Garden Crop Monitoring, Classification & Weed Detection Using CV & ML

Presented by :

GIDEON MPUNGU 2300704617 2023/HD05/04617U

YouTube Video Link: https://youtu.be/IVQ76VWOdhI

MAKERERE UNIVERSITY

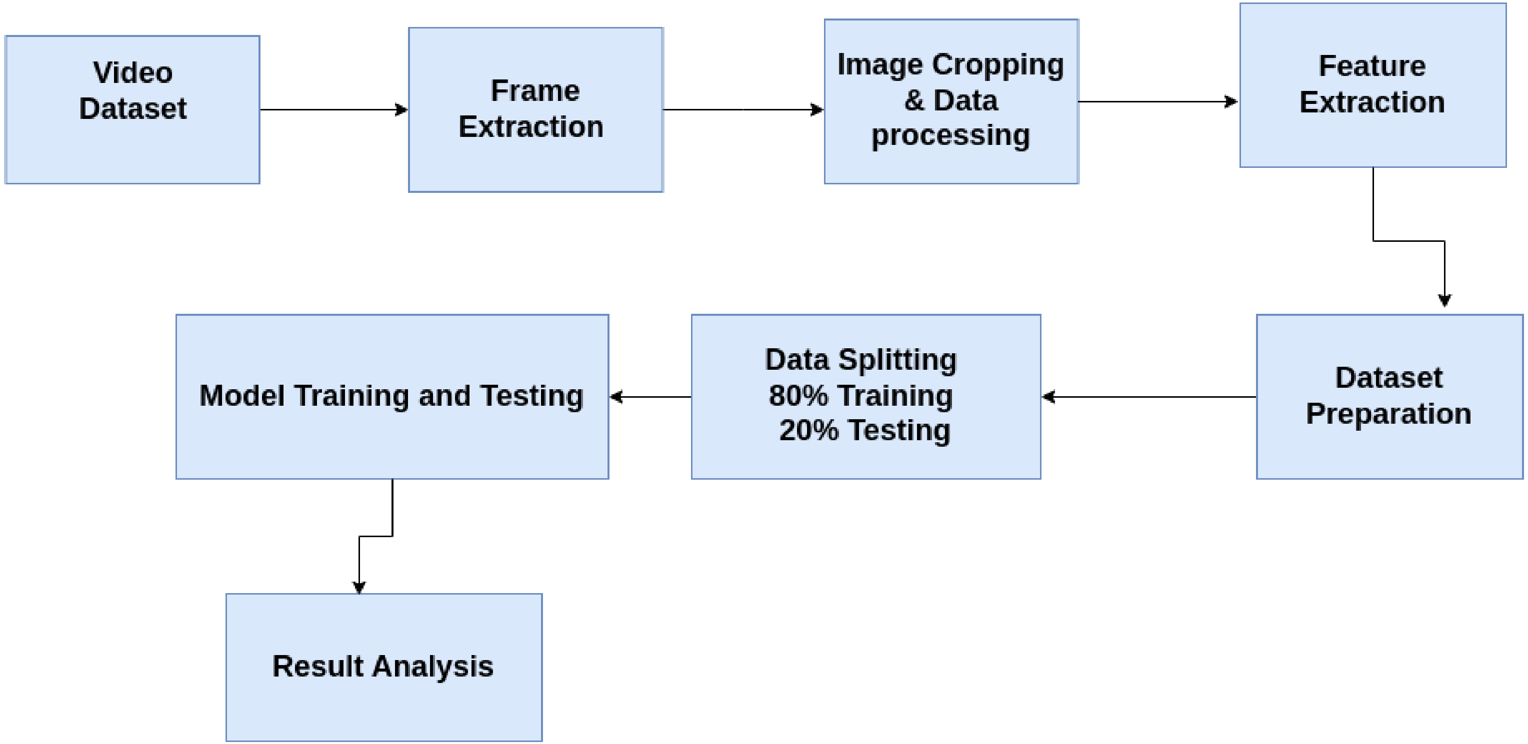
Overview

- Introduction
- Methodology
- Results and Discussion
- Conclusion
- References

Introduction

- Ever growing world population set to hit almost 10 billion by 2050.
- Crop fields face a challenge of weed control that cause {sun, nutrients, water} competition and thus low yields.
- Calls for measures to ensure efficient and continuous food production for sustenance.
- Use CV and ML to monitor gardens and detect weeds.
- Current models show promise but mostly focus on one feature at a time.
- Focus of research: To build accurate traditional ML models to properly classify crops(cassava, maize, sugarcane) and grass(weeds) in a garden based on a combination of features for optimal performance.

