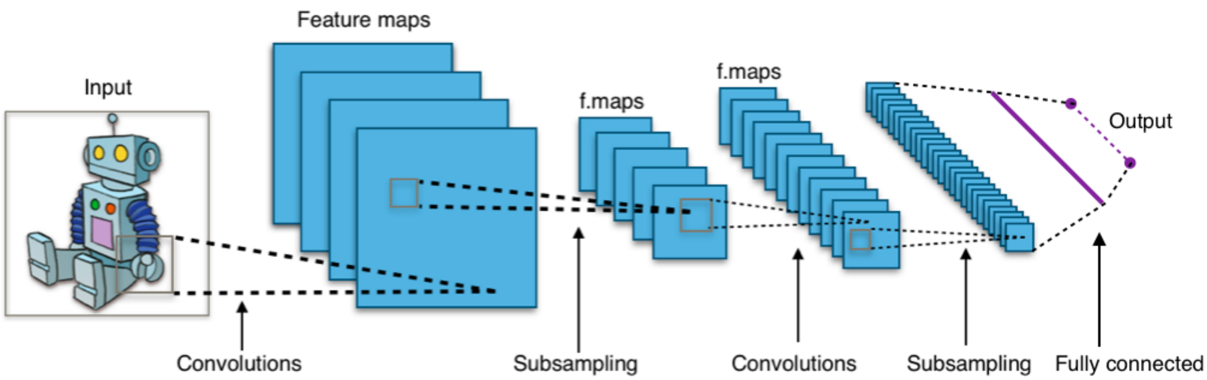
Computer vision systems have been applied in agriculture for various use cases. From better productivity to lowering production costs with automation, it has improved the agricultural sector’s overall functioning with its automation and detection capabilities.

The field of computer vision includes a set of main problems such as image classification, localization, Image segmentation and object detection. Image classification in our use case is identifying and labelling the pathology affecting the maize crop.

We identified and made use of the following algorithms to classify the maize diseases based on our maize crop disease datasets:

4.1.1 Convolution Neural Networks

Convolutional neural networks are a type of deep learning technology that has grown popular in computer vision and is receiving attention in a range of fields, including agriculture. A convolutional neural network is made up of numerous layers, such as convolution layers, pooling layers, and fully connected layers, and it uses a backpropagation algorithm to learn spatial hierarchies of data automatically and adaptively.



How a typical CNN works

In our case, the pixel features of the image are used as the input for the convolutional neural network (CNN) and trained and the maize crop disease identification as the output of the model.

Building the baseline CNN model

We implemented the sequential model by keras to build our baseline convolutional neural networks. For our model architecture there are three Convolution and Pooling pairs followed by a Flatten layer which serves as a connection between Convolution and Dense layers.

