Republic of the Philippines
Western Mindanao State University
College of Computing Studies
DEPARTMENT OF COMPUTER SCIENCE
Zamboanga City

Analyzing Above-Ground Root Crop Features For Non-Invasive Monitoring Using Deep Learning

A Thesis Presented to the Faculty of

Department of Computer Science

College of Computing Studies

In Partial Fulfillment of the Requirements for the Degree of

Bachelor of Science in Computer Science

Ronald L. Aljas

Aljamer H. Tajala

Kristina Marie V. Gadem

Kristin Mae V. Delos Santos

Researchers

Ceed Jezreel B. Lorenzo

Adviser

May 2026

APPENDIX A

Data Collected

Data Gathering Instrument

Different instruments will be used to gather and process data for the non-invasive monitoring of sweet potato crops based on deep learning. The major data gathering instruments are:

- Camera System – A high-resolution camera captured photos of sweet potato and potato plants above-ground features in different conditions. The camera will be kept at the same height and angle to capture images with consistency.
- Lighting Equipment – Professional lighting will be used to eliminate shadows and provide even illumination for all photos, thereby reducing variations due to differences in environmental light.

Data Gathering Technique and Procedures

Sweet Potato Farming Questionnaire

1. How long have you been planting sweet potatoes?

2. What variety of sweet potato do you usually grow?

3. How many times do you harvest sweet potatoes in a year?

4. How many days or weeks to take into consideration usually take from planting to harvest?

5. Which part of the sweet potato plant is most and least sensitive

6. How do you handle problems like pests, weather, or soil issues while growing sweet potatoes?

7. Can sweet potatoes be overripe or overgrown if not harvested on time? If Ues, What will happen?

8. How do you usually check if the sweet potato is ready for harvest?

9. Are you using any tools or technology to monitor your crops? If yes, what kind?

10. Would you be interested in using our system that can help you check your plant specifically your sweet potato?

These questions will help achieve the study's goal. Responses of respondents will guide image data needs, confirm visual signs matter most, shape CNN training, and ensure the mobile app works reliably in real fields.

The following data collection procedures were used within this study:

- Image Acquisition – A series of images of sweet potato plants were obtained with a high-resolution camera in controlled conditions.

- Feature Extraction – Crop features above ground like plant structure, color, and leaf size were extracted from the images.

- Dataset Preprocessing – Image processing methods were applied, including noise removal, contrast adjustment, and resizing, for improving dataset quality.

- Annotation and Labeling – Images were manually annotated with annotation tools to point out salient crop features to be employed in the course of CNN-based model training.

The following step-by-step data collection process is described:

Site Selection and Setup

- Conduct the imaging in a natural setting, specifically in the researcher’s backyard, rather than in a controlled environment.
- A mobile phone camera will be used to capture images of each crop from different angles to ensure varied perspectives for analysis.

Image Capturing

- Take photographs under different illuminating conditions and at different times to collect differentiated data sets.
- Capture multiple images from various angles to enhance the overall generalization ability of the model.

Preprocessing and Data Cleaning

- Resize images to default resolution and format.
- Use preprocessing methods like normalization and denoising. (include: image flipping and rotation)

Annotation and Labeling

- Annotation tools (LabelImg, Roboflow) are used to annotate notable features in all the images.
- Classify images by crop growth stage, health, and other relevant features.

Data Storage and Organization

- Store images in labeled folders according to classes.
- Upload data sets in cloud storage for easy access and backup.

Validation and Quality Control

- Check images gathered for inconsistency or error verification. The cleanliness and neatness of images should be ensured at this stage.
- Maintain balance in datasets to prevent model bias.
- Farmers that are expert’s in harvesting sweet potatoes.

Data Analysis & Model Training

The data collected was put through a controlled process of analysis to provide high-quality inputs to the deep model. The process included data cleaning, preprocessing, feature extraction, and CNN-based classification.

1. Data Preprocessing and Cleaning

   Raw images were preprocessed and cleaned before analysis for quality improvement and elimination of inconsistencies:

   - Image Standardization – The images were standardized to a particular resolution to provide consistent input sizes to the CNN model.
   - Noise Reduction – OpenCV filters were used to minimize image noise and increase clarity.
   - Contrast Adjustment – Histogram equalization methods were used to increase feature saliency.
   - Augmentation Techniques – Rotation, flipping, and brightness adjustments were used to augment the dataset and increase model robustness.

2. Feature Extraction and Dataset Preparation

   The images were processed to provide significant features utilizing Jupyter Notebook and TensorFlow:

   - Color Analysis – Color histograms extraction to study leaf pigmentation variations.
   - Shape and Size Analysis – Contour detection and edge detection were used to capture plant growth indicators.

- Texture Analysis – Gray-Level Co-occurrence Matrix (GLCM) was employed to analyze leaf texture for disease identification.
- The features extracted were stored in SQLite, whereas raw and processed images were stored in the local file system for offline retrieval.

3. Training of the Deep Learning Model (CNN-Based Analysis)

The above-ground crop features were classified and analyzed using a Convolutional Neural Network (CNN). The training procedure was as follows:

- Train-Test – Partitioning images into training (80%), validation (10%), and testing (10%) sets.
- Model Architecture – A multi-layer CNN model was created using convolutional, pooling, and fully connected layers to learn hierarchical features.
- Loss Function and Optimizer - The Categorical Cross-Entropy loss function was also used, along with a well calibrated Adam optimizer.
- Model Training –TensorFlow/Keras was used to train the models, with early stops to prevent overfitting.

4. Model performance and evaluation metrics.

The equivalent CNN model was trained using the following parameters:

- Accuracy is defined as the proportion of cropped photos that are correctly classified.
- Precision & Recall – Model performance assessment in recognizing certain crop conditions.

- F1-Score – Precision-recall tradeoff for improved performance reporting.

  Confusion Matrix – Displaying errors in classification to spot misclassifications.

5. Deployment and Front-End Integration

After the CNN model reached peak performance, it was deployed to TensorFlow Lite (TFLite) for integration into the Android studio app. The images after processing and model predictions were cached locally in SQLite for complete offline capability.