Methodology

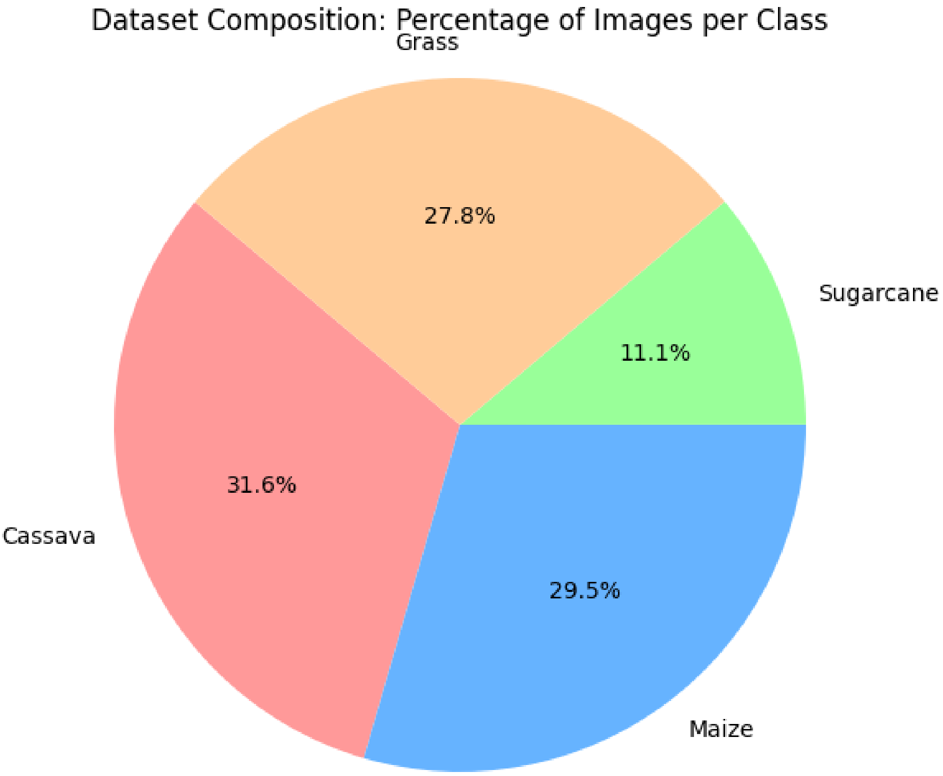


Dataset Description & Extraction

- Obtained frames by passing the garden video through a cv2 package "cv2.VideoCapture" and specifying the framerate(0.5), i.e. 2 frames per second.
- Saved frames were further manually cropped using snipping tool to obtain datasets of the chosen crops, i.e. cassava, maize, grass, sugarcane.
- The resultant 1034 frames were saved in crop specific folders to form a dataset for our feature(s) extraction.

EDA & Data Processing

- **Image resizing:** This was accomplished by resizing the frames to 500x500.
- Dataset Distribution was fair among the crops apart from sugarcane.

