CELEBRATING A HUNDRED YEARS OF AGRICULTURAL RESEARCH AND INNOVATION: PERSPECTIVES ON TROPICAL FOOD AND AGRICULTURAL SYSTEMS

Generating an open Caribbean-based UAV dataset of crops, weeds and landscape level data for precision agriculture in Small Island Developing States

Gabriela Sewdhan, Michael Ali, Roganci Fontelera, Ishala Gosine, Omar Mohammed, Professor Patrick Hosein

Motivation

- Caribbean Small Island Developing States (SIDS) are extremely vulnerable to food insecurity, further intensified by climate variability risks.
- Despite diversity, 60-100% food supply is imported.
- The adoption of **machine learning** in **precision agriculture** can build capacity for farm system productivity and resilience, helping farmers make better informed decisions and identify efficiencies, leading to higher productivity and profitability.
- To achieve agricultural resilience, crop, weed, and landscape level data, collected via an **Unmanned Aerial Vehicle** (UAV), as well as educational sources for emerging technologies, are needed.

Solution

To provide an open-sourced Caribbean-based annotated dataset of UAV images, documenting the process of data acquisition, data preprocessing and outlining its application for weed, plant health and water stress detection, and land use and cover classification.

Machine Learning in Precision Agriculture

The data collected will be used to train machine learning algorithms. Outlined below are some ways Machine Learning can assist agriculture.

Crop Yield Prediction
Identifies the output yield of crops for the next few weeks can help farmers obtain max profit

Intelligent Spraying
Detect weed affected areas and precisely spray herbicides in the right region, reducing the usage of herbicides

Crop and Soil Monitoring
Allows for monitoring of crop health, and diagnosing pests/soil defects, soil nutrient deficiencies etc.

Agriculture Robots
Autonomous robots to harvest huge volume of crops at a higher volume and faster pace

Disease Diagnose
Classification of Plant diseases helps farmers control the disease through proper strategies

Predictive Insights
Gather insights on the right time to sow, and impacts by weather conditions

Crops Used

Brassica rapa subsp. Chinensis (Pak Choi)

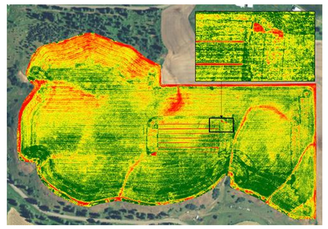
- Easy, fast growth phases
- Has moisture retentive, nutrient rich soil
- Phenotypically sensitive species, thus good for water stress testing
- Huge supply

Hot Peppers

- Can't outgrow weeds, great for collecting weed data

Normalised Difference Vegetation Index (NDVI) Imagery

NDVI provides accurate information relating the biomass levels. The **NDVI** values can provide useful conclusions regarding plant health, crop diseases, water stress, pest infestation, nutrient deficiencies and other conditions that affects crop productivity. Given this, NDVI images will be obtained from a UAV for the data collection process.

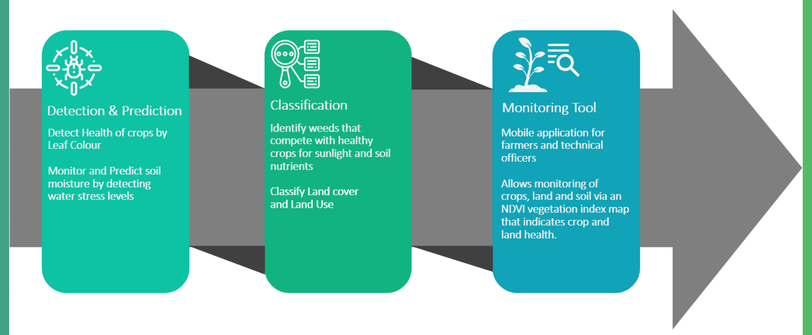


Data Collection Methods

- UAVs will be used to collect aerial images of **Pepper** and **Pak Choi** farms on flat land at a height of 5 meters or below.
- 4000 images consisting of both dry and wet season data, will be collected for each problem to be addressed (more details below).
- Data on farms with weeds (for **weed detection**), and data without (for **plant health detection**) will be collected.
- NDVI images will be collected for **plant health detection models**.
- Water sensor data on the same farm, with irrigated, moderately irrigated and non-irrigated (control) experiments will be collected and conducted. For these experiments, random sampling will be used of soil moisture content (for **water stress level detection**).
- Land use and cover data on different types of lands will be obtained.
- Data will be labelled manually with **MATLAB**, to be used as training data.
- Before using data for training the machine learning models, they will be cleaned and pre-processed.

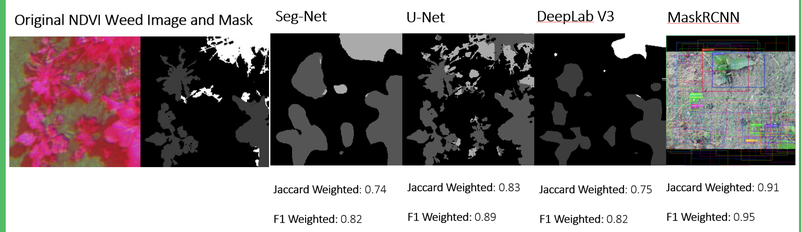
Problems to be addressed

Using the collected dataset, we will develop machine learning models to examine the following:



Work Done

- Data collection at **Pepper** and **Pak Choi** farms are in progress. The images above contains photos of farm visits completed.
- Using **WeedNet Dataset**, classification models were implemented, results are shown below. A higher **Jaccard** and **F1 score** means **higher accuracy** of weed detection. MaskRCNN and U-Net currently has the highest.



- Water Stress Level experiments are in progress.
- Plant Health Model implementation and data collection are in progress.

Remarks

Our custom dataset, developed from the Drone imagery and labelled with **MATLAB**, will be managed with **GeoNode**, a web-based application to upload and share our geospatial data. Our dataset will be freely available to the public and future agricultural and environmental developers. Furthermore, the pretrained models, and source code for the model architecture, will be made freely available on **GitHub** under the MIT license.