The activation functions used for the convolution layers were Relu and Softmax for the final dense layers which gives the probability of each class. The model makes predictions based on the class with the highest probability. Training the model on 20 epochs and three convolutional layers gav the best accuracy of 82.05%.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
rescaling (Rescaling)        (None, 224, 224, 3)       0
_________________________________________________________________
conv2d (Conv2D)              (None, 224, 224, 16)      448
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 112, 112, 16)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 112, 112, 32)      4640
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 56, 56, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 56, 56, 64)        18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 28, 28, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 50176)             0
_________________________________________________________________
dense (Dense)                (None, 128)               6422656
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
```

4.2.2 Transfer Learning

When elements of a pre-trained machine learning model are reused in a new machine learning model, this is known as transfer learning. Generalized information can be transferred across the two models if they are designed to accomplish similar tasks. To train models, this type of machine learning uses tagged training data.

For our image classification, we implemented the following pretrained models.

a) *MobileNet*

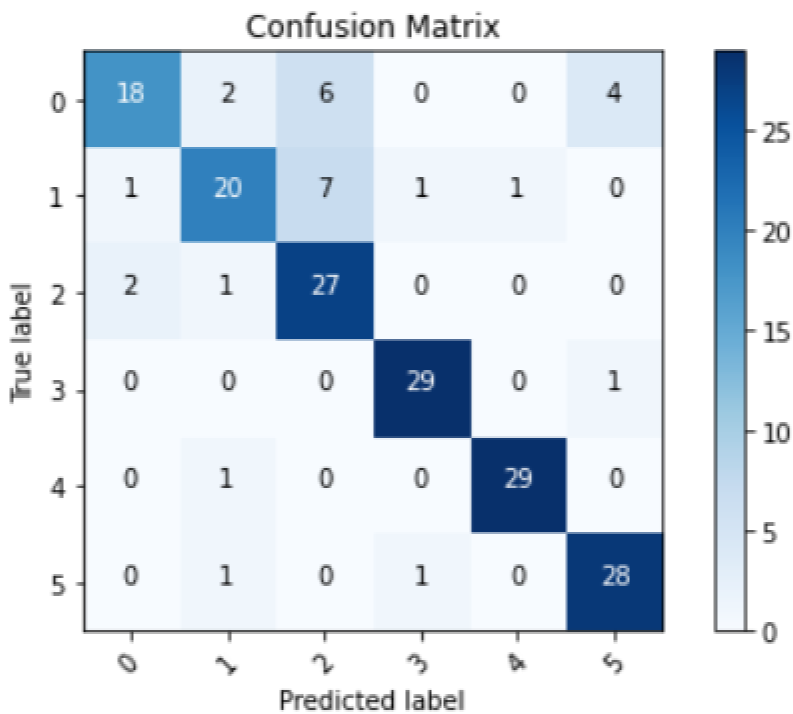
MobileNets is TensorFlow's first mobile computer vision model, and it is based on a streamlined design that leverages depth-wise separable convolutions to generate light weight deep neural networks.The following are the parameters that were used in this model:

- Learning_rate = 0.0001,
- Epochs = 50,
- Loss_function =categorical_crossenthropy,
- Fine tuned layers = 10,
- Optimizer = Adagrad,
- Architecture = MobileNet

Metrics

For the six disease classes, the MobileNet model has accuracy of 84 percent, precision of 68 percent to 97 percent, recall of 60 percent to 97 percent, and F1 score of 71 percent to 97 percent.

This was the most effective of the Transfer Learning Models



b) ResNet50

ResNet50 is a pre-trained model from the ImageNet database that has 50 layers and a 224 × 224 image patch size. The following are the parameters that were used in this model:

- Learning rate = 0.001,
- Epochs = 10,
- Batch_size = 32,
- Loss function = sparse_categorical_crossenthropy -this is because we have an index of exclusive (sparse) classes we are trying to predict against
- Optimizer = Adam,
- Architecture = CNN four layers.

Metrics

For the six disease classes, the ResNet50 model has accuracy of 12 percent, precision of 7 percent to 23 percent, recall of 7 percent to 18 percent, and F1 score of 7 percent to 21 percent.

c) InceptionV3

InceptionV3 is a convolutional neural network that was developed as a Googlenet module to aid in picture processing and object detection. Google's Inception Convolutional Neural Network is now in its third edition. Convolutions, average pooling, max pooling, concats, dropouts, and fully linked layers are among the symmetric and asymmetric building components in the model.The following are the parameters that were used in this model:

- Learning rate = 0.01
- Epochs = 10,
- Batch_size = 32,
- Loss_function = sparse_categorical_crossenthropy,
- Architecture = InceptionV3,
- Optimizer = Adam.

Metrics

For the six disease classes, the InceptionV3 model has accuracy of 74% percent, precision of 60 percent to 93 percent, recall of 50 percent to 93 percent, and F1 score of 56 percent to 93 percent.

d) VGG16

The VGG16 is a 16-layer deep convolutional neural network. You can load a pre-trained version of the network from the ImageNet database, which has been trained on over a million photos. In ImageNet, a dataset of over 14 million images belonging to 1000 classes, the model achieves 92.7 percent top-5 test accuracy. The model has a 224 × 224 image patch size.The following are the parameters that were used in this model:

- Learning rate = 0.001,
- Epochs = 10,
- Batch_size = 32,
- Loss_function = categorical_crossenthropy,
- Architecture = VVG16,
- Optimizer = Adam.

Metrics