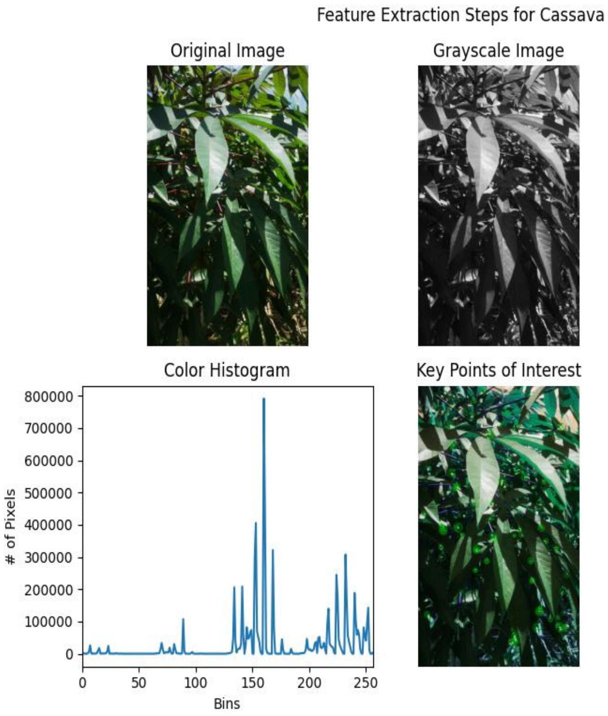


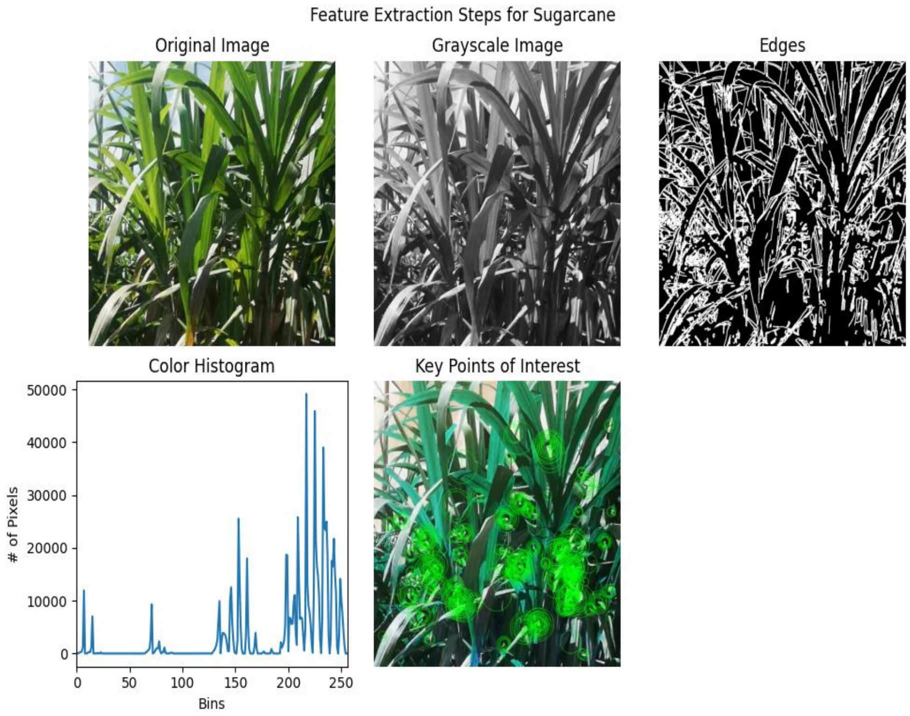
Feature Extraction and Combination

- Extracted 3 kinds of features;
  - Color Histogram
  - Edge Detection
  - Key points of Interest using ORB
- Wrote functions for each of the above.
- Combined all the features to create a training dataset.
- Dataset split: 80% Training , 20% Testing.

```
Training set shape: (827, 81312)
Test set shape: (207, 81312)
```

Sample Extracted Features





Models Trained

- 4 models were trained including;
    - Random Forest
    - Naïve Bayes
    - Decision Tree
    - SVM
- The 3 best performers (Random Forest, Naïve Bayes, Decision Tree) were chosen to create a Voting Classifier ensemble model
- All models were trained on the extracted training dataset of combined features.

Model Evaluation

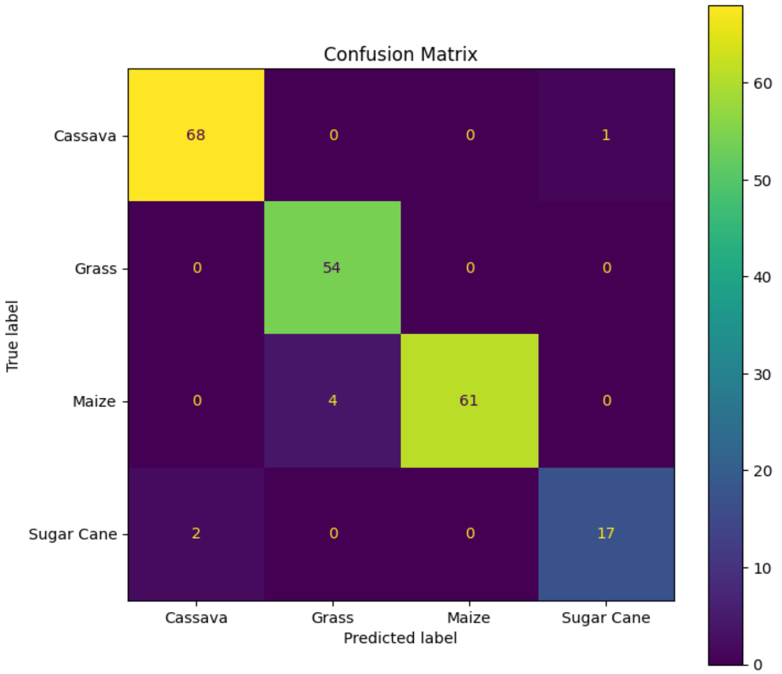
- Used the following evaluation criteria:
  - Accuracy
  - Classification Report {precision, recall, f1 score}
  - Confusion Matrix

Results & Discussion

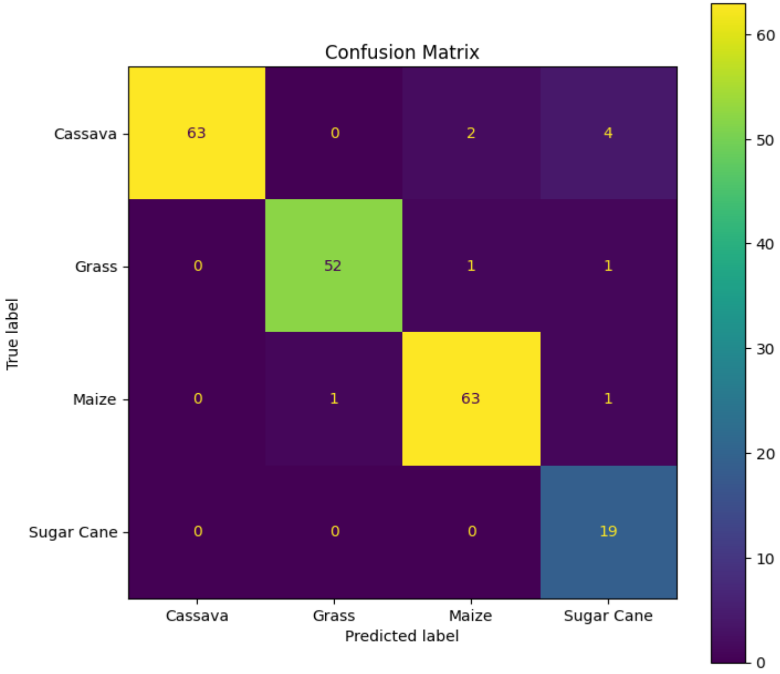
Model	Accuracy (%)	F1 score	Precision	Recall
Random Forest	97	0.96	0.95	0.96
Decision Tree	95	0.92	0.96	0.94
Naïve Bayes	93	0.93	0.90	0.91
SVM	92	0.94	0.87	0.90
Voting Classifier	98	0.97	0.97	0.97

Results: Confusion Matrices – RF & DT

Random Forest



Decision Tree



Results Discussion

- **Random Forest:** Best performance achieving an accuracy score of 97% demonstrating a high percentage of correctly predicted crops.
  - **Cassava:** Model had a high precision & F1 score, and perfect recall score showing a high sensitivity for cassava
  - **Grass:** The model had a good precision and perfect recall score for grass. It correctly predicted all the grass instances.
  - **Maize:** The model had a perfect precision for maize meaning that out of all the instances predicted as maize, 100% were actually maize.
  - **Sugarcane:** The model had a good precision for sugarcane meaning that out of all the instances predicted as sugarcane, most were actually sugarcane.