For the six disease classes, the VGG16 model has accuracy of 18% percent, precision of 5 percent to 26 percent, recall of 3 percent to 30 percent, and F1 score of 4 percent to 28 percent.

4.2.3 Vision Transformers

The Vision Transformer, or ViT, is a model for image classification that employs a Transformer-like architecture over patches of the image. An image is split into fixed-size patches, each of them are then linearly embedded, position embeddings are added, and the resulting sequence of vectors is fed to a standard Transformer encoder. In order to perform classification, the standard approach of adding an extra learnable “classification token” to the sequence is used.

We implemented the Vision Transformer model using the hugging face library and Pytorch. For the top linear layer, we used a pretrained model for feature extraction; google/vit-base-patch-224-in21k. The image was resized into a 224 by 224 sized image and parameters defined.

The parameters for the model were:

- Learning rate = 2e-5,
- Epochs = 3, 10
- Batch_size = 32,
- Loss function = cross entropy loss function,
- Optimizer = Adam,

Metrics

For the six disease classes, the Vision Transformer model has accuracy of 94% percent, precision of 81 percent to 100 percent, recall of 74 percent to 100 percent, and F1 score of 77 percent to 96 percent.

	precision	recall	f1-score	support
0	0.81	0.74	0.77	23
1	1.00	0.93	0.96	29
2	0.91	0.94	0.93	34
3	1.00	1.00	1.00	33
4	1.00	1.00	1.00	29
5	0.91	1.00	0.96	32
accuracy			0.94	180
macro avg	0.94	0.94	0.94	180
weighted avg	0.94	0.94	0.94	180

Classification report for the best performing model:Classification report for the best performing

5.0 Evaluation

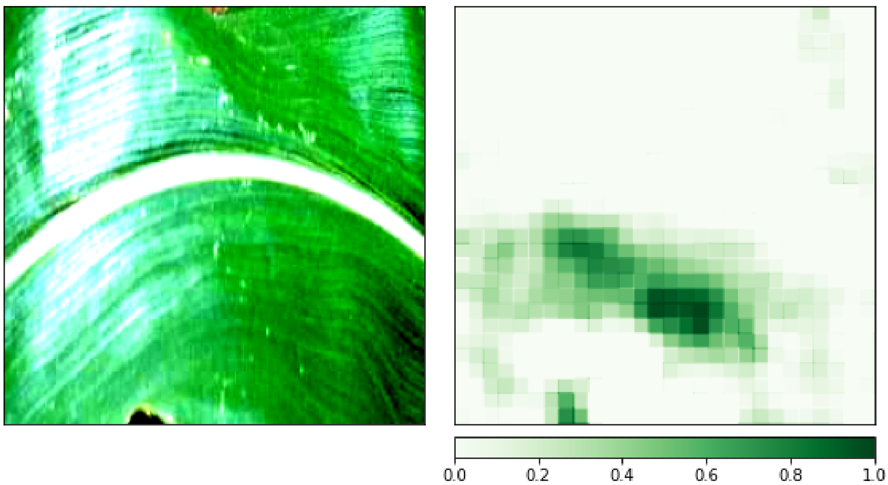
In order to give us an idea of how well our model is working, we specified different metrics to use in addition to loss for our models.

Model Performance

Model	**Accuracy(%)**	**Precision (%)**	**Recall(%)**
MobileNet	91.21	97	97
VGG16	18	26	30
ResNet50	12	23	18
InceptionV3	68	93	93
Vision Transformers	94	94	94

Model Interpretability with Captum

Captum helps one to understand how the data features impact your model predictions or neuron activations, shedding light on how your model operates.
Using Captum we applied feature attribution algorithms such Integrated Gradients in a unified way to attribute the predictions of an image classifier to their corresponding image features and visualize the attribution results.



This is a visualization example of showing the RGB attribution results of a healthy leaf image.

Conclusion

Of all the models worked on, the one that emerged the best in drawing inferences was the MobileNet model followed by the Hugging Face Vision Transformer (ViT) which was used for deployment.

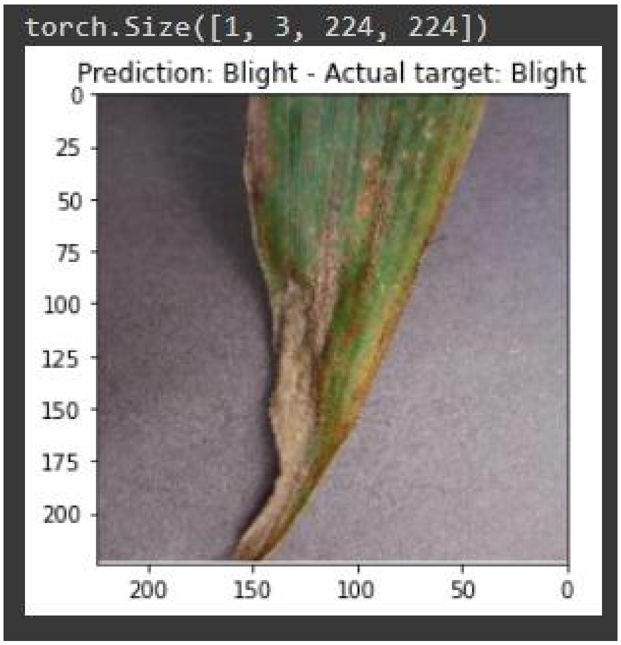


Image: Correct inference made by the ViT model predicting Blight disease in maize crops.

Challenges

- Generating predictions for foliar diseases such as anthracnose leaf blight, eyespot, gray leaf spot and common rust posed a challenge as lesions presenting themselves on the leaves at different stages are ambiguous hence they are easily confused with others.
- The torch library used was too large to have the app hosted on Heroku.

Recommendations

- Planting should take place within the first two weeks of rainy weather. Before sowing farmers should ensure that there is at least 30 cm of damp soil across the soil profile and that no stones or heavy soil clods cover the seeds during planting.
- Crop rotation is highly recommended to prevent the spread of maize diseases and insect pests. Depending on the area, rotation can be done with beans, cowpeas, peas, or potatoes, but rotation of maize with other cereal crops such as sorghum and millet should be avoided.
- Weeds serve as an alternate host for pests and diseases, and by competing for moisture, nutrients, space, and light, they diminish maize yields. The first four to six weeks after the crop emerges are the most critical for weed competition in the maize plant's life.Weeding can be done by hand or with herbicides.
- Planting using commercially obtained certified seeds that have fungicide and insecticide as normal procedure to keep diseases and insects at bay.
- Since most farmers in Kenya are small scale farmers who carry out mixed crop farming, the application can be extended to consider crops grown alongside the maize crop.