Accuracy: 0.966183574879227

Classification Report:

	precision	recall	f1-score	support
cassava	0.97	0.99	0.98	69
grass	0.93	1.00	0.96	54
maize	1.00	0.94	0.97	65
sugarcane	0.94	0.89	0.92	19
accuracy			0.97	207
macro avg	0.96	0.95	0.96	207
weighted avg	0.97	0.97	0.97	207

- **Decision Tree:** The model's overall accuracy is very high at approximately 95%, indicating excellent performance across the dataset.
  - **Cassava:** Showed very high precision, recall, and F1 score, suggesting the model is highly effective at identifying cassava. Only few instances of cassava were identified as maize or sugarcane.
  - **Grass:** Excellent recall and good precision, indicating the model correctly identifies almost all grass instances but with a few false positives.
  - **Maize:** Very high precision but slightly lower recall, suggesting the model is good at identifying maize when it predicts it but misses some actual maize instances and classifies them as grass or sugarcane.
  - **Sugarcane:** High precision and recall, indicating a good balance, though slightly lower than other classes, which suggests that the model performs well on sugarcane but with room for slight improvement

Accuracy: 0.9516908212560387

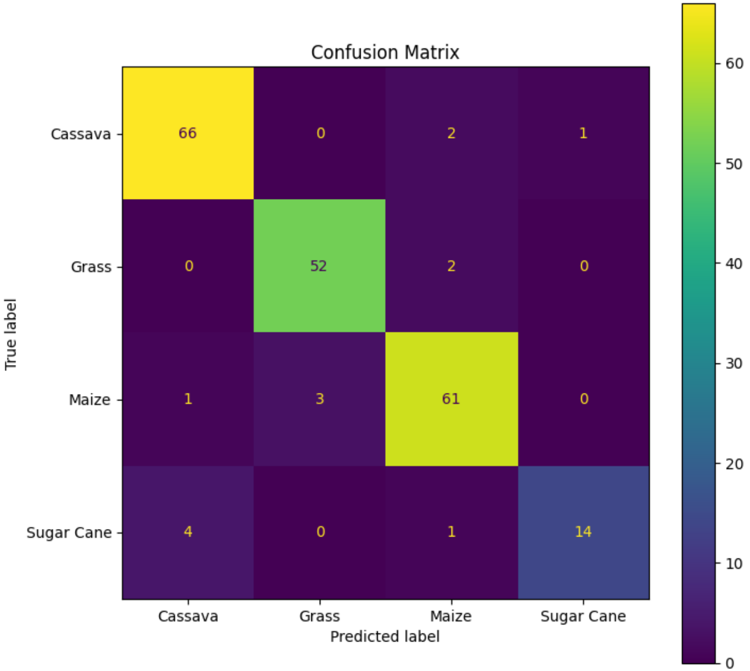
Classification Report:

	precision	recall	f1-score	support
cassava	1.00	0.91	0.95	69
grass	0.98	0.96	0.97	54
maize	0.95	0.97	0.96	65
sugarcane	0.76	1.00	0.86	19
accuracy			0.95	207
macro avg	0.92	0.96	0.94	207
weighted avg	0.96	0.95	0.95	207

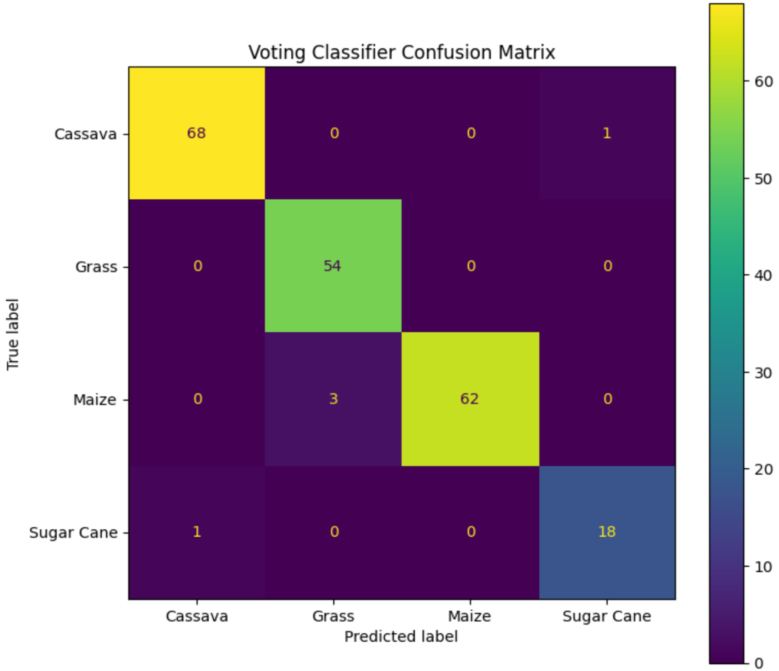
MAKERERE UNIVERSITY

Results: Confusion Matrices – NB & VT Ensemble

Naïve Bayes



Ensemble Voting Classifier



Results Discussion Cont'd

- **Naive Bayes:** Naive Bayes had an accuracy score of 93% indicating good performance across the dataset.
  - **Cassava:** High precision, recall, and F1 score, suggesting the model is very effective at identifying cassava.
  - **Grass:** Slightly lower precision but excellent recall, indicating the model correctly identifies almost all grass instances but with some false positives where it would classify grass as sugarcane.
  - **Maize:** High precision but slightly lower recall, suggesting the model is good at identifying maize when it predicts it, but misses some actual maize instances and classifies them as cassava or sugarcane.
  - **Sugarcane:** Has the lowest precision and recall among the classes, indicating the model struggles more with sugarcane predictions, both in terms of false positives and false negatives.

```
Accuracy: 0.9323671497584541
Classification Report:
              precision    recall  f1-score   support

     cassava       0.93      0.96      0.94        69
       grass       0.95      0.96      0.95        54
       maize       0.92      0.94      0.93        65
   sugarcane       0.93      0.74      0.82        19

    accuracy                           0.93       207
   macro avg       0.93      0.90      0.91       207
weighted avg       0.93      0.93      0.93       207
```

- **Ensemble Model – Voting Classifier:** The model attained the best overall accuracy of approximately 98%, indicating nearly perfect performance across the dataset.
  - **Cassava:** Showed perfect precision and near-perfect recall, suggesting the model was highly effective at identifying cassava. It only incorrectly classified one cassava instance as sugarcane.
  - **Grass:** Had perfect recall and high precision, indicating the model correctly identified all grass instances but with a few false positives.
  - **Maize** Had perfect precision and slightly lower recall, suggesting the model was good at identifying maize when it predicted it but missed a few actual maize instances and classified them as grass.
  - **Sugarcane:** Had very high precision and perfect recall, indicating the model performed excellently on sugarcane with very few false positives.
  - The excellent performance of the ensemble model showed the ability of this model to combine the strengths of all the individual models to obtain the best classification performance.

```
Voting Classifier Results:
Voting Classifier Accuracy: 0.9758454106280193
Voting Classifier Classification Report:
              precision    recall  f1-score   support

     cassava       0.99      0.99      0.99        69
       grass       0.95      1.00      0.97        54
       maize       1.00      0.95      0.98        65
   sugarcane       0.95      0.95      0.95        19

    accuracy                           0.98       207
   macro avg       0.97      0.97      0.97       207
weighted avg       0.98      0.98      0.98       207
```

Conclusion

- Traditional Machine learning models show potential in crop and weed monitoring, detection, and classification.
  - Adequately processing and nomalizing the dataset affects results.
  - Combining a variety of feature extraction techniques significantly impacts model performance.
  - Ensemble based models like Random Forest & the Voting Classifier yielded better results than other traditional ML models on the same training dataset

- Future Directions
  - Experiment with Deep Learning models like CNNs on the same dataset to compare performance between traditional and deep learni ng models on the monitoring, detection, and classification of garden crops and weeds.
  - Given the superior performance of the ensemble models, further research into other ensemble models is recommended.

References

[1] "World population to reach 8 billion this year, as growth rate slows", news.un.org, (Accessed 1 Jul. 2023).
[2] J. H. Westwood, R. Charudattan, S. O. Duke, S. A. Fennimore, P. Marrone, D. C. Slaughter, and R. Zollinger, "Weed management in 2050: Perspectives on the future of weed science," Weed Science, vol. 66, no. 3, pp. 275-285, 2018.
[3] M. Fontanelli, C. Frasconi, L. Martelloni, M. Pirchio, and R. Foster, "Weed management in autumn fresh market spinach: A nonchemical alternative," HortTechnology, vol. 25, pp. 177-184.
[4] J. H. Westwood et al., "Weed management in 2050: Perspectives on the future of weed science," Weed Science, vol. 66, no. 3, pp. 275-285, 2018.
[5] T. Pun, A. Neupane, R. Koech, "Quantification Of Root-knot Nematode Infestation In Tomato Using Digital Image Analysis", Agronomy, vol. 11, no. 12, p. 2372, 2021. https://doi.org/10.3390/agronomy11122372
[6] M. Guzel, B. Turan, İ. Kadıoğlu, B. Şin, A. Basturk, K. Ragab, "How To Affect the Number Of Images On The Success Rate For Detection Of Weeds With Deep Learn- ing", Turkish JAF Sci.Tech., vol. 10, no. 8, p. 1441- 1446, 2022. https://doi.org/10.24925/turjaf.v10i8.1441- 1446.5183
[7] Z. Wu, Y. Chen, B. Zhao, X. Kang, Y. Ding, "Re- view Of Weed Detection Methods Based On Computer Vision", Sensors, vol. 21, no. 11, p. 3647, 2021. https://doi.org/10.3390/s21113647
[8] J. Tang, X. Chen, R. Miao, and D. Wang, "Weed de tection using image processing under different illumination for site-specific areas spraying," Comput. Electron. Agric., vol. 122, pp. 103-111, 2016.
[9] L. Pereira, R. Nakamura, G. Souza, D. Martins, and J. Papa, "Aquatic weed automatic classification using ma chine learning techniques," Comput. Electron. Agric., vol. 87, pp. 56-63.
[10] Y. Zhang, S. Wang, G. Ji, and P. Phillips, "Fruit classification using computer vision and feedforward neu ral network," J. Food Eng., vol. 143, pp. 167-177.
[11] A. Bakhshipour, A. Jafari, S. Nassiri, and D. Zare, "Weed segmentation using texture features extracted from wavelet sub-images," Biosyst. Eng., vol. 157, pp. 1-12, 2017.
[13] F.-M. De Rainville et al., "Bayesian classifica tion and unsupervised learning for isolating weeds in row crops," Pattern Anal. Appl., vol. 17, no. 2, pp. 401–414, May 2014, doi: 10.1007/s10044-012-0307-5.
[14] F. Ahmed, H. Al-Mamun, A. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," Crop Prot., vol. 40, pp. 98-104.
[15] T. Rumpf, C. Romer, M. Weis, M. Sökefeld, R. Ger- hards, and L. Plumer, "Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsium arvense and Galium aparine," Comput. Electron. Agric., vol. 80, pp. 89-96.
[16] Y. Chen, P. Lin, Y. He, and Z. Xu, "Classification of broadleaf weed images using Gabor wavelets and Lie group structure of region covariance on Riemannian manifolds," Biosyst. Eng., vol. 109, pp. 220-227.

GitHub Link

- https://github.com/giddy-mpungu/Garden-Monitoring-Tool-Using-Computer-Vision